Universidade Federal de São Carlos

Centro de Ciências Biológicas e da Saúde

Programa de Pós-graduação em Ecologia e Recursos Naturais

GUSTAVO HENRIQUE DE CARVALHO

RELAÇÕES ENTRE AMBIENTE, TRAÇOS, COMPOSIÇÃO E FUNCIONAMENTO DE COMUNIDADES VEGETAIS DE CERRADO

Orientador: Dr. Marco Antônio Batalha

São Carlos

Março de 2013

Universidade Federal de São Carlos
Centro de Ciências Biológicas e da Saúde
Programa de Pós-graduação em Ecologia e Recursos Naturais

GUSTAVO HENRIQUE DE CARVALHO

RELAÇÕES ENTRE AMBIENTE, TRAÇOS, COMPOSIÇÃO E FUNCIONAMENTO DE COMUNIDADES VEGETAIS DE CERRADO

Tese apresentada ao Programa de Pós-Graduação em Ecologia e Recursos Naturais, do Centro de Ciências Biológicas e da Saúde, da Universidade Federal de São Carlos, sob a orientação do Prof. Dr. Marco Antônio Batalha.

São Carlos
Março de 2013

Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária/UFSCar

C331ra

Carvalho, Gustavo Henrique de.
Relações entre ambiente, traços, composição e funcionamento de comunidades vegetais de cerrado / Gustavo Henrique de Carvalho. -- São Carlos : UFSCar, 2013.
151 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2013.

1. Cerrados. 2. Comunidades vegetais. 3. Funcionamento de comunidades. 4. Modelos de equações estruturais. I. Título.

CDD: 574.5247 (20ª)

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Ecologia e Recursos Naturais

Relatório de Defesa de Tese
Candidato: Gustavo Henrique de Carvalho

Aos 21/03/2013, às 14:00, realizou-se na Universidade Federal de São Carlos, nas formas e termos do Regimento Interno do Programa de Pós-Graduação em Ecologia e Recursos Naturais, a defesa de tese de doutorado sob o título: RELAÇÕES ENTRE AMBIENTE, TRAÇOS, COMPOSIÇÃO E FUNCIONAMENTO DE COMUNIDADES VEGETAIS DE CERRADO, apresentada pelo candidato Gustavo Henrique de Carvalho. Ao final dos trabalhos, a banca examinadora reuniu-se em sessão reservada para o julgamento, tendo os membros chegado ao seguinte resultado:

Participantes da Banca	Função	Instituição	Conceito
Prof. Dr. Marco Antonio Portugal Luttembarck Batalha	Presidente	UFSCar	A
Profa. Dra. Maria Inês Salgueiro Lima	Titular	UFSCar	A
Profa. Dra. Maria Elina Bichuette	Titular	UFSCar	A
Prof. Dr. Rodrigo Augusto Santinelo Pereira	Titular	USP	A
Prof. Dr. Tadeu de Siqueira Barros	Titular	UNESP	A

Resultado Final: A

Parecer da Comissão Julgadora*:

Aprovado com distinção e louvor.

Encerrada a sessão reservada, o presidente informou ao público presente o resultado. Nada mais havendo a tratar, a sessão foi encerrada e, para constar, eu, João Augusto da Silva Affonso, representante do Programa de Pós-Graduação em Ecologia e Recursos Naturais, lavrei o presente relatório, assinado por mim e pelos membros da banca examinadora.

Prof. Dr. Marco Antonio Portugal Luttembarck Batalha

Maria Inês Salgueiro Lima
Profa. Dra. Maria Inês Salgueiro Lima

Maria Elina Bichuette
Profa. Dra. Maria Elina Bichuette

Prof. Dr. Rodrigo Augusto Santinelo Pereira

Tadeu Siqueira
Prof. Dr. Tadeu de Siqueira Barros

Representante do PPG: João Augusto da Silva Affonso

(X) Não houve alteração no título da tese ( ) Houve. O novo título passa a ser:

*Obs: Se o candidato for reprovado por algum dos membros, o preenchimento do parecer é obrigatório.

Para gozar dos direitos do título de Doutor em Ecologia e Recursos Naturais, o candidato ainda precisa ter sua tese homologada pelo Conselho de Pós-Graduação da UFSCar.

Dedico aos meus pais, Jerônimo e Sueli, e à minha noiva,

Cláudia, pelo amor e paciência.

AGRADECIMENTOS

Agradeço ao Prof. Dr. Marco Antônio Batalha, pela sempre pronta disponibilidade e orientação. Aos meus pais, por terem me dado os meios para chegar até aqui. Ao CNPq, CAPES e FAPESP pelas bolsas e auxílios financeiros. À direção do Parque Nacional das Emas, pela licença concedida. Aos companheiros de laboratório, pela ajuda nos trabalhos de campo e amizade. Finalmente, à minha noiva, Cláudia, por ter ficado ao meu lado nos momentos mais difíceis.

SUMÁRIO

ABSTRACT

Understanding how biodiversity and ecosystem functioning will respond to changes in the environment is fundamental to avoid the loss of species and ecosystem function. In realistic scenarios, the biodiversity-ecosystem functioning pathway may account for only a small share of all factors determining ecosystem function. . In the last chapter, we described how the use of latent variables and structural equation models could be a useful tool for environmental research. In the second chapter, we investigated the strength to which variations in environmental characteristics in a Neotropical savanna affected functional diversity and litter decomposition. We sought an integrative approach, using structural equation modelling to connect fire frequency, soil fertility, exchangeable aluminium, water availability, functional diversity of woody plants, and litter decomposition rates in a causal chain. By expressing these hypotheses simultaneously, we revealed a number of direct and interactions. We found significant effects of soil nutrients, water availability, and aluminium on functional diversity and litter decomposition. Fire did not have a significant direct effect on functional diversity or litter decomposition. However, fire was connected to both variables through soil fertility. In the third chapter, we tested if we could predict local abundances using a pool of species and traits. To test if traits improved the predictions generated by the information present in the pool, we used maximum entropy models coupled with permutation tests. We could accurately predict local abundances of the 73 species in the pool. Dispersal limitation was the main factor assembling communities at all the scales we studied, but the importance of stochasticity increased in local scales. Traits explained little of the uncertainty present in local abundances, but coupled with pool frequencies they yielded large coefficients of determination. In the fourth chapter, we showed how fire and soil fertility influence different sets of traits in different ways, which, in turn, influence community composition and density.

Keywords: cerrado, community assembly, community functioning, structural equation modelling.

RESUMO

Entender como a biodiversidade e o funcionamento dos ecosistemas vão responder às mudanças nas condições ambientais é essencial para a manutenção das interações que influenciam as propriedades dos ecossistemas. Os sistemas ecológicos respondem à mudanças nas condições ambientais não apenas por meio da interação direta com essas condições, mas também de maneira indireta, via organismos. No primeiro capítulo, descrevemos o uso de variáveis latentes e modelos de equações estruturais em ecologia. No segundo capítulo, nós investigamos como variações em características ambientais resultam em variações na diversidade funcional e funcionamento de uma comunidade de arbustos e árvores de cerrado no Parque Nacional das Emas (GO). Nós usamos modelagem de equações estruturais para quantificar os efeitos da fertilidade do solo, alumínio, disponibilidade de água e diversidade funcional na decomposição de serapilheira. Nós encontramos efeitos diretos entre nutrientes do solo, disponibilidade de água e alumínio na diversidade funcional e funcionamento da comunidade. O fogo não teve um efeito direto, mas sim caminhos indiretos pelos quais o fogo influencia a diversidade funcional e o funcionamento. No terceiro capítulo, nós procuramos identificar a importância de processos determinísticos e estocásticos na composicão da comunidade vegetal do cerrado em Emas. Nós testamos, por meio de modelos de máxima entropia e testes de permutação, se os traços das espécies adicionariam informação relevante para a previsão das abundâncias além da informação já presente no repositório de espécies. Nossos modelos tiveram alto poder de previsão para as 73 espécies do repositório. Limitação de dispersão foi o principal processor compondo as comunidades. Processos estocásticos também tiveram grande importância, principalmente na escala local. Sem a informação prévia sobre as frequências das espécies, modelos com os traços tiveram pouco poder de explicação. Entretanto, ao combinarmos traços e frequências no repositório, nossos modelos resultaram em altos coeficientes de determinação. No último capítulo, nós mostramos como fogo e fertilidade do solo influenciam diferentes grupos de traços e como esses traços influenciam na composição e densidade das comunidades.

