

Métodos multidimensionais baseados em modelos

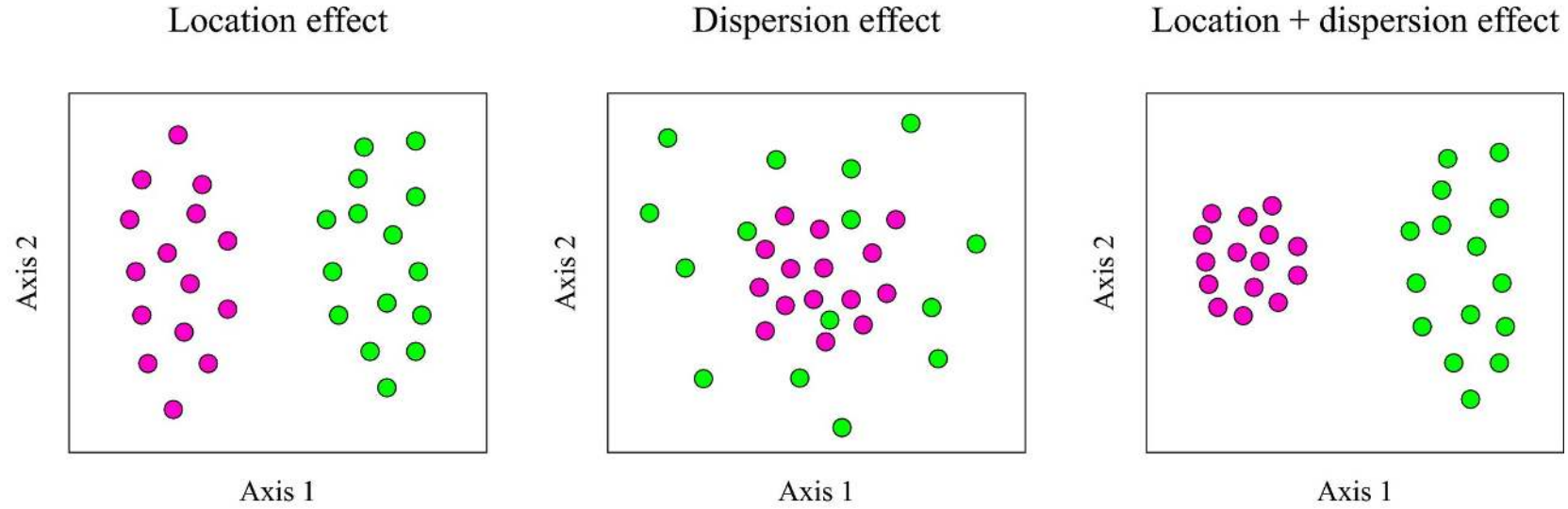
Tipos de métodos

- Agrupamento
 - EcoMix
- Ordenação
 - ecoCopula
 - boral
 - gllvm
 - Constrained quadratic ordination (CQO)
- Joint-species distribution models
- Inferencial
 - HMSC
 - mvabund

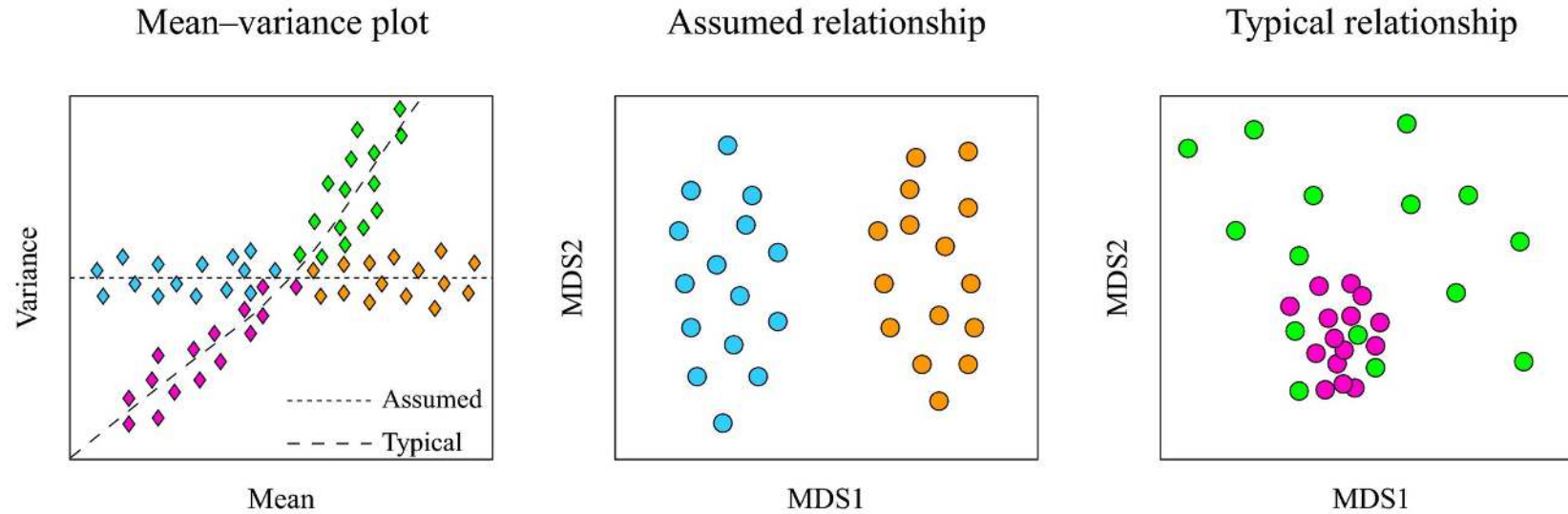
Motivação

- Problemas de abordagens baseadas em distância em lidar com a relação média-variância da abundância
 - Confundem posição e dispersão (multivariada)
- Mal desempenho com grandes conjuntos de dados
 - eDNA, microbioma, metabarcoding etc
- Utiliza mvGLMs para ajustar modelos para cada espécie individualmente
 - Utiliza alguma forma de aleatorização para gerar estatística do teste
 - Permite testar o efeito das covariáveis em cada espécie
 - Permite realizar previsões utilizando coeficientes do modelo

(A) ————— Types of between-group effects —————



(B) ————— Influence of mean–variance relationship on ordination —————



Motivação

- Problemas de abordagens baseadas em distância em lidar com a relação média-variância da abundância
 - Confundem posição e dispersão (multivariada)
- Mal desempenho com grandes conjuntos de dados
 - eDNA, microbioma, metabarcoding etc
- Utiliza mvGLMs para ajustar modelos para cada espécie individualmente
 - Utiliza alguma forma de aleatorização para gerar estatística do teste
 - Permite testar o efeito das covariáveis em cada espécie
 - Permite realizar previsões utilizando coeficientes do modelo

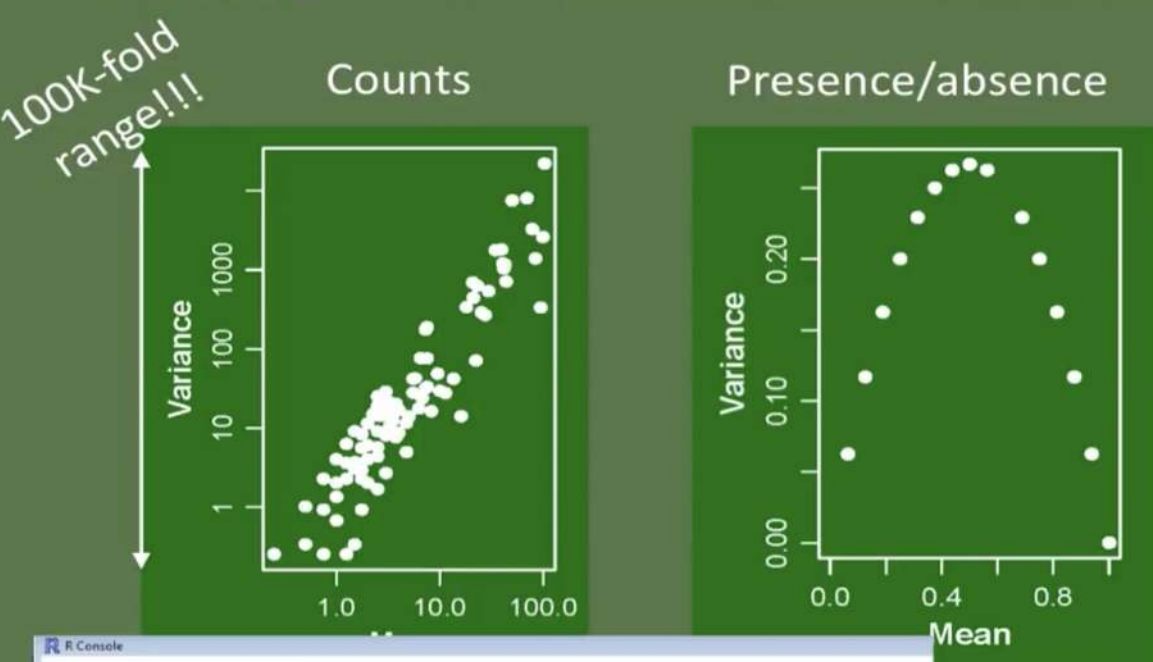
O que são métodos baseados em modelos?

- Métodos estatísticos paramétricos que capturam propriedades chaves de dados multivariados
 - Abundância
 - GLMs
 - Variáveis correlacionadas
 - Reamostragem
 - Fatores aleatórios

Distance-based multivariate analyses confound location and dispersion effects

David I. Warton^{1*}, Stephen T. Wright¹ and Yi Wang^{1,2}

2. Strong mean-variance relationship



```
> meanvar.plot(copepods~block)
```

- Métodos “tradicionais” baseados em distância (algoritmos) focam apenas nas espécies que têm maior variância
- Ignoram propriedade da *média-variância* da abundância (zeros e SADs)

2015

Review

So Many Variables: Joint
Modeling in Community
Ecology

David I. Warton,^{1,*} F. Guillaume Blanchet,² Robert B. O'Hara,³
Otso Ovaskainen,^{4,5} Sara Taskinen,⁶ Steven C. Walker,² and
Francis K.C. Hui⁷

Plant Ecol (2015) 216:669–682

DOI 10.1007/s11258-014-0366-3

Model-based thinking for community ecology

**David I. Warton · Scott D. Foster ·
Glenn De'ath · Jakub Stoklosa · Piers K. Dunstan**

FORUM

Distance, dissimilarity, and mean–variance ratios in ordination

David W. Roberts

- Transformação pelas linhas não resolve o problema
- Mesmo em dados que não têm grandes diferenças de abundância total entre locais (diferentes apenas na posição), métodos baseados em distância confundem posição com variância e dizem que há diferença entre locais

- Resultados de Warton et al. 2012 foram enviesados por simularem comunidades com alta diferença de abundância e igualmente usarem um conjunto de dados que tem essa característica (corais)
- Usam dados de abundância não transformados
- Transformação dos dados resolve o problema

FORUM

The central role of mean-variance relationships in the analysis of multivariate abundance data: a response to Roberts (2017)

David I. Warton^{*1}  and Francis K. C. Hui²

Vantagens de abordagens baseadas em modelos

- Controlar propriedades espúrias dos dados
- Diagnose do modelo
- Seleção de modelos e inferência (AIC/BIC)
 - Medidas de ajuste (Log likelihood) e 95% IC
 - Usa AIC para selecionar o melhor número de cluster ou de variáveis latentes (eixos pra representar os dados)
- Velocidade para lidar com grandes quantidade de dados
- Eficiência
 - Usa uma distribuição de probabilidade que permite modelar explicitamente o processo que gerou os dados
- Controla para diferença de esforço de amostragem entre locais (offset=site)

Methods in Ecology and Evolution



Methods in Ecology and Evolution 2015, 6, 399–411

doi: 10.1111/2041-210X.12236

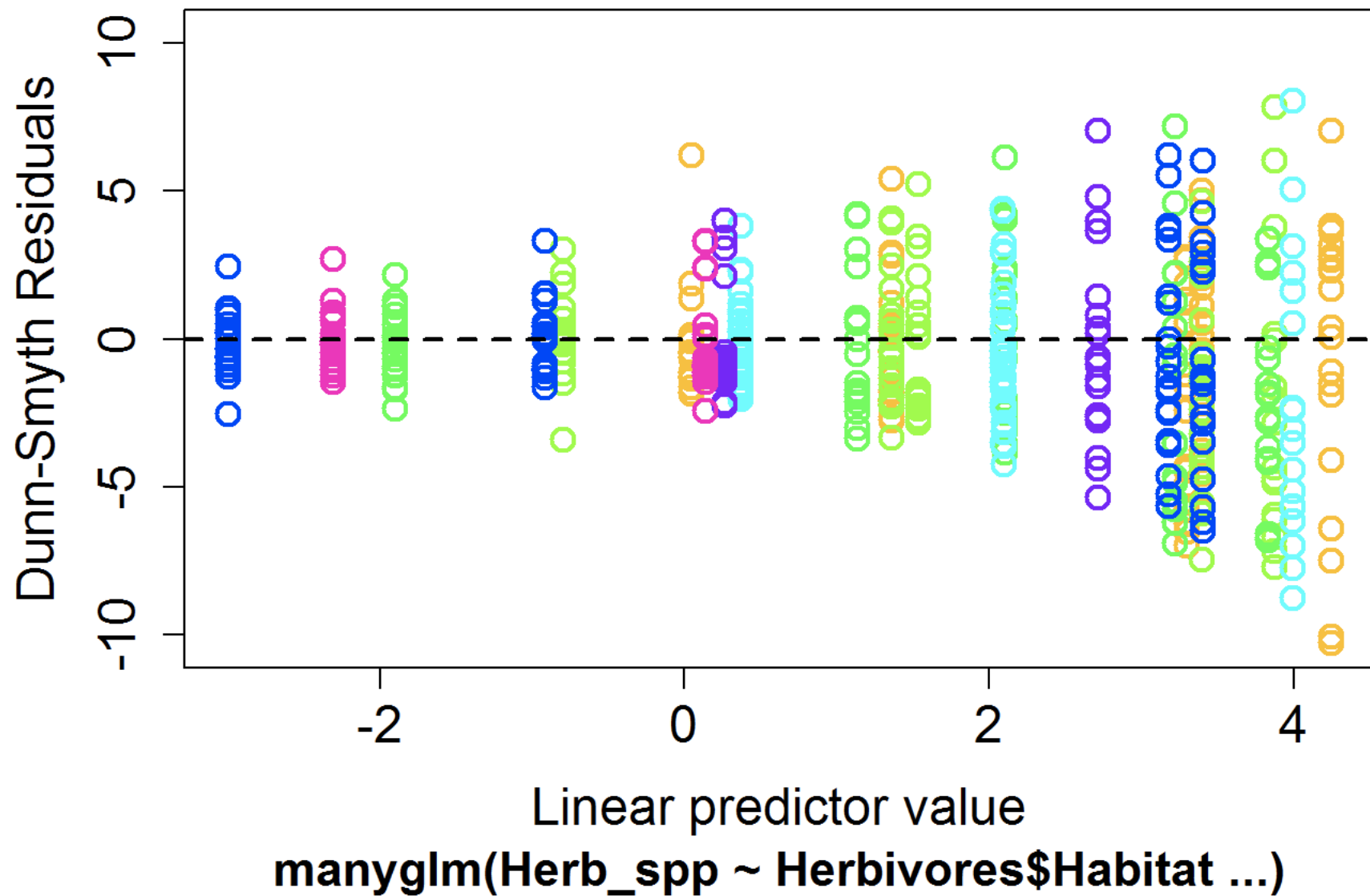
SPECIAL FEATURE

NEW OPPORTUNITIES AT THE INTERFACE BETWEEN ECOLOGY AND STATISTICS

Model-based approaches to unconstrained ordination

Francis K.C. Hui^{1,2,*}, Sara Taskinen³, Shirley Pledger⁴, Scott D. Foster² and David I. Warton^{1,5}

Residuals vs Fitted



Problemas de abordagens baseadas em modelos

- Impossível fazer previsões, porque não são modelos estatísticos formais
- Transformações (e.g., Hellinger) não lidam com distribuições extremamente enviesadas (zero inflated e forte dominância)

