SPRINGER BRIEFS IN ENVIRONMENTAL SCIENCE

Laura Pla · Fernando Casanoves
Julio Di Rienzo

Quantifying Functional Biodiversity

Springer

SpringerBriefs in Environmental Science

For further volumes:
http://www.springer.com/series/8868

Laura Pla · Fernando Casanoves
Julio Di Rienzo

Quantifying Functional Biodiversity

Springer

Prof. Laura Pla
Department of Agricultural Technology
Francisco de Miranda National
Experimental University (UNEFM)
Complejo Docente El Hatillo UNEFM
Apartado 7434
Coro 4101, Falcón
República Bolivariana de Venezuela
e-mail: laura@reacciun.ve

Dr. Fernando Casanoves
Tropical Agriculture Research and
Higher Education Center (CATIE)
Turrialba 30501
Costa Rica
e-mail: casanoves@catie.ac.cr

Dr. Julio Di Rienzo
Faculty of Agricultural Sciences
The National University of Cordoba
Avenida Valparaíso s/n CC 509
Ciudad Universitaria
5000 Córdoba
Argentina
e-mail: dirienzo@agro.uncor.edu

All files referenced throughout chapters 2 and 4 are available for download at Springer's Extra Materials website: http://extras.springer.com/

ISSN 2191-5547
ISBN 978-94-007-2647-5
DOI 10.1007/978-94-007-2648-2
Springer Dordrecht Heidelberg London New York

e-ISSN 2191-5555
e-ISBN 978-94-007-2648-2

Library of Congress Control Number: 2011939407

Cover design: eStudio Calamar, Berlin/Figueres

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgements

We thank the partial support given by the Inter-American Institute for Global Change Research, IAI-CRN 2015 (supported by National Science Foundation, Grant GEO-0452325) and Dra. Sandra Díaz who leads the Núcleo DiverSus.

Contents

Abbreviations

CWM	Community weighted mean
DGC	Univariate partitive mean comparison method
ES	Ecosystem services
FAD1	Functional attribute diversity, index 1
FAD2	Functional attribute diversity, index 2
FDc	Functional diversity index community based
FDis	Functional dispersion index (multi-trait)
FDiv	Functional divergence index (multi-trait)
FDvar	Functional divergence (single-trait)
FEve	Functional evenness index (multi-trait)
FRic	Functional richness index (multi-trait)
FRO	Functional regularity (single-trait)
GFD	Generalized functional diversity index
gDGC	Multivariate partitive mean comparison method
LA	Leaf area
LDMC	Leaf dry matter content
LMA	Leaf mass per area
LNC	Leaf nitrogen content
LTS	Leaf tensile strength
masl	Meters above sea level
MEA	Millennium Ecosystem Assessment
MFAD	Modified functional attribute diversity index
NPC	Leaf phosphorus content
PFT	Plant functional types
RTQ	Mexican land use system ‘roza-tumba-quema’
SLA	Specific leaf area
TRY	International database of plant traits
WD	Wood density
wFDc	Weighted functional diversity index community based
wFDp	Weighted functional diversity index plot based

Chapter 1
Introduction

Abstract Functional diversity is an increasing used concept to address changes in biodiversity. It is an emerging concept which summarizes key properties of ecosystems of special interest in global climate change studies and in the evaluation of the effects of land management in the preservation of ecosystem services for human wellbeing. In this chapter we introduce the main notions associated with functional diversity approach, including definition of functional diversity, ecosystem processes, and ecosystem services and linking these concepts to species traits. We highlight the importance of functional diversity approach using some examples to show the relationship between ecosystem services with species traits.

Keywords Ecosystem services · Functional traits · Functional diversity assessment · Millennium ecosystem assessment · Functional ecology

1.1 Functional Diversity Approach to Quantify the Biodiversity

Functional ecology establishes principles and tools to forge links between the characteristics of communities, and ecosystem functions and services (Cornelissen et al. 2003; Lavorel et al. 2007). For example, the energy and materials flow through the biotic and abiotic components of an ecosystem is directly related to productivity, while resistance and resilience are measures of the ability of a system to respond before the disturbance or adapting to change (Díaz and Cabido 2001).

The functional approach allows simplify the floristic complexity and the effects of vegetation to understand the responses, in terms of key ecological processes. It also provides tools to identify and monitor global change effects and other consequences of human activity, emphasizing ecosystem services (ES). This functional

L. Pla et al., *Quantifying Functional Biodiversity*, SpringerBriefs in Environmental Science, DOI: 10.1007/978-94-007-2648-2_1, © The Author(s) 2012

approach transcends the descriptive analysis. It can be done in a relatively easy, inexpensive and standardized way, allowing the comparison among communities and between community properties and environmental variables.

According to Grime (1998) three groups of species may be identify related to its contribution to the community performance: dominants, subordinates and transients. Dominant species are the most important species as determinants of ecosystem properties such as productivity, carbon sequestration, water relations, nutrient cycling and storage, litter quality and resistance and resilience to perturbations. Ecosystem functions are likely to be closely predictable from the most abundant species, those which contribute highly to the total plant biomass. This is known as mass ratio hypothesis (Grime 1998). The contributions to ecosystem functions are dictated by the laws of physics and chemistry. They state that the greater the effects of large autotrophs within the ecosystem, there will be a greater participation in processes like photosynthesis, resources inputs, nutrient cycling, and hydrology cycle, among others. This implies that ecosystem properties should be determined mainly by dominants species and some subordinates, and much less by transients' species.

Application of the mass ratio hypothesis is restricted to autotrophs in ecosystem processes. In animals, when attention is turned to trophic elements, like parasites, herbivores, and predators, impact on ecosystem functions is less related to abundance (Grime 1998).

Functional diversity approach using plants is based on the most abundant species, which implies the inclusion of all the species necessary to account for the 80% of the total biomass. When species' biomass is not available, other measures like cover, basal area or abundance may be used as surrogate for biomass (Díaz et al. 2007a; Lavorel et al. 2008). The protocols applied for the functional characterization comply with this recommendation discarding the less represented species in the community.

Ecosystem services are the benefits that humans obtain from ecosystems for support their survival and quality of life. The benefit may be directly associated to survival like food production or to effects indirectly related to quality of life, like energy provision (MEA 2005). ES are also used to link the ecological concept of functional diversity with the social concept of social actor strategies (Díaz et al. 2011). Going deeper into the links among biodiversity, ES, and social actors it is necessary to consider the contributions that biodiversity provides to an ES, the social actors perception, their needs, access, and management capability of the ES (Carpenter et al. 2009).

The ecosystem services depend on ecosystem properties which in turn are determined by ecosystem functions and ecosystem processes. For example, soil fertility (as service that ecosystem provide) depends on textural composition, organic matter accumulation and nutrient cycling. Not all ES depend directly upon ecosystem processes; some are associated to aesthetic or spiritual value of species (Díaz et al. 2007a; de Bello et al. 2010). For example, the aesthetic value of flowers from *Rafflesia arnoldii*, a parasitic species, with flowers up to more than 1 m, the largest in the world, growing in Sumatra (Beaman et al. 1988), or the

presence of a relic species of dolphins, *Lipotes vexillifer*, in the Yangtze river in China (Zhou et al. 1998), which is threatened by the dam harbor the largest hydropower plant in the world (López-Pujol 2008).

The ecosystem functions are determined by the role of different species in maintaining ecosystem processes. Changes in species composition and changes in the relative abundance have a direct implication over ecosystem structure in terms of community dynamics. Ecosystem properties related to ES would be referred as a function or process. As emphasis of functional diversity is placed on the services that an ecosystem can provide, we will use ecosystem properties to describe collectively the ecosystem processes and functions.

1.2 Functional Diversity Assessment

Functional diversity is defined as the value, range, distribution and relative abundance of the functional characteristics of organisms in a community (Chapin et al. 2000; Loreau and Hector 2001; Hooper et al. 2005). In contrast to the taxonomic biodiversity, based only on the relative abundance of species in the community, functional diversity summarizes various aspects of the biological composition and hence the role of populations in the community. Functional diversity may be linked directly to the ecosystem services (Díaz et al. 2007c).

As functional diversity states for characteristics of individuals of species in the community, a set of characteristics has to be evaluated. A trait is a well-defined, measurable property of organisms, usually measured at the individual level and used comparatively across species. A functional trait is one that strongly influences organismal performance in the community (Lavorel and Garnier 2002; Cornelissen et al. 2003; Violle et al. 2007). Trait values influence growth, reproduction and survival of organisms, and affect relationship among organisms of different species. These, in turn, drive the properties and services that ecosystem may provide (Luck et al. 2009).

The best subset of traits are those that provided the most complete information related to an ecosystem service under study and that, simultaneously, may be easily measured with the least sample effort and at a low cost. For example, to study photosynthesis capacity, measurement of area and weight of leafs may be used to estimate specific leaf area, meanwhile, maximum high or diameter at breast height registered at two or more times may be used to study growth rate.

There is empirical evidence that specific leaf area is positively correlated with photosynthetic potential and hence growth rates, recruitment and mortality, and negatively correlated with longevity and investment in defenses. For example, Garnier et al. (2004) found that the 58% of variation ($r^2 = 0.58$) of specific above-ground net primary productivity (g kg^{-1} d^{-1}) in 12 plots of vegetation in south France may be estimated using specific leaf area (m^2 kg^{-1}). Also, leaf dry matter content, and leaf tensile strength are negatively correlated with photosynthetic potential and hence growth rates, recruitment and mortality, and are positively

correlated with longevity and investment protection and defenses (Almeida-Cortez et al. 1999).

There is international consensus around the importance to follow a protocol to measure traits and there are international efforts to have information on traits values for as many species as possible. The project TRY (www.try-db.org) is an effort to collate existing plant functional trait data set into a communal repository (Kattge et al. 2011a). This initiative have now three million trait records for about 69000 plant species and about 50 scientific projects are using plant trait data via TRY. There are some guidelines to make your own data base part of TRY (Kattge et al. 2011b) and also to use data base from TRY. You may learn more about the data sharing policy within TRY going to the web page.

1.3 Classification of Ecosystem Services

According to the Millennium Ecosystem Assessment (MEA 2003) ecosystem services may be classify in four main groups. The classification differentiates among provisioning services, regulating services, cultural services and supporting services. Production of food, availability of fresh water, provision of fuel-wood, fiber, biochemical and genetic resources are provisioning services. Regulating services refer to the benefits obtained from regulation of ecosystem properties, like climate regulation, disease regulation, water regulation, water purification, pollination. Cultural services join those nonmaterial benefits obtained from ecosystems. Spiritual and religious services, recreation and ecotourism, aesthetic, inspirational, educational, sense of place and cultural heritage are examples of cultural ecosystem services (MEA 2003).

Supporting services, those necessary for the production of all other ecosystem services, like soil formation, nutrient cycling or primary production were considered as the fourth group in Millennium Ecosystem Assessment classification. Some ecosystem services included in this group are nowadays considered as part of the regulating services, or as ecosystem properties like primary production, oxygen production and nutrient cycling (Carpenter et al. 2009; Díaz et al. 2011; Polania et al. 2011).

1.4 Selection of Traits According to Ecosystem Service

To evaluate functional diversity at a community or assemblage, traits associated to main ecosystem properties has to be identify. Several studies identify the traits that have more prediction capability of ecosystem properties. For example, primary productivity, carbon accumulation in vegetation, soil carbon accumulation and decomposition rate are used to evaluate climatic regulation through carbon sequestration. To evaluate these properties we use traits like growth form and

growth rate, plant height, plant longevity, wood density, dry matter, lignin, leaf nitrogen content, leaf longevity, toughness of leaves, specific leaf area (SLA) and leaf mass per area (LMA), potential decomposition rate of leaves, and stems and specific root length (Lavorel and Garnier 2002; Díaz et al. 2004, 2009; Wardle et al. 2004; De Deyn et al. 2008).

Ecosystems may provide services to control water erosion. This service depends on water retention in soil and sediment, litter and standing vegetation, and balance between infiltration and runoff, properties that may be evaluate considering traits like growth form and growth rate of the plant, plant longevity, crown architecture, clonality, longevity of leaves, dry matter, lignin and nitrogen content in leaves, potential decomposition rate of leaves and stems, root architecture and deep and underground stems (Brauman et al. 2007).

Production of forage for herbivores like livestock, wild species or symbolic species is an ecosystem process associated with food provision. Traits associated with food provision are growth form and growth rate of the plant, plant longevity, plant high, regrowth, position of the buds of renewal, longevity of leaves, dry matter, lignin and nitrogen content in leaves, phosphorus and active toxic compounds in leaves, leaf toughness, specific leaf area (SLA) and leaf mass per area (LMA), symbiosis with nitrogen fixer microorganisms or insects (Wright et al. 2002; Díaz et al. 2007b; Quétier et al. 2007).

1.5 Functional Diversity Quantification

There are several methods to quantify functional diversity, the preference for one or another relays on type of available information and is related to the aims of research. All methods are based on data of functional traits measured, at least, at species level. The following chapters focus on methods to quantify functional diversity, how to relate functional diversity with environmental variables and its relation with ecosystem services. Numerical examples are analyzed using a free specialized software: FDiversity (Casanoves et al. 2011), which can be downloaded from www.fdiversity.nucleodiversus.org.

One option to quantify functional diversity is to estimate the number of functional groups in the community. This is a measure of functional richness. A functional group is a subset of the species present at the assemblage that shared similar trait profiles. Composition of functional group in a given community may vary according to the service being investigated. Functional groups are identified by cluster analysis of trait's profiles. In Chap. 2 we present this methodology and use one example to defined plant functional groups and other to define bird functional groups (guilds).

Functional diversity may also be summarized using functional indices. These indices are based on traits' values evaluated at species level. They may also incorporate weights by some measure of the species importance in the community. Chapter 3 is entirely dedicated to functional indices definitions. We included

taxonomy of the indices based upon the information they used and the output they offer. For each index we will include its definition, the information needed to estimate it, it's statistical and ecological properties, and some reference to explore its application to real cases.

The last chapter of the book, Chap. 4 is a tutorial to estimate the indices using FDiversity (Di Rienzo et al. 2008). Three step by step examples are presented. For each one we calculate the functional diversity indices and compare results from several communities to determine statistical differences among them, or explore the relationship with environmental variables. The data sets and key detailed output of the analysis are available at Springer's Extra Materials website: http://extras.springer.com/.

References

Almeida-Cortez JS, Shipley B, Arnason JT (1999) Do plant species with high relative growth rates have poorer chemical defences? Func Ecol 13(6):819–827

Beaman RS, Decker PJ, Beaman JH (1988) Pollination of Rafflesia (Rafflesiaceae). Am J Bot 75:1148–1162

Brauman KA, Daily GC, Duarte TK, Mooney HA (2007) The nature and value of ecosystem services: an overview highlighting hydrologic services. Annu Rev Environ Resour 32:67–98

Carpenter SR, Mooney HA, Agard J, Capistrano D, DeFries R, Díaz S, Dietz T, Duriappah A, Oteng-Yeboah A, Pereira HM, Perrings C, Reid WV, Sarukhan J, Scholes RJ, Whyte A (2009) Science for managing ecosystem services: beyond the millennium ecosystem evaluation. Proc Natl Acad Sci USA 106:1305–1312

Casanoves F, Pla L, Di Rienzo JA, Díaz S (2011) FDiversity: a software package for the integrated analysis of functional diversity. Methods Ecol Evol 2:233–237

Chapin FS III, Zavaleta ES, Eviner VT, Naylor R, Vitousek PR, Reynolds HL, Hooper DU, Lavorel S, Sala OE, Hobbie SE, Mack MC, Díaz S (2000) Functional and societal consequences of changing biotic diversity. Nature 405:234–242

Cornelissen JHC, Lavorel S, Garnier E, Díaz S, Buchmann N, Gurvich DE, Reich PB, Ter Steege H, Morgan HD, Heijden MGA, van der Pausas JG, Poorter H (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Aust J Bot 51:335–380

De Bello F, Lavorel S, Díaz S, Harrington R, Cornelissen JHC, Bardgett RD, Berg MP, Cipriotti P, Feld CK, Hering D, Marins da Silva P, Potts SG, Sandin L, Sousa JP, Storkey J, Wardle DA, Harrison PA (2010) Towards an assessment of multiple ecosystem processes and services via functional traits. Biodivers Conserv 19:2873–2893

De Deyn GB, Cornelissen JHC, Bardgett RD (2008) Plant functional traits and soil carbon sequestration in contrasting biomes. Ecol Lett 11:516–531

Di Rienzo JA, Casanoves F, Pla L (2008) FDiversity, Software to estimate functional diversity. RDNDA, Argentina Register Number 702841

Díaz S, Cabido M (2001) Vive la différence: plant functional diversity matters to ecosystem processes. Trends Ecol Evol 16(11):646–655

Díaz S, Hodgson JG, Thompson K, Cabido M, Cornelissen JHC, Jalili A, Montserrat-Martí G, Grime JP, Zarrinkamar F, Asri Y, Band SR, Basconcelo S, Castro-Díez P, Funes G, Hamzehee B, Khoshnevi M, Pérez-Harguindeguy N, Pérez-Rontomé MC, Shirvany FA, Vendramini F, Yazdani S, Abbas-Azimi R, Bogaard A, Boustani S, Charles M, Dehghan M, de Torres-Espuny L, Falczuk V, Guerrero-Campo J, Hynd A, Jones G, Kowsary E,

Kazemi-Saeed F, Maestro-Martínez M, Romo-Díez A, Shaw S, Siavash B, Villar-Salvador P, Zak MR (2004) The plant traits that drive ecosystems: evidence from three continents. J Veg Sci 15:295–304

Díaz S, Lavorel S, De Bello F, Quétier F, Grigulis K, Robson M (2007a) Incorporating plant functional diversity effects in ecosystem service assessments. Proc Natl Acad Sci USA 104:20684–20689

Díaz S, Lavorel S, McIntyre S, Falczuk V, Casanoves F, Milchunas DG, Skarpe C, Rush G, Sternberg M, Noy-Meir I, Landsberg J, Zhang W, Clark H, Campbell BD (2007b) Plant trait responses to grazing a global synthesis. Glob Change Biol 13:313–341

Díaz S, Lavorel S, Stuart Chapin F, Tecco PA, Gurvich DE, Grigulist K (2007c) Functional diversity—at the crossroads between ecosystem functioning and environmental filters. In: Canadell JG, Pataki DE, Pitelka LF (eds) Terrestrial ecosystems in a changing world. Springer, New York

Díaz S, Hector A, Wardle DA (2009) Biodiversity in forest carbon sequestration initiatives: not just a side benefit. Curr Opin Environ Sustain 1:55–60

Díaz S, Quétier F, Cáceres DM, Trainorc SF, Pérez-Harguindeguy N, Bret-Harted MS, Finegan B, Peña-Claros M, Poorter L (2011) Linking functional diversity and social actor strategies in a framework for interdisciplinary analysis of nature's benefits to society. Proc Natl Acad Sci USA 108:895–902

Garnier E, Cortez J, Bille′s G, Navas ML, Roumet C, Debussche M, Laurent G, Blanchard A, Aubry D, Bellmann A, Neill C, Toussaint JP (2004) Plant functional markers capture ecosystem properties during secondary succession. Ecology 85(9):2630–2637

Grime JP (1998) Benefits of plant diversity to ecosystems: immediate filter and founder effects. J Ecol 86:902–910

Hooper DUF, Chapin S, Ewel JJ, Hector A, Inchausti P, Lavorel S, Lawton JH, Lodge DM, Loreau M, Naeem S, Schmid B, Setälä H, Symstad AJ, Vandermeer J, Wardle DA (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol Monogr 75:3–35

Kattge J, Díaz S, Lavorel S, Prentice IC, Leadley P et al (2011a) TRY—a global database of plant traits. Glob Change Biol 17:2905–2935

Kattge J, Ogle K, Bönisch G, Díaz S, Lavorel S, Madin J, Nadrowski K, Nöllert S, Sartor K, Wirth C et al (2011b) A generic structure for plant trait databases. Meth Ecol Evol 2:202–213

Lavorel S, Garnier E (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Funct Ecol 16:545–556

Lavorel S, Díaz S, Cornelissen JH, Garnier E, Harrison SP, McIntyre S, Pausas J, Pérez N, Roumet C, Urcelay C (2007) Plant functional types: Are we getting any closer to the Holy Grail? In: Canadell JG, Pataki D, Pitelka L (eds) Terrestrial ecosystems in a changing world. Springer, New York

Lavorel S, Grigulis K, McIntyre S, Williams NSG, Garden D, Dorrough J, Berman S, Quétier F, Thébault A, Bonis A (2008) Assessing functional diversity in the field—methodology matters!. Func Ecol 16:134–147

López-Pujol J (2008) Impactos sobre la biodiversidad del embalse de las Tres Gargantas en China. Ecosistemas 17:134–145

Loreau M, Hector A (2001) Partitioning selection and complementarity in biodiversity experiments. Nature 412:72–76

Luck GW, Harrington R, Harrison PA, Kremen C, Berry PM, Bugter R, Dawson TP, de Bello F, Díaz S, Feld CK, Haslett JR, Hering D, Kontogianni A, Lavorel S, Rounsevell M, Samways MJ, Sandin L, Settele J, Sykes MT, van den Hove S, Vandewalle M, Martin Zobel M (2009) Quantifying the contribution of organisms to the provision of ecosystem services. Bioscience 59:223–235

MEA: Millennium Ecosystem Assessment (2003) Ecosystems and human well-being: a framework for assessment. World Resources Institute, Island Press, Washington

MEA: Millennium Ecosystem Assessment (2005) Ecosystems and human well-being: Biodiversity Synthesis. World Resources Institute, Washington

Polania C, Pla L, Casanoves F (2011) Diversidad funcional y servicios ecosistémicos. In: Casanoves F, Pla L, Di Rienzo JA (eds) Valoración y análisis de la diversidad funcional y su relación con los servicios ecosistémicos. Serie Técnica 384, CATIE, Turrialba

Quétier F, Thebault A, Lavorel S (2007) Plant traits in a state and transition framework as markers of ecosystem response to land-use change. Ecol Monogr 77:33–52

Violle C, Navas M, Vile D, Kazakou E, Fortunel C, Hummel I, Garnier E (2007) Let the concept of trait be functional. Oikos 116:882–892

Wardle DA, Walker LR, Bardgett RD (2004) Ecosystem properties and forest decline in contrasting long-term chronosequences. Science 305(5683):509–513

Wright IJ, Westoby M, Reich PB (2002) Convergence towards higher leaf mass per area in dry and nutrient-poor habitats has different consequences for leaf life span. J of Ecol 90:534–543

Zhou K, Sun J, Gao A, Würsig B (1998) Baiji (*Lipotes vexillifer*) in the lower Yangtze River: movements, numbers threats and conservation needs. Aquat Mammals 24(2):123–132

Chapter 2
Functional Groups

Abstract The set of species co-existing in a given community constitute a functional group if they have similar functional characteristics related to one ecosystem service. This dependence on ecosystem service is defined by theoretical framework or by empirical evidence. Functional groups in vegetation science are known as plant functional types and in animal science as guilds. Functional groups may be defined externally using categories for key traits or generated from several traits using cluster techniques. In this chapter we show how to identify functional groups, selecting the appropriate measures to evaluate species similarity based on trait profiles, and choosing linkage algorithms to conform the functional groups. Changes in the relative abundance of each group in a sample may be used to interpret the relationship of community composition with environmental conditions.

Keywords Cluster analysis · Number of functional groups · Distance measures · Similarity measures · Trait types

2.1 Selecting Trait and its Relation With Ecosystem Services

Because a functional group is a collection of organisms with similar suites of co-occurring functional attributes they have similar responses to external factors and/or effects on ecosystem processes (de Bello et al. 2010). A functional group is often referred as plant functional type (PFT) in vegetation sciences or as a 'guild' when referring to animals. Ecosystem properties or processes determine the services an ecosystem provides. These properties are associated to functional attributes of individuals (or population): the traits. Thus, the PFTs or the guilds are defined based on sets of species traits useful to explain ecosystem properties.

Several species of organisms within a trophy chain or trophy network with similar feeding types have the same function and are considered as a guild. Several plant

L. Pla et al., *Quantifying Functional Biodiversity*, SpringerBriefs in Environmental Science, DOI: 10.1007/978-94-007-2648-2_2,

species within an assemblage with similar photosynthetic strategy and foliar nutrient content (N and P) are considered as a PFT, possible related to wood density, which in turn affect the carbon sequestration service. Both, guild and PFT are associated to ecosystem services which make more suitable the human existence.

The idea of creating functional groups is to obtain a set of species having the same role in the ecosystem. These clusters of species are performed using a set of traits directly related to the ecosystem service. Several authors have summarized the relationship between traits and ecosystem services (MEA 2005; Carpenter et al. 2009; Lavorel et al. 2011; Polania et al. 2011).

The selected traits may include quantitative (continuous and discrete) and/or qualitative (nominal and ordinal) variables and the clusters are obtained by mean of a hierarchical algorithm. Hierarchical techniques are based on a dissimilarity matrix between species and a join procedure known as linkage strategy. The resulting hierarchy may be represented by a dendrogram, which allows grouping the species taking into account the level of the hierarchy and the aim of the study.

2.2 A Guide for Data Arrangement

The usual way to store data for further statistical processing is to arrange them in an $S \times t$ data table, where S represents the number of cases (in this context species) and t the number of traits. The traits may be continuous variables like leaf area, discrete quantitative variables like number of leaflets by leaf, or qualitative variables. If traits are qualitative, we should recognize if they are present-absent variables like evergreen or not, if they have more than two categories, or if they are ordinals. Additionally we should recognize if the categories are exclusive or not.

To calculate a dissimilarity matrix from a set of mixed type of traits it is convenient to express all of them in such a way that can be treated as quantitative. So the problem is how to re-express categorical variables. If a trait is categorical, having k exclusive categories, we may represent it as a set of $k - 1$ dummy variables (Box 2.1). To include the trait Reproductive system we have to use the two variables: Rep_Monoic and Rep_Dioic, which include all the information in the pairs (1,0) for Monoic, (0,1) for Dioic, and (0,0) for Hermaphrodite.

Box 2.1: Example of dummy transformation for an exclusive categorical trait

Species	Reproductive system	Rep_Monoic	Rep_Dioic
sp1	Monoic	1	0
sp2	Dioic	0	1
sp3	Dioic	0	1
sp4	Hermaphrodite	0	0
sp5	Hermaphrodite	0	0

When categories are not exclusive, like dispersal type (hydrochory, autochory, dispersed by mammals, etc.) we may represent the trait as a set of indicator variables (equals 1 if the category is present or 0 in contrary case), one for each observed category (Box 2.2). In future analysis we include the resulted four variables, in this example Hydrochory, Autochory, Zoochory and Wind.

Box 2.2: Example of indicator variables to identify nonexclusive categorical traits

Species	Dispersion	Hydrochory	Autochory	Zoochory	Wind
sp1	Hydrochory, Autochory, Zoochory	1	1	1	0
sp2	Hydrochory, Autochory	1	1	0	0
sp3	Autochory, Zoochory	0	1	1	0
sp4	Zoochory, Wind	0	0	1	1
sp5	Autochory, Wind	0	1	0	1

Ordinal variables can be numerically codified by a sequence of integers which relate to the rank of the categories they represent. For example canopy strata in a forest with four layers may be codifying as: 1, 2, 3, 4; this codification is equivalent to apply rank transformation to a quantitative variable. However, if the assumption that categories are representing equally spaced points of a scale do not hold, the numerical coding using a sequence of integers could be discarding important biological information about the trait expression. In that case it should be preferable to represent those traits through indicator variables as in the case of non-exclusive categorical variable, coding the category observed and all ones below it in the ordinal scale as present (Box 2.3). See for example, that a species of the highest strata has a one in all the categories; while sp3 from lowest strata has a one only in the lowest or sp2 belonging to the medium strata has ones in lowest and in medium variables. The three variables Lowest, Medium and Highest should be used to represent the ordinal variable in further analysis.

Box 2.3: Example of indicator variables to identify non equidistance ordinal categorical traits

Species	Strata	Lowest	Medium	Highest
sp1	Highest	1	1	1
sp2	Medium	1	1	0
sp3	Lowest	1	0	0
sp4	Highest	1	1	1
sp5	Lowest	1	0	0

In some cases it is possible to recover the original quantitative scale. Researches should avoid the use of categorical variable when a quantitative trait is possible to measure, even when the precision of the measure is not quite well. For example the layers in the forest canopy can be defined as: lower layer from 0 to 10 m, middle layer from 10 to 30 m, high layer from 30 to 40 m, and emergent layer from 40 to 60 m. In this case coding as 1, 2, 3 and 4 indicates the relative order of altitude of layers, but this coding is based on the assumption of equidistance between layers. It should be better to take the mean point of each layer to represent the strata and represent the layer by 5, 20, 35 and 50 m. If the higher strata were defined as 'more than 40 m', there is not a mean value to represent the layer unless an estimation of highest species is available.