Palavras-chave: assembleia de comunidades vegetais, cerrado, funcionamento da comunidade, modelos de equações estruturais.

I - INTRODUÇÃO GERAL

INTRODUÇÃO GERAL

Prever como a biodiversidade e o funcionamento dos ecossistemas responderão às mudanças nas condições ambientais é essencial para a manutenção das propriedades dos ecossistemas e dos serviços que eles provêm (Loreau *et al.* 2001). Em comunidades vegetais, tal entendimento pode levar à políticas de conservação mais efetivas, especialmente aquelas que referem-se ao manejo de agentes de perturbação, como o fogo, para minimizar a perda de espécies e serviços. Vários estudos tiveram como objetivo investigar as interações entre fatores abióticos, diversidade biológica e funcionamento dos ecossistemas (Tilman *et al.* 1997; Hooper *et al.* 2005). Entretanto, a maioria desses estudos analisou as relações entre apenas dois dos componentes mencionados, desconsiderando, principalmente, a influência de variações do ambiente na interação entre biodiversidade e funcionamento dos ecossistemas (Tilman *et al.* 1997; Hooper & Vitousek 1997; Hector *et al.* 1999).

Os sistemas ecológicos respondem às mudanças nas condições ambientais não apenas por meio de interação direta, mas também de maneira indireta, via organismos (Chapin *et al.* 1997; Cardinale *et al.* 2000). Dessa forma, ao considerarmos as relações diretas e indiretas entre ambiente, biodiversidade e funcionamento, teremos dado um importante passo rumo à identificação e quantificação das vias que, ultimamente, influenciam nas propriedades dos ecossistemas (Srivastava & Vellend 2005). A investigação das interações entre organismos e funcionamento em condições ambientais flutuantes é uma das questões da ecologia de ecossistemas que necessita de mais atenção (Srivastava & Vellend 2005).

Os mecanismos pelos quais o ambiente determina quais espécies ocorrem nas comunidades têm recebido atenção dos pesquisadores há muito tempo (Macarthur & Levins 1967; Connel & Slatyer 1977). Principalmente, dois grupos de teorias procuram explicar o processo de formação de comunidades à partir de um repositório de espécies (Chase *et al*. 2005): o da teoria neutra, segundo a qual fatores aleatórios governam a formação de comunidades, uma vez que todas as espécies do repositório regional são funcionalmente equivalentes, ou seja, "neutras" (Hubbell 2005, 2006) e o das teorias baseadas nos nichos das espécies, segundo as quais fatores determinísticos são os principais responsáveis pela composição das comunidades (Weiher *et al.* 1998).

Nos modelos de formação de comunidades baseados em nichos, vários mecanismos de exclusão de espécies foram propostos. Um desses mecanismos é dos filtros ambientais, segundo o qual fatores abióticos como fertilidade do solo e fogo determinam que espécies possuem as características necessárias para sobreviverem em um determinado local (Keddy 1992). De maneira semelhante, o processo de limitação da similaridade de nichos, causada principalmente pela interação entre as espécies de uma comunidade, também deixa uma marca nas comunidades. Por exemplo, espécies com atributos semelhantes têm mais chance de terem alta sobreposição de nichos, o que faz com que elas tenham que competir por recursos (Fridley 2001; Kraft *et al.* 2008; Cornwell & Ackerly 2009). Porém, a competição por recursos faz com que as espécies limitem sua similaridade para que tenham menor sobreposição de nichos, atuando em direção oposta aos filtros ambientais.

Os organismos influenciam taxas e processos dos ecossistemas de diversas formas (Loreau *et al*. 2001; Hooper *et al*. 2005). Por exemplo, quanto mais espécies em uma comunidade, maior a chance de que espécies-chave de algum processo estejam presentes nessa comunidade (Loreau 2000). Porém, uma vez que as espécies podem ter papéis semelhantes para o funcionamento dos ecossistemas, muitas vezes a riqueza não é o índice mais adequando para investigar a relação entre biodiversidade-funcionamento (Wardle *et al.* 1997; Hooper *et al*. 2005). Dessa forma, os ecólogos têm procurado por outras formas de medir a diversidade das comunidades e melhor entender como os organismos respondem ao ambiente e influenciam no funcionamento dos ecossistemas (Petchey & Gaston 2002; Pavoine *et al.* 2011). Uma maneira alternativa de se medir a biodiversidade é olhando para a diversidade de traços funcionais de uma comunidade. Foi sugerido que comunidades com maior diversidade de traços funcionais operam de maneira mais eficiente devido à maior complementaridade de nichos, o que leva ao particionamento de recursos (Díaz & Cabido 2001; Hooper *et al.* 2005). Ainda, a diversidade funcional pode abordar diferentes facetas do funcionamento das comunidades, uma vez que é medida por meio de vários traços funcionais das espécies (Cadotte 2011). Petchey & Gaston (2002, 2006) propuseram um índice de diversidade funcional (FD) que estima a complementaridade de traços funcionais em comunidades. Maiores diferenças nos traços indicam maior complementaridade. Posteriormente, uma extensão do índice para levar em conta a variação intraespecífica foi proposta (Cianciaruso *et al.* 2009).

Prever as abundâncias das espécies que podem ocorrer em uma comunidade sob certas condições ambientais é um dos objetivos mais frequentes na ecologia (Laughlin *et al.* 2012). Caso fatores determinísticos, como os filtros ambientais e interações, tenham papel importante na formação das comunidades, é esperado que a distribuição dos traços funcionais dessa comunidade indique a ação de processos determinísticos (Kraft *et al.* 2008; Mayfield & Levine 2010). Por outro lado, caso fatores aleatórios, como a estocasticidade demográfica, sejam os principais responsáveis pela formação das comunidades, a distribuição dos traços nessas comunidades não será diferente daquela esperada pelo acaso (Shipley *et al.* 2012). Ainda, mesmo que processos determinísticos deixem uma marca significativa na distribuição de traços de uma comunidade, fatores aleatórios também podem responder por uma proporção significativa da capacidade dos pesquisadores de prever bem as abundâncias das espécies (Tilman 2004; Chase 2005; Gravel *et al*. 2006; Stokes & Archer 2010). A identificação dos processos por trás da formação das comunidades é uma das questões centrais da ecologia (Sutherland *et al.* 2013).

No cerrado, a fertilidade do solo, baixo pH, altas concentrações de alumínio e incidência de queimadas se apresentam como possível fatores ambientais que limitam a ocorrência de espécies. Dessa forma, filtros ambientais provavelmente governam a ocorrência de plantas em áreas de cerrado. Por exemplo, é notável a variação na densidade de plantas na vegetação do cerrado, indo desde fisionomias campestres, mais abertas, até fisionomias mais fechadas, com alta densidade de indivíduos arbóreos. Teorias clássicas sugeriram que o cerrado é um gradiente de fertilidade. Assim, a densidade de indivíduos acompanharia a fertilidade do solo (Goodland & Pollard 1973). Estudos recentes tanto corroboraram (Amorim & Batalha 2008; Silva & Batalha 2008), quanto não corroboraram (Ruggiero *et al.* 2002) essas teorias. Além do solo, o fogo tem sido indicado como fator ambiental que influencia a diversidade fenotípica (Batalha *et al.* 2011; Cianciaruso *et al.* 2012) e filogenética (Silva & Batalha 2010; Cianciaruso *et al.* 2012) da vegetação do cerrado.

