Ecological Modelling 466 (2022) 109880

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

A new parametric measure of functional dissimilarity: Bridging the gap between the Bray-Curtis dissimilarity and the Euclidean distance

Carlo Ricotta[a,*], Sandrine Pavoine[b]

[a] Department of Environmental Biology, University of Rome 'La Sapienza', Rome, Italy
[b] Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, Paris, France

ARTICLE INFO

Keywords:
Branching requirement
Complementarity
Dissimilarity profile
Species abundances
Species commonness

ABSTRACT

Community ecologists usually consider the Euclidean distance inappropriate to explore the multivariate structure of species abundance data. This is because the Euclidean distance may lead to the counterintuitive result for which two sample plots with no species in common may be more similar to each other than two plots that share the same species list. To overcome this paradoxical situation, the species abundances need to be normalized in some way. Among the many coefficients used by ecologists for the analysis of assemblage data, the Bray-Curtis dissimilarity is certainly the most commonly used. This measure entails normalization of species-wise differences between two plots by the total species abundance in both plots. By highlighting the relationship between the Bray-Curtis dissimilarity and the Euclidean distance, we propose a parametric dissimilarity measure that is appropriate for handling data on community composition. We also show how the new parametric measure can be generalized to the measurement of functional dissimilarity between two plots. A small dataset on the species functional turnover along a chronosequence on Alpine grasslands is used to illustrate the behavior of the proposed measure.

1. Introduction

Ecologists frequently use multivariate dissimilarity measures between pairs of sampling units (or plots, quadrats, sites, etc.) for investigating the ecological processes that drive community assembly. Many authors have proposed a set of guidelines to help navigate the multitude of existing dissimilarity coefficients for the analysis of ecological data (e.g. Gower and Legendre 1986; Podani 2000; Legendre and De Cáceres 2013; Lengyel and Botta-Dukát 2021). However, selecting an appropriate question-specific coefficient is not always a simple operation.

In this framework, the Euclidean distance is a natural benchmark for assessing any other dissimilarity coefficient because it corresponds to the standard notion of distance in our everyday physical world (Podani 2000). Given two plots U and V, let x_{Uj} and x_{Vj} be the abundances of species j $(j = 1, 2, ..., N)$ in both plots. The Euclidean distance between U and V is defined as:

$$E = \sqrt{\sum_{j=1}^{N} \left(x_{Uj} - x_{Vj}\right)^2} \tag{1}$$

Note that in this paper we generally use the term distance for all measures that have metric properties; otherwise, the term dissimilarity is used (see Gower and Legendre 1986).

In multivariate analysis of assemblage data, a well-known limitation of the Euclidean distance, which is usually known as the 'Orlóci paradox', is that two plots with no species in common may result more similar than two plots which share the same species (Orlóci 1978). This counterintuitive situation occurs because with the Euclidean distance differences in species abundances are much more relevant than agreement in species presences and absences (Ricotta and Podani 2017). Accordingly, an important prerequisite for any meaningful measure of community dissimilarity is what Clarke et al. (2006) have termed 'complementarity'. This means that the measure reaches its maximum value when two plots have no species in common.

To overcome the 'Orlóci paradox' we need to normalize the species abundances in some way. The dissimilarity coefficient of Bray and Curtis (1957), one of the most popular measures of multivariate dissimilarity in community ecology, entails normalization of species-wise differences in U and V by the total species abundance in both plots:

---

* Corresponding author.
E-mail address: carlo.ricotta@uniroma1.it (C. Ricotta).

https://doi.org/10.1016/j.ecolmodel.2022.109880
Received 16 November 2021; Received in revised form 17 December 2021; Accepted 12 January 2022

$$D = \frac{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|}{\sum_{j=1}^{N} \left(x_{Uj} + x_{Vj} \right)} \tag{2}$$

The Bray-Curtis dissimilarity thus calculates the fraction of the total species abundances in which the two plots differ.

The aim of this paper is twofold: first, by highlighting the relationship between the Euclidean distance and the Bray-Curtis dissimilarity, we propose a parametric formulation of Eq. (2) that is adequate for handling species abundances. Next, we will show how this new parametric measure can be further generalized to summarize the functional dissimilarity between two plots. A worked example with data on the species functional turnover along a chronosequence on Alpine grasslands is used to show the behavior of this new measure.

2. Methods

2.1. A new parametric measure of dissimilarity

We start by observing that the Euclidean distance is the second order ($\alpha = 2$) of the Minkowski parametric distance:

$$M^{\alpha} = \sqrt[\alpha]{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|^{\alpha}} \tag{3}$$

Unlike the Euclidean and the Bray-Curtis coefficients which are single-point pictures of community dissimilarity, the Minkowski distance provides a vector description of the differences in species abundance between U and V. For $\alpha \geq 1$, the Minkowski distance is a metric, thus conforming to the triangle inequality (see Gower and Legendre 1986).

In the formulation of the Minkowski distance, the parameter α is related to the distinctness between sampling units, such that increasing the value of α increases the relevance of large differences in species abundances between U and V compared to small differences. For α tending to infinity, M^{∞} tends to $\max|x_{Uj} - x_{Vj}|$. As a result, parametric dissimilarity can be thought of as a scaling process that occurs in abstract data space of species abundances (Podani 1992).

By setting $\alpha = 1$ in Eq. (3), we obtain the Manhattan (or city-block) distance:

$$M^{1} = \sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right| \tag{4}$$

which is the sum of absolute differences in species abundances between U and V.

From Eqs. (2) and 4, it follows that the Bray-Curtis dissimilarity is nothing else than the Manhattan distance normalized by the total abundance of all species in both plots:

$$D = \frac{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|}{\sum_{j=1}^{N} \left(x_{Uj} + x_{Vj} \right)} = \frac{M^{1}}{\sum_{j=1}^{N} \left(x_{Uj} + x_{Vj} \right)} \tag{5}$$

This provides a direct connection between the Bray-Curtis dissimilarity and the Minkowski parametric family. For species abundance data x_{Uj}, the observation that the Bray-Curtis dissimilarity is essentially a normalized version of the first order Minkowski distance can be generalized to the entire parametric family in one of two ways:

$$D^{\alpha} = \frac{\sqrt[\alpha]{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|^{\alpha}}}{\sqrt[\alpha]{\sum_{j=1}^{N} \left(x_{Uj} + x_{Vj} \right)^{\alpha}}} = \sqrt[\alpha]{\frac{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|^{\alpha}}{\sum_{j=1}^{N} \left(x_{Uj} + x_{Vj} \right)^{\alpha}}} \tag{6a}$$

or, since $\sum_{j=1}^{N} (x_{Uj} + x_{Vj}) = \sum_{j=1}^{N} x_{Uj} + \sum_{j=1}^{N} x_{Vj}$

$$\Delta^{\alpha} = \frac{\sqrt[\alpha]{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|^{\alpha}}}{\sqrt[\alpha]{\sum_{j=1}^{N} x_{Uj}^{\alpha} + \sum_{j=1}^{N} x_{Vj}^{\alpha}}} = \sqrt[\alpha]{\frac{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|^{\alpha}}{\sum_{j=1}^{N} x_{Uj}^{\alpha} + \sum_{j=1}^{N} x_{Vj}^{\alpha}}} \tag{6b}$$

For $\alpha = 1$, Eq. (6a) and 6b both recover the Bray-Curtis dissimilarity, while for $\alpha = 2$ we obtain two equally admissible normalized versions of the classical Euclidean distance: $D^{2} = \sqrt[2]{\sum_{j=1}^{N} (x_{Uj} - x_{Vj})^{2}} / \sqrt[2]{\sum_{j=1}^{N} (x_{Uj} + x_{Vj})^{2}}$ and $\Delta^{2} = \sqrt[2]{\sum_{j=1}^{N} (x_{Uj} - x_{Vj})^{2}} / \sqrt[2]{\sum_{j=1}^{N} x_{Uj}{}^{2} + \sum_{j=1}^{N} x_{Vj}{}^{2}}$, respectively. Eq. (6a) and 6b thus represent two normalized expressions of the Minkowski distance in the range $[0, 1]$ that conform to the complementarity requirement. If U and V have no species in common D^{α} and Δ^{α} are both equal to 1, whereas if for all N species $x_{Uj} = x_{Vj}$, we have $D^{\alpha} = \Delta^{\alpha} = 0$.

