Exploratory Analysis of Traits for the Study of the Seed Dispersal by Birds

Sergio Camiz[1], Leandro D.S. Duarte[2], Gastão C. Gomes[3]

[1]*Dipartimento di Matematica Guido Castelnuovo, SapienzaUniversità di Roma, CNR–Istituto di Acustica e Sensoristica "O.M. Corbino".* sergio.camiz@uniroma1.it
[2]*Departamento de Ecologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.* duarte.ldas@gmail.com
[3]*Departamento de Matemática e Estadística – Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.* gastao@im.ufrj.br

RESUMO
Em mosaicos de floresta e pastagem, arbustos e árvores isoladas modificam a disponibilidade de recursos e condições ambientais sob as copas de suas árvores, criando habitats adequados para o estabelecimento de espécies florestais. Os mosaicos também dependem da dispersão de diásporas vegetais (frutos e / ou sementes) das florestas aos locais abertos. Árvores estabelecidas em pastagens podem atrair aves frutívoras que as usam como poleiros, aumentando a distribuição de sementes sob suas copas. Plantas colonizadoras com grande produção de sementes deverão ter vantagem sobre espécies com poucas sementes aumentando a área do mosaico. No entanto, a colonização também deverá ser influenciada pelas atividades dos agentes de dispersão, especialmente pássaros frutívoros. Foi avaliada a relevância funcional de traços de diáspora como indicadores de colonização de manchas de floresta com Araucária no sul do Brasil. Diásporos de gêneros de plantas colonizadoras foram caracterizados por tipo, tamanho e cor. Nesta comunicação analisaremos conjuntamente matrizes de dados coletados em trabalho de campo, a fim de estabelecer relações mais profundas entre os diferentes itens levando em conta: amostras de gênero de pássaro e gênero de plantas. Em particular, Análise de Correspondência foi usada, a fim de obter uma primeira visão dos dados, que será útil no planejamento de análises estatísticas mais adequadas. Este trabalho foi apresentado no "International Center for Earth Sciences – ICES-8", out 2012. Durante a visita do prof. Camiz a UFRJ, em abril 2013, apoiada pela FAPERJ (APV-E-26/110.018/2013), este trabalho foi finalizado e enviado para o ICES-8 e será publicado em suas atas.

ABSTRACT
In grassland-forest mosaics, isolated shrubs and trees modify the resource availability and environmental conditions beneath their crowns, creating suitable habitats for the establishment of forest species, leading to the nucleation of forest patches. Nucleation of forest patches in grassland also depends on dispersal of plant diaspores (namely fruits and/or seeds) from forests to open sites. Woody plants established on grassland attract frugivorous birds that use them as perches, increasing seed rain beneath the plant crown. Traditionally, traits of vertebrate-dispersed diaspores, such as size and color have been used as indicators of dispersal by distinct dispersers. Vegetation mosaics formed by Araucaria forest and Campos grassland constitute the main natural landscape found in southern Brazilian highlands. Colonizer plants with large, reserve-rich seeds are expected to have advantage over small-seeded species as patch area increases. Nonetheless, colonization is also expected to be influenced by the activities of dispersal agents, especially frugivorous birds. We evaluated the functional relevance of diaspores traits as indicators of plant-disperser mutualisms in the colonization of Araucaria forest patches in southern Brazil. Diaspores of colonizer plant genera were characterized by diaspore type, size and color.
In this communication we jointly analyze data matrices of data collected in a field work, in order to establish deeper relations among the different items taken into account: relevés characters, bird genus, plant genus, and traits. In particular, Correspondence Analysis was taken into account, in order to get a first insight of the data, helpful in planning more appropriate statistical analyses.

INTRODUCTION
A tendency of woody plant encroachment in grass-dominated ecosystems has been currently observed in several parts of the world (Bond & Midgley, 2000; Bond et al., 2003; Scholes &

Archer, 1997). In grassland-forest mosaics, isolated shrubs and trees modify the resource availability and environmental conditions beneath their crowns, creating suitable habitats for the establishment of forest species (Bruno et al., 2003; Callaway, 1995), leading to the nucleation of forest patches (Duarte et al., 2007). Nucleation of forest patches in grassland also depends on dispersal of plant diaspores (namely fruits and/or seeds) from forests to open sites, and the mechanisms underlying this process provide a basis for the understanding of grassland-forest boundary dynamics. In woody plant communities exposed to high rainfall levels, a major proportion of plant species is dispersed by animals, especially vertebrates (Herrera, 2002), whose behavior and habitat preferences may determine the deposition site of seeds (Nathan & Muller-Landau, 2000). Woody plants established on grassland attract frugivorous birds that use them as perches for resting and/or for foraging activities (Shiels & Walker, 2003), increasing seed rain beneath the plant crown, and thus favoring the colonization of open areas by forest species (Guevara et al., 1992).

