

ECOGRAPHY

Software notes

mFD: an R package to compute and illustrate the multiple facets of functional diversity

Camille Magneville, Nicolas Loiseau, Camille Albouy, Nicolas Casajus, Thomas Claverie, Arthur Escalas, Fabien Leprieur, Eva Maire, David Mouillot and Sébastien Villéger

C. Magneville (<https://orcid.org/0000-0003-0489-3822>), N. Loiseau (<https://orcid.org/0000-0002-2469-1980>), T. Claverie (<https://orcid.org/0000-0002-6258-4991>), A. Escalas (<https://orcid.org/0000-0002-6450-5716>), F. Leprieur (<https://orcid.org/0000-0001-6869-342X>), D. Mouillot (<https://orcid.org/0000-0003-0402-2605>) and S. Villéger (<https://orcid.org/0000-0002-2362-7178>) ✉ (sebastien.villeger@cnrs.fr), MARBEC, Univ. Montpellier, CNRS, Ifremer, IRD, Montpellier, France. TC also at: Centre Univ. de Formation et de Recherche de Mayotte, France. FL and DM also at: Inst. Univ. de France, Paris, France. – E. Maire (<https://orcid.org/0000-0002-1032-3394>), Lancaster Environment Centre, Lancaster Univ., Lancaster, UK. – C. Albouy (<https://orcid.org/0000-0003-1629-2389>), DECOD (Ecosystem Dynamics and Sustainability), IFREMER, INRAE, Inst. Agro – Agrocampus Ouest, Lorient, France. – N. Casajus (<https://orcid.org/0000-0002-5537-5294>), FRB-CESAB, Montpellier, France.

Ecography

44: 1–15, 2021

doi: 10.1111/ecog.05904

Subject Editor: Brody Sandel
Editor-in-Chief: Miguel Araújo
Accepted 4 November 2021



Functional diversity (FD), the diversity of organism attributes that relates to their interactions with the abiotic and biotic environment, has been increasingly used for the last two decades in ecology, biogeography and conservation. Yet, FD has many facets and their estimations are not standardized nor embedded in a single tool. *mFD* (multifaceted functional diversity) is an R package that uses matrices of species assemblages and species trait values as building blocks to compute most FD indices. *mFD* is firstly based on two functions allowing the user to summarize trait and assemblage data. Then it calculates trait-based distances between species pairs, informs the user whether species have to be clustered into functional entities and finally computes multidimensional functional space. To let the user choose the most appropriate functional space for computing multidimensional functional diversity indices, two *mFD* functions allow assessing and illustrating the quality of each functional space. Next, *mFD* provides 6 core functions to calculate 16 existing FD indices based on trait-based distances, functional entities or species position in a functional space. The *mFD* package also provides graphical functions based on the *ggplot* library to illustrate FD values through customizable and high-resolution plots of species distribution among functional entities or in a multidimensional space. All functions include internal validation processes to check for errors in data formatting which return detailed error messages. To facilitate the use of *mFD* framework, we built an associated website hosting five tutorials illustrating the use of all the functions step by step.

Keywords: alpha-diversity, beta-diversity, functional entities, functional space, functional traits, Hill numbers

Background

Functional diversity (FD) is the diversity component (Pollock et al. 2020) measuring the diversity of organism characteristics that relates to their interactions with their



www.ecography.org

© 2021 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

abiotic and biotic environment, i.e. functional traits (Petchey and Gaston 2006, Naeem et al. 2012). FD has been increasingly considered for the last two decades in all fields of ecology such as community ecology, conservation and biogeography (McGill et al. 2006, Violle et al. 2014). FD indeed allows assessing the response of species assemblages to natural or anthropogenic pressures (Mouillot et al. 2013a, Trindade-Santos et al. 2020) and unravels the effect of species on ecosystem functioning (Diaz and Cabido 2001, McLean et al. 2019). FD is multifaceted, embedding within (alpha-diversity) and between (beta-diversity) assemblages components, which gathers complementary features to describe the distribution of species trait values and hence answers various ecological questions (Mouillot et al. 2013a, Villéger et al. 2013). To measure those multiple FD facets, several frameworks (i.e. group-, distance-, dendrogram- or multidimensional-based) and indices (e.g. richness-, density-, entropy-like) have been proposed over the last two decades and some of them have been increasingly used (Villéger et al. 2008 and Laliberté and Legendre 2010 cited > 100 times each year from 2015 to 2020).

Most of these FD indices are already computable using publicly available R functions that are disseminated in several packages. For example, the *dbFD* function of the *FD* package (Laliberté and Legendre 2010, Laliberté et al. 2014) computes seven multidimensional FD indices based on Botta-Dukát (2005), Petchey and Gaston (2006), Lavorel et al. (2008), Villéger et al. (2008), Laliberté and Legendre (2010). However, this package computes the multidimensional space (or dendrogram) and FD indices within the same function while they correspond to independent steps (Mouchet et al. 2008, Maire et al. 2015), thus potentially inducing a misunderstanding of the outputs or limiting the reproducibility of analysis (Villéger et al. 2017).

The *betapart* package (Baselga and Orme 2012, Baselga et al. 2021) contains a set of functions to compute functional beta-diversity indices based on the overlap of species assemblages in a multidimensional space, but it does not permit the graphical representation of indices.

Two other R packages *hypervolume* and *BAT* (Blonder et al. 2018, Blonder and Harris 2019, Cardoso et al. 2020) propose a density-based FD framework using organism coordinates in a multidimensional space and the selection of bandwidth for computing kernels around observed organisms.

Lastly, some R functions have been provided as publication supplementary materials or on online repositories, including functions not only to select the best multidimensional space in which FD indices are computed (Maire et al. 2015) but also to compute group-based indices (Mouillot et al. 2014) or entropy-like indices (Chao and Ricotta 2019). Embedding these functions in a single and comprehensive R package is needed not only to facilitate analyses and ease interpretation and visualization of the results but also to increase their reproducibility.

Therefore, we designed the *mFD* R package to provide a wide set of ‘user-friendly’ functions covering all the steps of FD-based analyses, from checking of input data (Supporting information), through the identification of functional groups,

computation of trait-based distances, and building the functional space required to compute 16 functional diversity indices, including the graphical representation of key outputs (Fig. 1). Graphic aesthetics can be customized by the user and outputs can be displayed or saved as high-resolution files. *mFD* improves diagnostics, error reporting, and data visualization in a standardized way, and its website (<<https://cmlmagneville.github.io/mFD/>>) provides a set of tutorials to help newcomers in the field with their analyses. *mFD* is currently available through Github (<<https://github.com/CmlMagneville/mFD>>) and is available on CRAN (<<https://cran.r-project.org/web/packages/mFD/index.html>>).

In this paper, we first described the methods used and the features of the *mFD* package and then explained step by step the process of computing and plotting FD indices with the *mFD* package. Lastly, we illustrated the most basic workflow of the *mFD* package using a data set gathering fruits in several baskets (Box 1).

Methods and features

Functional diversity can be measured at various ecological grains, from local communities (e.g. 1 m² quadrat; (De Bello et al. 2009)) to continental or oceanic flora/fauna, sampled through various time scales. In all cases, quantifying FD first requires measuring a set of functional traits on all the species of interest. Functional traits should be, by definition, measurable at the individual level (Violle et al. 2007), but in practice, trait values are often aggregated at the species level. The selection of traits depends on the goal of the study and has thus to be carried out by experts of the studied taxa and ecosystems (Violle et al. 2007).

Most of the FD indices are designed to account for the dominance of organisms through quantitative weights. Depending on the type of organisms studied and on the aim of the study, weighting of organisms dominance could be based on biomass or abundance per unit of area, or percentage of ground coverage (e.g. for plants or corals, (McWilliam et al. 2020)). When only species occurrence data are available (as often in functional biogeography), weight-based FD indices are computed by considering that all species present in a given assemblage have the same weight (i.e. 1/number of species).

Hereafter, for simplicity, we refer to a set of ‘assemblages’ where ‘weights’ and ‘traits’ have been measured at the ‘species’ level but all statements remain true for other sampling approaches (e.g. traits and weights measured on individuals, clades above the species level or trophic groups). Species pool refers to all the species present in the set of assemblages.

mFD package allows the user to measure FD indices using three different frameworks based on either 1) groups of species; 2) pairwise trait-based distances between species; or 3) species coordinates in a multidimensional space (Fig. 1). It relies on a species-assemblages matrix with rows being assemblages and columns being species featuring species occurrence, densities, or biomass and a species-traits data frame with rows being species and columns being traits.