Métodos para Agrupamento

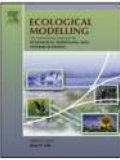
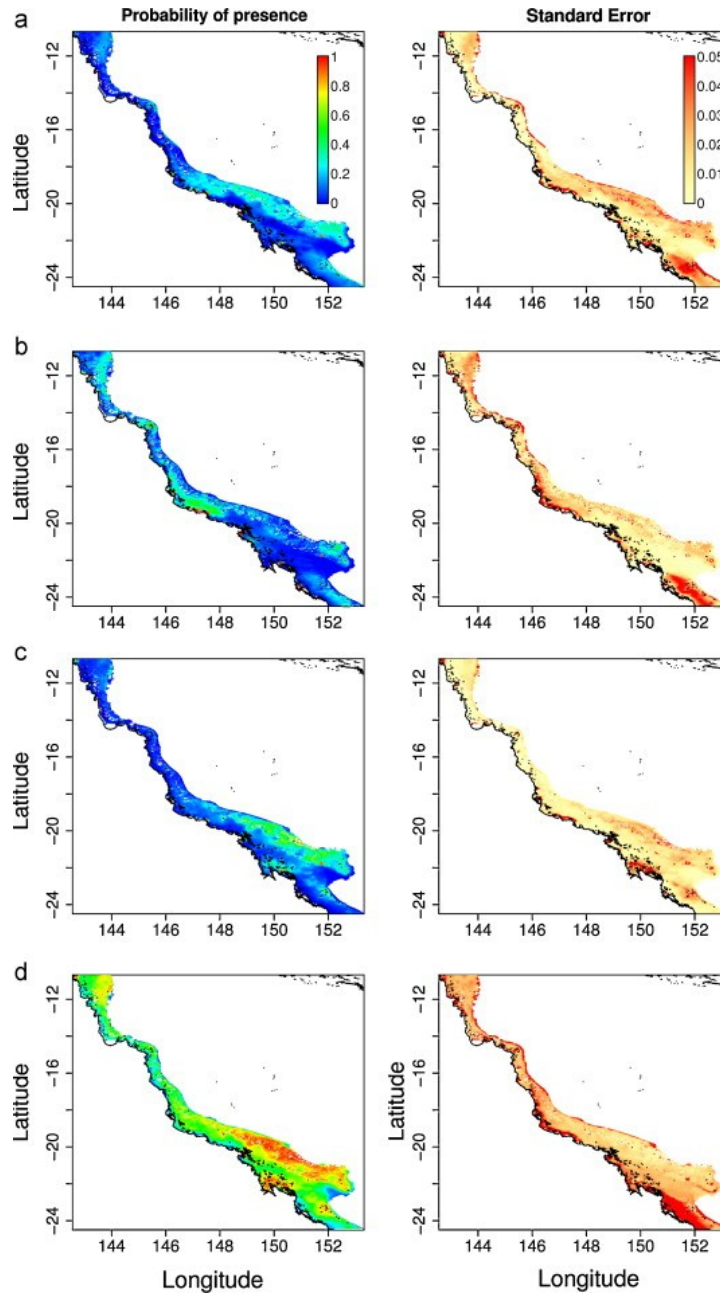
Species Archetype Models (SAMs)



Agrupa espécies baseado na sua resposta à variáveis ambientais



Útil para encontrar descontinuidades em eco/bioregiões



Model based grouping of species across environmental gradients

Piers K. Dunstan^{a,*}, Scott D. Foster^b, Ross Darnell^b

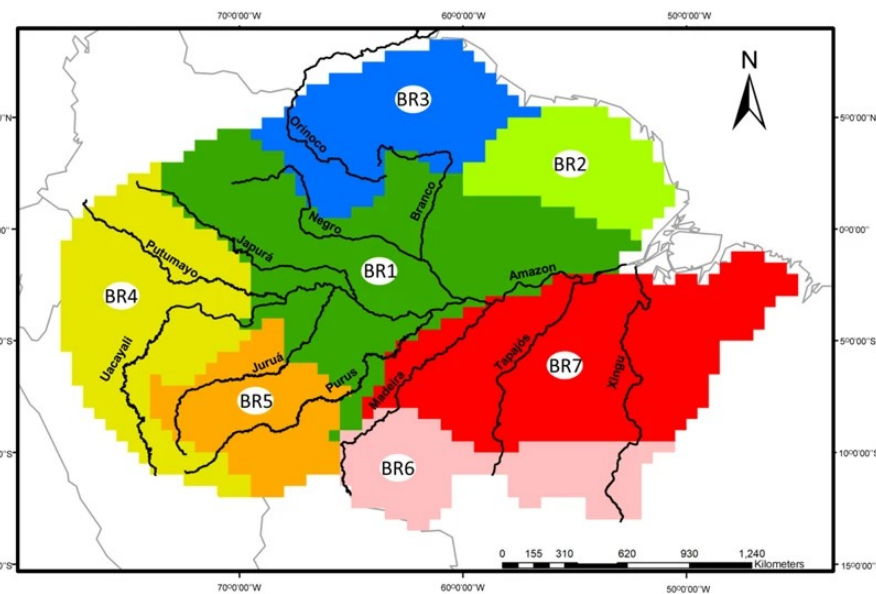
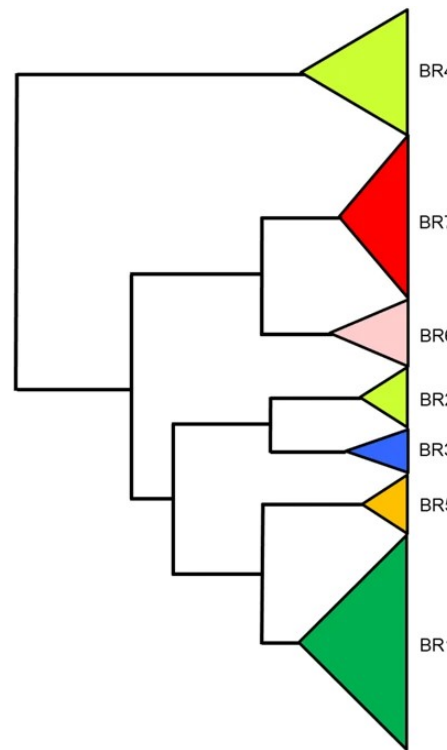
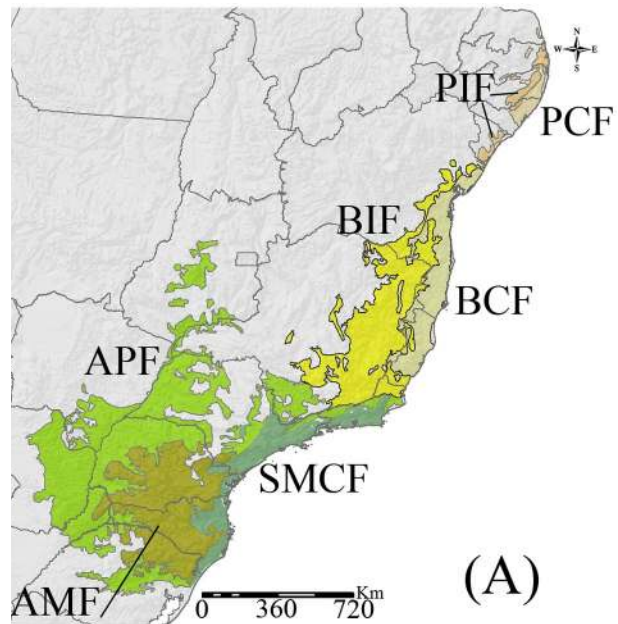
Ecology, 94(9), 2013, pp. 1913–1919
© 2013 by the Ecological Society of America

To mix or not to mix: comparing the predictive performance of mixture models vs. separate species distribution models

FRANCIS K. C. HUI^{1,2,5} DAVID I. WARTON^{1,3} SCOTT D. FOSTER^{2,4} AND PIERS K. DUNSTAN⁴

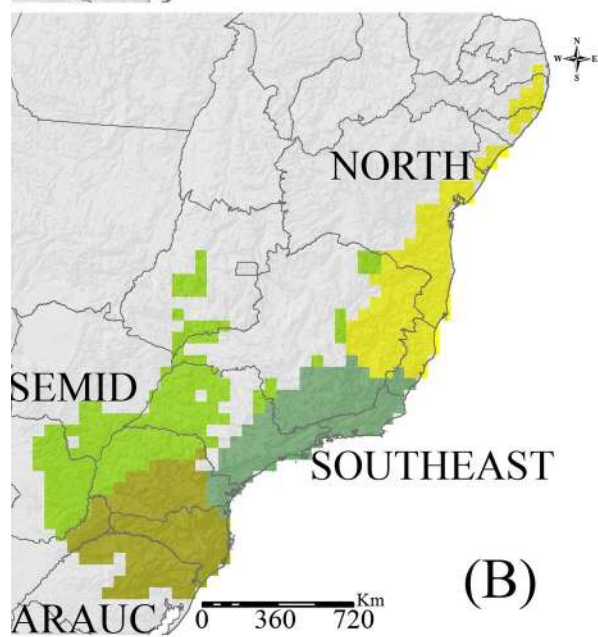
Pacote **ecoMix**

Distribuição dos 4 diferentes arquétipos encontrados



Godinho & da Silva 2018 Sci Rep

Vasconcelos et al.
2014 PLoS ONE

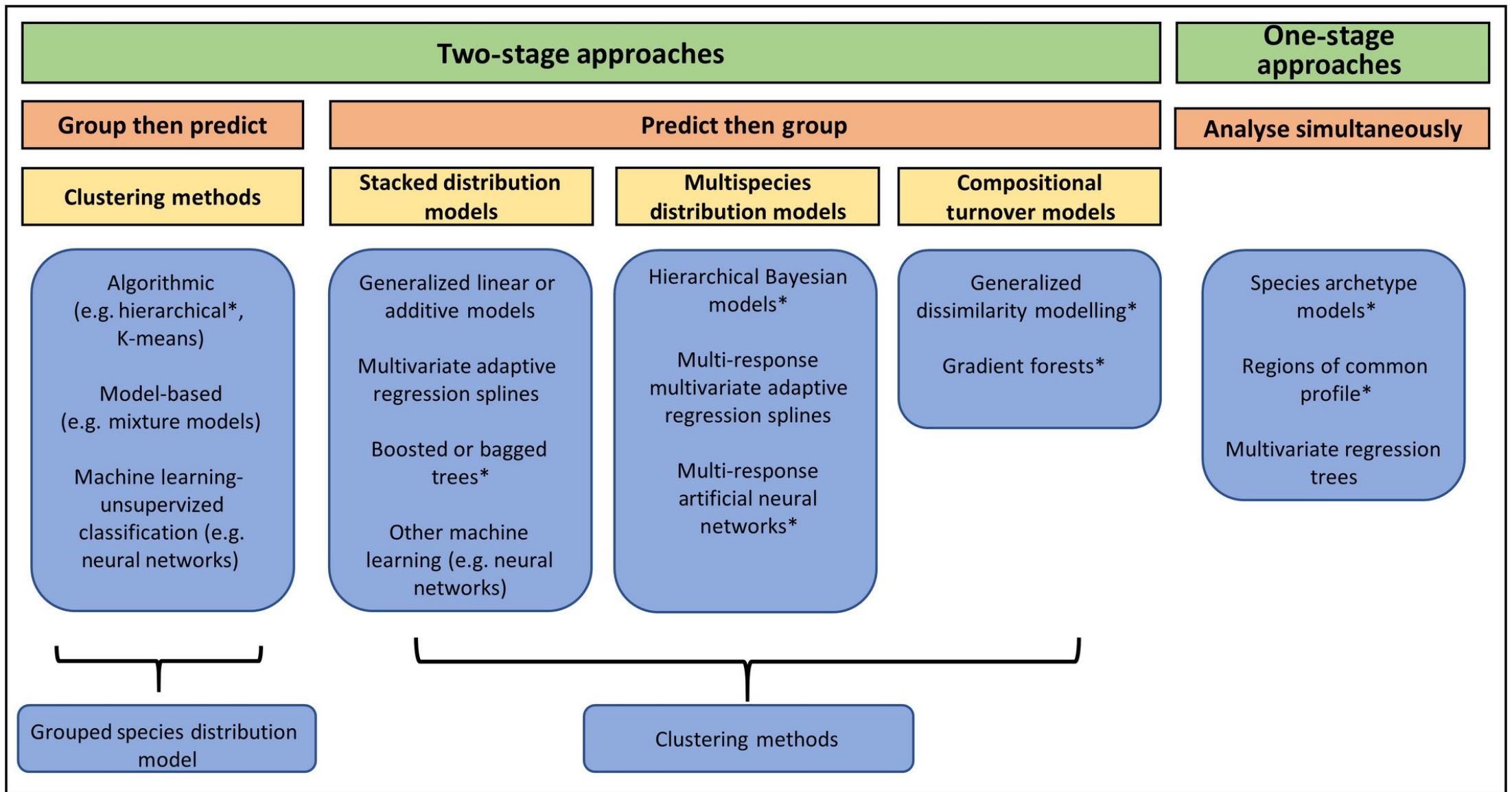


(B)

RESEARCH ARTICLE

Determining marine bioregions: A comparison of quantitative approaches

Nicole Hill¹  | Skipton N. C. Woolley²  | Scott Foster²  | Piers K. Dunstan²  |
John McKinlay³ | Otso Ovaskainen⁴  | Craig Johnson¹ 



Conclusões

- Métodos que analisam explicitamente a resposta de espécies ao ambiente para encontrarem grupos fornecem melhores resultados do que métodos que agrupam primeiro e depois predizem, ou predizem primeiro e depois agrupam
- Permite incluir informações sobre as regiões

Métodos para Ordenação

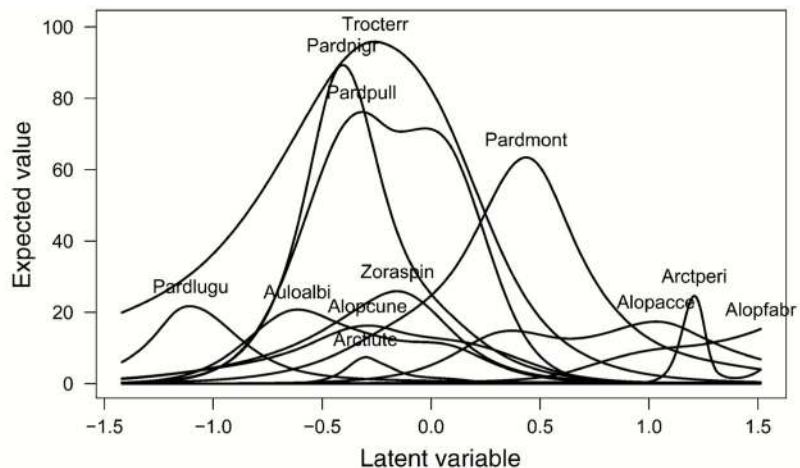


FIG. 2. Perspective plot of hunting spiders rank-1 Poisson constrained additive ordination (CAO) model. This is Fig. 1 combined into one plot and uses the original response scale. The species are labeled with the following abbreviations: Alopacce, *Alopecosa accentuata*; Alopuncne, *Alopecosa cuneata*; Alopfabr, *Alopecosa fabrilis*; Arctlute, *Arctosa lutetiana*; Arctperi, *Arctosa perita*; Auloalbi, *Aulonia albimana*; Pardlugu, *Pardosa lugubris*; Pardmont, *Pardosa monticola*; Pardnigr, *Pardosa nigriceps*; Pardpull, *Pardosa pullata*; Trocterr, *Trochosa terricola*; Zoraspin, *Zora spinimana*.

