

# On the identification of optimal plant functional types

Pillar, Valério DePatta

*Departamento de Ecologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 91540-000, Brazil;  
Fax +55513191568; E-mail vpillar@ecologia.ufrgs.br*

**Abstract.** The study of vegetation response to environmental change on a global scale cannot rely on species because most plant species have geographically limited distributions. To allow ecological predictions beyond the scale of the floristic region, models have to rely on vegetation descriptions using plant types other than the species. The crucial problem is how to define the types. Since types are described by traits, the problem translates into one of optimal trait selection. The best plant traits are those that when used to define plant types optimize the perception of association between vegetation and environmental (e.g., climate, disturbance) variation. I consider trait selection as a two-step procedure. The first step is the selection of a larger trait set based on past experience and known practicality, which is used for community description. The second step, for which the paper describes new methods, is accomplished on the data analytically by suitable computer algorithms that can find the optimal subset among the pre-selected traits. This subset defines optimal plant functional types (PFTs). The methods involve a fuzzy set approach and community description by plant types. The optimization algorithms described are tested with data from plant communities in South and North America. The utility of the approach in the evaluation of convergence of phylogenetically distant plant communities is discussed.

**Keywords:** Chaco; Climate; Convergence; Fuzzy set; Leaf functional type; Monte; Optimal PFT set; Sonoran Desert.

**Nomenclature:** Prado (1993).

## Introduction

To describe communities we need a taxonomy to dissect the assemblage of organisms into populations. We take these populations as community components. Pioneer studies in vegetation recognized plant types by morphology and function (see Du Rietz 1931 and references therein). Vegetation science has been essentially based on species, but this view has been criticized (e.g. Grime 1979; Ghiselin 1987). It is clear why we need vegetation descriptions with plant types defined by traits and not only by species. On a large scale, predictions based on plant species are geographically bound (Woodward & Cramer 1996). On a small scale, species are in

some instances so broad and variable that by describing communities by species composition we may not perceive relevant patterns occurring below the resolving power of species (Díaz et al. 1992). The problem has been studied in connection with the IGBP (Steffen et al. 1992), where these types are designated as 'plant functional types' (PFTs).

It is well known that regions of the world with similar climates tend to support structurally similar vegetation (e.g. Naveh 1967). This led to the proposition that community evolution is convergent (e.g. Mooney & Dunn 1970; Barbour & Diaz 1973; Orians & Solbrig 1977a), but suitable analytical techniques were lacking. The existence of PFTs with global validity is particularly grounded on this hypothesis.

There are different approaches to analyse vegetation data based on traits or plant types. A prevalent analytical scheme (e.g. Feoli & Scimone 1984; Díaz et al. 1992; Díaz & Cabido 1997) is the multiplication of the matrix of traits by species by the matrix holding presences/absences or performances of the species in the sampled sites, yielding a matrix of traits by sites. The latter is then analysed by conventional multivariate methods. A limitation of this approach is that the traits must be binary or quantitative. A trait which is qualitative multistate may be coded into as many binary traits as the number of states, but then problems of unequal weighting and non-independence among traits may arise (Feoli & Scimone 1984; Digby & Kempton 1987). PFTs may be groups of species defined by cluster analysis of the matrix of species by traits. We may instead define plant types as a combination of trait states, in which case the classification is monothetic. Orlóci (1991) used the term 'character set type' for trait combinations. He suggested accordingly an analytical scheme in which the traits are considered hierarchically based on a data matrix of character set types by sites (see Orlóci & Orlóci 1985; Orlóci et al. 1986 and the review by Pillar & Orlóci 1993a). The analysis may adopt a fuzzy set approach (Pillar & Orlóci 1991).

The crucial problem, however, is how to select the traits so that the PFTs will likely be 'functional'. The traits must be observable expressions of forms or behaviors

defining plant types that are responsive, in terms of occurrence or performance, to changes in ecosystem conditions. A prevalent approach in the search for PFTs is the selection of a trait set a priori, which is then used for description and interpretations thereafter. Box (1981) for example chooses traits based on perceived form-function relationships. Others use selection criteria that are consistent with theoretical principles (Thompson et al. 1996; Grime et al. 1997) or base the selection on known or suspected trends of trait variation in ecological space (e.g. Díaz et al. 1992; Skarpe 1996). Feasibility and cost of observation are also trait selection criteria (Skarpe 1996; Díaz & Cabido 1997). Traits may also be selected because they show correlation with major components in the ordination space (Thompson et al. 1996; Chapin et al. 1996; Grime et al. 1997), or because they define plant groups that are stable after different analytical techniques and repeatable in different sites or different times (Gitay & Noble 1997). How functional or ecologically relevant are the plant types so derived? This can be answered by comparing derived ecological trends to empirical evidence (Thompson et al. 1996; Skarpe 1996; Chapin et al. 1996; Díaz & Cabido 1997; Gitay et al. 1999; Kleyer 1999). But, in the most recent literature on PFTs, checking these trends is a posteriori and not part of the analytical search for PFTs per se. Furthermore, it is clear at the outset that we can hardly expect to find a set of traits or PFTs that would be optimal for all purposes. The selected traits may, for instance, define PFTs that are good indicators for macro-climatic change but may not be so good for changes in grazing regime (Skarpe 1996).

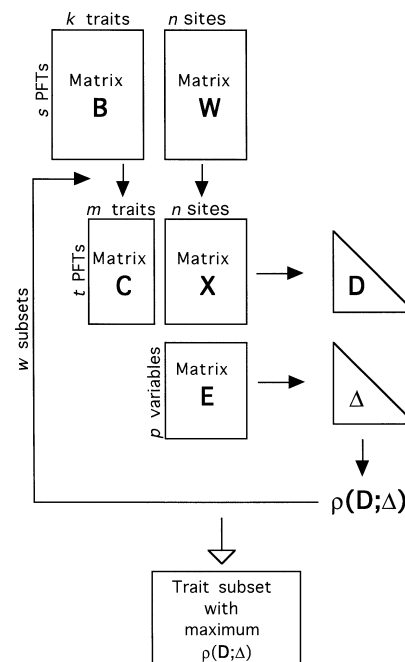
In this paper, I offer an analytical approach for trait selection, which is complementary to the a priori approach. By definition, community descriptions based on PFTs should produce data in which the compositional variation is highly congruent with environmental variation. Consider that we have a data matrix describing plant types by traits that were pre-selected. Consider also that we have two other matrices, one describing the communities by the composition of the plant types and the other the sites by environmental variables. With these three data matrices we can evaluate ecological relevance of the traits and accordingly we can rank the traits. Based on this we can choose an optimal subset of the initial trait set. I discuss in the sequel a general approach for the analysis, describe in some detail algorithms for finding optimal PFTs, and describe examples using data from the Chaco and Monte formations of Argentina and the Sonoran Desert of the United States. The specific methodology has evolved from earlier methods described in Pillar & Orlóci (1991, 1993a, b).