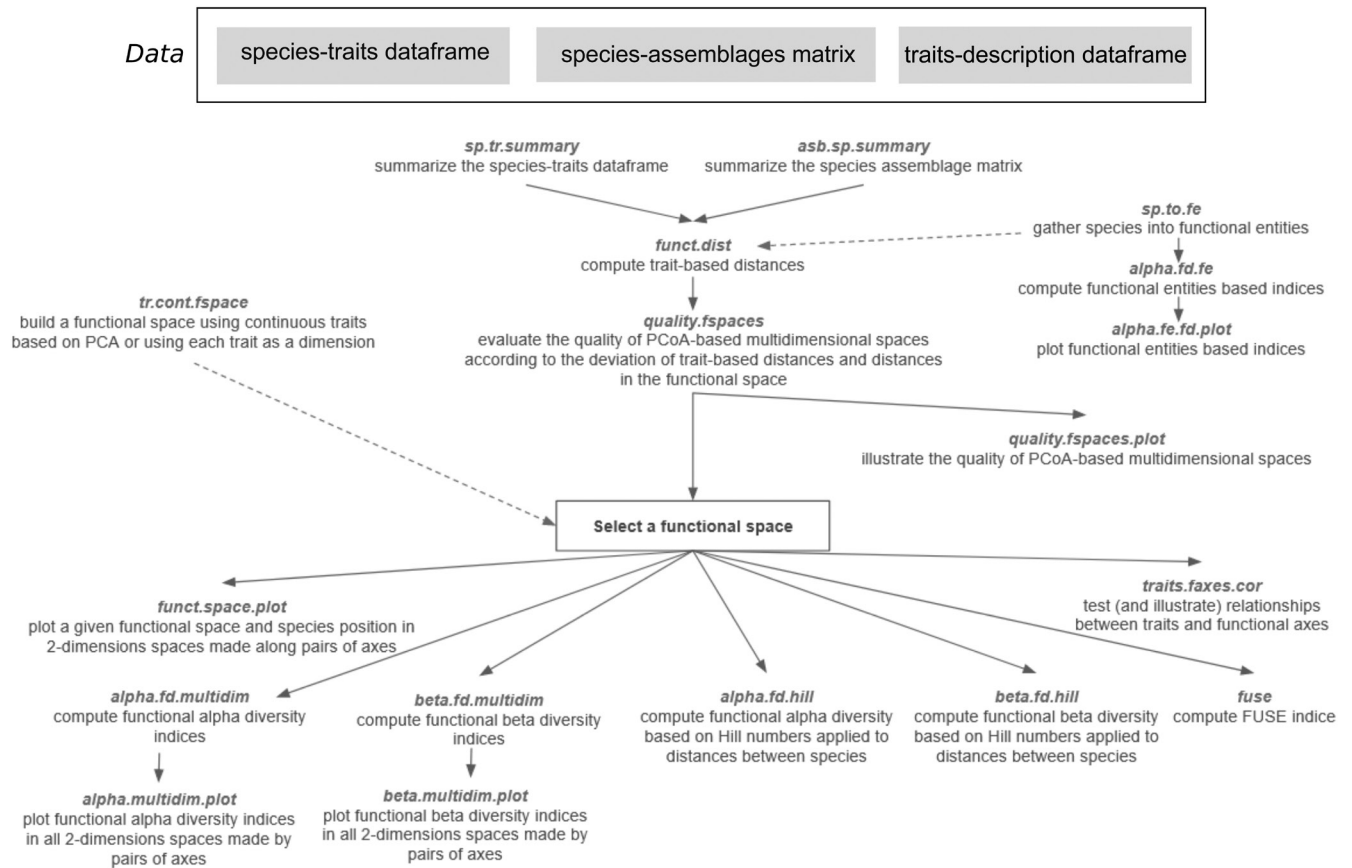


Figure 1. Workflow of the 18 key functions of the *mFD* package.

The list of key functions and their connections with main inputs/outputs are summed up in Fig. 1 below. The names of the functions present on the graph reflect the terminology used in our GPL-2 licensed R package.

Measuring FD based on functional entities

When only categorical and ordinal traits are used to describe species, there is a finite number of combinations of trait values. It is hence likely that some species from the pool will share the same trait values and can thus be clustered into functional entities (FE; Mouillot et al. 2014). This clustering can be done by applying the *sp.to.fe* function to the species-trait data frame. This function provides not only names and numbers of species in each FE but also the number of unique values per trait and the potential maximum number of FEs given the number and coding of traits.

Clustering species from the pool into functional entities is mandatory to compute the FD indices proposed by Mouillot et al. (2014) in each assemblage. First, functional entity richness (FEr) is simply the number of FE present. Second, functional redundancy (FRed) is computed as the ratio between species richness and FEr and represents the average number of species in FE present in a given assemblage. However, this average value could be similar in two

contrasting cases, that is with both even and skewed distributions of species number among FE. Therefore, Mouillot et al. (2014) proposed to measure functional over-redundancy (FOR) based on the proportion of the species that are in the FE with more species than average redundancy. FOR is low, close to 0, when all FE have the same number of species and close to 1 when most species are packed in the richest functional entity. Finally, functional vulnerability (FVuln) is measured as the proportion of FE in a given assemblage that is represented by a single species. The function *alpha.fd.fe* computes these four indices for each assemblage using as inputs the output of the *sp.to.fe* function and a matrix with species occurrences in assemblages.

The function *alpha.fd.fe.plot* is designed to illustrate how species are clustered in functional entities. Using the outputs of the *alpha.fd.fe* function, this function returns a barplot with FE ranked in decreasing order of species richness for each assemblage. Moreover, the user can select which of the indices described above could then be illustrated on this barplot, with *FRed* shown as a horizontal line, *FVuln* depicted as a horizontal double-arrow over the vulnerable entities (i.e. FE with a single species) which are further identified using colors, and *FOR* depicted by coloring the top of the bars corresponding to FE with more species than average (Supporting information).

Computing trait-based distance between species

The first step for computing the functional diversity indices that are not based on functional entities is to calculate the functional distance for each pair of species. Here, we provide the function ***funct.dist*** which computes pairwise trait-based distances for all sets of trait types (i.e. categorical and continuous traits can be mixed). In addition to the species-traits data frame, ***funct.dist*** requires a table describing the type of each trait and the chosen distance metric. Three distance metrics can be computed through the *mFD* package: the Euclidean distance, the Gower distance, and the Gower distance applied to fuzzy traits (Gower 1971, De Bello et al. 2021). The Euclidean distance should be used only when all traits are continuous. In this case, all traits must be standardized beforehand, either to a null mean and/or a standard deviation of one, or to the same range (Legendre and Legendre 2012). For matrices containing only categorical traits or several types of traits, the ***funct.dist*** computes Gower distance with the function *gawdis()* from the package *gawdis*. ***funct.dist*** can also be applied to functional traits coded as fuzzy groups (e.g. diet described with a proportion of several resources), and it allows the user to give different weights to each trait (e.g. to ensure that biological functions described with varying numbers of traits have eventually equal contribution to the distance metric). The Gower distance is able to deal with missing trait values (coded as NA in R objects) by accounting only for traits measured on both species of each pair. However, this feature could lead to situations where two species are identical (Gower=0) while having different Gower distances to a third species, thus influencing FD indices values. Hence, if the species-traits data frame contains NA, ***funct.dist*** returns a warning (this check could be overridden by the user), and it is recommended to infer missing traits values (Johnson et al. 2020) or delete species with missing trait values. For more information about distances, the user can refer to the *mFD* general workflow tutorial (<https://cmlmagneville.github.io/mFD/articles/mFD_general_workflow.html>).

Distance-based functional diversity indices

The first index describing the diversity of functional trait values within an assemblage was based on the pairwise trait-based distance between species present in an assemblage (functional attribute diversity, FAD (Walker et al. 1999)). The Rao's quadratic index (Q) went one step further by accounting for both species distance as well as species weight (i.e. relative biomass or abundance (Rao 1982, Botta-Dukát 2005)) and was later adapted by Ricotta and Szeidl (2009) to match the Hill's number framework (i.e. diversity expressed as an equivalent number of species) as proposed by Jost (2006) for taxonomic diversity. However, the standardization of Rao's Q by the maximum distance between two species from the regional pool makes the index poorly sensitive to differences in the functional structure of given assemblages, i.e. most assemblages have very low diversity values (Chao et al. 2019).