Nós coletamos os dados que utilizamos nos trabalhos desta tese no Parque Nacional das Emais, Goiás. O Parque Nacional das Emas é uma das mais importantes reservas de cerrado do Brasil, com aproximadamente 133,000 ha. No parque, a vegetação de cerrado vai desde o campo limpo, onde predomina o componente herbáceo, até o cerrado *sensu stricto,* onde predominam os indivíduos arbustivo-arbóreos. O solo e a topologia do parque apresentam grande variação, contribuindo para estudos que relacionam ambiente e vegetação, como este. Além disso, o histórico de queimadas no Parque Nacional das Emas mostra uma configuração parecida com um mosaico. A fragmentação causada por aceiros e estradas de manutenção faz com que mesmo áreas próximas tenham históricos de fogo diferentes, contribuindo para a importância do parque nos estudos de formação e funcionamento de comunidades vegetais.

No primeiro capítulo da tese, nós investigamos os papéis direto e indireto do fogo e do solo na diversidade funcional e funcionamento de uma comunidade de espécies arbóreas de cerrado no Parque Nacional das Emas. Nós buscamos uma abordagem integrada, quantificando todas as relações, diretas e indiretas, entre os fatores abióticos, bióticos e funcionamento. Assim, propusemos um modelo de equações estruturais com uma representação plausível de como as variáveis de interesse se relacionam.

No segundo capítulo da tese, utilizamos modelos de máxima entropia para prever a abundância das espécies arbóreas de cerrado no Parque Nacional das Emas por meio de alguns de seus traços funcionais. Fazendo isso, pudemos particionar a importância de processos determinísticos e aleatórios formação dessas comunidades. Ainda nesse capítulo, determinamos quais foram os traços mais contribuiram para as previsões, indicando, assim, aqueles traços que mais influenciam as abundâncias.

No terceiro capítulo, procuramos explicar a riqueza e a densidade de indivíduos arbóreos no Parque Nacional das Emas por meio da relação entre traços fisiológicos e de resposta e o histórico de fogo e disponibilidade de nutrientes no solo.

Finalmente, no quarto capítulo, descrevemos o uso da modelagem de equações estruturais nas ciências ambientais, focando no uso de variáveis latentes, ainda pouco utilizadas. Procuramos apresentar os conceitos mais importantes da técnica, suas premissas, como estimar parâmetros e interpretar os resultados.

Escolhemos apresentar a tese em capítulos, que formatamos de acordo com as normas das revistas científicas a que foram submetidos. Como as revistas que escolhemos para publicá-los exigem a sua redação em inglês, nossos capítulos foram escritos nesse idioma. O primeiro capítulo foi submetido ao periódico *Methods in Ecology and Evolution*, o segundo capítulo foi submetido ao periódico *Oecologia,* o terceiro capítulo será submetido ao periódico *Community Ecology* e o último capítulo foi submetido ao periódico *Biology Letters*. A divisão em capítulos se justifica, porque torna a publicação dos artigos científicos menos trabalhosa e mais rápida, ainda que repetições sejam muitas vezes inevitáveis devido à independência dos capítulos.

REFERÊNCIAS BIBLIOGRÁFICAS

Amorim, P.K. & Batalha, M.A. (2007) Soil-vegetation relationships in hyperseasonal cerrado, seasonal cerrado, and wet grasslands in Emas National Park (central Brazil). *Acta Oecologica*, **32**, 319-327.

Batalha, M.A., Silva, I.A., Cianciaruso, M.V., França, H. & Carvalho, G.H. (2011) Phylogeny, traits, environment, and space in cerrado plant communities at Emas National Park (Brazil). *Flora*, **206**, 949-956.

Cadotte, M.W. (2011) The new diversity: management gains through insights into the functional diversity of communities. *Journal of Applied Ecology*, **48**, 1067-1069.

Cardinale, B.J., Nelson, K. & Palmer, M.A. (2000) Linking species diversity to the functioning of ecosystems: on the importance of environmental context. *Oikos*, **91**, 175-183.

Chapin, F.S., III., Walker, B.W., Hobbs, R.J., Hooper, D.U., Lawton, J.H., Sala, O.E. & Tilman, D. (1997) Biotic control over the functioning of ecosystems. *Science*, **277**, 500-504.

Chase, J.M. (2005) Towards a really unified theory for metacommunities. *Functional Ecology*, **19**, 182-186.

Cianciaruso, M.V., Batalha, M.A., Gaston, K.J. & Petchey, O.L. (2009) Including intraspecific variability in functional diversity. *Ecology*, **90**, 81-89.

Cianciaruso, M.V., Silva, I.A., Batalha, M.A., Gaston, K.J. & Petchey, O.L. (2012) The role of fire on phylogenetic and functional structure of woody savannas: moving from species to individuals. *Perspectives in Plant Ecology, Evolution and Systematics*, **14**, 205-216.

Connel, J.H. & Slatyer, R.O. (1977) Mechanisms of succession in natural communities and their role in community stability and organization. *The American Naturalist*, **111**, 1119-1144.

Cornwell, W.K. & Ackerly, D.D. (2009). Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs*, **79**, 109-126.

Díaz, S. & Cabido, M. (2001) Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, **16**, 646-655.

Fridley, J.D. (2001) The influence of species diversity on ecosystem productivity: how, where, and why? *Oikos*, **93**, 514-523.

Goodland, R. & Pollard, R. (1973) The Brazilian cerrado vegetation: a fertility gradient. *Journal of Ecology*, **61,** 219-224.

Gravel, D., Canham, C.D., Beaudet, M., Messier, C. (2006) Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters,* **9**, 399-409.

Hooper, D.U., Chapin, F.S.III., Ewel, A.H., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J. & Wardle, D.A. (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, **75**, 3-35.

Hooper, D.U. & Vitousek, P.M. (1997) The effects of plant composition and diversity on ecosystem processes. *Science*, **277**, 1302-1305.

Keddy, P.A. (1992) Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science*, **3**, 157-164.

Kraft, N.J.B., Valencia, R. & Ackerly, D.D. (2008) Functional traits and niche-based tree community assembly in an Amazonian forest. *Science*, **322**, 580-582.

Laughlin, D.C., Joshi, C., van Bodegom, P.M., Bastow & Fulé, P.Z. (2012) A predictive model of community assembly that incorporates intraspecific trait variation. *Ecology Letters*, **15**, 1291-1299.

Loreau, M. (2000) Biodiversity and ecosystem functioning: recent theoretical advances. *Oikos*, **91**, 3-17.

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J.P., Hector, A., Hooper, D.U., Huston, M.A., Raffaelli, D., Schmid, B., Tilman, D. & Wardle, D.A.

(2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, **294**, 804-808.

Macarthur, R. & Levins, R. (1967) The limiting similarity, convergence, and divergence of coexisting species. *The American Naturalist*, **101**, 377-385.

Mayfield, M. & Levine, J.M. (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters,* **13**, 1085-1093.

Pavoine, S., Vela, E., Gachet, S., Bélair, G. & Bonsall, M.B. (2011) Linking patterns in phylogeny, traits, abiotic variables and space: a novel approach to linking environmental filtering and plant community assembly. *Journal of Ecology*, **99**, 165-175.

Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters,* **5**, 402-411.

Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters,* **9**, 741-758.

Ruggiero, P.G.C., Batalha, M.A., Pivello, V.R. & Meirelles, S.T. (2002) Soil-vegetation relationships in cerrado (Brazilian savanna) and semideciduous forest, Southeastern Brazil. *Plant Ecology*, **160**, 1-16.

Shipley, B., Paine, C.E.T. & Baraloto, C. (2012) Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly. *Ecology*, **93**, 760-769.

Silva, D.M. & Batalha, M.A. (2008) Soil–vegetation relationships in cerrados under different fire frequencies. *Plant and Soil*, **311**, 87–96.

Silva, I.A. & Batalha, M.A. (2010) Woody plant species co-occurrence in Brazilian savannas under different fire frequencies. *Acta Oecologica*, **36**, 85-91.

Srivastava, D.S & Vellend, M. (2005) Biodiversity-ecosystem functioning research: is it relevant to conservation? *Annual Review of Ecology, Evolution, and Systematics*, **36**, 267-294.

Stokes, C.J. & Archer, S.R. (2010) Niche differentiation and neutral theory: an integrated perspective on shrub assemblages in a parkland savanna. *Ecology,* **91**, 1152-1162.

Sutherland, W.J., Freckleton, R.P., Godfray, H.C.J., Beissinger, S.R., Benton, T., Cameron, D., Carmel, Y., Coomes, D.A., Coulson, T., Emmerson, M.C., Hails, R.S., Hays, G.C., Hodgson, D.J., Hutchings, M.J., Johnson, D., Jones, J.P.G., Keeling, M.J., Kokko, H., Kunin, W.E., Lambin, X., Lewis, O.T., Malhi, Y., Mieszkowska, N., Milner-Gulland, E.J., Norris, K., Phillimore, A.B., Purves, D.W., Reid, J.M., Reuman, D.C., Thompson, K., Travis, J.M.J., Turnbull, L.A., Wardle, D.A. & Wiegand, T. (2013) Identification of 100 fundamental ecological questions. *Journal of Ecology*, **101**, 58-67.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences*, **101**, 10854-10861.

Tilman, D., Lehman, C.L. & Thomson, K.T. (1997) Plant diversity and ecosystem productivity: theoretical considerations. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 1857-1861.

Wardle, D.A., Bonner, K.I. & Nicholson, K.S. (1997) Biodiversity and plant litter: experimental evidence which does not support the view that enhanced species richness improves ecosystem function, *Oikos*, **79**, 247-258.

II - CAPÍTULO 1

Artigo submetido ao periódico Methods in Ecology and Evolution

Basics of structural equation modelling in Ecology

Running title: Structural equation modelling in Ecology

Gustavo H. Carvalho[1*], Marco A. Batalha[1], and Owen L. Petchey[2]

[1] Plant Ecology Laboratory, Department of Botany, Federal University of São Carlos, PO Box 676, São Carlos, 13565-905, SP, Brazil.

[2] Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057, Zurich, Switzerland.

---

[*] Correspondence author. E-mail: gustavo.bio@gmail.com

Summary

The statistical methods ecologists commonly use to analyse their data are often unsuitable for testing hypotheses were variables could act as cause and effect. Also, most of these methods do not take into account the multidimensionality of some common concepts in ecology, such as biodiversity and body size. Structural equation modelling (SEM) provides the means to quantitatively test hypotheses that represent alternative causal structures of any level of complexity. SEM allows researchers to analyse their data from a system perspective, with less emphasis on bivariate relationships and more focus on the system of interacting variables. Theoretical variables, such as biodiversity, are included in structural models as latent variables. Latent variables cannot be directly measured, but arise from the variance shared by a set of observed variables called indicators. Latent variable modelling also takes into account our imperfections in taking measurements and the effects of variables not included in our models on those included. These are modelled as error variances and measurement error in SEM and help on building more faithful mathematical translations of our hypotheses. SEM, thus, yields a powerful tool for gleaning knowledge on ecological theories, but requires a lot of attention from users to render trustworthy results. From model specification to fit, passing through data collection and screening, the researcher using SEM has to mind a number of assumptions. In this article, we explain the basic concepts and assumptions of SEM analysing worked data, focusing on avoidance of the common pitfalls. When used with the required caution, SEM is a powerful set of tools that can be used to generate models about the functioning of a system and, ultimately, reproduce, corroborate, and fine-tune these frameworks to generalise our knowledge and increase our understanding of such systems.

Key-words: structural equation modeling, latent variable analysis, path analysis, causal modelling.

Introduction

Ecologists commonly use standard univariate (for example, ANOVA, regression, GLMs) and multivariate (such as principal component analysis, cluster analysis, and canonical correspondence analysis) methods to analyse their data. These methods are often unsuitable for testing hypotheses in which a variable can act as cause and effect, and hypotheses about chains and networks of causation (Fig. 1). Furthermore, standard methods usually give estimators of net effects instead of causal relationships and cannot accommodate theoretical ideas that are not directly measurable (Kline 2010). Structural equation modelling (SEM), in contrast, provides the means to quantitatively test hypotheses that represent alternative causal structures, be them simple, with only a handful of variables, or more complex, with several variables, some of which being both causes and effects. The hypotheses that can be tested via SEM are often described using diagrams that describe the causal assumptions of the researcher regarding a set of variables (Fig. 1). SEM allows researchers to look at their data from a systems perspective, with less emphasis on bivariate relationships and more focus on the system of interacting variables (Kline 2010).

Sewall Wright proposed the first type of SEM in the beginning of the 20th century (Wright 1918, 1921, 1937). At that time, Pearson's school of statistics and correlation dominated: causation was seen as a useless concept (Shipley 2000). Critics of Wright's proposal thought he was trying to infer causation from correlations (Niles 1922). Wright, however, was trying to estimate the strength of associations between variables in a system that may include causal knowledge and beliefs (Wright 1921). A few years later, Sir Ronald Fisher developed methods for inferring causation based on randomisation and experimental control (Fisher 1925). Fisher's methods became dominant, and Wrights methods were relatively forgotten.

SEM has recently gained popularity, especially in the social sciences where it is now well established. In ecology and evolution, SEM is less established, even though studies show its efficacy for gleaning knowledge about the prior causal relationship assumed to result in observed data. It has been used to analyse causes of, for example, plant species richness (Grace & Pugesek 1997), plant species succession (Vile *et al.* 2006), the abundance of rabbits and track numbers of Iberian lynx and Egyptian mongoose (Palomares *et al.* 1998), and the overnight survival of house sparrows (Pugesek & Tomer 1996). Furthermore, several books and papers provide excellent guides to the philosophy, principles, and methods involved with SEM (Grace 2006, Kline 2010, Schumacker & Lomax 2004, Shipley 2000).

Unfortunately, SEM can be complex and has numerous pitfalls. Possible source of problems go from not respecting the underlying assumptions of the method to failure to properly report results. Some of the common pitfalls have serious impacts on the validity and reproducibility of the analyses. Here we provide an overview to the basic concepts of SEM, focused on avoidance of the common and likely pitfalls. This overview we illustrate with reference to a worked example, with data and complete R code available as a supplement. We regularly indicate sources of the more in depth topics that we cannot cover.