Note that the so-called Minkowski inequality: $\sqrt[\alpha]{\sum_{j=1}^{N} |x_{Uj} - x_{Vj}|^{\alpha}} \leq \sqrt[\alpha]{\sum_{j=1}^{N} x_{Uj}^{\alpha}} + \sqrt[\alpha]{\sum_{j=1}^{N} x_{Vj}^{\alpha}}$ could also be used to construct a parametric family of normalized dissimilarities

$$L^{\alpha} = \frac{\sqrt[\alpha]{\sum_{j=1}^{N} \left| x_{Uj} - x_{Vj} \right|^{\alpha}}}{\sqrt[\alpha]{\sum_{j=1}^{N} x_{Uj}^{\alpha}} + \sqrt[\alpha]{\sum_{j=1}^{N} x_{Vj}^{\alpha}}} \tag{6c}$$

This measure has been previously proposed by Yuan et al. (2016) to quantify the biodiversity turnover from species proportions. However, unlike D^{α} or Δ^{α}, Eq. (6c) does not always assign maximum dissimilarity (i.e. $L^{\alpha} = 1$) to a pair of completely distinct assemblages with no species in common.

2.2. Extending the measure to functional differences between plots

Functional differences between species are usually represented by a $N \times N$ matrix of pairwise dissimilarities d_{ij} between species i and j such that $d_{ij} = d_{ji}$ and $d_{ii} = 0$. If d_{ij} is bounded in the range $[0, 1]$, a corresponding functional similarity coefficient can be easily derived as the complement of d_{ij} (i.e. $s_{ij} = 1 - d_{ij}$).

According to Leinster and Cobbold (2012), the functional abundance/commonness of species j in plot U can be defined as the abundance of all species in U that are functionally similar to j (including j itself):

$$c_{Uj} = \sum_{i=1}^{N} x_{Ui} s_{ij} \tag{7}$$

Therefore, assuming that species with similar traits are likely to support similar functions (Villéger et al., 2013), c_{Uj} summarizes the abundance of all individuals in plot U that support the functions associated with species j. For details, see Pavoine and Ricotta (2019). If all species in U are maximally dissimilar from j such that $s_{ij} = 0$ for all $i \neq j$, we have $c_{Uj} = x_{Uj}$, meaning that the abundance of all species similar to j cannot be less than the abundance of j itself. At the other extreme, if all species are functionally identical to j such that $s_{ij} = 1$, we have $c_{Uj} = \sum_{j=1}^{N} x_{Uj}$ (i.e. the total species abundance in plot U).

In principle, we can thus derive a family of parametric measures of functional dissimilarity between plots by substituting in Eq. (6a) and 6b the species abundances x_{Uj} with their commonness c_{Uj}:

$$F^{\alpha} = \sqrt[\alpha]{\frac{\sum_{j=1}^{N} \left| c_{Uj} - c_{Vj} \right|^{\alpha}}{\sum_{j=1}^{N} \left(c_{Uj} + c_{Vj} \right)^{\alpha}}} \tag{8a}$$

and

$$\Phi^{\alpha} = \sqrt[\alpha]{\frac{\sum_{j=1}^{N} \left| c_{Uj} - c_{Vj} \right|^{\alpha}}{\sum_{j=1}^{N} c_{Uj}^{\alpha} + \sum_{j=1}^{N} c_{Vj}^{\alpha}}} \tag{8b}$$

where the summation is taken over all species that are actually present in at least one of the two plots (i.e. over all species for which $x_{Uj} + x_{Vj} > 0$).

Eq. (8a) and 8b provide a parametric version of the Bray-Curtis dissimilarity that includes functional differences between species. However, this solution is not entirely satisfactory for two main reasons: first, most researchers usually try to assess how the species functional strategies are apportioned within the plots, irrespective of the species absolute abundances in each plot (Ricotta et al., 2021a). This aspect can be adjusted by calculating functional dissimilarity from the species *relative* abundances p_{Uj} instead of absolute abundances x_{Uj}. In this case, species commonness becomes:

$$\omega_{Uj} = \sum_{i=1}^{N} p_{Ui} s_{ij} \tag{9}$$

thus representing the *relative* abundance of all species in plot U that are functionally similar to j with $p_{Uj} = x_{Uj} / \sum_{j=1}^{N} x_{Uj}$ and $0 \leq \omega_{Uj} \leq 1$.

Second, Eq. (8a) and 8b do not conform to the requirement that dissimilarity remains unchanged if two species that are functionally identical in every way are merged into a single species (Leinster and Cobbold 2012; Pavoine and Ricotta 2019). The essence of this branching requirement is that a measure of functional dissimilarity should be able to highlight differences in ecosystem functioning between sampling units regardless of the species that sustain these functions. For throughout discussion of this aspect, see Leinster and Cobbold (2012); Botta-Dukát (2018); Ricotta et al. (2021a).

According to Pavoine and Ricotta (2019), this additional aspect can be fixed by multiplying all terms of the summations in Eqs. (8a) and 8b by a species-specific weighting factor

$$\lambda_j = \frac{(p_{Uj} + p_{Vj})}{\sum_{j=1}^{N} (p_{Uj} + p_{Vj})} = \frac{1}{2} (p_{Uj} + p_{Vj}) \tag{10}$$

which represents the pooled abundance of species j in U and V relative to the total species abundance in both plots.

Therefore, by substituting the species absolute abundances x_{Uj} with their relative abundances p_{Uj} and by introducing the weighting factors λ_j in the calculation of parametric dissimilarity, we can derive two parametric measures that conform to our intuitive notion of functional dissimilarity better than the previous ones:

$$f^{\alpha} = \sqrt[\alpha]{\frac{\sum_{j=1}^{N} \lambda_j |\omega_{Uj} - \omega_{Vj}|^{\alpha}}{\sum_{j=1}^{N} \lambda_j (\omega_{Uj} + \omega_{Vj})^{\alpha}}} \tag{11a}$$

and

$$\phi^{\alpha} = \sqrt[\alpha]{\frac{\sum_{j=1}^{N} \lambda_j |\omega_{Uj} - \omega_{Vj}|^{\alpha}}{\sum_{j=1}^{N} \lambda_j \omega_{Uj}^{\alpha} + \sum_{j=1}^{N} \lambda_j \omega_{Vj}^{\alpha}}} \tag{11b}$$

Both measures conform to the requirement that the functional dissimilarity between U and V remains unchanged if two species in U or V that are functionally identical in every way are merged into a single species. In addition, if $s_{ij} = 0$ for all $i \neq j$ and $\lambda_j = 1/N$ for all N species in the assemblage, Eqs. (11a) and 11b recover their abundance-based versions D^{α} and Δ^{α}, respectively.

3. Worked example

In this paper, data on Alpine vegetation sampled by Caccianiga et al. (2006) along a chronosequence at the foreland of the Rutor glacier (northern Italy) were used. The same data were also used by Ricotta et al. (2021a) to investigate the behavior of a different parametric measure of functional dissimilarity. This allows us to compare our results with those of Ricotta et al. (2021a). The data set (available in Ricotta et al., 2016: Appendix S2) is composed of a community composition matrix with the abundances of 45 species collected in 59 plots. The size of each plot was approximately 25 m^2; all species abundances were measured with a five-point ordinal scale transformed to ranks. Based on the age of the moraine ridges, plots were assigned to three distinct successional stages: early successional vegetation (ESV, 17 plots), mid successional vegetation (MSV, 32 plots), and late successional vegetation (LSV, 10 plots).

Six functional traits available in Caccianiga et al. (2006) related to the species global spectrum of form and function (Díaz et al., 2016) were used: leaf dry matter content (LDMC; %), leaf dry weight (LDW; mg), specific leaf area (SLA; mm^2 × mg^{-1}), leaf carbon content (LCC; %), leaf nitrogen content (LNC; %), and canopy height (CH; mm). First, the traits were linearly rescaled to zero mean and unit standard deviation. Next, the scaled traits were used to calculate a matrix of functional Euclidean distances d_{ij} between the 45 species in the data set. The Euclidean distances were finally normalized in the unit range by dividing each distance by the maximum value in the distance matrix.

To compute parametric dissimilarity, the species abundances of all plots within each stage were averaged and the species relative abundances within each stage were computed. The species relative abundances within each stage were then used, together with the functional distances, to compute the parametric dissimilarity between the three successional stages according to Eqs. (11a) and 11b. To this end, a new R script available in the Supporting information (Appendix 1 and 2) has been produced.

The dissimilarity profiles of f^{α} vs. α for $\alpha > 0$ among the three successional stages are shown in Fig. 1. The profiles of ϕ^{α} are very similar to those of f^{α}. Therefore, they are shown in Appendix 1. Caccianiga et al. (2006) and Ricotta et al. (2020) showed that the different successional stages of the chronosequence are characterized by a gradual substitution of ruderal species by stress tolerator species. From a functional viewpoint, this pattern is associated to a progressive reduction of leaf nitrogen content and specific leaf area and a corresponding increase of leaf carbon content and leaf dry matter.