After having all variables in an appropriate numerical scale, the data-table will be $S \times k$ where k will be equal or greater than t because of coding. Moreover the data table will only contain numerical representations of the traits in such a way that a common procedure can be applied to the whole table in order to obtain a dissimilarity matrix between species.

2.3 Statistical Procedures to Define Functional Groups

2.3.1 The Selection of a Dissimilarity Measure

The selection of dissimilarity measures depends on the type of variables in the data set. If all the traits have been measured in continuous or discrete scale, Euclidean distance will be appropriated to represent differences between pair of species. In case of dichotomous variables (0–1 variables) there are several similarity measures that can be used to derive a dissimilarity matrix. The most widely used are Jaccard (1908); Simple matching (Sokal and Michener 1958) and Dice (1945) (Box 2.4).

Cluster algorithms are usually based on dissimilarity measures, so when likeness between species is obtained from similarity measures, they must be transformed into dissimilarities. There is more than one way to convert similarities into dissimilarities, but when similarity ranges in the interval [0,1], the simplest one is

to calculate d_{ij} (the dissimilarity between species i and j) as $d_{ij} = 1 - s_{ij}$, were s_{ij} is the corresponding measure of similarity. Some specialized software like InfoStat automatically select the best (or most widely used) transformation for a given similarity measure.

When the data set has quantitative and qualitative variables the Gower similarity is one of the options (Gower 1971). This measure combines Euclidean distance with Jaccard similarity in a new similarity measure which can be converted into a dissimilarity using an appropriate transformation.

Box 2.4: Similarity measures for categorical variables

		Sp 1	
		Present (1)	Absent (1)
Sp 2	Present (1)	a	b
	Absent (0)	c	d

$$\text{Jaccard} = \frac{a}{a+b+c} \qquad \text{Dice} = \frac{2a}{2a+b+c}$$

$$\text{Single matching} = \frac{a+d}{a+b+c+d}$$

Another alternative to handle data sets with continuous and qualitative variables is to perform multidimensional scaling methods to summarize the qualitative variables in a set of new continuous variables (principal coordinates). In this case, it is possible to apply Euclidean distances to the set of continuous variables plus principal coordinates. One of the advantages using principal coordinates to summarize categorical traits is the possibility to use a multivariate analysis of variance or a cluster based mean comparison method (gDGC, Valdano and Di Rienzo 2007) to determine significance among resulted groups.

2.3.2 Standardization

Generally the traits values are expressed in their original scale of measurement, as for example: biomass (kg), leaf area (cm^2), wood density (mg cm^{-3}), maximum height (m), leaf carbon content (%), number of leaflets by leaf. When dissimilarity

measure is involve into the analysis, the scale and unit of measure will affect the results. Those variables having the largest scale will have the greatest impact on dissimilarity calculation. Standardization is the usual way to avoid the scale effect. Statistical software offers options to standardize data before performing analysis which are affected by the scale, like cluster analysis. Standardization can be also useful even when all the traits are expressed in the same units, because some traits can have more variance than others.

When the data set has mixture of quantitative and qualitative variables, previously transformed to zero–one variables, other options to avoid scale effects can be considered. For example, transform the variables to map the values into the zero–one interval [0, 1]. Thus, the minimum value in the original scale will be zero and the maximum will be one in the transformed scale.

2.3.3 Choosing the Linkage Algorithm Method

Widely used linkage clustering methods for cluster analysis are average linkage, single linkage, complete linkage and Ward, among others. Which of these methods is the best has not an easy answer. Although it depends on the purpose of clustering, the experience shows that the average linkage and Ward are the preferred methods. Average linkage is a linkage algorithm that maximized the cophenetic correlation. This means that the resulting hierarchy preserves in the best way possible the original dissimilarity structure. On the other hand, Ward method produces more clearly defined clusters which facilitate the definition of functional groups.

2.3.4 Assessing the Number of Functional Groups

Despite the existence of several proposals, Calinski and Harabasz (1974), Hartigan (1975), Sarle (1983), Kaufman and Rousseeuw (1990), Tibshirani et al. (2001), Pollard and van der Laan (2005), Fraley and Raftery (2002, 2006), Valdano and Di Rienzo (2007) and Pollard et al. (2009), to assessing the number of clusters in a data set, there are not statistical procedures or generally accepted rules to determine that number. Some times the number of clusters depends on the aim of the study, and it is determined by a heuristic criteria. For example if the main ecosystem property under study depends on acquisitive or conservative strategy of plants, two groups will be enough for the purpose of defining the corresponding PFTs. On the other hand, if the purpose of the research was to evaluate the effect of altitude on the composition of functional groups, a larger number of groups will be necessary.

As previously stated, each species in the data-set is represented by a single row-vector of traits values. Thus, no replicates are available at species level. This limit the use of assessing-number-of-clusters algorithms to those which do not need replicates. Within this class of algorithms, an approach to assessing the number of clusters in a data set is the comparison of the resulting clusters. This idea was

explored by Calinski and Harabasz (1974), who defined the number of clusters on the maximization of the between/within-cluster, of a generalized sum of squares ratio. Also Hartigan (1975), used the ratio between the within-cluster generalized sums of squares of k and (k + 1) clusters suggesting the selection of $k \geq 1$ as the minimum k for which the ratio is lesser than 10. A model-based approach to the determination of the number of cluster is found in the MCLUS algorithm by Fraley and Raftery (2002, 2006). Unpublished simulation results (Di Rienzo, personal communication) shown that MCLUS is the best choice when no replicate are available. It must be warning that the same simulation study also shown that all algorithms tend to underestimate the true number of clusters in the data. Previous discussed procedures for assessing the number of clusters assume that traits are continuous variables, and in case of MCLUS, that they follow a normal-multivariate distribution. When variables are categorical or a mixture of continuous and categorical, the previous methods could not be appropriate.

The analysis of molecular variance (AMOVA) was introduced by Excoffier et al. (1992).The method implements a multivariable analysis of variance like analysis for haplotypes data which are usually coded as 0–1 variables. Hypothesis testing is based on the permutation test principles. Because the method operates on a distances matrix, it is ease to extend its application to more general cases provided a distance matrix can be derived from the original data. Another approach, based on similar ideas is the analysis of similarities (ANOSIM). This method provides a way to test statistically whether there is a significant difference between two or more groups of sampling units (Clarke 1993).

2.4 Functional Characterization of Coastal Sandy Plain Vegetation in Southeast Brazil

Coastal sandy plain vegetation (Restinga) grows on sandy plains along the Brazilian coast. Because its proximity to the sea and flat to undulate plains, it is a preferential zone to human occupation and it is being degraded rapidly in the last two decades. To illustrate the functional group definition, we use a data set collected in 2010 by Dra. Leda Lorenzo in Ilha do Cardoso State Park, SP, Southeast Brazil. Foliar traits were selected because they are related to plant strategies of acquisition and use of resources. These traits are associated to services such as provision of fruits and medicinal plants, soil formation and fertility (that may leads to a more complex ecosystem in some centuries or decades), land fixation and control of sea erosion, carbon sequestration, and indirectly climate regulation (Díaz et al. 2007).

2.4.1 The Data Set

The five selected traits were leaf dry matter content (LDMC), leaf area (LA), leaf tensile strength (LTS), specific leaf area (SLA), and leaf nitrogen content (LNC).

In this survey, these leaf traits were evaluated according to Cornelissen et al. (2003) in two leaves of ten individuals by species.

One aim of the study is to defined plant functional groups, and then compares the abundance of each PFT in a gradient from the ocean coast. In Sect. 2.4.2 we present the analysis to cluster the 22 species founded in the study area, the procedures to select final number of clusters, and the characterization of PFTs.

2.4.2 Plant Functional Types from a Restinga Vegetation

All foliar trait measures were in a continuous scale and then the Euclidean distance is appropriate to evaluate the differences between species. We chose Ward as linkage criterion because the resulted dendrogram has compact groups and take into account the possible correlation among traits. We used InfoStat (Di Rienzo et al. 2010) to perform the analysis but similar results may be obtained with other statistical software.

From the dendrogram obtained (Fig. 2.1) it is clear that there are two main groups, one with six species and the other with 16 species. If we explore the mean trait values for each one (Table 2.1) we may recognize two plant strategies: group one with species allowing rapid acquisition of resource that have higher values of SLA, LA and LNC, and lower values of LTS and LDMC; and group two with species that conserve resources within well-protected tissues.

It could be interesting to interpret changes from sea coastal having more than two groups. Further inspection of the dendrogram of Fig. 2.1 allows differentiation among the 16 species of second group (see vertical line showing distance at which the dendrogram is splitting in Fig. 2.1). The resulting five clusters are characterized based on mean trait values (Table 2.2), and are referred as Plant Functional Type (PFT).

Plant functional types are named according to the species attributes. One group has three legume species with acquisitive attributes (Acq-Leg), other group has three non-legumes species with acquisitive attributes (Acq-non Leg). There are two groups of species with conservative attributes, one with six species in the transition to forest (Con-Forest) and the other with nine species in the transition to dunes (Con-Dune). The fifth PFT has only one species, *Clusia criuva* with high leaf tensile strength and latex contain, having maximum investment of resources in defense. Relationship among PFT and trait mean values are shown in a biplot (Fig. 2.2) resulted from a principal component analysis.

The biplot allows to relate traits mean values and species groups. The first two principal components explained 73.4% of the total variability; therefore the biplot presents a good synthesis for the relation among traits, groups and the interrelationships between PFT and traits.

The first principal component shows the maximum differentiation, in this case acquisitive strategies in the right, with high values of LA, SLA and LNC, and conservatives in the left, with low values for these traits. Between the both PFT

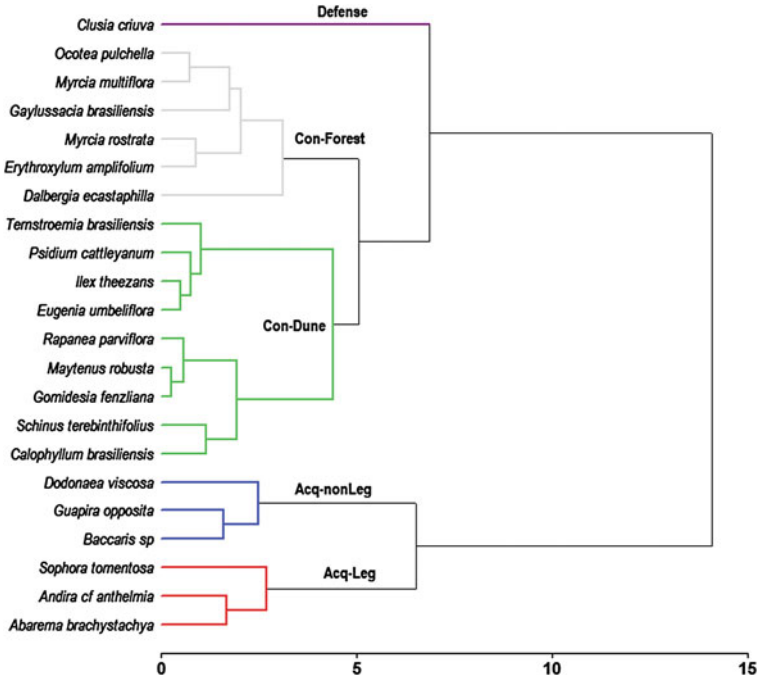


Fig. 2.1 Plant functional types in a Restinga vegetation survey. Five groups of species were defined using Euclidean distance and Ward linkage. *Acq-Leg* acquisitive legumes', *Acq-nonLeg* acquisitive non legumes', *Con-Dune* conservative of dune transition, *Con-Forest* conservative of forest transition, and *Defense* invested in defense

Table 2.1 Trait mean values when the species are splitting in two groups using Euclidean distance and Ward linkage algorithm

Trait	Cluster	S	Mean	SD	Min	Max
LTS (N.mm)	1	6	0.63	0.21	0.36	0.95
	2	16	1.07	0.47	0.67	2.62
LDMC (% dry weight)	1	6	36.05	9.50	26.25	47.20
	2	16	43.21	6.25	27.30	52.49
SLA (mm^2/mg)	1	6	9.28	1.09	7.70	10.63
	2	16	5.82	1.13	3.90	7.63
LA (cm^2)	1	6	67.76	63.24	9.02	153.24
	2	16	33.08	23.66	6.88	75.88
LNC (mg/g)	1	6	22.72	3.84	18.08	26.81
	2	16	12.35	3.35	8.49	19.83

S number of species in each group, *SD* standard deviation, *Min* minimum, *Max* maximum, *LTS* leaf tensile strength, *LDMC* leaf dry matter content, *SLA* specific leaf area, *LA* leaf area, *LNC* leaf nitrogen content

Table 2.2 Mean traits values for plant functional types (PFT) in a Restinga vegetation survey

PFT	S	LTS	LDMC	SLA	LA	LNC
Acq-Leg	3	0.60	41.55	9.81	123.22	24.41
Acq-nonLeg	3	0.66	30.54	8.75	12.30	21.03
Con-Dune	6	0.89	49.02	6.45	20.79	14.84
Con-Forest	9	1.02	41.11	5.49	38.56	11.05
Defense	1	2.62	27.30	5.04	57.53	9.24

S number of species in each PFT, *LTS* leaf tensile strength, *LDMC* leaf dry matter content, *SLA* specific leaf area, *LA* leaf area, *LNC* leaf nitrogen content, *Acq-Leg* acquisitive legumes', *Acq-nonLeg* acquisitive non legumes', *Con-Dune* conservative of dune transition, *Con-Forest* conservative of forest transition, *Defense* invested in defense

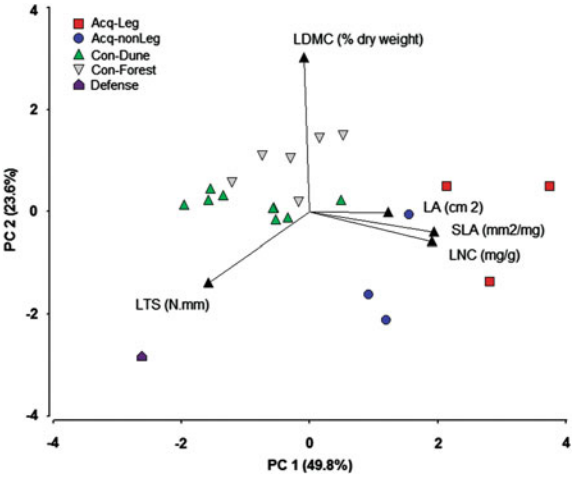


Fig. 2.2 First and second principal component using five traits and five plant functional types in a Restinga vegetation survey. *Acq-Leg* acquisitive legumes', *Acq-nonLeg* acquisitive non legumes', *Con-Dune* conservative of dune transition, *Con-Forest* conservative of forest transition, *Defense* invested in defense. *LTS* leaf tensile strength, *LDMC* leaf dry matter content, *SLA* specific leaf area, *LA* leaf area, and *LNC* leaf nitrogen content

with acquisitive strategy, the legumes are more acquisitive than the non legumes. The acquisition of resources is relatively cheaper for legumes since they have symbiosis with nitrogen fixers.

For the PFTs with conservative strategy, the second principal component allows to separate between Con-Forest and Con-Dune through LDMC values, which are higher on the Con-Forest group. The Defense PFT is defined by extremely high values of LTS, associated to investments in defense against mechanical damage.

2.5 Functional Groups for Bird Species in Nicaragua

In this example we used a database with bird traits to define and characterize functional groups in tropical landscapes in Nicaragua to assess patterns of functional diversity in different land uses. We will define and characterize the

functional groups, and after that we will study the relative frequency of each group in six landscapes with different human intervention.

The concept of functional diversity links bird species diversity to ecosystem processes through resource-use patterns (Petchey and Gaston 2002; Tscharntke et al. 2007). In this case we are looking for patterns with respect to body mass, beak measures, diet, habitat and status (resident or migratory).

2.5.1 The Data Set

Six land-uses were identifying, from forest to pastureland: secondary forest, riparian forest, forest fallows, live fences, pastureland with high tree cover, and pastureland with low tree cover. Data in riparian forest and life fence were collected in four counting points every 100 m along linear transect, while data in other land uses were collected in 1 ha plots following the methodology by Vilchez et al. (2007). The data were collected for the project FRAGMENT (Developing methods and models for assessing the impacts of tree on farm productivity and regional biodiversity in Fragmented Landscapes).

The database used for this example comprises 56 species. Individuals of each species were inspected to record beak features (nares, width, and depth), wing-spread and body weights. After identification, diet source were classified as primary, secondary, tertiary and fourth preference. Each species was classified as migratory or resident and the habitat in the study area were recorded using three categories: generalist, open areas, and covered areas (forests). These traits are associated to the provision of regulatory ecosystem services due to bird participation in biodiversity conservation and ecosystem services in fragmented landscapes.

Foraging guild classifications of each species is also part of the important information needed to interpret the role of functional groups. This variable was not used in the cluster procedure, which is based on individual bird characteristics like beak features that are associated with guilds.

To perform functional groups we used InfoStat software (website: http://www.infostat.com.ar, Di Rienzo et al. 2010). File 'Traits by bird sp Nicaragua.IDB2' (available for download via Springer's Extra Materials website: http://extras.springer.com/) has trait information for 56 species recorded, including status and feeding guild. The status variable is nominal with two exclusive categories, and habitat variable is also nominal with three exclusive categories. We transform them to dummy variables (see Sect. 2.2 and Box 2.2). For status we need one column (status_R), and for habitat we need two columns (habitat_G and habitat_OA). File 'Res traits by bird sp Nicaragua.IDB2' has extra columns with these new variables (available for download via Springer's Extra Materials website: http://extras.springer.com/).

File 'Bird sp by use Nicaragua.IDB2' (available for download via Springer's Extra Materials website: http://extras.springer.com/) has the abundance (number of

individuals) of each species in the six land uses. Double click the label of the column 'Land_use' to see the codification used for land-use.

2.5.2 Bird Functional Types from Nicaragua

In this example we have three types of numerical variables, those derived from the dummy transformation, those indicating the ordinal value of feeding preference and those measures in a continuous scale. To perform cluster analysis we used the same procedure as in previous example. We select Gower as a measure of similarity because we have continuous and dummy variables. The software selects automatically the appropriate transformation to distance measure and evaluate the distance matrix between species. We also select Ward as the linkage algorithm and ask for two clusters.

The resulted dendrogram has a clear difference between two main branches, but heterogeneity among species within the groups is still high. We run again asking for five groups and obtained the dendrogram shown the five groups left to the vertical line in Fig. 2.3. File 'Res traits by bird sp Nicaragua.IDB2' (available for download via Springer's Extra Materials website: http://extras.springer.com/) has the identification to which cluster each species belong.

2.5.3 Characterization of Bird Functional Types of Nicaragua

Traits nares, wingspread and weight are important to differentiate functional types, also the migratory species are cluster together and with the resource-use patterns we complete the characterization of groups. Mean values for continuous variables (Table 2.3), proportion of nominal categories and mean importance values for feeding categories (Table 2.4) allows the full characterization of the bird functional groups.

The functional groups are:

- Nectarivorous: Mainly nectar-feeding birds of the family Trochilidae (all species of hummingbirds) with beak and body size very small with the largest less nares.
- Migratory generalist (Migr-Gen): Birds of small and medium size, consisting mainly of migratory species with the larger ratio wing/weight but Nectarivorous. They prefer habitat generalist, these species that can live in forest edges, secondary growths but not in open areas. They feeding mainly on insects. All migratory species are in this group.
- Insectivorous specialists (Ins-Spec): Birds with small to medium sizes, composed of birds only of covered habitat. Most of these species feed on insects in the understory.

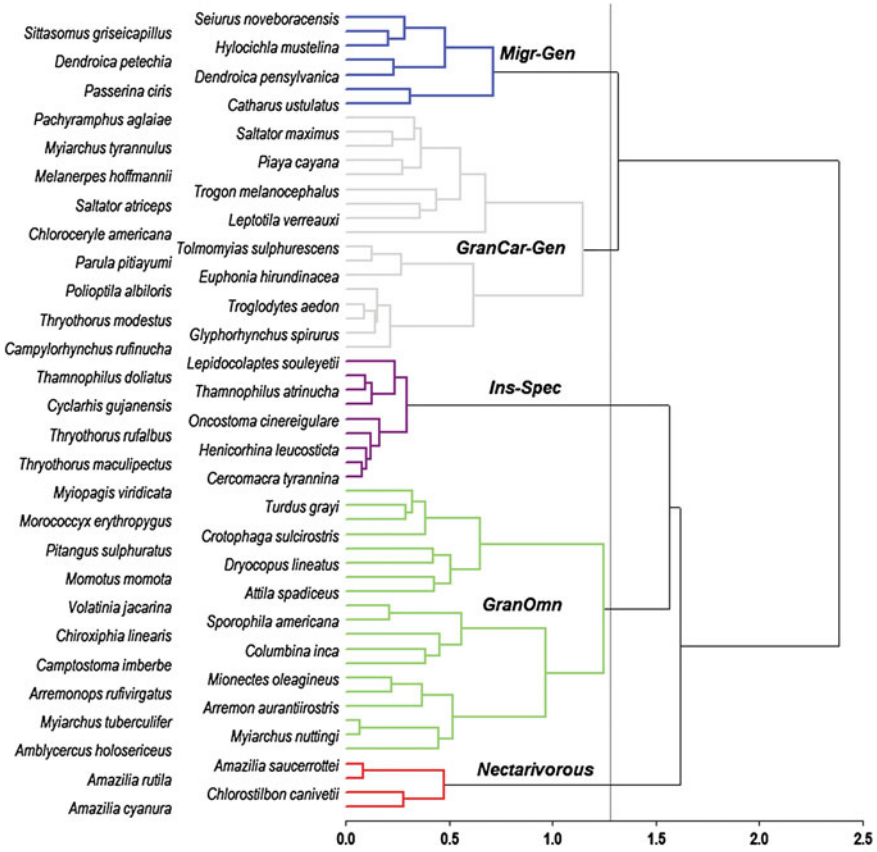


Fig. 2.3 Bird functional types in Nicaragua. Five groups of species were defined using Gower similarity coefficient transform to distance as square root of one minus Gower and Ward linkage algorithm. *Ins-Spec* insectivorous specialists, *GranCar-Gen* granivorous and carnivorous generalists, *GranOmn* granivorous and omnivorous, *Migr-Gen* migratory generalists, and *Nectarivorous*

- Granivorous and omnivorous (GranOmn): Bird species with the highest weight, wing and beak measures, except the nares; they are resident species with omnivorous feeding habits. This functional group has bird species foraging in open areas like pastures and crops, or species foraging in cover areas like forest.
- Granivorous and carnivorous generalists (GranCar-Gen): The birds of this group are generalist species of medium to large size, they can live in forest edges, secondary growths and advanced youth and scattered trees. The species of this group may be carnivores (fish), granivorous and insectivorous.

Table 2.3 Mean trait values of the five bird functional types in Nicaragua

Functional type	Richness	Wing	Weight	Nares	Width	Depth
Ins-Spec	9	64.72	20.97	12.35	5.53	4.97
GranCar-Gen	17	87.89	42.09	14.53	6.42	6.81
GranOmn	19	89.07	50.11	14.29	10.35	8.18
Migr-Gen	7	76.85	20.76	9.37	5.12	4.21
Nectarivorous	4	52.38	4.06	17.41	3.06	2.04

Ins-Spec insectivorous specialists, *GranCar-Gen* granivorous and carnivorous generalists, *GranOmn* granivorous and omnivorous, *Migr-Gen* migratory generalists, and *Nectarivorous*. Weights in grams, and beak measures and wing in millimeters

Table 2.4 Mean feeding preference for bird functional types in Nicaragua

FTypes	Fish	Invertebrates	Seeds	Small_fruits	Amphibians	Reptiles	Nectar
Ins-Spec	0.00	4.00	0.00	0.00	0.00	0.00	0.00
GranCar-Gen	0.24	3.29	0.65	2.12	0.00	0.12	0.12
GranOmn	0.00	3.26	1.42	1.89	0.42	0.53	0.00
Migr-Gen	0.00	3.14	1.00	2.29	0.00	0.00	0.00
Nectarivorous	0.00	0.00	0.00	0.00	0.00	0.00	4.00

Ins-Spec insectivorous specialists, *GranCar-Gen* granivorous and carnivorous generalists, *GranOmn* granivorous and omnivorous, *Migr-Gen* migratory generalists, and *Nectarivorous*. Scale of mean feeding preference follows variable codification: the highest value is four, when all the species of the functional type have the food category as first preference; lowest value is zero when no species eat that food in the functional type

2.5.4 Relationship of Functional Types with Land Uses

We performed a correspondence analysis to explore relationships between five functional types of birds with six land uses categories (Fig. 2.4). This analysis allows showing bivariate observations in a plane and identify the heaviest associations between patterns of two qualitative variables, in our case the functional types with the land uses. Correspondence analysis evaluates which are the combinations of modalities that have more inertia, which contributes most to reject the hypothesis of independence between the two variables. Points on the graph having a similar profile for land use are very close; those having similar functional type profiles are also very close. The distances from the origin indicate the discrepancy between the functional types from the average profile. The same applies to land uses categories. The distances between functional types and land use category's has no direct interpretation, but points in the graph fall in the same direction (relative to the origin) are positively correlated, while those that fall in opposite directions are negatively correlated. To perform this analysis we have to concatenate information from two data tables. The concatenated file 'Concat bird sp by use Nicaragua.IDB2' (available for download via Springer's Extra Materials website: http://extras.springer.com/) has the information of the species present in

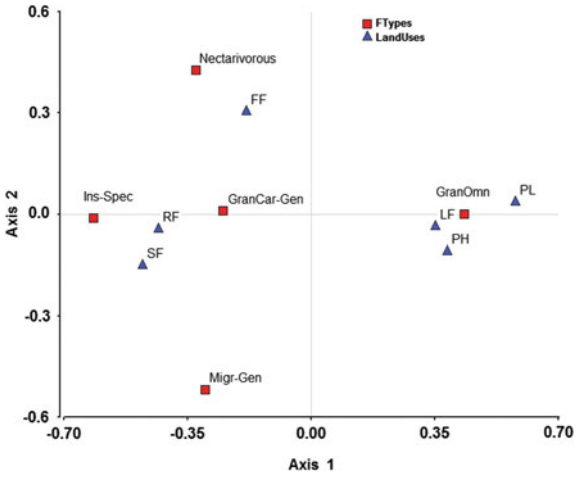


Fig. 2.4 First two axes from correspondence analysis between functional types of birds and land uses conditions in Nicaragua. The first axis explains the 85.50% of total variability and the second axis explains 11.65% of total variability. Five functional types: *Ins-Spec* insectivorous specialists, *GranCar-Gen* granivorous and carnivorous generalists, *GranOmn* granivorous and omnivorous, *Migr-Gen* migratory generalists, and *Nectarivorous*. Six land use conditions: *RF* riparian forest, *SF* secondary forest, *FF* forest fallows, *LF* live fences, *PH* pastureland with high tree cover, *PL* pastureland with low tree cover

each land use and the functional group to which the species belong. We used abundance (number of individuals) to weight the presence of each species in the land use categories.

The first two axes explain 97.15% (85.5% for axis one and 11.65% for axis two). There is a strong association of group GranOmn with pastures having high and low density of trees, and with live fences. Nectarivorous group has a strong association with secondary growths vegetation like forest fallows because this habitat provides pioneer plant species, increasing the availability of flowers for bird species of this group. Species of the functional type frugivorous and insectivorous specialists (Ins-Spec) are associated with secondary forests and riparian forests as expected because these bird species need to forage at understory. The GranCar-Gen type is mainly associated with riparian forests and species of Migr-Gen functional type prefer forest than pasture lands.