- Baseado em Vector generalised additive models (VGAMs)
 - Não é restrito à família exponencial
 - Pode modelar outros parâmetros além da média
 - Assim como nMDS, requer que o usuário especifique o número de dimensões a priori

A NEW TECHNIQUE FOR MAXIMUM-LIKELIHOOD CANONICAL GAUSSIAN ORDINATION

THOMAS W. YEE¹

CONSTRAINED ADDITIVE ORDINATION

THOMAS W. YEE¹

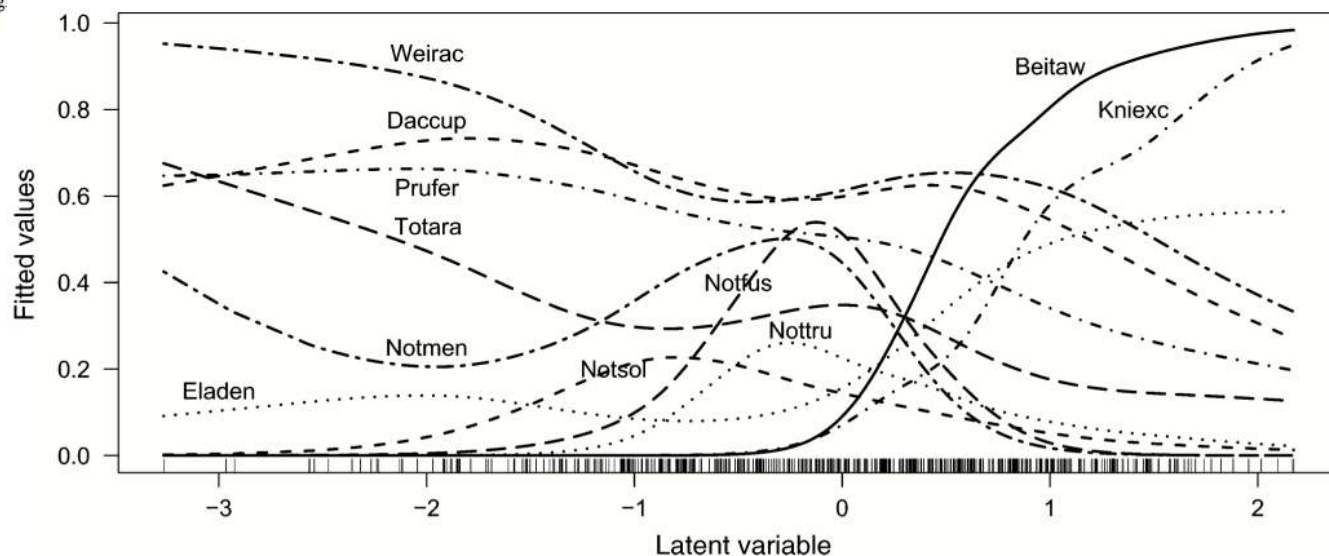


FIG. 3. Constrained additive ordination (CAO) diagram of 11 native New Zealand tree species (presence/absence data). There are 500 sites, and each species has three nonlinear degrees of freedom. The rugplot gives the sites scores $\hat{\nu}_i$. The species are labeled with the following abbreviations: Beitaw, *Beilschmiedia tawa* or tawa; Daccup, *Dacrydium cupressinum* or rimu; Eladen, *Elaeocarpus dentatus* or hinau; Kniexc, *Knightia excelsa* or rewarewa; Nottfus, *Nothofagus fusca* or red beech; Notmen, *Nothofagus menziesii* or silver beech; Notsol, *Nothofagus solandri* or black beech; Notttru, *Nothofagus truncata* or hard beech; Prufer, *Prumnopitys ferruginea* or miro; Totara, *Podocarpus totara* or totara; Weirac, *Weinmannia racemosa* or kamahi.

Random-effects ordination: describing and predicting multivariate correlations and co-occurrences

STEVEN C. WALKER¹ AND DONALD A. JACKSON

Pacote *reo* (R-forge)

- Assume que abundância de espécies têm relação unimodal com gradientes ambientais
- Usa ML/LASSO para estimar parâmetros das espécies e locais ao mesmo tempo
- Só permite usar 2 dimensões

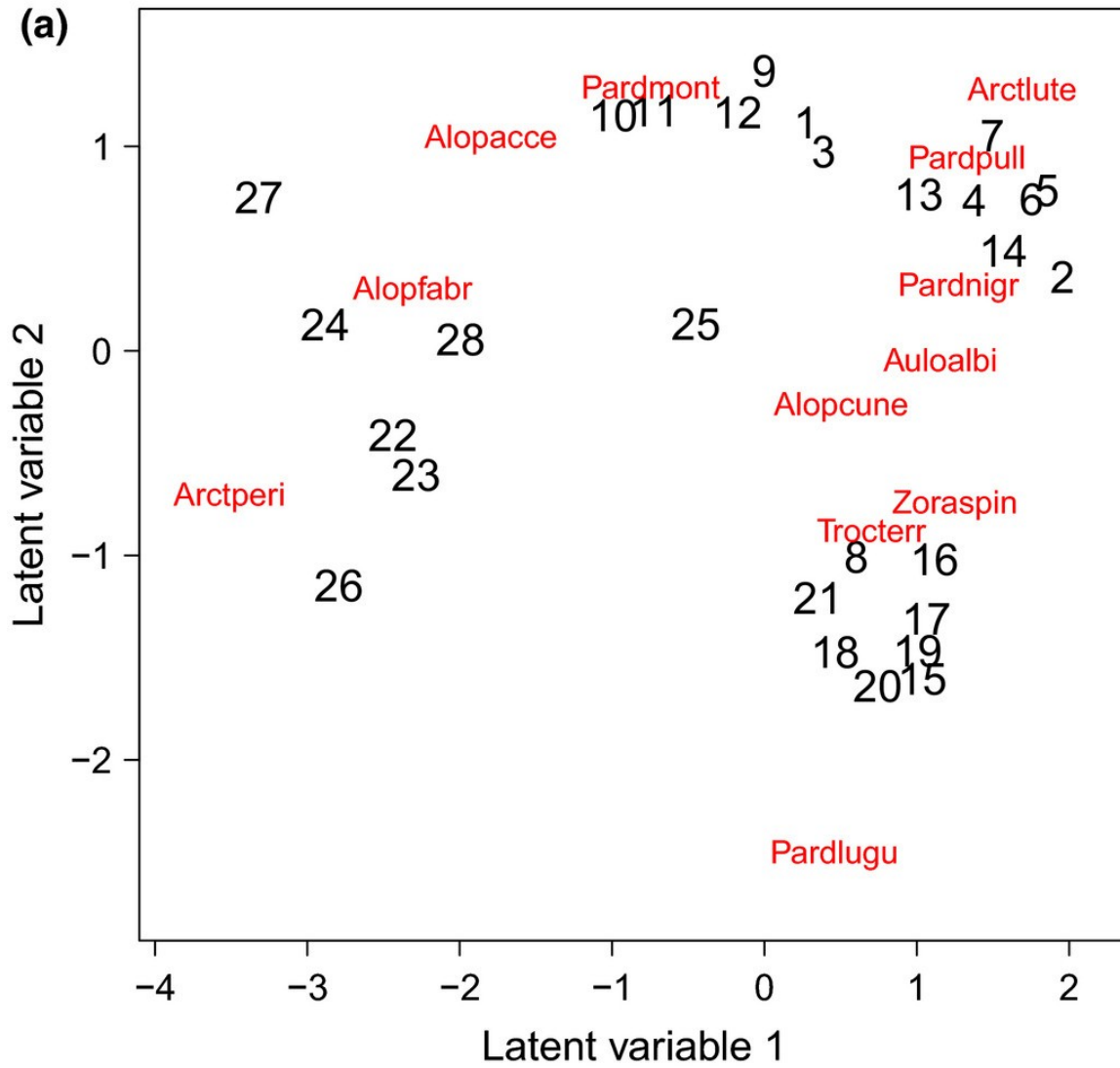
APPLICATION

BORAL – Bayesian Ordination and Regression Analysis of Multivariate Abundance Data in R

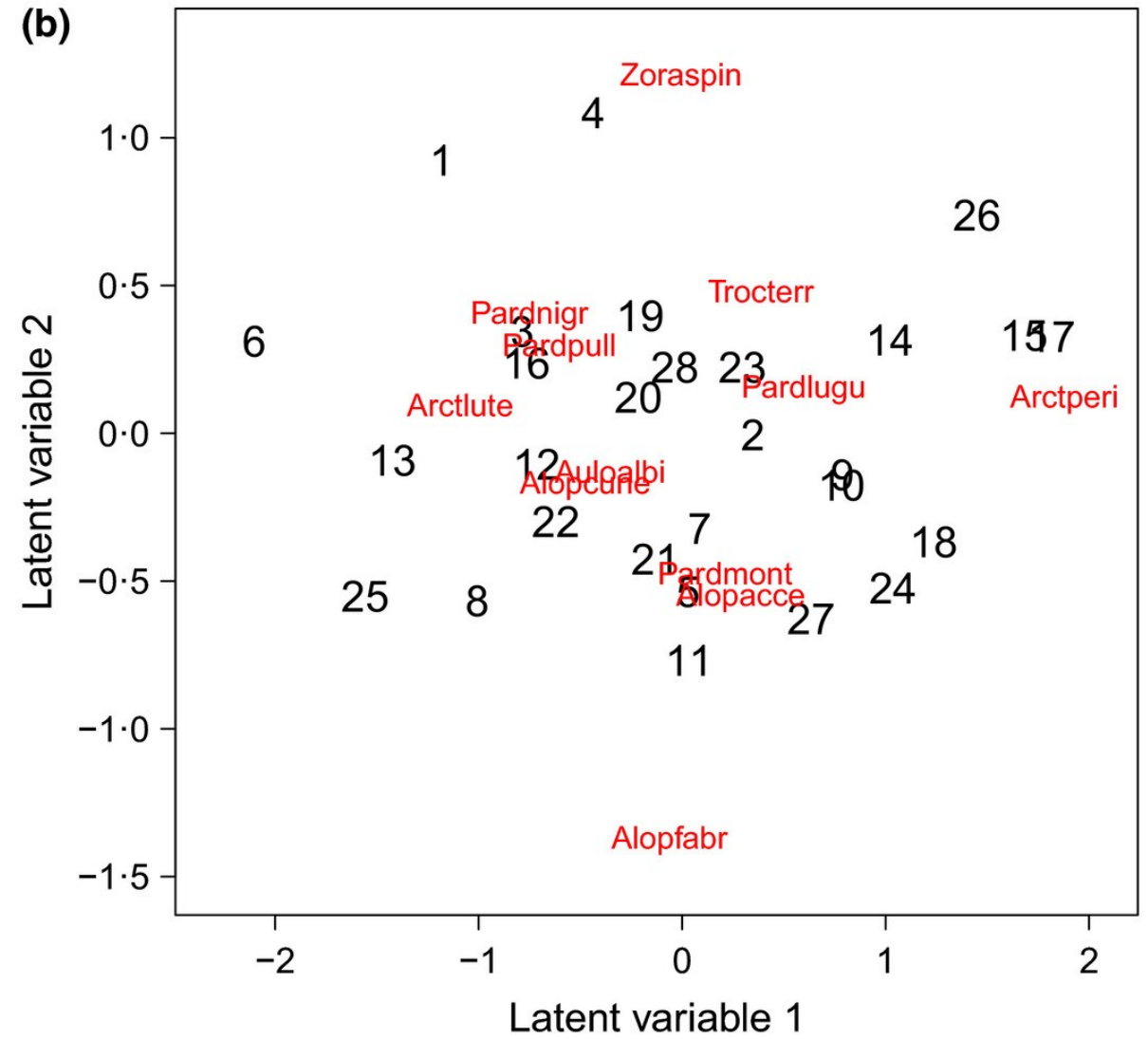
Francis K.C. Hui*

- Método de ordenação irrestrita Bayesiano
- Usa MCMC para modelar dados de abundância, incidência, ordinal (classes) e biomassa e variáveis latentes para modelar correlação entre espécies
 - Precisa de JAGS no ‘background’ para rodar a MCMC
- Permite levar em conta correlação entre espécies dado por variáveis não medidas (latentes)
 - Estima correlação entre espécies dada por similaridades na resposta à variáveis ambientais, e também ao remover respostas similares ao ambiente (variação “biótica”)
- Permite diagnose do modelo

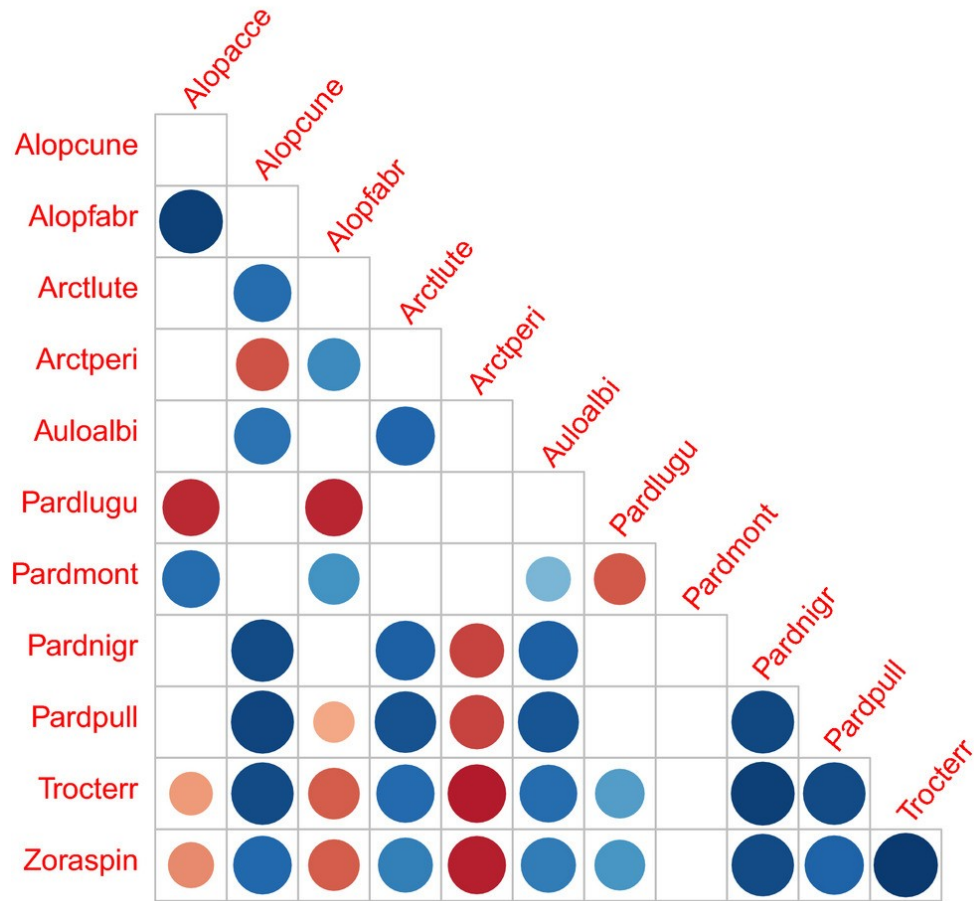
Unconstrained biplot



Residual biplot

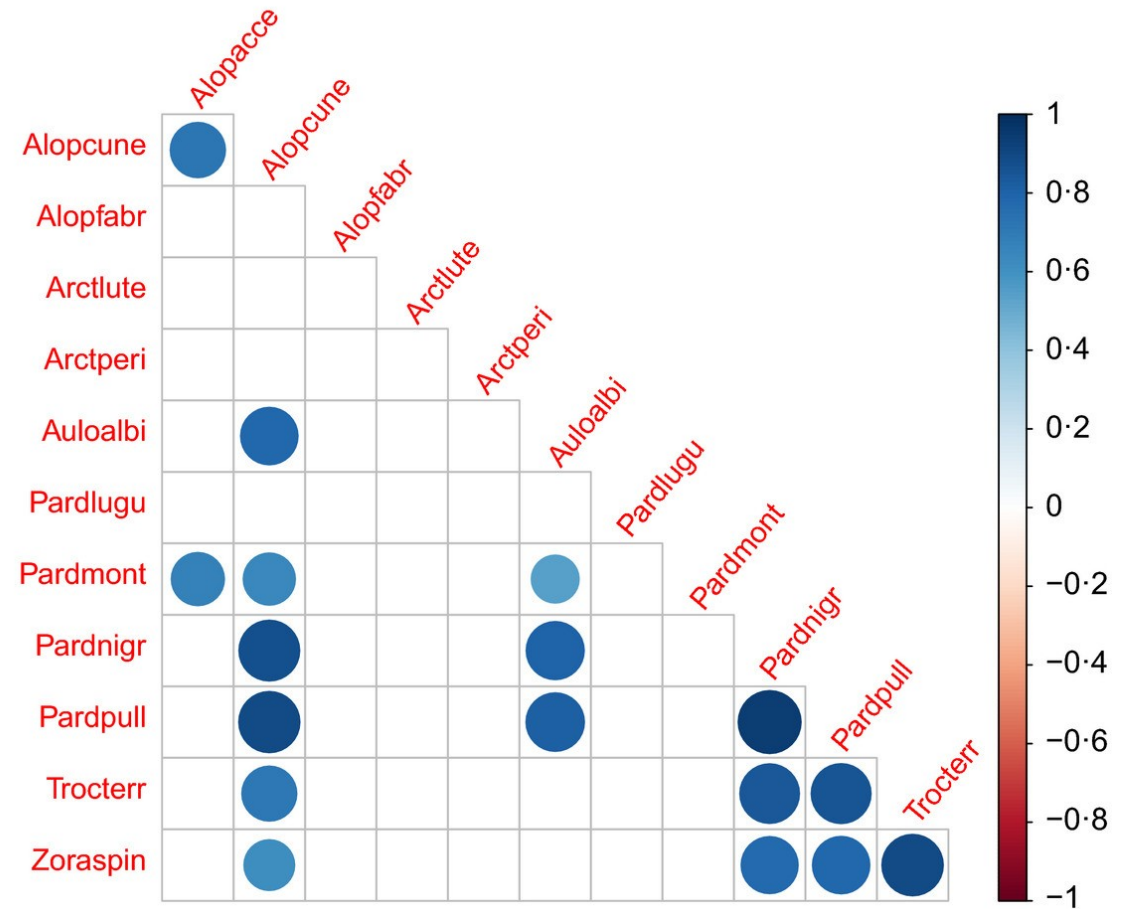


Correlations due to covariates



Correlação entre espécies dada por respostas similares a fatores ambientais

Residual correlations



Correlação entre espécies ao se remover (residual) o efeito comum do ambiente.
Correlação positiva: “facilitação” e negativa: “Competição”

Comparison of distance-based and model-based ordinations

DAVID W. ROBERTS¹

Received: 3 October 2019

Revised: 9 December 2019

Accepted: 20 December 2019

DOI: 10.1002/ece3.6059

ORIGINAL RESEARCH

Ecology and Evolution

Open Access

WILEY


Should ecologists prefer model- over distance-based multivariate methods?