Traditionally, attributes of vertebrate-dispersed diaspores, such as size, color and smell, have been used as indicators of dispersal by distinct disperser groups(Van der Pijl, 1972). Hereafter these diaspores traits are referred as *Disperser Attraction Traits* (*DAT*). The adaptive value of *DAT* has been vastly demonstrated by empirical studies (e.g. Gauthier-Hion et al. 1985).

Nevertheless, interspecific variation in *DAT* is expected to show some degree of phylogenetic signal (Fischer and Chapman 1993). An alternative to prevent inflated co-variation in diaspores traits due to phylogenetic signal consists of using higher taxonomic levels (genus, family, etc.) instead of species to characterize diaspores in relation to ecologically relevant attributes (Voigt et al. 2004). Vegetation mosaics formed by Araucaria forest and Campos grassland constitute the main natural landscape found in southern Brazilian highlands (Rambo, 1994). Araucaria forest expansion over grassland in southern Brazil was already demonstrated by paleopollen evidence (Behling & Pillar, 2007), and it occurs either gradually by edge dynamics (Oliveira & Pillar, 2004), or by nucleation of forest patches in the grassland (Duarte et al., 2006a). The recruitment of woody saplings in Araucaria forest patches is strongly associated with patch area. As patch size increases, the relative abundance of colonizer species decreases, while the abundance of local recruiters increases (Duarte et al. 2007). Therefore, a positive correlation between patch size and space occupancy by resident species is observed. In such a scenario, colonizer plants with large, reserve rich seeds are expected to have advantage over small-seeded species (Leishman 2001, Moles and Westoby 2006). Indeed, we found a significant positive association between patch size and mean seed diameter of colonizer species (Duarte et al. 2007). Most species colonizing Araucaria forest patches were found to show vertebrate-dispersed diaspores (Duarte et al. 2007). Thus, although the higher seed reserve content conferred advantage for plants establishing in large patches, colonization patterns seemed to be also influenced by processes related to attributes and/or behavior of dispersal agents.

MATERIALS AND METHODS

Data are taken from Duarte et al. (2007) and refer to 24 Araucaria forest patches scattered in a grassland site surrounded by a continuous Araucaria forest in the southern Brazilian highlands (29°28'S and 50°13'W, at ca. 900ma.s.l.): the full description of plant composition is reported there. The data have been organized in the following matrices:

1) Matrix *W*. This binary matrix records the presence/absence of a plant genus among 13 genera (*Miconia, Myrsine, Nectandra, Ficus, Myrceugenia, Ocotea, Cabralea, Eugenia,*

Leandra, Psidium, Solanum, Drimys, and *Rudgea*) in each of 24 observed patches. Each patch analyzed contained, at least, two colonizer plant genera.

2) Matrix *E*.   This binary matrix attributes the 24 patches to one of 3 groups according to area and structural development (Duarte et al., 2006b): *i)* Small patches (*NP, N = 14*) of *Araucaria angustifolia* individuals isolated in the grassland (mean crown area ± standard error = 35 m^2 ± 5.2). *ii)* Medium patches (*SP, N = 5*) correspond to groups of forest woody plants with at least one adult tree and accompanying shrubs forming a continuous canopy, surrounded by a continuous herb layer (mean patch area ± standard error = 195.3 m^2 ± 37.9). *iii)* Large patches (*LP, N = 5*, data from (Machado, 2004) with higher structural development (mean patch area ± standard error = 1899.6 m^2 ± 255.7).