## Methods

### Data acquisition and handling

Community description is accomplished by recording the presence or more often the quantities of plant populations. In the present case, the plant populations are PFTs defined as sets of trait states. A PFT (or any other designation for plant types) is identified as a population containing plants that are similar with regard to the pre-selected traits. This 'population' does not necessarily coincide with a species. But, as a convenience for field description, the plants belonging to the same species in a community may form a PFT if the variability of the traits within that species in that community is acceptable. In this case, even if two or more species have identical descriptions they may be recorded separately. If the intra-species variability for the selected traits is also acceptable across communities, the use of floras or new information on traits (e.g. from Grime et al. 1997) coupled with community descriptions from published records and data bases is allowed.

The description of the PFTs, accordingly, yields a matrix **B** of PFTs by traits. The description of the communities yields a matrix **W** of PFTs by sites holding species presences/absence or quantitative values. The traits may be binary, qualitative multistate, and



**Fig. 1.** A general algorithm for the selection of an optimal subset of traits to maximize congruence  $\rho(\mathbf{D}; \Delta)$  of vegetation and environmental variation. Matrix **B** of PFTs by traits, **W** of PFTs by sites, and **E** of environmental variables by sites result from description. See the main text for explanation.

quantitative with a limited number of classes. If the description of PFTs is achieved as explained in the previous paragraph, and depending on the trait set, identical PFTs may appear in the raw data across and within sites. For the analysis, identical PFTs will have their performances or presences/absences within a site pooled, leaving only distinct PFTs in the rearranged data matrices **B** and **W**. Depending on the performance measuring scale, implicit weighting may result from the pooling of identical PFTs within a site. If performances are densities or cover percentages, there is no differential weighting. But if they are measured e.g. with the usual 1-9 cover-abundance scale (van der Maarel 1979), PFTs formed by dominant species will have less weight than PFTs formed by other species. The weighting is not only a direct consequence of the measuring scale but because scarce species are more numerous (Pillar & Orlóci 1993a: 53). A similar but more extreme effect is involved in pooling presences/absences. Note that depending on the data source and traits, a PFT may include different species, and plants of the same species may belong to different PFTs. The environmental information is summarized by matrix **E** of variables by sites.

#### *Optimization of the trait set*

The optimization algorithm is based on the magnitude of the congruence between vegetational variation and environmental variation. To measure this I use a matrix correlation  $\rho(\mathbf{D}; \Delta)$ , similarly as a standardized Mantel (1967) statistic. **D** is a dissimilarity matrix of sites computed from matrix **X** of PFTs by sites (explained below).  $\Delta$  is a dissimilarity matrix of the same sites based on environmental data (matrix **E**). Any dissimilarity functions are applicable for **D** and  $\Delta$ . I define  $\rho$  as the Pearson's product moment correlation involving  $n(n-1)/2$  off-diagonal pairs of corresponding dissimilarities in **D** and  $\Delta$ . The larger  $\rho(\mathbf{D}; \Delta)$ , the more likely that the PFTs are in fact 'functional' for the environmental factors taken into account.

Consider that a description of communities used  $k$  pre-selected traits defining  $s$  distinct PFTs and that the data are arranged in matrices **B** and **W** as explained. A subset of the pre-selected trait set may suffice to define optimal PFTs. This is the foundation of the optimization algorithm (Fig. 1). A short numerical example is presented in App. 1. Accordingly, matrix **C**, initially with  $s$  PFTs by  $m$  traits, is extracted from matrix **B**. Similarly, matrix **X** with  $s$  PFTs by  $n$  sites is extracted from matrix **W**. Identical PFTs may appear since  $m \leq k$ . Pooling of identical PFTs will leave  $t$  distinct PFTs in matrix **C**, with corresponding quantities in matrix **X**. With these elements and matrix  $\Delta$  the congruence  $\rho(\mathbf{D}; \Delta)$  is evaluated.

Different subsets belonging to the pre-selected set define new matrices **C** and **X** and congruencies  $\rho(\mathbf{D}; \Delta)$ . By recursion, the subset that produces maximum congruence will be found, which will be called the 'optimal traits' subset. Two different recursive algorithms are used. The full algorithm will search all possible subsets with sizes varying from  $m = 1$  to  $k$ . Hence, the number of subsets evaluated is

$$w = \sum_{m=1}^k \frac{k!}{m!(k-m)!} \quad (1)$$

However, this alternative may be too demanding in computations. A stepwise algorithm (as in Pillar & Orlóci 1993b) reduces the computation load, where the number of trait subsets is

$$w = \sum_{m=1}^k (k-m+1) \quad (2)$$

In this alternative the algorithm starts evaluating subsets with size  $m = 1$ , until among the  $k$  traits the one that maximizes  $\rho(\mathbf{D}; \Delta)$  is found. Then evaluation of subsets with  $m = 2$  follows, until finding among the remaining  $k-1$  traits the one that when added to the first maximizes the function. The process continues up to  $m = k$ . It is revealing that for an initial  $k = 16$ , there will be 65 535 different trait subsets to check in the full algorithm and only 136 in the stepwise algorithm.

#### *Fuzzy PFTs*

If the notion of identity is absolute, no matter how similar two PFTs might be, they are not to be pooled unless their records are exactly identical regarding the selected traits. This may aggravate the problem of analytical indeterminacy caused by an excessive number of absences (zeros) in community comparisons (Pillar & Orlóci 1993a: 80). An alternative is offered by the fuzzy approach (Pillar & Orlóci 1991, 1993a). The theory of fuzzy sets was proposed by Zadeh (1965) and has been applied in ecology (e.g. Roberts 1986, 1989; Feoli & Zuccarello 1988; Marsili-Libelli 1989; Banyikwa et al. 1990; Pillar & Orlóci 1991). It is an extension of classical set theory. The belonging of an object to a set is described by a degree of membership. Only if the object belongs entirely to a set (degree = 1) or not at all (degree = 0), the set is called an ordinary (crisp) set. In fuzzy sets the degree of membership is between 0 and 1.

The fuzzy approach will take the partial similarity of PFTs into account (Fig. 2). From matrix  $\mathbf{C}_{t \times m}$  a symmetric matrix  $\mathbf{U}_{t \times t}$  is computed, which contains similarities on the 0 to 1 scale for all pairs of the  $t$  distinct PFTs based on the  $m$  traits. Any similarity measure is acceptable. But, when the traits are of mixed types (e.g. qualitative and

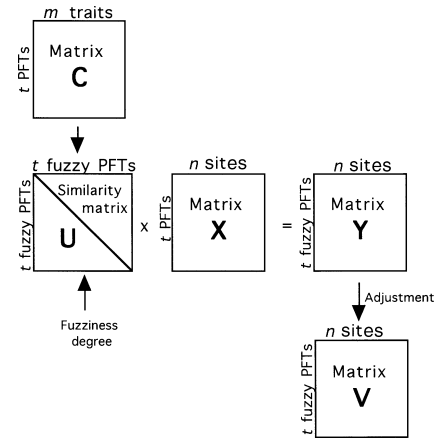
quantitative traits in the same set), indexes proposed by Goodall (1966) and Gower (1971) may be convenient choices. The rows in matrix **U** define fuzzy sets. That is, row  $d$  holds the grades of membership  $u_{de}$  of each PFT  $e$  in the fuzzy set formed by PFTs similar to PFT  $d$ . All PFTs belong to this fuzzy set defined by the properties (trait states) of PFT  $d$ , with a certain grade of membership varying from 0 to 1. Only PFT  $d$  though has  $u_{dd} = 1$  to this fuzzy set since only distinct PFTs originate matrix **U**. The elements in **U** may be adjusted by  $u_{de}^{\xi-1}$ , where  $0 > \xi \leq 1$  expresses a degree of fuzziness chosen by the researcher. If  $\xi$  is one, there will be no change in the fuzzy set. If  $\xi$  approaches zero,  $u_{de}$  will be zero if the similarity  $u_{de} < 1$  and one if  $u_{de} = 1$ , that is, the set will become not-fuzzy (crisp) and **U** will be an identity matrix.

