Aplicações da Linguagem R em Análises de Vegetação

Écio Souza Diniz
Pedro Manuel Villa
(Organizadores)

Atena Editora
Ano 2020

Aplicações da Linguagem R em Análises de Vegetação

Écio Souza Diniz
Pedro Manuel Villa
(Organizadores)

Atena Editora
Ano 2020

2020 by Atena Editora
Copyright © Atena Editora
Copyright do Texto © 2020 Os autores
Copyright da Edição © 2020 Atena Editora
Editora Chefe: Profª Drª Antonella Carvalho de Oliveira
Diagramação: Geraldo Alves
Edição de Arte: Lorena Prestes
Revisão: Os Autores

Todo o conteúdo deste livro está licenciado sob uma Licença de Atribuição *Creative Commons*. Atribuição 4.0 Internacional (CC BY 4.0).

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores. Permitido o download da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Conselho Editorial
Ciências Humanas e Sociais Aplicadas
Profª Drª Adriana Demite Stephani – Universidade Federal do Tocantins
Prof. Dr. Álvaro Augusto de Borba Barreto – Universidade Federal de Pelotas
Prof. Dr. Alexandre Jose Schumacher – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná
Prof. Dr. Antonio Gasparetto Júnior – Instituto Federal do Sudeste de Minas Gerais
Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília
Prof. Dr. Carlos Antonio de Souza Moraes – Universidade Federal Fluminense
Prof. Dr. Constantino Ribeiro de Oliveira Junior – Universidade Estadual de Ponta Grossa
Profª Drª Cristina Gaio – Universidade de Lisboa
Profª Drª Denise Rocha – Universidade Federal do Ceará
Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia
Prof. Dr. Edvaldo Antunes de Farias – Universidade Estácio de Sá
Prof. Dr. Eloi Martins Senhora – Universidade Federal de Roraima
Prof. Dr. Fabiano Tadeu Grazioli – Universidade Regional Integrada do Alto Uruguai e das Missões
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná
Profª Drª Ivone Goulart Lopes – Istituto Internazionele delle Figlie de Maria Ausiliatrice
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense
Profª Drª Keyla Christina Almeida Portela – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
Profª Drª Lina Maria Gonçalves – Universidade Federal do Tocantins
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Marcelo Pereira da Silva – Universidade Federal do Maranhão
Profª Drª Miranilde Oliveira Neves – Instituto de Educação, Ciência e Tecnologia do Pará
Profª Drª Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa
Profª Drª Rita de Cássia da Silva Oliveira – Universidade Estadual de Ponta Grossa
Profª Drª Sandra Regina Gardacho Pietrobon – Universidade Estadual do Centro-Oeste
Profª Drª Sheila Marta Carregosa Rocha – Universidade do Estado da Bahia
Prof. Dr. Rui Maia Diamantino – Universidade Salvador
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande
Prof. Dr. William Cleber Domingues Silva – Universidade Federal Rural do Rio de Janeiro
Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

Ciências Agrárias e Multidisciplinar
Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano
Prof. Dr. Antonio Pasqualetto – Pontifícia Universidade Católica de Goiás
Profª Drª Daiane Garabeli Trojan – Universidade Norte do Paraná

Profª Drª Diocléa Almeida Seabra Silva – Universidade Federal Rural da Amazônia
Prof. Dr. Écio Souza Diniz – Universidade Federal de Viçosa
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul
Prof. Dr. Fágner Cavalcante Patrocínio dos Santos – Universidade Federal do Ceará
Profª Drª Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia
Prof. Dr. Júlio César Ribeiro – Universidade Federal Rural do Rio de Janeiro
Profª Drª Lina Raquel Santos Araújo – Universidade Estadual do Ceará
Prof. Dr. Pedro Manuel Villa – Universidade Federal de Viçosa
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará
Profª Drª Talita de Santos Matos – Universidade Federal Rural do Rio de Janeiro
Prof. Dr. Tiago da Silva Teófilo – Universidade Federal Rural do Semi-Árido
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

Ciências Biológicas e da Saúde
Prof. Dr. André Ribeiro da Silva – Universidade de Brasília
Profª Drª Anelise Levay Murari – Universidade Federal de Pelotas
Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás
Prof. Dr. Edson da Silva – Universidade Federal dos Vales do Jequitinhonha e Mucuri
Profª Drª Eleuza Rodrigues Machado – Faculdade Anhanguera de Brasília
Profª Drª Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina
Prof. Dr. Ferlando Lima Santos – Universidade Federal do Recôncavo da Bahia
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria
Prof. Dr. Igor Luiz Vieira de Lima Santos – Universidade Federal de Campina Grande
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará
Profª Drª Magnólia de Araújo Campos – Universidade Federal de Campina Grande
Profª Drª Mylena Andréa Oliveira Torres – Universidade Ceuma
Profª Drª Natiéli Piovesan – Instituto Federacl do Rio Grande do Norte
Prof. Dr. Paulo Inada – Universidade Estadual de Maringá
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande

Ciências Exatas e da Terra e Engenharias
Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto
Prof. Dr. Alexandre Leite dos Santos Silva – Universidade Federal do Piauí
Prof. Dr. Carlos Eduardo Sanches de Andrade – Universidade Federal de Goiás
Profª Drª Carmen Lúcia Voigt – Universidade Norte do Paraná
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará
Prof. Dr. Juliano Carlo Rufino de Freitas – Universidade Federal de Campina Grande
Prof. Dr. Marcelo Marques – Universidade Estadual de Maringá
Profª Drª Neiva Maria de Almeida – Universidade Federal da Paraíba
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