Overview of the method

SEM emphasises estimating the strengths of causal effects between variables through path coefficients (Grace *et al.* 2010). A deeper discussion about causation can be found in Shipley (2000), where he suggests that the common adage that ‘correlation does not imply causation’ should be revised to ‘correlation implies an unknown causative structure’. Pearl (2000) further explains causality, its statistical foundations, and its link with SEM. He also presents the technique of causal diagrams for testing causal assumptions from correlations. Here we use the term ‘causal assumptions’ as a reference to all the prior knowledge a researcher has about certain phenomena. Accordingly, we use ‘causal effects’ or simply ‘effects’ as the strength of such causal assumptions. The purpose of SEM is not to establish cause from association. Instead, it provides ways of statistically testing the strength of support for causal assumptions.

The first step in SEM is model specification: formulation of a hypothesis about the causal relationships among the variables of interest (for instance, Fig. 1), with each causal relationship representing a plausible ecological process. The variables might be directly observable or unmeasured theoretical variables (see below). Similarly to more standard analyses, hypotheses are ideally created before data collection. The second step in SEM is model identification, whereby one examines if the specified model can be solved. The third step is checking the requirements of the observed data, including examination of sample size, distribution, collinearity, and outliers in and among variables. The fourth step is estimation of model (= hypothesis) fit and parameter values. Model fit informs about how well the proposed model explains the observed data and can be used to assess the support among a set of causal structures. Parameter values indicate how much response variables change with changes in explanatory variables.

Throughout this article, we focus on maximum likelihood estimation (MLE) to estimate parameter values and maximum likelihood chi-squared (MLχ^2) to assess model fit. Used this way, SEM gives estimates of how well the models proposed explain the observed relationships among variables, of the direct and indirect effects of one variable on another, and of how much of the variances and covariances are not explained by the model.

If models are not specified in advance, SEM can assist in data exploring by searching for a causal structure (model) that best explains the data. The usual dangers associated with exploratory analyses apply and interpretations of data produced by exploratory SEM should be tested using new, independent data (Schumacker & Lomax 2004).

SEM Software

A number of statistical software packages are capable of implementing structural equation models. Some of the most popular packages include LISREL (Jöreskog & Sörbom 2006), EQS (Bentler 1995), Mplus (Muthén & Muthén 1998-2010), R (R Development Core Team 2012), and SPSS Amos (Arbuckle 1995-2009). The software is quite user-friendly and can analyse models of varying complexity using different techniques to assess model fit and to estimate parameters (Schumacker & Lomax 2004). The R environment for statistical analysis provides a few choices of SEM package. OpenMX (Boker *et al.* 2011) is one of the most actively developed. It has fewer parameter estimators than commercial packages, but is perfectly capable in most usage scenarios and can be expanded to include more estimators. Other SEM packages for R include sem (Fox *et al.* 2012) and lavaan (Rosseel 2012). lavaan is quite new, but already offers several different types of estimation methods and is, feature-wise, on a par with commercial packages. It also has useful supporting functions, including bootstrapping and simulation of data, and is one of the easiest to use. The package sem has fewer features, but was the first package to offer SEM support for R and was recently rewritten. Due to the popularity of R, we have chosen it and the lavaan package to present a step-by-step guide to performing structural equation modelling (supplementary files). The package stremo (Carvalho *et al.* 2011) for R has a few published datasets that can be used alongside lavaan and is a good starting point for those who wish to learn SEM in R.

Model specification

Specification of the structural part of an structural equation model

Translating previous knowledge, theories, ideas, and hypotheses into a causal structure to be tested using SEM is probably the most difficult part of using the method. The researcher can do this translation by starting with variables and then identifying hypotheses, theories, or processes that might connect them. This initial step is greatly aided by drawing the causal structure in a graphical representation (Fig. 1). In such a drawing, arrows (hypothetical causal effects) connect variables (shapes) to give a causal structure to the model. Variables in rectangles are termed observed variables because they are directly measurable. For instance, Shannon diversity (shannon rectangle), and nitrogen content (nitrogen rectangle) are observed variables because they are directly measurable. A structural equation model may contain only directly measurable variables. Grace (2006) argues that *path analysis* (structural equation model with only directly measurable variables) considers only one level of the theory being tested, the relationships between the components of interest, as they do not account for error variances and variables that cannot be directly measured (i.e., *latent variables*). Thus, latent variables add a powerful dimension to structural equation models.

Biodiversity can be represented as a latent variable, for example, since it is a multidimensional concept that cannot be reduced to a single value (Purvis & Hector 2000). Concepts that are not directly measurable (latent variables, *constructs* or *factors*) are graphically represented by ellipses (soil fertility, biodiversity, and community functioning in Fig. 1). Since latent variables cannot be measured, we use other variables as their *indicators* or *proxies*. Each indicator variable can be seen as a dimension of a multidimensional concept. In Fig. 1, biodiversity has three indicators: shannon, which corresponds to Shannon's diversity index values; FD, a functional diversity index (Cianciaruso *et al.* 2009); and phylodiversity, a phylogenetic diversity index (Allen *et al*. 2009). Each of these indicators represents a different facet of biological diversity. Shannon diversity takes into account the identity and proportions of the species in a given sample, FD accounts for the trait variation among individuals, and phylodiversity is a measure of the phylogenetic diversity in a community. The two other latent variables in Fig. 1, soil fertility and community functioning, each have indicators that correspond to aspects of the hypothesised construct.

After drawing the causal network, some variables will have arrows only going out (termed an *exogenous variable*, e.g., soil fertility in Fig. 1), some will have arrows going in and some will have arrows going in an out of them. The latter two are termed *endogenous variables* (e.g., biodiversity and community functioning in Fig. 1).

Any two variables can be directly (e.g., soil fertility and biodiversity in Fig. 1) or indirectly connected (soil fertility and community functioning through biodiversity). Variables are sometimes both directly and indirectly connected (soil fertility and community functioning). Variables only indirectly linked are *fully mediated* whereas variables with both direct and indirect causal paths are *partially mediated*. The absence of causal links between two variables is a very strong assumption. When causally connected, the path linking two variables is usually free to take on any value during estimation. However, if two variables are not causally connected, the researcher imposes a fixed path coefficient of zero. This will lead to a predicted bivariate covariance equal to zero if there are also no indirect paths between them.

Incorrectly translating previous knowledge into a model may lead to rejection of the proposed model due to its poor fit. In SEM, making as few as one parameter fixed or free can cause drastic changes in fit. The directionality of the paths is equally influential. Models where all variables are only at one end of a causal chain (for instance, $A \rightarrow B \rightarrow C$) are called *recursive*. Models where a given causal chain can have the same variable at its both ends (for instance, $A \rightarrow B \rightarrow C \rightarrow A$) are called *nonrecursive*. Nonrecursive models often arise from uncertainty about directionality leading to inclusion of bi-directional causality or covariances between endogenous variables. Nonrecursive models require special attention and tend to be problematic, so they are best avoided unless the nonrecursiveness is strongly backed by theory, which is seldom the case.

Careful translation of theory and previous knowledge into a structural equation model at this step is crucial since everything following assume that the model is correct (Kline 2010). Sometimes the knowledge driving the specification of a model provides room for alternative causal structures. This can be accommodated with the specification of competing models or with a list of possible modifications of the causal paths to improve model fit later on. It is important to come up with competing equivalent models at this point to make sure that theory and not an exploratory search for the model with the best fit is driving specification. In cases where exploratory latent variable analysis is indeed the aim of the researcher, he or she must make the rationale behind this choice very clear.

Finally, more often than not data will already have been collected before model specification and identification. In this situation, the researcher has to be extra careful and propose models that are consistent with theory instead of trying to arrange the variables in a way that looks plausible enough and hope for the best. Here, we create data and specify a very simple yet plausible model to use as example. Thus, it is not our intention to thoroughly test theories regarding causal links between the environment, biological diversity, and community functioning. A very detailed discussion on the specification of ecological structural equation models can be found in Grace *et al.* (2010). Shipley (2000) and Grace (2006) also address this topic in great detail.