3) Matrix *B*.   This matrix describes diaspore traits of the 13 colonizer plant genera. For each trait, described by diaspore type, size and color (*DAT,* Duarte et al., 2007), the genus value is the proportion of species in this genus with the trait present. Berries (*BERR*) and drupes (*DRUP*) represent 85% of diaspores, while the remaining 15% are arillate seeds (*ARIL*, 8%) or figs (*FIGS,* 8%). We considered three diaspore size classes: most diaspores (ca. 31%) are small (*SMDI,* diameter smaller than 10 mm.), 48% are medium (*MEDI,* diameter within 10 and 35 mm.), while only 21% are large (*LADI,* diaspores larger than 35 mm.). Bright colors (*black, violet, red*, and *orange*) are present in 82% of diaspores, while dull colors (*brown* and *yellow*) are present in the remaining 18% of colonizer diaspores.

4) Matrix *N*.   Data on plant-bird interactions were obtained from the available literature on bird diet and frugivory (see the Electronic Appendix of Duarte et al., 2011). We defined a plant-frugivore table in which the presence/absence of interactions between the 13 plant genera (rows) and 30 fruit-consumer bird species (columns) is recorded.

From these given data tables, other tables were obtained by multiplication of the original matrices:

5) Matrix *WE.* This matrix results as product of *W* and *E* gathering the patches according to their size. Thus, *WE* crosses 13 plant-genuses with the 3 patch sizes. It contains the frequencies of plant genera in each patch size type.

6) Matrix *BN.* This matrix is the product of *B* and *N:* since *B* crosses plant traits x plant genera, and *N* is a plant genera by bird species binary interaction matrix, their product will result in a plant traits by bird species ccontingency table. That is, *BN* contains the trait frequencies in the plants that are picked by each bird species. It results in a 13 x 30 matrix that represents the interest of a bird species to pick diaspores with specific plant traits.

For the analysis, we submitted all matrices, but the *E* that is trivial, to *Correspondence Analysis* (*CA*, Benzécri et al., 1973-82; Greenacre, 1983). This method allows representing graphically the items that cross in the tables, which makes easier the representation of the relations between tables' rows and columns. In addition, two tests, based on the chi-square, allow establishing the significance of the dimensions, according to the corresponding significance of the table obtained by the partial reconstruction due to a reduced number of factors. This is especially useful if one may associate to each factor some external character that may influence the table structure.

Indeed, as the chi-square is a multiple of the table's inertia, expressed in the chi-square metrics, it results that it is decomposed proportionally to each factor's eigenvalue:

$$\chi^2 = \chi_1^2 + \chi_2^2 + \ldots = n\lambda_1 + n\lambda_2 + \ldots = n \sum_{\alpha=1}^{\min(r,c)-1} \lambda_\alpha$$

In the reconstruction formula of *CA,* each table's cell is worth

$$n_{ij} = nr_i c_j \left(1 + \sum_{\alpha=1}^{\min(r,c)-1} \lambda_\alpha u_{i\alpha} v_{j\alpha} \right) = nr_i c_j + nr_i c_j \lambda_1 u_{i1} v_{j1} + nr_i c_j \lambda_2 u_{i2} v_{j2} + \ldots$$

so that the table may be seen as a sum of *layers*, that is, a rank-1 table, whose first corresponds to the expected value under independence and each other corresponds to the departure from expectation due to the α-th factor. Each layer inertia equals the corresponding eigenvalue, so that its chi-square is proportional to their sum.

On this basis, for each α-th layer a test may be done for significance of its χ^2, with max($r;c$)-1 degrees of freedom, and another for the significance of the α-thresidual, that is the difference between the α-dimensional reconstruction and the original table. Indeed, the following decomposition holds for any α:

$$n_{ij} = nr_i c_j \left(1 + \sum_{\beta=1}^{\alpha} \lambda_\beta u_{i\beta} v_{j\beta} + \sum_{\beta=\alpha+1}^{\min(r,c)-1} \lambda_\beta u_{i\beta} v_{j\beta} \right)$$

Thus, the table's chi-square may be decomposed as a sum of independent chi-squares, the first with ($r+c$-2)α-1 and the second with (r-α-1)(c-α-1) degrees of freedom, respectively.
Indeed, the latter statistics approximates the goodness of fit statistics (Kendall and Stewart, 1961) as showed by Malinvaud (1987). Should this test be positive, the residuals carry a significant deviation from the expectation and must be examined; otherwise, there would be no interest to go further. In analogous way, each layer may be tested for significance with ($r+c$-2α−1) degrees of freedom: should it be positive, it would mean that the layer's deviation from expectation is in itself significant, thus the associated factor has a significant linear influence on the table's meaning (Orlóci, 1978).

