

In Search of Potential Ecological Communities

Sergio Camiz

Dipartimento di Matematica, Sapienza Università di Roma

E-mail: sergio.camiz@uniroma1.it

Valério D. Pillar

Departamento de Ecologia, Universidade Federal do Rio Grande do Sul

E-mail: vpillar@ufrgs.br

Abstract

A problem in plant community ecology is to define the potential community expected to be found in a given sampling unit. For the same reason, the distinction between potential and actual community allows to analyze the difference between actual and potential community, whose structure deserves being studied. With this objective in mind, we thought reasonable to ground our study on the plant traits, that are characteristics of species that might influence their distribution, and we define a methodology able to estimate potential communities based on the observed species co-occurrences and on the species trait similarities. The so defined potential communities may be compared by means of different measures, such as sums of element to element deviations, the RV coefficient and Procrustes correlations, which may be tested against proper null models relating species traits and species abundances in communities. The tools involved estimating potential communities by using (1) Beals smoothing of the species \times communities data table and (2) a fuzzy weighting of the same community data based on traits similarity among species. As an example, we apply the method to plant community data from grassland vegetation, considering species cover recorded at two scales of observation (0.2×0.2 m plots, and 0.2×1.0 m plots). The species were described by a set of 12 morphological traits. With both small and large plots we found a subset of traits that maximized the RV coefficient between potential metacommunity estimated by smoothing and by traits, which in both cases was significant by permutation test. When all traits were used instead, the RV was not significant in any case. The results with this example suggest that the selected traits are critical in the assembly of the studied plant communities and that these traits are good proxies of plant adaptations to the prevailing environmental conditions and biotic interactions.

Keywords: Potential Communities, Species traits, Beal's smoothing, Fuzzy sets

1 Introduction

In the framework of plant community studies, we want to study the issue of *potential community* as the kind of community one may imagine to find somewhere, should not random events occur that may favor a species in respect to another. In particular, we aim at identifying potential communities, based on observed ones, and try to interpret their structure, based on the species traits. This allows to study also the deviations of the observed communities from the potential ones, and investigate the relations, if any, that may exist among them and the environmental factors. The study may be performed at two different levels, say considering a community individually, or considering a metacommunity, that is a set of communities, as a whole.

As usual, we start with a metacommunity, that is a sample represented by a data table A , in which n rows represent species and p columns represent communities. Each cell entry

$a_{ij}, i = 1, \dots, n, j = 1, \dots, p$ is a measure of abundance of the species i in the community j . Note that each column of A , $\mathbf{A}^j, j = 1, \dots, p$ is a (hopefully random) vector extracted by the population under study: if the population from which the sample A was extracted is homogeneous, all \mathbf{A}^j s are realizations of the same random vector $\boldsymbol{\alpha}$, whose n components are the absolute frequencies of the species present in the population; thus, we may estimate $\boldsymbol{\alpha}$ through the average of the \mathbf{A}^j s, that is $\hat{\boldsymbol{\alpha}} = \frac{1}{p} \sum_{j=1}^p \mathbf{A}^j$.

In the present study, we are not really interested in species abundance, but only in composition of communities. Thus, we transform A to a matrix W of column *profiles*, that is proportions of abundance of each species in the specific community. We obtain profiles by dividing each entry $a_{ij} \in A$ by the total abundance of the j -th community, that is by the j -th column total: $W = \{w_{ij} = a_{ij}/a_{.j}\}$, where the dot indicates summation on the corresponding index. Now, again, each column of W , \mathbf{W}^j may be an independent realization of the population profile $\boldsymbol{\varphi}$, this too estimated by $\hat{\boldsymbol{\varphi}} = \frac{1}{p} \sum_{j=1}^p \frac{\mathbf{A}^j}{a_{.j}} = \frac{1}{p} \sum_{j=1}^p \mathbf{W}^j$. In the following, the matrix A may be a presence/absence one: should it be the case, W values would equal to the inverse of the number of species present in each community.