Conselho Técnico Científico
Prof. Msc. Abrãao Carvalho Nogueira – Universidade Federal do Espírito Santo
Prof. Msc. Adalberto Zorzo – Centro Estadual de Educação Tecnológica Paula Souza
Prof. Dr. Adaylson Wagner Sousa de Vasconcelos – Ordem dos Advogados do Brasil/Seccional Paraíba
Prof. Msc. André Flávio Gonçalves Silva – Universidade Federal do Maranhão
Profª Drª Andreza Lopes – Instituto de Pesquisa e Desenvolvimento Acadêmico
Profª Msc. Bianca Camargo Martins – UniCesumar
Prof. Msc. Carlos Antônio dos Santos – Universidade Federal Rural do Rio de Janeiro
Prof. Msc. Claúdia de Araújo Marques – Faculdade de Música do Espírito Santo
Prof. Msc. Daniel da Silva Miranda – Universidade Federal do Pará
Profª Msc. Dayane de Melo Barros – Universidade Federal de Pernambuco

Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)

A642 Aplicações da linguagem R em análises de vegetação [recurso eletrônico] / Organizadores Écio Souza Diniz, Pedro Manuel Villa. – Ponta Grossa, PR: Atena, 2020.

Formato: PDF
Requisitos de sistema: Adobe Acrobat Reader
Modo de acesso: World Wide Web
Inclui bibliografia
ISBN 978-65-86002-35-5
DOI 10.22533/at.ed.355200903

1. Desenvolvimento sustentável. 2. R (Linguagem de programação de computador). 3. Recursos vegetais – Manejo. I.Diniz, Écio Souza. II. Villa, Pedro Manuel.
CDD 333.7511

Elaborado por Maurício Amormino Júnior – CRB6/2422

Atena Editora
Ponta Grossa – Paraná - Brasil
www.atenaeditora.com.br
contato@atenaeditora.com.br

Atena Editora
Ano 2020

APRESENTAÇÃO

Os diferentes tipos de vegetação ao redor do globo, principalmente as florestas tropicais, se destacam por fornecer importantes bens e serviços ecossistêmicos para a humanidade como, por exemplo, regulação climática, provisão de alimentos e diversas fontes de energia. Contudo, as crescentes e rápidas mudanças no meio ambiente causadas por sua intensa exploração no século 21 têm promovido reduções drásticas de importantes vegetações distribuídas em distintos Biomas. O Brasil como um país de dimensão continental e rico em recursos vem atravessando profundas transformações em seus Biomas, o que é destacadamente devido aos usos intensos da terra sem técnicas adequadas de manejo para a sua exploração.

Diante desse panorama de significativas transformações do meio natural, se faz necessário e urgente o estudo de diferentes tipos de comunidades vegetais e ecossistêmicas para definir estratégias de manejo e conservação, assim como pesquisas que visem a otimização de produções agrícolas de forma sustentável. A união de compreensão ecológica precisa e adequadas técnicas de manejo permitem uma exploração sustentável a longo-prazo dos recursos vegetais, assegurando manutenção de diversidade e provisões para o futuro.

A execução de estudos robustos para alcançar essa interface entre conservação e exploração demanda o uso de eficientes ferramentas analíticas. Dentre essas ferramentas, as linguagens de programação têm se sido importantes aliadas para obtenções de predições e resultados estatísticos confiáveis e informativos. A linguagem contida no software R é a mais amplamente utilizada para processamento de dados e análises de vegetação. O R engloba diversos pacotes importantes para análises de dados de plantas em diversos contextos ecológicos e agrários. Com seus diversos pacotes, o R permite a busca mais apurada pela compreensão de padrões e processos ecológicos, avaliação de impactos antrópicos sobre vegetação, monitoramentos e previsões de condições do solo para plantios e predições de efeitos de mudanças climáticas em florestas. Essa gama de possibilidades analíticas amplifica o acerto em tomadas de decisão com relação ao uso dos nossos recursos naturais de forma geral.

Este livro tem como objetivo trazer uma compilação de algumas potencialidades do software R para análise de vegetação, contribuindo para o aumento da capacidade técnica de diversos profissionais das áreas de Ciências da Terra ou Naturais no uso dessa poderosa ferramenta analítica. Para tal, os capítulos aqui presentes discorrem de forma aplicada sob temas em contextos ecológicos e agrários. Todos os capítulos possuem links de compartilhamento livre de dados e scripts com códigos para execução das análises que eles abordam no R. Assim, desejamos que o conteúdo aqui presente auxilie você leitor (a) em sua tarefa analítica, amplificando a obtenção de resultados informativos e potenciais de aplicação prática.

Écio Souza Diniz
Pedro Manuel Villa

SUMÁRIO

CAPÍTULO 10

EFFECTS OF SPATIAL SCALE ON PEQUI ENTOMOFAUNA

Data de aceite: 12/02/2020

Data de submissão: 28/11/2019

Gustavo Amorim Santos
Universidade Federal de Goiás, Instituto de Química
Goiânia – Goiás
http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4210157E7

Deomar Plácido da Costa
Instituto Federal Goiano,
Ceres – Goiás
http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4744387Y9

Suzana da Costa Santos
Universidade Federal de Goiás, Instituto de Química
Goiânia – Goiás
http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4793945U9

Pedro Henrique Ferri
Universidade Federal de Goiás, Instituto de Química
Goiânia – Goiás
http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4785526A5

ABSTRACT: Spatial distribution and abundance of entomofauna can alter plant community structure by affecting the competitive ability of individual plants. Spatial influence on the diversity of leaf entomofauna in eleven populations of *Caryocar brasiliense* was determined at large, intermediate and small scales using Moran's eigenvector maps (MEMs) and multivariate variograms. The insect community was best captured by MEMs derived from the Gabriel graph which was linearly weighted in terms of distance between sampling sites. Spatial dependencies were measured in the phenolic contents of pequi leaves and soil nutrients, which suggests that insect diversity is spatially dependent on these variables. Large and intermediate spatial scales were mainly related to the highest Mn^{2+} and Cu^{2+} soil contents and to the lowest total phenol contents of pequi leaves. MEMs not only enabled the modelling of multiscale spatial influence in the study area, but also controlled the spatial variance not explained by phenolic contents and soil nutrients.