These functional differences are mirrored by the dissimilarity profiles of the three successional stages: in good agreement with Ricotta et al. (2021a), the dissimilarity profiles of the intermediate stages of the chronosequence (ESV vs. MSV and MSV vs. LSV) show less pronounced functional differences, whereas the largest functional differences are shown by the dissimilarity profile between the early and the late successional stages of the Alpine vegetation (ESV vs. LSV). By increasing the values of the parameter α, these differences tend to become increasingly larger, thus showing that the dominant species in the successional stages at the opposite ends of the chronosequence tend to be functionally well distinct from each other.

One of the criticisms sometimes leveled at the Minkowski parametric distance is that, for $\alpha > 2$, it is highly influenced by large species-wise differences, which is not always justified. However, in Fig. 1 we can see that the dissimilarity profiles reach a plateau for values of α far beyond this threshold. Therefore, the dissimilarity values calculated for $\alpha > 2$ carry an important amount of ecological information on the pattern of functional differences between the dominant species in different sampling units. This pattern would be overlooked if the calculation of the dissimilarity profiles were limited to lower values of α.

4. Discussion

In this paper, we introduced a new parametric measure of community dissimilarity that extends the normalization method inherent in the Bray-Curtis dissimilarity to the whole Minkowski parametric family. Parametric dissimilarity has been used in community ecology for a long time (e.g. Grassle and Smith 1976; Jost 2007; Chao et al., 2014; Chao and Ricotta 2019) and its use is due to the consciousness that no single index is able to adequately summarize all facets of the multivariate dissimilarity among sampling units. Therefore, one uses a parametric family of dissimilarity measures whose members have increasing

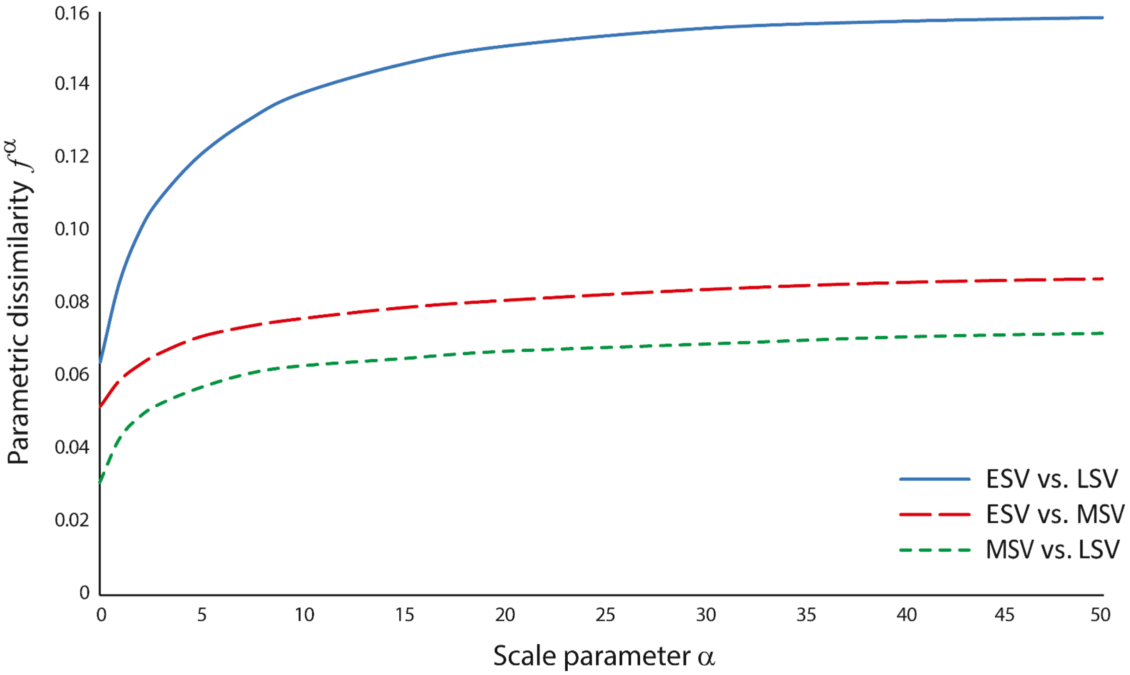


Fig. 1. Functional dissimilarity profiles f^α vs. α among the three successional stages of the Rutor chronosequence. ESV = early successional vegetation; MSV = mid successional vegetation; LSV = late successional vegetation.

sensitivities to large species-wise differences between plots for increasing values of the selected scale parameter. With parametric functions, such as D^α or Δ^α, dissimilarity can be plotted against the corresponding scale parameter and the resulting profiles can be compared for the various communities under study (Taillie 1979).

A desirable property of D^α and Δ^α is that for both measures, a few characteristic values of the parameter α recover more traditional dissimilarity coefficients. For example, for $\alpha = 1$ both parametric functions reduce to the Bray-Curtis dissimilarity, while for $\alpha = 2$, two normalized versions of the classical Euclidean distance, D^2 and Δ^2, are obtained. Therefore, D^α and Δ^α are not just a mere addition to the dissimilarity measures already available in the ecologist's toolbox. Rather, an interesting novelty of this work is that such measures provide an explicit relationship between the Bray-Curtis dissimilarity and the Minkowski family that can be further extended to the measurement of functional dissimilarity. Hence, by providing a unifying perspective for a number of seemingly unrelated dissimilarity coefficients, both parametric measures help organize different aspects of species resemblance into a higher-order coherent framework. For a detailed analysis of the relationships between the newly proposed parametric measures of community dissimilarity and a number of classical single-point measures of resemblance, see Appendix 3.

Note that the normalized Euclidean distances D^2 and Δ^2 are both S-decomposable, meaning that their squared values can be additively decomposed into species-level contributions (Ricotta et al., 2021b). Accordingly, with D^2 and Δ^2 we can decompose the within- and between-group components of distance-based multivariate ANOVA into additive species-level values. In this way, we are able to identify the species that contribute most to the compositional differentiation among the various groups of plots (Ricotta et al., 2021b).

By substituting the species abundances x_{Uj} with their commonness ω_{Uj}, parametric dissimilarity can be generalized to account for functional differences among species. Unlike most of the functional dissimilarity measures published to date (e.g. Cardoso et al., 2014; Chao et al., 2014; Chiu and Chao 2014; Pavoine and Ricotta 2014), the resulting parametric measures f^α and ϕ^α are not based on Whittaker's (1972) classical model of diversity decomposition into alpha, beta and gamma components. Therefore, they can be calculated from virtually any available interspecies dissimilarity measure without restrictions on their geometrical properties (for additional details, see e.g. Pavoine and Ricotta 2014). This high flexibility in the choice of the interspecies dissimilarity coefficients is a desirable aspect of the proposed framework.

If the interspecies dissimilarities d_{ij} are in the range $[0, 1]$, the corresponding similarities $s_{ij} = 1 - d_{ij}$ can be interpreted as the fuzzy degree of functional compatibility of species i with species j. Likewise, the commonness of species j, $\omega_{Uj} = \sum_{i=1}^{N} p_{Ui} s_{ij}$ can be seen as the mean fuzzy compatibility of all species in U with species j (including j itself). Since most parametric and non-parametric functional dissimilarity and diversity measures are calculated from interspecies dissimilarities, this relationship opens the way for a more general mathematical description of biodiversity in terms of fuzzy set theory (Cross and Sudkamp 2002; Feoli 2018).

To conclude, while how to summarize functional dissimilarity remains an open question, we think that the major advantage of parametric measures is that by comparing different pairs of plots by their dissimilarity profiles, it is possible to trace how dissimilarity changes as a function of the scale parameter α. This operation may help understanding the effects of large and small species-wise differences on ecological processes in a more general way than by simply using a scalar measure.

A familiar problem related to the differential weighting of large and small species-wise differences for different values of the parameter α is that two dissimilarity profiles can cross, such that we cannot unambiguously say which of the two profiles shows the largest functional differences (Ricotta et al., 2021a). In this paper, we used numerical simulations to show that this is also the case for the newly proposed parametric measures f^α and ϕ^α.

This potential inconsistency in the ordering of two parametric profiles was at the basis of Hurlbert's (1971) critique on the 'nonconcept' of diversity. However, as emphasized by Patil and Taillie (1982) and Leinster and Cobbold (2012), inconsistent ordering is a common problem in multivariate analysis and should not be the case for undue pessimism. On the contrary, when two dissimilarity profiles cross, the positions of the intersections provide relevant ecological information on the compositional and functional differences between the communities under scrutiny.