References

Calinski RB, Harabasz J (1974) A dendrite method for cluster analysis. Commun Stat 3(1):1–27

Carpenter SR, Mooney HA, Agard J, Capistrano D, DeFries R, Díaz S, Dietz T, Duriappah A, Oteng-Yeboah A, Pereira HM, Perrings C, Reid WV, Sarukhan J, Scholes RJ, Whyte A (2009) Science for managing ecosystem services: beyond the millennium ecosystem evaluation. Natl Acad Sci USA 106:1305–1312

Clarke KR (1993) Non-parametric multivariate analysis of changes in community structure. Aust J Ecol 18:117–143

Cornelissen JHC, Lavorel S, Garnier E, Díaz S, Buchmann N, Gurvich DE, Reich PB, Ter Steege H, Morgan HD, Heijden MGA, van der Pausas JG, Poorter H (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Aust J Bot 51:335–380

De Bello F, Lavorel S, Díaz S, Harrington R, Cornelissen JHC, Bardgett RD, Berg MP, Cipriotti P, Feld CK, Hering D, Marins da Silva P, Potts SG, Sandin L, Sousa JP, Storkey J, Wardle DA, Harrison PA (2010) Towards an assessment of multiple ecosystem processes and services via functional traits. Biodivers Conserv 19:2873–2893

Di Rienzo JA, Casanoves F, Balzarini MG, Gonzalez L, Tablada M, Robledo CW (2010) InfoStat versión 2009. Grupo InfoStat, FCA, Universidad Nacional de Córdoba

Díaz S, Lavorel S, Stuart Chapin F, Tecco PA, Gurvich DE, Grigulist K (2007) Functional diversity—at the crossroads between ecosystem functioning and environmental filters. In: Canadell JG, Pataki DE, Pitelka LF (eds) Terrestrial ecosystems in a changing world. Springer, New York

Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131(2):479–491

Fraley C, Raftery AE (2002) Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc 97:611–631

Fraley C, Raftery AE (2006) MCLUST Version 3 for R: normal mixture modelling and model-based clustering, technical report no. 504. Department of Statistics, University of Washington, Seattle, WA, USA

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857–874

Hartigan J (1975) Clustering algorithms. Wiley, New York

Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull Soc Vaudoise Sci Nat 44:223–270

Kaufman L, Rousseeuw P (1990) Finding groups in data: an introduction to cluster analysis. Wiley, New York

Lavorel S, Grigulis K, Lamarque P, Colace MP, Garden D, Girel J, Pellet G, Douzet R (2011) Using plant functional traits to understand the landscape distribution of multiple ecosystem services. J Ecol 99:135–147

MEA (Millennium Ecosystem Assessment) (2005) Ecosystems and human well-being: biodiversity synthesis. World Resources Institute, Washington

Petchey OL, Gaston KJ (2002) Functional diversity (FD), species richness and community composition. Ecol Lett 5:402–411

Polania C, Pla L, Casanoves F (2011) Diversidad funcional y servicios ecosistémicos. In: Casanoves F, Pla L, Di Rienzo JA (eds) Valoración y análisis de la diversidad funcional y su relación con los servicios ecosistémicos. Serie Técnica 384, CATIE, Turrialba

Pollard KS, van der Laan MJ (2005) Cluster analysis of genomic data. In: Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S (eds) Bioinformatics and computational biology solutions using R and bioconductor. Springer, New York

Pollard KS, van der Laan MJ, Wall G (2009) Hopach: hierarchical ordered partitioning and collapsing hybrid (HOPACH). R-package version 2.4.0. http://CRAN.R-project.org/package=hopach

Sarle WS (1983) The cubic clustering criterion. SAS technical report A-108. SAS Institute, Cary

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. Univ Kansas Sci Bull 38:1409–1438

Tibshirani R, Walther G, Hastie T (2001) Estimating the number of clusters in a dataset via the gap statistic. J R Stat Soc B Methodol 63(2):411–423

Tscharntke T, Sekercioglu CH, Dietsch TV, Sodhi NS, Hoehn P, Tylianakis JM (2007) Landscape constraints on functional diversity of birds and insects in tropical agroecosystems. Ecology 89:944–951

Valdano SG, Di Rienzo JA (2007) Discovering meaningful groups in hierarchical cluster analysis. An extension to the multivariate case of a multiple comparison method based on cluster analysis. Available via InterStat Journal. http://interstat.statjournals.net subordinate document http://interstat.statjournals.net/YEAR/2007/articles/0704002.pdf. Accessed 13 May 2011

Vilchez S, Harvey C, Sánchez D, Medina A, Hernández B, Taylor R (2007) Diversidad y composición de aves en un agropaisaje de Nicaragua. In: Harvey C, Sáenz JC (eds) Evaluación y conservación de la biodiversidad en paisajes fragmentados de mesoamérica. INBio, Heredia

Chapter 3
Functional Diversity Indices

Abstract Functional diversity may be summarized using indices based on trait values and species importance in the community, like abundance, cover and biomass. This chapter includes taxonomy of the indices based upon the information they used and the output they offer. For each index we have included its definition, the information needed to estimate it, their statistical and ecological properties, and some reference to explore its application to real cases. To facilitate the comprehension of all indices and diversity measures we used homogeneous notation.

Keywords Single-trait indices · Weighted diversity indices · Multiple-trait indices · Taxonomic biodiversity indices · FDiversity software

3.1 About Functional Diversity Indices and Measures

In this chapter, several indices and their definitions will be introduced. To avoid redundancy, Box 3.1 shows the notation used in this and other chapters. When deviations from this notation occur, it will be made clear in the text. The code name for the indices, as well as the author's reference will be mentioned in each definition. Even though we are focused on functional diversity and functional diversity indices, Sect. 3.2 has a brief presentation and one example of species diversity indices. We include the most commonly used indices and those that are often compared with functional diversity indices.

L. Pla et al., *Quantifying Functional Biodiversity*, SpringerBriefs in Environmental Science, DOI: 10.1007/978-94-007-2648-2_3, © The Author(s) 2012

Box 3.1: Notation

Variable	Meaning
ij	Used as subscripts to identify species; $i, j = 1, \ldots, S$
t	Used as subscript to identify traits; $t = 1, \ldots, T$
S	Number of species
T	Number of traits
a_i	Absolute abundance of the *i*th species (with units; i.e. m^2 for cover)
w_i	Relative abundance of the *i*th species; $w_i = a_i \Big/ \sum_{i=1}^{S} a_i$
x_{ti}	*t*th trait value of *i*th species; for single-trait indices the subscript t is not necessary

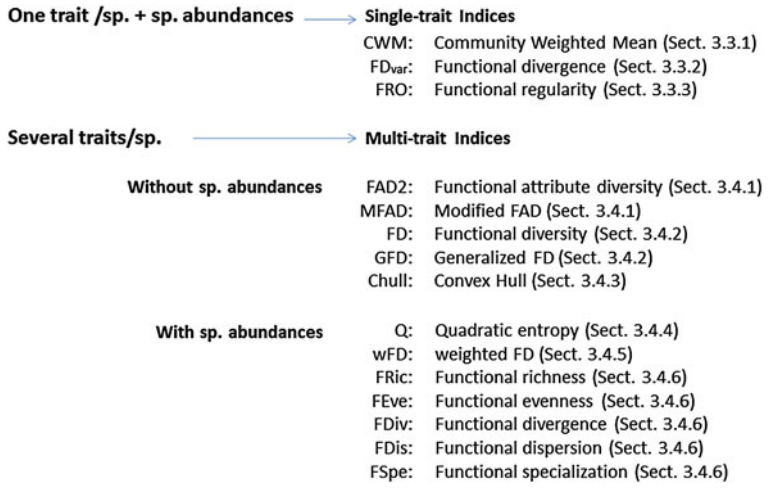


Fig. 3.1 Index classification with reference to sections in this chapter

There are single-trait metrics and indices that include some measurement of abundance to load the contribution of each species to the diversity aspect to be summarized. For example, the community weighted mean (CWM) reflects the functional mean of the trait, the functional divergence (FDvar) reflects the variance of the trait, while the functional regularity index (FRO) measures the functional evenness (Fig. 3.1). The classification criterion in Sects. 3.3 and 3.4 is if the metrics are single-trait or multi-trait. CWM is included as single-trait measure, although there is a fundamental difference between the CWM and the diversity indices: there is not any ecological principle to propose that the CWM of one trait would have any correlation with the functional diversity definition of the community.

Undoubtedly, CWM is a measure that allows knowing the best value to represent the state of a trait in the community; it is a metric of functional composition. Definitions and examples of single-trait indices are presented in Sect. 3.3.

The indices based on multi-trait profiles, may or may not take into account species abundance. In the last decade of the twentieth century, Faith (1996) proposed a functional diversity index based on the application of the index of environmental diversity (ED) to a functional space constructed using phylogenetic information, and Walker et al. (1999) defined a functional diversity index by identifying the number of different combinations of trait values (FAD1). Most of the later approaches are based on dissimilarity among species in trait space, the T-dimensional space defined by the T traits (Fig. 3.1). In Sect. 3.4 we define the indices and present an example. The multi-trait indices that may be estimated without abundance information are based only on presence/absence data; to estimate the other multi-trait indices it is needed to incorporate species abundances. There is some controversy in current literature about how the impact of summarizing functional diversity and its relationship with the variable is used to measure abundance. Posed questions focus on the relative contribution of abundance upon the functional diversity components, like the functional richness, the functional evenness and the functional divergence (Schmera et al. 2009; Poos et al. 2009; Mouchet et al. 2010).

To estimate an index we used abundance-trait profiles corresponding to the species present at the community or assemblage. This profile includes one value for each trait and one value for each variable used to quantify the relative contribution of each species to the pool. At least values of one trait by species are needed to compute single-trait indices and values for two or more traits to compute multi-trait indices. Nevertheless, it is analytically possible to calculate some multi-trait indices with values for only one trait. For example, the FD (Petchey and Gaston 2002) diversity index (Sect. 3.3.1) may be calculating with species values for one trait.

3.2 Species Diversity Indices

Richness (*S*): richness is the total number of species in a community. It is the most simple biodiversity index and it does not take into account any characteristic of species or their relative abundance. Several estimators to avoid bias due to unseen or unrecorded species are currently used like those proposed by Chao et al. (for details see for example Magurran 2004). In a functional diversity context it is common practice to include enough species to account for near 80% of the abundance (Grime 1998). This approach does not need to adjust observed species number because functional diversity is not affected by rare species.

Shannon Index (*H*): this index assumes that heterogeneity depends on both, the number of species in a community and their proportional abundances. Conceptually, it is a measure of uncertainty degree related to a random selection of

individuals from the community. In a homogeneous community with S species, in which only one is dominant (highly abundant), the uncertainty degree of selecting a given species is lower than if all species would have the same abundance. When species abundances are uneven, the probability that a random individual taken from the population belongs to the dominant species approximates one; conversely, in a heterogeneous community any randomly drawn individual has the same probability (1/S) to belong to any species. The Shannon index is one of the biodiversity measures most widely used.

The Shannon index assumes that individuals are randomly sampled from an "indefinitely large" (i.e., an effectivelly infinite) population. The index also assumes that all species are represented in the sample. It is calculated from the equation (Shannon and Weaver 1949)

$$H = -\sum_{i=1}^{S} w_i \ln(w_i).$$

Eveness (*E*): the maximum of Shannon index is attended when all the species has the same relative abundance, and it reduces to

$$H_{max} = \ln(S).$$

Based on this maximum it is possible to derive an evenness index, using the maximum as reference for the actual value (Pielou 1975) as the ratio

$$E = H/H_{max},$$

where E is the evenness index.

Simpson Index (*D*): Simpson index of biodiversity equals the probability of drawing without replacement two individuals of different species from a given collection. There is more than one form to express the index. The expression widely used offers the index as a measure of dominance

$$D = 1 - \sum_{i=1}^{S} w_i^2,$$

where w_i^2 is the squared of the proportion of the *i*th species (Simpson 1949).

3.3 Single-Trait Metrics and Indices: Properties and Estimation

3.3.1 Community Weighted Mean

Community weighted mean (CWM) is a good indicator to represent the expected functional value of one trait in a random community sample. Also defined as aggregate values of plant traits by Garnier et al. (2004) it is extensively used as

community weighted mean (Díaz et al. 2007; Lavorel et al. 2008). As CWM is based on single-trait, each trait has its CWM value in the assemblage. To calculate it we need one trait value to represent each species, so if we have 10 species and we want to calculate CWM of specific leaf area (SLA) we need 10 SLA's values (x_i in Box 3.1). We also need one variable representing the contribution of each species to the community. Suppose that for each species we have evaluated the cover in square meters (a_i in Box 3.1) then we may express the cover in relative form (w_i in Box 3.1). The community weighted mean is

$$\mathrm{CWM} = \sum_{i=1}^{S} w_i \, x_i. \tag{3.1}$$

It is strongly recommend following standard protocols to assign trait values to species (Cornelissen et al. 2003). Depending on the trait variability it could be necessary to measure 5–10 representative individual for each species. After having the data set for the whole community the mean value for continuous variables or median for discrete ones may represent each species in the index calculation. There are several categories already defined in the literature for particular traits like flammability (Cornelissen et al. 2003; Jaureguiberry et al. 2011), or nodule type for nitrogen-fixing species (Cornelissen et al. 2003).

To apply the formula (3.1) to the example in Box 3.2 we have to calculate the relative contribution of each species from the cover data in its original scale divided by the total coverage, in this case it is 10,042 m^2, doing so we obtained the values of the relative cover column. Having all the data in the appropriate scale, we multiply each SLA value by the corresponding relative value and sum all results to obtain the CWM. In the example its value is 19.01 $mm^2\ g^{-1}$. This value is greater than the arithmetic mean (18.26 $mm^2\ g^{-1}$) because it incorporates a loaded factor that in this case favors the *sp2* and *sp8* with greater SLA values. To summarize community performance related to ecosystem processes the CWM represents the best single value to link with other variables and look for relationships with a given ecosystem service.

Box 3.2: Trait profiles for ten species used to estimate CWM

Species	SLA (mm^2g^{-1})	Cover (m^2)	Relative cover	Species	SLA (mm^2g^{-1})	Cover (m^2)	Relative cover
sp1	19.30	245	0.0235	sp6	13.81	312	0.0300
sp2	19.53	2540	0.2439	sp7	9.94	780	0.0749
sp3	15.64	34	0.0033	sp8	21.93	3545	0.3405
sp4	18.44	2045	0.1964	sp9	31.65	108	0.0104
sp5	17.37	35	0.0034	sp10	14.98	768	0.0738

3.3.2 Functional Divergence

Functional divergence index (FD_{var}) is essentially the variance in the attribute values of the species present at a site, with the squared residuals weighted by the abundance of the species involved (Mason et al. 2003). It is defined as

$$FD_{var} = \frac{2}{\pi} \arctan(5V) \tag{3.2}$$

where 5 is a scaling factor used to define the index over a range of 0–1; V is the weighted variance of trait X, expressed as:

$$V = \sum_{i=1}^{S} w_i \left(\ln x_i - \overline{\ln x}\right)^2. \tag{3.3}$$

This index considers one trait at a time and used the relative abundance of each species (w_i) to load its contribution to the variability in the community (Box 3.1). The mean of $\ln x_i$ is weighted by the abundance as

$$\overline{\ln x} = \sum_{i=1}^{S} w_i \ln x_i. \tag{3.4}$$

Using data from Box 3.2 the FDvar is 0.15 and this index has no units because the trait values, originally expressed in squared millimeters of leaf area divided by dry weight in grams has been transformed to a logarithm scale and expressed in the range zero–one. So this value of 0.15 corresponds to a small variability for SLA. If we interchange in Box 3.2 the cover of species *sp7* and *sp8* but keeping the SLA values and recalculate the FDvar, we obtain FDvar = 0.30. This is twice the first estimation, and it is the consequence of assigning a cover of 3,545 m^2 to *sp7* with SLA = 9.94 $mm^2\ g^{-1}$ (one of the smallest values for SLA). The variability of this trait increases due to that more abundant species now bear the more extreme values for SLA (9.94, 19.53 and 18.44 $mm^2\ g^{-1}$).

This index has also been defined to handle more than one value of the trait by species (Mason et al. 2003) using the character values (x_i) and the abundance of these values in all the species of the community. For this functional divergence formulation the sum is over the total possible values of the trait under consideration.

3.3.3 Functional Regularity

Functional regularity index (FRO) has been defined for one trait with only one value of the trait by species, like the mean or the median. FRO was introduced to capture a neglected aspect of functional diversity as is the regularity or evenness of the trait values in the observed range (Mouillot et al. 2005). As the other single-trait indices it also used the relative abundance of each species. The procedure to

calculate the index needs to sort the observations. It is as follows: (a) the species are ranked by increasing values of the trait (x_i); (b) we calculate the weighted difference $\left(EW_{i,i+1}\right)$ of trait values of two consecutive species loaded by the abundance difference as

$$EW_{i,i+1} = \frac{|x_{i+1} - x_i|}{|w_{i+1} + w_i|} \tag{3.5}$$

where w_i is the relative abundance of the ordered *i*th species; (c) with these values, we calculate the percentage of the weighted difference ($PEW_{i,i+1}$) in trait values for the pair of species as

$$PEW_{i,i+1} = \frac{EW_{i,i+1}}{\sum_{i=1}^{S-1} EW_{i,i+1}}; \tag{3.6}$$

and (d) the FRO index results from the summation of all $S - 1$ pair comparison, choosing the minimum between the percentage of the weighted difference and the equally probable space $1/(S - 1)$

$$FRO = \sum_{i=1}^{S-1} \min\left(PEW_{i,i+1}, \frac{1}{S-1}\right). \tag{3.7}$$

The maximum of FRO is obtained when each pair of nearest neighbors equals $1/(S - 1)$ and each species has the same abundance, case of maximum functional regularity and FRO = 1. In all other cases FRO is less than one; and quantifies how the observed community differs from other communities with the same richness, where all species have the same abundance, and its trait values are regularly distributed resembling to the outcome of the uniform probability distribution.

Using the data of Box 3.2 we obtained a FRO = 0.39 (Fig. 3.2a), approximately one third of a community where the ten species have equal abundance and its trait values are uniformly distributed in the range of SLA (9.94 to 31.65 $mm^2\ g^{-1}$). Suppose that due to human modification or to environmental process the relative abundance of each species is near 0.10 (Fig. 3.2b). In this case the FRO increase to 0.60. FRO may also increase if the trait values are more evenly distributed. If we interchange abundance of species *sp7* and *sp8*, and also between *sp9* and *sp10*, we obtained FRO = 0.47 (Fig. 3.2c).

All these changes in FRO have happened with the same set of trait values, but with modifying the matching between trait value and the relative abundance. The index may also be affected by changes in the trait values. For example, if the species *sp1*, *sp5* and *sp10* change their values to 29.30, 27.37 and 24.98 $mm^2\ g^{-1}$, the FRO with the observed abundance would be 0.65; and if we considered all species with the same abundance it would be 0.89, closer to one, due to a very uniformly distribution of equally abundance trait values (Fig. 3.2d).

Even though FRO is defined for one trait at a time, the authors (Mouillot et al. 2005) suggested two options to extend the index to multiple traits. One is to estimate FRO for each trait and then take the mean value to represent the

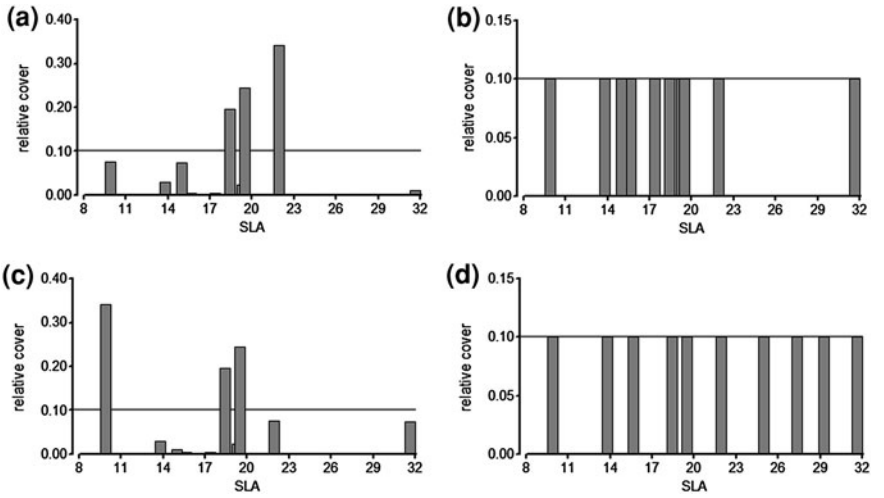


Fig. 3.2 Relative abundance for specific leaf area (SLA) trait values of species in Box 3.2. **a** Original data, corresponding to FRO = 0.39; **b** species with the same abundance (relative cover 0.10), corresponding to a FRO = 0.60; **c** interchange abundance values between sp7 and sp8, and between sp9 and sp10, corresponding to FRO = 0.47; **d** change the SLA values of sp1, sp5 and sp10 to 29.30, 27.37 and 24.98 mm^2g^{-1} with the same species abundance (0.10), corresponding to a FRO = 0.89. Note y-axis scale in **a** and **c** is different from y-axis scale in **b** and **d**. Horizontal line is at relative cover 0.10 in the four plots

community; this mean value may be weighted by trait importance if desirable. The second one is to compute a Principal Component Analysis on the S × T matrix and use the principal components to compute T values of FRO and then sum using eigenvalues or standard deviation of eigenvectors as weights to compute the overall functional regularity (OFRO)

$$\mathrm{OFRO} = \sum_{t=1}^{T} \mathrm{SD}_t \times \mathrm{FRO}_t,$$

where SD_t the standard deviation and FRO_t the functional regularity for each principal component.

3.4 Multi-Trait Indices: Properties and Estimation

3.4.1 Functional Attribute Diversity

The index FAD (Functional Attribute Diversity, Walker et al. 1999; Walker and Langridge 2002) has two expressions. FAD1 is the number of different attribute combinations that occur in the community and it is always less than or equal to richness. When traits are in a continuous scale it always coincides with richness

and its use is not recommended. When the traits are categorical, with few levels it may resume a crude functional richness; even though, there are more appropriate functional indices as further described.

The second expression (FAD2) is the sum of the standardized distance between all pairs of species in the trait space. The authors recommend expressing the trait values in a five-point scale. As an ecological distance (ED), they use the Euclidean distance between two species. The sum over all pairs of species gives the FAD2. The ecological distance between species 'i' and 'j' may be expressed as (Eq. 1a, from Walker et al. 1999)

$$ED_{ij} = \sqrt{\sum_{t=1}^{T} \left(x_{tj} - x_{ti}\right)^2}$$

where T is the total number of attributes, and x_{ti} and x_{tj} are the values of the *t*th trait of species 'i' and 'j'. Using ED_{ij}, which is the Euclidean distance between two species, FAD2 is defined as

$$FAD2 = \sum_{i=1}^{S} \sum_{j>1}^{S} ED_{ij}$$

in a community with S species.

In an attempt to make FAD2 comparable among communities with different number of species, the authors propose to standardize the index dividing by the number of interspecies comparisons. With S species the total number of distances among a pair of species is $S \times (S-1)/2$, and the index for each community to be compared may be transformed to comparable scales using

$$FAD2_{(Z)} = \frac{FAD2}{S(S-1)/2},$$

being $FAD2_{(Z)}$, the standardize expression.

MFAD is another modified version of FAD2 proposed by Schmera et al. (2009) to overcome the violation of monotonicity criteria. For a given assemblage with S species and T traits they first defined the so called functional units. The number of functional units results from combining the species with exactly the same trait profiles into only one functional unit. The number of entities in the data matrix will be reduced from S to N ($N \leq S$), and dimensions of the distance matrix will be reduced from $S \times S$ to $N \times N$. N itself is a measure of functional richness, already proposed by Walker et al. (1999) as FAD1.

To calculate MFAD, the dissimilarity metric must be defined in the range [0; 1], and the authors proposed the use of the Marczewski-Steinhaus index (Marczewski and Steinhaus 1958) or a distance defined in the same interval, like Gower distance (Gower 1971). The index is estimated as:

$$\mathrm{MFAD} = \frac{\sum_{i=1}^{N} \sum_{j>1}^{N} d_{ij}}{N},$$

where d_{ij} is the dissimilarity between functional units 'i' and 'j', and N is the number of functional units.

MFAD measures the dispersion of species in the trait space in comparable scales if the set of traits are the same. So it provides a simple numerical tool to compare several communities. Several authors have claimed that its contribution to functional diversity approach is not significant due to high correlation with richness (Mouchet et al. 2010; Pavoine and Bonsall 2011).

3.4.2 Functional Diversity Based on Dendrograms

There is a family of functional diversity indices based on dendrograms. The first one was proposed by Petchey and Gaston (2002) and has been used in functional ecology as 'the' index. Even a decade later this index and its code name FD is one of the most used in applied functional ecology. Several adjustments have emerged from scientific controversy between Petchey and Gaston (2002, 2006, 2007) and Podani and Schmera (2006, 2007); others from the inclusion of abundance to weight the relative contribution of each branch tree to the index (wFD, Pla et al. 2008; Casanoves et al. 2008, 2011).

FD is the total length of the branches of the dendrogram constructed from information on species functional traits (Petchey and Gaston 2002). Different measures of dissimilarities, and different strategies used to define the dendrogram lead to different values of FD. As in the hierarchical clustering algorithm several linking strategies can be used and the number of distance measurement can be calculated to quantify the distance between species, Mouchet et al. (2008) have proposed iteratively select the best combination of linkage and distance to reproduce the original distance between the species pool. The selection is based on the cophenetic correlation and the index is known as generalized-FD (GFD).

The first definition of FD (Petchey and Gaston 2002) proposed to make a single dendrogram for the so-called 'regional community', with all the species that have been observed in any sample of the study area. Based on this maximum value for FD, any particular sample of the study area will have its own FD resulting from the sum of branch length of the species present at the considered sample, required to connect these species to the root of the dendrogram. This expression was criticized because the index does not equal zero when only one species is present. It is zero only when no species is present (Podani and Schmera 2006).

These authors have proposed to recalculate the dendrogram for each sample, but in doing so the desirable property of 'set monotonicity' does not hold. That is, the index may be greater for a community after one (or more) species is lost; and

conversely the FD may be smaller for a community which has gained one or more species.

The second definition of FD as the sum of lengths of all branches of the dendrogram made by Petchey and Gaston (2006) addresses the lack of monotonicity that arises when using a particular dendrogram for each sample as Podani and Schmera had proposed. To calculate FD in each sample of the community, this version of FD sums the lengths of the branches on the dendrogram needed to connect the present species, no including the length of the branch to get to the root (Petchey and Gaston 2007).

Both versions of the FD are based on a single trait value per species. However, it is possible to incorporate intraspecific variability in the estimation when trait values are available at individual level. This functional diversity index incorporates intraspecific variability and it is less correlated with richness (iFD, Cianciaruso et al. 2009). The authors claim that important ecosystem processes operate at individual level, like competition for resources, niche occupancy and so natural selection. A particular value taken by the trait is an attribute of the individual; within a species the trait may show different attributes along environmental gradients or through time, or among different land use practices.

One way to avoid subjectivity in the selection of the distance measurement and the linkage strategy is to compare the ultrametric matrix computed to make the dendrogram with the distance matrix calculate from the functional trait profiles. Even though, no clustering procedure perfectly fits data distribution in multidimensional space. Mouchet et al. (2008) propose a systematic procedure to calculate combinations of distances (Euclidean and Gower) and several cluster linkages: single linkage, complete linkage, UPGMA (unweighted pair group method using arithmetic averages), WPGMA (weighted pair group method using arithmetic averages), UPGMC (unweighted pair group centroid method), WPGMC (weighted pair group centroid method) and Ward's method; and then, build the consensus tree that optimally represents the clustering methods.

This index is called Generalized FD (GFD) and the selection of the best combination is based on the comparison between dissimilarity matrix and cophenetic matrix. The cophenetic matrix is an $S \times S$ symmetrical matrix that quantifies the distance between species in the dendrogram. The less difference between the corresponding elements of these two matrices, the better the cluster procedure resembles the diversity in trait space. The authors used the cophenetic correlation (Pearson correlation computed from pairs of distances) to select the combination that best fits. If the cophenetic correlation is large, the distance portrayed in the dendrogram is a good representation of distances between species, in the trait space. R script may be downloaded from Ecolag author's site (http://www.ecolag.univ-montp2.fr) and FDiversity software also calculates it from the same information used for the other functional diversity indices.

Further discussion about dissimilarity measurements, cluster strategies and comparison among communities can be found in Petchey et al. (2004) and Petchey and Gaston (2007), the response of Podani and Schmera (2007) and Poos et al. (2009).

FDiversity team[1] (Pla et al. 2008; Casanoves et al. 2011) proposed using abundance to quantify species contribution to the community when a hierarchical clustering approach is used for functional diversity (wFD, Sect. 3.4.5).

3.4.3 Convex Hull Hyper-Volume

The dispersion of species in the trait space is a crude multivariate representation of the functional diversity in a community. Cornwell et al. (2006) proposed to synthesize this dispersion by quantifying the best shape hyper-volume with an appropriated volume model. Among the candidates are hyper-cube or hyper-sphere models, but these do not reduce the amount of empty space. A better option is to use the convex hull, defined as the smallest convex set enclosing the points (Barber et al. 1996).

A convex hull hyper-volume (CHull) in a multivariate space is defined, based on the irregular form yielded by species occupancy in the trait space. Taking two species from one community, any third species with traits inside the range of traits is included in the CHull. If only two traits are involved, the CHull may be represented with a surface in 2D (Fig. 3.3a–c); when there are three traits the CHull is a volume in 3D (Fig. 3.3d), and with four or more traits CHull is a hypervolume. The convex hull is a multivariate measure of the range of trait space (trait values that may be found in a given assemblage). The sequence from Fig. 3.3a–c shows increasing trait ranges; the three graphs have the same units because the trait values were standardized to have zero mean and unit variance. T3, the third trait has the widest range, so the combination T3–T2 has the highest surface, all expressed in standard deviations from the corresponding mean.

3.4.4 Quadratic Entropy

Functional diversity may be expressed as the average of the species differences when some measurement of pairwise differences between species and relative frequency data are available. The index proposed by Rao (1982) is derived from entropy theory and is expressed as a quadratic form using the matrix of distances among species and the vector of relative abundance of species. To compute the index it is necessary to calculate the Euclidean distance between species in the trait space as

$$d_{ij} = \sum_{t=1}^{T} \left(x_{tj} - x_{ti}\right)^2$$

[1] FDiversity team is integrated by the authors of this book and is the developer team of FDiversity, statistical software to calculate functional diversity with extended capabilities (Di Rienzo et al. 2008).