This weakness has been overcome by Chao et al. (2019), who proposed a general framework for measuring FD using the Hill numbers approach with two parameters: *q* defines the relative importance given to species weights compared to species distances, and *tau* defines the threshold level applied to functional distances between species to identify functionally indistinct sets of species. The ***alpha.fd.hill*** function uses species distances and assemblage data to compute indices for varying values of *tau*, and *q*. *tau* could be set to minimum, mean (default value), or maximum computed over all pairwise distances in the species pool. *q* could be set to 0, 1, or 2, to give increasing importance to species weights compared to trait-based distances. Setting *q*=2 and *tau*='maximum' computes Rao's Q from Ricotta and Szeidl (2009) while setting *tau*='minimum' yields taxonomic diversity indices (i.e. *q*=0 is species richness, *q*=1 is exponential of Shannon index) that could be compared to functional diversity (Chao et al. 2019).

The Hill numbers framework has the advantage of remaining valid with all distance metrics, including the Gower distance computed on mixed trait values as well as the Euclidean distance computed on continuous trait values or on species coordinates in a multidimensional functional space (see next sections below).

This framework also allows a multiplicative partitioning of diversity into gamma, alpha, and beta-components (Chao et al. 2019) following either Sorensen- or Jaccard-like formulas. The ***beta.fd.hill*** function uses the same inputs as ***alpha.fd.hill*** and allows computing Hill number-based functional beta-diversity indices with varying values of *q* and *tau* for all pairs of species assemblages.

Building functional space from continuous traits

Most FD indices assessments have been based on the multivariate indices proposed by Villéger et al. (2008, 2013) and Laliberté and Legendre (2010) that account for the distribution of species in a Euclidean space. When all functional traits are continuous, it is possible to build such a space where each axis is a trait (Villéger et al. 2008). When traits have different units, or contrasted distributions, it is necessary to scale their values to ensure they have similar range, mean, and/or variance. The ***tr.cont.fspace*** function performs this trait scaling according to a user-defined scaling method and if required, computes a principal component analysis (PCA) on the Euclidean distance between species (internally computed using ***funct.dist***).

The main output of the ***tr.cont.fspace*** function is a matrix of species coordinates along a set of axes (PCs or traits), and if required by the user, Pearson correlation coefficients between traits, trait-based distance between species, and eigenvalues of PCA axes are also returned.

Assessing the quality of functional space

When at least one trait is not continuous, building a Euclidean space requires computing a principal coordinates

analysis (PCoA) based on distances between species pairs and using its principal components to build the Euclidean space. Several studies have demonstrated that the number of axes selected markedly influences FD patterns (Villéger et al. 2011, 2017, Maire et al. 2015). More specifically, Maire et al. (2015) proposed a framework to measure the quality of all possible functional spaces. It is based on the lowest deviation between the original trait-based distances and the final Euclidean distances in the functional space.

Here, we provide the function *quality.fspaces* that extends this framework by allowing the user to measure the quality of spaces with a combination of two criteria. The first criterion determines whether distances in the functional space are scaled or not so that their maximum value equals the one for trait-based distance. This procedure (applied in Maire et al. 2015) hypothesizes that since most FD indices are unitless, these are the ratios between distances, not the distances themselves, that have to be faithfully represented.

The second criterion determines how to weigh a deviation of distance. Maire et al. (2015) proposed to square transform the values to strongly penalize large deviations. However, this mean squared-deviation index does not have the same unit as initial distances, so we suggest using the square-root of mean squared deviations (RMSD) to have the same unit as initial distances thus getting a metric easier to interpret. In addition, the function allows the user to measure the mean of absolute deviations (MAD) that better reflects the inaccuracy of species distances in the Euclidean space that will then affect the FD computation.

The *quality.fspaces* function thus computes up to four quality metrics based on the combination of the two criteria (absolute or squared deviation computed on raw or scaled distances) for all spaces built using increasing numbers of PCoA axes. The maximum number of PCoA axes to consider is set a priori by the user, but it is eventually constrained by the number of PCoA axes that present positive eigenvalues. Indeed, we recommend not applying square-root transformation before computing PCoA because distances in the Euclidean space will end up reflecting transformed trait-based distances that are not linearly related to the ones measured (see the ‘Compute and Interpret Quality of Functional Spaces’ tutorial on *mFD* website (<https://cmlmagneville.github.io/mFD/articles/Compute_and_interpret_quality_of_functional_spaces.html>)).

Besides multidimensional spaces, the *quality.fspaces* function allows the user to compute the quality of a dendrogram (clustering algorithm being specified by user). However, most often dendrograms are biased representations of the trait-based distance between species with some pairs of species with close trait values being distantly related in the dendrogram (see ‘Compute and Interpret Quality of Functional Spaces’ tutorial on *mFD* website (<https://cmlmagneville.github.io/mFD/articles/Compute_and_interpret_quality_of_functional_spaces.html>)). Thus, dendrogram-based FD indices (Petchey and Gaston 2002, Cardoso et al. 2014), available in other R packages such as *FD* (Laliberté and Legendre 2010, Laliberté et al. 2014), *BAT* (Cardoso et al. 2020), and *adiv*

(Pavoine 2020, 2021), could lead to erroneous ecological conclusions (Loiseau et al. 2017, Villéger et al. 2017) and should be used after cautionary checking of dendrogram quality.

Quality of functional spaces can be illustrated using the *quality.fspaces.plot* function that takes as input the output of *quality.fspaces* and returns a figure where for each space, three panels show 1) the relation between trait-based distances and distances in the space; 2) difference between the former and latter distances; and 3) deviation between those distances as accounted for in the select quality metrics (i.e. raw or scaled distance in the functional space and absolute or squared deviation). Plot panels are produced with *ggplot2* library and arranged into a single figure with the *patchwork* library, and colours used to highlight gradients in deviation are customizable by users.

These two functions allow the user to make a parsimonious selection of the axes number keeping enough axes to have a space that faithfully represents the original trait-based distance while allowing computation of FD indices and an easier graphical interpretation of their patterns. Indeed, convex hull-based indices require a space with less axes than the number of species number, and their computation time increases with the number of axes (to the point that functional beta-diversity indices are hardly computable in more than five dimensions). So if, for example, the best space is the one with six axes while the quality index of the 4D and 5D spaces are close, keeping the 4D space will be a pragmatic choice.

Illustrating the distribution of species and traits values in the functional space

Once the ‘best’ functional space has been selected, and species coordinates along each axis are extracted as an output of *quality.fspaces*, *mFD* package provides two dedicated functions for further analyses.

The *funct.space.plot* function is designed to illustrate the position of all species from the pool along functional axes (all pairs or a subset). The user selects the colors and shapes of points to represent species as well as to display additional features, such as the convex hull shaping all species and names of (some) species.

In addition, the *traits.faxes.cor* function is designed to test relationships between values of each trait and positions along each functional axis using r^2 and p-value from linear regression and η^2 and p-value from Kruskal–Wallis test for continuous and non-continuous traits, respectively. An option of the function allows the user to illustrate those relationships in a multi-panel plot (trait \times axes) with scatterplots and boxplots for continuous and non-continuous traits, respectively.

Computing multidimensional functional alpha-diversity

The *alpha.fd.multidim* function computes nine multidimensional functional alpha-diversity indices (Table 1).

Table 1. Multidimensional functional alpha-diversity indices computed by the *alpha.fd.multidim* function.