KEYWORDS: *Caryocar brasiliense*, insects, chemical variability, soil nutrients, spatial variation.

EFEITOS DA ESCALA ESPACIAL NA ENTOMOFAUNA DO PEQUI

RESUMO: A distribuição espacial e abundância da entomofauna podem alterar a estrutura de uma comunidade de plantas, afetando a capacidade competitiva de plantas individuais. A influência espacial na diversidade da entomofauna foliar em onze populações de *Caryocar brasiliense* foi determinada em escalas grande, intermediária e pequena, utilizando mapas de autovetores de Moran (MEMs) e variogramas multivariados. A comunidade de insetos foi melhor representada pelos MEMs derivados do grafo de Gabriel ponderado linearmente com a distância entre os locais de amostragem. Dependências espaciais foram observadas quanto aos teores fenólicos das folhas de pequi e nutricionais do solo, sugerindo que a diversidade de insetos é espacialmente dependente dessas variáveis. As estruturas espaciais de escalas grande e intermediária foram relacionadas principalmente aos maiores teores de Mn^{2+} e Cu^{2+} do solo e ao menor teor de fenóis totais das folhas. Os MEMs não apenas permitiram modelar a influência espacial em múltiplas escalas na área de estudo, como também controlaram a variação espacial não explicada pelos conteúdos fenólicos e nutrientes do solo.

PALAVRAS-CHAVE: *Caryocar brasiliense*, insetos, variabilidade química, nutrientes do solo, variabilidade espacial.

1 I INTRODUCTION

Insect richness results from the balance of several ecological factors (ABDALA-ROBERTS et al., 2019). Insects, which amount to nearly 60% of all animals, are especially useful as bioindicators of environmental changes (ZEQUI et al., 2019). Climate, plant-insect interaction, habitat complexity, and host plant quality have generally been used to explain variations in insect diversity within and between habitats (TORRES & MADI-RAVAZZI, 2006). Plant metabolites can also affect entomofauna distribution and, to some extent, result from responses to the local abiotic and biotic environment (MUTIKAINEN et al. 2000), despite being genetically determined and potentially inducing the presence of herbivores (MEAUX & MITCHELL-OLDS, 2003). Soil nutrients can also affect plant metabolites (CUNNINGHAM et al., 1999), although the consequences for plant-insect interactions are still little explored (MEINDL et al., 2013).

Food resources and soil nutrients vary in space, directly affecting the diversity and distribution of insect populations (BISPO & OLIVEIRA, 2007). Although several studies have assessed spatial influence on insect and plant distribution, few have simultaneously analyzed entomofauna distribution with the host plant's metabolite variability and abiotic factors such as soil nutrients. The joint effects of these factors have recently been described on leaves of *Caryocar brasiliense* Cambess (Caryocaraceae), a species tree found in several phytophysiognomies of the Brazilian Cerrado and popularly referred as pequi.

The diversity of pequi entomofauna, when restricted to its leaves (FERREIRA et al., 2009), shows variation fractions better explained by edaphic (64.5%) and spatial

(60.6%) predictors than phenolic contents (22.7%) in the leaves (COSTA et al., 2019). The spatial pattern observed by Costa et al. used polynomial expressions of spatial coordinates as predictors, which cannot accurately model the spatial patterns on intermediate and small scales (JOMBART et al., 2009). Thus, empirically observed spatial influence may stem from a multiscale process, whose identification is a crucial issue in ecological studies (DUNGAN et al., 2002).

In this study, spatial influence on diversity of insects in pequi was partitioned into the spatial components of large, intermediate and small scales, by applying Moran's eigenvector maps (MEMs) as well as multiscale ordination (MSO) using multivariate variograms (WAGNER, 2004). Thus, the aim was to produce a more effective description of the associations between the entomofauna and phenolic contents of pequi leaves, in addition to the soil nutrients sampled in the study sites.

2 I MATERIAL AND METHODS

Data Collection and Analysis: Leaf entomofauna was collected from 11 pequi populations located in the cities of Alvorada do Norte (ALV), Araguapaz (ARA), Damianópolis (DAM), Faina (FAI), Formoso (FOR), Hidrolândia (HID), Iporá (IPO), Mambaí (MAM), Morrinhos (MOR), Mutunópolis (MUT), and Paraúna (PAR), all in the central Brazilian Cerrado, in Goiás state. Detailed information on leaf collection and phenolic analysis (total phenols, total tannins, flavonoids, condensed tannins and hydrolysable tannins), as well as on the collection and analysis of 17 physicochemical soil parameters, has been previously described (COSTA et al., 2019). Canonical correspondence analysis (CCA) was applied to the response matrix formed by occurrence (presence/absence) of leaf insects (41 sampling sites x 95 species), conditioned by edaphic (41 x 17 soil parameters) and leaf chemicals (41 x 5 phenolic contents) defined as explanatory variables (COSTA et al., 2019). Successive CCAs allowed the selection of significant soil (Cu^{2+}, Fe^{3+}, Mn^{2+}, Zn^{2+}, P) and leaf (total phenols, flavonoids, total tannins and hydrolysable tannins) variables. In all CCAs, the variance inflation factor ($VIF < 4$) with Holm-Bonferroni correction was used to guide the selection of explanatory variables, avoiding multicollinearity in multivariate regressions. VIF values > 20 are considered strongly multicollinear (LEPS & SMILAUER, 2007).