Jonathan F. Jupke  | Ralf B. Schäfer 

Conclusões

- dbRDA e mvabund foram melhores do que CCA e CQO
- t-SNE é melhor que nMDS, mas os dois foram melhores que REO e BORAL para construir ordenações representando tanto espécies quando locais
- BORAL é muito mais lento comparado à COPULA

A pathway for multivariate analysis of ecological communities using copulas

Marti J. Anderson^{1,2}  | Perry de Valpine³ | Andrew Punnett² | Arden E. Miller⁴

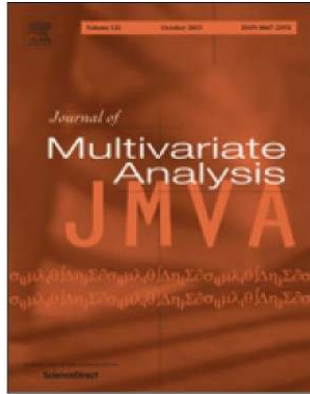
- Permite criar um modelo conjunto onde cada espécie é modelada separadamente
- Propõe uma combinação de abordagens usando ***copulas*** com aquelas baseadas em dissimilaridade
- Possibilidade de melhor representar incertezas em ordenações e modelos preditivos
- Bastante usada em Economia há vários anos, mas só recentemente aparecem em aplicações de ecologia



Contents lists available at [ScienceDirect](#)

Journal of Multivariate Analysis

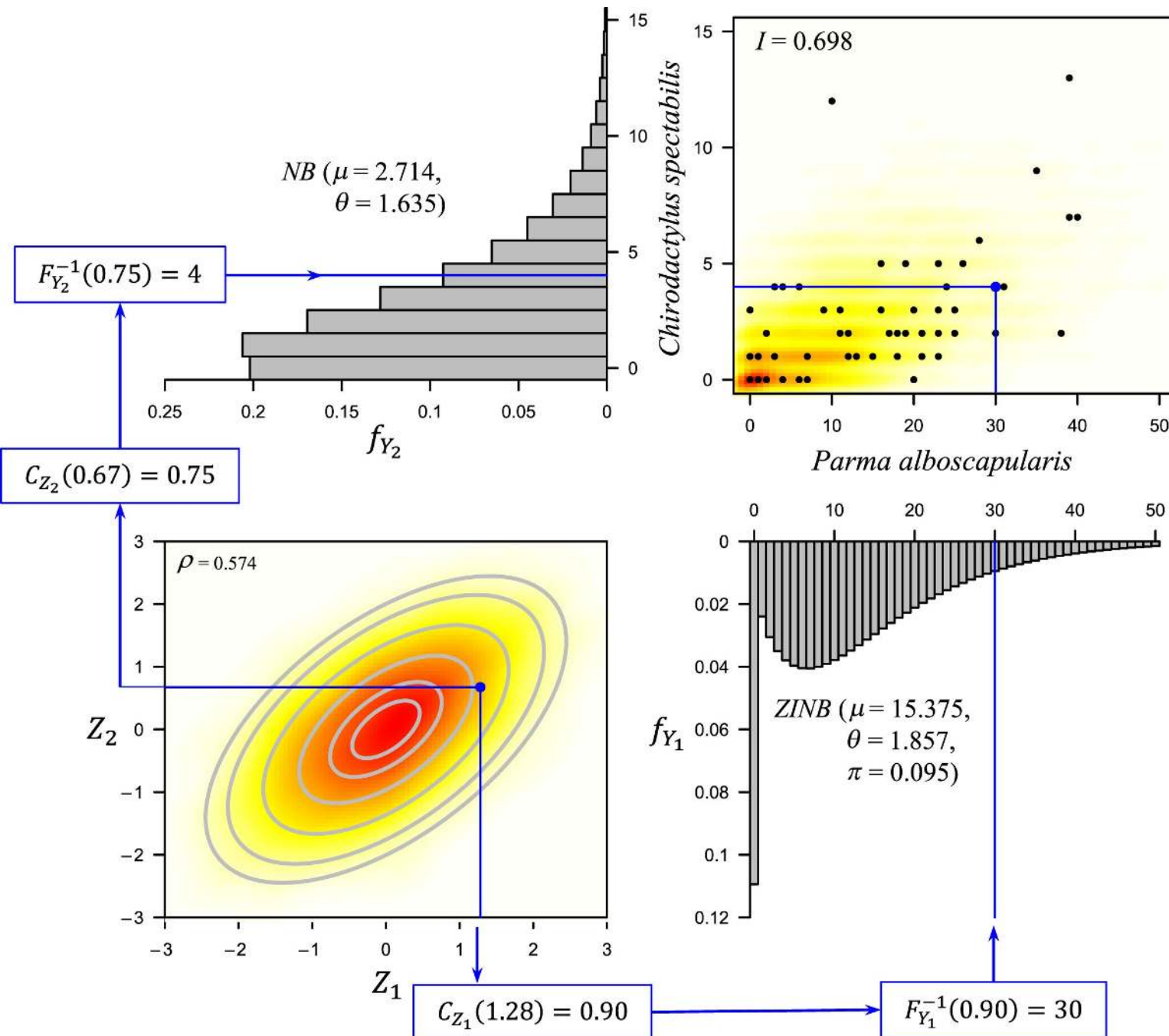
journal homepage: www.elsevier.com/locate/jmva



A general algorithm for covariance modeling of discrete data

Gordana C. Popovic^{a,*}, Francis K.C. Hui^b, David I. Warton^{a,c}

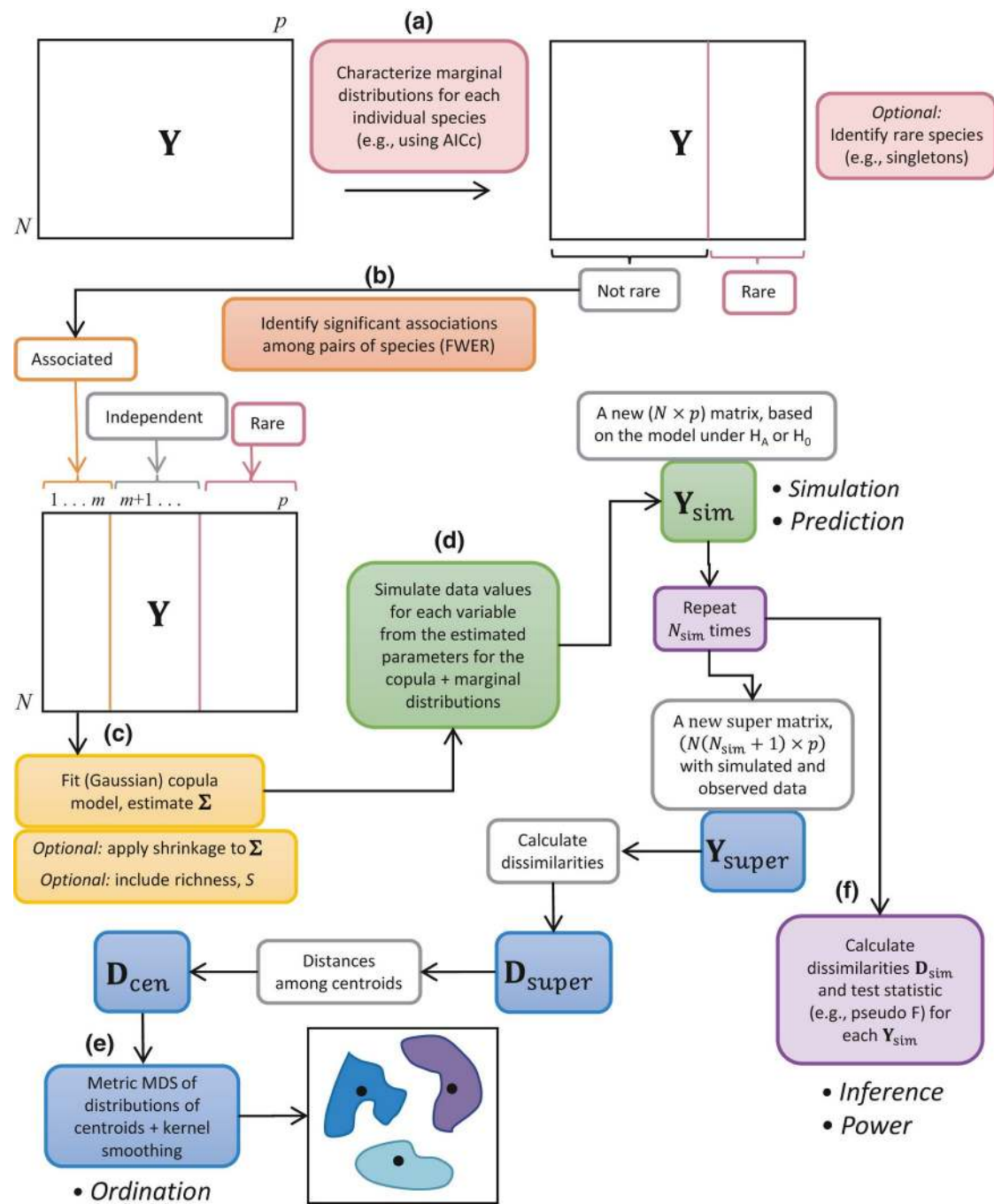


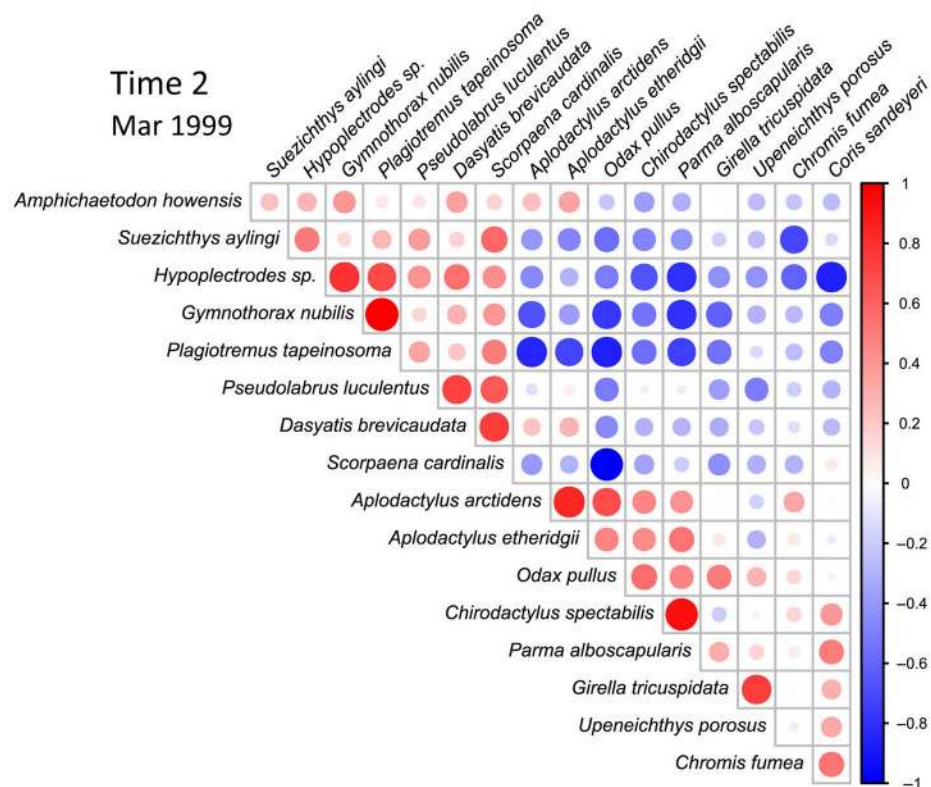
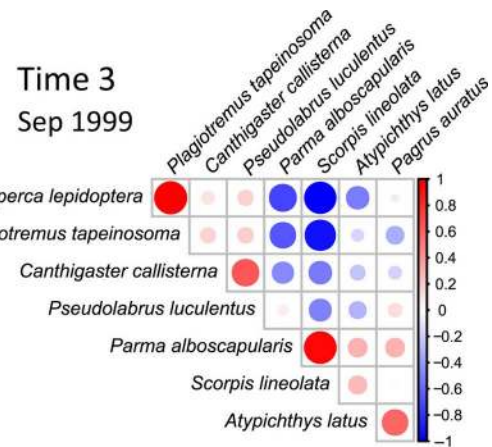
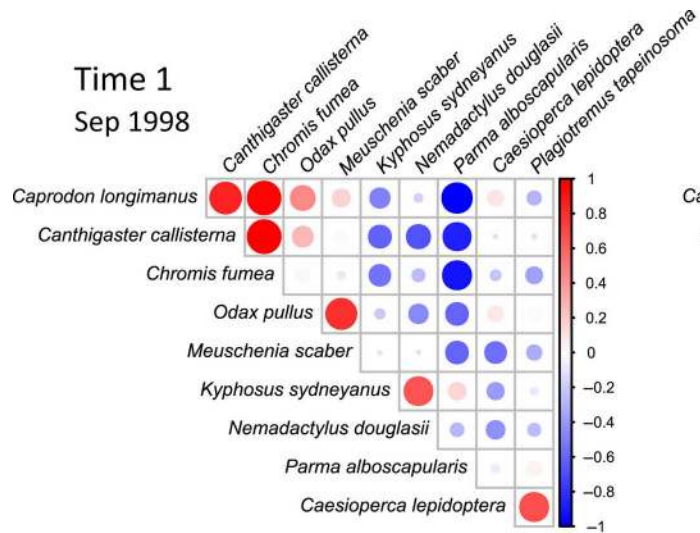


Acoplamento de duas variáveis quaisquer via função de distribuição cumulativa

Distribuição de abundância de duas espécies de peixes, cada uma modelada usando uma distribuição diferente (binomial negativa e binomial negativa zero-inflada) e depois combinadas para obter a distribuição conjunta

Permite calcular um índice de associação interespecífica (I)





Também permite estimar estruturas de correlação entre espécies



São estimadas separadamente da relação entre variáveis, diferentemente de abordagens de usam variáveis latentes (e.g., gllvm, boral)

Received: 28 March 2021

Accepted: 7 September 2021

DOI: 10.1111/2041-210X.13733

RESEARCH ARTICLE

Methods in Ecology and Evolution  BRITISH
ECOLOGICAL
SOCIETY

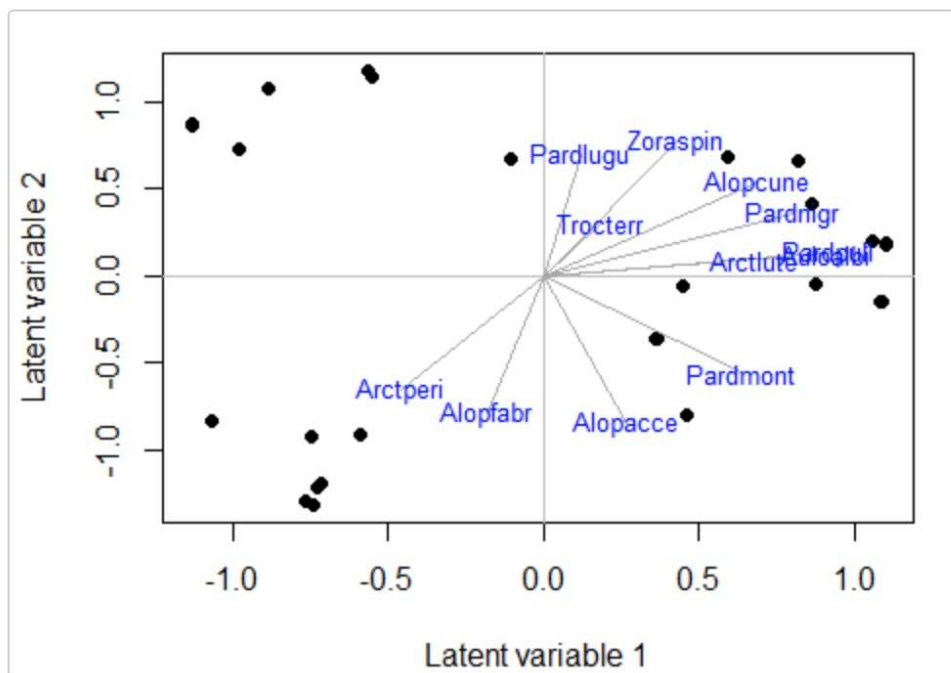
Fast model-based ordination with copulas