A difference must be observed in the estimation of either $\boldsymbol{\alpha}$ or $\boldsymbol{\varphi}$: whereas the n species abundances in each \mathbf{A}^j are totally free, in \mathbf{W}^j an entry is constrained by the others, in that they all sum to 1, so that only $n - 1$ are free. We express this by saying that the estimation of $\boldsymbol{\alpha}$ has n *degrees of freedom* (df), whereas for the estimation of $\boldsymbol{\varphi}$ $df = n - 1$: this will be important in testing for significance.

2 Beals smoothing

The issue of potential community seems present in plant community studies since long, in the past discussed within the problems raised by the presence of zeros in a data table. As such, it was already considered in the 1960's: "where a species does not occur, its absence may be a matter of chance only, the habitat being favorable for its occurrence. Alternatively, site conditions may be entirely unfavorable for the occurrence of the species. Since each unrecorded species is equally absent in the data, the zeros hide this part of the dialogue between species and habitat" (Swan 1970, quoting Lambert and Dale 1964). In order to adjust his models to real data, Swan (1970) proposes to substitute zeros according to an average association of the species missing from a community with those present, considering all communities, and adjusting accordingly all cover values.

A better approach to potential community is proposed by Beals (1984), which argues that the scales of presence and absence used by Swan (1970) are not comparable, so that he defines a *sociological favorability index* generated entirely from presence/absence data, valid for both present and absent species in each community. Let A^0 a presence/absence species \times communities data table. His *smoothed* matrix Y is obtained by calculating, for each species i , the average over all species in each community of the joint occurrences of i with all others, each divided by the frequency of the other, in formula:

$$y_{ij} = \sum_{k=1}^n \left(\frac{a_{kj}^0}{a_{.j}^0} \sum_{h=1}^p \frac{a_{ih}^0 a_{kh}^0}{a_{.k}^0} \right) = \sum_{k=1}^n \left(\frac{a_{kj}^0}{a_{.k}^0 a_{.j}^0} \sum_{h=1}^p a_{ih}^0 a_{kh}^0 \right). \quad (1)$$

Indeed, such a table raises the number of species in a community and at the same time introduces the idea of *potential presence* of species, that may be there since there could be conditions that allow their presence and growth in the community. Thus, Y may be considered a *potential*

meta-community matrix derived by A . The table smoothing occurs since the presence/absence of species is transformed into a probability of occurrence of a species at a given site on the basis of its joint occurrences with the remaining species in the data table (De Cáceres and Legendre, 2008); thus, each 1 in A^0 is fractioned in probabilities of occurrence of absent species and a table smoothing results. The last formulation in (1) shows that y_{ij} is based on a double normalization of A^0 , according to both row and column totals, and on its cross-product AA' . On the other side, is interesting to note that we may assume the inner sum in the first formulation in (1) as

$$p(i|k) = \sum_{h=1}^p \frac{a_{ih}^0 a_{kh}^0}{a_{k \cdot}^0} \quad (2)$$

and interpret $p(i|k)$ as the probability of the i -th species to be in a community conditioned by the k -th species presence, since $\sum_i p(i|k) = 1$.

Beals smoothing is highly appreciated by plant ecologists (McCune, 1994; Schnittler et al., 2006) and currently used in applications (see De Cáceres and Legendre, 2008, for citations). Indeed, the latter authors warn about its misuse, and revisit smoothing to establish its reliability, in particular suggesting a statistical test to either accept or reject the null-hypothesis of random species co-occurrences. De Cáceres and Legendre (2008) propose also to remove the i -th species itself from the estimation, thus proposing an alternative matrix Z , whose elements are, according to: (2)

$$z_{ij} = \sum_{k=1, k \neq i}^n \frac{a_{kj}^0}{a_{\cdot j}^0} p(i|k) \quad (3)$$

De Cáceres and Legendre (2008) suggest that conditional probabilities $p(i|k)$ may be estimated by considering the joint distribution of species of another metacommunity data matrix referring to the same species. The question whether to use Equations (1) or (3) remains open: the inclusion of the present species in the conditional probabilities computation could raise them in sites in which the species is present, thus causing a problem should inference be done (Münzbergová and Herben, 2004); on the other side, removing it could get higher probabilities for a species to occur where it is not present than where it is (Oksanen et al., 2008). In all cases, zeros are removed from the smoothed matrices.