Moran's eigenvector maps (MEMs) were generated by means of a spatial weighting matrix (**W**) obtained by the Hadamard product of a connectivity matrix (**B**) with weighting matrices (**A**) (DRAY et al., 2006; DRAY et al., 2019). Matrix **B** was generated using three graph-based connection schemes (Gabriel = *gab*, relative neighbourhood = *rel* and minimum spanning tree = *mst*). Weighting matrices **A** were derived from three functions: f_{bin} = binary, based only on topology and reflecting a neutrality in the weight; $f_{lin} = 1-(d/dmax)$, linear, and $f_{con} = 1-(d/dmax)^{0.5}$, nonlinear, both weighted by the inverse of the power of Euclidean distances (d and dmax) between two sampling sites (BAUMAN et al., 2018). The selection of matrix **W** among the nine candidate matrices and the subset of significant MEMs was conducted using the optimization method (*listw.candidates* function) available in the *adespatial* package (DRAY et al., 2019), considering only MEMs associated with positive eigenvalues.

Significant MEMs were incorporated into the explanatory matrix containing previously selected soil nutrients and leaf phenolic contents. They were submitted to stepwise selection by CCA, whose best model was based on the VIF < 8 and the adjusted coefficient of determination (R^2adj), allowing the comparison of different models. Pure and overlapping variance fractions of predictors were assessed by total variance partitioning via partial CCA (pCCA). In pCCA, four subsets of predictors were considered as explanatory: environmental variables, incorporating chemical and edaphic variables (total phenols, total tannins, flavonoids, hydrolysable tannins, Cu^{2+}, Fe^{3+}, Mn^{2+}, Zn^{2+}, and P) and spatial predictors representing the large (MEM1, MEM3), intermediate (MEM5, MEM6) and small (MEM8, MEM14) scales previously selected by the CCA.

Multiscale ordination (MSO) was used to model spatial structures (WAGNER, 2004), with semivariance multiple comparisons adjusted by the Bonferroni correction ($k = 0.05/n$, where n is the number of classes and k the adjusted significance), avoiding the increase of error type I. Envelopes describing the confidence limits around the mean semivariance were used to assess the independence of scale over distance. The number of classes ($h = 11$) was determined by Sturges' rule, which avoids arbitrary inflation of spatial variation explained as a function of class size. All analyses were conducted using *vegan* (OKSANEN et al., 2017), *spdep* (BIVAND & PIRAS, 2015) and *adespatial* (DRAY et al., 2019) packages, in R version 3.4.4 (R CORE TEAM, 2018).

3 I RESULTS AND DISCUSSION

All spatial weighting matrix **W** displayed high statistical power, regardless of connectivity matrix **B** (*gb*, *rel*, *mst*) between the sampling sites; weighting matrices (**A**) were either binary (without weight) or weighted by functions that decreased linearly or nonlinearly with the distance between two sites (Table 1).

W[a]	R^2adj (global)[b]	R^2adj (subset)[c]	No. of MEMs	*P*[d]
gb-binary	0.86660	0.85700	17	0.00089
***gb*-linear**	**0.90096**	**0.89929**	**17**	**0.00089**
gb-nonlinear	0.87552	0.86779	17	0.00089
rel-binary	0.85395	0.85133	18	0.00089
rel-linear	0.89861	0.89214	17	0.00089

rel-nonlinear	0.87579	0.87116	17	0.00089
mst-binary	0.85395	0.85133	18	0.00089
mst-linear	0.89861	0.89214	17	0.00089
mst-nonlinear	0.87579	0.87116	17	0.00089

Table 1. Spatial weight matrices (**W** = **B*****A**) with adjusted coefficients of determination to the global model and subset of MEMs.

[a]Spatial weight matrix via graphs (*gb* = Gabriel, *rel* = relative neighbourhood, *mst* = minimum spanning tree). [b]R^2adj for the global model and [c]subset of MEMs. [d]Probability with Sidak correction.

Among the nine candidate **W** matrices, the one that used the linear weighting function with the Gabriel graph captured the highest variability of pequi leaf entomofauna. Optimization of **W** and selection of the MEM subset were based on maximising the R^2adj value and controlling error type I rate by the Sidak correction, $Ps = 1-(1-P)^m$, where m = number of tests. The graphs used to represent the connectivity matrix are inclusive, so that all *mst* connections are contained in those of *rel*, which in turn are those of *gb*, with the number of connections increasing along these graphs (LEGENDRE & LEGENDRE, 2012). However, Table 1 shows the same representation for *mst* and *rel*, probably as a result of sites' irregular arrangement. Simulations indicated that the performance of MEMs under irregular sampling is highly dependent on the connectivity matrix and, consequently, on an appropriate neighbourhood relationship (BRIND'AMOUR et al., 2018). In other simulated scenarios, connectivity matrices based on a Delaunay triangulation performed poorly, a finding associated with connections at the edges of sampling sites by Delaunay criterion (BAUMAN et al., 2018). These edge effects artificially connect distant sites, hence misspecifying matrix **W** and causing the observed underestimations. Due to the sample's considerable irregularity, we chose not to incorporate the Delaunay graph among connectivity matrices.

The analysis of CCA using previously selected leaf chemical and soil nutrient variables allowed the selection of six significant MEMs (MEM1, MEM3, MEM5, MEM6, MEM8 and MEM14), based on VIF < 8 and R^2adj values. CCA results show that 91.7% (R^2adj = 86.7%, $F = 18.3$, $p = 0.001$) of total variation in entomofauna can be explained by all selected predictors (total phenols, total tannins, flavonoids, hydrolysable tannins, Cu^{2+}, Fe^{3+}, Mn^{2+}, Zn^{2+}, P, MEM1, MEM3, MEM5, MEM6, MEM8 and MEM14) which showed strong correlations ($R_1 = 0.962$; $R_2 = 0.995$). Modelling took into account the occurrence (presence/absence) of entomofauna species, which seems to better reflect the community compared to the species richness index or distance-based approaches (LEGENDRE & LEGENDRE, 2012). The latter may overlap dispersion and local effects, leading to a failure to detecting a multivariate effect in ordination techniques (WARTON et al., 2012).