Matrix  $\mathbf{Y}_{\text{tot}}$  is the product **UX**. To keep the totals in the sites unchanged, matrix **Y** is standardized by the site totals in **Y** divided by the corresponding totals in matrix **X**, yielding matrix **V**. Matrix **V** contains the quantities of the  $t$  fuzzy sets in the  $n$  sites. Matrix **V** is then used in the optimization of the trait set as explained (identically as for matrix **X**). In this case the dissimilarity matrix **D** in  $\rho(\mathbf{D}; \Delta)$  will be computed from matrix **V**. Matrix **V** may also be analysed for other purposes in like manner as conventional species by sites tables. Note that when the degree of fuzziness  $\xi$  is set to approach zero, matrix **V** is identical to **X** and there is no real need to compute **U** and **V**.

#### Testing the methods with field data

The stepwise and the full optimization algorithms, both with crisp and fuzzy PFTs, are tried with existing field data from the Chaco and Monte formations of NW Argentina and Sonoran Desert in Arizona (USA). The objective is to find an optimal subset of the pre-selected traits to be used in the description of vegetation-climate trends. The data set contains 19 site descriptions from a precipitation gradient in Argentina and 10 from an elevation gradient in Arizona.

The sites in Argentina are located between latitudes 27° and 28° S in the provinces of Santiago del Estero and Catamarca, nine sites from the dry Chaco, and 10 sites from the Monte desert. Estimated average annual precipitation for each site varies from 168 mm in the Monte to 700 mm in the Chaco. The altitude of the sites ranged from 190 m in the Chaco to 1120 m in the Monte. The climate is characterized by dry and relatively cold winters, with precipitation concentrated in the warm season. Maximum summer temperatures are extremely high in the Chaco. The sampling avoided stands in the initial secondary successional stages and the unit boundaries were defined to have internal homogeneity



**Fig. 2.** A fuzzy set approach for the analysis of community data described by plant types (Pillar & Orlóci 1991, 1993a).

in the stand. The sites sampled in the Chaco were located on a transect from near the city of Santiago del Estero to the west in the direction of the Subandean ranges in Tucumán and Catamarca. The most widespread vegetation type in this part of the Chaco is thorn woodland dominated by *Schinopsis quebracho-colorado* and *Aspidosperma quebracho-blanco* (Prado 1993). The stands were characterized by a first stratum with trees (> 10 m high) of these two species. A second stratum was formed by 2.5–10 m high trees and shrubs such as

and arboreal stem-succulents:

*Cereus validus*      *Opuntia quimilo*      *Stetsonia coryne*

a lower shrub layer comprised among others

*Capparis atamisquea*      *Celtis pallida*      *Larrea divaricata*.

To the west of the Chaco, the Monte is an open xerophytic shrubland formation that extends through western Argentina. The Monte sites are from the Salar de Pipanaco in Catamarca, to the west from the Subandean ranges, which is the same area used in studies of vegetation convergence by Barbour & Diaz (1973) and Orians & Solbrig (1977a). The vegetation in the sampled areas was dominated by shrub species 1–2 m high, mostly

*Larrea cuneifolia*      *Prosopis* spp.

Other species were shrubs, 1–2 m high, such as

*Bulnesia retama*      *Senna aphylla*      *Tricomaria usillo*

and stem-succulents: *Opuntia quimilo*      *Trichocereus* spec.

Sites with extremely high soil salinity in both Chaco and Monte were not considered in the analysis. For additional information on the Argentina sites, see Pillar & Orlóci (1993a), for the complete data set, see Pillar (1992).

The stands in Arizona (L. & M. Orlóci unpubl. 1991) are from a moisture/elevation gradient from typical Sonoran Desert (800 m a.s.l.), near Tucson, at the foothills of the Santa Catalina Mountains, to *Abies* forest at the highest elevation (3000 m a.s.l.). The gradient was

studied earlier by Whittaker & Niering (1964). The same procedure as in Argentina was used to delimit the stands. Average annual precipitation, estimated from normal precipitation maps, varies among the sites from approximately 300 to 960 mm. The stands at the driest parts of the gradient were characterized by 2-5 m tall trees:

*Prosopis juliflora*,  
by shrubs 1-2 m tall: *Larrea tridentata*  
low shrubs: *Celtis pallida* *Calliandra eriophylla*  
*Jatropha cardiophylla* *Hibiscus coulteri* *Encelia farinosa*  
and stem-succulents:

*Carnegiea gigantea* *Mammillaria microcarpa*.  
At intermediate portions stands were typically formed by trees 5-10 m tall: *Pinus cembroides*  
trees/shrubs 1-5 m tall: *Amorpha californica*  
*Arctostaphylos pungens* *Quercus arizonica* *Rhamnus californica* ;  
rosette shrubs:

*Agave palmeri* *Nolina microcarpa* *Yucca schottii*;  
stem-succulents: *Echinocereus triglochidiatus*.

At moist sites of the gradient, 5-10 m tall trees:

*Populus tremuloides* *Robinia neomexicana* *Salix scouleriana*.

In both regions the species were described in the field, for each site, using 14 morphological traits (Table 1, more details given below). The records in Argentina included woody and stem-succulent species, while those in Arizona also included herbs. In the present analysis, the herb component was ignored. The records in Argentina include species cover-abundance values, while the ones in Arizona are species presence/absence. The analysis used cover-abundance values only when the Argentina data set was considered in isolation; otherwise, the analysis used presence/absence records.

The pre-selected traits are easy to measure or record in the field and are known to contain variation associated with ecosystem conditions (Pillar & Orlóci 1993a: 22-28). Some of the pre-selected traits have a counterpart in the traits described by Díaz & Cabido (1997) for species in central-western Argentina. The existence of ecological trends on leaf traits, related to trade-offs of carbon gain and water loss is well documented (Givnish 1987 and references therein). Leaf size tends to decrease and/or leaf thickness to increase on gradients of decreasing rainfall. Leaf life span and leaf texture strategies are related to soil nutrient supply and to seasonal variation in favorable conditions for photosynthesis (Orlans & Solbrig 1977b). Compound leaves favor deciduousness and show increased frequency with aridity. Modified leaf surfaces (glaucous, trichomous) increase leaf reflectance and are more common in sunnier or arid sites. Succulence is a surrogate for the CAM photosynthetic pathway. Thorniness appears to be related to herbivore avoidance. Biological type encompasses a suit of other traits, such as growth form, leaf type, photosynthetic pathway. Plant height reflects carbon immobilization in

**Table 1.** Traits used for PFT-description in plant communities in Argentina and Arizona. Stem refers to a stem-like structure and leaf to a leaflet or leaf-like structure.