Beals (1984) proposed the method only for presence/absence data. Thus, should A be a count matrix, it might be reduced to a presence/absence one, say A^0 prior to compute conditional probabilities and perform the smoothing. It must be noted the way De Cáceres and Legendre (2008) define the conditional probabilities $p(i|k)$: indeed, they define the vectors of conditional probabilities

$$\mathbf{p}(|\mathbf{k}) = (\text{diag}(A^0 A^0))^{-1} A^0 \mathbf{A}_k^0 \quad (4)$$

with \mathbf{A}_k^0 the vector corresponding to the k -th column of A^0 . With this formulation, we may, as they do in their electronic appendix, use Equation 4 to generalize Beals smoothing to the case in which A contains abundances. Indeed, in this case, one must distinguish A (with abundance data) from A^0 (with presence/absence data); thus, Equation 4 becomes

$$\mathbf{p}(|\mathbf{k}) = (\text{diag}(A' A^0))^{-1} A' \mathbf{A}_k^0 \quad (5)$$

and the resulting $p(i|k)$ s may be used in both Equations 1 and 3.

The choice of a smoothing, whether to exclude or not the present species from the conditional probabilities and whether to use abundance or presence/absence, should be the consequence of a thorough experimentation. Thus, in the following we shall refer with Y to a smoothed matrix,

regardless how it was built.

Indeed, smoothing reduces the differences among sampled communities, so that the quest for sub-populations may be done on Y better than on A : in fact, two different sampled communities may have identical potential community structure and thus they would result two realization of the same potential population φ . This is a crucial point that deserves being studied in detail, since it could lead to the identification of subpopulations in the metacommunity at hand better than by classifying the original data matrix. A cluster analysis may do the job, but, in order to give usable results, one should ascertain that all the communities gathered in a group are estimates of the same potential community. The comparison may be done according to several methods:

- The Kullback-Leibler divergence (DKL , Kullback and Leibler, 1951; Kullback, 1959): it is a non-symmetric measure of the difference between two probability distributions $P = p_1, \dots, p_r$ and $Q = q_1, \dots, q_r$. Specifically, the DKL of Q from P , denoted $DKL(P \parallel Q)$, is a measure of the information lost when Q is used to approximate P :

$$DKL(P \parallel Q) = \sum_{i=1}^r p_i \log_2 \frac{p_i}{q_i}$$

In information terms, the $DKL(P \parallel Q)$ measures the expected number of bits required to complete the code necessary to describe P once it is approximated through Q . Typically P represents the empirical (observed) data distribution and Q a model or an approximation of P . DKL properties are:

1. $DKL(P \parallel Q)$ is defined only if, for every i , $q_i = 0 \Rightarrow p_i = 0$. For our purposes this should not be a problem, since if species do not appear in either distribution, we are driven to consider them two different potential communities.
2. $DKL(P \parallel Q) \geq 0$ (Gibbs' inequality).
3. $DKL(P \parallel Q) = 0$ if and only if for every i , $q_i = p_i$.

It is important to warn against the use of Kullback as distance, since: *i*) it is not symmetric and *ii*) the triangular inequality does not hold. For our purposes, we may decide that two potential communities are equal if the probability to get a random $DKL(P \parallel Q)$ larger than the one we want to check is sufficiently high (larger than 5%) under permutation, randomization, montecarlo or any other suitable method.