CCA1 (17.7%, $F = 4.8$, $p = 0.001$) was associated mainly with an increase in total phenols, tannins and hydrolysable tannins, as well as with a reduction in flavonoid contents (Figure 1). These trends were observed in MEM5 and MEM3, respectively, which suggests that CCA1 is linked to metabolite variability. Moreover, CCA2 (14.8%, $F = 4.8$, $p = 0.001$) was mainly associated with a decrease in soil micronutrients mainly correlated to MEM1. CCA also show greater insect diversity in pequi populations located in micronutrient-rich soils (COSTA et al., 2019), such as those in Morrinhos (MOR), Araguapaz (ARA), Formosa (FOR), and Mutunópolis (MUT).

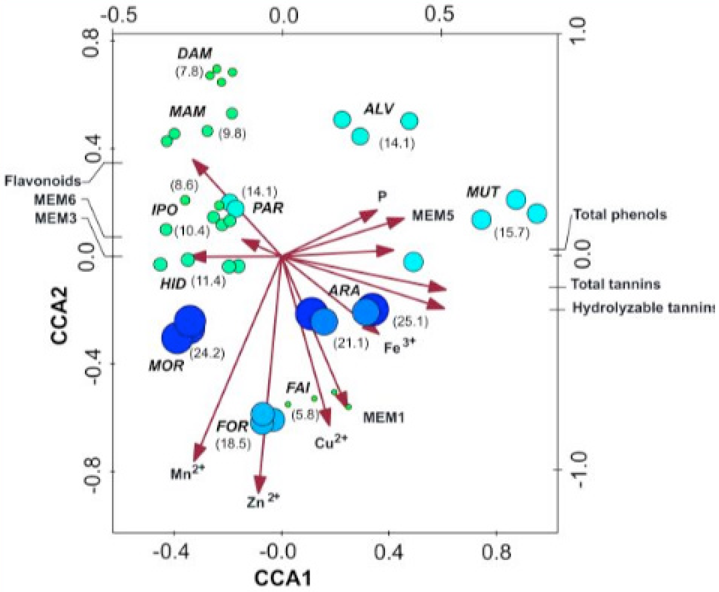

Figure 1. CCA triplot of pequi entomofauna with explanatory variables (red arrows). Pequi populations are represented by circles whose areas are proportional to Hill's measured diversity index (values in parentheses). MEM8 and MEM14 are not shown

Partitioning of the explained the variation provided R^2adj values in pure fractions ($p < 0.05$), without overlap of predictor effects, of 1.8% (leaf chemical), 43.8% (edaphic) and 33.2% (spatial), which were similar to 1.4%, 37.0% and 33.0%, respectively, obtained with polynomial functions of geographic coordinates (COSTA et al., 2019). MEMs represent a spectral decomposition of spatial relationships between observations in all possible variation ranges; thus, they represent different spatial scales. Those obtained by CCA adequately modelled spatial influence in large (MEM1, MEM3), intermediate (MEM5, MEM6) and small (MEM8, MEM14) scales (Figure 2).

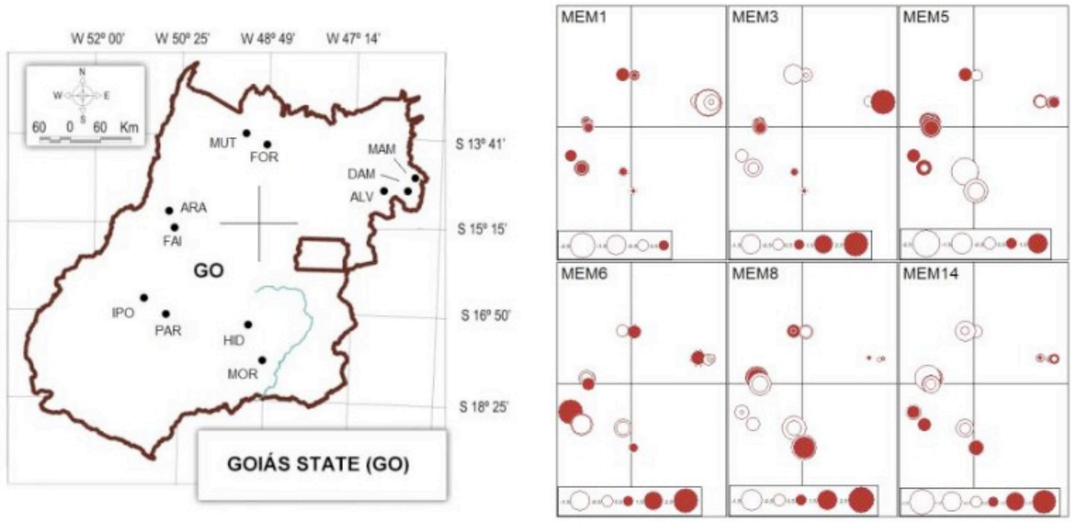

Figure 2. Pequi sampling sites and MEMs that modelled the large (MEM1, MEM3), intermediate (MEM5, MEM6) and small (MEM8, MEM14) scales of leaf entomofauna. MEMs with high positive values have larger brown bubbles.

Variance partitioning of MEMs by pCCA of leaf entomofauna indicated a larger R^2adj for the intermediate scale (15.2%), followed by large (11.5%) and small scales (8.0%), which suggests pure spatial influences ($p \leq 0.03$). Successive CCAs for each spatial scale resulted in canonical axes representing the three spatially structured patterns of entomofauna variation. Subsequent regression of site scores, viewed by significant canonical axes, on phenolic contents and soil nutrients made it possible to assess environmental influence on spatial patterns (Table 2).