We thus hope that in spite of all its biological and statistical complexity, this work will help awaken some interest to parametric dissimilarity functions and their ecological applications.

Supporting information

Appendix 1. R scripts for the calculation of parametric dissimilarity: manual and examples.

Appendix 2. R scripts in text format.

Appendix 3. On the relationship between the newly proposed parametric measures of community dissimilarity and a number of classical single-point measures of resemblance.

CRediT authorship contribution statement

Carlo Ricotta: Conceptualization, Methodology, Data curation, Writing – original draft. **Sandrine Pavoine:** Methodology, Software, Data curation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2022.109880.

References

Botta-Dukát, Z., 2018. The generalized replication principle and the partitioning of functional diversity into independent alpha and beta components. Ecography 41, 40–50.
Bray, J., Curtis, J., 1957. An ordination of the upland forest communities in southern Wisconsin. Ecol. Monogr. 27, 325–349.
Caccianiga, M., Luzzaro, A., Pierce, S., Ceriani, R.M., Cerabolini, B., 2006. The functional basis of a primary succession resolved by CSR classification. Oikos 112, 10–20.
Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J., Schmera, D., 2014. Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. J. Biogeogr. 41, 749–761.
Chao, A., Chiu, C.H., Jost, L., 2014. Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through Hill numbers. Annu. Rev. Ecol. Evol. Syst. 45, 297–324.
Chao, A., Ricotta, C., 2019. Quantifying evenness and linking it to diversity, beta diversity, and similarity. Ecology 100, e02852.
Chiu, C.H., Chao, A., 2014. Distance-based functional diversity measures and their decomposition: a framework based on Hill numbers. PLoS ONE 9, e100014.
Clarke, K.R., Somerfield, P.J., Chapman, M.G., 2006. On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. J. Exp. Mar. Biol. Ecol. 330, 55–80.
Cross, V.V., Sudkamp, T.A., 2002. Similarity and Compatibility in Fuzzy Set Theory. Physica Verlag, Heidelberg.
Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., et al., 2016. The global spectrum of plant form and function. Nature 529, 167–171.
Feoli, E., 2018. Classification of plant communities and fuzzy diversity of vegetation systems. Commun. Ecol. 19, 186–198.
Gower, J.C., Legendre, P., 1986. Metric and Euclidean properties of dissimilarity coefficients. J. Classif. 3, 5–48.
Grassle, J.F., Smith, W., 1976. A similarity measure sensitive to the contribution of rare species and its use in investigation of variation in marine benthic communities. Oecologia 25, 13–22.
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., De Cáceres, M., 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. Ecol. Lett. 16, 951–963.
Leinster, T., Cobbold, C.A., 2012. Measuring diversity: the importance of species similarity. Ecology 93, 477–489.
Lengyel, A., Botta-Dukát, Z. (2021) Review and performance evaluation of trait-based between-community dissimilarity measures. bioRxiv, doi: 10.1101/2021.01.06.425560.
Orlóci, L., 1978. Multivariate Analysis in Vegetation Research. The Hague, Junk.
Patil, G.P., Taillie, C., 1982. Diversity as a concept and its measurement. J. Am. Stat. Assoc. 77, 548–561.
Pavoine, S., Ricotta, C., 2014. Functional and phylogenetic similarity among communities. Methods Ecol. Evol. 5, 666–675.
Pavoine, S., Ricotta, C., 2019. Measuring functional dissimilarity among plots: adapting old methods to new questions. Ecol. Indic. 97, 67–72.
Podani, J., 1992. Space series analysis: processes reconsidered. Abstracta Bot. 16, 25–29.
Podani, J., 2000. Introduction to the Exploration of Multivariate Biological Data. Backhuys, Leiden, NL.
Ricotta, C., Acosta, A.T.R., Caccianiga, M., Cerabolini, B.E.L., Godefroid, S., Carboni, M., 2020. From abundance-based to functional-based indicator species. Ecol. Indic. 118, 106761.
Ricotta, C., de Bello, F., Moretti, M., Caccianiga, M., Cerabolini, B., Pavoine, S., 2016. Measuring the functional redundancy of biological communities: a quantitative guide. Methods Ecol. Evol. 7, 1386–1395.
Ricotta, C., Pavoine, S., Cerabolini, B.E.L., Pillar, V., 2021b. A new method for indicator species analysis in the framework of multivariate analysis of variance. J. Veg. Sci. 32, e13013.
Ricotta, C., Podani, J., 2017. On some properties of the Bray-Curtis dissimilarity and their ecological meaning. Ecol. Complex. 31, 201–205.
Ricotta, C., Szeidl, L., Pavoine, S., 2021a. Towards a unifying framework for diversity and dissimilarity coefficients. Ecol. Indic. 129, 107971.
Taillie, C., 1979. Species equitability: a comparative approach. In: Grassle, J.F., Patil, G. P., Smith, W.K., Taillie, C. (Eds.), Ecological Diversity in Theory and Practice. International Cooperative Publishing House, Fairland, MD, pp. 51–62.
Villéger, S., Grenouillet, G., Brosse, S., 2013. Decomposing functional β-diversity reveals that low functional β-diversity is driven by low functional turnover in European fish assemblages. Global Ecol. Biogeogr. 22, 671–681.
Whittaker, R., 1972. Evolution and measurement of species diversity. Taxon 21, 213–251.
Yuan, Y., Buckland, S.T., Harrison, P.J., Foss, S., Johnston, A., 2016. Using species proportions to quantify turnover in biodiversity. J. Agric. Biol. Environ. Stat. 21, 363–381.

Appendix 1. R scripts: manual and examples

Functions `spedisparam` and `fundisparam` available in Appendix 2 allow calculating Eq. 3, 6a, 6b, and 6c (function `spedisparam`) and 11a, 11b + the equivalents of 3 and 6c applied to species commonness (function `fundisparam`). The functions will be later integrated in package adiv of R (Pavoine 2020a,b).

Usage:

```
spedisparam(comm, method = c("D", "Delta", "L", "M"), abundance =
c("relative", "absolute"), alpha = 2, tol = 1e-8, ...)

fundisparam(comm, dis, method = c("D", "Delta", "L", "M"), abundance
= c("relative", "absolute"), alpha = 2, tol = 1e-8, ...)
```

Arguments:

`comm`: a matrix or a data frame of N plots × S species containing the relative or absolute abundance of all species. Columns are species and plots are rows. Species must have names.

`dis`: a matrix or an object of class dist providing the functional dissimilarities between species (dissimilarities are nonnegative, symmetric, and the dissimilarity between a species and itself is zero). ***All species in comm must be included as row and column names in dis.***

`method`: a character string or a vector of character strings (if a vector is used only the first value is retained). Possible values are `"D"` for Eq. 11a of the main text, `"Delta"` for Eq. 11b of the main text, `"L"` for Eq. 6c (applied to species commonness rather than abundance in function `fundisparam`), and `"M"` for Eq. 3 (applied to species commonness rather than abundance in function `fundisparam`).

`abundance`: a character string or a vector of character strings (if a vector is used only the first value is retained). Possible values are `"relative"` for using relative abundances, and `"absolute"` for using absolute abundances.

`alpha:` a positive numeric for parameter alpha of Eq. 3, 6a,b,c, 11a,b of the main text.

`tol:` a numeric tolerance threshold. A value between –tol and tol is considered as zero.

Example of use:

Install and load package `adiv`.

```
install.packages("adiv") # Pavoine (2020a,b)

library(adiv)
```

Load in R functions `spedisparam` and `fundisparam` available in Appendix 2 (a .txt file). For that, use the following script to source the Appendix:

```
source(file.choose())
```

Load the dataset named RutorGlacier, available in package adiv.

```
data(RutorGlacier)
```

`RutorGlacier$Abund` gives the abundance of plant species in plots and `RutorGlacier$Traits2` the values of traits for all observed plant species. In addition, `RutorGlacier$Fac` indicates which level of the successional gradient a given plot belongs to ("early" = early-successional stage, "mid" = mid-successional stage and "late" = late-successional stage).