Gordana C. Popovic¹  | Francis K. C. Hui²  | David I. Warton¹ 

Ordination

To plot an ordination biplot for presence-absence data `pa`, we first fit a marginal model using `stackedsdm` or `manyglm` with `family="binomial"`, then use the copula ordination function `cord` on the resulting object to get the output.

```
# fit marginal model
spider_pa <- stackedsdm(pa,~1, data = X, family="binomial", ncores = 2) #equiv. many
# fit copula ordination
spid_lv=cord(spider_pa)
# biplot
plot(spid_lv,biplot = TRUE)
```

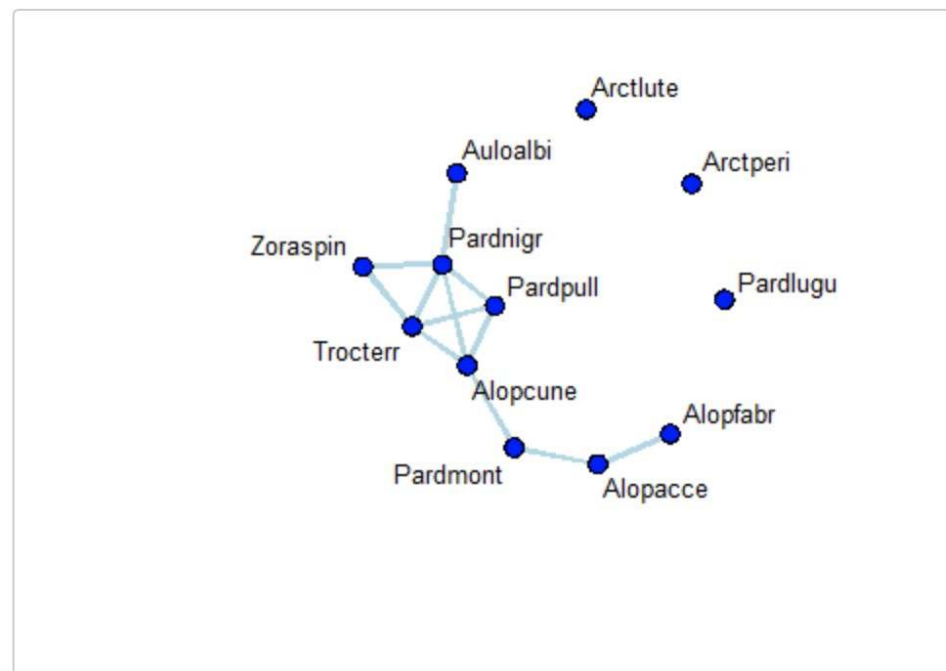


To colour the sites by a predictor, we can create a colour variable.

Graphical model

To plot a *graph* of the spider counts `abund`, we again first fit a marginal model using `stackedsdm` or `manyglm` with `family="negative.binomial"` (for counts). For graphical models it makes the most sense to first control for environmental variables, so that associations between taxa control for those. Then we use the copula graph function `cgr` on the resulting object, and `plot` the output.

```
# fit marginal model
spider_nb <- stackedsdm(abund,~, data = X, family="negative.binomial", ncores = 2) #equiv.
# fit copula ordination
spid_gr=cgr(spider_nb, seed=3)
# biplot
plot(spid_gr, pad=1)
```



APPLICATION

gllvm: Fast analysis of multivariate abundance data with generalized linear latent variable models in R

Jenni Niku¹  | Francis K. C. Hui² | Sara Taskinen¹ | David I. Warton³ 

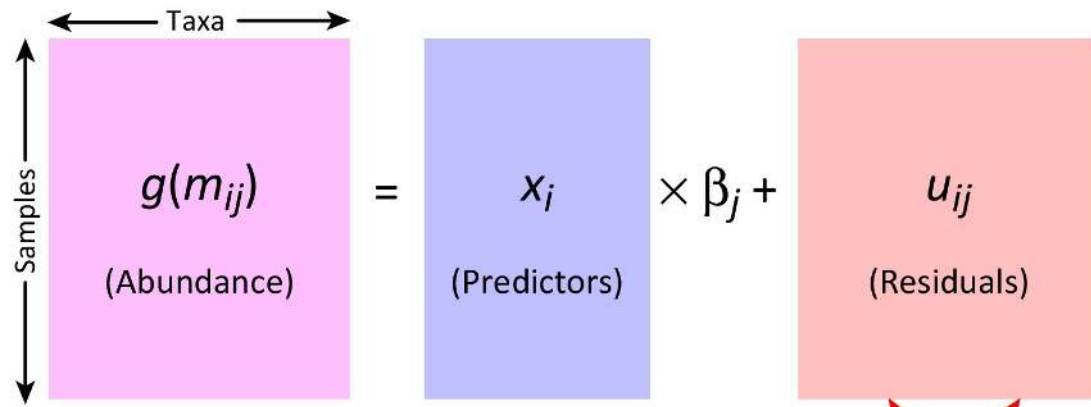
mvabund usa GLM pra modelar espécies ~ ambiente e GEE pra modelar $\text{vcov}(\text{espécies})$

Usa uma estratégia diferente para modelar a correlação entre espécies

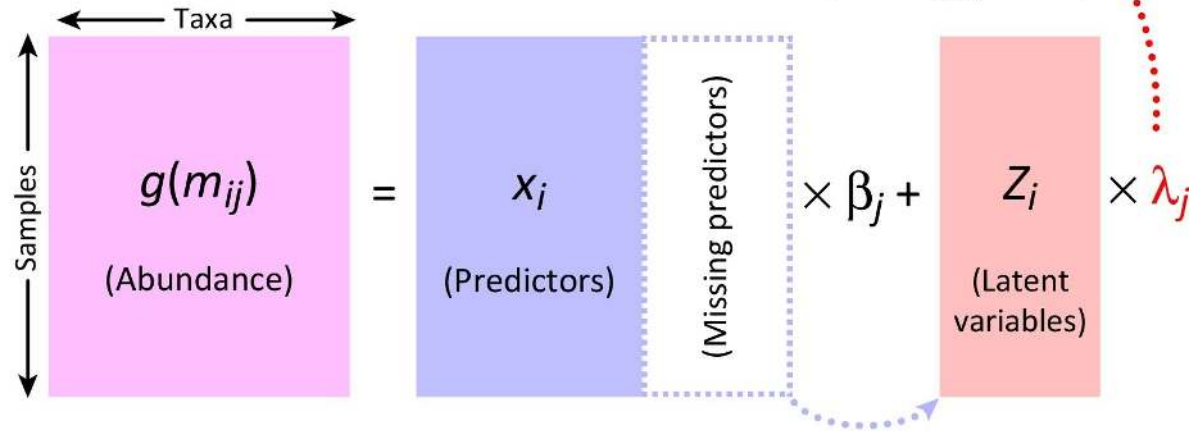


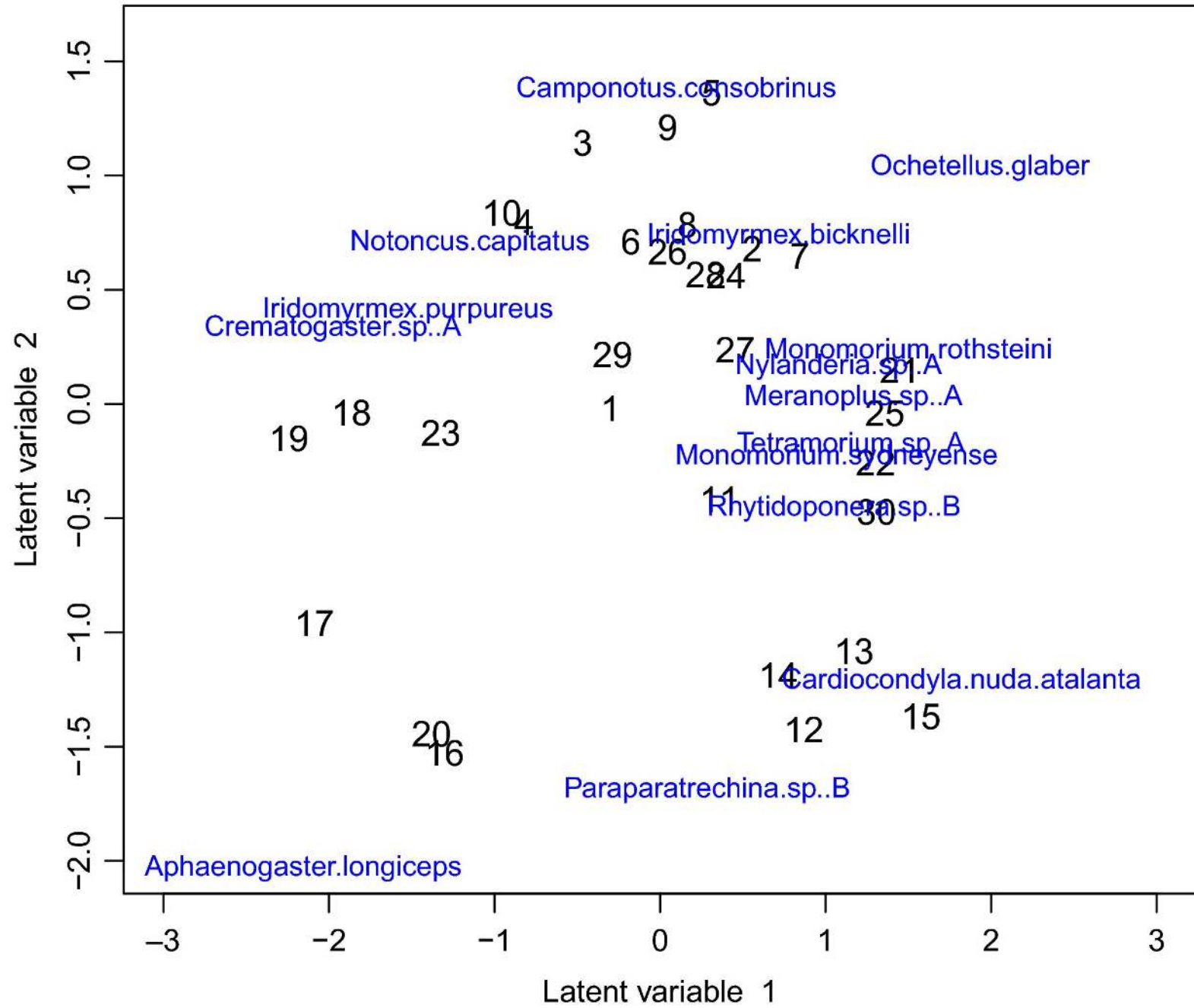
Inclui variável latente aleatória como preditora num GLM

(A) Multivariate generalised linear mixed model (GLMM)



(B) Latent variable model (LVM)

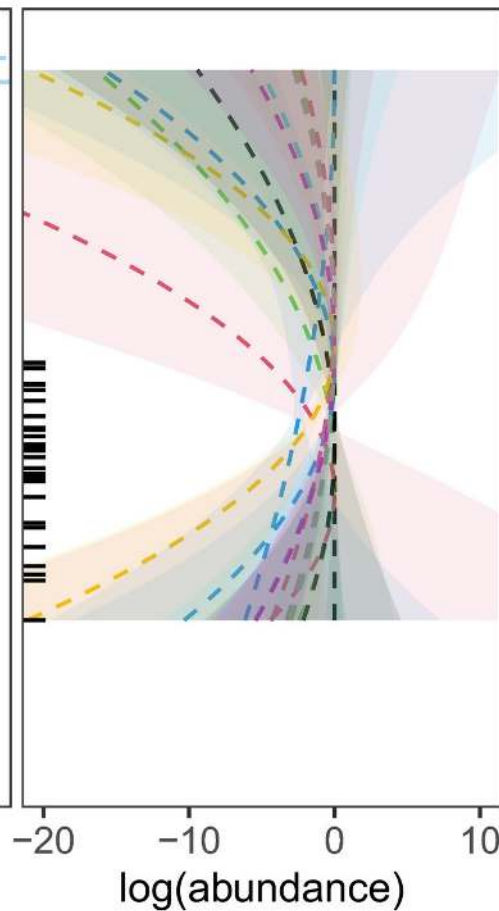
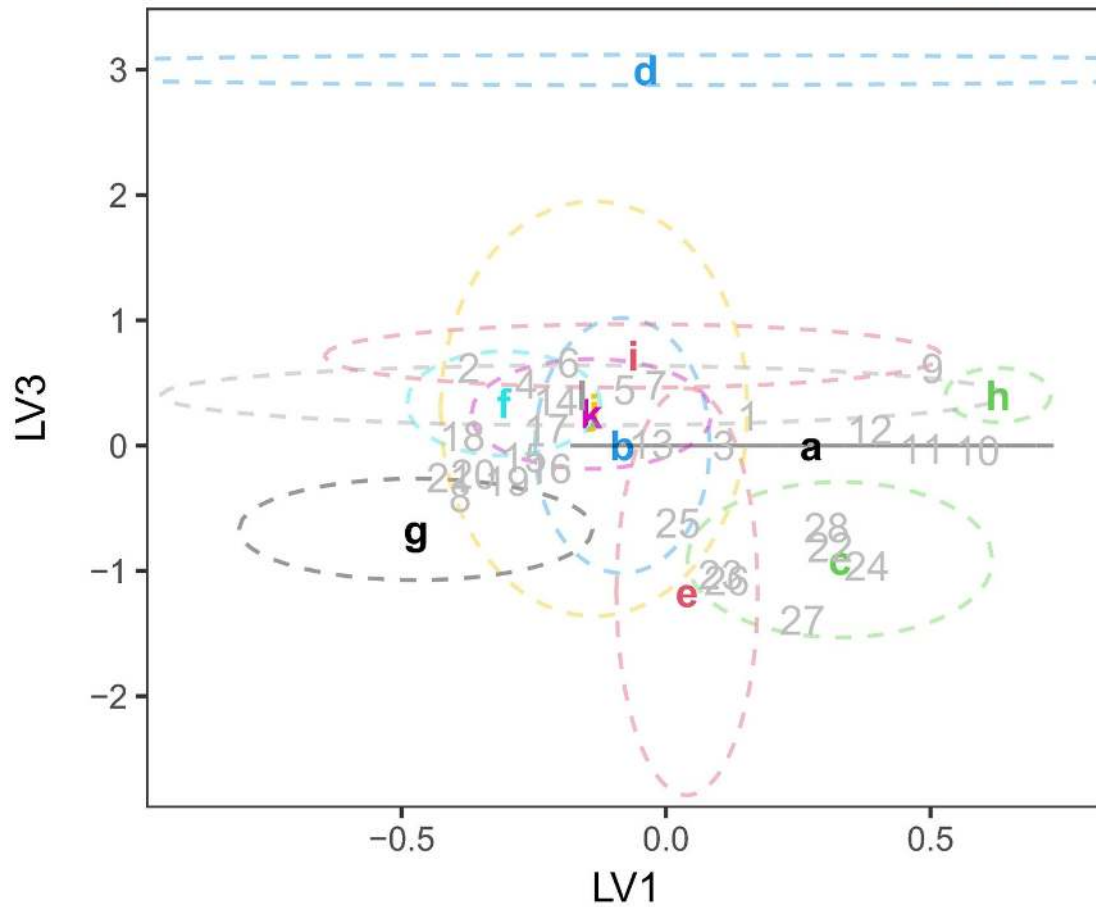
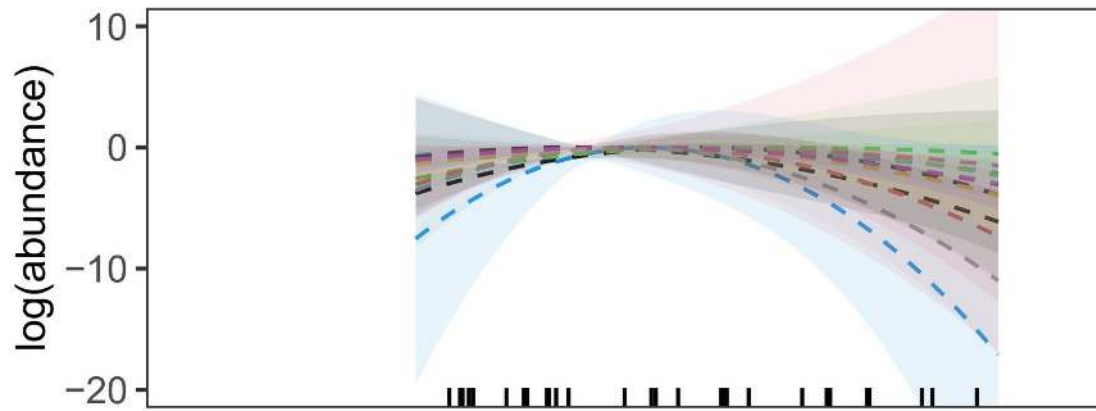