Name	Acronym	Definition	Details
Functional richness (Cornwell et al. 2006, Villéger et al. 2008)	FRic	The volume of the convex hull shaping the species present in the assemblage	Computed using the 'convhulln' function of the 'geometry' package. Computed only if the number of species (i.e. points) is strictly higher than the number of functional axes. If points are coplanar, the convex hull can not be computed and the function returns NA.
Functional identity (Garnier et al. 2004, Mouillot et al. 2013a)	Flde	The weighted average position of species of the assemblage along each axis	None
Functional dispersion (Laliberté and Legendre 2010)	FDis	The weighted deviation to center of gravity (i.e. defined by the Flde values) of species in the assemblage	Flde is always computed with FDis.
Functional divergence (Villéger et al. 2008)	FDiv	The deviation of biomass-density to the center of gravity of the vertices shaping the convex hull of the studied assemblage	FDiv requires computing first vertices of the convex hull so it has the same constraints as FRic (see above).
Functional evenness (Villéger et al. 2008)	FEve	The regularity of biomass-density distribution along the minimum spanning tree (i.e. the tree linking all species of the assemblage with the lowest cumulative branch length) for the studied assemblage	There must be at least three species to compute the FEve index. The minimum spanning tree is computed using the 'mst' function of the 'ape' package.
Functional originality (Mouillot et al. 2013a)	FOri	The weighted mean distance to the nearest species from the global species pool	None
Functional specialisation (Bellwood et al. 2006, Mouillot et al. 2013a)	FSpe	The weighted mean distance to the centroid of the global species pool (i.e. center of the functional space)	None
Functional mean pairwise distance (Weiher et al. 1998)	FMPD	The mean weighted distance between all pairs of species	None
Functional mean nearest neighbor distance (Weiher et al. 1998)	FNND	The weighted distance to the nearest neighbor within the assemblage	None

Required inputs are the matrix of species coordinates in the multidimensional space built using *quality.fspaces* or *tr.cont.fspace* functions and the matrix with species weights in studied assemblages. Users can select a subset of the nine FD indices to be computed (see package functions help for details).

An optional logical argument allows the user to scale a posteriori the values of all FD indices but FEve and FDiv (whose ranges are already constrained to 0–1) by dividing raw values by the maximum value possible given coordinates of species present in the pool of all assemblages.

The user *alpha.fd.multidim* returns a table with values of FD indices (columns) for all assemblages (rows). If required by the user, details about indices computation are also returned as a list of objects (e.g. distances to centroid/nearest neighbor, names of species being vertices). These outputs could help interpret FD indices and are required for illustrating them.

The multidimensional framework described above could be applied to functional entities using their traits values and their occurrences (or weight computed as a sum of weights of species belonging to each FE) in assemblages.

Illustrating multidimensional functional diversity

The *alpha.multidim.plot* function is designed to illustrate FD indices for one or two assemblages, using the distribution of species (and of their weights) along axes of the functional

space with key geometric features accounted for by each index. For example, functional richness (FRic) is illustrated by a projected convex hull filled by the species assemblage, functional evenness (FEve) is illustrated as the minimum spanning tree linking species of the assemblage, and functional divergence (FDiv) is illustrated through the gravity center of the vertices of the convex hull.

The function requires output from the *alpha.fd.multidim* function in addition to matrices containing species coordinates and species weights in assemblages. The user can customize classic graphical parameters such as point or line sizes and point or line colours. A list of ggplot figures is returned for each index and each (pair of) assemblage(s), and as with other graphical functions, figures can be locally saved as jpeg files.

The *mFD* package also features graphical functions returning *ggplot* layers for each of these indexes allowing users to draw more complex graphs, for example, with more than two assemblages or with several indices on the same plot (Supporting information).

Computing and illustrating multidimensional functional beta-diversity

The *beta.fd.multidim* function computes the pairwise functional beta-diversity indices proposed by Villéger et al. (2013) following the framework proposed by Baselga (2012) for decomposing Jaccard and Sorensen taxonomic beta-diversity

indices into their respective ‘turnover’ and ‘nestedness-resultant’ additive components. More precisely, functional beta-diversity indices are computed for a pair of assemblages based on the overlap between the convex hulls shaping their respective species. Functional beta-diversity decreases when the volume of the intersection between the two convex-hulls increases relative to the volume of the union. The functional turnover component accounts for the proportion of functional richness represented by the functionally poorest assemblage that is not shared, and it is thus independent of the difference in functional richness between assemblages (Villéger et al. 2013).

The ***beta.fd.multidim*** function takes as inputs the matrix of species coordinates, the matrix of species occurrences in assemblages, and a character string referring to the family of indices to compute (Jaccard and/or Sorensen). ***beta.fd.multidim*** uses functions from the *betapart* package (ver. 1.2) which allows internal parallel computing.

The ***beta.fd.multidim*** function returns a list of distance objects with values for all pairs of assemblages. The function also returns a list containing a vector with the volume of the convex hull shaping each assemblage (i.e. *FRic*) and a list of vectors with names of species being vertices of the convex hull per assemblage.

The ***beta.multidim.plot*** function is designed to illustrate functional beta-diversity indices by plotting the overlap between the convex hulls shaping two given species assemblages projected along pairs of functional space axes. It takes as inputs the list returned by the ***beta.fd.multidim*** function, a vector containing the name of the two assemblages to represent, a vector with names of axes to plot, and a character string referring to the values of beta-diversity indices to print (i.e. Jaccard or Sorensen family). The names and range of functional axes as well as shapes, colours, and sizes of points could be customized by the user.

Computing the functionally unique, specialized, and endangered indices

Beyond FD indices estimated at the scale of species assemblages, the field of functional ecology also has a long history of assessing species-specific indices in relation to their trait values (Olden et al. 2006, Violle et al. 2017).

Recently, Pimiento et al. (2020) introduced the FUSE (functionally unique, specialized, and endangered) index that combines functional uniqueness, specialization, and global endangerment to identify threatened species of particular importance for functional diversity. The *mFD* package provides for the first time a function to calculate this index. The ***fuse*** function allows calculating the FUSE index as well as several metrics at the species level describing their functional specialization and uniqueness (Moullot et al. 2013b). It takes as inputs 1) the functional distance between each pair of species (*dist* object) provided by the ***funct.dist*** function; 2) a data frame with the coordinates of the species on a multidimensional space based on a selected number of axes derived from a principal coordinate analysis (PCoA) that can be retrieved through ***quality.fspaces*** or ***tr.cont.fspace*** functions; and 3) the level of endangerment expressed as either ranking values or IUCN extinction probabilities in a numerical vector. The number of neighbouring species considered to calculate species functional uniqueness (or originality) is fixed to five by default, as in Pimiento et al. (2020), but the user can define different numbers of nearest neighbours to consider with a minimum of one species. The ***fuse*** function returns a data frame with species in rows and the five metrics used to characterize species’ contributions to functional diversity and its combination with the level of global endangerment in columns.

Box 1. Example

We illustrate the most common workflow of the *mFD* package using a data set gathering 10 baskets made of 8 types of fruits (from a pool of 25) with varying biomass (*fruits_baskets* data frame, Supporting information). Each fruit is characterized by 5 traits (*fruits_traits* data frame) whose types are described in the *fruits_traits_cat* data frame (Table 2).

The *mFD* package is based on three main objects: a matrix describing the distribution of species into assemblages, the *fruits_baskets* matrix in this example, a first data frame containing the traits values for each species, the *fruits_traits* data frame here, and a second data frame with details about each trait, the *fruits_traits_cat* data frame.

The first step after importing these three objects in R is to check their content using the summary functions:

sp.tr.summary function summarizes the *fruits_traits* data frame and ***asb.sp.summary*** function summarizes the *fruits_baskets* matrix, which returns, among others, the species richness per baskets.

Table 2. Table ‘*fruits_traits_cat*’ summarizing traits used in the example.