- The Chi-square test:

$$CHI(P, Q) = \sum_{i=1}^r \frac{(p_i - q_i)^2}{p_i + q_i}$$

it is the usual test to compare two distributions: indeed, to be applied, it is necessary to deal with the original counts and not with the profiles. Note that if the observations are equal the degrees of freedom should be reduced by 1 (Press et al., 1992).

2.1 Comparison

Y may be compared to W . Which is the purpose of such a comparison? If we think that Y is a better estimate of φ , we may wonder which is the meaning of the residuals, that is of differences between W^j and Y^j . Are they due to some early arrival of a species? to its random presence? is it the winner of a competition?

Let us consider the residuals $(\mathbf{W} - \mathbf{Y})_{ij}$ of species i in the community j . We have several choices:

1. $SS_{e,i} = \sum_j ((\mathbf{W} - \mathbf{Y})_{ij})^2$: measures the total sum of squares of residuals for each species in the whole metacommunity.
2. $SS_{e,j} = \sum_i ((\mathbf{W} - \mathbf{Y})_{ij})^2$: measures the total sum of squares of residuals for each community considering the found set of species.
3. $SS_e = \sum_j \sum_i ((\mathbf{W} - \mathbf{Y})_{ij})^2$: it is a total sum of squares of residuals in the metacommunity.

As an alternative to the ordinary sum of squares, two other measures may be taken into account: the absolute differences, with a similar meaning

1. $SAD_{e,i} = \sum_j |(\mathbf{W} - \mathbf{Y})_{ij}|$.
2. $SAD_{e,j} = \sum_i |(\mathbf{W} - \mathbf{Y})_{ij}|$.
3. $SAD_e = \sum_j \sum_i |(\mathbf{W} - \mathbf{Y})_{ij}|$.

that have the well known advantages of the L_1 -metrics on the L_2 - *Euclidean* one: in this particular case, the sum of squares reduces dramatically the small differences, that may in practice disappear. The simple differences $SD_{e,i} = \sum_j (\mathbf{W} - \mathbf{Y})_{ij}$ have no issue dealing with both communities and metacommunity, since they sum up to zero, but may be of relevance when studying species across the metacommunity at hand: some may have systematically larger (or smaller) presence than expected everywhere, an information surely interesting to study in detail.

For exploratory purposes, it may be of interest to get an overall evaluation of the quality of the prediction of W through the potential metacommunity Y . At least, two methods may be proposed:

- The *RV* coefficient among them:

$$RV(W, Y) = \frac{\text{trace}(\mathbf{W}\mathbf{W}'\mathbf{Y}\mathbf{Y}')}{\sqrt{\text{trace}(\mathbf{W}\mathbf{W}')\text{trace}(\mathbf{Y}\mathbf{Y}')}}.$$

RV (Escoufier, 1973; Robert and Escoufier, 1976) acts as a squared correlation coefficient, since it ranges $[0, 1]$, with the ordinary meaning. As Josse et al. (2008) propose tests for significance of *RV*, no particular computations are necessary to get an overall examination of the relation between real and potential communities.

Assuming *RV* as correlation coefficient, *Statis* method (Lavit, 1988) is a three-way *RV*-based *PCA*, that allows the representation of the tables in principal *interstructure* spaces, the definition of a *compromise* in which to represent both the *intrastructure*, that is all characters and principal components of all two-way data tables involved, and the units *trajectories*, that is their position according to each data table at hand, together with the compromise position, a kind of centroid of trajectories. For further details, see Lavit (1988); Abdi et al. (2012).

- The *Procrustes correlation*

$$R^2 = \text{trace}(\Delta^2)$$

with $\Delta = E'\tilde{Y}\tilde{W}'F$, the diagonal matrix of the singular values of the product of Y W , transformed as explained in the following, as resulting by the *Singular Value Decomposition* (*SVD*, Abdi, 2007) of $\tilde{Y}\tilde{W}' = E\Delta F'$ with E, F unit matrices.