RESEARCH ARTICLE

Model-based ordination for species with unequal niche widths

Bert van der Veen^{1,2,3}  | Francis K. C. Hui⁴  | Knut A. Hovstad^{3,5}  | Erik B. Solbu¹  |
Robert B. O'Hara^{2,3} 



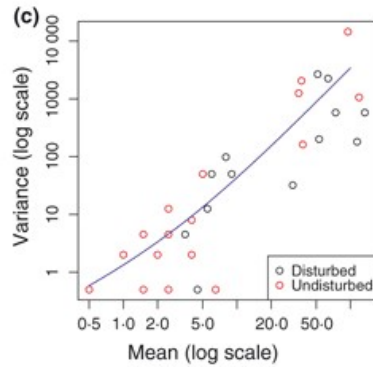
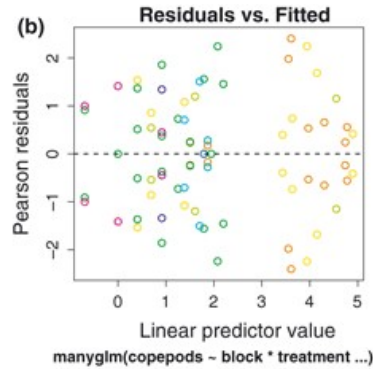
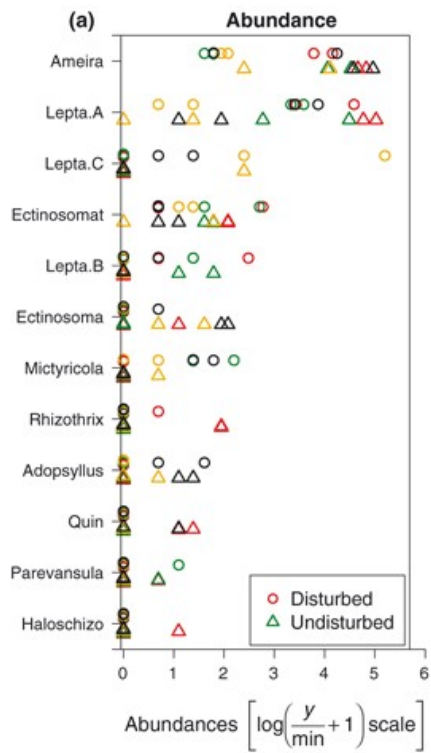
- Modelo do gllvm que relaxa o pressuposto de largura de nichos iguais das espécies.
- Permite fazer ordenações

Métodos para inferência

mvabund – an R package for model-based analysis of multivariate abundance data

Yi Wang^{1,2}, Ulrike Naumann¹, Stephen T. Wright¹, and David I. Warton^{1,3*}

- GLMs para cada espécie
- Lida com natureza correlacionada das variáveis por re-amostragem/permutação de resíduos (PIT residuals)



(d) Analysis of Deviance Table

Model: `manyglm(copepods ~ block * treatment, family = "negative binomial")`

Multivariate test:

	Res.Df	Df.diff	Dev	Pr(>Dev)
blocks	12	3	326.1	0.001 ***
treatment	11	1	106.5	0.008 **
blocks:treatment	8	3	48.5	0.063 .

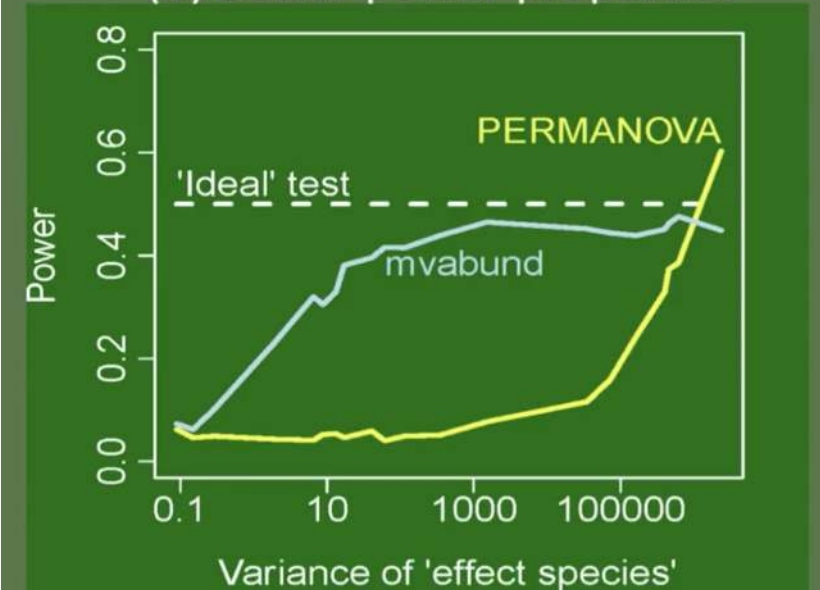
Univariate Tests:

	Ameira		Adopsyllus		Ectinosoma		...	Mictyricola		...
	Dev	Pr(>Dev)	Dev	Pr(>Dev)	Dev	Pr(>Dev)	...	Dev	Pr(>Dev)	...
blocks	18.066	0.037	23.793	0.018	23.113	0.019	...	19.15	0.033	...
treatment	29.98	0.045	0.091	0.986	21.072	0.073	...	21.072	0.071	...
blocks:treatment	7.581	0.493	1.294	0.946	0.833	0.946	...	5.269	0.493	...

manyglm
manyyany

Advantages of *mvabund*

(1) Better power properties



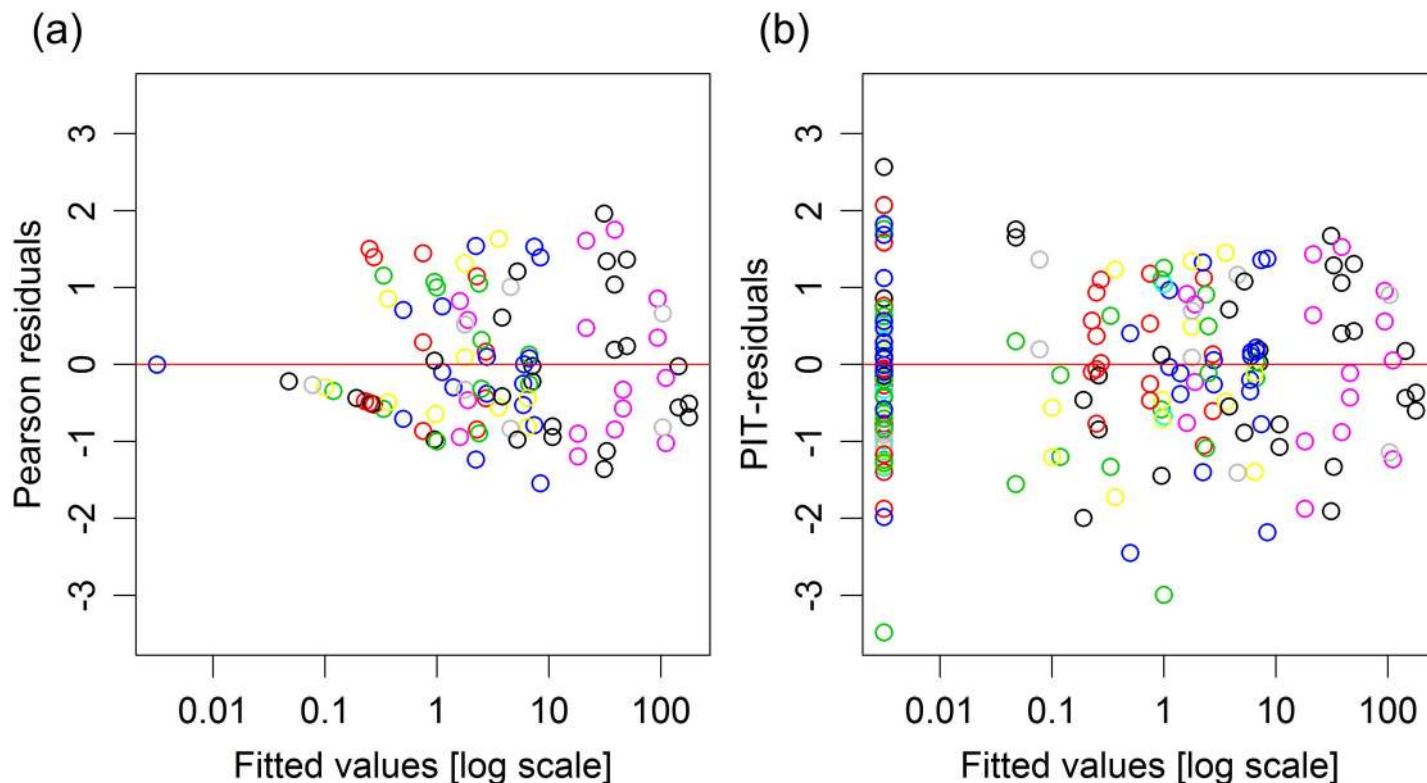
Citation: Warton DI, Thibaut L, Wang YA (2017) The PIT-trap—A “model-free” bootstrap procedure for inference about regression models with discrete, multivariate responses. PLoS ONE 12(7): e0181790. <https://doi.org/10.1371/journal.pone.0181790>

RESEARCH ARTICLE

The PIT-trap—A “model-free” bootstrap procedure for inference about regression models with discrete, multivariate responses

David I. Warton^{1*}, Loïc Thibaut², Yi Alice Wang³

- “model-free bootstrap”, adapted to the problem of discrete and highly multivariate data.
- Bom poder ao reamostrar dados multivariados com muitos zeros



Vantagens do mvabund

- Permitem adicionar o efeito do local explicitamente ao fazer ordenações, controlando assim efeitos de abundância total do local na composição de espécies

Composition - partitioning α and β diversity

We can include row effects in the model, to account for differences across sites in total abundance. All remaining the model then look at relative abundance or *composition*.

```
ft_comp=manyglm(Abundances~treatment+offset(log(pitfalls)),
                data=data.frame(reveg[-1]),composition=TRUE)

anova(ft_comp, nBoot=99)
#> Time elapsed: 0 hr 0 min 20 sec
#> Analysis of Deviance Table
#>
#> Model: Abundances ~ cols + treatment + offset(log(pitfalls)) + rows + cols:(treatment + offset(log(pitfalls)))
#>
#> Multivariate test:
#>
#>           Res.Df Df.diff   Dev Pr(>Dev)
#> (Intercept)      239
#> cols             216      23 361.2   0.01 **
#> treatment        215       1  14.1   0.01 **
#> rows             206       9  25.5   0.02 *
#> cols:treatment   184      23  56.7   0.02 *
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#> Arguments: P-value calculated using 99 iterations via PIT-trap resampling.
```

We can return to the anova and add a *p.uni* argument...

```
an.reveg = anova(NBft.reveg,p.uni="adjusted")
#> Time elapsed: 0 hr 0 min 7 sec
print(an.reveg$uni.p)
#>
#>      Acarina Amphipoda Araneae Blattodea Coleoptera Collembola
#> (Intercept)      NA      NA      NA      NA      NA      NA
#> treatment    0.185    0.154    0.985    0.109    0.129    0.31
#>
#>      Dermaptera Diotocardia Diplura Diptera Formicidae Haplotaxida
#> (Intercept)      NA      NA      NA      NA      NA      NA
#> treatment    0.985    0.985    0.852    0.378    0.983    0.79
#>
#>      Hemiptera Hymenoptera Isopoda Larvae Lepidoptera Polydesmida
#> (Intercept)      NA      NA      NA      NA      NA      NA
#> treatment    0.97     0.57    0.983    0.985    0.983    0.959
#>
#>      Pseudoscorpionida Scolopendrida Seolifera Soleolifera Thysanoptera
#> (Intercept)      NA      NA      NA      NA      NA      NA
#> treatment    0.983    0.983    0.983    0.57     0.983
#>
#>      Tricladida
#> (Intercept)      NA
#> treatment    0.793
#> attr(,"title")
#> [1] "Univariate Tests:"
```

Testa quais espécies são responsáveis pelo efeito

Equivalente à espécies indicadoras

err_dist	err_class
bernoulli	binary
binomial.count	binomial
binomial.prop	binomial
poisson	count
negbin	count
zip	count
zinb	count
zipl	count
zinbl	count
zipl.mu	count
zinbl.mu	count
gaussian	abundance
tweedie	abundance
zig	abundance
zigl	abundance
zigl.mu	abundance
ziig	abundance
ziigl	abundance
ziigl.mu	abundance
tab	percentage
zitag	percentage

Using the `senlm` package

Source: vignettes/senlm.Rmd

Installing the senlm package

```
install.packages("devtools")
devtools::install_github("PRIMER-e/senlm")
```

Models

The models (mean functions and error distributions) described here are for non-negative data. If it is possible to simulate negative data from these models (e.g. Gaussian error) then it is understood that those data points will be set to zero. Also, if a mean function takes a negative value, then the mean function should be set to zero at that location. Any data simulated from a mean function equal to zero, must be zero.

If the data is Bernoulli or percentage data then the data is bounded inclusively between 0 and 1.

If the data is binomial count data, then then the data is bounded inclusively between 0 and n .

The parameters for the mean function are stored in θ_M , the parameters for the error distribution are stored in θ_E . The binomial n value, and the δ parameter of the tail-adjusted beta are stored in θ_C .

Modelos conjuntos de distribuição de espécies

Extending SDMs to communities and BIs

Florian Hartig

P/A ~ Env

P/A ~ Env

P/A ~ Env

P/A ~ `glmmTMB(abundance ~ 0 + genus
+ env + (env|genus)`
...

- SDM = Species Distribution Model

- Stacked SDM if species are independent

- Multi-Species SDM if estimates for species are connected (e.g. via random slopes)

P/A ~ Env X

P/A ~ Env + X X

P/A ~ `glmmTMB(abundance ~ 0 + genus
+ env:genus + (0+genus|site)`
...

- joint Species Distribution Model (jSDM): add covariance between species

- Fits so-called "species association" – species are more/less often found together than expected by the environment alone

Características

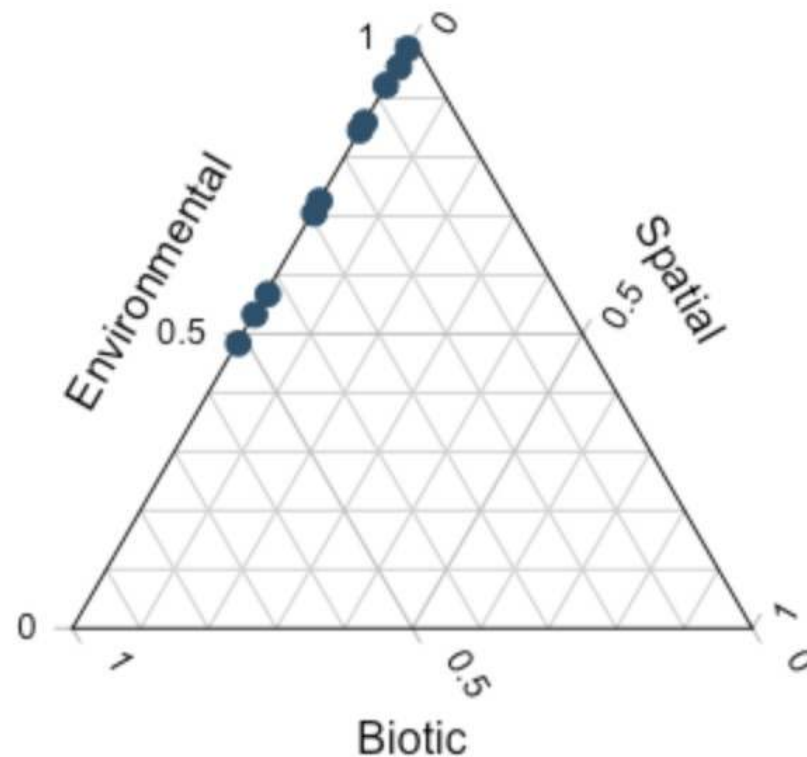
- Porque *joint*?
- Possibilidade de modelar cada espécie separadamente testando o efeito das covariáveis na abundância/incidência
- Partição da variação
- Inclusão de fatores aleatórios para controlar efeitos do local

A new joint species distribution model for faster and more accurate inference of species associations from big community data