Scale (MEMs)	**Variable**	**Estimate**	**Standard error**	***t*-Values**	***P* (>l *t* l)**
Large	Fe^{3+}	3.5886	0.9195	3.903	0.0005
(MEM1, MEM3)	Mn^{2+}	1.0241	0.3733	2.743	0.0100
	Total phenols	-3.4196	1.5186	-2.252	0.0316
Intermediate	Cu^{2+}	8.3212	1.6331	5.095	1.6e-05
(MEM5, MEM6)	Fe^{3+}	-1.9963	0.7908	-2.524	0.0169
	Total phenols	-2.9849	1.4805	-2.016	0.0525

Table 2. Regression parameters of spatial variation of significant canonical axes from leaf entomofauna with environmental variables.

As shown in Table 2, the large-scale spatial pattern of leaf entomofauna is clearly related to the highest Fe^{3+} and Mn^{2+} contents in the soil and the lowest total phenol contents in pequi leaves. The space-environment relationship observed at the intermediate scale, in turn, indicates an association with the highest and lowest levels of Cu^{2+} and Fe^{3+}, respectively. The absence of total phenols had only a weak influence in this scale. As it commonly observed, the structured small-scale spatial pattern had no relation with the environment. When significant, small-scale MEMs stem mainly from local spatial correlations generated within the community (LEGENDRE & LEGENDRE, 2012). These results are useful in explaining higher insect richness in pequi populations located in micronutrient-rich soils, as well as high phenolic content correlated with higher insect richness of Hymenoptera and Hemiptera in some populations, whereas lower values are mainly related to sites with a large number of beetles (COSTA et al., 2019).

Quantification of the influence of each subset in the space-environment relationship was obtained by pCCA, and results can be seen in the Venn diagram in Figure 3, whose values correspond to those of R^2adj.

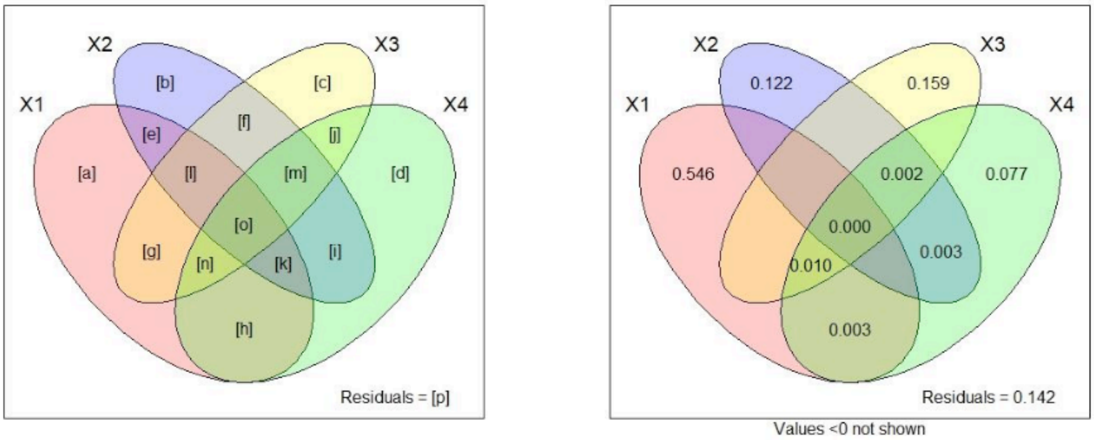

Figure 3. Partitioning of the explained variation of pequi leaf entomofauna by environmental descriptors (X1) (leaf phenolic contents and soil nutrients) and MEMs for large (X2), intermediate (X3) and small (X4) spatial scales. Negative values of R^2adj are not shown.

Results show that 91.1% (R^2adj = 85.8%) of the entomofauna variation ([a-o]) was explained by predictive matrices, resulting in a model with a residual of 14.2% ([h]). It was observed that 54.6% ([a]) of insect variability can be explained solely by soil characteristics and leaf phenolic contents. Pure spatial variation contributed with 35.8% ([b + c + d]), with major contributions to intermediate ([c] = 15.9%) and large ([b] = 12.2%) scales, and all pure fractions were regarded as significant ($p = 0.001$). Partitioning generally assumes only additivity in predictor variables and can produce negative components of variation due to interactions between variables. In fact, spatial scale-related MEMs presented a nonzero intersection ([j + m + n + o] = 1.2%), despite being orthogonal. This is due to the fact that environmental variables involved in the pCCA are not orthogonal to MEMs, and the technique involves subtracting R^2adj values which depend on the number of explanatory variables (LEGENDRE & LEGENDRE, 2012).

A solution to this problem was the combination of ordination techniques (CCA) with variograms, i.e. multiscale ordination (MSO). The MSO enables us to evaluate not only mean and variance stationarity in the study area, but also the absence of spatial correlation in regression residuals, hence allowing regression coefficients to be interpreted, since residuals are spatially independent (WAGNER, 2004).

Initially, total variation in distance classes (*h*) was obtained without conditioning, that is, via correspondence analysis (CA) of the response matrix (entomofauma), followed by eigenvalue decomposition over distance (Figure 4A). Each point in the multivariate variogram represented the sum of all contributions of environmental variables to spatial covariance, weighted by sample pairs within the distance class and plotted as the mean of distance between the pairs. Total variance showed a sharp increase with distance, and spatial autocorrelation in two classes was statistically significant, which challenged the assumption of stationarity along distance *h*.

The entomofauna matrix was then conditioned by leaf phenolic contents and soil nutrients using CCA, whose eigenvalues were also decomposed along the distance classes (Figure 4B).