Procrustes Analysis (PA) is a method that provides the best adjustment of a set of points, called *test cloud*, to a given set, called *target cloud*, according to transformations that do not change, up to a scale factor, the reciprocal distances among the points of the test cloud. Originally proposed by Mosier (1939), its name is due to Hurley and Cattell (1962) and further developments are due to Gower (1971b). The method may be applied to all situations in which direct comparisons among configurations of the same objects under different representations are requested.

We suppose that Y and W are the two tables with dimensions (n, p) whose columns correspond to the two sets of corresponding points we aim at comparing. Then a transformation of Y is sought that minimizes the sum of the distances between the pairs of corresponding points: the transformation is composed by three steps: a rotation T , a translation a , and a scaling s , that may be described as (see Borg and Groenen, 2005):

1. center by columns and standardize both matrices to have both their total variance equal to 1 (Gower, 1975), getting the matrices \tilde{Y} and \tilde{W} ;
2. compute the product $\tilde{Y}\tilde{W}'$;
3. compute the *SVD* of $\tilde{Y}\tilde{W}' = E\Delta F'$.

The Procrustes residual sum-of-squares statistic (Gower, 1971a, 1975) is $m_{12}^2 = 1 - \text{trace}(\Delta^2)$, whose complement to 1 is the said Procrustes correlation.

3 Traits and neighborhoods

Let us now consider species traits, that are characteristics of species that one wishes to take into account, like leaf area and thickness, plant height, seed size, etc. Its use may be of help in understanding the structure of the population at hand better than the catalog of present species. Indeed, for our purposes, we may think that the potential presence and the abundance of a species in a site may depend upon some traits the species has and that shares with others that were actually found in the same site (Pillar et al., 2009). Thus, we may use what we know about species traits to build a potential community matrix, analogous to the one we built by considering the simple species association.

Referring to the same population we are dealing with, we may consider the trait matrix B , in which the n rows represent species and the t columns represent traits. Each cell entry b_{ik} , $i = 1, \dots, n, j = 1, \dots, t$ is a trait level, depending upon the kind of the trait. Indeed, the choice of the traits depends upon the specific interests of both the researcher and the investigation. Note that matrix B is independent from the specific population at hand, since it depends upon only the traits characteristic of a species (unless in a specific site a species shows some particular traits). Thus, B usually may be built based on other databases or literature and according to the traits that the researcher thinks may be of interest (Pillar et al., 2009).

An interesting use of B is to transform the community matrix W into a traits \times communities one. For this task, it is advisable that B be an indicator matrix; thus, nominal characters should be split into presence/absence of each level, and quantitative ones may be left as they are. Indeed, the $t \times p$ matrix $T = B'W$, allows a more compact representation of the population, in terms of traits (Pillar et al., 2009). Depending on the kind of data in B , the entries in T are either a relative frequency of a binary trait in a community or the weighed average of a quantitative trait of the species present in the community (Pillar et al., 2009).

Additionally (and independently from the communities at hand), from B , a similarity (symmetric) $n \times n$ matrix S among species may be built, by using any similarity index. Considering traits of different kind, the one proposed by Gower (1971a) may be a good choice, but any index may do the job, see also Goodall (1966). Thus, S measures how two species are either different ($s_{ij} \rightarrow 0$) or alike ($s_{ij} \rightarrow 1$).

S may be used in data analysis as usual for both ordination and classification, if any, with limitations due to the nature of the used index. Here, we use it to create *fuzzy sets* (Zadeh, 1965; Pillar et al., 1991) to which species may be attributed according to their *degree of belonging*. Unlike ordinary sets, whose attribution of elements is ascertained without doubt, thus either they belong to a set or do not, the degree of belonging establishes how an element belongs to fuzzy set elements: it is a real number ranging $[0, 1]$, where 0 means “does not belong” and 1 means “does belong”, and intermediate values indicate intermediate conditions.