Label	Trait name and states
bt	Biological type: 1 = bryoid; 2 = lichen; 3 = pteridophyte; 4 = conifer; 5 = graminoid; 6 = cactoid; 7 = other angiosperm.
st	Stem tissue type: 1 = succulent; 2 = herbaceous; 3 = woody; 4 = no stem.
fu	Stem function: 1 = regular; 2 = twin-purpose; 3 = no stem.
es	Spininess: 1 = thorn/spine; 2 = none; 3 = plant with no stem.
ls	Leaf life span: 1 = aseasonal deciduous; 2 = seasonal deciduous; 3 = withering; 4 = persistent; 5 = plant leafless.
su	Leaf succulence: 1 = succulent; 0 = else.
tx	Leaf texture: 1 to 5: scale from herbaceous to sclerophyllous; 6 = else.
sh	Leaf shape: 1 = scale; 2 = filiform/needle; 3 = other; 4 = plant leafless.
ar	Leaf arrangement: 1 = simple; 2 = compound; 3 = plant leafless.
ed	Leaf dorsal hairiness: 1 = glabrous; 2 = glaucous; 3 = trichomous sparse; 4 = trichomous dense; 5 = plant leafless.
ev	Leaf ventral hairiness: 1 = glabrous; 2 = glaucous; 3 = trichomous sparse; 4 = trichomous dense; 5 = plant leafless.
si	Leaf size: index computed as $\log(\text{leaf width} \times \text{leaf length})$ rescaled from 1 - 5.
th	Leaf thickness: 1 = < 1 mm; 2 = 1 - 3; 3 = 3 - 5; 4 = > 5; 5 = leafless.
he	Plant height: 1 = < 5 cm; 2 = 5 - 25; 3 = 25 - 75; 4 = 75 - 125; 5 = 125 - 250; 6 = 250 - 500; 7 = 500 - 1000; 8 = > 1000.

support tissue, which may be an important indicator of soil water and nutrient supply (Tilman 1988). Plant height was estimated as the height at which most of the photosynthetic biomass was concentrated.

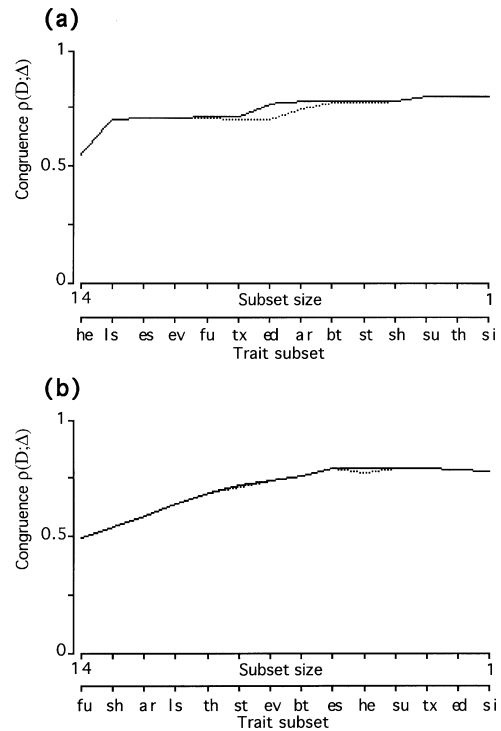
The two data sets shared only one species among 215 (*Celtis pallida*) and 13 genera among 98: *Acacia* *Atriplex* *Baccharis* *Celtis* *Cercidium* *Cereus* *Condalia* *Jatropha* *Larrea* *Lycium* *Mimosa* *Opuntia* *Prosopis*. The sites in Argentina recorded 58 species and in Arizona 158 non-herbaceous species. A preliminary inspection of the data revealed that, despite the great difference between Arizona and Chaco/Monte data sets in terms of species richness, the PFTs of one region in general tended to have counterparts in PFTs of the other region, except for conifers, occurring only in the Arizona data.

The optimization analysis followed a similar approach as in Pillar & Orlóci (1993a: 222-235). In this, PFTs defined by the same trait set are evaluated for response to the same factor in two regions with different floras. That is, first, optimal trait subsets for precipitation were found for the data sets of Argentina and Arizona separately, with crisp and fuzzy PFTs and using the sequential and the full algorithm. Second, the optimal trait subset for each one data set was validated using the other data set in isolation or jointly in terms of congruence  $\rho(\mathbf{D}; \Delta)$  under the same fuzziness level. The analyses were performed with the software SYNCSA (Pillar 1998) for automatic pooling of PFTs, fuzzy set adjustments, resemblance measures, ranking of traits and ordination. An early version of this software is described by Pillar & Orlóci (1993a).