Specifying the measurement part of structural equation model

The measurement model consists of the latent variables and their indicators (Fig. 2). Indicators of latent variables are also known as *manifest variables*. The arrows pointing from a latent variable to its indicators are termed *loadings*. This terminology arises from the idea that the immeasurable latent variable loads or manifests itself on the indicators, thus causing them.

A latent variable arises from the variance shared by its indicators. Take, for example, the measurement model in Fig. 2a. The grey area in Fig. 2b represents the latent variable in Fig. 2a. Variance not shared by indicators (unshaded areas in Fig. 2b) does not go into the definition of the latent variable. This unique variability, which is represented by the Greek letter *epsilon* (ε, Fig. 1, 2, and 4), is referred to as *measurement error*. Measurement error may represent random error and imprecisions during data collection and also the effects of omitted variables that have some effect on the indicators. The same omitted variable can affect multiple indicators, making error terms correlated. Researchers can include this correlation by linking variables with two-headed arrows (δ_1, δ_2, and δ_3 in Fig. 2). Explicitly thinking about and modelling these error terms provides SEM a very important advantage over other methods.

Latent variables also have error terms. For exogenous latent variables (e.g., soil fertility in Fig. 1), they are called variance and are represented by a two-headed arrow beginning and ending at the latent variable. The variance of an exogenous latent variable is a proxy for the influence of all the omitted variables, much like the error of an indicator. Error associated with an endogenous latent variable (biodiversity and community functioning in Fig. 1) is termed disturbance and is the amount of factor variance not explained by the model. Disturbances often are represented by the letter D (Fig. 1 and 2a).

Take, for instance, the latent variable biodiversity in Fig. 1. It is well acknowledged that there is no single number that serves as a perfect measure of the biological diversity of a sample (Purvis & Hector 2000). Instead, ecologists rely on several indices to represent the multidimensionality of this concept. So, we used three indices as indicators of biodiversity, each representing a dimension. The values we get for the indices cannot be fully explained by the latent variable though. This unaccounted indicator variance is known as measurement error in SEM jargon. As an example, an index of phylogenetic diversity does not perfectly explain the phylogenetic relationships between the species of a community. These method imperfections (e.g, poorly resolved phylogenies) are modelled as measurement error in SEM.

Careful consideration of measurement error and of the distinction between latent and observed variables may contribute to the refinement and maturation of ecological theories. Instead of including biodiversity as a latent variable, we might have included only Shannon diversity. This would have resulted in loss of information about the multidimensionality of the concept of biodiversity, and loss of information about our imprecisions in measuring this concept. Both losses would result in a less realistic mathematical translation of current theory.

Since we assume indicators to be caused by the latent variable, all bivariate correlations between indicators are expected to be somewhat high. Indicators with too little or too much shared variance are likely to be improper and ought to be reconsidered. The number of indicators of each latent variable is also of great importance. Models with latent variables with fewer than three indicators are more prone to problems during identification and estimation. When possible, structural equation models should include at least three indicators for each latent variable. If a latent variable has only one indicator, the error variance of the single indicator is not modelled. Unless the error variance is pre-set due to previous knowledge, measurement error is not considered. Finally, contrarily to manifest variables, latent variables do not have scales. To set a scale to a latent variable, the researcher has to fix either the variance of the latent variable or one of the loadings. This choice does not affect model fit.

Model identification

A model is identifiable if it is possible, in theory at least, to identify a unique solution (i.e, one best set of parameter estimates). The maximum number of free parameters in a model is given by the *t-rule:* $t = n(n+1)/2$, where n is the number of observed (i.e., not latent) variables. The model in Fig. 1 has nine observed variables, thus $t = 9(9+1)/2 = 45$. The number of parameters to be estimated is the number of arrows without fixed values. The model in Fig. 1 has 21 arrows (the variance of soil fertility and the loadings of shannon and leaf N are fixed to 1 to set the scales of the latent variables), and therefore 21 free parameters. The number of degrees of freedom of a model is the maximum number of free parameters that could be estimated minus the number that will be estimated. Thus, the model in Fig. 1 has 45 – 21 = 24 degrees of freedom. When a model has a positive number of degrees of freedom, it is *under-identified*, meaning that the covariance matrix provides more information than the model needs. A *just-identified* model has zero degrees of freedom, so it is saturated. Saturated models always have perfect fit to the data, which makes them often pointless. *Over-identified* models have negative number of degrees of freedom, no unique solution, and cannot be tested.

Positive number of degrees of freedom does not guarantee identifiability. Empirical under-identification (Kline 2010) arises during parameter estimation and can be caused by problems in data. For instance, two highly correlated variables reduce the amount of information available in the data. Specification issues may also lead to an under-identified model, but these are hard to predict beforehand. When such problem arises, one has to make small changes to the model to narrow the source of under-identification. Kline (2010) provides further guidance for models with identifications problems, including empirical under-identification. For example, one might fix the variances and covariances of all exogenous variables in a model, freeing up degrees of freedom and possibly solving identification problems.

Checking data requirements

SEM with MLE compares an observed covariance matrix with a model-implied one. The smaller the differences between the two matrices, the better the model. When using MLE, data can be provided as a correlation matrix (upper triangle in Table 1) with standard deviations (diagonal in Table 1), as a covariance matrix (lower triangle in Table 1), or raw (a matrix with sample units in rows and variables in columns). Even though covariances are used in the modelling, the raw data must not deviate from the following assumptions:

Type

MLE requires normally distributed, and therefore continuous variables. Other types of data common in ecology, like categorical and count data, can be estimated with alternative estimators, especially from the weighted least squares family. Most introductory books give guidance when models have non-continuous variances.

Sample size

SEM needs large samples to provide accurate standard errors of parameter estimates. As a rule of thumb for MLE, sample size should be > x20 the number of free parameters. A model with 15 parameters to be estimated ideally would have 300 cases. Bentler (1995) suggests that the minimum sample size is x5 the number of free parameters (100 for a model with 20 free parameters). For a small discussion on the implications of samples of different sizes on parameter estimates, see Shipley (2000). In situations where it is not possible to collect a sufficient amount of data for MLE, the researcher can resort to MLE coupled with bootstrapping methods to generate unbiased standard errors of parameter estimates (Shipley 2000).

Collinearity

As usual, redundant variables are problematic and they should be eliminated or combined. The same applies for variables with strong bivariate correlations. Scatterplots are useful for identifying strong correlation. Multivariate collinearity is harder to spot. Kline (2010) suggests building several multiple regression models, each with a different variable as the response and all other variables as explanatory variables. A high coefficient of determination is an indication of multicollinearity. Another way to detect multicollinearity in linear models is by using the variance inflation factor (VIF; Fox & Monette 1992). VIF values greater than 10 suggest that there is multicollinearity (Quinn & Keough 2002). Multicollinearity is dealt with very much the same way as bivariate collinearity. One can either remove one of the variables that contribute to the issue or combine related variables into a composite one.

Normality

MLE requires normally distributed variables in multivariate space. Univariate normality can be measured by the skewness and kurtosis of the distribution, can be viewed with histograms and density plots, and can be tested for with Kolmogorov-Smirnov test of goodness-of-fit or the Shapiro-Wilk test (Legendre & Legendre 1998). Univariate normality does not guarantee multivariate normality (Legendre & Legendre 1998). The generalised Shapiro-Wilk's test proposed by Villasenor Alva & Estrada (2009) can check for multivariate normality.

Highly skewed or kurtotic data can be transformed. See Zar (2009) for further information. Alternatively, one can use robust maximum likelihood estimators (e.g., Satorra & Bentler 1994), asymptotic distribution-free estimators, or bootstrapping. Shipley (2000) discusses the issue of non-normality in biological datasets and provides guidance to alternative estimation methods.