For our purposes (see Pillar et al. 2009; Pillar and Duarte 2010), we normalize the columns of S to their respective total, we obtain a matrix U whose columns $U^j = S^j/s_{.j}, j = 1, \dots, n$, are a kind of similarity profile of species j with all other species. We may interpret its values as probabilities that the species j be replaced by another one, based on its normalized similarity, under the assumption that the higher the similarity among species, the higher the chance to find another species in the place of the found one. Note that this interpretation is merely based on the fact that each profile elements sum to 1, not on empirical surveys, that may be done to ascertain the reliability of the method.

Let us now look at U by rows: in every $U_i, i = 1, \dots, n$ for the corresponding i -th species, its normalized similarities with all others (including itself) appear. Note that if a species is similar only to itself, then $u_{jj} \rightarrow 1$ and all other values tend to zero; if it is highly similar to another, say k , both u_{jj} and u_{kj} will tend to 0.5, and so on, the column maximum progressively lowering with the increase of the number of high similar species. This is evident, since the replacement may be with several species, each with a lower probability. We may try to attribute a meaning to the rows $U_i, i = 1, \dots, n$ of U : their sums are the sum of the similarities normalized by columns; thus, the most similar are the other species to some, the largest this sum. Thus, we may not consider them probabilities, as they do not sum to 1, but *degrees of belonging* of all species to a fuzzy set defined by the i -th species itself, that we call the *neighborhood* of species i in the traits space. Indeed, the more likely a species may be replaced by species i the larger is its degree of belonging to i -th neighborhood. Thus, a neighborhood represents the capacity of a species to replace others, as said under the assumption that the similarity of traits favours the replacement.

Now, let us right multiply the matrix U by W , giving an $n \times p$ matrix $X = UW$ (Pillar et al., 2009; Pillar and Duarte, 2010). This matrix has a special meaning, since each element $x_{ij} = \sum_k u_{ik}w_{kj}$ represents an estimation of the abundance of the neighborhood of species i in the j th community. Indeed, x_{ij} is based on relative frequencies w_{kj} and since the columns of U are already standardized to unity sum, each element x_{ij} is also a relative frequency.

Once again, the columns of $X, X^j, j = 1, \dots, p$, represent the profile of a potential community, this time built through the fuzzy sets defined by each species based on the traits similarity among species.

Thus, matrix X describes a potential metacommunity, under the hypothesis that similarity of species traits may favour their reciprocal replacement in a given community. As said, this hypothesis ought to be ascertained through a specific experimentation.

3.1 Comparison

Matrix X may be compared to other matrices in the same ways as suggested in Section 2.1, i.e. sum of squared or absolute residuals, the RV coefficient or Procrustes correlation, with analogous interpretation, with the difference that this time our fitting depends upon the choice of the traits at hand. Thus, we must adopt a model-like approach, that means that we must evaluate how X fits the target matrix. Depending upon the target, either W or Y , the meaning of the results, in particular the interpretation of the residuals, may be different, as well as the choice of the most suitable traits for defining X . As said, we get two alternatives:

1. Find traits that maximize the fit to W . The procedure is identical to the addition of independent predictors in linear models, that each time improve the model, with the consequence that one may end with a perfect fit of W , thus losing the idea of potential community.
2. Find traits that maximize the fit to Y . This choice would mean that we assume as potential community that obtained by smoothing and we search the traits that best fit it. Indeed, should the assumption be correct, that is that common traits may define potential communities, this seems the best way to proceed.

It is evident that the second alternative is better, because Y may be assumed as an upper limit of the adjustment based on traits. Indeed, one may require a good fit, in order to conclude that the communities obtained by smoothing correspond to some specific traits structure.

4 Example

We apply the method to plant community data from (Blanco et al. 2007, see also Pillar et al. 2013), which were collected in a 3-ha grassland site of the south Brazilian campos vegetation. There were 70 0.2×0.2 m plots located systematically in 14 sets of 5 contiguous plots, in which the estimated cover of plant species was recorded. The species were described by a matrix B of 12 morphological traits. The traits refer to plant height (HE), leaf tensile strength (LT), woody biomass proportion (WB), maximum height (MH), senescent leaves proportion (SL), upper leaf density proportion (UL), leaf width (LW), leaf thickness (TH), all quantitative, and presence of vegetative propagation by rhizomes ($V1$), presence of vegetative propagation by stolons ($V2$), presence of smooth leaf surface ($L1$), presence of prickles on leaf surface ($L2$), as binary traits. These traits were measured in each plot and here we considered their average value for each species across the plots.