Indeed, the problem tied to both tests is that the matrices we are dealing with are in general incidence matrices. Thus, on one side they are rather sparse and on the other, their values are very small. Thus, a true chi-square might not be performed. Nevertheless, it is customary to apply Correspondence Analysis the same, so that no particular attention is drawn to these tests. We used them the same, in order to have an idea (maybe too conservative) of the intensity of the relations we are studying with them.

RESULTS

The exploration of the data tables through correspondence analysis gave the following results:

1) Matrix *W*
The *CA* of the genera-patches table *W*, albeit not significant per se, reveals one factor significant at over 1.3% level, to which over 40% of the total table inertia may be attributed. The second factor, albeit attributed over 20% of the total inertia, does not result significant at 5% level. In Figure 1 the pattern of both genera and patches is clearly in arch form. This is usually interpreted as the presence of an important main gradient that influences the data most than the others. It is well known from the theory (see Camiz, 2005) that in this case the

following factors adjust the elements' position along the gradient. Indeed, on the positive side
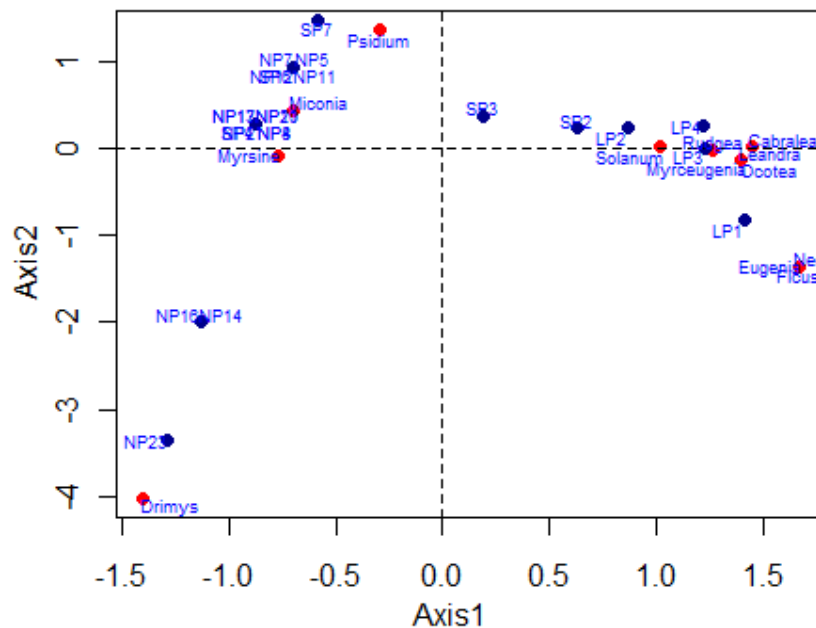

Figure 1 - Genera and patches on the first factor plane of CA of matrix W.

of the first axis, the largest patches are set, and most of the small ones are set on the other side, with the medium patches are in the middle, somehow mixed with the small ones. For what concerns the genera, four groups appear very clearly: the Drimys is set alone at the left end of the gradient, then Myrsine, Miconia, and Psidium are set on the left side closer to the center; on the right site all other genera are found, with Eugenia, Nectandria, and Ficus at the very end of the gradient. This pattern may be interpreted as a factor determined by the patches size, concerning in particular the distinction between small and large patches.

2)   Matrix E
      This matrix only classifies the patches in the three classes, as specified. For this reason
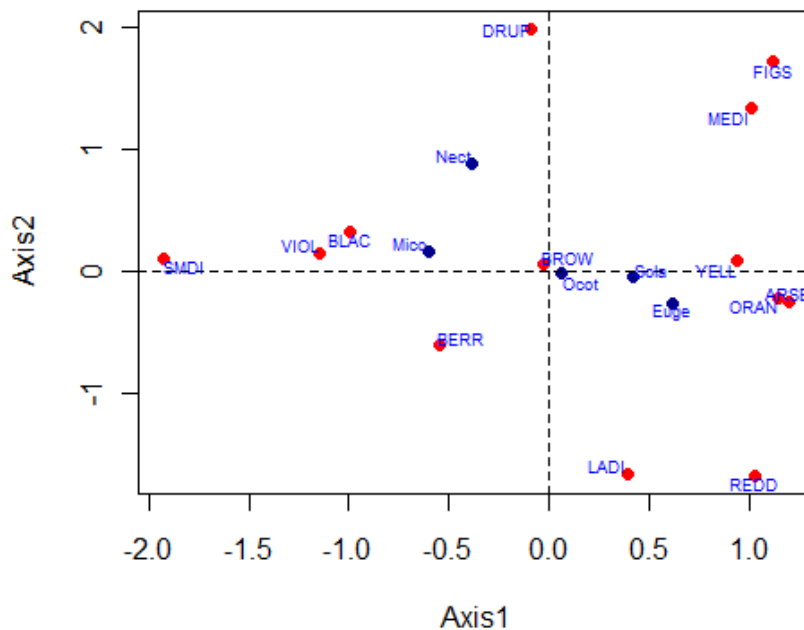