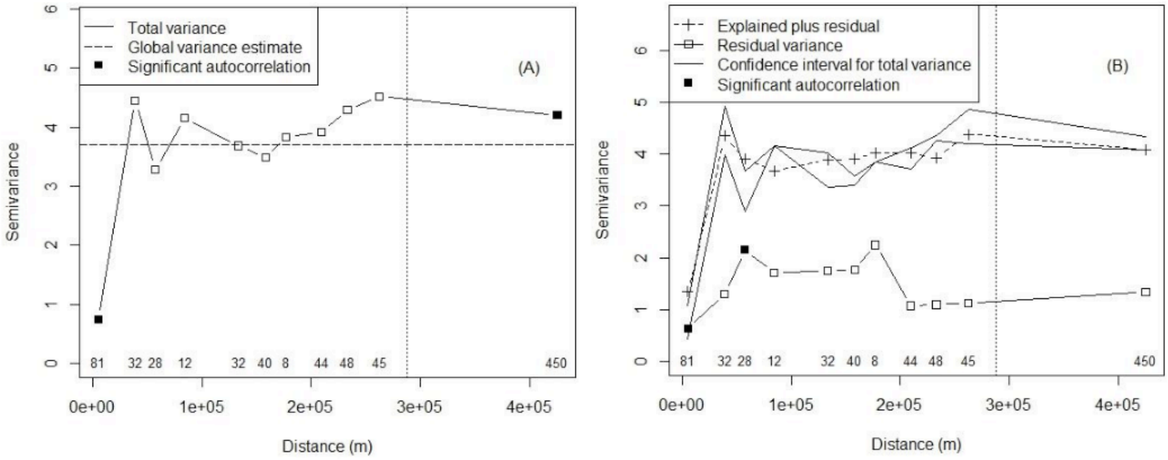

Figure 4. Multivariate variogram of (A) total variation of pequi entomofauna and (B) explained variation by environmental variables (leaf phenolic contents and soil nutrients).

Residual variance showed spatial autocorrelation in the first (distance < 5.1 km) and third (57.4 km) distance classes. CCA residues were spatially independent between and above these classes. Furthermore, when correlations between CCA matrices do not vary with the scale, the sum of explained and residual variances remains within a 5% probability envelope. In Figure 4B, only four classes remained within this confidence interval, and environmental predictors cannot be assumed to be implicitly invariant over a geographic scale.

In an attempt to control the regionalization of environmental variables caused by spatial autocorrelation, MSO was performed by introducing MEMs into the set of predictor variables (Figure 5A). Results show that the model controlled spatial autocorrelation in the residuals at distances of up to 60 km. This control can be attributed to the variation fraction captured by pCCA (fractions [n + h], Figure 3), suggesting an environmental effect on the small and intermediate spatial distances. This is important because genotypic variability among neighboring trees is often associated with a spatial structure, regardless of environmental conditions (COVELO et al., 2011). However, it could result from phenotypic characteristics of the individuals, which are strongly associated with autocorrelation on a small scale (LEGENDRE & LEGENDRE, 2012). However, this does not seem to be the case in this study.

Despite spatial autocorrelation control in the residuals of the regression models, the sum of explained and residual variances exceeded the confidence envelope in the fourth (85 km) and seventh (177 km) distance classes (Figure 5A). These results, compared with those given by variation partitioning (Figure 3), allow this finding to be attributed to the intersection of the spatial structure at different scales, represented by fractions [i + m] (R^2adj = 0.5%). Lastly, CCA was performed with the explanatory environmental variables, i.e. phenolic contents and soil nutrients, now controlled by spatial structure, using MEMs as covariates. Figure 5B shows MEMs controlling scale dependence, so that the model was suitable for the environmental predictor variables.

In the multivariate variogram there is no spatial autocorrelation in regression residuals, and the sum of explained and residual variances after elimination of the spatial structure represented by MEMs remained within the confidence interval of total variance, including the fourth and seventh distance classes. In addition, MEMs removed the trend of total variance gradient of leaf entomofauna, thus resulting in a globally oscillating multivariate empirical variogram. Therefore, MEMs were able to control spatial variation in pequi leaf entomofauna, which was not explained by the predictive variables represented by leaf phenolic contents and soil nutrients.

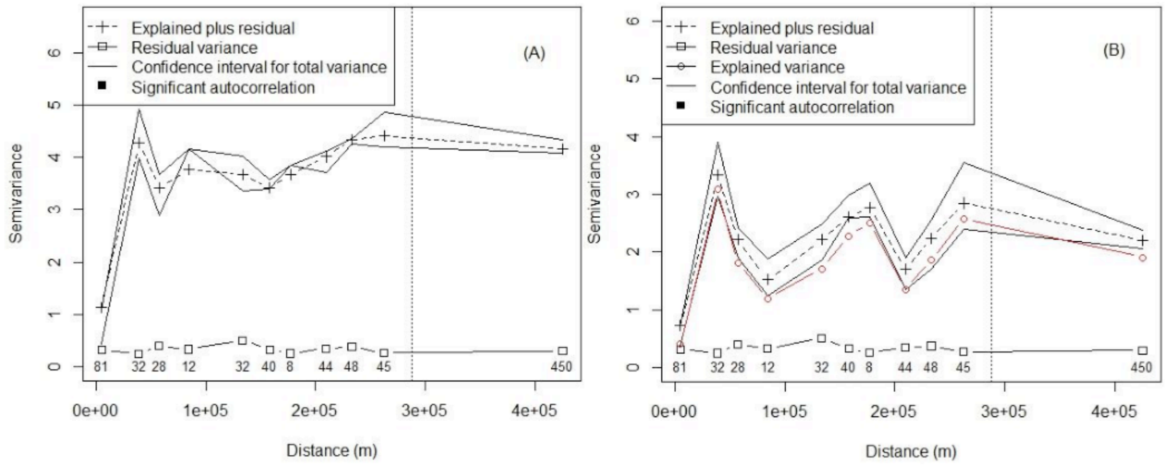

Figure 5. Multivariate variogram of pequi entomofauna (A) explained by environmental and spatial variables (MEMs) and (B) after removal of spatial structure explained by MEMs.