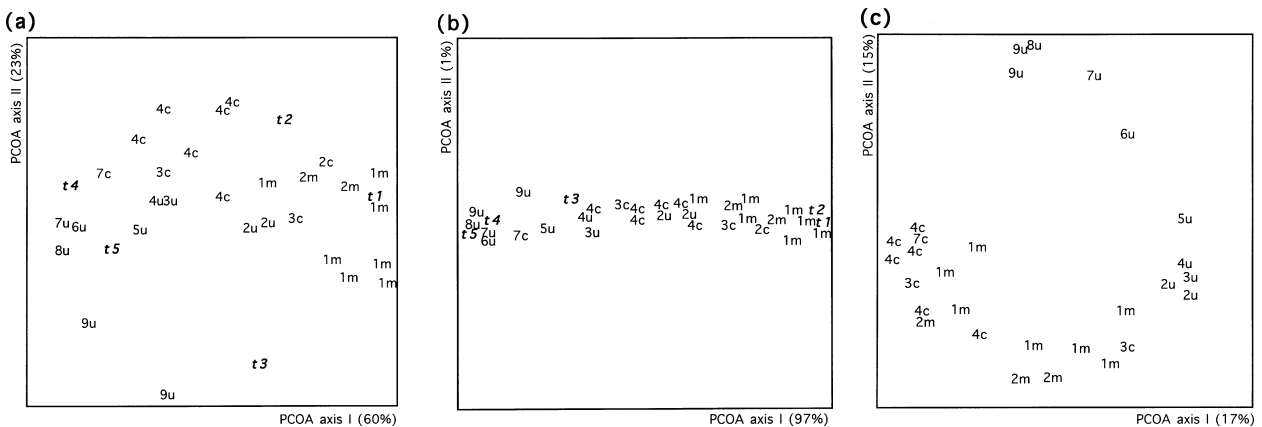
**Results**

The analysis of the Chaco/Monte data (Fig. 3) found optimal subsets of traits for response to the rainfall gradient. The optimal subset with crisp PFTs, in both stepwise and full algorithms, was formed by ‘leaf size’ only:  $\rho(\mathbf{D}; \Delta) = 0.806$ . This allowed the definition of five PFTs. It is important to note that PFTs defined using the complete set of 14 traits revealed a much lower environmental congruence value:  $\rho(\mathbf{D}; \Delta) = 0.545$ , (Table 2). For validation, this optimal trait subset was used to define the PFTs in the analysis of the Arizona stands, which resulted in a close congruence with precipitation variation:  $\rho(\mathbf{D}; \Delta) = 0.788$ . A joint analysis of the stands from Argentina and Arizona, described by similarly defined PFTs, also revealed a close congruence with precipitation variation:  $\rho(\mathbf{D}; \Delta) = 0.796$ . Furthermore, the stands analysed jointly showed a clear, though non-linear vegetation pattern in ordination space (Fig. 4a). Also, the PFTs on the same ordination space are separated between smaller and larger leaf-size classes. An ordination with fuzzy PFTs portrays a more linear trend, from drier sites, characterized by small leaf sizes, to moister sites characterized by larger leaf sizes (Fig. 4b). Fig. 4c depicts the joint ordination of sites from both areas described by crisp PFTs defined by the suboptimal, complete set of pre-selected traits. The optimal subset of traits for fuzzy PFTs (Fig. 3b) contains ‘leaf size’, ‘leaf dorsal hairiness’, ‘leaf texture’ and ‘leaf succulence’. The analysis revealed a slightly lower congruence in validation than the one for crisp PFTs (Table 2).

The sequence of optimization was then reversed, by finding an optimal trait subset for the Arizona data, which was in turn used in similar analyses as above involving the Chaco/Monte sites. See Table 2. Although the congruence in the Arizona data set was very high, it



**Fig. 3.** Profile of maximum congruence  $\rho(\mathbf{D}; \Delta)$  at different trait subset sizes, with crisp (a) and fuzzy (b) PFTs. The data include 19 sites from the Chaco and Monte.  $\Delta$  is based on precipitation data. Maximum congruence found by the full algorithm is at the solid line. The dotted line indicates maximum congruence found by the stepwise algorithm. Vegetation composition is defined by plant types based on traits, which for the stepwise algorithm are shown on the horizontal axis cumulatively from right to left (see labels in Table 1). Note that when subset size = 14, the complete trait set is used. When the lines meet the trait subset is identical in both algorithms. Data from Pillar (1992).



**Fig. 4.** Principal coordinates analysis based on chord distances for 29 sites from Chaco (c), Monte (m) and Arizona (u). Numbers identify the average annual precipitation at the sites (scale 1-9). **a, b.** sites described by five PFTs based on ‘leaf size’. PFTs t1-t5, in italics, from smaller to larger leaf sizes, plotted by their correlation (rescaled) with the axes. The PFTs in **a.** are crisp, in **b.** fuzzy. **c.** sites described by 223 crisp PFTs defined by all 14 traits. Data from Pillar (1992) and L. & M. Orlóci (unpubl.).

**Table 2.** Summary of optimization results, by the stepwise algorithm, with different data sets and degrees of fuzziness, DF, (crisp, C:  $\xi$  approaching 0, fuzzy, F:  $\xi = 1$ ).  $\rho(\mathbf{D}; \Delta)$  is the congruence of vegetation and precipitation variation considering all pre-selected traits or only the optimal trait subset. The ‘optimal trait subset’ was found by the optimization algorithm aiming at maximal congruence  $\rho(\mathbf{D}; \Delta)$ . The elements in  $\mathbf{D}$  are chord distances of sites and in  $\Delta$  the absolute differences of annual precipitation between the sites. ‘ $\rho(\mathbf{D}; \Delta)$  in validation’ is the congruence, based on PFTs defined by the optimal trait subset, in the analysis (a) with the other data set (Arizona or Chaco/Monte) or (b) with Chaco/Monte/Arizona jointly, to which ‘Number of PFTs’ refers. Results of the full algorithm were identical to the ones presented, except in the cases in which they are indicated in parentheses.

Data set	DF	$\rho(\mathbf{D}; \Delta)$ All 14 traits	Opt. trait subset	$\rho(\mathbf{D}; \Delta)$ Opt. traits	$\rho(\mathbf{D}; \Delta)$ in validation		No. of PFTs	
					a	b	a	b
Chaco/M	C	0.545	si	0.806	0.788	0.796	5	5
	F	0.495	si ed tx su	0.793	0.751	0.672	52	64
Arizona	C	0.933	he ed st (he ed at sh)	0.980	0.470	0.601	35	46
	F	0.845	he ls tx su	0.945	0.422	0.539	40	76
Joint sets	C	0.498	si sh ar su th	0.873	-	-	-	31
	F	0.637	si he	0.860	-	-	-	33

dropped to much lower values when validated with the Chaco/Monte data set in isolation or with the two data sets jointly. That is to say, PFTs that were optimal for the Arizona analysis were not so for the Chaco/Monte data set. It is important to note that the optimization had very little advantage for the Arizona data set, since the congruence  $\rho(\mathbf{D}; \Delta)$  with the complete set of pre-selected traits was already high, with either crisp or fuzzy PFTs.

The analysis also included optimization involving all 29 Chaco/Monte/Arizona stands. The optimal trait subsets for crisp and fuzzy PFTs contain traits that were optimal for the Arizona and Chaco/Monte data sets in isolation and accordingly produced intermediate levels of congruence (Table 2). The congruence  $\rho(\mathbf{D}; \Delta)$  when using the complete set of pre-selected traits was suboptimal (Fig. 4c), but greater with fuzzy than with crisp PFTs.

## Discussion

The results suggest that the technique works. There seems to be no advantage in using the more computationally demanding full optimization algorithm instead of the stepwise algorithm. Both found nearly identical optimal PFTs. When solutions diverged it was at suboptimal subsets or when the alternative subsets compared at a given subset size had very close congruence values.

The results of optimization with crisp and fuzzy plant types differed. Whether to build or not the notion

**Table 3.** Congruence  $\rho(\mathbf{D}; \Delta)$  for one-trait defined crisp PFTs, with Arizona and Chaco/Monte data sets. Other definitions are in the caption of Table 2.