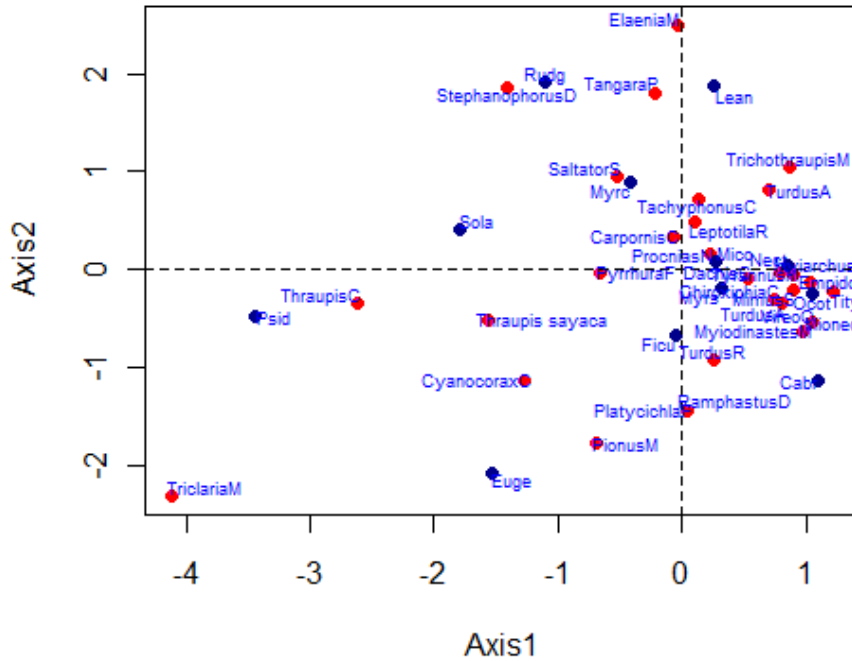Figure 2 - Diasporas traits and genera on the first factor plane of CA of matrix B.

Figure 3- Plant genera and Bird species on the first factor plane of CA of matrix N.

its correspondence analysis does not show particularly interesting results, since its rigid geometry, with the patches set on the vertices of a regular triangle, was implicit in the data structure. On the opposite, it will be used in the following to synthesize the results concerning the patch size.

3)   Matrix *B:*   This matrix is a composition of three matrices, each one referring to a
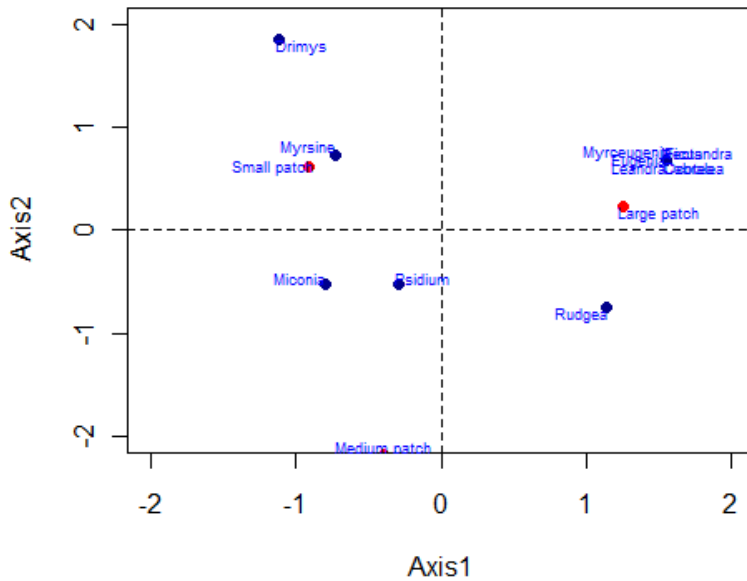