Trait name	Trait measurement	Trait type	Number of classes	Classes code	Unit
Size	Maximal diameter	Ordinal	5	0–1; 1–3; 3–5; 5–10; 10–20	cm
Plant	Growth form	Categorical	4	Tree; shrub; vine; forb	NA
Climate	Climatic niche	Ordinal	3	Temperate; subtropical; tropical	NA
Seed	Seed type	Ordinal	3	None; pip; pit	NA
Sugar	Sugar	Continuous	NA	NA	g kg ⁻¹

Functional distance between all species from the 10 baskets is computed using the *funct.dist* function. Here, we applied the Gower distance since traits were not all continuous, and we scaled center continuous traits according to their mean.

```
sp_dist_fruits <- mFD::funct.dist(
  sp_tr = fruits_traits,
  tr_cat = fruits_traits_cat,
  metric = "gower")
```

Then, we assessed the quality of PCoA-based functional spaces with up to 10 dimensions and of the UPGMA dendrogram. For that, we used the *quality.fspaces* function and computed the *mad* index (mean absolute deviation):

```
fspaces_quality_fruits <- mFD::quality.fspaces(
  sp_dist = sp_dist_fruits,
  maxdim_pcoa = 10,
  deviation_weighting = "absolute",
  fdendro = "average")
```

Here, according to the *mad* index, the best functional space is the one with four dimensions as it has the lowest index value:

```
##          mad
## pcoa_1d    0.150
## pcoa_2d    0.073
## pcoa_3d    0.047
## pcoa_4d    0.040
## pcoa_5d    0.049
## pcoa_6d    0.055
## pcoa_7d    0.060
## pcoa_8d    0.064
## pcoa_9d    0.065
## pcoa_10d   0.065
## tree_average 0.082
```

The quality of functional spaces built with up to six PC axes and UPGMA dendrogram are illustrated using the *quality.fspaces.plot* function (Fig. 2):

```
mFD::quality.fspaces.plot(
  fspace_quality = fspace_quality_fruits,
  quality_metric = "mad",
  fspace_plot = c("tree_average", "pcoa_2d", "pcoa_3d",
    "pcoa_4d", "pcoa_5d", "pcoa_6d"))
```

This figure illustrates the distortion of distances on the dendrogram and in 2D spaces. As the 4D space has the best quality, we hereafter consider the coordinates of species along these four axes for computing FD indices. The matrix *sp_faxes_coord* is one of the outputs of the *quality.fspaces* function:

```
sp_faxes_coord_fruits <- fspace_quality_fruits$details_fspaces$sp_pc_coord
```

Positions of species in the 4D multidimensional functional spaces are plotted using the *funct.space.plot* function.

```
mFD::funct.space.plot(
  sp_faxes_coord = sp_faxes_coord_fruits[, c("PC1", "PC2", "PC3", "PC4)],
  faxes = NULL,
  faxes_nm = NULL)
```

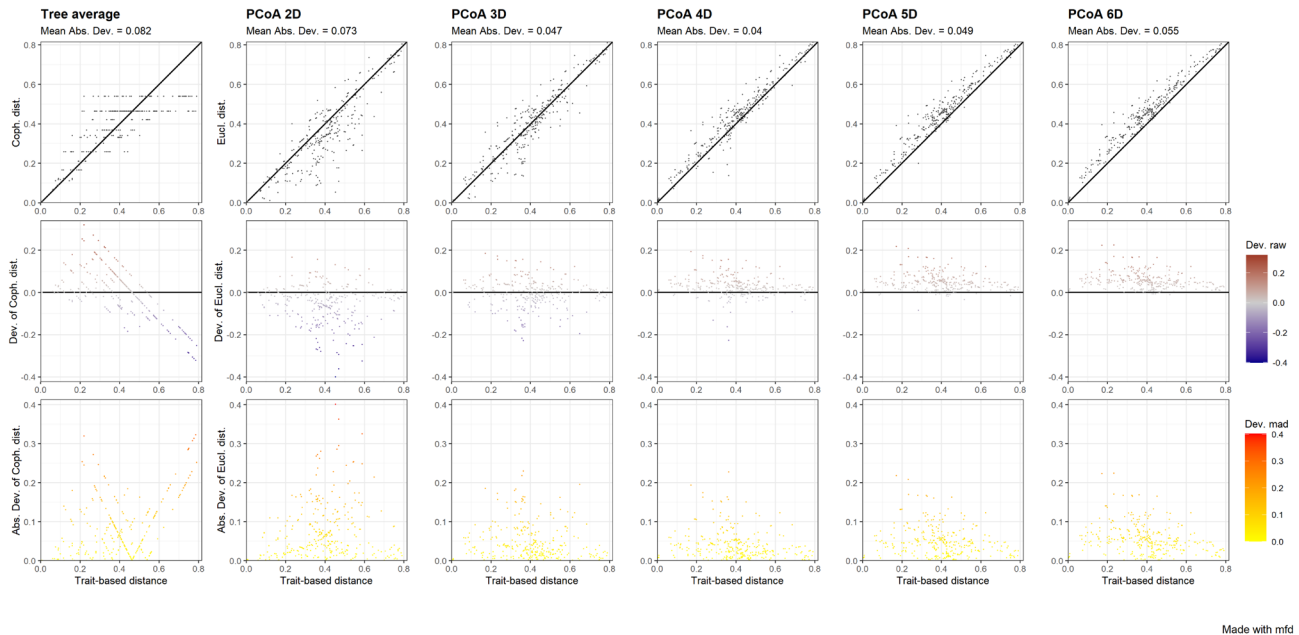



Figure 2. Illustrating the quality of functional spaces and UPGMA dendrogram with a single plot with three panels per space. Each column represents a functional space and the x-axis of all panels represents trait-based distances. The y-axis is different for each row: on the first row, it represents species functional distances in the multidimensional space; on the second row, it shows the raw deviation of species distances in the functional space compared to trait-based distances; and on the third row, it shows the absolute or squared deviation of the ('scaled') distance in the functional space.

In this plot (Fig. 3), fruits are represented either in purple if they are vertices of the convex hull shaping the pool of 25 fruits or in green if they are not vertices.

Using the *traits.faxes.cor* function, correlations between traits and axes of the functional space are computed:

```
fruits_tr_faxes <- mFD::traits.faxes.cor(
  sp_tr          = fruits_traits,
  sp_faxes_coord = sp_faxes_coord_fruits[, c("PC1", "PC2", "PC3",
                                              "PC4")],
  plot           = TRUE)
```

We can thus see on the table below that PC1 is mostly driven by climate, plant type, and size with the weaker influence of seed ($\eta^2 < 0.25$). PC2 is mostly driven by seed and plant type. PC3 is driven by only one trait, size. And finally, PC4 is driven by sugar with a weaker influence of plant type.

##	trait	axis	test	stat	value	p.value
## 1	Size	PC1	Kruskal-Wallis	eta2	0.308	0.0377
## 3	Size	PC3	Kruskal-Wallis	eta2	0.326	0.0325
## 5	Plant	PC1	Kruskal-Wallis	eta2	0.471	0.0049
## 6	Plant	PC2	Kruskal-Wallis	eta2	0.382	0.0116
## 8	Plant	PC4	Kruskal-Wallis	eta2	0.264	0.0360
## 9	Climate	PC1	Kruskal-Wallis	eta2	0.731	0.0001
## 13	Seed	PC1	Kruskal-Wallis	eta2	0.201	0.0402
## 14	Seed	PC2	Kruskal-Wallis	eta2	0.593	0.0005
## 20	Sugar	PC4	Linear Model	r2	0.682	0.0000