Trait	Congruence $\rho(\mathbf{D}; \Delta)$			
	Arizona		Chaco/Monte	
bt	0.856	-0.101	st	0.722
fu	0.904	0.216	es	0.727
ls	0.852	0.242	su	0.259
tx	0.838	0.538	sh	0.751
ar	0.724	0.313	ed	0.793
ev	0.844	0.664	si	0.788
th	0.855	0.393	he	0.955

of fuzziness into the analysis is a decision analogous to other decisions with respect to data analysis, as usual with data transformations. Either with crisp PFTs or fuzzy PFTs a high congruence is indicative of linear correlation between *community* and environmental dissimilarities. With crisp sets, the absence of a PFT in a stand is absolute, while in the fuzzy approach absence is a matter of degree and zeros may be nearly eliminated from the data table (Pillar & Orlóci 1991). As a high degree of analytical indeterminacy is linked to non-linearity, the fuzzy approach may reduce non-linearity in the data. However, an analysis with fuzzy PFTs conveys to the community data (by matrix multiplication) the similarities of the PFTs based on the traits. Consequently, with fuzzy PFTs, high congruence  $\rho(\mathbf{D}; \Delta)$  and association of *traits* to environmental factors are necessarily linked, while these may not be so with crisp PFTs. We do not know whether the use of other measures of congruence between community and environmental variation in the optimization algorithm, e.g. a Procrustean statistic (Schönemann & Carroll 1970; Jackson 1995), could better handle non-linear responses.

Associated traits increase redundancy in the matrix describing the PFTs by traits. The elimination of one of two associated traits will cause little change in the resulting matrix  $\mathbf{X}$  of PFTs by sites and in the congruence  $\rho(\mathbf{D}; \Delta)$ ; the elimination of a trait that is nearly constant across PFTs will have a similar effect. The optimal trait subset with fuzzy PFTs for the Chaco/Monte data set (Table 2) contains some degree of redundancy. This is shown by the low increase in congruence when the traits ‘leaf dorsal hairiness’, ‘leaf texture’ and ‘leaf succulence’ are added to the subset containing only ‘leaf size’ (Fig. 3b). Therefore, in this case there is little advantage in defining PFTs by the three traits in addition to ‘leaf size’. The optimal trait subsets revealed by the Arizona analysis also contain redundancies. In this case, there was, e.g. a tiny increase (from 0.976 to 0.980) in the congruence when the trait ‘stem tissue type’ was included in the subset with ‘plant height’ and ‘leaf dorsal hairiness’. In fact, due to redundancy among traits in this data set, different subsets of traits were very

close in terms of congruence, which warns against neglecting suboptimal traits under these circumstances. This also explains discrepancies between the results of full and stepwise algorithms.

Redundancy among traits in the Arizona analysis is also evident from the small differences in congruence comparing the optimal and the complete set of traits (Table 2), and comparing the different traits in isolation (Table 3). Redundancy may be used as a criterion for further optimization of the trait subset, to achieve a parsimonious representation of the data. An algorithm is described by Pillar & Orlóci (1993a).

The optimal traits subsets for the Chaco/Monte gradient revealed a high congruence  $\rho$  ( $\mathbf{D}$ ;  $\Delta$ ) when validated with the communities in Arizona, but the reverse was not so. All traits were nearly equally relevant defining community patterns highly associated to precipitation for the Arizona data set, but only some of the traits were good for the Chaco/Monte data set. Especially the congruence for PFTs defined by 'plant height' was much greater for the Arizona data set. It is not clear whether the fact that the gradient in Arizona encompasses much greater species richness would explain these results, for most PFTs of one region in general tended to have counterparts in PFTs of the other region. This may explain why 'biological type', which is a complex of other traits, was not relevant for the Chaco/Monte data set. It is clear, however, that the approach involving optimization followed by cross validation is capable of seeking traits that are optimal in more than one floristic region. This is the case of 'leaf size', which determined high congruence with both data sets. Incidentally, the relation between 'leaf size' and rainfall gradients is well known (e.g. Givnish 1987).

The results highlight the potential for analytical optimization in PFT-based vegetation analysis. The results clearly show that a selection of traits from an initial set, may improve the ecological relevance of the PFTs. However, *a priori* selection of traits is critical, because the optimization will be useless if the traits are irrelevant at the scale of the study and intended objectives. In this regard, analytical optimization is an important complementary tool. Whether PFTs are really functional should not be a question of definition, but something that can be decided by measurement in a given data set.

Due to phylogenetic constraints (Givnish 1987), organisms tend to evolve towards possessing common sets of attributes. It is not difficult to find plant traits that maximize the similarity of phylogenetically distant communities. We should rather examine the convergence of vegetation responses to similar environmental gradients in phylogenetically distant floristic regions, e.g. the semi-arid areas of South and North America included here. The analysis found optimal traits that captured

vegetation variation maximally congruent with regional precipitation gradients in Argentina and Arizona. Thus, an analysis involving communities on climatic or disturbance gradients from different regions is a good framework to find optimal PFTs. These PFTs will be more responsive and likely predictive to the environmental gradient. If under the foregoing conditions the assumption of phylogenetic independence is acceptable, the question of community convergence under similar environmental gradients can be objectively evaluated.

Some attempts to seek PFTs by data analysis (i.a., Thompson et al. 1996; Chapin et al. 1996; Grime et al. 1997; Gitay & Noble 1997) have been based on the data matrix of taxa by traits only, but community level- and environmental data were not explicitly considered in the analysis seeking PFTs, as in the present paper. Gitay & Noble (1997) defined PFTs as plant groups that are stable after different analytical techniques, with different traits, and repeatable at different sites or different times. However, ecologically relevant PFTs may not necessarily comply with these conditions. PFTs may be fuzzy in terms of the defining traits but show consistent responses to changes in ecosystem conditions in different floristic regions. Furthermore, the fact that the groups are repeatable at different sites is not sufficient to conclude that they form similar community patterns at the different sites. Repeatability is better defined in the same way as I defined community convergence, that is, the consistency of community trends or responses, in terms of PFTs, to similar gradients in different sites.

**Acknowledgements.** Grants and fellowships from CNPQ, CAPES and FAPERGS (Brazil) supported the research. I recognize the important contribution of L. Orlóci in the development of the ideas presented here, and thank him and M. Orlóci for permission to use their data. I am very grateful to J.P. Lewis for invaluable cooperation in data collection in Argentina. S. Lavorel, S. Hartz, S. Díaz and an anonymous reviewer gave many helpful suggestions to improve the paper.

## References

- Banyikwa, F.F., Feoli, E. & Zuccarello, V. 1990. Fuzzy set ordination and classification of Serengeti short grasslands, Tanzania. *J. Veg. Sci.* 1: 97-104.
- Barbour, M.G. & Diaz, D.V. 1973. Larrea plant communities on bajada and moisture gradients in the United States and Argentina. *Vegetatio* 28: 335-352.
- Box, E.O. 1981. *Macroclimate and plant forms*. Junk, The Hague.
- Cabrera, A.L. 1971. Fitogeografía de la República Argentina. *Bol. Soc. Argent. Bot.* 14: 1-42.
- Chapin, F.S., III, Bret-Harte, M.S., Hobbie, S.E. & Zhong, H. 1996. Plant functional types as predictors of transient responses of arctic vegetation to global change. *J. Veg. Sci.* 7: 347-358.