The following commands provide the average abundance of each species in each successional stage:

```
FAC <- factor(RutorGlacier$Fac, levels = c("early","mid","late"))
MeanAbund <- apply(RutorGlacier$Abund, 2, function(x) tapply(x, FAC, mean))
```

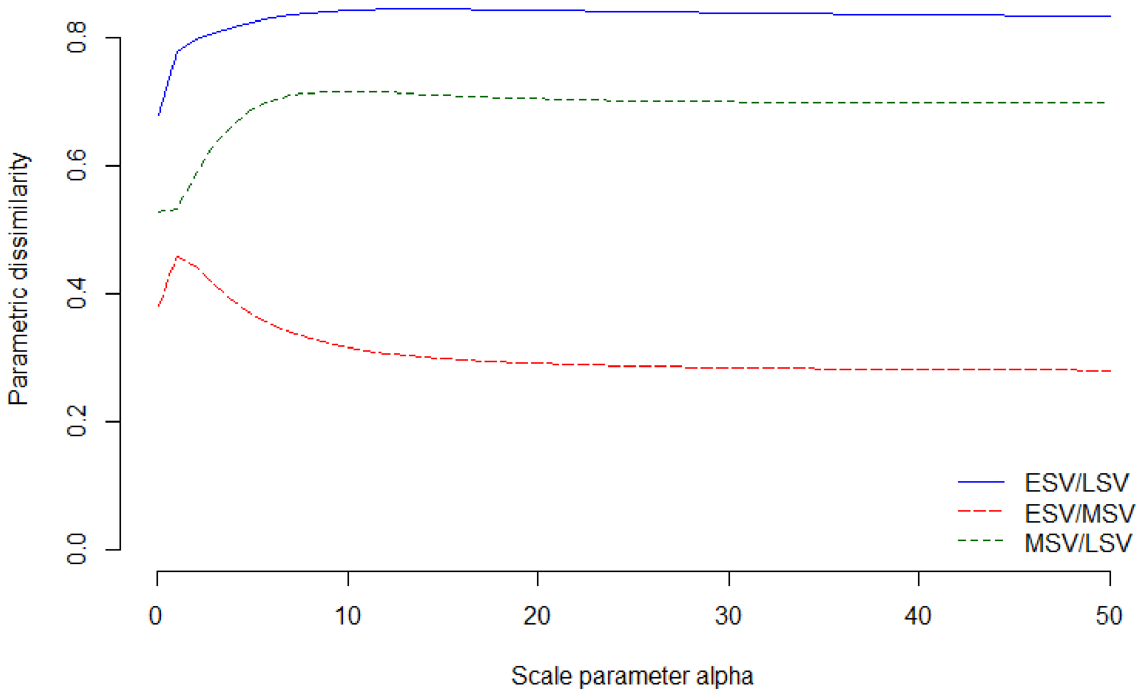
Pairwise species-based distances between successional stages were then obtained as follows:

```
# with Eq. 6a of the main text applied with relative abundances
D_EM  <-  unlist(lapply(c(0.05,1:50),  function(a)  spedisparam(MeanAbund[
c("early","mid"), ], method="D", alpha = a)))

D_EL  <-  unlist(lapply(c(0.05,1:50),  function(a)  spedisparam(MeanAbund[
c("early","late"), ], method="D", alpha = a)))

D_ML  <-  unlist(lapply(c(0.05,1:50),  function(a)  spedisparam(MeanAbund[
c("mid","late"), ], method="D", alpha = a)))

plot(c(0.05,1:50), D_EM, ylim = c(0,max(c(D_EM, D_EL, D_ML))), col="red",
type="l",   lty=5,   xlab="Scale   parameter   alpha",   ylab="Parametric
dissimilarity", bty="n")
lines(c(0.05,1:50), D_EL, col="blue")
lines(c(0.05,1:50), D_ML, col="darkgreen", lty=2)
legend("bottomright",    col=c("blue","red","darkgreen"),    lty=c(1,5,2),
legend=c("ESV/LSV", "ESV/MSV", "MSV/LSV"), bty="n")
```

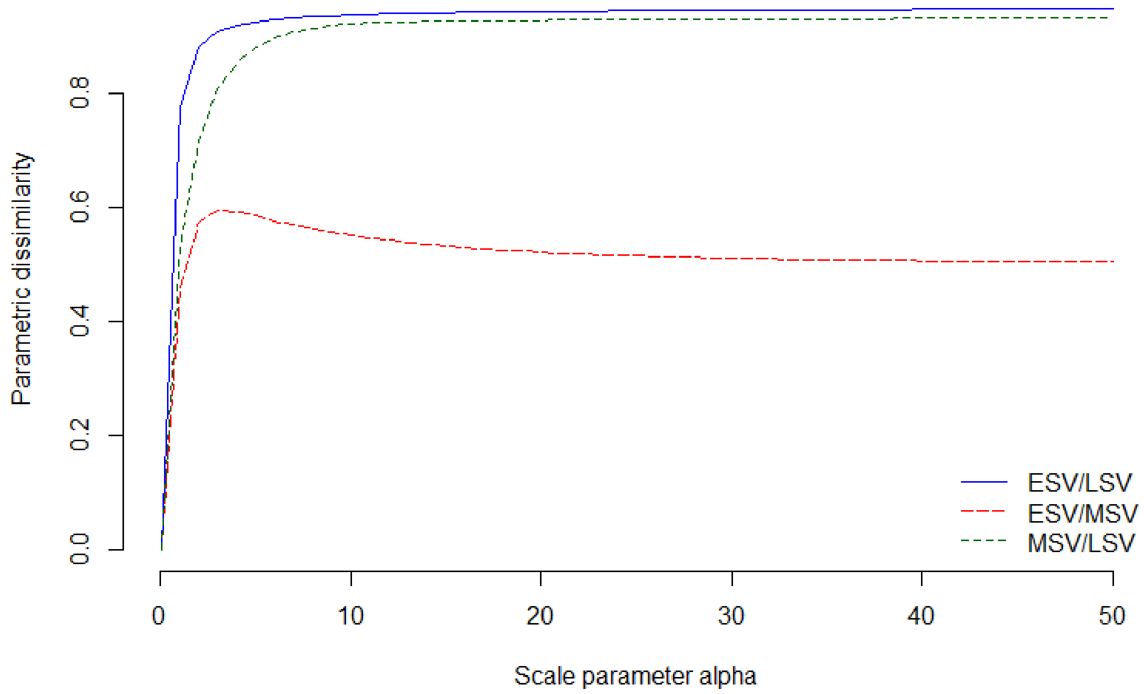


```
# with Eq. 6b of the main text applied with relative abundances
Delta_EM <- unlist(lapply(c(0.05,1:50), function(a) spedisparam(MeanAbund[
c("early","mid"), ], method="Delta", alpha = a)))

Delta_EL <- unlist(lapply(c(0.05,1:50), function(a) spedisparam(MeanAbund[
c("early","late"), ], method="Delta", alpha = a)))

Delta_ML <- unlist(lapply(c(0.05,1:50), function(a) spedisparam(MeanAbund[
c("mid","late"), ], method="Delta", alpha = a)))

plot(c(0.05,1:50), Delta_EM, ylim = c(0,max(c(Delta_EM, Delta_EL,
Delta_ML))), col="red", type="l", lty=5, xlab="Scale parameter alpha",
ylab="Parametric dissimilarity", bty="n")
lines(c(0.05,1:50), Delta_EL, col="blue")
lines(c(0.05,1:50), Delta_ML, col="darkgreen", lty=2)
legend("bottomright", col=c("blue","red","darkgreen"), lty=c(1,5,2),
legend=c("ESV/LSV", "ESV/MSV", "MSV/LSV"), bty="n")
```



Like in Ricotta et al. (2021), the functional distances between species are calculated using the Euclidean distance applied to species' traits, each standardized to zero mean and unit standard deviation:

```
fdis <- dist(scale(RutorGlacier$Traits2[1:6]))
```

Then the resulting functional distances among species are scaled to the unit range by dividing each distance by the maximum value in the distance matrix:

```
fdis <- fdis/max(fdis)
```

Pairwise functional distances between plots were then obtained as follows:

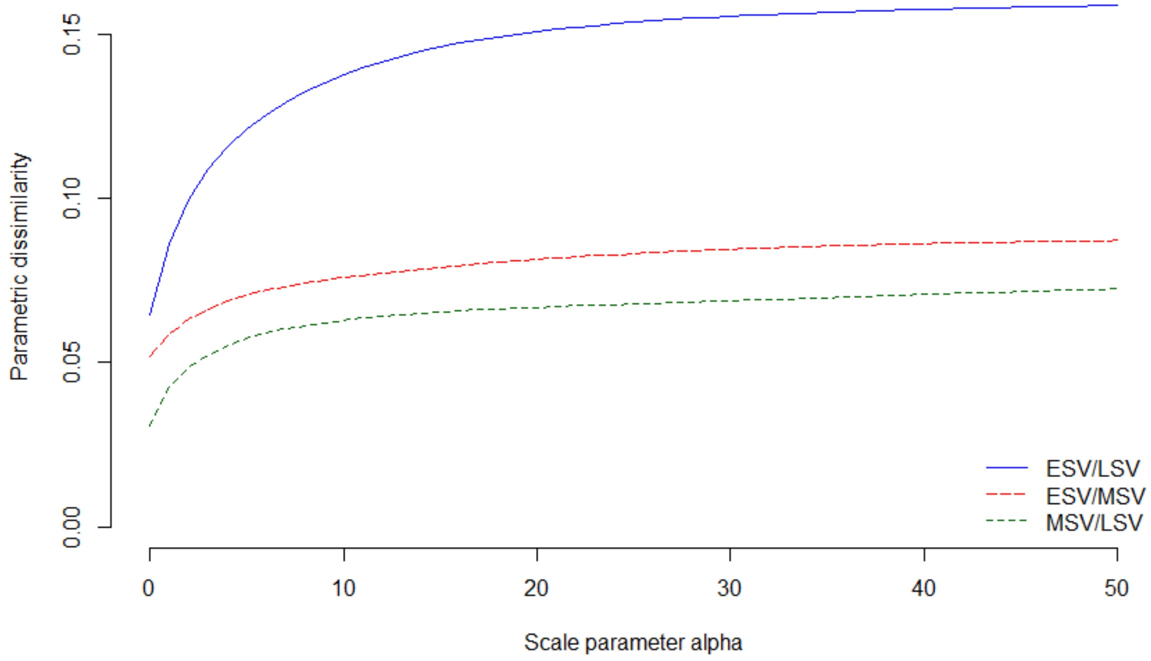
```
# with Eq. 11a of the main text
Dfun_EM <- unlist(lapply(c(1e-8,1:50), function(a) fundisparam(MeanAbund[
c("early","mid"), ], fdis, method="D", alpha = a)))

Dfun_EL <- unlist(lapply(c(1e-8,1:50), function(a) fundisparam(MeanAbund[
c("early","late"), ], fdis, method="D", alpha = a)))

Dfun_ML <- unlist(lapply(c(1e-8,1:50), function(a) fundisparam(MeanAbund[
c("mid","late"), ], fdis, method="D", alpha = a)))

plot(c(1e-8,1:50), Dfun_EM, ylim = c(0,max(c(Dfun_EM, Dfun_EL, Dfun_ML))),
col="red", type="l", lty=5, xlab="Scale parameter alpha", ylab="Parametric
dissimilarity", bty="n")
lines(c(1e-8,1:50), Dfun_EL, col="blue")
lines(c(1e-8,1:50), Dfun_ML, col="darkgreen", lty=2)
```

```
legend("bottomright", col=c("blue","red","darkgreen"), lty=c(1,5,2),
legend=c("ESV/LSV", "ESV/MSV", "MSV/LSV"), bty="n")
```

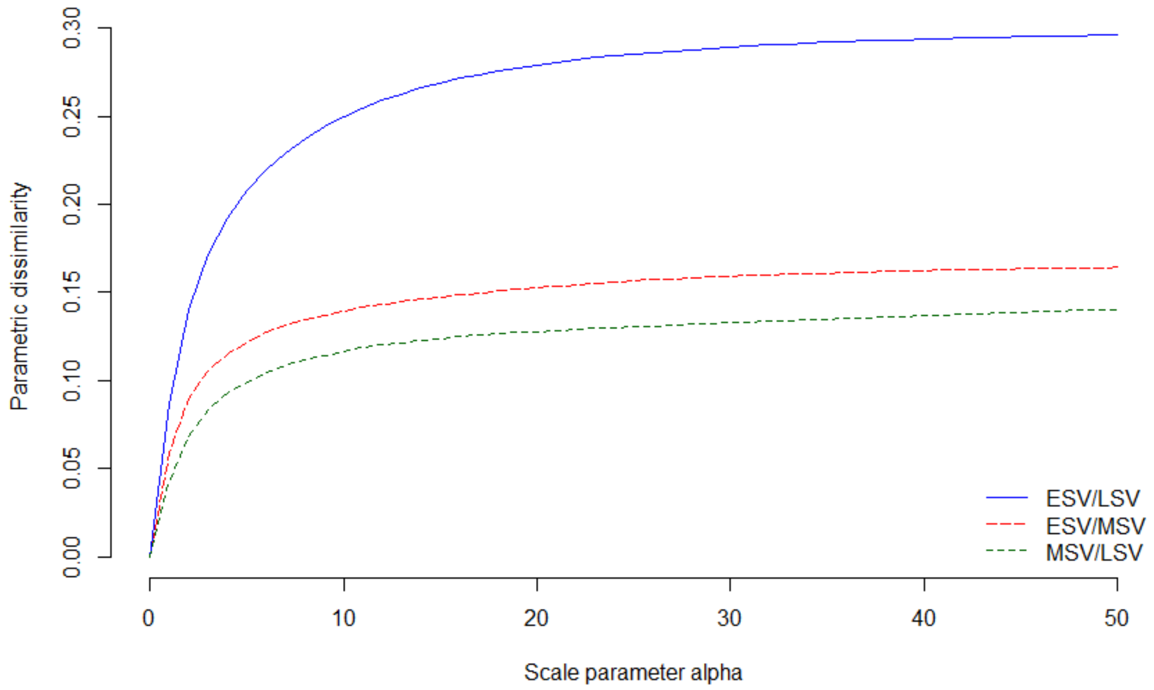


```
# with Eq. 11b of the main text
Deltafun_EM <- unlist(lapply(c(1e-8,1:50), function(a)
fundisparam(MeanAbund[ c("early","mid"), ], fdis, method="Delta", alpha =
a)))

Deltafun_EL <- unlist(lapply(c(1e-8,1:50), function(a)
fundisparam(MeanAbund[ c("early","late"), ], fdis, method="Delta", alpha =
a)))

Deltafun_ML <- unlist(lapply(c(1e-8,1:50), function(a)
fundisparam(MeanAbund[ c("mid","late"), ], fdis, method="Delta", alpha =
a)))

plot(c(1e-8,1:50), Deltafun_EM, ylim = c(0,max(c(Deltafun_EM, Deltafun_EL,
Deltafun_ML))), col="red", type="l", lty=5, xlab="Scale parameter alpha",
ylab="Parametric dissimilarity", bty="n")
lines(c(1e-8,1:50), Deltafun_EL, col="blue")
lines(c(1e-8,1:50), Deltafun_ML, col="darkgreen", lty=2)
legend("bottomright", col=c("blue","red","darkgreen"), lty=c(1,5,2),
legend=c("ESV/LSV", "ESV/MSV", "MSV/LSV"), bty="n")
```



References

Pavoine, S. (2020a) adiv: Analysis of Diversity. R package version 2.0. https://CRAN.R-project.org/package=adiv

Pavoine, S. (2020b) adiv: an R package to analyse biodiversity in ecology. Methods in Ecology and Evolution 11: 1106–1112. https://doi.org/10.1111/2041-210X.13430