Position of species along pairs of functional axes

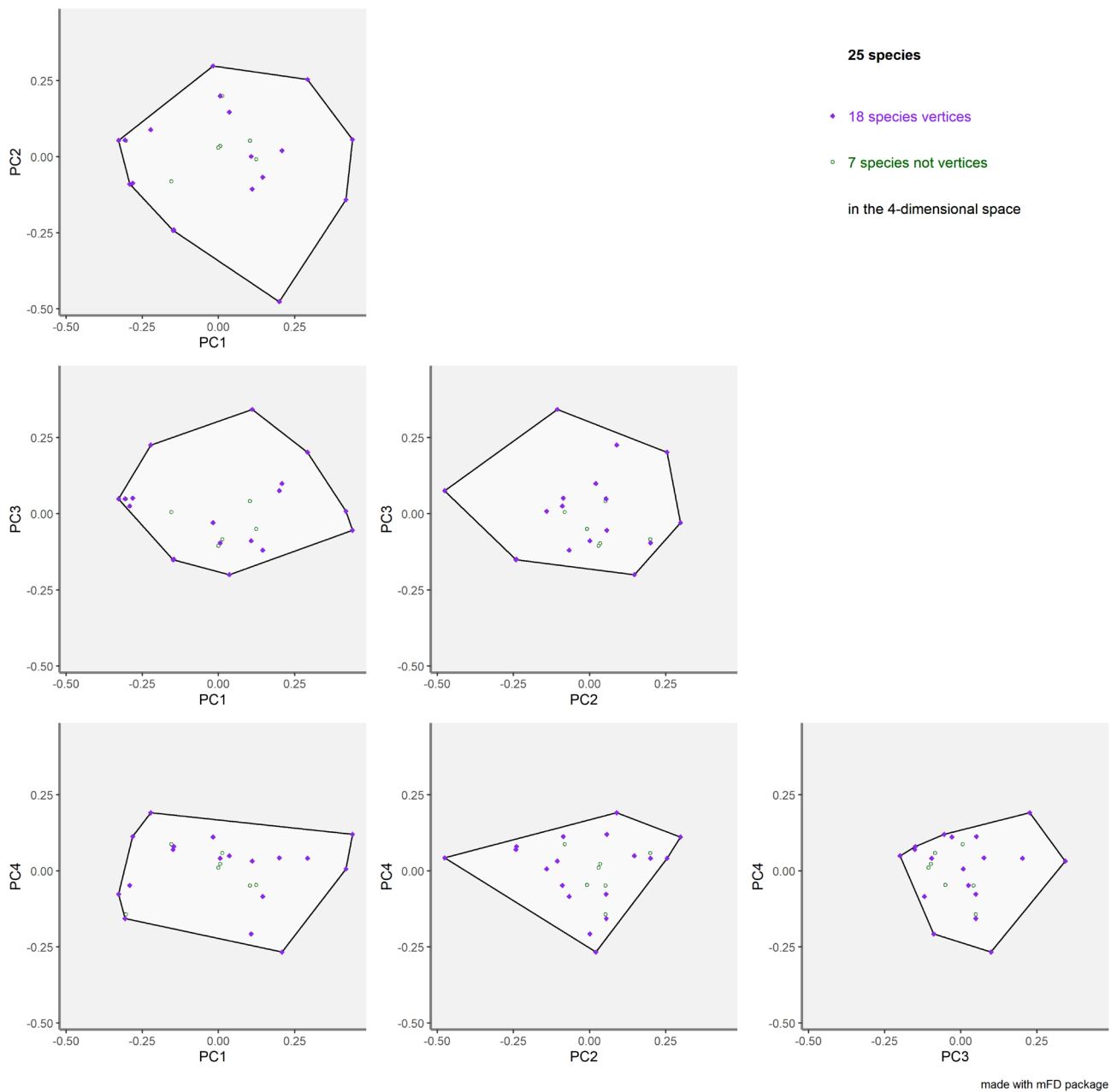


Figure 3. Distribution of species in four-dimensional space.

Relation between traits and PCoA axes can then be plotted using the same function (Supporting information).

Then three functional diversity indices are computed in the 4D functional space for all fruits baskets using the *alpha.fd.multidim* function:

```
alpha_fd_indices_fruits <- mFD::alpha.fd.multidim(
  sp_faxes_coord = sp_faxes_coord_fruits[, c("PC1", "PC2", "PC3", "PC4")],
  asb_sp_w       = baskets_fruits_weights,
  ind_vect       = c("fric", "feve", "fdiv"),
  scaling        = TRUE,
  check_input    = TRUE,
  details_returned = TRUE)
```

Species richness, functional richness, functional evenness, and functional divergence values per fruits baskets are summarized in the table below:

	sp_richn	fric	feve	fdiv
basket_1	8	0.163	0.565	0.551
basket_2	8	0.163	0.755	0.806
basket_3	8	0.163	0.806	0.809
basket_4	8	0.008	0.760	0.608
basket_5	8	0.008	0.843	0.729
basket_6	8	0.148	0.815	0.894
basket_7	8	0.148	0.827	0.899
basket_8	8	0.036	0.566	0.711
basket_9	8	0.036	0.676	0.779
basket_10	8	0.026	0.660	0.774

All computed indices are then plotted using the *alpha.multidim.plot* function for the basket 1 and basket 5:

```
fruits_plots_alpha <- mFD::alpha.multidim.plot(
  output_alpha_fd_multidim = alpha_fd_indices_fruits,
  plot_asb_nm              = c("basket_1", "basket_5"),
  ind_nm                   = c("fric", "feve", "fdiv"),
  faxes                    = NULL,
  faxes_nm                 = NULL)
```

In Fig. 4, functional richness is plotted along the six pairs of functional axes making the 4D space. The white polygon reflects the global pool of fruits, the grey crosses represent distribution of the fruits in the functional space, green points are fruits from basket 1, and yellow points are fruits from basket 5. The green and yellow surfaces respectively show a two-dimensional projection of the multidimensional convex hull shaped by the fruits from basket 1 and basket 5.

Beta-diversity indices are computed using the *beta.fd.multidim* function with Jaccard family indices here:

```
beta_fd_indices_fruits <- mFD::beta.fd.multidim(
  sp_faxes_coord = sp_faxes_coord_fruits[, c("PC1", "PC2", "PC3", "PC4")],
  asb_sp_occ      = asb_sp_fruits_occ,
  beta_family     = c("Jaccard"))
```

Turnover and nestedness-resultant components of dissimilarity indices are shown below only for pairs involving basket 1.

	x1	x2	jac_diss	jac_turn	jac_nest
basket_1	basket_2		0.000	0.000	0.000
basket_1	basket_3		0.000	0.000	0.000
basket_1	basket_4		0.965	0.432	0.533
basket_1	basket_5		0.965	0.432	0.533
basket_1	basket_6		0.870	0.863	0.007
basket_1	basket_7		0.870	0.863	0.007
basket_1	basket_8		0.980	0.943	0.037
basket_1	basket_9		0.980	0.943	0.037
basket_1	basket_10		0.915	0.599	0.316

Beta indices are then plotted using the *beta.multidim.plot* function for the basket 1 and basket 5:

```
beta_plot_fruits <- mFD::beta.multidim.plot(
  output_beta_fd_multidim = beta_fd_indices_fruits,
  plot_asb_nm              = c("basket_1", "basket_5"),
  beta_family              = c("Jaccard"),
  plot_sp_nm               = c("apple", "lemon", "pear"),
  faxes                    = paste0("PC", 1:4),
  check_input              = TRUE)
```

Functional Richness of 'basket_1' and 'basket_5'