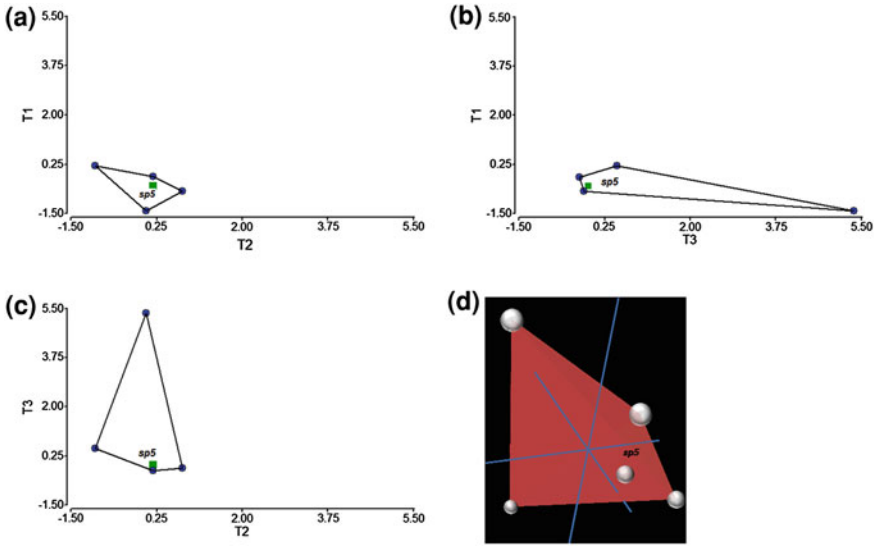


Fig. 3.3 Convex hull (CHull) in two and three dimensions. The data are for three traits and five species; **a** surface delimited by the outer species in plane T1-T2 (CHull = 1.140), **b** surface in plane T1-T3 (CHull = 2.941), **c** surface in plane T3-T2 (CHull = 4.947), **d** volume in the three dimensional space T1-T2-T3 (CHull = 0.462); sp5 has intermediate trait values and it is identify within the surfaces and also within the volume. Data are standardized and CHull values showed at the top of each graph. A similar limit for x–y axes of the first three graphs allows visual comparison among CHull values. Trait values were standardized by trait

and estimate Rao index as (Rao 1982)

$$\text{Rao} = \sum_{i=1}^{S-1} \sum_{j>1}^{S} d_{ij}\ w_i w_j = \frac{1}{2}\mathbf{w}'\,\mathbf{D}\,\mathbf{w}$$

where d_{ij} is the distance between species 'i' and 'j'. In matrix notation, **D** is a distance matrix with elements d_{ij}, and **w** is a column vector with the relative abundances.

Botta-Dukát (2005) has suggested using Euclidean distance divided by the number of traits used to define it. To be compared, functional diversity indices has to be evaluated over the same set of traits, so all the distances are calculated over the same number of traits and this adjustment has no effect for comparison among samples or communities. The new expression is the original divided by a constant.

The Rao index may also be seen as the expected value of the conflict among species (Ricotta and Szeidl 2006). As the species abundances are expressed as relative values, it sums to one $\sum_{i=1}^{S} w_i = 1$ and the frequency of any species may be expressed as $1 - \sum_{j \neq i}^{S} w_j$. So, the conflict between species 'i' and the remaining $(C_d(w_i))$ may be express as

$$C_d(w_i) = \sum_{j \neq i}^{S} d_{ij} w_j.$$

The functional diversity is then

$$\text{Rao} = \sum_{i=1}^{S} w_i \, C_d(w_i) = \sum_{i=1}^{S} w_i \left[\sum_{j \neq i}^{S} d_{ij} \, w_j \right],$$

the second summation equals $\sum_{j=1}^{S} d_{ij} \, w_i$, since the distance of a species with itself is zero, $d_{ii} = d_{jj} = 0$. This expression proves that the Rao index is also a measure of the conflict among species (Ricota and Szeidl 2006).

The unbiased estimator of Rao when the abundance of species is expressed as number of individuals is

$$\hat{\text{Rao}} = \frac{n}{n-1} 2 \sum_{i>j}^{S} d_{ij} \frac{n_i n_j}{n^2}, \quad \hat{\text{Rao}} = 2 \sum_{i>j}^{S} d_{ij} \frac{n_i n_j}{n(n-1)},$$

where n_i is the number of individuals of species 'i' and $n = \sum_{i=1}^{S} n_i$ is the total number of individuals. If the sample is big enough the correction term $n/(n-1)$ is almost one and the index may be calculated without correction. The variance of Râo may be estimated as (Shimatani 2001)

$$\text{Var}(\hat{\text{Rao}}) = \frac{4}{S(S-1)} \left\{ \begin{array}{l} (3-2S) \left(2 \sum_{i>j}^{S} d_{ij} \frac{n_i n_j}{n^2} \right)^2 + \\ \qquad + (S-2) \sum_{i,j,k}^{S} d_{ij} d_{ik} \frac{n_i n_j n_k}{n^2} + \sum_{i>j}^{S} d_{ij} \frac{n_i n_j}{n^2} \end{array} \right\}.$$

The variance is useful with large samples and when differences between communities have to be tested based on one sample for each community. The distribution model for Rao index is not known and depends on the distance measure thus, non parametric estimation may be preferred to build confidence intervals and to test hypothesis.

Useful information may be extracted from the symmetric matrix $Q = \text{diag}(\mathbf{w})\mathbf{D}(\text{diag})\mathbf{w}$, where 'diag' states for diagonal matrix with relative abundances of each species. This matrix has dimension $S \times S$ and its *ij*th element is $q_{ij} = d_{ij} \, w_i w_j$, its main diagonal is zero, and it is known as the species contribution matrix. The absolute species contribution for each species may be evaluated summing along the columns of the contribution matrix, and a relative expression of this contribution is obtained dividing these values by 2×Rao index (Box 3.3). The reference must be twice the index because each distance, between two species, counts twice, one when sum is over the column for the first one of the pair and the other when the sum is over the second.

Box 3.3: Distance matrix and abundance vector used to estimate Rao index

$$X = \begin{array}{c} \begin{matrix} t_1 & t_2 & t_3 \end{matrix} \\ \begin{bmatrix} 0.20 & -1.00 & 0.51 \\ -1.41 & 0.03 & 5.35 \\ -0.20 & 0.18 & -0.27 \\ -0.70 & 0.77 & -0.18 \\ -0.52 & 0.18 & -0.08 \end{bmatrix} \end{array} \begin{matrix} sp1 \\ sp2 \\ sp3 \\ sp4 \\ sp5 \end{matrix}$$

X is the data matrix with
3 columns, one for each trait; and
5 rows, one for each species.
The trait values has been standardize (see Box 4.1 for details).

$$A = \begin{bmatrix} 70 & 27 & 1 & 140 & 17 \end{bmatrix}$$

A is the vector of observed frequencies.

$$\mathbf{w} = [0.274509\ 0.105882\ 0.003922\ 0.549020\ 0.066667]$$

W is the vector of relative frequencies.

With this data the distance matrix D may be calculated using the Euclidean distance

D =	*sp1*	*sp2*	*sp3*	*sp4*	*sp5*
sp1	0.0000	27.0786	2.1608	4.4190	2.2589
sp2	27.0786	0.0000	33.0710	31.6326	30.2995
sp3	2.1608	33.0710	0.0000	0.6062	0.1385
sp4	4.4190	31.6326	0.6062	0.0000	0.3905
sp5	2.2589	30.2995	0.1385	0.3905	0.0000

Each element of the contribution matrix Q may be calculated multiplying the distance between two species by the corresponding relative frequencies. For examples, for the first element of the second column the value 0.7871 was obtained as

$$q_{12} = 27.0786 \times 0.274509 \times 0.105882.$$

Q =	*sp1*	*sp2*	*sp3*	*sp4*	*sp5*
sp1	0.0000	0.7871	0.0023	0.6660	0.0413
sp2	0.7871	0.0000	0.0137	1.8389	0.2139
sp3	0.0023	0.0137	0.0000	0.0013	0.0000
sp4	0.6660	1.8389	0.0013	0.0000	0.0143
sp5	0.0413	0.2139	0.0000	0.0143	0.0000
Total	1.4967	2.8546	0.0173	2.5205	0.2695
Contribution (%)	20.9	39.9	0.2	35.2	3.8

Two species (*sp2* and *sp4*) have high contribution, other two has very low contribution (*sp3* and *sp5*) and *sp1* has a medium contribution to the functional diversity.

Both, absolute and relative contributions of each species are highly related to abundance. To explore the effect of species contribution upon the index, the partial derivative of the species contribution matrix (Q) with respect to the *i*th species contribution (n_j) may be computed as

$$\partial Q / \partial n_j = \frac{2}{n}\left(\sum_{i=1}^{S} d_{ij}\frac{n_i}{n} - \text{Rao}\right).$$

If $\sum_{i=1}^{S} d_{ij} n_i \Big/ n$ is greater than Q, a small increment of species 'j' increases the Rao diversity index (Shimatani 1999). This is because the first term is the average distance from an individual of species 'j' to all the others (see numerical example in Box 3.4), and the second term is the average distance over all the pairs; so, if the species 'j', has a greater average and its frequency increases, the overall average increases accordingly. On the other hand, if a species with very small contribution increases its relative abundance the index may decrease due to a negative difference between the species average distance and the overall average distance that is the Rao index.

Box 3.4: Species relative abundance changes affect the Rao index

Partial derivative of the species contribution matrix (Q)

	sp1	*sp2*	*sp3*	*sp4*	*sp5*
sp1	0.000000	2.867146	0.008474	2.426118	0.150593
sp2	7.433341	0.000000	0.129690	17.366918	2.019967
sp3	0.593161	3.501635	0.000000	0.332816	0.009233
sp4	1.213059	3.349334	0.002377	0.000000	0.026033
sp5	0.620090	3.208182	0.000543	0.214392	0.000000
Sum	9.859651	12.926298	0.141084	20.340243	2.205827

The sums for species *sp1*, *sp2* and *sp4* are greater than 2 × Rao (2 × 3.579 = 7.158), so if they increase its contribution to the assemblage, the Rao index will increase. For example, if *sp2* change from the original 27 individuals to 60 individuals the new Rao index goes to 5.540. You only need to recalculate relative values and recalculate the index and the partial derivative for each species.

The partial derivative is useful for understanding graphics of changes in diversity when one species is virtually removed from the assemblage. The index may increase or decrease in relation to the average distance of the removed species with the overall average distance.

The expression of quadratic entropy as an absolute value is not useful when the comparisons have to be done between communities with very different numbers of species or when different sets of traits were used to define the distance matrix. To get a relative expression the maximum has to be estimated from the data. The distance matrix does not depend on the abundance of species and is fixed for a given set of species, but changes in the relative abundance of these species may lead to the maximum diversity index (Rao_{max}). There are two types of abundance vectors that define two subclasses of maximum: (a) weak maximization, when some of the w_i abundances that maximize Rao_{max} are zero; and (b) strong maximization, when all the w_i values that maximize Rao_{max} are positives.

The maximization process relies on the dissimilarity matrix and on any ultrametric matrix that belongs to the strong subclass (Pavoine et al. 2005). The drawback arising from having only some species to maximize the Rao's quadratic index when dissimilarity between species are based on functional traits is the absence of distance measures that guaranty the ultrametric condition and then ecological meaningful expression of the functional diversity using relative Rao index. Taxonomic or phylogenic dissimilarity trees may have ultrametric distances and give a maximum value of Rao that relies on total abundance distributed among all the species presents.

In the *ade4* software, also available as an R library, the sentence 'divc' may be used to calculate this index. The algorithm return Rao when absolute value is required (scale = FALSE), and the relative value (scale = TRUE) in the range [0, 1] when scaling is required. With FDiversity software both expressions may be obtained simultaneously.

3.4.5 Extended FD

The FD proposed by Petchey and Gaston (2006) is based on a dissimilarity matrix computed with one mean value by trait and species and no importance measure to weigh the species abundance in the community. It is possible to load each entry of this dissimilarity matrix with a relative measure of abundance (frequency, coverage, biomass, basal area, or other) before performing the dendrogram. If d_{ij} denotes the dissimilarity measurement between species 'i' and 'j', and w_i and w_j denote the relative abundance of each species, the weighted FD (wFD) is computed from a matrix with entries $d'_{ij} = d_{ij}\left(\sqrt{w_i w_j}\right)$. The resulting weighted dissimilarity matrix is symmetric with zeros in the diagonal. The wFD is computed as the total branch length of the functional dendrogram derived from this symmetric matrix (Pla et al. 2008, Casanoves et al. 2008, 2011). The scaled form of wFD multiplies the sum by the total number of species to put it in the same metric as FD.

As in the FD case, the dendrogram may be computed only with the set of species present in each plot (wFDp plot based), or may be derived from one dendrogram including the species community pool (wFDc community based). With equi-abundance wFD equals FD. In case study 1 in Sect. 4.2 we compare FD and wFD to show how changes in these indices can be used to explore the relationship between functional diversity and changes in abundance and trait values.

3.4.6 Functional Richness, Evenness, Divergence and Dispersion

Villéger et al. (2008) argued that functional diversity cannot be summarized by a single number because it has to include components of richness, evenness and

divergence taken into account the trait values and their abundance. They proposed a framework where functional diversity comprises three components: functional richness, functional evenness, and functional divergence. The three independent components provide more detail in examining the mechanisms linking biodiversity to ecosystem functioning. Mason et al. (2005) were one of the first to call the attention about the importance of these three facets of functional diversity to understand its relationship with ecosystem processes and ecosystem services. Villéger et al. defined FRic (functional richness), FEve (functional evenness) and FDiv (functional divergence) using multiple traits.

To complement these three measurements of functional component Villéger et al. (2010) have proposed an index of functional specialization (FSpe) that quantify the relative positions of species respect to the gravity center calculated from the regional pool of species. The index is based on Bellwood et al. (2006) relative distance of a species from the centroid of the principal component space account for at least 85% of the variability observed.

Functional richness (*FRic*): FRic represents the trait space filled by the community. In the one trait case it is represented by the range (maximum–minimum), but with more than one trait it is represented by the volume filled by the community in the trait space. The procedure is like the convex hull hyper-volume (Cornwell et al. 2006). The algorithm identifies the extreme species and then estimates the volume in the trait space. It is recommended to standardize the traits to avoid scale effects. To calculate FRic the number of species must be greater than the number of traits and the species must not relay on a line. The maximum value of FRic in a T dimensional trait space is attained when 2^T species have a combination of extreme trait values.

An option to estimate FRic when the number of species is less than the number of traits is to synthesize the trait space using a multivariate technic to reduce the dimensions. If all the traits are in a quantitative scale principal component analysis may be applied and the resulted component used as new 'trait synthesis'. The number of components retained depends on the proportion of variability explained and are limited to the number of species minus one. With categorical or nominal traits the reduction may be derived using principal coordinate analysis (also known as classical multidimensional scaling) and retained the appropriate coordinate values. In the R-scrip of FD-library written by Laliberté and Legendre (2010) to perform this calculation, this procedure is used by default. In FDiversity, there are two separate indices, one for Convex Hull (equal to FRic when $S > T$, and no values when $S \leq T$ or at least two species are distributed in a line) and other for FRic calculated using linear combination of traits resulted from ordination technics.

Functional evenness (*FEve*): FEve measures the regularity of spacing between species in the trait space as did the univariate FRO and also the evenness in the distribution of the species abundance. The authors do not use the overall FRO proposed by Mouillot et al. (2005, see Sect. 3.3.3) because the method depends on ordination techniques and some information may be lost. Villéger et al. (2008)

used the minimum spanning tree (MST) to transform a multidimensional space to a distribution on a single axis. The MST links points in the T-dimensional space with the minimum sum of branch lengths. As in FRO this new functional evenness index measures both the regularity of branch lengths in the MST and the evenness in species abundances. There are a total of $S-1$ branches in the MST of S species and each of the b branch length is divided by the sum of the abundances of the species linked

$$EW_b = \frac{d_{ij}}{w_i + w_j},$$

where EW_b is the weighted evenness, d_{ij} is the Euclidean distance between species i and j, those involved in the branch b, and w_i and w_j are the relative abundance of these species.

In case of perfect regularity of abundance all weighted evenness would be equal, but otherwise it is useful to compute the partial weighted evenness PEW_b dividing by the sum of the EW_b across the $S-1$ branches

$$PEW_b = \frac{EW_b}{\sum_{b=1}^{S-1} EW_b}.$$

When the PEW_b value differs among branches, the final index will decrease. To quantified the discrepancy they compared PEW_b with $1/(S-1)$, the index is

$$FEve = \frac{\sum_{b=1}^{S-1} \min\left(PWE_b, \frac{1}{S-1}\right) - \frac{1}{S-1}}{1 - \frac{1}{S-1}}$$

with an standardization similar to that suggested by Bulla (1994).

This index does not correlate with species richness and ranges from 0, complete unevenness, to 1, complete evenness and it is independent of the convex hull. At least three species have to be present in the sample to enable the calculation because at least three points are needed to define the MST, no matter the number of traits. The index value decreases when relative abundance of species is less evenly distributed and when distances among species are irregular.

Functional divergence (*FDiv*): FDiv quantify how the trait values are spread along the range of a trait space. For only one trait Mason et al. (2003) have defined FD_{var} (see Sect. 3.3.2) but when there are more than one trait the linear range is replace by a multidimensional range, like the convex hull. So functional divergence is related to how abundance is distributed within the volume of functional trait space. The first step in index calculation is defining the gravity center of the V species that form the vertices of the convex hull $G_v = \{g_1, g_2, \ldots, g_T\}$, being

$$g_t = \frac{1}{V} \sum_{i \in Sv}^{V} x_{ti}$$

where Sv is the subset of all the V species forming the vertices of the convex hull, x_{ti} is the coordinate (trait value) of species 'i' on the 't' trait, T is the total number of traits, and g_t is the coordinate of the gravity center for trait 't'. Knowing the coordinate of gravity center, we compute Euclidean distance of each species from this point as

$$dG_i = \sqrt{\sum_{t=1}^{T} (x_{ti} - g_t)^2}$$

and the mean distance $\overline{dG}$ of the S species to the gravity center is

$$\overline{dG} = \frac{1}{S} \sum_{i=1}^{S} dG_i.$$

These distances are computed only with trait values and do not include the species abundance, so they reflect the shape and the volume of the convex hull. To take the abundance into account it is necessary to compute the abundance-weighted deviances (Δd) and the absolute abundance-weighted deviances ($\Delta|d|$) as

$$\Delta d = \sum_{i=1}^{S} w_i \times \left(dG_i - \overline{dG}\right)$$

and

$$\Delta|d| = \sum_{i=1}^{S} w_i \times \left|dG_i - \overline{dG}\right|$$

being w_i the relative abundance of species 'i'. The functional divergence index is then

$$FDiv = \frac{\Delta d + \overline{dG}}{\Delta|d| + \overline{dG}}.$$

Adding $\overline{dG}$ to the numerator and denominator makes that the index belongs to interval 0 to 1, because dG_i are Euclidean distance and so are positive or zero, thus Δd is bounded between $\overline{dG}$ and $\Delta|d|$.

Functional dispersion (*FDis*): Functional dispersion (FDis) is a multi-dimensional index based on multi-trait dispersion (Laliberté and Legendre 2010). FDis is the average distance of individual species to the centroid of all species in the community trait space taken into account the relative abundances of species by computing the weighted centroid. It is calculated from the 'species × trait' matrix as

$$\mathbf{c} = \{c_1, c_2, \ldots, c_T\}$$

where the vector **c** has the coordinates of the weighted centroid in the T-dimensional space, and c_t for $t = 1, \ldots, T$, is estimated for each dimension (trait) as

$$c_t = \sum_{i=1}^{S} w_i \, x_{ti}$$

being w_i the relative abundance of species 'i', and x_{ti} the value of the 't' attribute of species 'i'. This formulation implies that $x_{(.)}$ represent a quantitative trait, but the authors generalize the distance measure to include semi-quantitative and qualitative traits through principal coordinate analysis (PCoA).

FDis, the weighted average distance $\bar{z}$ from each species to the weighted centroid **c**, is then computed as

$$\mathrm{FDis} = \sum_{i=1}^{S} w_i \, z_i$$

where w_i is the abundance of species 'i' and z_i is the distance of species 'i' to the weighted centroid **c**. This procedure essentially shift the position of the centroid towards the more abundant species and weigh distances of individual species to this weighted centroid by their relative abundances. It has been suggested that communities with only one species should have FDis = 0, but there is no upper limit for this index.

Functional specialization (*FSpe*): Functional specialization is defined using the multidimensional trait space of the regional pool of species and quantifying how apart the species are from the gravity center. To estimate the index for each plot the distance is loaded by the relative contribution of each species.

The first step in index calculation is defining the gravity center of the S species in the T-dimensional space of the traits $G = \{g_1, g_2, \ldots, g_T\}$, being

$$g_t = \frac{1}{S} \sum_{i=1}^{S} x_{ti}$$

where x_{ti} is the coordinate (trait value) of species 'i' on the 't' trait, T is the total number of traits, and g_t is the coordinate of the gravity center for trait 't'. Knowing the coordinate of gravity center, we compute Euclidean distance of each species from this point as

$$dG_i = \sqrt{\sum_{i=1}^{S} (x_{ti} - g_t)^2}.$$

If the traits are standardized the gravity center has coordinates (0, …, 0). The FSpe is computed at plot level (or local level) as the weighted sum of the dG_i of species present at the plot using its relative abundance (w_i) to load

$$FSpe = \sum_{i \in plot}^{S} (dG_i \times w_i).$$

3.5 Ability of Indices to Detect some Ecological Processes

The assembly process is one of the ecological mechanisms that model the way how species coexist in a community. Even though the set of species that form a given community depends on available species, it is also strongly affected by the main assembly process that operates during early stages. The pattern generated combine environmental stress and biological competition that operates on interactions among species and between species and the environmental conditions. The combination may affect morphological, physiological and functional properties that characterize the species in the community. These properties can be grouped to associate with an environmental ecological service or process. Several authors have studied patterns of assembling species assuming that the traits are phylogenetically conserved (Kraft et al. 2007; Mouchet et al. 2010): limiting similarity (MacArthur and Levins 1967) to produce a uniform dispersion of traits values (Stubbs and Wilson 2004), habitat filtering (Zobel 1997) to produce cluster of traits values (Perez-Neto 2004) and neutral assembly (Gotelli and Graves 1996) to produce random scattering.

The functional diversity indices have different abilities to reflect these processes, and their dependence on species richness varies. Ideally, functional diversity indices have to be able to reflect other aspect of community than crude richness in order to be useful to link trait expression to functional performance. Changes in trait community weighted mean may be used to associate community function to a given ecological services; for example, wood density (wd) is associate to carbon sequestration: as wd-CWM increases the expected amount of carbon sequestered in vegetal tissues increases. Shifts in CWM for key traits may be combining with functional diversity indices to trace changes in community succession, for example restoration after fire events (Ricotta and Moretti 2011).

Mouchet et al. (2010) examined the performance of several functional diversity indices using increasing richness (from 10 to 100 species with intervals of 10) and simulation data set using three assembly processes (limiting similarity, environmental filtering and neutral or random). They compared five indices that do not include abundance (FAD2, MFAD, FD, GFD and FRic) and three indices including abundance (Rao, FDiv and FEve). Spearman correlation coefficient was not significant between FEve and FAD2, among the others the correlations were

highly significant. Using this set of indices the authors identified four groups of indices related to the three orthogonal functional components.

The functional components (Villéger et al. 2008) have also been referred as facets of functional diversity in the single-trait approach (Mason et al. 2005) and have to be interpreted as properties emerging of the set of species in a given community. The facets or components of functional diversity are not associated to species, it is not the sum of species attributes. They are community characteristics that depend on the interaction among species, and between species pool and the environment. The main facets or components of functional diversity are: functional richness, functional evenness and functional divergence.

The four groups were identified using the algorithm K-means based on principal component axes calculated with indices values in the communities. Two of these groups were identified with functional richness and include FAD2, MFAD, FD, GFD and FRic. The functional divergence is associated with FDiv and also with Rao, and functional evenness is associated with FEve.

The ability of indices to differentiate assembly processes showed that FRic, FEve, FDiv and Rao are much more sensitive than FAD2, MFAD, FD and GFD. Indices with values higher than expected by chance are associated to limiting similarity and with values lower than expected by chance are associated to environmental filtering. Whichever the index selected, relationship of functional diversity to community assemblage processes has to be investigated comparing the observed value with that expected by random. When no evidence is found to reject the random process, both environmental filtering and the competition may be operated sequentially or simultaneously. The best subset of indices includes FRic, FEve and FDiv because each is able to reflect one component of functional diversity. We have to mention that in this work FDis, the one proposed by Laliberté and Legendre (2010) to complement the other three was not included; neither the FSpe proposed by Villéger et al. (2010).