- Díaz, S. & Cabido, M. 1997. Plant functional types and ecosystem function in relation to global change. *J. Veg. Sci.* 8: 463-474.
- Díaz, S., Acosta, A. & Cabido, M. 1992. Morphological analysis of herbaceous communities under different grazing regimes. *J. Veg. Sci.* 3: 689-696.
- Díaz, S., Cabido, M. & Casanoves, F. 1998. Plant functional traits and environmental filters at a regional scale. *J. Veg. Sci.* 9: 113-122.
- Digby, P.G.N. & Kempton, R.A. 1987. *Multivariate analysis of ecological communities*. Chapman & Hall, London.
- Du Rietz, G.E. 1931. Life-forms of terrestrial flowering plants. *Acta Phytogeogr. Suec.* 3: 1-95.
- Feoli, E. & Scimone, M. 1984. A quantitative view of textural analysis of vegetation and examples of application of some methods. *Arch. Bot. Biogeogr. Ital.* 60: 72-94.
- Feoli, E. & Zuccarello, V. 1988. Syntaxonomy: a source of useful fuzzy sets for environmental analysis? *Coenoses* 3: 141-147.
- Ghiselin, M.T. 1987. Species concepts, individuality, and objectivity. *Biol. Philos.* 2: 127-143.
- Gitay, H. & Noble, I.R. 1997. What are functional types and how should we seek them? In: Smith, T.M., Shugart, H.H. & Woodward, F.I. (eds.) *Plant Functional Types: their relevance to ecosystem properties and global change*, pp. 3-19. Cambridge University Press, Cambridge.
- Gitay, H., Noble, I.R. & Connell, J.H. 1999. Approaches to defining functional types in a rainforest. *J. Veg. Sci.* 10: 641-650. (This issue.)
- Givnish, T.J. 1987. Comparative studies of leaf form: assessing the relative roles of selective pressures and phylogenetic constraints. *New Phytol.* 106 (Suppl.): 131-160.
- Goodall, D.W. 1966. A new similarity index based on probability. *Biometrics* 22: 882-907.
- Gower, J.C. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27: 857-871.
- Grime, J.P. 1979. *Plant strategies and vegetation processes*. John Wiley & Sons, New York, NY.
- Grime, J.P. et al. 1997. Integrated screening validates primary axes of specialization in plants. *Oikos* 79: 259-281.
- Jackson, D.A. 1995. PROTEST - A Procrustean randomization test of community environment concordance. *Ecoscience* 2: 297-303.
- Kleyer, M. 1999. The distribution of plant functional types on gradients of disturbance intensity and resource supply in an agricultural landscape. *J. Veg. Sci.* 10: 697-708. (This issue.)
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209-220.
- Mooney, H.A. & Dunn, E.L. 1970. Convergent evolution in mediterranean-climate evergreen sclerophyll shrubs. *Evolution* 24: 292-303.
- Naveh, Z. 1967. Mediterranean ecosystems and vegetation types in California and Israel. *Ecology* 48: 445-459.
- Orians, G.H. & Solbrig, O.T. (eds.) 1977a. *Convergent evolution in warm deserts*. Dowden, Hutchinson and Ross, Stroudsburg, PA.
- Orians, G.H. & Solbrig, O.T. 1977b. A cost-income model of leaves and roots with special reference to arid and semiarid areas. *Amer. Nat.* 111: 677-690.
- Orlói, L. 1991. On character-based community analysis: choice, arrangement, comparison. *Coenoses* 6: 103-107.
- Orlói, L., Feoli, E., Lausi, D. & Nimis, P. 1986. Estimation of character structure convergence (divergence) in plant communities; a nested hierarchical model. *Coenoses* 1: 11-20.
- Orlói, L. & Orlói, M. 1985. Comparison of communities without the use of species: model and example. *Ann. Bot. (Roma)* 43: 275-285.
- Pillar, V.D. 1992. *The theory of character-based community analysis*. Ph.D. Thesis, University of Western Ontario, London.
- Pillar, V.D. 1998. *SYNCSA software for character-based community analysis*. Dep. de Ecologia, UFRGS, Porto Alegre.
- Pillar, V.D. & Orlói, L. 1991. Fuzzy components in community level comparisons. In: Feoli, E. & Orlói, L. (eds.) *Computer assisted vegetation analysis*, pp. 87-93. Kluwer, Dordrecht.
- Pillar, V.D. & Orlói, L. 1993a. *Character-based community analysis: theory and application program*. SPB Academic Publishing, The Hague.
- Pillar, V.D. & Orlói, L. 1993b. Taxonomy and perception in vegetation analysis. *Coenoses* 8: 53-66.
- Prado, D.E. 1993. What is the Gran Chaco vegetation in South America? I. A review. Contribution to the study of flora and vegetation of the Chaco. *V. Candollea* 48: 145-172.
- Roberts, D.W. 1986. Ordination on the basis of fuzzy set theory. *Vegetatio* 66: 123-131.
- Roberts, D.W. 1989. Fuzzy systems vegetation theory. *Vegetatio* 83: 71-80.
- Schönemann, P.H. & Carroll, R.M. 1970. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika* 35: 245-256.
- Skarpe, C. 1996. Plant functional types and climate in a southern African savanna. *J. Veg. Sci.* 7: 397-404.
- Steffen, W.L., Walker, B.H., Ingram, J.S.I. & Koch, G.W. 1992. *Global change and terrestrial ecosystems; the operational plan*. International Geosphere-Biosphere Programme. IGBP Report, No. 21, Stockholm.
- Thompson, K., Hillier, S.H., Grime, J.P., Bossard, C.C. & Band, S.R. 1996. A functional analysis of a limestone grassland community. *J. Veg. Sci.* 7: 371-380.
- Tilman, D. 1988. *Plant strategies and the dynamics and structure of plant communities*. Princeton University Press, Princeton, NJ.
- van der Maarel, E. 1979. Transformation of cover-abundance values in phytosociology and its effects on community similarity. *Vegetatio* 39: 97-114.
- Whittaker, R.H. & Niering, W.A. 1964. Vegetation of the Santa Catalina Mountains, Arizona. I. Ecological classification and distribution of species. *J. Ariz. Acad. Sci.* 3: 9-34.
- Woodward, F.I. & Cramer, W. 1996. Plant functional types and climatic changes: Introduction. *J. Veg. Sci.* 7: 306-308.
- Zadeh, L.A. 1965. Fuzzy sets. *Inform. Control* 8: 338-353.

Received 13 October 1998;  
Revision received 3 May 1999;  
Accepted 1 June 1999.

For App. 1, see next page

**App. 1.** Numerical example illustrating steps of the PFT-optimization algorithm. Data set with three sites using PFTs defined by pre-selected traits: leaf size, *si*; leaf life span, *ls*; plant height, *he* (Table 1) and precipitation as environmental variable. See main text.