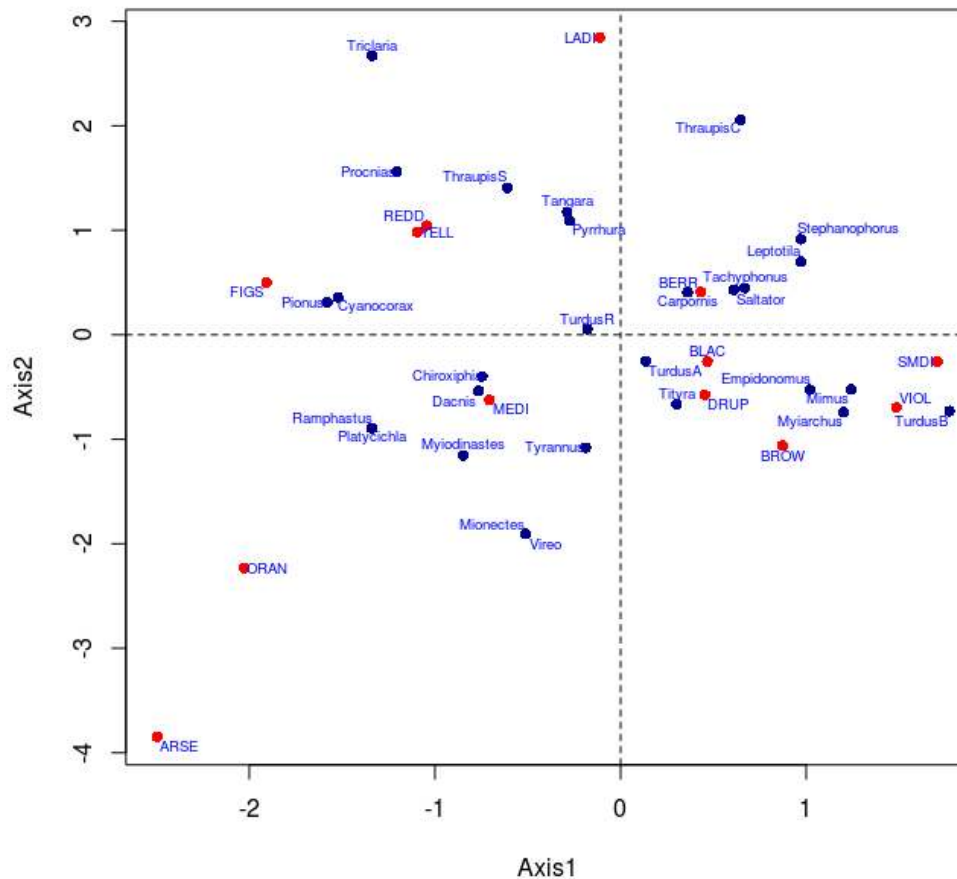*Figure 4 - Plant genera and Patches size on the first factor plane of CA of matrix WE.*

Figure 5 - Traits of diaspores and Bird species on the first factor plane of CA of matrix BN.

specific trait of diaspores: the first two, kind and size, are typical of each genus, whereas the colour, is more distributed among genera, so that the table results not significant. Nevertheless, its representation on the first factor plane of *CA* (Figure 2) describes sufficiently well the relation between genera and their diaspores traits of belonging.

4)   Matrix *N*

Even this matrix is non-significant, thus would not deserve being analyzed, unless a certain low significance of the first factor (11%), that suggests to give it a glance. The scattering of both plant genera and bird species on the first factor plane of *CA* is represented in Figure 3.

5)   Matrix *WE*

This matrix represents the pattern of plant genera according to patch size. Unlike the matrix *W*, in this case a highly significant factor results, summarizing most of the matrix inertia (87%). Indeed, looking at Figure 4, an arch effect is evident, and an association between patch size and groups of genera seems readable. Thus, *Drymis* and *Myrsinae* seem to be preferably in small patches, *Miconia, Psydium,* and *Rudgea* seem to be rather ubiquitous (but with a tendence toward large patches, according to this sequence, and all others present only in large patches.