Outliers

Outliers can cause problems during parameter estimation. Identifying them and their cause is often useful, therefore. Bivariate scatterplots, boxplots, and dotplots help find outliers. Another method is calculation of z-scores ($z = (x - \overline{x})/s$), where x is the raw score to be standardised, $\overline{x}$ is the mean of the sample, and s is the standard deviation of the sample) for each variable in the set. In a normal distribution, 0.27% of the absolute standard scores are expected to have values greater than 3.00. Deviations from this expected frequency (for instance, 1 when $n = 370$) indicate the presence of possible outliers. MLE requires unstandardised data, so z-scores should be used only to detect outliers.

Multivariate outliers are not necessarily outliers in the univariate frequency distributions. When data are normally distributed in multivariate space, robust multivariate Mahalanobis distances are expected to follow a χ^2_{df} distribution (Filzmoser *et al.* 2005). Data that fall outside of the expected distribution may be outliers.

If outliers are present, one should carefully consider their source (sampling from a different population, wrong data entry, etc) and then take action, such as data transformation, outlier removal, or alternative estimators.

Covariance matrix

The covariance matrix used as input in SEM software must be positive-definite. If you are passed a covariance matrix or are using one published in the literature, it is well worth checking this. Some of the characteristics of positive-definite matrices are: 1) all covariances and correlations are within bounds; 2) all eigenvalues (λ_n) are positive; and 3) the determinant of the matrix is positive. These requirements are related to the iterative process during MLE, where the matrix is inverted several times.

Please see the supplementary files for more information on how to detect and deal with data related problems. Grace (2006) and Shipley (2000) further discuss the common issues posed by biological data in the context of SEM.

Model estimation, fit, interpretation, and modification

At this point, the model is specified and identified, there are sufficient data, and it meets the assumptions of the estimator (here MLE). The next step is to pass to the estimation software the covariance matrix and the proposed model structure. Although some software accept a correlation matrix, the analysis of correlations requires constraints in the model and should be avoided whenever possible. If raw data are available, it is best to analyse covariances.

Testing the measurement model using CFA

Instead of immediately estimating the full structural equation model (one-step method), it is best to follow a multi-step approach. Multi-step approaches can help in finding sources of poor fit. We present the two-step method proposed by Anderson & Gerbing (1988). The first step is testing the measurement part of the structural equation model with a *confirmatory factor analysis* (CFA) (e.g, Fig. 3). CFA is a type of SEM that focuses on the relationships between latent variables and their indicators. In a CFA, all factors, that is are normally free to covary (e.g, two-headed arrows between latent variables in Fig. 3). The objective of this step is to find an adequate measurement model before moving on to the structural model, so re-specification due to poor fit is justified. A CFA indicates the need for model re-specification when indicators have low standardised loadings, which may lead to poor fit with data.

One of the most important steps in parameter estimation is choosing adequate starting values of free parameters. If the programme is given bad starting values, even the most robust routines will fail to converge to an acceptable solution. Most software will automatically propose adequate starting values. When this automation fails, it is up to the researcher to analyse each causal relationship and to try to guess a plausible starting estimates. Kline (2010) gives information on a way of manually guessing an adequate set of starting values.

From the starting parameter values, the estimator selects a set of close-by parameter guesses, builds the expected covariance matrix, ($\tilde{S}$) and compares it to the observed covariance matrix (S, Table 1). Programmes make this comparison by feeding both matrices to a fitting function. For MLE, the fitting function is $F_{ML} = \ln \tilde{S} - \ln S + \mathrm{tr}\left[S\tilde{S}^{-1}\right] - (NX + NY)$, where NX is the number of observed exogenous variables and NY is the number of observed endogenous variables. The closer S and $\tilde{S}$ are, the lower is the value of F_{ML}. When the estimator finds the set of estimates that gives the minimum value of F_{ML}, it stops and returns the parameters.

The standard error of each parameter estimate is calculated using a Hessian matrix (a by-product of the iterative fitting process). Standard errors provide access to the significance of each estimate with simple t-tests. The value of fitting function F_{ML} provides the MLχ^2, which is a measure of model fit with data. ML$\chi^2 = N \times F_{ML}$ where N is the number of observations in the dataset.

When a model has adequate fit, the value of MLχ^2 has a P value greater than 0.05: the observed and model-implied covariance matrices are not statistically different. The model in Fig. 3 had a non-significant MLχ^2 value ($n = 210$, ML$\chi^2 = 22.98$, $P = 0.52$, and $df = 24$).

Before testing the structural part of the model, it is sensible to compare observed and model implied bivariate correlations via the *residual matrix* (Table 2), obtained by subtracting the observed correlation matrix (S) from the model implied one ($\hat{S}$). Too large absolute differences [Kline (2010) suggests a threshold of 0.10] indicate that the model does not properly accommodate certain bivariate correlations, and can occur even when the MLχ^2 statistic suggests good fit with data. In such cases, re-specification is warranted.

If the model does not have an acceptable fit, the researcher can consider alternative models by freeing or setting some parameters. However, significant changes to the structure of the model following poor fit should be reported in detail and supported by sound theory, as it is easy to obtain good fit by adding or removing causal paths.

Following a satisfactory CFA, one should interpret the coefficients and assess latent variable validity. Indicator loadings are interpreted as regression coefficients. For instance, in Fig. 3, the unstandardised loading of the latent variable on the indicator phosphorus is 2.470, so an increase of 1 standard deviation in the latent construct will cause an expected 2.470 standard deviation increase in the indicator. Unstandardised path coefficients may not be easily compared due to differences in the scales of variables. Standardised effects, on the other hand, are more easily compared since they are all in the same scale. Path coefficients can be either positive or negative, indicating the nature of the relationships between two variables. Negative error variances and disturbances are, however, an indication of *Heywood cases*, in which the software converges to a solution without any apparent problems, but closer inspection reveals interpretable results such as negative error variances. Possible causes of Heywood cases include outliers, misspecification, nonidentification, poor starting values, and small sample size in combination with too few indicators per latent variable (Kline 2010).

In the model in Fig. 3, only the covariance between soil fertility and community functioning was not significantly different from zero ($P < 0.05$). Since we were interested only in testing factor loadings before fitting the structural model, we left the non-significant parameter in the model. If we were interested in this relationship, two routes are available: keeping the parameter or removing it to try to improve model fit. However, non-significant parameters are often best left, because even though they are not statistically different from zero, they contribute to lower bivariate correlational residuals, rendering better fit with data (Kline 2010).

In the measurement model, in addition to the significance of each loading, one should examine the *reliability* of the indicators: the amount of indicator variance explained by the latent variable. For instance, the reliability of the variable phosphorus is given by one minus its standardised error variance (1 – 0.375 = 0.625; Fig. 3). Thus, soil fertility accounts for 62.5% of the variance of the manifest phosphorus. Reliability of an indicator is can also be calculated by squaring its standardised loading. One should always consider all reliabilities in a model and weigh up if the amount of variance explained is consistent with the theory driving the specification. When reliability is low, one should reconsider one's choice of latent or manifest variables. See Grace (2006) and Grace *et al*. (2010) for more on reliability in ecological models.

Fitting the structural equation model

Following a satisfactory CFA, one proceeds to solving the structural equation model proposed. The only difference between the CFA and the structural equation model is the relationships between latent variables: the causal relations are used. Model fit and parameter estimates are calculated and interpreted similarly. Any changes made to the measurement model resulting from the CFA should be incorporated into the full structural equation model in the second step. Since our CFA model was adequate, we did not make any changes to our measurement model.