Maximilian Pichler  | Florian Hartig 

Pacote s-jSDM

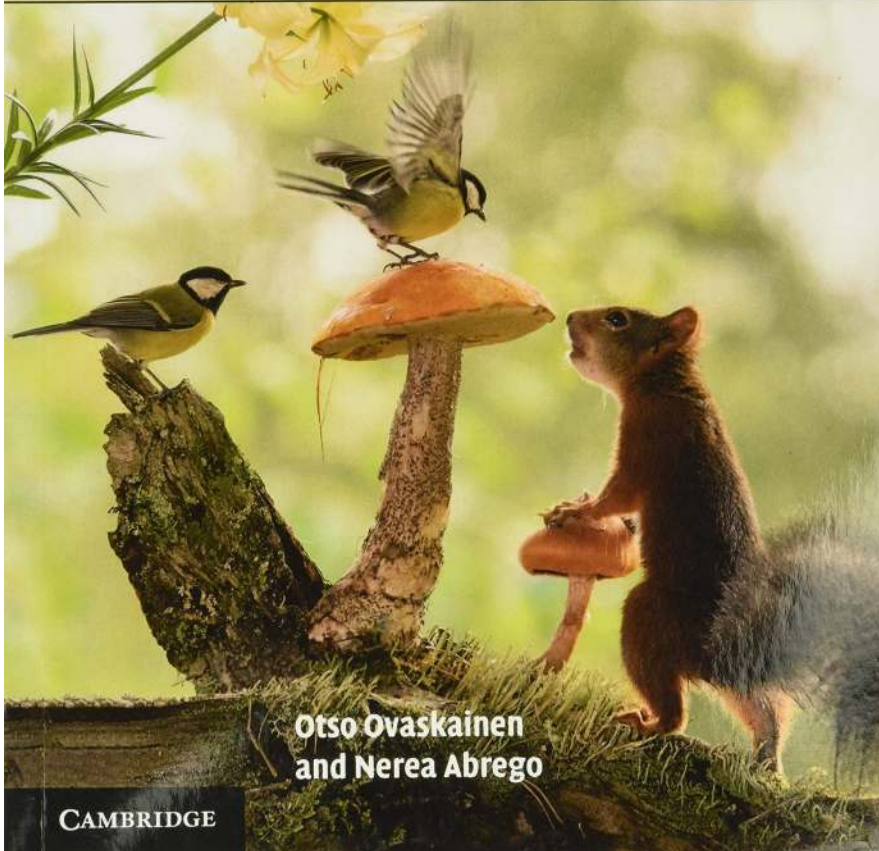
- Eficiente para lidar com grandes quantidades de dados
- Ao invés de usar variáveis latentes usa uma modificação do Monte Carlo





Joint Species Distribution Modelling

With Applications in R



Otso Ovaskainen and Nerea Abrego

CAMBRIDGE

2020

IDEA AND PERSPECTIVE

How to make more out of community data? A conceptual framework and its implementation as models and software

Abstract

Community ecology aims to understand what factors determine the assembly and dynamics of species assemblages at different spatiotemporal scales. To facilitate the integration between conceptual and statistical approaches in community ecology, we propose Hierarchical Modelling of Species Communities (HMSC) as a general, flexible framework for modern analysis of community data. While non-manipulative data allow for only correlative and not causal inference, this framework facilitates the formulation of data-driven hypotheses regarding the processes that structure

Otso Ovaskainen,^{1,2*}
Gleb Tikhonov,¹ Anna Norberg,¹
F. Guillaume Blanchet,^{3,4}
Leo Duan,⁵ David Dunson,⁵
Tomas Roslin⁶ and
Nerea Abrego^{2,7}

Received: 23 September 2019

Accepted: 16 December 2019

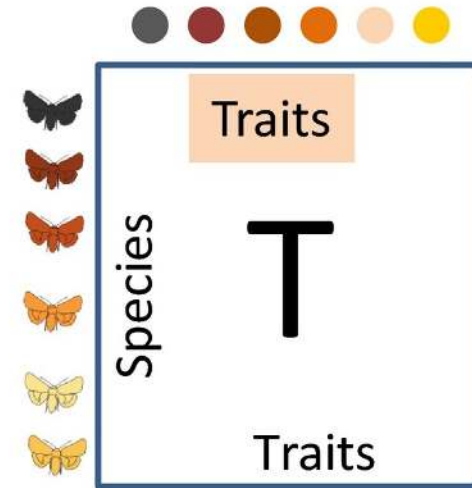
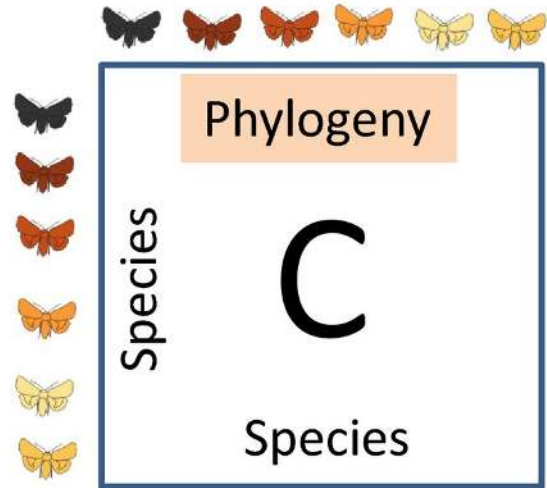
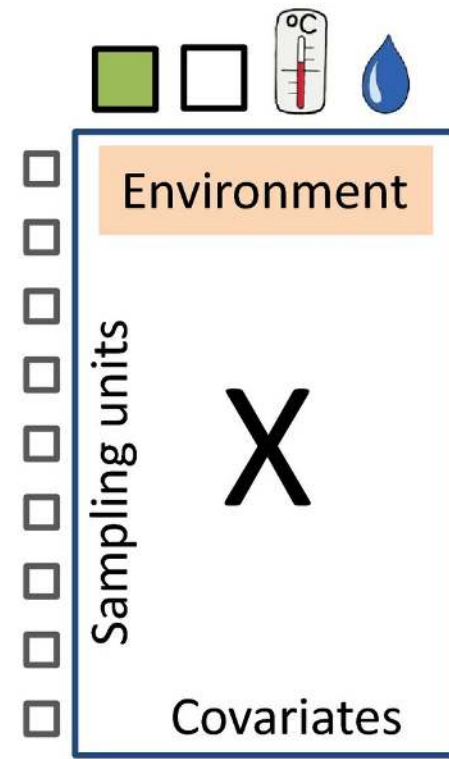
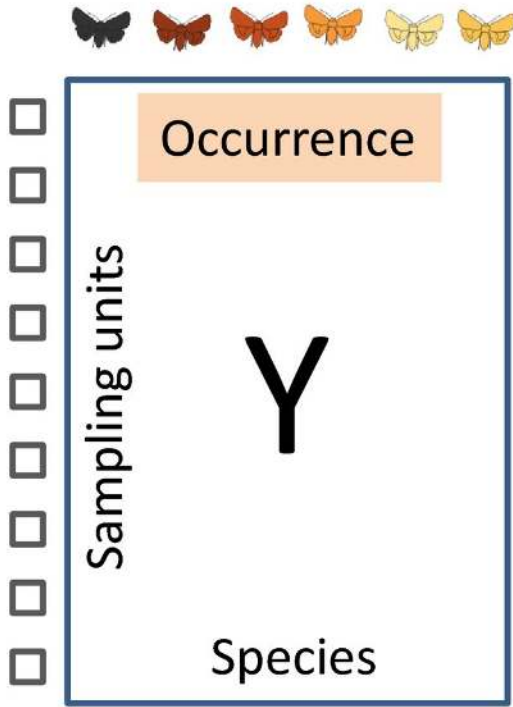
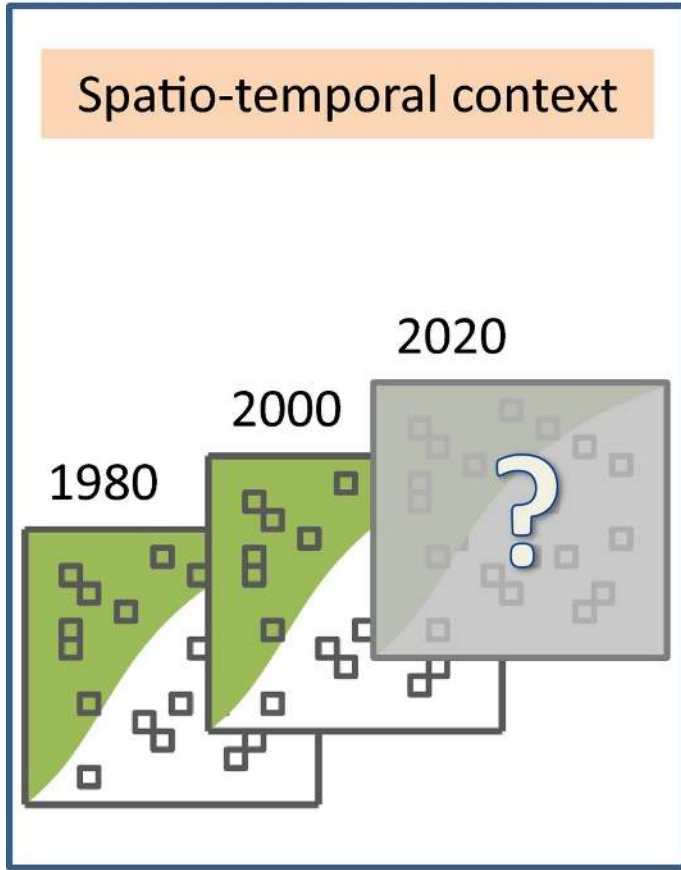
DOI: 10.1111/2041-210X.13345

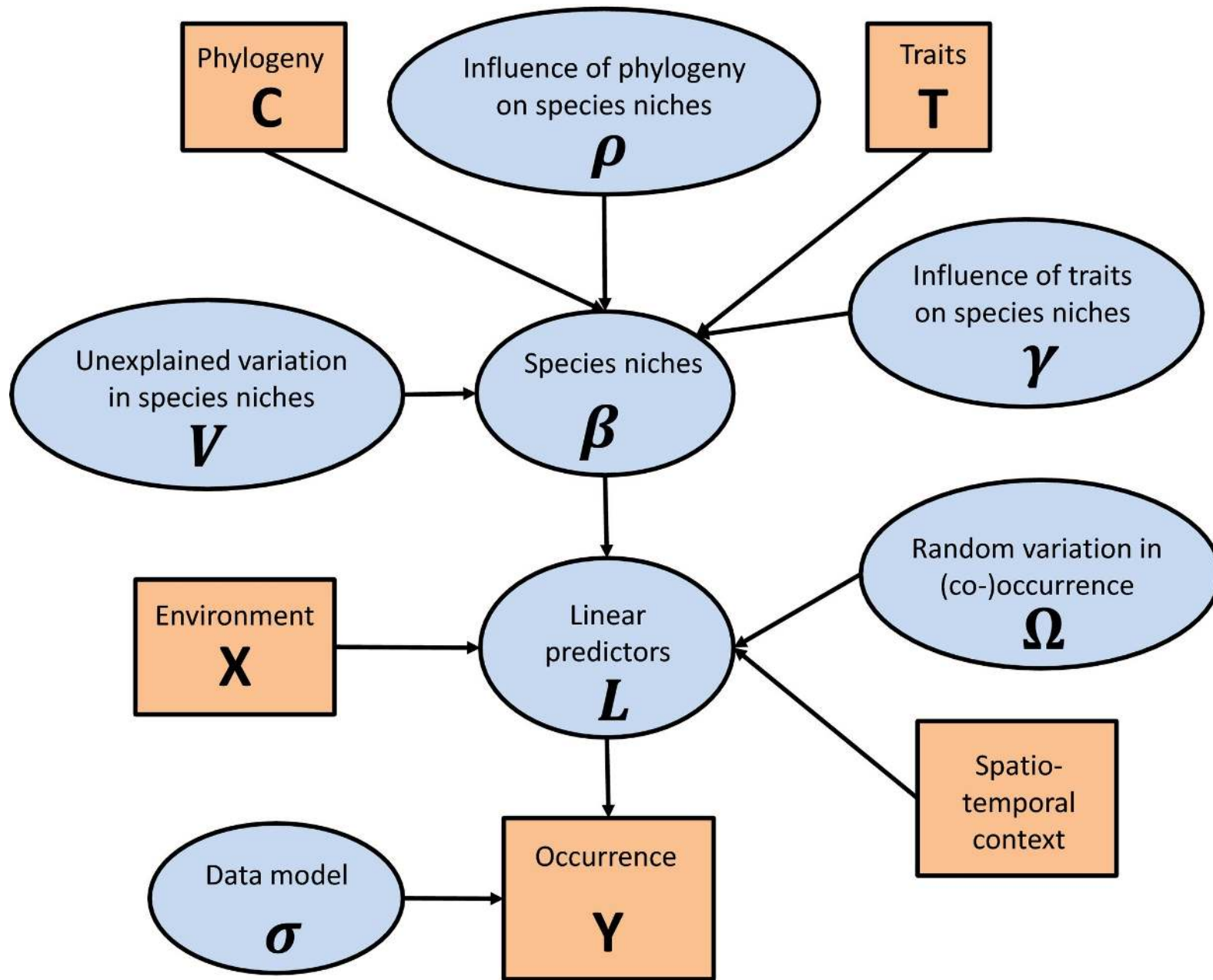
APPLICATION

Methods in Ecology and Evolution

Joint species distribution modelling with the R-package Hmsc

Gleb Tikhonov^{1,2} | Øystein H. Opedal^{2,3} | Nerea Abrego⁴ | Alekski Lehikoinen⁵ |
Melinda M. J. de Jonge⁶ | Jari Oksanen⁷ | Otso Ovaskainen^{2,3}





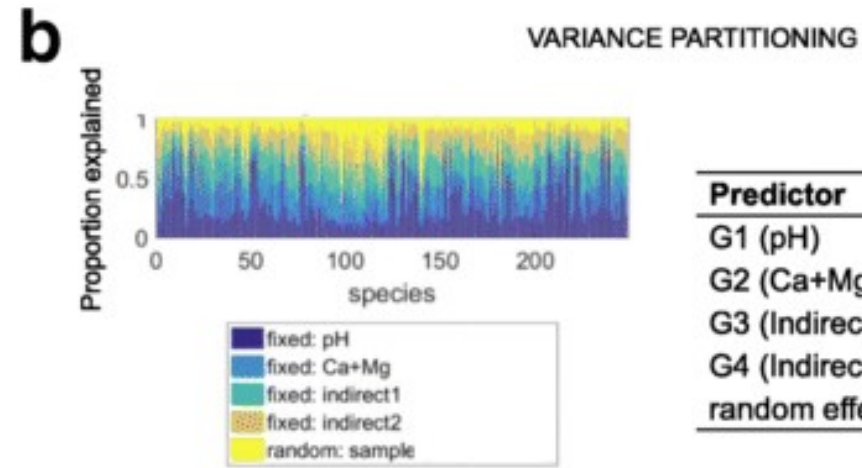
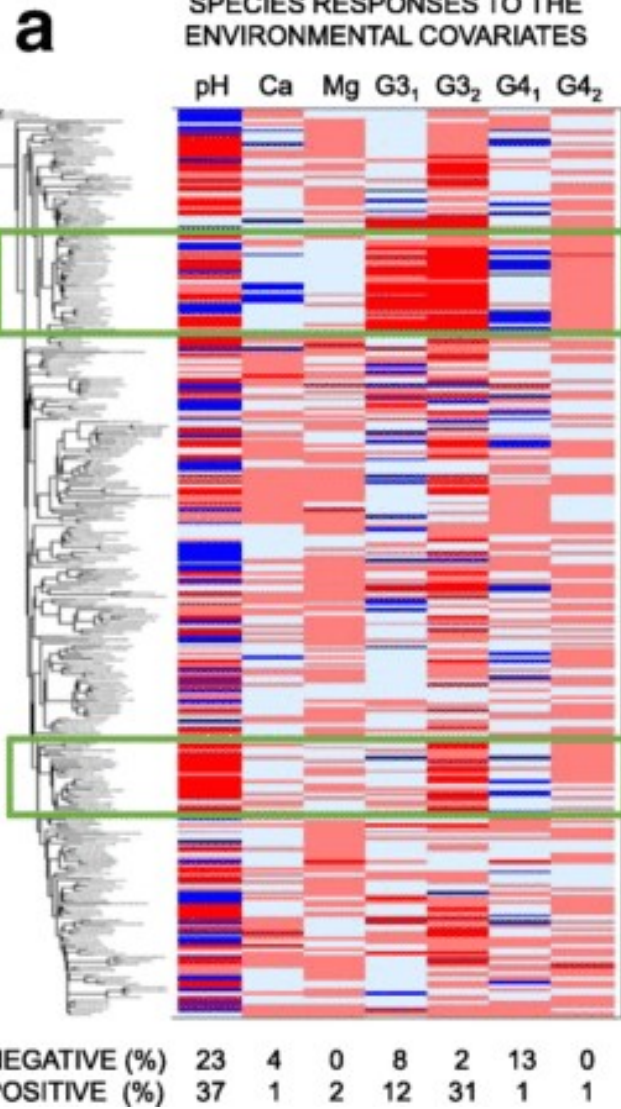
RESEARCH