References

Barber CB, Dobkin DP, Huhdanpaa H (1996) The Quickhull algorithm for convex hulls. ACM T Math Softw 22:469–483

Bellwood DR, Wainwright PC, Fulton CJ, Hoey AS (2006) Functional versatility supports coral reef biodiversity. Proc R Soc B 273:101–107

Botta-Dukát Z (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. J Veg Sci 16:533–540

Bulla L (1994) An index of evenness and its associated diversity measure. Oikos 70:167–171

Casanoves F, Di Rienzo JA, Pla L (2008) User manual FDiversity: statistical software for the analysis of functional diversity, 1st edn. Argentina. http://www.fdiversity.nucleodiversus.org

Casanoves F, Pla L, Di Rienzo JA, Díaz S (2011) FDiversity: a software package for the integrated analysis of functional diversity. Methods Ecol Evol 2:233–237

Cianciaruso MV, Batalha MA, Gaston KJ, Petchey OL (2009) Including intraspecific variability in functional diversity. Ecology 90:81–89

Cornelissen JHC, Lavorel S, Garnier E, Diaz S, Buchmann N, Gurvich DE, Reich PB, ter Steege H, Morgan HD, van der Heijden MGA, Pausas JG, Poorter H (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Aust J Bot 51:335–380

Cornwell WK, Schwilk DW, Ackerly DD (2006) A trait-based test for habitat filtering: convex hull volume. Ecology 87:1465–1471

Di Rienzo JA, Casanoves F, Pla L (2008) FDiversity, software to estimate functional diversity. RDNDA, Argentina, 702841

Díaz S, Lavorel S, de Bello F, Quétier F, Grigulis K, Robson M (2007) Incorporating plant functional diversity effects in ecosystem service assessments. Proc Natl Acad Sci USA 104:20684–20689

Faith DP (1996) Conservation priorities and phylogenetic pattern. Conserv Biol 10:1286–1289

Garnier E, Cortez J, Billès G, Navas M-L, Roumet C, Debussche M, Laurent G, Blanchard A, Aubry D, Bellmann A, Neill C, Toussaint J-P (2004) Plant functional markers capture ecosystem properties during secondary succession. Ecology 85:2630–2637

Gotelli NJ, Graves GR (1996) Null models in ecology. Smithsonian Institution, Washington DC

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857–874

Grime JP (1998) Benefits of plant diversity to ecosystems: immediate filter and founder effects. J Ecol 86:902–910

Jaureguiberry P, Bertone G, Díaz S (2011) Device for the standard measurement of shoot flammability in the field. Austral Ecol. doi:10.1111/j.1442-9993.2010.02222.x

Kraft NJB, Cornwell WK, Webb CO, Ackerly DD (2007) Trait evolution, community assembly, and phylogenetic structure of ecological communities. Am Nat 170:271–283

Laliberté E, Legendre P (2010) A distance-based framework for measuring functional diversity from multiple traits. Ecology 91:299–305

Lavorel S, Grigulis K, McIntyre S, Williams NSG, Garden D, Dorrough J, Berman S, Quétier F, Thébault A, Bonis A (2008) Assessing functional diversity in the field—methodology matters! Func Ecol 16:134–147

MacArthur RH, Levins R (1967) The limiting similarity, convergence, and divergence of coexisting species. Ame Nat 101:377–385

Magurran AE (2004) Measuring biological diversity. Blackwell, UK

Marczewski E, Steinhaus H (1958) On a certain distance of sets and the corresponding distance of functions. Colloq Math 6:319–327

Mason NWH, MacGillivray K, Steel JB, Wilson JB (2003) An index of functional diversity. J Veg Sci 14:571–578

Mason NWH, Mouillot D, Lee WG, Wilson JB (2005) Functional richness, functional evenness and functional divergence: the primary components of functional diversity. Oikos 111:112–118

Mouchet M, Guilhaumon F, Villéger S, Mason NW, Tomasini JA, Mouillot D (2008) Towards a consensus for calculating dendrogram-based functional diversity indices. Oikos 117:794–800

Mouchet MA, Villéger S, Mason NW, Mouillot D (2010) Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. Func Ecol 24:867–876

Mouillot D, Mason NW, Dumay O, Wilson JB (2005) Functional regularity: a neglected aspect of functional diversity. Oecologia 142:353–359

Pavoine S, Bonsall MB (2011) Measuring biodiversity to explain community assembly: a unified approach. Biol Rev. doi:10.1111/j.1469-185X.2010.00171.x

Pavoine S, Ollier S, Pontier D (2005) Measuring diversity from dissimilarities with Rao's quadratic entropy: are any dissimilarities suitable? Theor Popul Biol 67:231–239

Perez-Neto PR (2004) Patterns in the co-occurrence of fish species in streams: the role of site suitability morphology and phylogeny versus species interactions. Oikos 93:110–120

Petchey OL, Gaston KJ (2002) Functional diversity (FD), species richness and community composition. Ecol Lett 5:402–411

Petchey OL, Gaston KJ (2006) Functional diversity: back to basics and looking forward. Ecol Lett 9:741–758
Petchey OL, Gaston KJ (2007) Dendrograms and measuring functional diversity. Oikos 116:1422–1426
Petchey OL, Hector A, Gaston KJ (2004) How do different measures of functional diversity perform? Ecology 85:847–857
Pielou EC (1975) Ecological diversity. Wiley, New York
Pla L, Casanoves F, Di Rienzo JA, Fernandez F, Finegan B (2008) Confidence intervals for functional diversity indices considering species abundance (http://ibc2008abstracts.tibs.org/Multiple%20Testing/IBC2008_Paper_ref_571.pdf). Paper presented at the XXIV International Biometric Conference, University College, Dublin, 13–18 July 2008
Podani J, Schmera D (2006) On dendrogram-based measures of functional diversity. Oikos 115:179–185
Podani J, Schmera D (2007) How should a dendrogram-based measures of functional diversity function? A rejoinder to Petchey and Gaston. Oikos 116:1427–1430
Poos MS, Walker SC, Jackson DA (2009) Functional-diversity indices can be driven by methodological choices and species richness. Ecology 90:341–347
Rao CR (1982) Diversity and dissimilarity coefficients: a unified approach. Theor Popul Biol 21:24–43
Ricotta C, Moretti M (2011) CWM and Rao's quadratic diversity: a unified framework for functional ecology. Oecologia 167:181–188
Ricotta C, Szeidl L (2006) Towards a unifying approach to diversity measures: bridging the gap between the Shannon entropy and Rao's quadratic index. Theor Popul Biol 70:237–243
Schmera D, Erős T, Podani J (2009) A measure for assessing functional diversity in ecological communities. Aquat Ecol 43:157–167
Shannon C, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana
Shimatani K (1999) The appearance of a different DNA sequence may decrease nucleotide diversity. J Mol Evol 49:810–813
Shimatani K (2001) On the measurement of species diversity incorporating species differences. Oikos 93:135–147
Simpson EH (1949) Measurements of diversity. Nature 163:688
Stubbs WJ, Wilson JB (2004) Evidence for limiting similarity in a sand dune community. J Ecol 92:557–567
Villéger S, Mason NW, Mouillot D (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology 89:2290–2301
Villéger S, Ramos Miranda J, Flores Hernández D, Mouillot D (2010) Contrasting changes in taxonomic vs functional diversity of tropical fish communities after habitat degradation. Ecol Appl 20:1512–1522
Walker BH, Langridge JL (2002) Measuring functional diversity in plant communities with mixed life forms: a problem of hard and soft attributes. Ecosystems 5:529–538
Walker BH, Kinzig A, Langridge JL (1999) Plant attribute diversity, resilience, and ecosystem function: the nature and significance of dominant and minor species. Ecosystems 2:95–113
Zobel M (1997) The relative role of species pools in determining plant species richness. An alternative explanation of species coexistence? Trend Ecol Evol 12:266–269

Chapter 4
How to Estimate Functional Diversity Indices

Abstract There are several programs to estimate some subsets of functional diversity indices, and only one integrated tool with a friendly interface capable to estimate all indices here included: FDiversity. In the first part of this chapter we present the user interface to handle and summarize data using one simple example (Sect. 4.1, The FDiversity software: capabilities and data management). The other three sections show how to estimate functional diversity indices using real data sets with small random modification to allow data files availability for unpublished databases. In Sect. 4.2 we study relationships of functional indices and plant functional types with altitude; in Sect. 4.3 we study functional indices and test changes in a chronological sequence after stopping human intervention; in the last section (Sect. 4.4) we show how to use graphical tools to display results and transmit findings.

Keywords Trait by species matrix · Species by plot matrix · Standardized traits · Functional traits in a chronosequence · Functional traits in an altitude gradient

4.1 The FDiversity Software: Capabilities and Data Management

4.1.1 How to Install FDiversity

FDiversity implements a user-friendly interface to open source routines for the estimation and analysis of functional diversity indices. The open source platform is R with an interface written in Delphi® using DCOM-R (a way to run R in the background, due to Thomas Baier and Erich Neuwirth). DCOM-R is accessed via Delphi routines written by Dieter Menne.

L. Pla et al., *Quantifying Functional Biodiversity*, SpringerBriefs in Environmental Science, DOI: 10.1007/978-94-007-2648-2_4, © The Author(s) 2012

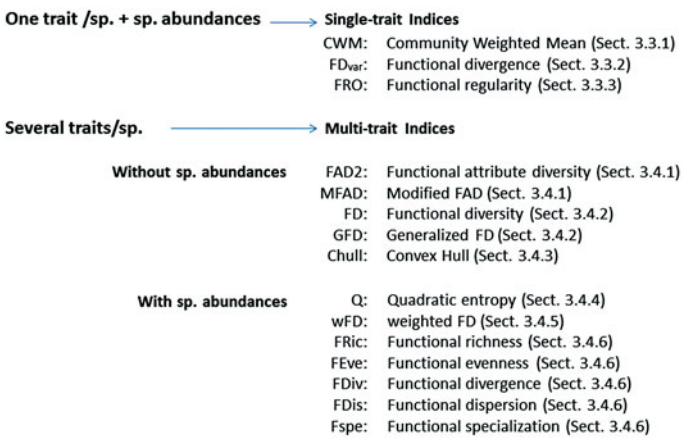


Fig. 4.1 User interface of FDiversity showing options of file menu and help menu

To install FDiversity (Di Rienzo et al. 2008; Casanoves et al. 2011) you have to access http://www.fdiversity.nucleodiversus.org and download the installer. From this link the user may also be re-direct to CRAN-R site to download R or access CRAN repository from (http://cran.r-project.org/) if it is not already installed in your computer, and download DCOM version DCOM 3.1-2B7.

In order for FDiversity to have access to R, DCOM and R must be previously installed in your system. Follow the following steps:

(a) Install DCOM 3.1-2B7 (for R versions R 2.12.x or later)
(b) Install R 13.0 or later
(c) Run R to verify its installation and quit (File-Exit)
(d) Install FDiversity using fdiversityinstaller.exe
(e) Run FDiversity and select Trying to connect to R from Help menu.

The program installs a library needed to make the link between FDiversity and R, and other libraries needed to calculate functional diversity indices. After doing these, a new icon in the upper toolbar appears ([R]) indicating that the connection with R has been established. In case this procedure fails, repeat the procedure from d) if R was successfully installed (step c), otherwise re-install R. In the User Manual there are some additional instructions that may help this procedure. The manual may be access from the web page or from the Help menu of FDiversity.

When FDiversity is open a blank window with a toolbar appears (Fig. 4.1). Eight menus offer general tools to data management and specific options to calculate functional diversity indices. The File menu has the common options to open an empty data table (New), to load a saved data table (Open table), to save a data table (Save table or Save table as …), to close or print the table (Close table, Print) and an option to open tables from a Data folder installed with the program. This folder has several files which may be used to test the program performance (see User Manual for instructions).

The User Manual was downloading and saved during installation process (Casanoves et al. 2008). From the Help menu you may access the Manual, a summary of what FDiversity may do (About FDiversity) and how to reference it (Citing FDiversity). If you have access to internet you may Check for new version, if you have the last version, a message will tell you that it is updated, otherwise it offers to update and redirects to the web page. The News offers information of the last changes that may be an update due to new indices added, to any bugs corrected, to a new database, or to improve User Manual. If you have any difficult to see the file or you only see texts and symbols, cut and past the link and use other web browser. You may subscribe to receive alerts in your email.

The last two entries are options to install R, the first one is a link to a webpage with instruction to do it manually (Installing R), and the last one is used to link R and FDiversity automatically, after the DCOM and R have been installed. If you have the [R] menu in the toolbar do not used these options.

Using the Edit menu it is possible to Cut, Copy, Paste and Undo changes in the data table or in the result windows. The Windows menu has options: Cascade, Align vertical, and Align horizontal, and additionally a list of the open windows during current session as in many software.

4.1.2 Data Menu

To fully explore the options in this menu, a data table has to be opened. To do it, go to File-Open test data folder and select DataSet01. The data file has 252 rows and 14 columns, this information is shown at the left-bottom corner of the table (Records 252 * 14). With the data file opened, click Data menu and several options are displayed (Fig. 4.2).

These procedures are intended to manage usual actions on a data sheets, such as: inserting, adding and deleting rows, activating and deactivating cases to allow or disallow participation in future calculations, and invert the selection (Actions on rows); inserting, adding and deleting columns, editing labels, view or change the data type, control alignment and number of decimals, categorize o re-categorize a variable, and generate a classification variable according cells color (Actions on columns), arranging rows according to different sorting criteria (Sort). Several transformations are implemented to make trait scales comparable, and there is a formula option to define new variables from the combination of those in the data base or from mathematical functions and constants (Formula).

It also contains links to more specialized procedures that allow to merge tables side by side according to matching criteria (Merge horizontally) and on how to merge tables appending one to the other (Merge vertically). This is very useful in functional diversity analysis because it is a common practice to have one table or spreadsheet with the list of species and its mean trait values and another table with the information on sample or plot richness with some variable indicating the

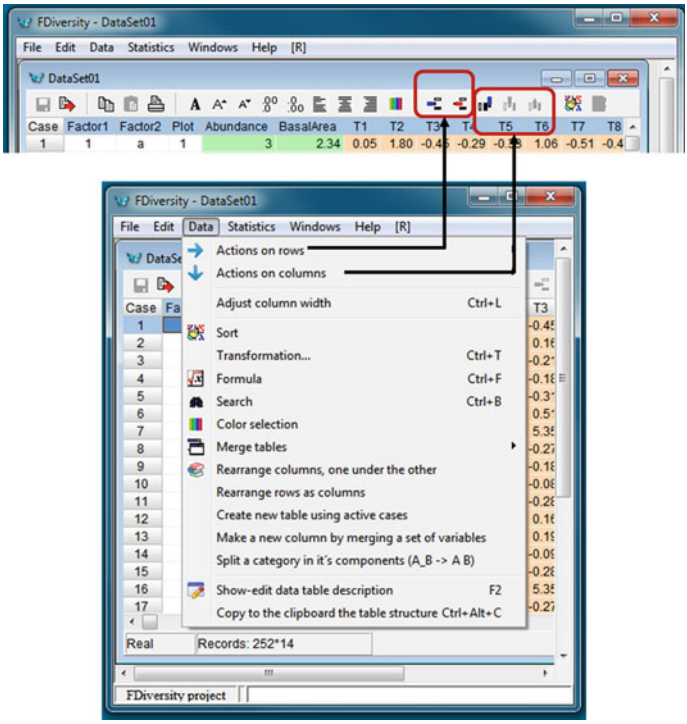


Fig. 4.2 User interface of FDiversity showing tool bar of data table and options for data menu

absolute or relative participation of each species in the assemblage (Merge tables). We shall see an example later in this Chapter (Sect. 4.2).

Most statistical programs used a 'clean' data table, which is one with no comments nor special warnings. The drawback of 'clean' data tables is that important information about the set of data, where, when and how they were collected or by whom are absent or has to be saved separately. FDiversity has a special note pad to enter text useful to describe the data set. This note pad is opened with F2. To add information about a variable, double clicking the label of any variable, then its name may be edit, and double clicking again to open a note pad to enter text; most useful information is related to measurement units or meaning of levels for ordinal variables or nominal variables.

4.1.3 Statistics and Output Menus

The most important procedures of FDiversity are grouped in Statistics menu. The Summary Statistics and the Functional diversity estimation and analysis

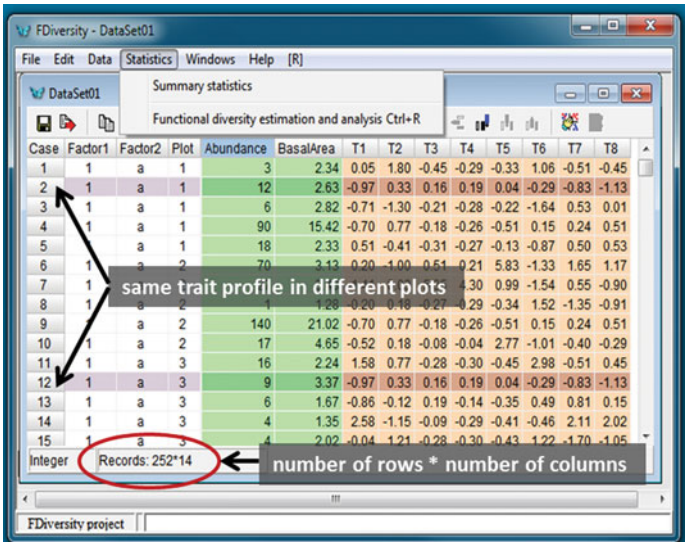


Fig. 4.3 User interface of FDiversity showing a data table

sub-menus, allow exploiting the main tools. Open the data file DataSet01. FDDB from test data folder (Fig. 4.3).

The file has information on species importance in each plot (*Abundance* in column 4 or *Basalarea* in column 5) and on eight traits (*T1* to *T8* standardized values with cero mean and unit variance). Each row represents a species (no names are shown to emphasize that it is not necessary to identify the species) with the trait value it bears in a given plot. Here, the trait values for the species are the same whichever the plot considered. For example, row 2 and row 12 has the same trait profile (−0.97, 0.33, …, −1.13) because they represent the same species, with basal area of 2.63 in plot 1 and 3.37 in plot 3 (Fig. 4.3). In this case both plots belong to level '1' of *Factor1* and level 'a' of *Factor2*.

Coexisting species were identified with the same combination of levels of *Factor1* with *Factor2*, and plot number. In this experiment there were three repetitions for the combination of four levels of *Factor1* with two levels of *Factor2*. So, the study has eight conditions (4 × 2) with three repetitions each. There are 24 plots with different richness. The functional diversity indices have to be calculated using species growing together in a plot; in this case we will get 24 values for each functional diversity index.

4.1.3.1 Summary Statistics Sub-Menu

When Summary statistics is selected, a new window appears. We call this window Selector of Variable. In the right sub-window appears the list of all

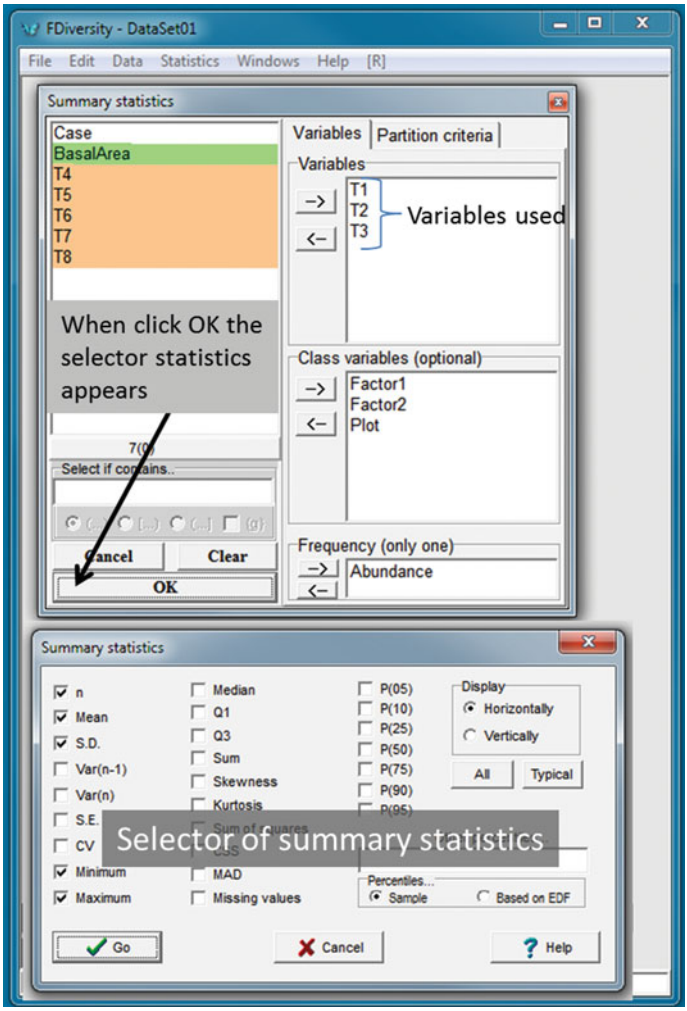


Fig. 4.4 User interface of FDiversity showing options of summary statistics

variables in the data table plus one variable named *Case*, which identifies the row (Fig. 4.4). To the right there are three sub-windows: Variables, Class variables (optional) and Frequency (only one). Highlight the variables to analyze with the mouse, and then move to the sub-window Variable using the arrow. If you want to remove one, double click in its name, and the variable will return to the original list; or highlight the desired set of variables and use the back arrow.

The classification variables are those used to identify the cases to analyze together. Move there the three variables used for *Factor1*, *Factor2*, and *Plot*. If you have no classification variable and want a summary for the whole database, leave this clear. It is possible to incorporate the weight of each species (case)

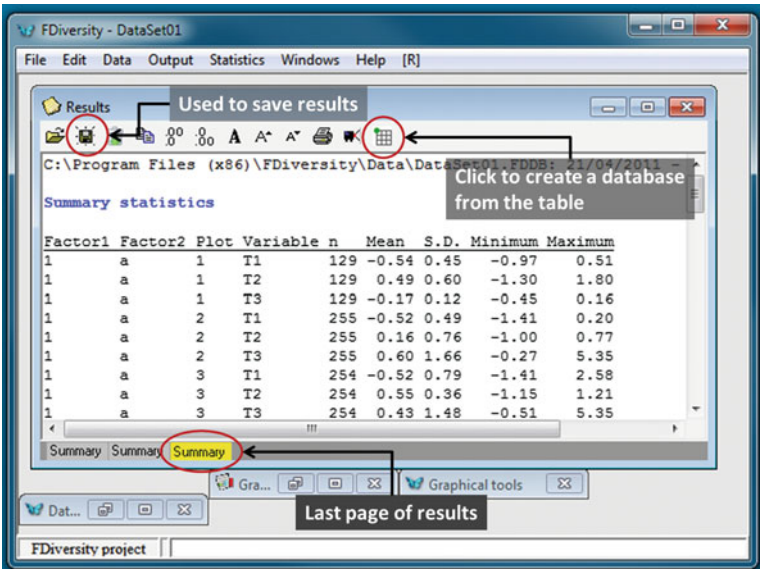


Fig. 4.5 User interface of FDiversity showing the results window

according to its contribution to the estimated value; in this file we may use *Abundance* or *Frequency*, or both but sequentially because only one is accepted at a time (Fig. 4.4). Classical statistics, confidence intervals for fixed probability and for user defined using parametric or nonparametric bootstrap estimation may be shown in two layouts at the result window. Field free to experiment and select the best subset to fulfill what you need.

Results are shown in the Result window. Each time you perform an analysis using FDiversity the results are shown in one page of the results window. This window may be copied (click right mouse and select copy as usual), or saved as a file with a desired name. If you have several analyses, one result page for each is accumulated and you may see one of them clicking at the bottom tab. If you have several tabs, and saved the result window, all of them will be saved together. The name of each page may be change right clicking in the tab name (Fig. 4.5). The table with the values for each class may be transform to a database for further analyses; use the last button on the toolbar in the results window (Generate data table from frames in results). Notice that when you slide the mouse over the buttons in any toolbar, a legend with its function appears.

4.1.3.2 Functional Diversity Estimation and Analysis Sub-Menu

This is the main menu in the program. From the different windows it is possible to select functional diversity indices to estimate, know the original reference to each

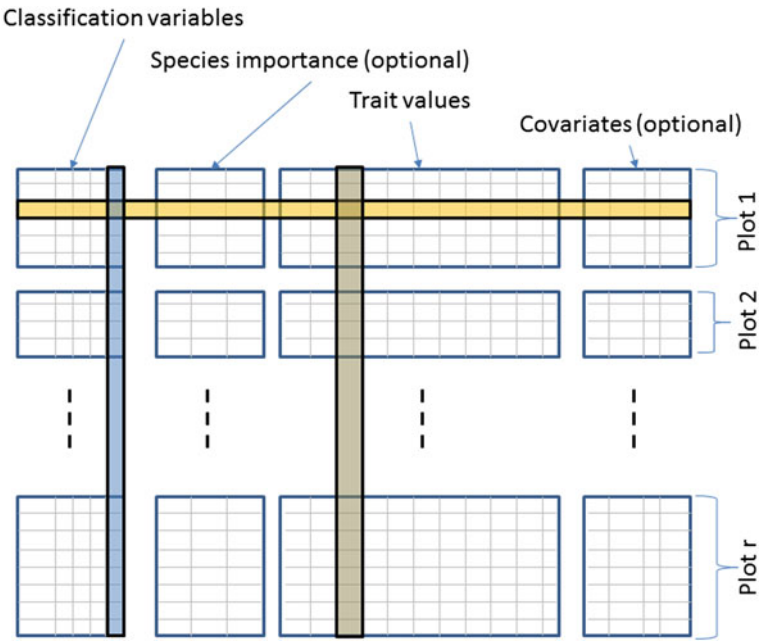


Fig. 4.6 Structure of a data table to calculate functional diversity indices using FDiversity

one, do analysis of variance and covariance, and choose appropriate distance measure (or similarity measure) to calculate distance matrices and decide which algorithm use to build dendrograms.

To calculate an index is necessary to have a data table with one row for each species in each plot (Fig. 4.6) and at least two types of variables. One to identify the plot, and classification variables associate to the plot that will be used for the post analysis. Here the scientific and common names of the species may be included even though they are not used in index calculations. The second set are the trait values that may be in its original units or with values standardized to avoid scale effect assigning more importance to traits with higher variance.

Additionally, information on absolute or relative contribution of each species in each plot allows the calculation of indices weighted by the abundance (like Rao, wFD, FEve, FDis and FDiv), and the information on covariables may be used to analyze dependence on plots characteristics that is not directly associated with the levels of classification variable used to test hypothesis.

Using DataSet01.FDDB we are going to calculate functional diversity indices and species indices. With FDiversity we can calculate 14 multi-trait indices, three single-trait indices and four species indices (richness, Shannon–Weaver, Evenness and Simpson). Select Statistics-Functional diversity estimation and analysis, when the selector of variable is displayed (Fig. 4.7), at left side is the list of all variables in the data table. To the right, the upper sub-window is the Traits window,

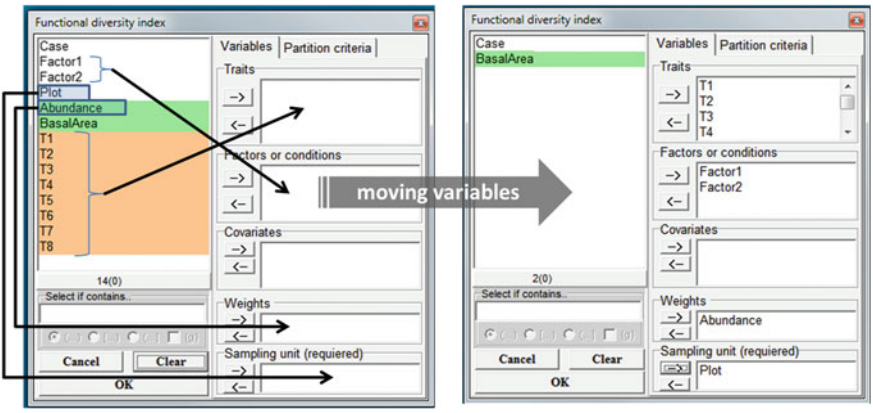


Fig. 4.7 Selector variable window use to calculate functional diversity indices

move the variables *T1* to *T8* using the arrow; move one or more variables that identify the levels of each condition to the sub-window Factors or conditions. For example, variable may be 'land use type' with four levels (unused, forestry, agroforestry, and forestry-livestock). The levels may be labeled with numbers, as in this data base, and also with names; move *Factor1* and *Factor2* there. The contribution of each species (each row) to the estimation of one value of the index goes in the Weights sub-window; in this case we use *Abundance*. Remember that only one weighted variable is accepted at a time. The bottom sub-window (Sampling unit) is mandatory, is reserves for variable which labeled the species to be considered together to estimate an index value, is the sample or plot identity. In our case, *Plot* is moved there (Fig. 4.7). There is not covariable in this case.

When we accept, the Functional diversity index menu is activated. In this dialogue, various tabs are shown: General, Model, and Comparisons (Fig. 4.8). The general tab allows selecting the diversity index desired. When you select one, reference to the original paper is showed and is copy to the clipboard in case you want to paste it in a word file. As several indices are based on distance matrices, to the right there are more than 25 distance measures (or similarity measures, transformed to dissimilarity by the program before applying) and for linkage procedures to form the dendrogram using by several indices. Only one distance measure to build the distance matrix can be selected at a time. If you want to try different distances, the analysis has to be run again with the second selection. In this tab there is a check box to generate a table with all the results. Use it! It is very helpful having a data table with all indices to use in further analysis and graphics.

At the right side of the Model tab, there is a list of the classification variables and the covariables; at the left side, in the window Fixed effects model terms there is a list of main effects (classification variables and covariables). You may include interactions in the model highlighting the desired variables and clicking

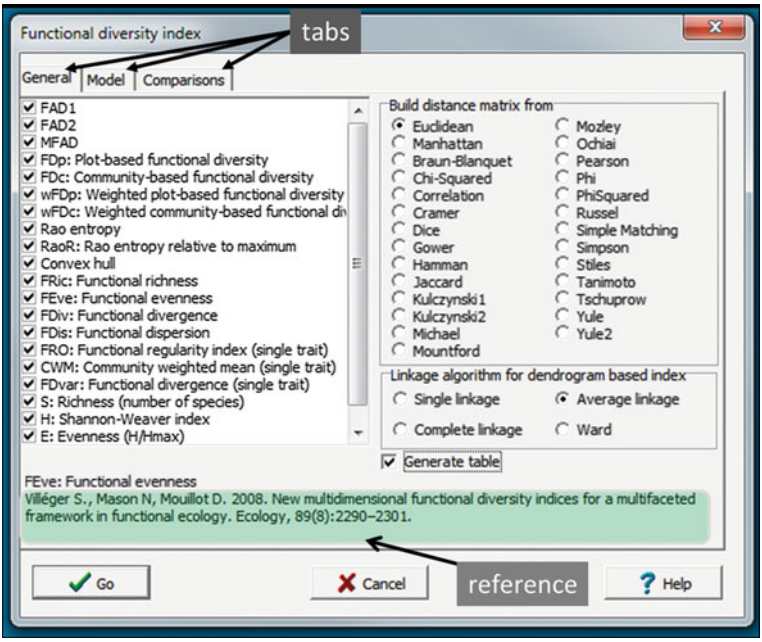


Fig. 4.8 Functional diversity index window selector

the bar Generate interactions terms or using the '*' (Fig. 4.9a). You may write a hierarchical model of variance using the '>' symbol, in case the levels of one factor do not represent the same condition in every levels of the other. In this example you may indicate that *Factor2* is nested in levels of *Factor1*, writing the fixed effects as: *Factor1* > *Factor2*. For nested models there is no interaction to evaluate because *Factor2* is nested in each level of *Factor1*. In the result window appears the ANOVA results followed by the means. If there are one or more covariables in the model, the mean values are adjusted.

In the Comparisons tab, you may choose mean comparison in the model main effects (Fig. 4.9b). There are two methods of mean comparisons: pair-wise comparisons with no correction for multiple pairs, like Fisher's LSD (based on t-distribution), and DGC (Di Rienzo et al. 2002) that implements cluster-based method for identifying groups of nonhomogeneous means. Significance level may be adjusted at any level in LSD (to adjust for multiple comparisons using Bonferroni approach), and to 0.05 or 0.01 with DGC. Select Dendrogram check box and one graph for each mean comparison is shown in popup windows from R, which may be copied.