Ricotta, C., Kosman, E., Laroche, F., Pavoine, S. (2021) Beta redundancy for functional ecology. Methods in Ecology and Evolution 12: 1062–1069. https://doi.org/10.1111/2041-210X.13587

```
spedisparam <-
function(comm, method = c("D", "Delta", "L", "M"), abundance =
c("relative", "absolute"), alpha = 2, tol = 1e-8, ...)
{
    q <- alpha[1]
    if(!is.numeric(q)) stop("q must be a numeric")
    if(!inherits(comm, "matrix") && !inherits(comm, "data.frame"))
stop("Incorrect definition of comm")
    method <- method[1]
    if(!method%in%c("D", "Delta", "L", "M")) stop("Incorrect definition
for method")
    abundance <- abundance[1]
    if(!abundance%in%c("relative", "absolute"))
stop("Incorrect definition for abundance")
    dataset <- t(comm)
    total <- colSums(dataset)
    if(abundance == "relative")
        abu <- sweep(dataset, 2, total, "/")
    else abu <- dataset
    num.plot <- dim(dataset)[2]
    num.sp <- dim(dataset)[1]
    names<-list(colnames(dataset), colnames(dataset))
    dis.matrix<-matrix(0, nrow=num.plot, ncol=num.plot, dimnames=names)
    for (i in 2:num.plot) {
    for (j in 1:(i-1)) {
        Zik <- abu[, j]
        Zih <- abu[, i]
        tabZ <- rbind.data.frame(Zik, Zih)
        garde <- apply(tabZ, 2, sum)>tol
         Zik <- Zik[garde]
        Zih <- Zih[garde]
        tabZ <- tabZ[, garde]

        NUM <- (sum(((abs(Zik-Zih))^q)))^(1/q)

        if(method == "D")
            DEN <- (sum(((Zik+Zih)^q)))^(1/q)
        else if(method == "Delta")
            DEN <- ( sum((Zik^q))+ sum((Zih^q)) )^(1/q)
        else if(method == "M")
            DEN <- 1
        else
            DEN <- (sum((Zik^q)))^(1/q) + (sum((Zih^q)))^(1/q)
        index <- NUM/DEN
        dis.matrix[i, j] <- index
    }
    }
    dis.matrix <- dis.matrix + t(dis.matrix)
return(as.dist(dis.matrix, ...))
}

fundisparam <-
function(comm, dis, method = c("D", "Delta", "L", "M"), abundance =
c("relative", "absolute"), alpha = 2, tol = 1e-8, ...)
{
    q <- alpha[1]
    if(!is.numeric(q)) stop("q must be a numeric")
    if(inherits(dis, "dist")) dis <- as.matrix(dis)
    if(!inherits(dis, "matrix")) stop("Incorrect definition of dis")
    if(any(dis< (-tol))) stop("Incorrect definition of dis")
```

```
    dis[dis<0] <- 0
    if(!inherits(comm, "matrix") && !inherits(comm, "data.frame"))
stop("Incorrect definition of comm")
    if(any(dis>1)){
        warning("dissimilarities in dis are not in the range 0-1. They have
been normalized by the maximum")
        dis <- dis/max(dis)
    }
    if(any(!colnames(comm) %in%rownames(dis))) stop("At least one species
in the matrix of abundances is missing in the matrix of dissimilarities")
    if(any(!colnames(comm) %in%colnames(dis))) stop("At least one species
in the matrix of abundances is missing in the matrix of dissimilarities")
    method <- method[1]
    if(!method%in%c("D", "Delta", "L", "M")) stop("Incorrect definition
for method")
    abundance <- abundance[1]
    if(!abundance%in%c("relative", "absolute"))
stop("Incorrect definition for abundance")
dis <- dis[colnames(comm), colnames(comm)]
    dataset <- t(comm)
    similarities <- 1-as.matrix(dis)
    total <- colSums(dataset)
    if(abundance == "relative")
        abu <- sweep(dataset, 2, total, "/")
    else abu <- dataset
    num.plot <- dim(dataset)[2]
    num.sp <- dim(dataset)[1]
    names<-list(colnames(dataset), colnames(dataset))
    dis.matrix<-matrix(0, nrow=num.plot, ncol=num.plot, dimnames=names)
    for (i in 2:num.plot) {
    for (j in 1:(i-1)) {
        mat_folk <- similarities*abu[, j]
        mat_folk2 <- similarities*abu[, i]
        Zik <- colSums(mat_folk)
        Zih <- colSums(mat_folk2)
        tabZ <- rbind.data.frame(Zik, Zih)
        garde <- apply(tabZ, 2, sum)>tol & apply(abu[,c(i,j)], 1, sum)>tol
         Zik <- Zik[garde]
        Zih <- Zih[garde]
        tabZ <- tabZ[, garde]
        wk <- (abu[, j]+abu[, i])/sum(abu[, c(i,j)])
        wk <- wk[garde]

        NUM <- (sum((wk*(abs(Zik-Zih))^q)))^(1/q)

        if(method == "D")
            DEN <- (sum((wk*(Zik+Zih)^q)))^(1/q)
        else if(method == "Delta")
            DEN <- ( sum(wk*(Zik^q))+ sum(wk*(Zih^q)) )^(1/q)
        else if(method == "M")
            DEN <- 1
        else
            DEN <- (sum(wk*(Zik^q)))^(1/q) + (sum(wk*(Zih^q)))^(1/q)
        index <- NUM/DEN
        dis.matrix[i, j] <- index
    }
    }
    dis.matrix <- dis.matrix + t(dis.matrix)
return(as.dist(dis.matrix, ...))
}
```

Appendix 3. On the relationship between the newly proposed parametric measures of community dissimilarity and a number of classical single-point measures of resemblance.

Given two plots *U* and *V*, let x_{Uj} and x_{Vj} be the abundances of species *j* $(j=1,2,...,N)$ in both plots. First, note that as x_{Uj} and x_{Vj} are both nonnegative, we have for $\alpha > 0$ $\sum_j \left| x_{Uj} - x_{Vj} \right|^\alpha \le \sum_j \left(\max_j (x_{Uj}, x_{Vj}) \right)^\alpha$ and thus $\sum_j \left| x_{Uj} - x_{Vj} \right|^\alpha \le \sum_j x_{Uj}^\alpha + \sum_j x_{Vj}^\alpha \le \sum_j \left(x_{Uj} + x_{Vj} \right)^\alpha$, and for $\alpha \ge 1$, $\sqrt[\alpha]{\sum_j \left(x_{Uj} + x_{Vj} \right)^\alpha} \le \sqrt[\alpha]{\sum_j x_{Uj}^\alpha} + \sqrt[\alpha]{\sum_j x_{Vj}^\alpha}$ (Minkowski's inequality).

Therefore, the following inequalities are satisfied: $D^\alpha \le \Delta^\alpha \le M^\alpha$ for $\alpha > 0$ and $L^\alpha \le D^\alpha$ for $\alpha \ge 1$ (for the definition of the parametric dissimilarities, see the main text).

For $\alpha = 2$, the dissimilarity index Δ^α is related to several resemblance measures published in the literature:

- Morisita-Horn similarity index (Morisita 1959; Horn 1966):

$$S_{MH} = \frac{\sum_j 2 x_{Uj} x_{Vj}}{\left[\left(\sum_j x_{Uj}^2 / X_U^2 \right) + \left(\sum_j x_{Vj}^2 / X_V^2 \right) \right] X_U X_V} \qquad \text{(A1)}$$

where $X_U = \sum_j x_{Uj}$ and $X_V = \sum_j x_{Vj}$ are the total abundances in plots *U* and *V*, respectively. The Morisita-Horn similarity index thus corresponds to

$$S_{MH} = \frac{\sum_j 2 p_{Uj} p_{Vj}}{\sum_j p_{Uj}^2 + \sum_j p_{Vj}^2} \qquad \text{(A2)}$$

where x_{Uj} and x_{Vj} are the *relative* abundances of species *j* in plots *U* and *V*, respectively.
Accordingly, a dissimilarity index can be defined as $D_{MH} = \sqrt{1 - S_{MH}}$:

$$\begin{aligned} D_{MH} &= \sqrt{1 - \frac{\sum_j 2 p_{Uj} p_{Vj}}{\sum_j p_{Uj}^2 + \sum_j p_{Vj}^2}} = \sqrt{\frac{\sum_j p_{Uj}^2 + \sum_j p_{Vj}^2 - \sum_j 2 p_{Uj} p_{Vj}}{\sum_j p_{Uj}^2 + \sum_j p_{Vj}^2}} \\ &== \sqrt{\frac{\sum_j \left(p_{Uj} - p_{Vj} \right)^2}{\sum_j p_{Uj}^2 + \sum_j p_{Vj}^2}} \end{aligned} \qquad \text{(A3)}$$

The index D_{MH} thus corresponds to index Δ^α (eq. 6b of the main text) computed from species relative abundances with $\alpha = 2$.

- Similarity ratio (e.g. Wishart 1969; Janssen 1972):

$$S_{SR} = \frac{\sum_j x_{Uj} x_{Vj}}{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 - \sum_j x_{Uj} x_{Vj}} \tag{A4}$$

A dissimilarity index can be defined as $D_{SR} = \sqrt{1 - S_{SR}}$:

$$\begin{aligned} D_{SR} &= \sqrt{1 - \frac{\sum_j x_{Uj} x_{Vj}}{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 - \sum_j x_{Uj} x_{Vj}}} = \sqrt{\frac{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 - \sum_j x_{Uj} x_{Vj} - \sum_j x_{Uj} x_{Vj}}{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 - \sum_j x_{Uj} x_{Vj}}} \\ &= \sqrt{\frac{\sum_j \left(x_{Uj} - x_{Vj} \right)^2}{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 - \sum_j x_{Uj} x_{Vj}}} \end{aligned} \tag{A5}$$

- A similarity index derived from Sokal and Sneath (1963) as shown in Pavoine and Ricotta (2014):

$$S_{SN} = \frac{\sum_j x_{Uj} x_{Vj}}{2\sum_j x_{Uj}^2 + 2\sum_j x_{Vj}^2 - 3\sum_j x_{Uj} x_{Vj}} \tag{A6}$$

with the corresponding dissimilarity index which is defined as $D_{SN} = \sqrt{1 - S_{SN}}$:

$$\begin{aligned} D_{SN} &= \sqrt{1 - \frac{\sum_j x_{Uj} x_{Vj}}{2\sum_j x_{Uj}^2 + 2\sum_j x_{Vj}^2 - 3\sum_j x_{Uj} x_{Vj}}} = \sqrt{\frac{2\sum_j x_{Uj}^2 + 2\sum_j x_{Vj}^2 - 3\sum_j x_{Uj} x_{Vj} - \sum_j x_{Uj} x_{Vj}}{2\sum_j x_{Uj}^2 + 2\sum_j x_{Vj}^2 - 3\sum_j x_{Uj} x_{Vj}}} \\ &= \sqrt{\frac{2\sum_j \left(x_{Uj} - x_{Vj} \right)^2}{2\sum_j x_{Uj}^2 + 2\sum_j x_{Vj}^2 - 3\sum_j x_{Uj} x_{Vj}}} \end{aligned} \tag{A7}$$

All these dissimilarity measures satisfy the complementarity requirement *sensu* Clarke et al. (2006). That is, the measure takes its maximum value (1) when the two plots *U* and *V* have no species in common.