6)   Matrix *BN*

This table, in which the traits of diaspores are crossed with bird species, indicate the bird preferences in picking some specific diaspores. As it is, this table is statistically non-

significant, that is, its chi-square is not significant. For this reason, it is questionable to consider the associations between birds and diaspores as reliable. In addition, in *CA*, albeit the first two table's factors are accounted for 60% of total inertia, only the first has a significance of around 15%. This is too low to be currently taken into account, so that only some descriptive attention to it may be given. It is possible that the relative positions of both traits and birds understate some preferences, but we are not in condition to argue anything about it. In Figure 5 the relative positions of both bird species and diaspores traits are represented, but no comment may be reliably provided.

CONCLUSION

In this study we found that *DAT* variation in the colonizer plants living in contrasting habitat types resulted from a complex interplay between phylogenetic relationships between plants, interactions with frugivorous birds, and habitat-specific characteristics related to patch size. Interactions with frugivorous birds helped to explain *DAT* variation across the patches. The ecological interplay between diaspore traits and vertebrate dispersers has been extensively demonstrated in the literature (Gauthier-Hion et al. 1985; Voigt et al. 2004), and is likely mediated by functional interactions between plant traits and bird traits (Moran et al. 2004). Morphological attributes of diaspores make them more appropriate for consumption by particular sets of frugivores than by others. For instance, bird gape width limits the ingestion of diaspores above a compatible size (Wheelwright 1985). Thus, some relevant correspondence between functional attributes of plants and dispersers would reveal meaningful interaction patterns at the patch level.

This first work is a part of a larger study, in which the diaspores' traits are studied with their interactions with the birds, in order to understand their dispersal. Indeed, this exploratory exam allows to understand to what extent the data tables taken into account may be suitable for the study. Some specific problems raised by the *BN* matrix gets questionable the information it may give concerning the interaction between diaspores traits and birds, at least, as far as the chi-square statistics may be considered reliable for this kind of data. Thus, some further deepening will be considered an important step in this study.

REFERENCES

Behling, H. and V.D. Pillar (2007), "Late Quaternary vegetation, biodiversity and fire dynamics on the southern Brazilian highland and their implication for conservation and management of modern Araucaria forest and grassland ecosystems". *Philosophical Transactions of the Royal Society B: Biological Sciences,* 362: pp. 243-251. doi:10.1098/rstb.2006.1984.

Benzécri, J.P. et coll., (1982), *L'Analyse des Données*, 2 tomes, Paris, Dunod.

Bond, W.J. and G.F. Midgley (2000), "A proposed CO_2-controlled mechanism of woody plant invasion in grasslands and savannas". *Global Change Biology,* 6: pp. 865-869.

Bond, W.J., G.F. Midgley, and F.I. Woodward (2003), "The importance of low atmospheric CO_2 and fire in promoting the spread of grasslands and savannas". *Global Change Biology,* 9: pp. 973-982.

Bruno, J.F., J.J. Stachowicz, and M.D. Bertness (2003), "Inclusion of facilitation into ecological theory". *Trends in Ecology and Evolution,* 18: pp. 119-125.

Callaway, R. (1995), "Positive interactions among plants". *The Botanical Review* 61: 306-349.

Camiz, S. (2005), "The Guttman Effect: its Interpretation and a New Redressing Method". Τετραδια Αναλυσησ Δεδομενων (Data Analysis Bulletin), 5: pp. 7-34.

Duarte, L.D.S., M. Dos-Santos, S. Hartz, and V.D. Pillar (2006a), "Role of nurse plants in Araucaria Forest expansion over grassland in south Brazil". *Austral Ecology,* 31: pp. 520-528. doi:10.1111/j.1442-9993.2006.01602.x.

Duarte, L.D.S., M.B. Carlucci, C.S. Fontana, S.M. Hartz, and V.D. Pillar (2011), "Plant diaspore traits as indicators of mutualistic interactions in woody vegetation patches developing into a grassland-forest mosaic". *Community Ecology,* 12: pp. 126-134. doi:10.1556/ComEc.12.2011.1.15.

Duarte, L.D.S., M.B. Carlucci, S.M. Hartz, and V.D. Pillar (2007), "Plant dispersal strategies and the colonization of Araucaria forest patches in a grassland-forest mosaic". *Journal of Vegetation Science* ,18: pp. 847-858. doi:10.1111/j.1654-1103.2007.tb02601.x.