Press Go and the software will calculate the indices, the analysis of variance for each index, the mean comparisons and will generate the data table if you have asked for it. The table may be saved with extensions *.FDDB to re-use with FDiversity or

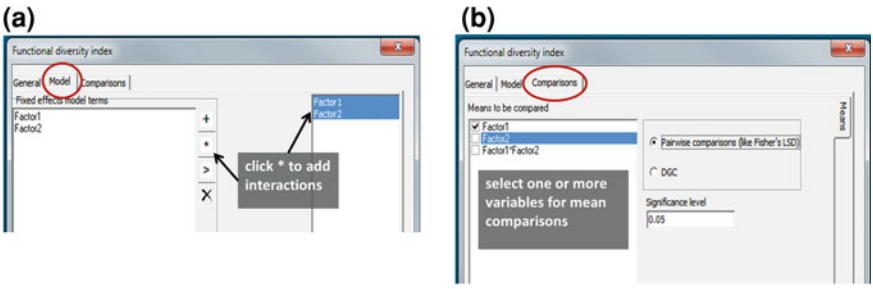


Fig. 4.9 Model specification to analyze indices. **a** Model tab, to select fixed effects and interactions; **b** Comparisons tab, to select the mean comparison method, the significant level and the desired effects

open it from InfoStat, or may be saved with several formats like *.xls or *.txt extensions, allowing compatibility with other statistical or graphical software.

All the indices values are shown in the Results window. The results begin with a list of the traits you used followed by information on the weighting variable, the linkage algorithm for dendrogram based index, and the distance measure. The first table is a list of the multi-trait indices for each plot with the levels of classification variable and the mean values of the co-variables. The next table is a list of the single-trait indices, trait by trait if you asked for FRO, CWM or FDvar, following by a list of the taxonomic diversity indices. All these tables are merging together in a new table called Diversity indices if you selected Generate table at the General tab. Multi-trait functional diversity indices appear in the first's columns, then the single-trait indices and taxonomic indices at the last columns. After the index tables, there are the ANOVA results for each index followed by the corresponding mean comparison.

4.2 Case Study: Changes in Functional Diversity in an Altitudinal Gradient

4.2.1 Sample Design and Trait Evaluation

This example presents a subset of field data taken at a tropical region with secondary forest formation. The sample units were located in an altitudinal gradient from 653 to 2,810 masl. Woody vegetation, including all trees, palms and ferns were identified and diameter at breast height (dbh) was evaluated. For a total of 38 sample units with an area of 0.25 ha (50 m × 50 m), the basal area of each individual was calculated for trees, palms and ferns with more than 10 cm of dbh (Bermeo 2010). Dominant species per plot which together accounted for 75% of basal area (Grime 1998) were identified to produce the species list to determine trait values.

Six traits were selected to relate with impacts of climate change. They were: specific leaf area (SLA mm^2 mg^{-1}), the leaf dry matter content (LDMC, mg g^{-1}), the concentration of nitrogen (LNC, mg g^{-1}), the concentration of phosphorus (LPC, mg g^{-1}), and the physical strength of the leaves (tough, N mm^{-1}) and wood density (WD, g cm^{-3}).

Measurements of functional traits in dominant species were made in the 106 species identified as dominant in the 38 sites evaluated. The sampling methods and categories of allocation of functional traits were based on the protocols proposed by Garnier et al. (2001) and Cornelissen et al. (2003).

4.2.2 The Database

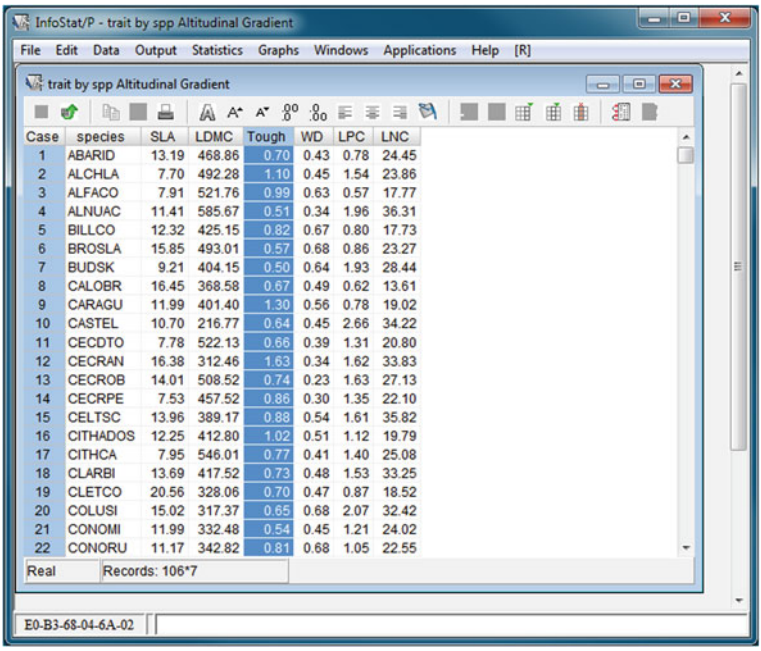
The data were synthesized in two files, one with the information about traits (Fig. 4.10) and the other with information about distribution of species in the observational units. File trait by spp Altitudinal Gradient.IDB2 (available for download via Springer's Extra Materials website: http://extras.springer.com/) has one row for each species which is identified by a code (*species* column) and the mean values for the six traits considered (*SLA*, *LDMC*, *Tough*, *WD*, *LPC* and *LNC*). The file spp by plot Altitudinal Gradient.IDB2 (available for download via Springer's Extra Materials website: http://extras.springer.com/) has information about the presence of each species at each observational unit. It also has information about the altitude of each plot. The importance of each species was quantified with the frequency (*freq*) and the total basal area (*ba*). The data base has 458 rows, one for each species present in each plot, by five columns.

Functional variation along the gradient included two main approaches, one to identify and characterize plant functional types (see Chap. 2 for complete details of statistical procedures) and the other to calculate functional indices and test its relationship with altitude.

4.2.3 Changes of Plant Functional Types with the Altitude

To define functional types we performed cluster analysis using trait information in the dominant species (trait by spp Altitudinal Gradient.IDB2). We used InfoStat statistical software (Di Rienzo et al. 2010) which has an interface comparable to FDiversity and databases that are interchangeable.

InfoStat may be downloaded from (www.infostat.com.ar), the free option has complete analytical functionality with the only restriction that the result and the graphic windows cannot be saved or copied. Open InfoStat and from the file menu, open the data file, or after install InfoStat double click the file name. It should look like in Fig. 4.10.



Case	species	SLA	LDMC	Tough	WD	LPC	LNC
1	ABARID	13.19	468.86	0.70	0.43	0.78	24.45
2	ALCHLA	7.70	492.28	1.10	0.45	1.54	23.86
3	ALFACO	7.91	521.76	0.99	0.63	0.57	17.77
4	ALNUAC	11.41	585.67	0.51	0.34	1.96	36.31
5	BILLCO	12.32	425.15	0.82	0.67	0.80	17.73
6	BROSLA	15.85	493.01	0.57	0.68	0.86	23.27
7	BUDSK	9.21	404.15	0.50	0.64	1.93	28.44
8	CALOBR	16.45	368.58	0.67	0.49	0.62	13.61
9	CARAGU	11.99	401.40	1.30	0.56	0.78	19.02
10	CASTEL	10.70	216.77	0.64	0.45	2.66	34.22
11	CECDTO	7.78	522.13	0.66	0.39	1.31	20.80
12	CECRAN	16.38	312.46	1.63	0.34	1.62	33.83
13	CECROB	14.01	508.52	0.74	0.23	1.63	27.13
14	CECRPE	7.53	457.52	0.86	0.30	1.35	22.10
15	CELTSC	13.96	389.17	0.88	0.54	1.61	35.82
16	CITHADOS	12.25	412.80	1.02	0.51	1.12	19.79
17	CITHCA	7.95	546.01	0.77	0.41	1.40	25.08
18	CLARBI	13.69	417.52	0.73	0.48	1.53	33.25
19	CLETCO	20.56	328.06	0.70	0.47	0.87	18.52
20	COLUSI	15.02	317.37	0.65	0.68	2.07	32.42
21	CONOMI	11.99	332.48	0.54	0.45	1.21	24.02
22	CONORU	11.17	342.82	0.81	0.68	1.05	22.55

Fig. 4.10 Trait data base for altitudinal gradient example. First column indicates the codes for the 106 species. The study includes measurement of six traits: specific leaf area (SLA), leaf dry matter contents (LDMC), leaf toughness (Tough), woody density (WD), leaf phosphorus and nitrogen contents (LPC and LNC) respectively

As all the traits are quantitative, Euclidean distance is appropriate. We explore two linkage algorithms: average linkage and Ward; and four to seven functional groups. With each combination (two linkage algorithms and four functional plant type partitions) we inspect the dendrogram looking for the functional explanations of successive partitions, and also perform multivariate mean comparisons to ensure that mean vector profiles are statistically different among plant functional types.

4.2.3.1 Formation of Plant Functional Types (PFT) for Altitudinal Gradient Example

In InfoStat, select Statistics and then Multivariate Analysis, Cluster analysis (Fig. 4.11). In the Cluster analysis window, select all the traits and move them to the Variable sub-window using the arrow. Use *species* as a classificatory variable (Fig. 4.12). Press Ok and in the next window select Average linkage (it is checked by default) as the linkage method, and Euclidean as the distance, check Standardize data (it is checked by default) and select 4 in the number of groups (Fig. 4.13). This option has two effects: one is to identify the species of each group

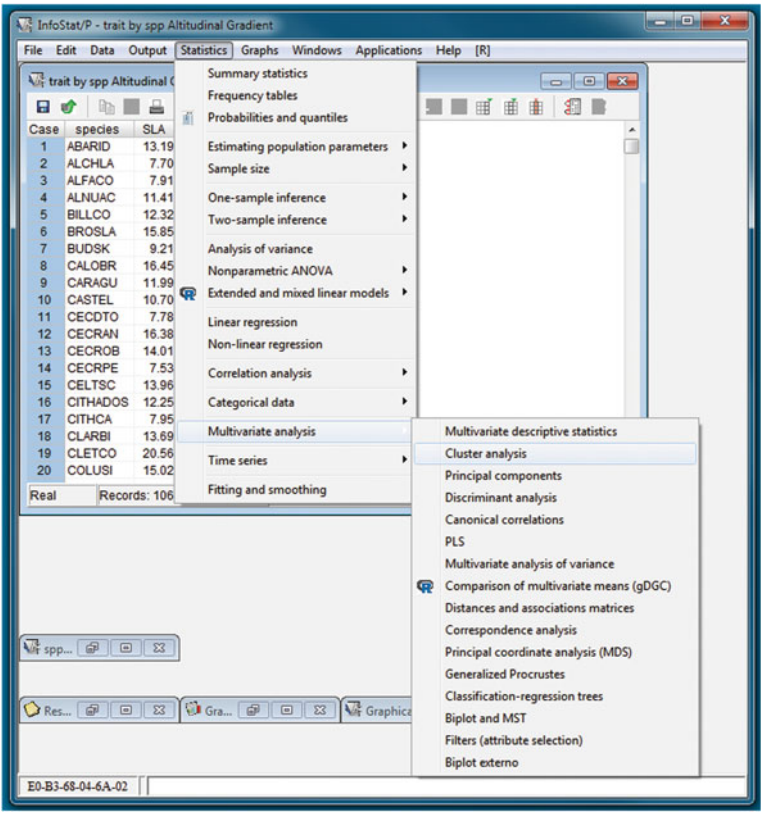


Fig. 4.11 Selection of multivariate procedure to perform cluster analysis using InfoStat

with different colors in the dendrogram, and the other is to add a new column at the end of the data file called *Cluster* indicating the group that each species belongs. Uncheck Overwrite classification if you want to save successive partitions criteria and new columns named *Cluster1*, will be added to data file.

Results of multivariate mean comparison for all partition criteria were synthesizing en Table 4.1. Size of species clusters using average linkage are very heterogeneous (from 1 to 32 with seven groups) and become even more heterogeneous when four groups are considered, going from a small group with two species to a big group with more than half the species present along the gradient (70 species). This kind of dendrogram, where hierarchy shows a unique nested structure is frequently obtained with average linkage (Fig. 4.14). On the other hand, using Ward we obtained clusters more compact with more homogeneous richness by cluster. Dendrogram has clear branches splitting the first two groups of species at high distance (Fig. 4.15), and then alternative split each main branch, which allows defining more compact plant functional types with species that has overall less dispersion. It is important to remember (see Chap. 2) that Ward

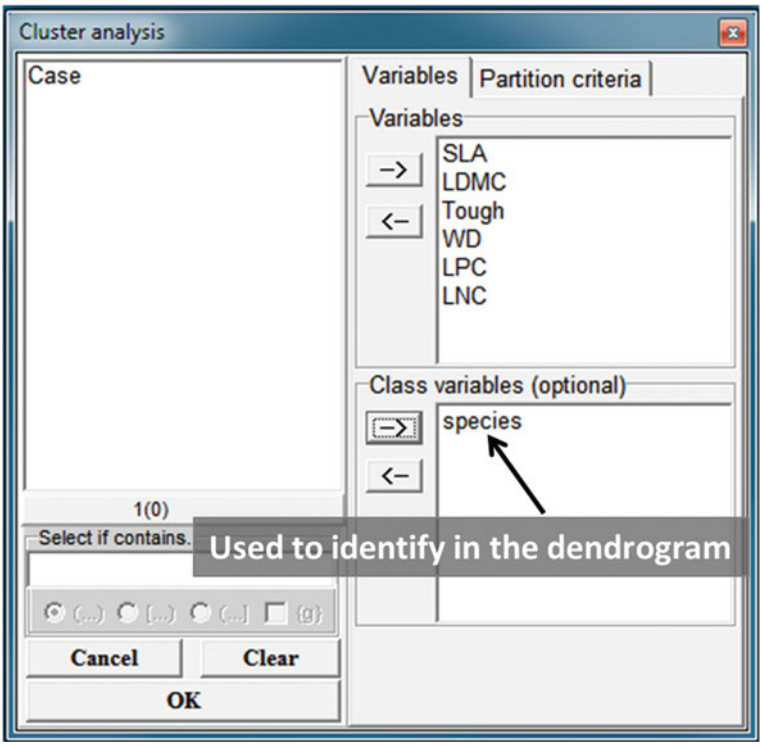


Fig. 4.12 Variable selector for cluster analysis to define plant functional groups for altitudinal gradient example

distance take into account variance and covariance between traits when calculate the criteria to merge two species or two groups of species.

Using Ward and beginning the partition with 4 groups, we obtained two groups very homogeneous. One of them with 19 species having high LNC and LPC and low WD, plant attributes that allow rapid acquisition of resources (acquisitive type) and other group with 26 species having low SLA, LPC and LNC but with high WD, attributes that permit conservation of resources at plant tissues (conservative or retentive type). None of these two groups are split in successive partitions.

The other two groups have 30 and 31 species. One has high SLA and low LDMC, WD and Tough, with intermediate LPC and LNC; the other has high WD but medium Tough, and very high SLA and very low LDMC. These two groups were splitting in two when six clusters were defined: one with attributes supporting conservative functions in the ecosystem and the other supporting acquisitive functions. The best option is to consider the six clusters defined using Ward algorithm because groups have significant differences among trait profiles

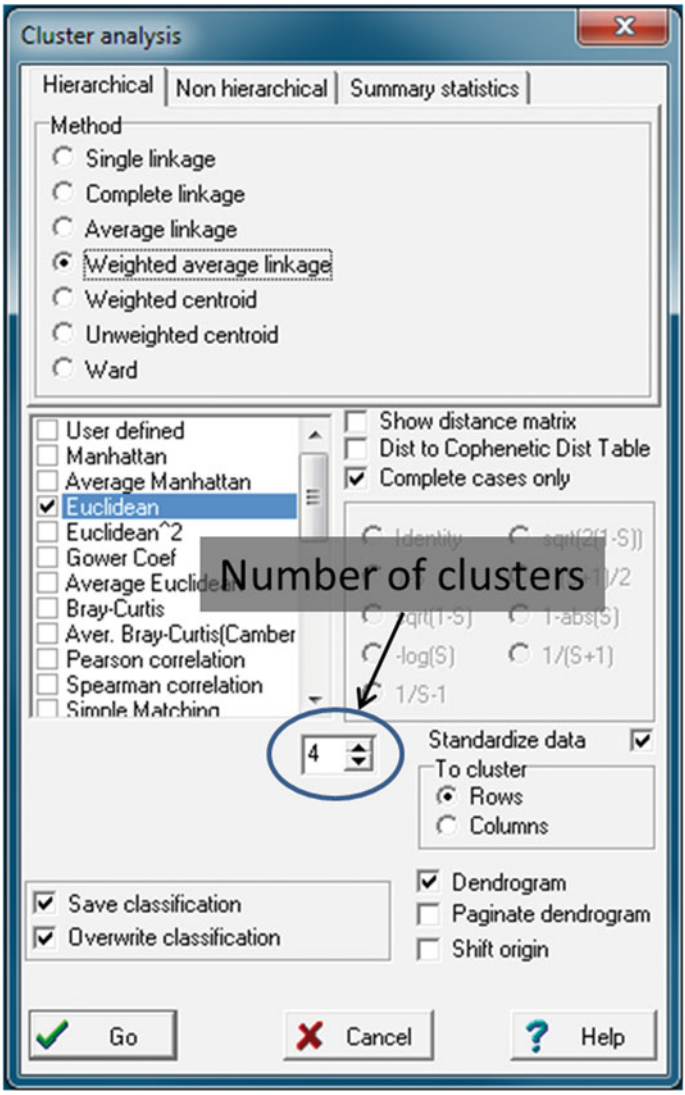


Fig. 4.13 Cluster methods to define plant functional groups for the altitudinal gradient example. When the number of clusters is fixed, the classification for each row is saved in a new column at the end of the data file. Uncheck over-write classification box if you want to keep different columns for each analysis you run

(Table 4.1), has almost similar richness, its mean profiles are interpretable and the range of variation of traits within the cluster is small allowing the ecological description of the group in the ecosystem.

Table 4.1 Results of mean vector comparisons for successive cluster partitions with average linkage and Ward methods using Hotelling test adjusted by Bonferroni and 0.05 significance level

Cl	Seven clusters		Six clusters		Five clusters		Four clusters	
	AvLin	Ward	AvLin	Ward	AvLin	Ward	AvLin	Ward
1	A(2)	A (26)	A(2)	A (26)	A(2)	A (26)	A(2)	A (26)
2	B(6)	B (19)	B(6)	B (19)	B(6)	B (19)	B(6)	B (19)
3	C(6)	C (13)	C(6)	C (13)	C(6)	C (13)	C(28)	C (30)
4	D(27)	D (15)	D(27)	D (15)	D(28)	D (30)	D(70)	D (31)
5	DE(1)	E (15)	D(1)	E (15)	E(64)	E (18)		
6	E(32)	F (12)	E(64)	F (18)				
7	F(32)	G (6)						

AvLin average linkage. Means with a common letter are not significantly different ($p \leq 0.05$). Number of species in each cluster in brackets

4.2.3.2 Characterization of Plant Functional Types for Altitudinal Gradient Example

Mean trait values for each plant functional type (PFT) where compared using analysis of variance followed by Fisher LSD mean comparison (Table 4.2) using the selected cluster method (file trait by spp Altitudinal Gradient with PFT.IDB2 is available for download via Springer's Extra Materials website: http://extras.springer.com/). The extreme PFT were Acquisitive 1 (ACQ1) with the highest values for LPC and LNC, with high SLA and the smallest WD with tree species with the smallest leaf toughness, and Conservative 1 (CON1) with the highest Tough, intermediate WD and the smallest SLA. The other Acquisitive PFT's (ACQ2 and ACQ3) has small WD and high LPC and LNC; while intermediate PFT (INT) has low values of LDMC and intermediate values for all the other traits. The sixth group, Conservative 2 (CON2) has the highest WD with intermediate Tough, with low values of LNC and LPC (Table 4.2).

4.2.3.3 Distribution of Plant Functional Types along the Gradient

We expect distributional changes along the altitudinal gradient. At low altitude acquisitive groups may be dominant and at higher altitude those PFT with species having conservative attributes would replace species with attributes associated to acquisition like high SLA, and leaf content of N and P. Two variables to estimate species abundance were recorded: basal area and number of individuals.

To study this relationship, we have to merge two data files: the trait by spp Altitudinal Gradient.IDB2 we have used to define PFT with the spp by plot Altitudinal Gradient.IDB2. This last file has information of frequency (*freq*) and basal area (*ba*) for each species in each plot. It also has the variable *altitude* for each observational unit (*plot*).

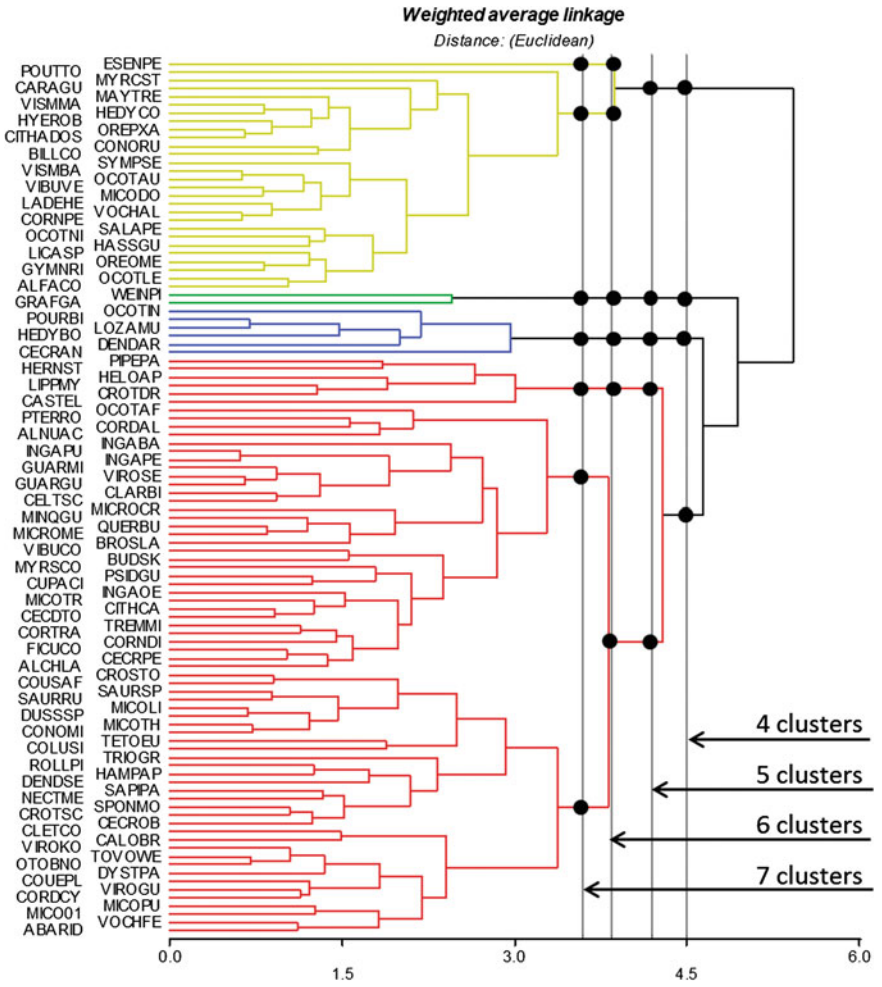


Fig. 4.14 Dendrogram for the altitudinal gradient example obtained using average linkage and Euclidean distance. *Vertical lines* indicate cutting distance to define 4–7 clusters. *Dots* were used to show interception of cutting distance with dendrogram branches in each partition. Codes for branches at the y-axis correspond to species

To merge these files using InfoStat or FDiversity (FDdiversity version of these databases with *.FDDB is available for download via Springer's Extra Materials website: http://extras.springer.com/), we have to open the file (or click in the last file if it is already open) and select Data, Merge tables, and then Merge Horizontal (Fig. 4.16). Then select the variable *species* as *concatenation variable* and press OK, in the new window select the file with the trait values and the classificatory variable used to define the PFT (variable *PFT* if you have assigned this name). Rename this new file for subsequent use. We have saved it as Concat

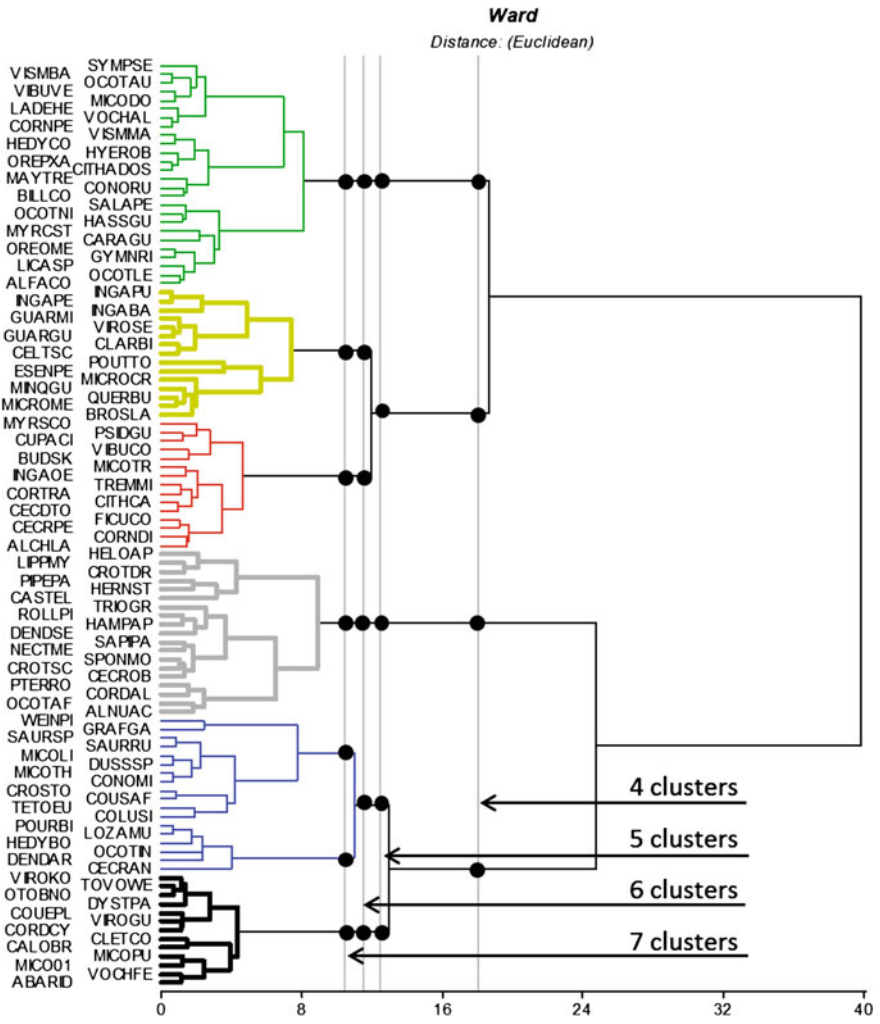


Fig. 4.15 Dendrogram for the example altitudinal gradient obtained using Ward and Euclidean distance. *Vertical lines* indicate cutting distance to define 4–7 clusters. *Dots* were used to show interception of cutting distance with dendrogram branches in each partition. Codes for branches at the y-axis correspond to species. Different widths for branch lines were used to improve visual distinction among groups

Altitudinal Gradient with PFT.FDDB (available for download via Springer's Extra Materials website: http://extras.springer.com/).

The relative contribution of each PFT along the gradient may be study estimating the total basal area by plot and PFT, or the total frequency. We calculate these using InfoStat and selecting Summary statistics from Statistics menu,

Table 4.2 Trait summary statistics for plant functional types (PFT) in altitudinal gradient example

PFT	Summary	SLA	LDMC	Tough	WD	LPC	LNC
ACQ1 (S = 19)	Mean	13.72b	431.87b	0.66a	0.31a	1.87c	32.59d
	Minimum	9.90	216.77	0.25	0.16	1.24	24.99
	Maximum	19.07	680.93	0.96	0.45	2.66	41.20
ACQ2 (S = 15)	Mean	8.77a	490.00c	0.76ab	0.48b	1.47b	24.47b
	Minimum	7.00	394.80	0.50	0.30	1.23	20.80
	Maximum	12.29	657.07	1.10	0.64	1.93	32.42
ACQ3 (S = 18)	Mean	16.56c	313.67a	0.82b	0.46b	1.44b	26.90bc
	Minimum	10.11	224.53	0.34	0.29	1.04	22.26
	Maximum	30.63	389.38	1.63	0.68	2.09	33.83
INT (S = 13)	Mean	12.76b	337.24a	0.70ab	0.46b	0.84a	19.54a
	Minimum	7.88	192.82	0.44	0.35	0.62	13.61
	Maximum	20.56	468.86	0.90	0.63	1.19	24.45
CON1 (S = 26)	Mean	9.32a	475.49bc	0.99c	0.58c	0.97a	18.84a
	Minimum	5.68	342.82	0.75	0.42	0.56	14.61
	Maximum	12.32	616.12	1.34	0.80	1.45	23.32
CON2 (S = 15)	Mean	12.99b	463.20bc	0.83b	0.67d	1.30b	28.80b
	Minimum	10.10	357.74	0.57	0.48	0.64	20.73
	Maximum	15.85	661.57	1.23	0.98	1.93	35.82

S number of species in each PFT; *SLA* specific leaf area ($mm^2\ mg^{-1}$); *LDMC* leaf dry matter content ($mg\ g^{-1}$); *LNC* concentration of nitrogen ($mg\ g^{-1}$); *LPC* concentration of phosphorus ($mg\ g^{-1}$); *Tough* physical strength of the leaves ($N\ mm^{-1}$); and *WD* wood density ($g\ cm^{-3}$); *ACQ* for acquisitive PFT; *CON* for conservative PFT and *INT* for intermediate PFT Means with a common letter are not significantly different ($p \leq 0.05$) using Fisher LSD

then select *freq* and *ba* as variables, *plot* and *PFT_6ward* as class variables, press OK and check only for the sum in the next window, press Go.