In the first analysis we considered as input such matrix B and a community matrix W of 70 plots by 61 species. We computed Beals smoothing based on the species relative cover and excluding the present species from the conditional probabilities (De Cáceres and Legendre, 2008), generating matrix Z of potential communities, which was also standardized to unit community total. By using a simple iterative algorithm (see Pillar and Sosinski 2003) we searched, among the 12 traits, for a trait subset maximizing the RV coefficient between matrices X and Z ($RV(XZ)$). Then we tested the significance of the RV coefficient against the null model described in Pillar et al. (2009), which is based on the random permutation among the rows of matrix B , keeping matrices W and Z fixed. Since matrices W and Z are not permuted, there is no risk of type I error inflation due to spatial autocorrelation, since if there is any, it is incorporated in the null model. The results indicated that maximum $RV(XZ)$ was obtained by using the following traits: $SL, MB, BH, V1, V2, LW$, and TH with $RV(XZ) = 0.574312$ ($p = 0.012$). When all traits were considered in the computation of X , the coefficient dropped to a non-significant $RV(XZ) = 0.247725$ ($p = 0.43$).

In another analysis we considered a matrix W with 14 pooled plots in which the five plots in each set were pooled in one larger plot of 0.2×1.0 m. Maximum $RV(XZ)$ was obtained by nearly the same trait subset: $SL, BH, V1, V2, LW$, and TH , with a higher $RV(XZ) = 0.826478$ ($p = 0.008$). Using all traits for defining X , the RV coefficient, again, dropped to a non-significant $RV(XZ) = 0.554034$ ($p = 0.182$).

The results suggest that the selected traits are critical in the assembly of the studied plant communities and that these traits are good proxies of plant adaptations to the prevailing environmental conditions and biotic interactions. With larger plots we should expect a more robust estimation of both X and Z , but the effects of environmental conditions and biotic interactions may be blurred, since environmental conditions and biotic interactions are in general spatially structured.

5 Conclusion

Some further consideration deserves being done: the choice of different traits in B may lead to different results, thus, the problem raises to choose the traits that best build a potential community. We do not close the discussion concerning these alternatives. Systematic deviations of one species, independently of the gradients or correlated to them, may help in establishing some particular features. In any case, chance would play a mayor role.

Acknowledgements

This work was carried out during the stay of Sergio Camiz at Laboratório de Ecología of Universidade Federal do Rio Grande do Sul (*UFRGS*) with a Special Visiting Researcher Fellowship under the Brazilian Scientific Mobility Program “Ciências sem Fronteiras” (Grant number: 401345/2014-9).

References

- Abdi, H. (2007). RV Coefficient and Congruence Coefficient. In Salkind, N., editor, *Encyclopedia of Measurement and Statistics*, pages 1–10. Sage, Thousand Oaks (CA).
- Abdi, H., Williams, L. J., Valentin, D., and Bannani-Dosse, M. (2012). STATIS and DISTATIS: Optimum multitable principal component analysis and three way metric multidimensional scaling. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2):124–167.
- Beals, E. W. (1984). Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. *Advances in Ecological Research*, 14:1–55.
- Blanco, C. C., Sosinski, E. E., Santos, B. R. C., Abreu da Silva, M., and Pillar, V. D. (2007). On the overlap between effect and response plant functional types linked to grazing. *Community Ecology*, 8:57–65.
- Borg, I. and Groenen, P. (2005). *Modern Multidimensional Scaling*. Springer Series in Statistics, Berlin.
- De Cáceres, M. and Legendre, P. (2008). Beals smoothing revisited. *Oecologia*, 156(3):657–669.
- Escoufier, Y. (1973). Le Traitement des Variables Vectorielles. *Biometrics*, 29(4):751–760.
- Fisher, W. D. (1958). On Grouping for Maximum Homogeneity. *Journal of American Statistical Association*, 53:789–798.
- Goodall, D. W. (1966). A new similarity index based on probability. *Biometrics*, 22(4):882–907.
- Gower, J. C. (1971a). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*, 27(4):857–871.