Duarte. L.D.S., R:E: Machado, S.M. Hartz, and V.D. Pillar (2006b), "What saplings can tell us about forest expansion over natural grasslands". *Journal of Vegetation Science,* 17: pp. 799-808. doi:10.1658/1100-9233(2006)17[799:wsctua]2.0.co;2.

Gauthier-Hion, A., J.M. Duplantier, R. Quris, F. Feer, C. Sourd, J.P. Decoux, G. Dubost, L. Emmons, C. Erard, P. Hecketsweiler, A. Moungazi, C. Roussilhon, and J.M. Thiollay (1985), "Fruit characters as a basis of fruit choice and seed dispersal in a tropical forest vertebrate community". *Oecologia,* 65: pp. 324-337.

Greenacre, M.J. (1983), *Theory and applications of correspondence analysis.* London, Academic Press.

Guevara, S., J. Meave, P. Moreno-Casasola, and J. Laborde (1992), "Floristic composition and structure of vegetation under isolated trees in neotropical pastures". *Journal of Vegetation Science,* 3: pp. 655-664.

Herrera, C.M. (2002), "Seed dispersal by vertebrates". In: C.M. Herrera and O. Pellmyr (Eds.) (2002), *Plant-animal interactions: an evolutionary approach,* Oxford, England, Blackwell Scientific Publications: pp. 185-208.

Kendall, M.G. and A. Stuart (1961), *The Advanced Theory of Statistics*, vol. 2. London, Griffin.

Leishman, M. (2001), "Does the seed size/number trade-off model determine plant community structure? An assessment of the model mechanisms and their generality". *Oikos,* 93: pp. 294-302.

Machado, R.E. (2004), "Padrões vegetacionais em capões de floresta com araucaria no planalto nordeste do Rio Grande do Sul, Brasil", Vol. MSc.: *Ecologia*, Brazil, Porto Alegre (RGDS), ed. UFRGS.

Malinvaud, E. (1987), "Data analysis in applied socio-economic statistics with special consideration of correspondence analysis". Marketing Science Conference, Joy en Josas: HEC-ISA.

Moles, A. and M. Westoby (2006), "Seed size and plant strategy across the whole life cycle". *Oikos,* 113: pp. 91-105.

Moran, C., C.P. Catterall, R.J. Green, and M.F. Olsen (2004). "Functional variation among frugivorous birds: implications for rainforest seed dispersal in a fragmented subtropical landscape". *Oecologia,* 141: pp. 584-595.

Nathan. R. And H.C. Müller-Landau (2000), "Spatial patterns of seed dispersal, their determinants and consequences for recruitment". *Trends in Ecology and Evolution,* 15: pp. 278-285.

Oliveira, J.M. and V.D. Pillar (2004), "Vegetation dynamics on mosaics of Campos and Araucaria forest between 1974 and 1999 in Southern Brazil". *Community Ecology,* 5: pp. 197-202.

Orlóci, L. (1978), *Multivariate Analysis in Vegetation Research*, 2nd ed.. Den Haag (Netherlands), Junk.

Rambo, B. (1994), *A Fisionomia do Rio Grande do Sul*. 3 edn, Brazil, São Leopoldo (RGDS), Editora Unisinos.

Scholes, R.J. and S.R. Archer (1997), "Tree-grass interactions in savannas". *Annual Review of Ecology and Systematics,* 28: pp. 517-544. doi:doi:10.1146/annurev.ecolsys.28.1.517.

Shiels, A.B. And L.R. Walker (2003), "Bird perches increase forest seeds on Puerto Rican landslides". *Restoration Ecology,* 11: pp. 457-465.

Van der Pijl, L. (1972), *Principles of dispersal in higher plants*. 2 edn. Berlin, Springer-Verlag.

Voigt, F.A., Bleher, J. Fietz, J.U. Ganzhorn, D., Schwab, and K. Böhning-Gaese (2004), "A comparison of morphological and chemical fruit traits between two sites with different frugivore assemblages".*Oecologia,* 141: pp. 94-104.

Wheelwright, N.T. (1985), "Fruit size, gape width and the diets of fruit-eating birds". *Ecology,* 66: pp. 808-818.