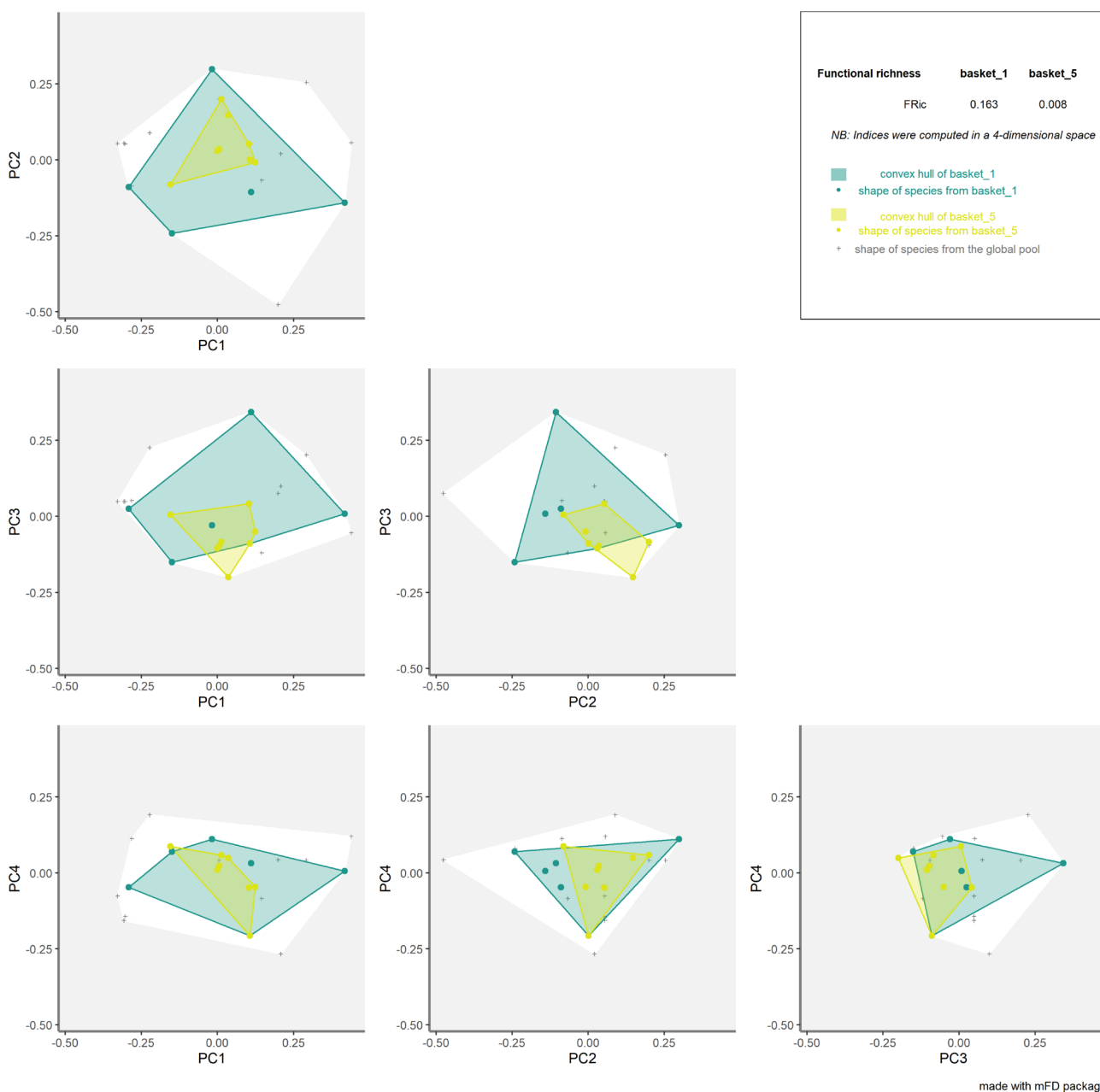
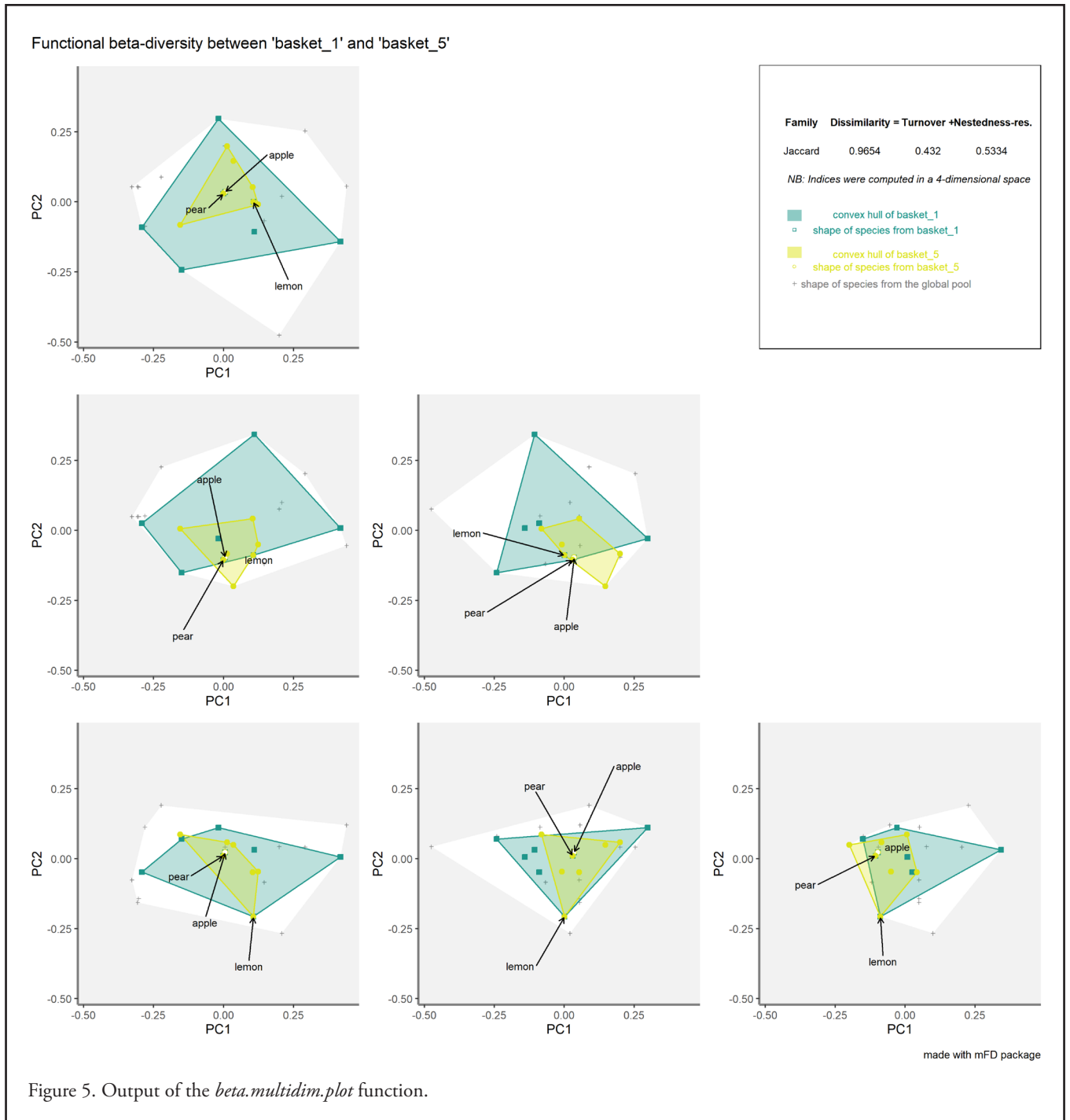


Figure 4. Output of the *alpha.multidim.plot* function for the FRic index.

In Fig. 5, the convex hulls filled by species belonging to basket 1 and basket 5 are plotted respectively in green and yellow. Further illustrations of the use of the functions listed above as well as of the functions for working with continuous traits and computing distance-based indices or entities-based indices of the *mFD* package are available as tutorials in the package vignette (<<https://cmlmagneville.github.io/mFD/index.html>>).



Conclusion

The *mFD* package allows assessing all FD indices available to date with all types of traits in an integrated, standardized, and comprehensive manner. Importantly, all methodological choices to be made by the user are included as parameters of the functions to ensure reproducibility. The functions use

common formatting and terminology for inputs and outputs which ease computing the successive steps within and between the three main approaches (entities-, distance-, and coordinates-based indices).

In addition, all functions include internal checks of inputs (e.g. match of species names between traits and assemblages matrix, absence of missing values) that return detailed error

messages if needed. All graphical functions can produce ready-to-publish figures with customizable parameters (e.g. shape, size, colors of points, and convex hulls). *mFD* package was developed in R.4.0, works for $R \geq 3.5$, and depends directly on 16 packages (Supporting information). It can be cited using the output of citation (package = ‘mFD’) and by citing this article.

The website associated with the package (<<https://cml-magneville.github.io/mFD/index.html>>) presents five reproducible tutorials as R markdown files.

The unique features of *mFD* functions and associated online materials will thus make it easier for students and researchers in functional ecology to compute FD in a reproducible way.

The functions from *mFD* could even be applied for researches outside functional ecology, such as to measure phenotypic variability within populations or between species, for example trophic position based on variables such as stable isotope ratios of carbon and nitrogen (Cucherousset and Villéger 2015) or to map vegetation status using telemetry data (Feret and de Boissieu 2019).

To cite mFD or acknowledge its use, cite this software note as follows, substituting the version of the application that you used for ‘version 1.0’:

Magneville, C. et al. 2021. mFD: an R package to compute and illustrate the multiple facets of functional diversity. – *Ecography* 44: XX–XX (ver. 1.0).

Funding – N. Loiseau was supported by a grant from University of Montpellier MUSE (BUBOT). D. Mouillot and N. Loiseau were part of the FREE group funded by the Centre for the Synthesis and Analysis of Biodiversity (CESAB) of the Foundation for Research on Biodiversity (FRB) and EDF. S. Villéger and D. Mouillot were supported by European Union’s Horizon 2020 (MaCoBioS project, grant agreement no. 869710). S. Villéger and A. Escalas were supported by ANR grant (EXOFISHMED project, grant agreement no. ANR-17-CE32-0003). C. Albouy was funded by an ‘étoile montante’ fellowship from the ‘Pays de la Loire’ region (no. 2020_10792). E. Maire was funded by the FAIRFISH project (ERC grant number: 759457).