Open Access



Direct and indirect effects of a pH gradient bring insights into the mechanisms driving prokaryotic community structures

Daniel R. Lammel^{1,2,3}, Gabriel Barth⁴, Otso Ovaskainen^{5,6}, Leonardo M. Cruz¹, Josileia A. Zanatta⁷, Masahiro Ryo³, Emanuel M. de Souza¹ and Fábio O. Pedrosa^{1*}



Predictor	Proportion Explained
G1 (pH)	33%
G2 (Ca+Mg)	6%
G3 (Indirect 1)	18%
G4 (Indirect 2)	10%
random effect	33%

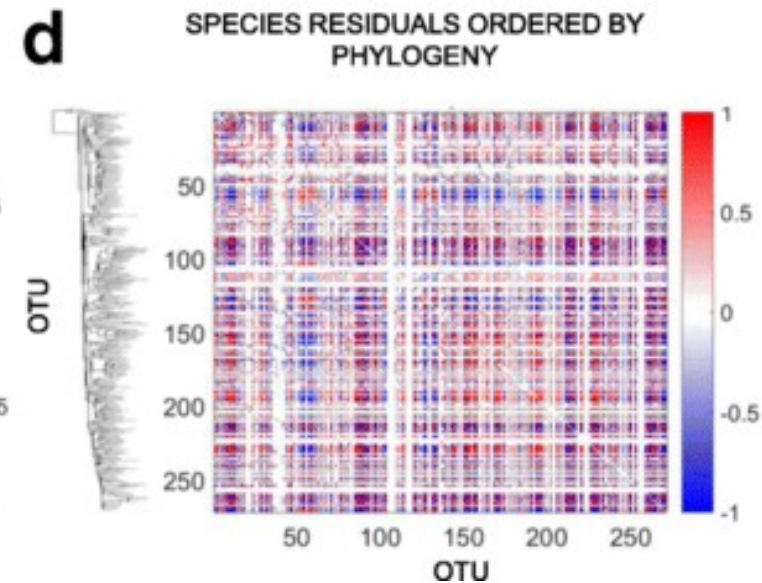
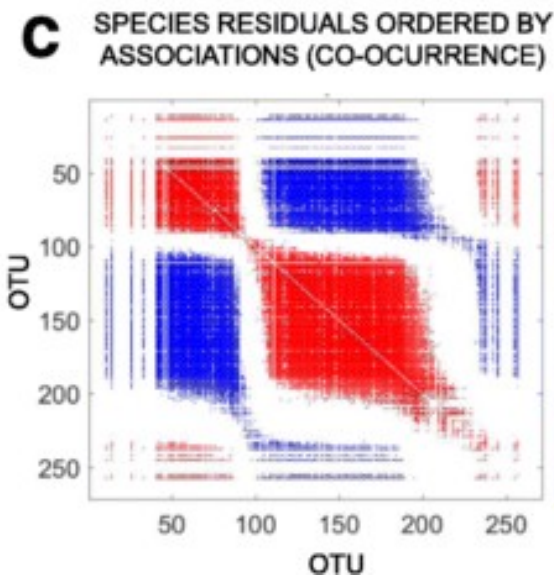


Table Table 2.. Predictive power to disentangle among simulated metacommunity scenarios. The values give AUC values for predicting each of the Choices C1–C10 based on 10-fold cross-validation, using as predictors the output metrics related to each of the five analyses types listed in Table 1. The column ALL shows the predictive power if using all output metrics of the five statistical approaches simultaneously. Values smaller than 75% are not shown to emphasize the cases where predictive power is substantial

Statistical approach	HAB	BETA	db-VP	db-RDA	JSDM	ALL
C1. presence versus absence of gradient	0.94		0.88	0.82	0.79	0.97
C2. patchy versus continuous landscape	1	0.99	0.78	0.97	0.92	1
C3. uniform versus varying patch quality						
C4. dynamic versus static patches					0.78	0.82
C5. slow versus fast patch turnover					0.8	0.93
C6. generalists versus specialists		0.76			0.9	0.96
C7. partial versus strict specialists						0.81
C8. variation versus no variation in dispersal strategy						
C9. short versus long dispersal		0.87		0.8	0.75	0.93
C10. missing versus not missing covariates					0.87	0.9

Simulação com 10 cenários diferentes



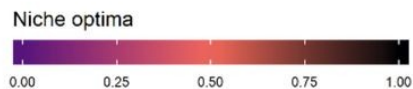
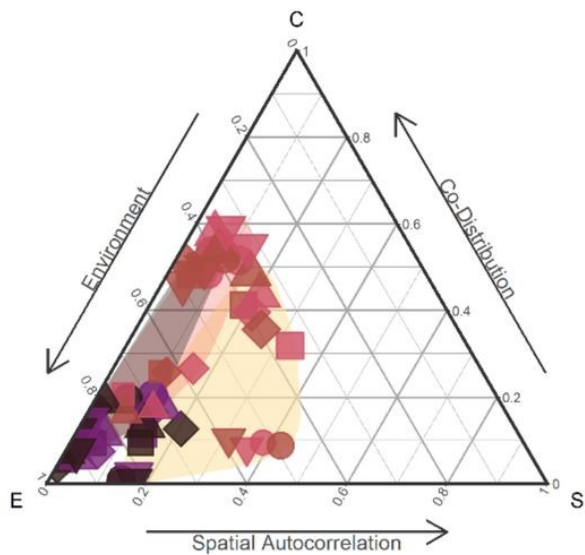
JSDMs foram os que melhor capturaram os processos geradores das comunidades

The internal structure of metacommunities

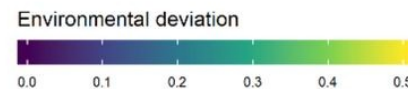
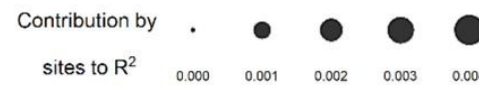
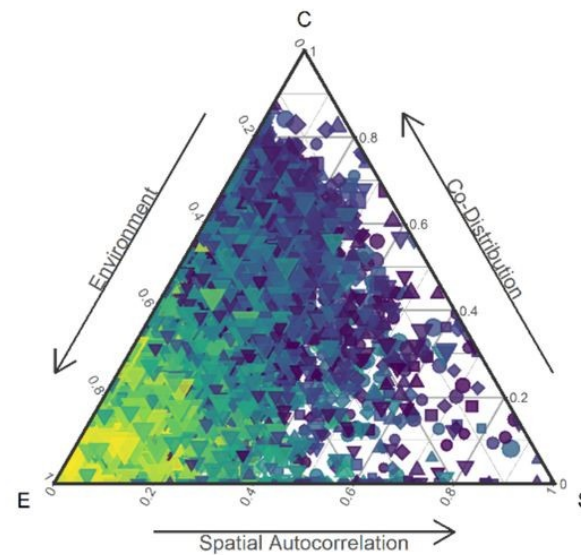
Mathew A. Leibold, F. Javiera Rudolph, F. Guillaume Blanchet, Luc De Meester, Dominique Gravel, Florian Hartig, Pedro Peres-Neto, Lauren Shoemaker and Jonathan M. Chase



Species



Sites



Investigação da contribuição de diferentes espécies e locais pra padrões em escala de metacomunidades usando HMSC

- Species can have distributions that vary in the degree to which they are determined by combinations of the basic community assembly processes depending on features of their ecology (e.g. dispersal and environmental preferences); and
- The predominant assembly processes that determine local communities can differ among adjacent sites in a metacommunity (e.g. sites that are occupied by species most fit for environmental conditions vs sites occupied by species in nearby sites that may differ in environmental conditions).

Como interpretar a co-distribuição
(associações entre espécies) derivada da
covariância dos resíduos

- Abordagens correlativas não substituem estudos mecanísticos
- Argumenta para a combinação de JSDM com experimentos e uso de atributos

Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions

Carsten F. Dormann¹  | Maria Bobrowski² | D. Matthias Dehling³ | David J. Harris⁴ | Florian Hartig^{1,5} | Heike Lischke⁶ | Marco D. Moretti⁷  | Jörn Pagel⁸ | Stefan Pinkert⁹  | Matthias Schleuning¹⁰ | Susanne I. Schmidt¹¹  | Christine S. Sheppard⁸  | Manuel J. Steinbauer^{12,13}  | Dirk Zeuss¹⁴  | Casper Kraan^{15,16} 

ECOLOGY LETTERS

Ecology Letters, (2020) 23: 1050–1063

doi: 10.1111/ele.13525

IDEAS AND PERSPECTIVES

Co-occurrence is not evidence of ecological interactions

Abstract

There is a rich amount of information in co-occurrence (presence–absence) data that could be used to understand community assembly. This proposition first envisioned by Forbes (1907) and then Diamond (1975) prompted the development of numerous modelling approaches (e.g. null

F. Guillaume Blanchet,^{1*} 
Kevin Cazelles²  and
Dominique Gravel¹ 

- Simples associação espacial entre espécies por si (ou a falta dela) só não é um bom indicativo de interações bióticas
- Desenho amostral, variáveis incluídas e escala espacial são essenciais para fazer inferência de relações entre espécies
- Competidora inferior pode já ter sido excluída do sistema

ECOGRAPHY

Forum

Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments?

Damaris Zurell, Laura J. Pollock and Wilfried Thuiller

Ecography

41: 1812–1819, 2018

doi: 10.1111/ecog.03315

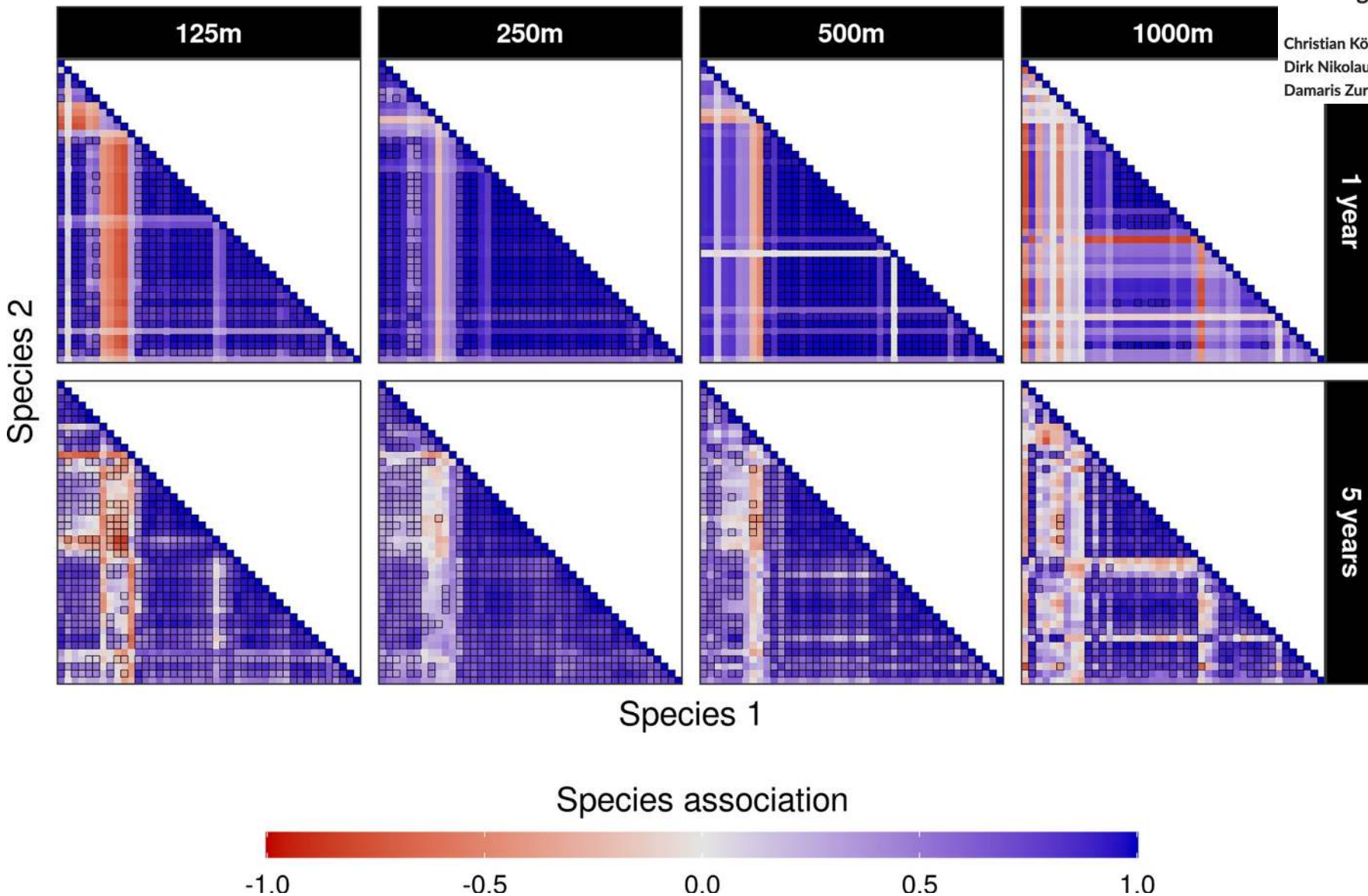
JSDM parecem detectar bem efeitos de competição e facilitação/mutualismo, mas não tanto relações predador-presa

Valor da correlação em si não reflete força de interação, pois depende da prevalência das espécies

No entanto, a detecção deste sinal depende da escala espacial...

Scale dependency of joint species distribution models challenges interpretation of biotic interactions

Christian König¹ | Rafael O. Wüest² | Catherine H. Graham² |
Dirk Nikolaus Karger² | Thomas Sattler³ | Niklaus E. Zimmermann^{2,4} |
Damaris Zurell^{1,2}



1 year

5 years

Species association

-1.0 -0.5 0.0 0.5 1.0

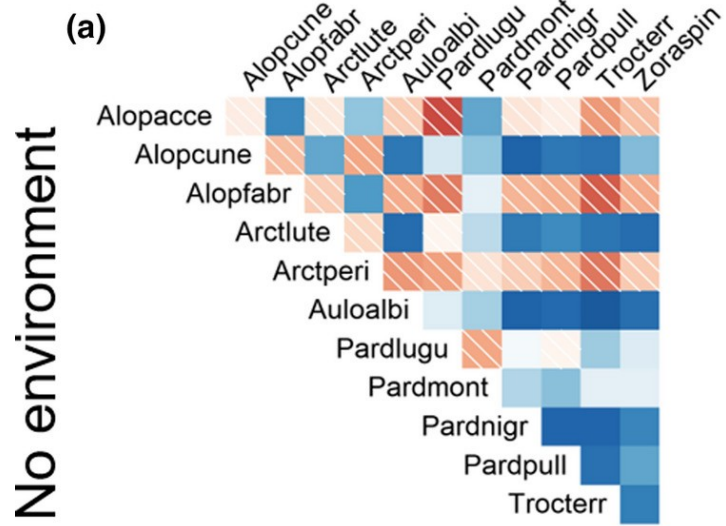
Métodos para estimar co-distribuição (co-ocorrência) derivada da covariância dos resíduos

Untangling direct species associations from indirect mediator species effects with graphical models

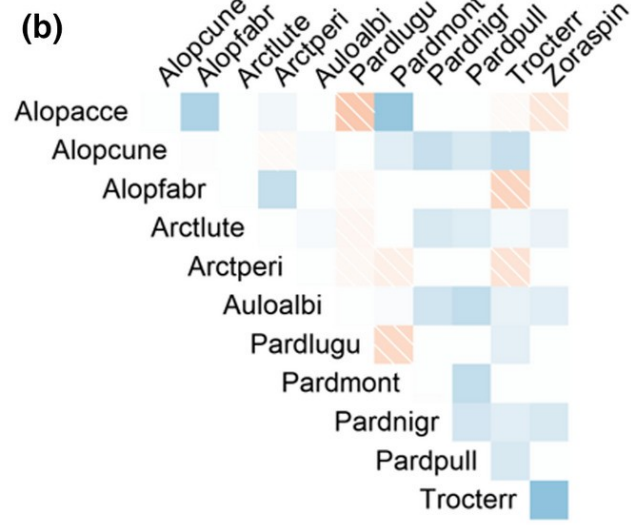
Gordana C. Popovic¹  | David I. Warton¹  | Fiona J. Thomson² | Francis K. C. Hui³  | Angela T. Moles⁴ 

- Nova aplicação de copulas
- Distinguir associação entre pares de espécies via uma “espécie mediadora”
 - Associação indireta via outra espécie ou variável ambiental
- Permite usar dados de abundância, biomassa, proporção e ordinal de forma paramétrica
- Melhor do que *gllvm* e abordagens baseadas em incidência e modelos nulos para diferenciar entre associação interespecíficas diretas ou indiretas

Correlation (indirect + direct)



Precision (direct only)



Graph