4 I CONCLUSIONS

Plant-soil interactions may influence entomofauna occurrence and diversity, despite being as yet largely unexplored. In this study, a significant spatial correlation was observed between pequi leaf entomofauna, on the one hand, and leaves' phenolic contents and soil mineral nutrients from sampling sites, on the other.

Given that several processes can create the same pattern of spatial structure, their identification alone is insufficient to elucidate the mechanisms underlying the spatial variability. In this study, MEMs not only enabled the modelling of spatial variability in the entomofauna, but also identified the associated spatial structure at large (MEM1 and MEM3), intermediate (MEM5 and MEM6) and small (MEM8 and MEM14) scales. Furthermore, multiscale ordination analysis showed that environmental influences, represented by leaves' chemical predictors and soil nutrients, are not only invariant but also congruent with the assumption of mean and variance stationarity between samples and of independence of residuals in multivariate regressions.

ACKNOWLEDGEMENTS

The authors wish to thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for its financial support (# 401188-2016-7).

SUPPORTING INFORMATION

Additional supporting information (R script and dataset) may be found at http://dx.doi.org/10.13140/RG.2.2.18335.92320

REFERENCES

ABDALA-ROBERTS et al. Bottom-up control of geographic variation in insect herbivory on wild cotton (*Gossypium hirsutum*) by plant defenses and climate. **American Journal of Botany**, v. 106, n. 8, p. 1059-1067, 2019.

BAUMAN, D. et al. Optimizing the choice of a spatial weighting matrix in eigenvector-based methods. **Ecology**, v. 99, n. 10, p. 2159-2166, 2018.

BISPO, P. C.; OLIVEIRA, L. G. Diversity and structure of Ephemeroptera, Plecoptera and Trichoptera (Insecta) assemblages from riffles in mountain streams of Central Brazil. **Revista Brasileira de Zoologia**, v. 24, n. 2, p. 283-293, 2007.

BIVAND, R.; PIRAS, G. Comparing Implementations of Estimation Methods for Spatial Econometrics. **Journal of Statistical Software**, v. 63, n. 18, 1-36, 2015.

BRIND'AMOUR, A. et al. Application of Moran's Eigenvector Maps (MEM) to irregular sampling designs. **Spatial Statistics**, v. 26, p. 56-68, 2018.

COSTA, D. P. et al. Effects of soil, spatial parameters and leaf phenolic contents on entomofauna variability in pequizeiro. In TULLIO, L. (Org.) **Características dos solos e sua interação com as plantas**. Atena: Ponta Grossa, Cap. 4, p. 28-42, 2019.

COVELO, F. et al. Temporal changes in the spatial pattern of leaf traits in a *Quercus robur* population. **Annals of Forest Science**, v. 68, n. 3, p. 453-460, 2011.

CUNNINGHAM, S. A. et al. Evolutionary divergences in leaf structure and chemistry, comparing rainfall and soil nutrient gradients. **Ecological Monographs,** v. 69, n. 4, p. 569-588, 1999.

DUNGAN, J. et al. A balance view of scale in spatial analysis. **Ecography**, v. 25, n. 5, p. 626-640, 2002.

DRAY, S. et al. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). **Ecological Modelling**, v. 196, n. 3/4, p. 483-493, 2006.

DRAY, S. et al. **Adespatial: multivariate multiscale spatial analysis**. R package version 0.3-7, 2019. https://CRAN.R-project.org/package=adespatial.

FERREIRA, G. A. et al. Biodiversidade de insetos em Pequizeiro (*Caryocar brasiliense*, Camb.) no Cerrado do Estado de Goiás, Brasil. **Agrociência**, v. 13, n. 2, p. 14-31, 2009.

JOMBART, T. et al. Finding essential scales of spatial variation in ecological data: a multivariate approach. **Ecography**, v. 32, n. 1, p. 161-168, 2009.

LEGENDRE, P.; LEGENDRE, L. **Numerical ecology**. 3th ed. Elsevier: Amsterdam, 2012.

LEPS, J.; SMILAUER, P. **Multivariate analysis of ecological data using Canoco**. Cambridge University Press: Cambridge, 2007.

MEAUX, J.; MITCHELL-OLDS, T. Evolution of plant resistance at the molecular level: ecological context of species interactions. **Heredity**, v. 91, n. 4**,** p. 345-352, 2003.

MEINDL, G. A. et al. Edaphic factors and plant–insect interactions: direct and indirect effects of serpentine soil on florivores and pollinators. **Oecologia**, v. 173, n. 4, p. 1355-1366, 2013.

MUTIKAINEN, P. et al. Herbivore resistance in *Betula pendula*: effect of fertilization, defoliation, and plant genotype. **Ecology**, v. 81, n. 1, p. 49-65, 2000.

OKSANEN, J. et al **Vegan: community ecology package**. R version 2.4-5, 2017. https://CRAN.R-project.org/package=vegan.

R CORE TEAM. **R: a language and environment for statistical computing**. R Foundation for Statistical Computing: Vienna, version 3.4.4, 2018.

TORRES, F. R. & MADI-RAVAZZI, L. Seasonal variation in natural populations of *Drosophila* spp. (Diptera) in two woodlands in the State of São Paulo, Brazil. **Iheringia: Série Zoologia**, v. 96, n. 4, p. 437-444, 2006.

WAGNER, H. H. Direct multi-scale ordination with canonical correspondence analysis. **Ecology**, v. 85, n. 2, p. 342-351, 2004.

WARTON, D. I. et al. Distance-based multivariate analyses confound location and dispersion effects. **Methods in Ecology and Evolution**, v. 3, n. 1, p. 89-101, 2012.

ZEQUI, J. A. C. et al. Aquatic insect communities in small stream in the south of Brazil. **Environmental Monitoring and Assessment**, v. 191, n. 7, p. 408, 2019.

ÍNDICE REMISSIVO

Atena Editora
2020