Precipitation in mm: R1 = 168; R2 = 360; R3 = 700.  
 Environmental distance matrix  $\Delta$  (squared Euclidean distances):

0	192	532
	0	340
		0

Vegetation data with 5 PFTs defined by the pre-selected traits (matrix **B**), with corresponding quantities in the sites (matrix **W**):

PFT	si	ls	he	R1	R2	R3
1	1	5	3	2	4	0
2	2	2	6	0	5	3
3	4	4	6	0	0	3
4	4	2	5	0	2	0
5	3	4	4	5	0	0

Then, different trait subsets extracted from the data above are used to define PFTs and examined for congruence  $\rho(\mathbf{D}; \Delta)$ . It starts with one-trait subsets:

1 *Trait subset is formed by leaf size (si):*

Vegetation data (matrices **C** and **X**) extracted from matrices **B** and **W**, with PFTs defined by leaf size. Note that PFTs 3 and 4 were pooled since they are identical for leaf size.

PFT	si	R1	R2	R3
1	1	2	4	0
2	2	0	5	3
3	4	0	2	3
4	3	5	0	0

PFT similarity matrix **U** (Gower index) based on leaf size; a quantitative trait, hence similarity of, e.g. PFTs 2 and 3 is  $[1 - |2 - 4| + (4 - 1)] = 0.333333$ , where  $(4 - 1) =$  leaf size range.

1	0.666667	0	0.333333
0.666667	1	0.333333	0.666667
0	0.333333	1	0.666667
0.333333	0.666667	0.666667	1

Matrix product  $\mathbf{Y} = \mathbf{UX}$ . In this, e.g. quantity of PFT 2 in R1 =  $(0.666667 \times 2 + 1 \times 0 + 0.333333 \times 0 + 0.666667 \times 5) = 4.666669$ .

	R1	R2	R3
	3.666665	7.333335	2.000001
	4.666669	8.333334	3.999999
	3.333335	3.666665	3.999999
	5.666666	6.000001	4.000002
Site total	17.33334	25.33334	14.000001

Fuzzy PFTs vegetation matrix **V** after adjustment by site totals; quantity of e.g. PFT 2 in R1, adjusted by totals in matrix **X** and **Y**:  $4.666669 \times 7 \div 17.33334 = 1.88462$ .

PFT	si	R1	R2	R3
1	1	1.48077	3.18421	0.857143
2	2	1.88462	3.61842	1.71429
3	4	1.34615	1.59211	1.71429
4	3	2.28846	2.60526	1.71429

Vegetational distance matrix **D** (squared chord distances) of sites based on matrix **V**:

0	0.0753016	0.0584932
	0	0.170737
		0

If we take crisp PFTs instead, matrix **D** is computed from **X** directly:

0	1.55709	2
	0	0.52427
		0

Congruence  $\rho(\mathbf{D}; \Delta)$  with PFTs defined by 'si': Fuzzy: -0.212227; Crisp: 0.36288.

2. *Trait subset is formed by leaf life span (ls):*

Vegetation data (matrices **C** and **X**) extracted from matrices **B** and **W**, with PFTs defined by leaf life span. Note that PFTs 2 and 4, and PFTs 3 and 5 in the original table were pooled since they are identical for leaf life span:

PFT	ls	R1	R2	R3
1	5	2	4	0
2	2	0	7	3
3	4	5	0	3

The PFT similarity matrix **U** based on leaf life span. The only trait is qualitative, thus there are no identical PFTs and all pair-wise similarities are zero.

1	0	0
0	1	0
0	0	1

Since **U** is an identity matrix, matrices **Y** and **V** are identical with matrix **X**.

Matrix **D** (squared chord distances) based on fuzzy (or crisp) PFTs:

0	1.63148	0.686936
	0	0.772119
		0

Congruence with PFTs defined by 'ls': Fuzzy: -0.86952; Crisp: -0.86952.

3. *Trait subset is formed by plant height (he):*

Similar steps as in (1), matrix **C** and **B** extracted and congruence computed accordingly.

Congruence with PFTs defined by 'he': Fuzzy: 0.930836; Crisp: 0.359799.

The trait that, when used alone, is maximizing congruence is plant height, if using fuzzy PFTs, and leaf size, if using crisp PFTs.

The next steps evaluate subsets of two traits. The full algorithm evaluates all combinations of two traits, the stepwise evaluates some of the subsets

4<sup>1</sup>. *Trait subset is formed by leaf size (si) and leaf life span (ls):*

Vegetation data (matrices **C** and **X**) extracted from matrices **B** and **W**.

PFT	si	ls	R1	R2	R3
1	1	5	2	4	0
2	2	2	0	5	3
3	4	4	0	0	3
4	4	2	0	2	0
5	3	4	5	0	0

PFT-similarity matrix **U** based on leaf size and leaf life span. The similarity of, e.g. PFTs 2 and 3 is  $\{[1 - |2 - 4| + (4 - 1)] + 0\} \div 3 = (0.333333 + 0) \div 2 = 0.166667$ , where 0.333333 is the similarity for leaf size and zero is the similarity for leaf life span (the PFTs do not coincide for leaf life span). The Gower similarity is the average of the partial indexes.

1	0.333333	0	0	0.166667
0.333333	1	0.166667	0.666667	0.333333
0	0.166667	1	0.5	0.833333
0	0.666667	0.5	1	0.333333
0.166667	0.333333	0.833333	0.333333	1

Fuzzy PFT-vegetation matrix **V** after computation of matrix  $\mathbf{Y} = \mathbf{UX}$ , adjustment by site totals.

PFT	si	ls	R1	R2	R3
1	1	5	1.21429	2.65248	0.4
2	2	2	1	3.58865	1.4
3	4	4	1.78571	0.858156	1.4
4	4	2	0.714286	2.49645	1.4
5	3	4	2.28571	1.40426	1.4

Vegetation distance matrix **D** (squared chord distances) of sites based on matrix **V**:

0	3.39575	0.556633
	0	2.83993
		0

If the analysis takes crisp PFTs instead, matrix **D** is computed from **X** directly:

0	1.55709	2
	0	0.945907
		0

Congruence with PFTs defined by 'si', 'ls': Fuzzy: -0.974323; Crisp: 0.484917.

5<sup>1,2</sup>. *Trait subset is formed by leaf size (si) and plant height (he):*

Similar steps are performed as before...

Congruence with PFTs defined by these traits: Fuzzy: 0.976425; Crisp: 0.484917.

6<sup>2</sup>. *Trait subset is formed by plant height (he) and leaf life span (ls):*

Similar steps are performed as before.

Congruence with PFTs defined by these traits: Fuzzy: 0.225571; Crisp: 0.484917.

The subset with two traits that is maximizing congruence is formed by plant height and leaf size if using fuzzy PFTs and any of the subsets if using crisp PFTs. The same is concluded by using the stepwise algorithm.

7<sup>1,2</sup>. *Trait subset is formed by all pre-selected traits:*

Similar steps are performed as before, but in this case matrices **C** and **X** are obviously identical with matrices **B** and **W**.

Congruence with PFTs defined by the three traits: Fuzzy: 0.455332; Crisp: 0.484917.

In conclusion, for this data set, congruence  $\rho(\mathbf{D}; \Delta)$  is maximized by using PFTs defined by leaf size and plant height, in both the fuzzy and crisp approaches, and with full or stepwise algorithm.

<sup>1</sup>crisp PFTs; <sup>2</sup>fuzzy PFTs.