We used these results to study changes in PFT with altitude. Regression of total basal area with linear and quadratic altitude terms shows that basal area for PFT CON1, the group with high WD and tough, has a quadratic significant relationship with altitude ($p < 0.0001$ for the model). The model indicates that the basal area for this functional type decrease up to 1,186 masl where rich the minimum of 4.51 m^2 per plot, and then increase up to the highest plot included in the study area at 2,810 masl (Fig. 4.17). The relationships of total basal area were not significant for the other functional groups.

When frequency was considered as a measure of abundance, we find significant relationship only in the distribution of ACQ3 PFT. At low altitudes as well as high altitude the total number of individuals are lower than at intermediate high (Fig. 4.18). No significant relationship with altitude was found for the other PFT's. There are some tendencies (Fig. 4.19), but the variability among plot is very high.

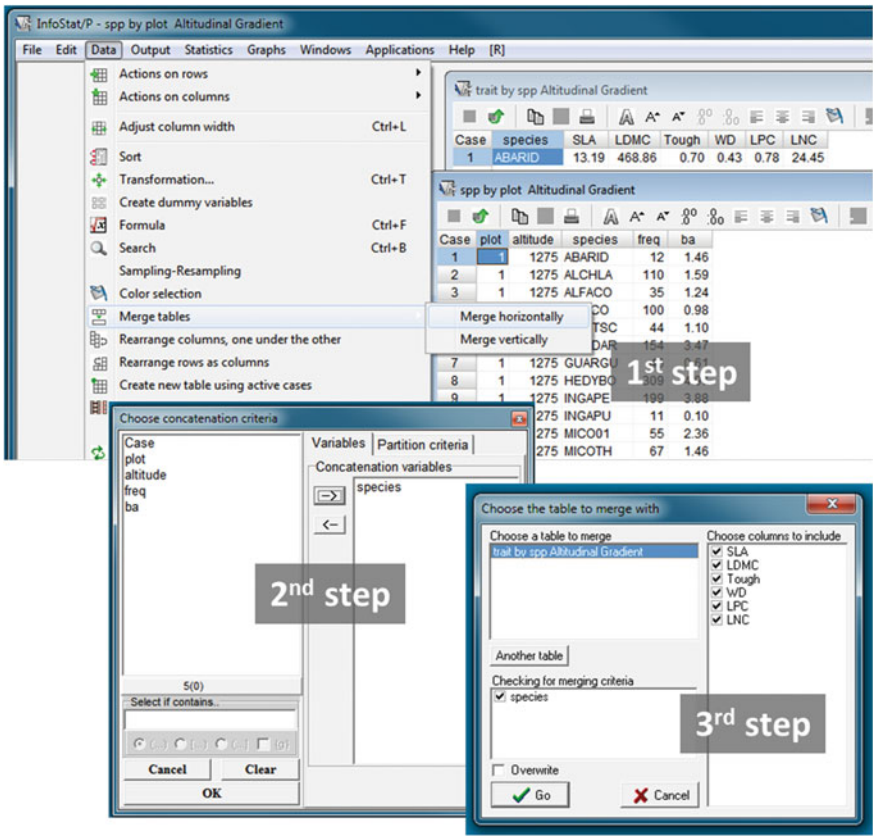


Fig. 4.16 Procedure to merge tables using InfoStat or FDiversity. Be aware and confirm that the concatenation variable has the same name in both files and the variables are of the same type. First step: select data, merge tables, merge horizontally. Second step: choose concatenation criteria, in our case it is species. Third step: choose the table to merge with, in our case 'trait by spp Altitudinal Gradient' and the variables in the right appears. In case you want to avoid merging one of them, uncheck it

4.2.4 Changes of Functional Diversity Indices with the Altitude

To calculate functional diversity indices we have to use the concatenated file with information on species abundance and on functional traits. When the traits have different variances or their measurement scales are different, standardization is necessary because calculations involved in indices estimation are scale dependent. Only single-trait indices like FRO and FDvar may be estimated from variables in

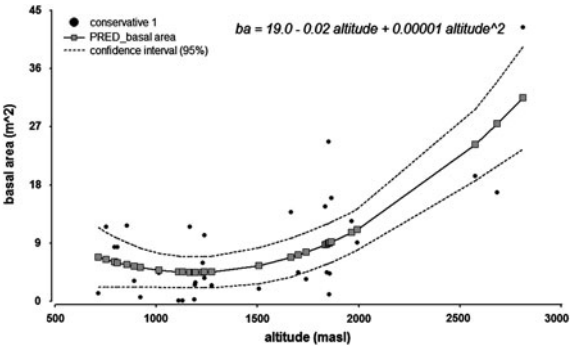


Fig. 4.17 Linear relationship of total basal area (ba) per plot for plant functional type conservative 1 with altitude. The minimum of basal area is at 1,186 masl

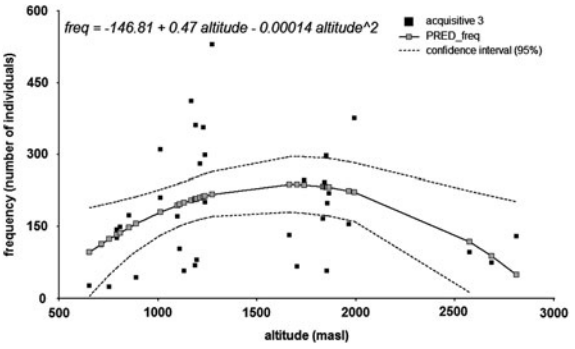


Fig. 4.18 Linear relationship of frequency (freq) per plot for plant functional type acquisitive 3 with altitude. The maximum frequency is attended at 1,653 masl

the original scale. CWM always has to be estimated using original values because the interpretation depends on the measurement scale.

The standardization must be done using a file with one row by species and as much as needed columns to include all variables. In this example we only have continuous variables, so no particular recodification is needed. In a given study we may have more than one value for species determined following the protocols. For example, if plasticity of species is to be investigated we would need to determine one trait value for each condition. This must not be confounding with the replication of individuals or samples by individual needed to fulfill the Cornelissen et al. (2003) protocol. In cases where more than one value is recorded, this could be called populations and one species may have several populations (Lavorel et al. 2008).

4.2.4.1 Trait Values Standardization and Merging Files

Standardization of trait values has to be done according to the field design. If only one value of each trait by species is calculated, the database to standardize the

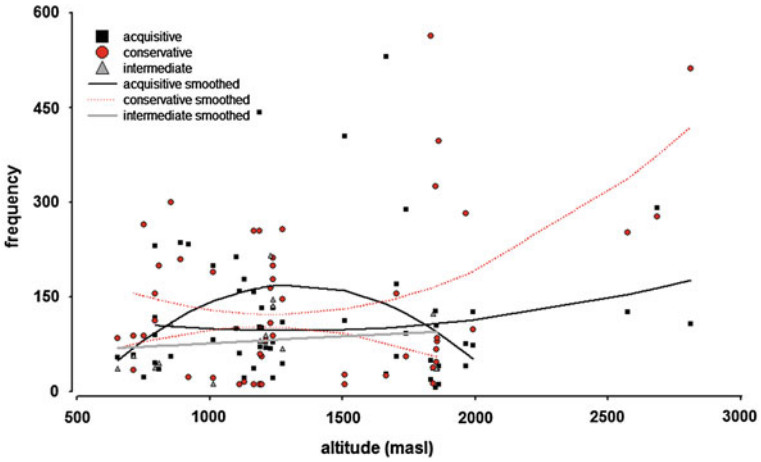


Fig. 4.19 Smoothed quadratic adjustment for frequency changes with altitude. Tendencies for acquisitive plant functional types (excluding acquisitive 3) show a maximum between 1,300 and 1,600 masl, and the relationship are not significant. One of the conservative plant functional groups shows an increase in frequency from 1,300 m up, but the other has no presence in the plots with more than 1,800 masl

traits must be the one with the species list. See Box 4.1 for details and one example of the possible options.

Box 4.1: Options to standardize trait values

Data table with four species (A, B, C, and D) in two conditions with two samples by condition. In case1, there is only one trait value for each species, in case2 there is one value of the trait for each condition, and in case3 there is one value of the trait in each sample**.**

Data table for two conditions with 2 samples each

sp	cond	sample	t1_case1	t1_case2	t1_case3	EST_t1_case3
A	1	1	5	5	5	−0.73
B	1	1	22	22	22	1.46
C	1	1	8	8	8	−0.34
A	1	2	5	5	4	−0.86
B	1	2	22	22	19	1.08
A	2	1	5	7	7	−0.47
B	2	1	22	18	18	0.95
D	2	1	4	4	4	−0.86
A	2	2	5	7	8	−0.34
B	2	2	22	18	24	1.72
C	2	2	8	12	6	−0.60
D	2	2	4	4	3	−0.99

Trait values for case one in variable *t1_case1*, for case two in variable *t1_case2* and for case three in variable *t1_case3*.
Case 1: one value by species

sp	t1_case1	EST_t1_case1	EST_wrong1
A	5	−0.57	−0.73
B	22	1.47	1.34
C	8	−0.21	−0.36
D	4	−0.69	−0.85

The standardization has to be done using the data set where each species appears only one time, because there is only one value for each species. If the transformation is doing with the complete data table, each different value appears as many times as samples where the species is present. Calculating the mean and variance of *EST_wrong1*, you get mean −0.15 and variance 1.03, instead of zero mean and variance of one, as it should be.
Case 2: one value by species for each condition

sp	condition	t1_case2	EST_t1_case2	EST_wrong2
A	1	5	−0.86	−0.85
B	1	22	1.63	1.55
C	1	8	−0.42	−0.42
A	2	7	−0.56	−0.56
B	2	18	1.04	0.99
C	2	12	0.17	0.14
D	2	4	−1.00	−0.99

The standardization has to be done using the data set where each species appears with its different values, one for each condition. If we make the standardization with the complete data table, we would obtain the values of EST_wrong2, which has mean values of −0.02 and variance value of 0.97, instead of zero and one.

Case 3: one value by species for each sample

This is the only case where the standardization is making from the whole data table. The column t1_case3 has the different values of this trait in different samples. The EST_t1_case3 column shows the standardized values; this has zero mean and unit variance.

For this example we have one trait value for species and six trait measures in a continuous scale with different units. Open the file trait by spp Altitudinal Gradient.FDDB using FDiversity software (see Sect. 4.1 for details). The data based has 106 species in rows and six traits in columns; it includes code for each species.

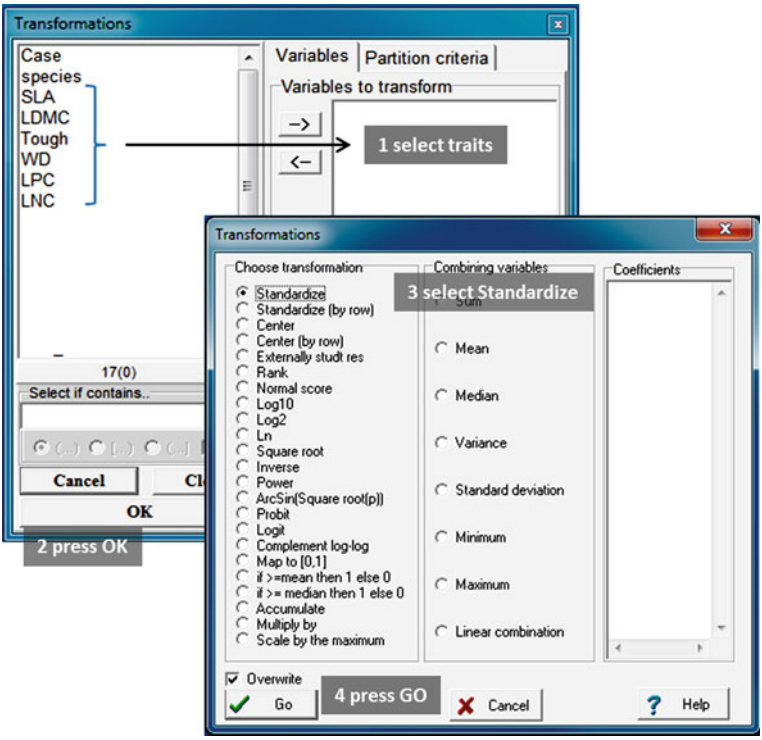


Fig. 4.20 Procedure to standardize traits to have unit variance and zero mean using FDiversity software

Open Data menu and select Transformation (or Ctrl-T), the Transformations window appears. Select the six traits used in this study and transfer them to the variable window, press OK. Another window appears to select the transformation desired, select Standardize and press GO (Fig. 4.20). Six new columns were added to data base, one for each trait with the prefix *EST_*, follow by the trait name.

Then open the data file spp by plot Altitudinal Gradient.FDDB. Open the Data menu and merge both tables as explained in Sect. 4.1 and Fig. 4.16 using the species as concatenated variable. Merge original and standardized sets of traits because the first will be used to estimate CWM and the second to estimate functional diversity indices. Save the resulted file with the same or other name for future analysis. You always may be able to open the file Concat spp by plot Altitudinal Gradient.FDDB (available for download via Springer's Extra Materials website: http://extras.springer.com/).

4.2.4.2 Estimation of Functional Diversity Indices

We follow the same steps explained in Sect. 4.1.3.2 to estimate all the functional diversity indices and obtained one new data table with 38 rows, one for each plot, and 31 columns. The first column identifies the plot; the others are for the indices selected. In our case there are 14 columns for multi-trait indices, six columns for single-trait index FRO and other six columns for single-trait FDvar, and the last four columns for the species biodiversity indices. We did a second run using the traits in its original scales to calculate CWM for each trait and then merge both resulted files to have all numerical results together. As we need the altitude of each plot, we also used merge option to copy the variable *altitude* from the file spp by plot Altitudinal Gradient.IDB2.

We have frequency and basal area to weigh the contribution of each species to the community, so we calculate the set of indices and CWM for both variables following the same procedure. At the end we should have two result files, one for basal area and the other for frequency with 38 rows and 38 columns. These two files have similar values for indices like FAD1, FAD2, FD (both versions), Convex Hull and FRic because the abundance is not involved in their calculations (files Res indices with freq.FDDB and Res indices with ba.FDDB are available for download via Springer's Extra Materials website: http://extras.springer.com/).

Note that Convex hull and FRic have no values for plots 37 and 38; this is due to restrictions of these indices that need at least more species than traits to allow estimation of the multivariate volume. These two plots have only four species and the number of traits is six. The species must not be distributed in a line, in which case the hull volume is zero, like in plots 19 and 20, that even though they have seven species, the convex hull has no volume.

4.2.4.3 Changes of Functional Diversity Indices with the Altitude

Linear regression of trait community weighted means (CWM) using altitude as independent variable (Table 4.3) show significant and positive relation of leaf dry matter content (LDMC) when CWM is calculated using basal area or frequency. Also wood density (WD) shows positive relation when CWM is calculated using basal area but do not show significant tendency when frequency is used (Fig. 4.21).

Specific leaf area (SLA, Fig. 4.22) and leaf nitrogen content (LNC) has significantly inverse relations with altitude, both when their CWM are calculate using basal area or frequency (Table 4.3). The relationships reflect that communities at low altitude are dominate by species with acquisitive attribute, and at higher altitude species with conservative attribute are more frequent and have individuals with larger basal area.

This type of results may be used to relate plant attribute and plant functional diversity to other variables associated with altitude, like precipitation or temperature. Often these relations are useful to predict community composition variations when climatic changes occur.

Table 4.3 Fitted linear models and p-values for relationship of community weighted mean (CWM) of traits values with altitude

CWM	Weight	p-value	Linear model
SLA	ba	0.0084	$14.816 - 0.0014 \times$ altitude
	freq	0.0317	$14.881 - 0.0011 \times$ altitude
LDMC	ba	0.0148	$376.576 + 0.023 \times$ altitude
	freq	0.0206	$362.199 + 0.025 \times$ altitude
Tough	ba	0.0929	
	freq	0.6710	
WD	ba	0.0055	$0.36483 + 0.00006 \times$ altitude
	freq	0.0654	
LPC	ba	0.7723	
	freq	0.6964	
LNC	ba	0.0195	$29.26044 - 0.00265 \times$ altitude
	freq	0.0266	$28.94511 - 0.0022 \times$ altitude

Weights of CWM are basal area (ba) and frequency (freq). *SLA* specific leaf area ($mm^2\ mg^{-1}$); *LDMC* leaf dry matter content ($mg\ g^{-1}$); *Tough* physical strength of the leaves ($N\ mm^{-1}$); *WD* wood density ($g\ cm^{-3}$); *LNC* concentration of nitrogen ($mg\ g^{-1}$); and *LPC* concentration of phophorus ($mg\ g^{-1}$). Estimated linear models are shown only for significant relationships ($p < 0.05$)

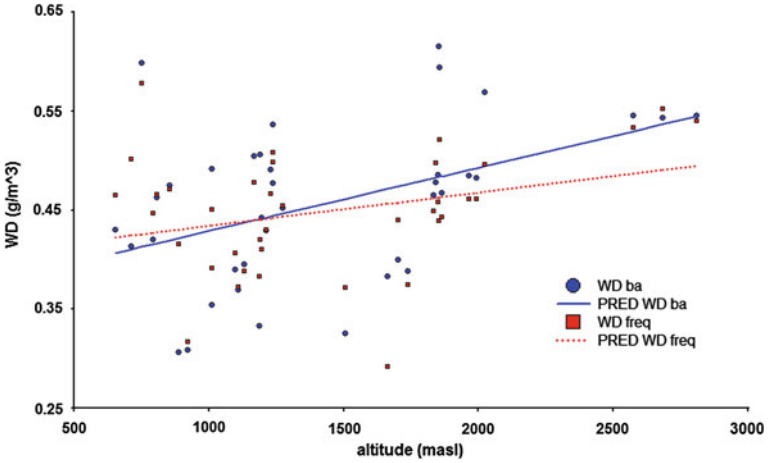


Fig. 4.21 Fitted model for community weighed mean of wood density (WD) using basal area (ba) and frequency (freq) as loaded variable. In both cases the WD increases as altitude increases showing dominance of species with conservative attributes at high altitude

We perform linear regressions for all functional diversity indices and using basal area and frequency for those that include weighted variable to account for relative importance of each species. We obtained significant relationship only for FDiv, both with basal area and with frequency, and also for rRao when it was estimated using frequencies.

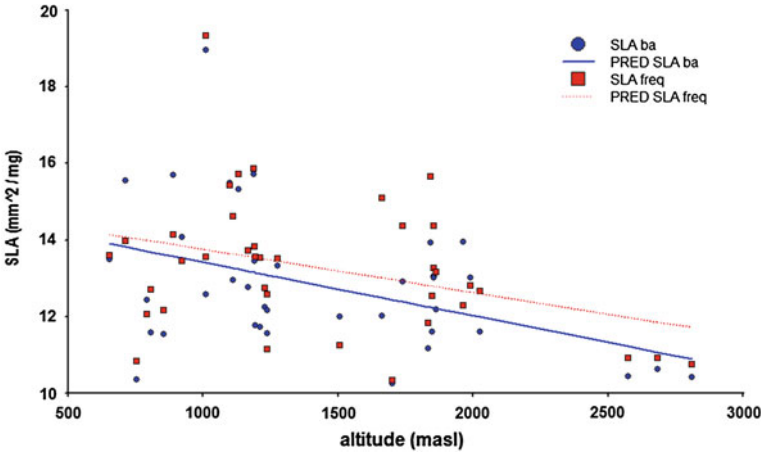


Fig. 4.22 Fitted model for community weighed mean of specific leaf area (SLA) using basal area (ba) and frequency (freq) as loaded variable. In both cases the SLA decreases as altitude increases showing dominance of species with acquisitive attributes at low altitude

Functional divergence, measure using FDiv (Villéger et al. 2008), quantify how the trait values are spread in the multivariate range of the trait space. It varies between zero and one being one the most evenly distributed abundance in the trait space. The linear model using basal area ($FDiv_{ba} = 0.864166 - 0.000064 \times$ altitude) was significant ($p = 0.0319$) and estimate a minimum of 0.684 at the highest altitude of 2,810 masl and a maximum value of 0.822 at the lower altitude of 653 masl. Both values are closer to one showing that the abundance of trait values are well distributed in the altitudinal gradient being significantly lower at highest altitude, this may be due to differences in a few traits or an overall effect. We will investigate these options using the single-trait index FDvar. Result with frequency is also significant ($FDiv_{freq} = 0.85887 - 0.00007 \times$ altitude) with $p = 0.0079$. The estimated minimum and maximum where similar: 0.662 at 2,810 masl and 0.813 at 653 masl.

Linear regression of relative Rao index was also significant ($p = 0.0495$) with frequency as loaded variable. The model ($rRao_{freq} = 0.59139 - 0.00008 \times$ altitude) allows estimation of a maximum 0.54 at 653 masl and a minimum of 0.37 at 2,810 masl. These are relative values obtained as the ratio between observed values (Rao index) and the highest value that may be obtained with the same set of species, which implies with the same set of trait values, varying only the relative abundance among species. As higher the rRao as closer the community is to the best distribution of individual (frequency) or biomass (basal area) among the species. In our case, as Rao do not change significantly with the altitude, we may conclude that dominance of some species (or even only one species) is responsible for the unbalance in the trait space.

Even more, we may say that at higher altitude relative abundance changes of same species present may increase functional diversity in a 63% (1 – 0.37 = 0.63). Arguments like this should be taken with care when the index is calculated using values of traits. To interpret the ratio rRao = Rao/Rao_{max}, the distance matrix between species must be ultrametric. When phylogenetic trees are used instead of traits, this condition is fulfilled. In some cases the use of continuous variables may lead to maximum value of Rao index with only two species of contrasting traits. If this happens the rRao has no functional ecological interpretation.

4.3 Case Study: Changes in Functional Diversity in a Chronosequence

This example presents a subset of field data taken in the south of Mexico, where the '*roza-tumba-quema'* (RTQ) practice is widely used. The RTQ consists in clear, cut the trees and burn them before use the land for cropping. After two or three years the soil fertility decrease and the land is abandoned.

4.3.1 Sample Design and Trait Evaluation

The sample units were located in a chronosequence of 5 times of abandonment (time 1 with average 5 years of abandon, time 2 with an average of 15 years, time 3 with an average of 20 years, time 4 with an average of 25 and time 5 is a natural control. For a total of 20 sample units with an area of 0.25 ha (50 × 50 m), the basal area of each individual was calculated for trees with more than 10 cm of diameter at breast height (Chan-Dzul 2010). Dominant species per plot which together accounted for 75% of basal area (Grime 1998) were identifying to produce the species list to determine trait values.

The objective was to determine and compare the functional diversity according to age of abandonment of forests, based on four functional traits: height (MaxHeight, m), wood density (WD, g cm^{-3}), leaf phenology (deciduous or evergreen), dispersing agent (anemochory, autochory and zoochory) and reproductive system (monoic, dioic and hermaphrodite).

Measurements of functional traits in dominant species were made in the 113 species identified as dominant in the 20 sites evaluated. The sampling methods and categories of allocation of functional traits were based on the protocols proposed by Cornelissen et al. (2003).

4.3.2 The Database

The data were synthesized in two files, one with the information about traits (Fig. 4.23) and the other with information about distribution of species in the

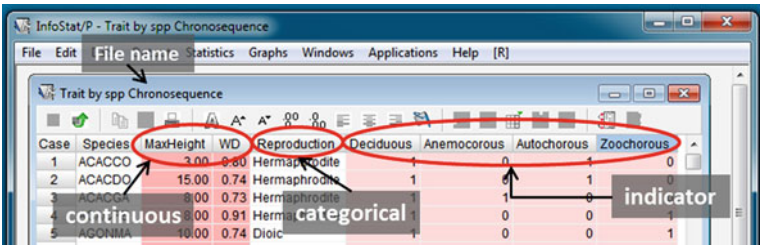


Fig. 4.23 Database of chronosequence example showing the traits by species information of file 'Trait by spp Chronosequence' using InfoStat

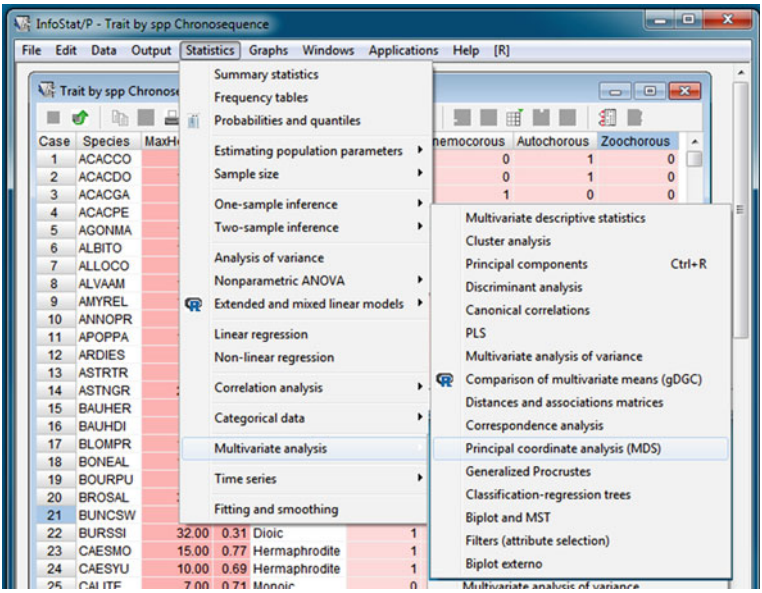


Fig. 4.24 Sequence to perform principal coordinate analysis to synthesize dummy and indicator variables of file 'Trait by spp Chronosequence' using InfoStat

observational units (Fig. 4.24). The trait data base trait by spp Chronosequence.IDB2 (available for download via Springer's Extra Materials website: http://extras.springer.com/) has one row for each species which is identify by a code (*species* column), the mean values for the two continuous traits considered (*MaxHeight* and *WD*), one column for reproductive system *(Reproduction*) with three categories, one column for phenology (*Deciduous*), and three other columns to indicate seed dispersion (file trait by spp Chronosequence.FDDB is available for download via Springer's Extra Materials website: http://extras.springer.com/).

The second database spp by plot Chronosequence.IDB2 (available for download via Springer's Extra Materials website: http://extras.springer.com/) has

information about the presence of each species at each observational unit (*Plot*), and the time of abandonment (*Time*). The importance of each species was quantified with the abundance (*Abundance*, number of trees) and with the total basal area (*BasalArea*, m^2). The database has 735 rows, one for each species present in each plot, by five columns (file spp by plot Chronosequence.FDDB is available for download via Springer's Extra Materials website: http://extras.springer.com/).

Functional variation in the chronosequence may be study defining functional groups and interpreting the relative contribution of each functional group at different states (see Chap. 2 for complete details of statistical procedures). To complement this approach we calculate functional indices and test its relationship with age of abandonment to explore changes in different facets of functional diversity.

4.3.3 Changes of Plant Functional Types in the Chronosequence

To define functional types we performed cluster analysis using trait information in the dominant species (trait by spp Chronosequence.IDB2) using InfoStat (Di Rienzo et al. 2010). Open InfoStat and from the file menu, open the data file, or after install InfoStat double click the file name. The database should look like in Fig. 4.23 with 113 rows and eight columns.

The traits used to link functional properties to ecosystem services are express in different scales. Wood density (*WD*) and maximum height (*MaxHeight*) are continuous variables, while phenology (*Deciduous*) is a binary variable (1 = deciduous, 0 = evergreen), reproductive system is a nominal variable with three categories, and the seed dispersion methods are indicative variables.

4.3.3.1 Formation of Plant Functional Types (PFT) for Chronosequence Example

One option to take full advantage of information in the continuous variables when there are also binary and nominal variables is to use principal coordinate analysis (PCoA, also known as multidimensional scaling) to synthesize binary variables, indicators and dummy variables in continuous indices called principal coordinates. As PCoA use numerical variables, nominal variable has to be transformed previously to dummy variables (*Reproduction_Hermaphrodite* and *Reproduction_Monoic*). Two columns will be added at the end of the file shown in Fig. 4.23 (file trait by spp Chronosequence for PCoA.IDB2 is available for download via Springer's Extra Materials website: http://extras.springer.com/).