In addition, starting from D^{α} (eq. 6a of the main text), for $\alpha = 2$ we can define a new dissimilarity index $D_{RP} = D^2$. The corresponding similarity index can be thus defined as:

$$\begin{aligned} S_{RP} &= 1 - D_{RP} = 1 - \frac{\sum_j \left(x_{Uj} - x_{Vj} \right)^2}{\sum_j \left(x_{Uj} + x_{Vj} \right)^2} \\ &= \frac{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 + 2\sum_j x_{Uj} x_{Vj} - \sum_j x_{Uj}^2 - \sum_j x_{Vj}^2 + 2\sum_j x_{Uj} x_{Vj}}{\sum_j \left(x_{Uj} + x_{Vj} \right)^2} \\ &= \frac{4\sum_j x_{Uj} x_{Vj}}{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 + 2\sum_j x_{Uj} x_{Vj}} \end{aligned} \tag{A8}$$

Generalizing the previous indices we obtain:

$$S\left(x_U, x_V; m\right) = \frac{\frac{m}{2}\left(2\sum_j x_{Uj}x_{Vj}\right)}{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 + \left(\frac{m}{2}-1\right)\left(2\sum_j x_{Uj}x_{Vj}\right)} \qquad \text{(A9)}$$

with

$S\left(x_U, x_V; 0.5\right) = S_{SN}$;
$S\left(x_U, x_V; 1\right) = S_{SR}$;
$S\left(x_U, x_V; 2\right) = S_{MH}$ (if relative abundances are used);
$S\left(x_U, x_V; 4\right) = S_{RP}$

and

$$D\left(x_U, x_V; m\right) = \sqrt{1 - S\left(x_U, x_V; m\right)} = \sqrt{\frac{\sum_j \left(x_{Uj} - x_{Vj}\right)^2}{\sum_j x_{Uj}^2 + \sum_j x_{Vj}^2 + \left(\frac{m}{2}-1\right)\left(2\sum_j x_{Uj}x_{Vj}\right)}} \qquad \text{(A10)}$$

where $D\left(x_U, x_V; 2\right)$ and $D\left(x_U, x_V; 4\right)$ correspond to eq. 6b and eq. 6a of the main text for $\alpha = 2$, respectively.

The parametric dissimilarities D^α, Δ^α and the other dissimilarity measures presented above can be further put into an even wider context with the general formulas:

$$S\left(x_U, x_V; m, \alpha\right) = \frac{\frac{m}{2}\left(\sum_j \left(x_{Uj} + x_{Vj}\right)^\alpha - \sum_j x_{Uj}^\alpha - \sum_j x_{Vj}^\alpha\right)}{\sum_j x_{Uj}^\alpha + \sum_j x_{Vj}^\alpha + \left(\frac{m}{2}-1\right)\left(\sum_j \left(x_{Uj} + x_{Vj}\right)^\alpha - \sum_j x_{Uj}^\alpha - \sum_j x_{Vj}^\alpha\right)} \qquad \text{(A11)}$$

$$D\left(x_U, x_V; m, \alpha\right) = \sqrt[\alpha]{\frac{\sum_j \left|x_{Uj} - x_{Vj}\right|^\alpha}{\sum_j x_{Uj}^\alpha + \sum_j x_{Vj}^\alpha + \left(\frac{m}{2}-1\right)\left(\sum_j \left(x_{Uj} + x_{Vj}\right)^\alpha - \sum_j x_{Uj}^\alpha - \sum_j x_{Vj}^\alpha\right)}} \qquad \text{(A12)}$$

$D\left(x_U, x_V; m, \alpha\right)$ satisfies the complementarity requirement as for two plots without shared species $\sum_j \left(x_{Uj} + x_{Vj}\right)^\alpha - \sum_j x_{Uj}^\alpha - \sum_j x_{Vj}^\alpha = 0$.
For $\alpha = 2$, we find again the generalizations given above ($S\left(x_U, x_V; m\right)$ and $D\left(x_U, x_V; m\right)$). Note also that $\sum_j \left(x_{Uj} + x_{Vj}\right)^\alpha - \sum_j x_{Uj}^\alpha - \sum_j x_{Vj}^\alpha = \sum_j \sum_{k=1}^{\alpha-1} C_\alpha^k x_{Uj}^k x_{Vj}^{\alpha-k}$.
In addition, $D\left(x_U, x_V; 2, \alpha\right)$ and $D\left(x_U, x_V; 4, \alpha\right)$ correspond to eq. 6b and eq. 6a of the main text, respectively.

For $m = 0.5$, we obtain a generalization of the Sokal and Sneath similarity index (presented above) and its associated generalized dissimilarity index:

$$S\left(x_U, x_V; 0.5, \alpha\right) = \frac{0.5\sum_j \sum_{k=1}^{\alpha-1} C_\alpha^k x_{Uj}^k x_{Vj}^{\alpha-k}}{2\sum_j x_{Uj}^\alpha + 2\sum_j x_{Vj}^\alpha - 3\left(0.5\sum_j \sum_{k=1}^{\alpha-1} C_\alpha^k x_{Uj}^k x_{Vj}^{\alpha-k}\right)} \quad \text{(A13)}$$

$$D\left(x_U, x_V; 0.5, \alpha\right) = \sqrt[\alpha]{\frac{2\sum_j \left|x_{Uj} - x_{Vj}\right|^\alpha}{2\sum_j x_{Uj}^\alpha + 2\sum_j x_{Vj}^\alpha - 3\left(0.5\sum_j \sum_{k=1}^{\alpha-1} C_\alpha^k x_{Uj}^k x_{Vj}^{\alpha-k}\right)}} \quad \text{(A14)}$$

Finally, for $m = 1$, we have a generalization of the similarity ratio (presented above) and its associated dissimilarity index:

$$S\left(x_U, x_V; 1, \alpha\right) = \frac{\frac{1}{2}\left(\sum_j \sum_{k=1}^{\alpha-1} C_\alpha^k x_{Uj}^k x_{Vj}^{\alpha-k}\right)}{\sum_j x_{Uj}^\alpha + \sum_j x_{Vj}^\alpha - \frac{1}{2}\left(\sum_j \sum_{k=1}^{\alpha-1} C_\alpha^k x_{Uj}^k x_{Vj}^{\alpha-k}\right)} \quad \text{(A15)}$$

$$D\left(x_U, x_V; 1, \alpha\right) = \sqrt[\alpha]{\frac{\sum_j \left|x_{Uj} - x_{Vj}\right|^\alpha}{\sum_j x_{Uj}^\alpha + \sum_j x_{Vj}^\alpha - \frac{1}{2}\left(\sum_j \sum_{k=1}^{\alpha-1} C_\alpha^k x_{Uj}^k x_{Vj}^{\alpha-k}\right)}} \quad \text{(A16)}$$

References

Clarke, K.R., Somerfield, P.J., Chapman, M.G. (2006) On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. Journal of Experimental Marine Biology and Ecology 330: 55–80.

Horn, H.S. (1966) Measurement of “overlap” in comparative ecological studies. American Naturalist 100: 419–424.

Janssen, J.G.M. (1972) Detection of some micropatterns of winter annuals in pioneer communities of dry sandy soils. Acta Botanica Neerlandica 21: 609–616.

Morisita, M. (1959) Measuring of interspecific association and similarity between communities. Memoirs of the Faculty of Science, Kyushu University Series E (Biology) 3: 65–80.

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

Sokal, R.R., Sneath, P.H.A. (1963) Principles of Numerical Taxonomy. W.H. Freeman, San Francisco.

Wishart, D. (1969) An algorithm for hierarchical classification. Biometrics 25: 165–170.