Author contributions

Camille Magneville: Conceptualization (lead); Methodology (equal); Software (lead); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (lead); Writing – review and editing (lead). **Nicolas Loiseau:** Conceptualization (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Camille Albouy:** Conceptualization (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Nicolas Casajus:** Conceptualization (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Thomas Claverie:** Conceptualization (equal); Software (equal); Visualization (equal); Writing – original draft

(equal); Writing – review and editing (equal). **Arthur Escalas:** Conceptualization (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Fabien Leprieux:** Conceptualization (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Eva Maire:** Conceptualization (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **David Mouillot:** Conceptualization (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Sébastien Villéger:** Conceptualization (lead); Methodology (equal); Project administration (lead); Software (lead); Visualization (equal); Writing – original draft (lead); Writing – review and editing (lead).

Transparent Peer Review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.05904>>.

Data availability statement

Code and data used for the example box are publicly available in the mFD package on CRAN (<<https://cran.r-project.org/web/packages/mFD/index.html>>).

Supporting information

Any supporting information associated with this article is available from the online version.

References

- Baselga, A. 2012. The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. – *Global Ecol. Biogeogr.* 21: 1223–1232.
- Baselga, A. and Orme, C. D. L. 2012. betapart: an R package for the study of beta diversity. – *Methods Ecol. Evol.* 3: 808–812.
- Baselga, A. et al. 2021. betapart: partitioning beta diversity into turnover and nestedness components. R package ver. 1.5.4. – <<https://cran.r-project.org/web/packages/betapart/index.html>>.
- Bellwood, D. R. et al. 2006. Functional versatility supports coral reef biodiversity. – *Proc. R. Soc. B* 273: 101–107.
- Blonder, B. and Harris, D. J. 2019. hypervolume: high dimensional geometry and set operations using Kernel density estimation, support vector machines and convex hulls. R package ver. 2.0.12. – <<https://cran.r-project.org/web/packages/hypervolume/index.html>>.
- Blonder, B. et al. 2018. New approaches for delineating n-dimensional hypervolumes. – *Methods Ecol. Evol.* 9: 305–319.
- Botta-Dukát, Z. 2005. Rao’s quadratic entropy as a measure of functional diversity based on multiple traits. – *J. Veg. Sci.* 16: 533–540.
- Cardoso, P. et al. 2014. A new frontier in biodiversity inventory: a proposal for estimators of phylogenetic and functional diversity. – *Methods Ecol. Evol.* 5: 452–461.

- Cardoso, P. et al. 2020. BAT – biodiversity assessment tool – R package ver. 2.7.1. – <<https://cran.r-project.org/web/packages/BAT/index.html>>.
- Chao, A. and Ricotta, C. 2019. Quantifying evenness and linking it to diversity, beta diversity and similarity. – *Ecology* 100: e02852.
- Chao, A. et al. 2019. An attribute-diversity approach to functional diversity, functional beta diversity and related (dis)similarity measures. – *Ecol. Monogr.* 89: e01343.
- Cornwell, W. K. et al. 2006. A trait-based test for habitat filtering: convex hull volume. – *Ecology* 87: 1465–1471.
- De Bello, F. D. et al. 2009. Partitioning of functional diversity reveals the scale and extent of trait convergence and divergence. – *J. Veg. Sci.* 20: 475–486.
- De Bello, F. D. et al. 2021. Towards a more balanced combination of multiple traits when computing functional differences between species. – *Methods Ecol. Evol.* 12: 443–448.
- Díaz, S. and Cabido, M. 2001. Vive la différence: plant functional diversity matters to ecosystem processes. – *Trends Ecol. Evol.* 16: 646–655.
- Garnier, E. et al. 2004. Plant functional markers capture ecosystem properties during secondary succession. – *Ecology* 85: 2630–2637.
- Gower, J. C. 1971. A general coefficient of similarity and some of its properties. – *Biometrics* 27: 857–871.
- Jost, L. 2006. Entropy and diversity. – *Oikos* 113: 363–375.
- Laliberté, E. and Legendre, P. 2010. A distance-based framework for measuring functional diversity from multiple traits. – *Ecology* 91: 299–305.
- Laliberté, E. et al. 2014. FD: measuring functional diversity from multiple traits, and other tools for functional ecology. – R package ver. 1, <<https://cran.r-project.org/web/packages/FD/index.html>>.
- Lavelle, S. et al. 2008. Assessing functional diversity in the field – methodology matters! – *Funct. Ecol.* 22: 134–147.
- Legendre, P. and Legendre, L. 2012. Numerical ecology. – Elsevier.
- Loiseau, N. et al. 2017. Performance of partitioning functional beta-diversity indices: influence of functional representation and partitioning methods. – *Global Ecol. Biogeogr.* 26: 753–762.
- Maire, E. et al. 2015. How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. – *Global Ecol. Biogeogr.* 24: 728–740.
- McGill, B. J. et al. 2006. Rebuilding community ecology from functional traits. – *Trends Ecol. Evol.* 21: 178–185.
- McLean, M. J. et al. 2019. Functional reorganization of marine fish nurseries under climate warming. – *Global Change Biol.* 25: 660–674.
- McWilliam, M. et al. 2020. Deficits in functional trait diversity following recovery on coral reefs. – *Proc. R. Soc. B* 287: 20192628.
- Mouchet, M. et al. 2008. Towards a consensus for calculating dendrogram-based functional diversity indices. – *Oikos* 117: 794–800.
- Mouillot, D. et al. 2013a. A functional approach reveals community responses to disturbances. – *Trends Ecol. Evol.* 28: 167–177.
- Mouillot, D. et al. 2013b. Rare species support vulnerable functions in high-diversity ecosystems. – *PLoS Biol.* 11: e1001569.
- Mouillot, D. et al. 2014. Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. – *Proc. Natl Acad. Sci. USA* 111: 13757–13762.
- Naeem, S. et al. 2012. The functions of biological diversity in an age of extinction. – *Science* 336: 1401–1406.
- Olden, J. D. et al. 2006. Life-history strategies predict fish invasions and extirpations in the Colorado River Basin. – *Ecol. Monogr.* 76: 25–40.
- Pavoine, S. 2020. adiv: an r package to analyse biodiversity in ecology. – *Methods Ecol. Evol.* 11: 1106–1112.
- Pavoine, S. 2021. adiv: analysis of diversity. R package ver. 2.1.1. – <<https://cran.r-project.org/web/packages/adiv/index.html>>.
- Petchey, O. L. and Gaston, K. J. 2002. Functional diversity (FD), species richness and community composition. – *Ecol. Lett.* 5: 402–411.
- Petchey, O. L. and Gaston, K. J. 2006. Functional diversity: back to basics and looking forward. – *Ecol. Lett.* 9: 741–758.
- Pimiento, C. et al. 2020. Functional diversity of marine megafauna in the Anthropocene. – *Sci. Adv.* 6: eaay7650.
- Pollock, L. J. et al. 2020. Protecting biodiversity (in all its complexity): new models and methods. – *Trends Ecol. Evol.* 35: 1119–1128.
- Rao, C. R. 1982. Diversity and dissimilarity coefficients: a unified approach. – *Theor. Popul. Biol.* 21: 24–43.
- Ricotta, C. and Szeidl, L. 2009. Diversity partitioning of Rao's quadratic entropy. – *Theor. Popul. Biol.* 76: 299–302.
- Trindade-Santos, I. et al. 2020. Global change in the functional diversity of marine fisheries exploitation over the past 65 years. – *Proc. R. Soc. B* 287: 20200889.
- Villéger, S. et al. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. – *Ecology* 89: 2290–2301.
- Villéger, S. et al. 2011. The multidimensionality of the niche reveals functional diversity changes in benthic marine biotas across geological time. – *Ecol. Lett.* 14: 561–568.
- Villéger, S. et al. 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.
- Villéger, S. et al. 2017. Functional ecology of fish: current approaches and future challenges. – *Aquat. Sci.* 79: 783–801.
- Violle, C. et al. 2007. Let the concept of trait be functional! – *Oikos* 116: 882–892.
- Violle, C. et al. 2014. The emergence and promise of functional biogeography. – *Proc. Natl Acad. Sci. USA* 111: 13690–13696.
- Violle, C. et al. 2017. Functional rarity: the ecology of outliers. – *Trends Ecol. Evol.* 32: 356–367.
- Walker, B. et al. 1999. Plant attribute diversity, resilience and ecosystem function: the nature and significance of dominant and minor species. – *Ecosystems* 2: 95–113.
- Weiher, E. et al. 1998. Community assembly rules, morphological dispersion and the coexistence of plant species. – *Oikos* 81: 309–322.