We run PCoA using InfoStat (Fig. 4.24) with distance measure Jaccard and save the first five principal coordinates. The principal coordinates appear at the end of the data file (Fig. 4.25). We use the first five axes accounting for 85% of the

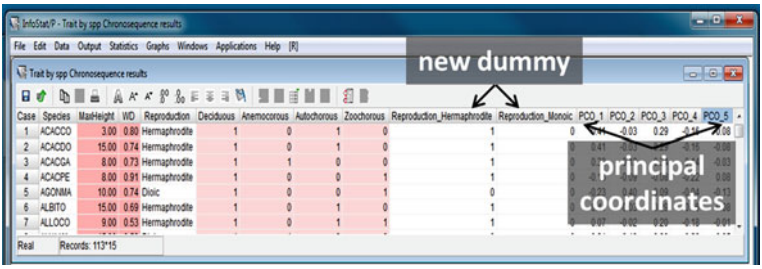


Fig. 4.25 Enlarged database of chronosequence example showing the traits by species information and the new added variables

variability (see information on the result windows). Then, having all variables in continuous scale we may use Euclidean distance to perform cluster analysis and Ward as linkage algorithm.

We continuous with InfoStat following the same steps as in altitudinal gradient example (see steps shown in Figs. 4.11 and 4.12). The dendrogram has six main branches (Fig. 4.26). We re-run the cluster analysis asking for six groups and then perform a multivariate analysis of variance with the same variables. Null hypothesis of equal vector means is rejected ($p < 0.0001$) and mean vector comparison shows significant differences among all the clusters (Box 4.2).

Box 4.2: Result of multivariate mean comparison in the chronosequence

Hotelling test Alpha:=0.05

Error: Pooled covariance matrix df: 107

PFT	MaxHeight	WD	PCO_1	PCO_2	PCO_3	PCO_4	PCO_5	n	
4	12.71	0.69	-0.16	-0.09	-0.08	-0.22	0.08	28	A
5	15.50	0.60	-0.21	0.39	-0.08	-0.03	-0.13	16	B
6	15.62	0.72	-0.25	-0.24	0.15	0.22	-0.07	21	C
1	14.21	0.69	0.33	0.01	0.27	-0.15	-0.07	19	D
2	10.10	0.62	0.07	0.28	0.15	0.24	0.30	10	E
3	13.53	0.64	0.32	-0.09	-0.32	0.13	-0.02	19	F

Means with a common letter are not significantly different (p<= 0.05)

4.3.3.2 Characterization of Plant Functional Types for the Chronosequence Example

To characterize the clusters we used average trait values. In this study there are continuous and categorical variables transformed in continuous principal coordinates. Mean values of woody density and maximum height are interpretable based on mean values shown in Box 4.2, but the rest (PCO_1–PCO_5) even though may be interpret through the coefficients loading the contribution of each original variable to each axis, it is strongly recommend to use contingency tables with the original variables to characterize clusters. Contingency table allows testing the null hypothesis of no association between two categorical variables based on frequencies of cross categories.

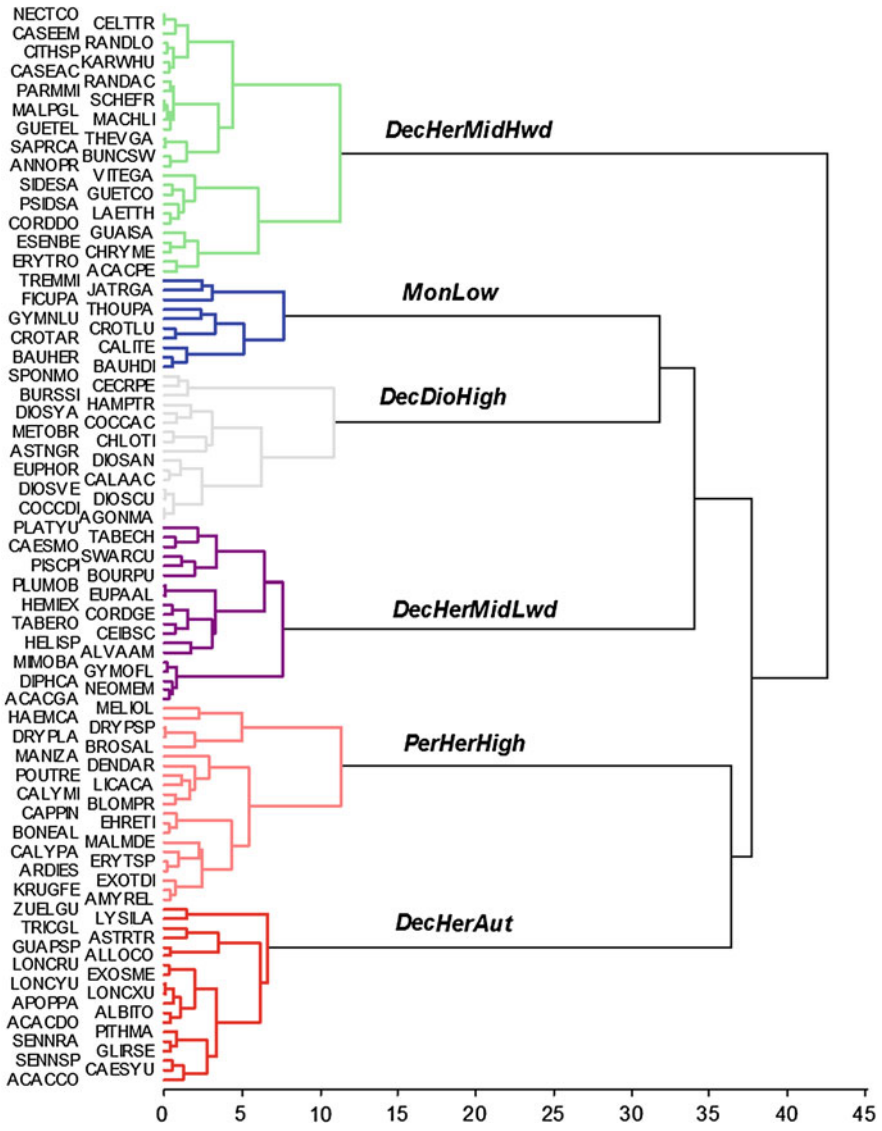


Fig. 4.26 Dendrogram obtained using Euclidean distance with five principal coordinates and the two continuous traits. Six plant functional types (PFT) were identified in the chronosequence, species of each PFT are indicated by code name at left side. Species of PFT are characterized as: deciduous-hermaphrodite-autochory (DecHerAut); perennial-hermaphrodite-high strata (PerHerHigh); deciduous-hermaphrodite-middle strata-low wood density (DecHerMidLwd); deciduous-dioic-high strata (DecDioHigh); monoic-autochory (MonLow); deciduous-hermaphrodite-middle strata-high wood density (DecHerMidHwd)

Three contingency tables are needed for seed dispersion because it is an indicator variable with three categories (categories of indicator variables are non-exclusives). None of them are independent of plant functional types ($p < 0.0001$ in Box 4.3) so we may use the proportion of each category to explain differences among plant functional types.

Box 4.3: Contingency table for PFT and seed dispersion methods

Absolute frequency
In columns:Anemocorous

PFT	0	1	Total
DecHerAut	19	0	19
DecHerMidLwd	0	19	19
DecHerMidHwd	28	0	28
DecDioHigh	15	1	16
PerHerHigh	20	1	21
MonLow	9	1	10
Total	91	22	113

Statistic	Value	df	p
Chi-square (Pearson)	95.21	5	<0.0001
Chi-square (ML-G2)	89.38	5	<0.0001
Contingency Coef. (Cramer)..	0.65		
Contingency Coef. (Pearson..	0.68		

Absolute frequency
In columns:Autochorous

PFT	1	0	Total
DecHerAut	19	0	19
DecHerMidLwd	0	19	19
DecHerMidHwd	0	28	28
DecDioHigh	1	15	16
PerHerHigh	2	19	21
MonLow	4	6	10
Total	26	87	113

Statistic	Value	df	p
Chi-square (Pearson)	83.95	5	<0.0001
Chi-square (ML-G2)	87.75	5	<0.0001
Contingency Coef. (Cramer)..	0.61		
Contingency Coef. (Pearson..	0.65		

Absolute frequency
In columns:Zoochorous

PFT	0	1	Total
DecHerAut	15	4	19
DecHerMidLwd	16	3	19
DecHerMidHwd	0	28	28
DecDioHigh	0	16	16
PerHerHigh	1	20	21
MonLow	5	5	10
Total	37	76	113

Statistic	Value	df	p
Chi-square (Pearson)	71.51	5	<0.0001
Chi-square (ML-G2)	84.88	5	<0.0001
Contingency Coef. (Cramer)..	0.56		
Contingency Coef. (Pearson..	0.62		

PFT: plant functional type

Contingency tables for phenology (Box 4.4) and for reproduction system (Box 4.5) show significant differences among species proportion of each category in plant functional types. Both original variables were transform to dummy variables before clustering so the relative frequencies at the contingency table must be interpret according codification: for phenology one column indicate deciduous condition with one and evergreen with zero; for reproduction system two columns were use in the clustering, but for contingency tables we may use the original variable *Reproduction* with the three categories.

Box 4.4: Contingency table for PFT and Deciduous

Absolute frequency
In columns:Deciduous

PFT	0	1	Total
DecDioHigh	0	16	16
DecHerAut	0	19	19
DecHerMidHwd	0	28	28
DecHerMidLwd	0	19	19
MonLow	3	7	10
PerHerHigh	21	0	21
Total	24	89	113

Statistic	Value	df	p
Chi-square (Pearson)	100.45	5	<0.0001
Chi-square (ML-G2)	104.65	5	<0.0001
Contingency Coef. (Cramer)..	0.67		
Contingency Coef. (Pearson..	0.69		

Box 4.5: Contingency table for PFT and reproduction system

Absolute frequency
In columns:Reproduction

PFT	Hermaphrodite	Dioic	Monoic	Total
DecHerAut	17	2	0	19
DecHerMidLwd	17	2	0	19
DecHerMidHwd	28	0	0	28
DecDioHigh	0	16	0	16
PerHerHigh	17	4	0	21
MonLow	0	0	10	10
Total	79	24	10	113

Statistic	Value	df	p
Chi-square (Pearson)	184.15	10	<0.0001
Chi-square (ML-G2)	133.39	10	<0.0001
Contingency Coef. (Cramer)..	0.74		
Contingency Coef. (Pearson..	0.79		

The characterization is based on mean values for continuous variables (Box 4.2) and on contingence tables for the rest of the variables (Box 4.3, 4.4 and 4.5). There are three groups with high wood density (clusters 1, 4 and 6) and other three with low wood density (clusters 2, 3, and 5).

Cluster 1 has deciduous species, hermaphrodite and autochoras, with some of them zoochoras (DecHerAut). Cluster 4 has deciduous species, hermaphrodites and from the middle strata (DecHerMidHwd). Cluster 6 will be called PerHerHigh, with all the perennial species, hermaphrodites from the highest strata. Cluster 2 will be called MonLow, it has only monoic species and they have autochory or zoochory. Cluster 3 has deciduous species hermaphrodite with low wood density growing in the low strata (DecHerMidLwd). Cluster 5 has only dioic species, deciduous from the highest strata (DecDioHigh). File trait by spp Chronosequence with PFT.IDB2 with the results of this analysis is available for download via Springer's Extra Materials website (http://extras.springer.com/).

4.3.3.3 Distribution of Plant Functional Types in the Chronosequence

We expect that distribution of plant functional types changes along the chronosequence. To study this relationship, we have to merge two data files: the trait by spp chronosequence with PFT.IDB2 we have used to define PFT with the spp by plot chronosequence.IDB2. To merge these files using InfoStat or FDiversity, we have to open the file (or click in the last file if it is already open) and select Data, Merge tables, and then Merge Horizontal. Then select the variable *species* as *concatenation variable* and press OK, in the new window select the file with the traits values and the classificatory variable used to define the PFT. Rename this new file for subsequent use like you did with the previous example. You should obtain a file like Concatenated Chronosequence.IDB2 (available for download via Springer's Extra Materials website: http://extras.springer.com/).

To show the relative contribution of each PFT to the total basal area we make a stack bar plot (Fig. 4.27) using relative values. Total basal area increases as time of abandoned increases: 6.76, 16.86, 17.58, 27.37, and 30.11 m^2 from time 1 to natural forest. Express the basal area of functional type using proportion to the total allows avoiding the effect of differences in total basal area in the chronosequence. Plant functional type DecHerAut decrease its abundance with time of abandonment while PerHerHigh clearly increase its abundance. This reflects the successional tendency due to replacement of deciduous species with low wood density (DecHerAut) by evergreen species of the highest strata with high wood density (PerHerHigh). There is small contribution of the species growing in the low strata and their contribution is almost constant in the chronosequence

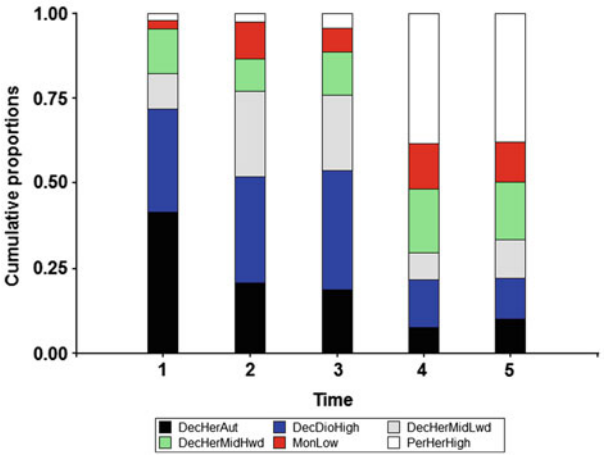


Fig. 4.27 Stack bar showing relative basal area of species in plant functional types of the chronosequence. *DecHerAut* deciduous-hermaphrodite-autochory; *PerHerHigh* perennial-hermaphrodite-high strata; *DecHerMidLwd* deciduous-hermaphrodite-middle strata-low wood density; *DecDioHigh* deciduous-dioic-high strata; *MonLow* monoic-autochory; *DecHerMidHwd* deciduous-hermaphrodite-middle strata-high wood density

(MonLow). The two groups characterize by deciduous species (DecHerMidLwd and DecHerMidHwd) contribute to the medium strata, while species with low wood density are more important in the first stages of succession those with high wood density increase with time. It is also clear that as time pass the proportion of basal area in the highest strata decrease for deciduous species (DecDioHigh). If we compare the distribution of plant functional types after 25 year of abandonment (time 4) with that of the natural forest (time 5) we notice that there are no important differences.

4.3.4 Changes of Functional Diversity Indices in the Chronosequence

To calculate functional diversity indices we have to use the concatenate file with information on species abundance and on functional traits. When the traits have different scales we have to use an appropriate distance like Gower. When we use Gower it is not necessary to standardize the traits because this similarity measure use the range (difference between maximum and minimum) to put all variables in the same scale avoiding effect of heterogeneity of variance and units). We used the

file Concatenated Chronosequence.IDB2, already obtained from InfoStat (file Concatenated Chronosequence.FDDB is available for download from Springer's Extra Materials website http://extras.springer.com/).

4.3.4.1 Estimation of Functional Diversity Indices

We have abundance and basal area to weigh the contribution of each species to the community. We follow the same steps explained in Sect. 4.1.3.2 to estimate all the functional diversity indices using Gower distance and Ward linkage algorithm. We obtained one new data table with 20 rows, one for each plot, and 36 columns. The first column identifies the time, the second one the plot, columns 3–13 have all the functional indices but rRao and Convex hull in the same order as FDiversity provides. From columns 14 to 32 we obtained the eight single-traits FRO indices, the eight CWM and the eight FDvar, but these single-trait indices only have ecological meaning for continuous variables, so we delete those corresponding to dummy and indicator variables retaining only six columns. The last four columns are for the taxonomic indices. As result we have a file with 20 rows and 24 columns using basal area and other file using abundance (files Results Chronosequence with basal area.FDDB and Results Chronosequence with abundance.FDDB are available for download via Springer's Extra Materials website: http://extras.springer.com/).

4.3.4.2 Changes of Functional Diversity Indices in the Chronosequence

We have explored the changes of functional diversity using PFT relative distribution, and in this section we explore the differences among stage of the chronosequence using continuous functional diversity indices. As has been shown in Sect. 4.3.2 FDiversity performs analysis of variance and mean comparisons for each index if we ask for in the Comparison tab (Fig. 4.9).

The output has the information that has been saved in a new table, results for analysis of variance and mean comparisons. We select some functional indices to show the software output (Box 4.5). There are statistical differences for FDp ($p = 0.0075$) and the same index including the importance of each species to the community (wFDp, using basal area) also shows differences ($p = 0.0010$).

Box 4.5: FDiversity output for functional diversity indices in the chronosequence using Gower distance and basal area

ANOVA (FDp: Plot-based functional diversity)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	43.92	10.98	5.26	0.0075
Residuals	15	31.30	2.09		

Adjusted means and standard error for Time

Pairwise comparisons (like Fisher's LSD) (alfa=0.05)

Time	Means	S.E.			
5	13.17	0.72	A		
4	12.49	0.72	A	B	
3	11.42	0.72	A	B	
2	10.40	0.72		B	C
1	9.01	0.72			C

Means with a common letter are not significantly different (p<= 0.05)

ANOVA (wFDp: Weighted plot-based functional diversity)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	64.96	16.24	8.19	0.0010
Residuals	15	29.74	1.98		

Adjusted means and standard error for Time

Pairwise comparisons (like Fisher's LSD) (alfa=0.05)

Time	Means	S.E.		
4	10.24	0.70	A	
5	9.49	0.70	A	
3	7.24	0.70		B
2	6.41	0.70		B
1	5.53	0.70		B

Means with a common letter are not significantly different (p<= 0.05)

ANOVA (FRic: Functional richness)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	0.02	4.2E-03	7.71	0.0014
Residuals	15	0.01	5.5E-04		

Adjusted means and standard error for Time

Pairwise comparisons (like Fisher's LSD) (alfa=0.05)

Time	Means	S.E.			
5	0.10	0.01	A		
4	0.09	0.01	A	B	
3	0.06	0.01		B	C
2	0.04	0.01			C
1	0.03	0.01			C

Means with a common letter are not significantly different (p<= 0.05)

ANOVA (FEve: Functional evenness)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	0.08	0.02	6.39	0.0033
Residuals	15	0.05	3.3E-03		

Adjusted means and standard error for Time

Pairwise comparisons (like Fisher's LSD) (alfa=0.05)

Time	Means	S.E.			
1	0.57	0.03	A		
4	0.57	0.03	A		
5	0.50	0.03	A	B	
2	0.45	0.03		B	C
3	0.40	0.03			C

Means with a common letter are not significantly different (p<= 0.05)

ANOVA (FDiv: Functional divergence)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	0.03	0.01	1.05	0.4128
Residuals	15	0.12	0.01		

ANOVA (FDis: Functional dispersion)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	3.11	0.78	0.30	0.8751
Residuals	15	39.27	2.62		

The set of four functional indices that allows quantifying richness (FRic), evenness (FEve), functional divergence (FDiv) and functional dispersion (FDis) show significant differences in richness and evenness. The last two times differ from the others in functional richness but the tendency is not so clear for evenness, nevertheless all values are around 0.50, indicating that the species in the communities has neither a clear dominance nor a complete uniform distribution. There is no significant differences for FDiv or FDis which indicate that even though the species composition changes the variability among plot of the same age of abandonment is not different from the variability among times ($p = 0.4128$ and $p = 0.8751$).

There were differences in CWM for MaxHeight ($p = 0.0466$) among times of abandonment, times 1 and 3 show the maximum height. In case of wood density (CWM_wd) there were differences among times ($p < 0.0001$) with times 4 and 5 with the highest values (Box 4.6). Taxonomic indices show differences among times (richness, $p = 0.0114$; Shannon, $p = 0.0085$). Times 5 and 4 have the highest values for both indices (Box 4.7).

Box 4.6: FDiversity output for trait community weighted means in the chronosequence using basal area

```
ANOVA (CWM: Community weighted mean (single trait):MaxHeight)
             Df    Sum Sq  Mean Sq  F value  Pr(>F)
Time          4     75.39    18.85     3.13  0.0466
Residuals    15     90.36     6.02

Adjusted means and standard error for Time
Pairwise comparisons (like Fisher's LSD) (alfa=0.05)
Time  Means   S.E.
1     22.77   1.23    A
3     20.60   1.23    A       B
5     18.20   1.23            B
2     17.90   1.23            B
4     17.84   1.23            B
Means with a common letter are not significantly different (p<= 0.05)

ANOVA (CWM: Community weighted mean (single trait):WD)
             Df    Sum Sq  Mean Sq  F value  Pr(>F)
Time          4      0.15     0.04    15.82  <0.0001
Residuals    15      0.04  2.4E-03

Adjusted means and standard error for Time
Pairwise comparisons (like Fisher's LSD) (alfa=0.05)
Time  Means   S.E.
5      0.73   0.02    A
4      0.73   0.02    A
2      0.62   0.02            B
3      0.55   0.02            B       C
1      0.53   0.02                    C
Means with a common letter are not significantly different (p<= 0.05)
```

Box 4.7: FDiversity output for taxonomic diversity indices in the chronosequence using basal area

```
ANOVA (S: Richness (number of species))
             Df       Sum Sq   Mean Sq  F value  Pr(>F)
Time          4       704.00    176.00     4.73  0.0114
Residuals    15       557.75     37.18

Adjusted means and standard error for Time
Pairwise comparisons (like Fisher's LSD) (alfa=0.05)
Time  Means    S.E.
5     44.25    3.05     A
4     41.25    3.05     A        B
3     37.25    3.05     A        B
2     33.75    3.05              B        C
1     27.25    3.05                       C
Means with a common letter are not significantly different (p<= 0.05)

ANOVA (H: Shannon-Weaver index)
             Df       Sum Sq   Mean Sq  F value  Pr(>F)
Time          4         2.26      0.56     5.09  0.0085
Residuals    15         1.66      0.11

Adjusted means and standard error for Time
Pairwise comparisons (like Fisher's LSD) (alfa=0.05)
Time  Means    S.E.
4      3.02    0.17     A
5      2.88    0.17     A        B
3      2.47    0.17              B        C
2      2.45    0.17              B        C
1      2.08    0.17                       C
Means with a common letter are not significantly different (p<= 0.05)
```

4.4 Multivariate Graphical Projection Methods

Principal component analysis is widely used to synthesize multivariate information and project the observations in a plane with the first two principal components. The scatter plot of observations in these two axes allows interpreting relationship among observations. If the variables are added to this graph we obtained a combination of spaces: one for the observations and the other for the variables plotting together, called biplot. This join representation allows interpreting relationship among variables, and also between variables and observations. This kind of graphical representation was used in Chap. 2 in the example of Restinga vegetation.

It is possible to add another set of variables to this space; in this case it is called triplot. The triplot results from the Partial Least Squares (PLS) multivariate analysis (Wold 1985). Like in regression, we have two set of variables: one set taken as response variables and the other as explanatory or regressor variables. Both sets have to be measures in the same experimental or observation units.

We use the chronosequence database to show this technique. With the triplot we may interpret relation among times (observational units), among functional diversity indices (regressor variables) and among traits (response variables). Even more, we may explore relationships between time and traits, between time and

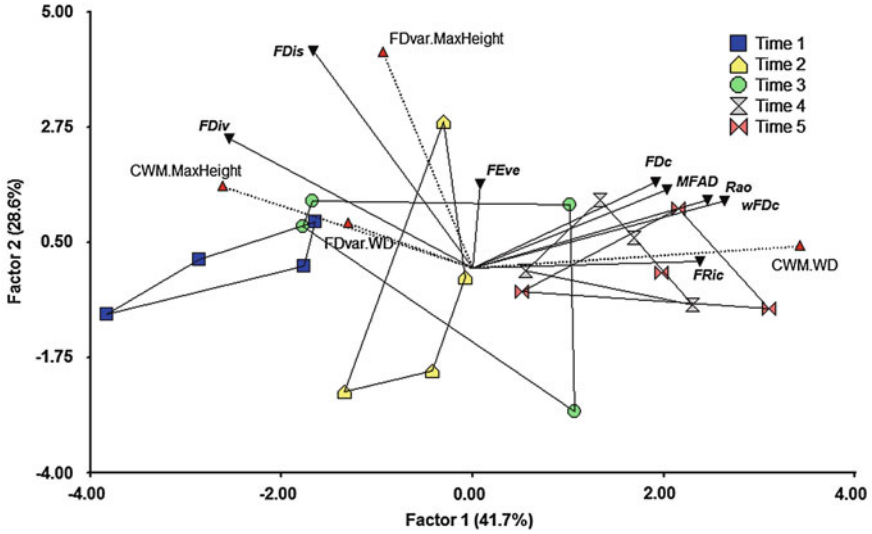


Fig. 4.28 Triplot of CWM and FDvar single-trait indices, functional indices and times of abandonment using partial least squares. *MaxHeight* maximum height; *WD* wood density; *CWM* community weighted mean; *FDvar* functional variability; *FDis* functional dispersion; *FDiv* functional divergence; *FDc* functional diversity base on community dendrogram; *wFDc* weighted FD based on community dendrogram; *FRic* functional richness; *Rao* Rao index; *FEve* functional evenness; *MFAD* modified functional attributes

functional diversity indices, between functional diversity indices and traits, and among time, traits and indices.

Using the example of the chronosequence we perform the PLS using InfoStat. The first two factors in the triplot explain 70% of variability (41.7% Factor 1 and 28.6% Factor 2). CWM for wood density is more or less negatively correlated with CWM of MaxHeight and with FDvar_wd, as CWM increases its variance decreases. MFAD and FDc are highly positively correlated and both are positively correlated with wFDc, FRic and Rao. FEve is not correlated with FRic nor with FDiv; and has moderate correlation with FDis.

We also may interpret the position of 20 sample plots (five times by four replicates) along the first axis. We put and envelop to show the four replicates (points) in each time. The centroid of the envelopes resembles the time tendency; from left to right we find the sequence of the abandonment age. This trend goes in the direction of wood density CWM as expected, and in the direction of FRic and wFDc showing maximum functional richness (Fig. 4.28).

Presenting the results using the triplot may be useful in same cases and could be unnecessary in others. In functional diversity studies, when almost always we have to interpret mean values of the traits (usually by the trait CWMs) jointly with functional indices, and especially if we have also treatments, having all relations in one graph that synthesize most of the variability (in the last example 70%) is very useful.

References

Bermeo DF (2010) Determinación y caracterización de tipos funcionales de plantas (TFPs) en bosques secundarios del Corredor Biológico Volcánica Central Talamanca (CBVC-T) y su relación con variables bioclimáticas. Tesis Mag. Sc. Turrialba, Costa Rica, CATIE

Casanoves F, Di Rienzo JA, Pla L (2008) User manual FDiversity: Statistical software for the analysis of functional diversity, 1st edn, Argentina. www.fdiversity.nucleodiversus.org

Casanoves F, Pla L, Di Rienzo JA, Díaz S (2011) FDiversity: a software package for the integrated analysis of functional diversity. Methods Ecol Evol 2:233–237

Chan-Dzul AM (2010) Diversidad florística y funcional a través de una cronosecuencia de la selva mediana subperennifolia en la zona de influencia de la Reserva de la Biosfera Calakmul, Campeche, México. Tesis Mag. Sc. Turrialba, Costa Rica, CATIE

Cornelissen JHC, Lavorel S, Garnier E, Díaz S, Buchmann N, Gurvich DE, Reich PB, Ter Steege H, Morgan HD, Heijden MGA, van der Pausas JG, Poorter H (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Aust J Bot 51:335–380

Di Rienzo JA, Guzman AW, Casanoves F (2002) A multiple comparisons method based on the distribution of the root node distance of a binary tree obtained by average linkage of the matrix of Euclidean distances between treatment means. J Agr Biol Envir St 7:129–142

Di Rienzo JA, Casanoves F, Pla L (2008) FDiversity, software to estimate functional diversity. RDNDA, Argentina. Number 702841

Di Rienzo JA, Casanoves F, Balzarini MG, Gonzalez L, Tablada M, Robledo CW (2010) InfoStat. Grupo InfoStat. FCA, Universidad Nacional de Córdoba, Argentina

Garnier E, Shipley B, Roumet C, Laurent G (2001) A standardized protocol for the determination of specific leaf area and leaf dry matter content. Func Ecol 15:688–695

Grime JP (1998) Benefits of plant diversity to ecosystems: immediate filter and founder effects. J Ecol 86:902–910

Lavorel S, Grigulis K, McIntyre S, Williams NSG, Garden D, Dorrough J, Berman S, Quétier F, Thébault A, Bonis A (2008) Assessing functional diversity in the field—ethodology matters!. Func Ecol 22:134–147

Villéger S, Mason NW, Mouillot D (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology 89:2290–2301

Walker BH, Kinzig A, Langridge JL (1999) Plant attribute diversity, resilience, and ecosystem function: the nature and significance of dominant and minor species. Ecosystems 2:95–113

Wold HOA (1985) Partial least squares. In: Kotz S, Johnson NL (eds) Encyclopaedia of statistical sciences, vol 6. Wiley, New York

Index

L. Pla et al., *Quantifying Functional Biodiversity*, SpringerBriefs in Environmental Science, DOI: 10.1007/978-94-007-2648-2, © The Author(s) 2012