- Gower, J. C. (1971b). Statistical methods of comparing different multivariate analyses of the same data. In Hodson, F., Kendall, D., and Tautu, P., editors, *Mathematics in the Archaeological and Historical Sciences*, pages 138–149. Edinburgh University Press.
- Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, 40(1):33–51.
- Hurley, J. R. and Cattell, R. B. (1962). The Procrustes Program: Producing Direct Rotation to Test a Hypothesized Factor Structure. *Computers in Behavioral Sciences*, 7:258–262.
- Josse, J., Pagès, J., and Husson, F. (2008). Testing the significance of the RV coefficient. *Computational Statistics and Data Analysis*, 53(1):82–91.
- Kullback, S. (1959). *Information Theory and Statistics*. John Wiley and Sons, New York.
- Kullback, S. and Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.
- Lambert, J. and Dale, M. B. (1964). The Use of Statistics in Phytosociology. *Advances in Ecological Research*, 2:59–99.
- Lavit, C. (1988). *Analyse Conjointe de Tableaux Quantitatifs*.
- Legendre, P. and Legendre, L. (2012). *Numerical Ecology*. 3rd ed. edition.
- McCune, B. (1994). Improving community analysis with the beals smoothing function. 1:82–86.
- Mosier, C. I. (1939). Determine a simple structure when loadings for certain tests are known. *Psychometrika*, 4(2):149–162.
- Münzbergová, Z. and Herben, T. (2004). Identification of suitable unoccupied habitats in metapopulation studies using co-occurrence of species. *Oikos*, 105(2):408–414.
- Oksanen, J., Kindt, R., Legendre, P., O’Hara, B., Simpson, G., and Stevens, M. (2008). *Vegan: community ecology package*. r package version 1.11-0.
- Pillar, V. and Duarte, L. d. (2010). A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters*, 13:587–596.
- Pillar, V. D., Blanco, C. C., Müller, S. C., Sosinski, E. E., Joner, F., and Duarte, L. D. S. (2013). Functional redundancy and stability in plant communities. *Journal of Vegetation Science*, 24:963–974.
- Pillar, V. D., Duarte, L. D., Sosinski, E. E., and Joner, F. (2009). Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *Journal of Vegetation Science*, 20(2):334–348.
- Pillar, V. D., Orloci, L., and Feoli, E. (1991). Fuzzy components in community level comparisons. pages 87–93. Kluwer, Dordrecht.
- Pillar, V. D. and Sosinski, E. E. (2003). An improved method for searching plant functional types by numerical analysis. *Journal of Vegetation Science*, 14(3):323–332.
- Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (1992). *Numerical Recipes in C - The Art of Scientific Computing*. Oxford University Press, New York, 2nd ed. edition.
- Robert, P. and Escoufier, Y. (1976). A Unifying Tool for Linear Multivariate Statistical Methods: The RV-Coefficient. *Applied Statistics*, 25(3):257–265.
- Schnittler, M., Unterseher, M., and Tesmer, J. (2006). Species richness and ecological characterization of myxomycetes and myxomycete-like organisms in the canopy of a temperate deciduous forest. *Mycologia*, 98(2):223–232.
- Sokal, R. and Rohlf, F. (1981). *Biometry: the principles and practice of statistics in biological research*. New York.
- Swan, J. (1970). An Examination of Some Ordination Problems By Use of Simulated Vegetational Data. *Ecology*, 51(1):89–102.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353.