The structural equation model in Fig. 1 had good fit with data ($n = 210$, $\mathrm{ML}\chi^2 = 22.98$, $P = 0.52$, and $df = 24$; Fig. 4). All parameters except the regression of soil fertility on community functioning ($\beta = -0.06$, $P = 0.66$; Fig. 4) were significant. Removing this parameter from the model (i.e., setting its value to zero) only marginally increases the value of the $\mathrm{ML}\chi^2$ statistic ($n = 210$, $\mathrm{ML}\chi^2 = 23.17$, $P = 0.57$, and $df = 25$). Thus, *nested models*, which are models with the same set of variables but different causal configurations, can be compared by subtracting the value of the $\mathrm{ML}\chi^2$ statistic of the more restrictive model from the $\mathrm{ML}\chi^2$ of the less restrictive model, with significance assessed in χ^2. In our case, the less restrictive model is the one in Fig. 1, since the path between soil fertility and community functioning is free (in the more restrictive model, the value of this parameter is fixed at zero). The χ^2 difference in our case was $\chi^2 = 23.17 - 22.98 = 0.19$, for one degree of freedom ($df = 25 - 24$). The P value associated with this difference is 0.66. Thus, both models have essentially the same fit with data. Analysis of residuals also suggested the absence of any noteworthy differences between models. We retained the nonsignificant parameter since it makes the model more consistent with previous knowledge, and possibly more generalisable to other datasets.

Good fit is an indication that the model is capable of reproducing observed bivariate relationships (Kline 2010), but does not necessarily indicate explanation of a large amount of the variability of endogenous variables. For instance, the model in Fig. 4 explains only about 3% of the variance of the latent variable biodiversity ($R^2 = 1 - 0.966 = 0.03$; Fig. 4). The R^2 of an endogenous variable is given by subtracting its standardised variance from 1. Such low explained variance suggests that soil fertility alone is not a major cause of biodiversity. If we are interested in explaining biodiversity, we should revisit the literature and include more predictors of biodiversity. All other variables had a high amount of their variances explained by the model ($R^2 > 0.5$).

The models in Fig. 3 and 5 are *equivalent models*. In the context of SEM, equivalent models are models with the same set of variables but with different path directionality. The main characteristic of equivalent models is that they produce the same value of fit statistics. For instance, the difference between the models in Fig. 3 and 5 is the paths between latent variables. The choice between mathematically equivalent models is, thus, based purely on theory and one should thoroughly explain motives behind this decision.

Alternative measures of fit

The MLχ^2 statistic is a test of the exact-fit hypothesis between the observed and predicted covariance matrices. Measures of approximate fit describe the degree of discrepancies between model and data. The rationale for reporting approximate fit measures is that the power of the chi-square statistics increases with large sample size, resulting in even the slightest deviations between observed and predicted covariance matrices causing failure to fit statistics. Shipley (2000) suggests that approximate fit measures are used after a model has failed exact-fit tests, to show how close the rejected model is from a baseline model that fully reproduces observed data. He adds that a major problem with these tests is the lack of evidence for the underlying assumption that small specification problems translate into small deviations between data and model-implied covariances. Also, there are no statistical tests to determine the cut-off values of approximate fit measures. Researchers have to rely on the results of simulations as thresholds of how "approximate" the model is from a correctly specified model.

Conclusion

When used with the required caution, SEM yields a powerful set of tools that can be used to generate models about the functioning of a system and, ultimately, reproduce, corroborate, and fine-tune these frameworks to generalise our knowledge and increase our understanding of such systems (Grace *et al*. 2010).

References

Allen, B., Kon, M., & Bar-Yam, Y. (2009) A new phylogenetic diversity measure to generalizing the Shannon index and its applications to phyllostomid bats. *The American Naturalist*, **174**, 236-243.

Anderson, J. C. & Gerbing, D. W. (1988) Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, **103**, 411-423.

Bentler, P. M. (1995). *EQS structural equations program manual*. Multivariate Software, California.

Boker, S. M., Neale, M. C., Maes, H. H., Wilde, M. J., Spiegel, M., Brick, T. R., Spies, J., Estabrook, R., Kenny, S., Bates, T. C., Mehta, P., & Fox, J. (2011) OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika*, **76**, 306-317.

Carvalho, G. H., Batalha, M. A., & Petchey, O. L. (2011) *stremo: Functions to help the process of learning structural equation modelling*. URL http://cran.r-project.org/web/packages/stremo/index.html [accessed on 12 April 2012]

Cianciaruso, M. V., Batalha, M. A., Gaston, K. J., & Petchey, O. L. (2009) Including intraspecific variability in functional diversity. *Ecology*, **90**, 81-89.

Filzmoser, P., Garrett, R. G., & Reimann, C. (2005) Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences*, **31**, 579-587.

Fisher, R. A. (1925) *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.

Fox, J. & Monette, G. (1992) Generalized collinearity diagnostics. *JASA*, **87**, 178-183.

Fox, J., Nie, Z., & Byrnes, J. (2012) *sem: Structural Equation Models*. URL http://cran.r-project.org/web/packages/sem/index.html [accessed on 12 April 2012]

Grace, J. & Pugesek, B. H. (1997) A structural equation model of plant species richness and its application to a coastal wetland. *The American Naturalist*, **149**, 436-460.

Grace, J. (2006) *Structural Equation Modeling and Natural Systems*. 1st edn. Cambridge University Press, New York.

Grace, J., Anderson, T. M., Olff, H., & Scheiner, M. (2010) On the specification of structural equation models for ecological systems. *Ecological Monographs*, **80**, 67-87.

Jöreskog, K. G. (1970) A general method for analysis of covariance structures. *Biometrika*, **57**, 239-251.

Jöreskog, K. G., & Sörbom, D. (2006) *LISREL 8.80 for Windows*. Scientific Software

International, Lincolnwood. URL http://www.ssicentral.com/ [accessed 12 February 2012]

Kline, R. (2010) *Principles and Practice of Structural Equation Modeling*. 3rd edn. The Guilford Press, New York.

Legendre P. & Legendre L. (1998) *Numerical Ecology*. 2nd edn. Elsevier Science, Amsterdam.

Muthén, L. K., & Muthén, B. O. (1998–2010) *Mplus user's guide.* 6th edn. Muthén & Muthén, Los Angeles.

Niles, H. E. (1922) Correlation, causation and Wright's theory of 'path coefficients'. *Genetics*, **7**, 258-273.

Palomares, F., Ferreras, P., Travaini, A. & Delibes, M. (1998) Co-existence between Iberian lynx and Egyptian mongooses: estimating interaction strength by structural equation modelling and testing by an observational study. *Journal of Animal Ecology*, **67**, 967-978.

Pearl, J. (2000) *Causality: Models, Reasoning, and Inference*. 1st edn. Cambridge University Press, Cambridge.

Pugesek, B. H. & Tomer, A. (1996) The bumpus house sparrow data: A reanalysis using structural equation models. *Evolutionary Ecology*, **10**, 387-404.

Purvis, A. & Hector, A. (2000) Getting the measure of biodiversity. *Nature*, **405**, 212-219.

Quinn, G. P. & Keough, M. J. (2002) *Experimental Design and Data Analysis for Biologists*. 1st edn. Cambridge University Press, Cambridge.

R Development Core Team (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. URL http://www.R-project.org [accessed 12 February 2012]

Rosseel, Y. (2012) lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, **48**, 1-36.

Satorra, A. & Bentler, P. M. (1994) Corrections to test statistics and standard errors on covariance structure analysis. *Latent variables analysis* (eds A. von Eye & C. C. Clogg), pp 399-419, Sage, California.

Shipley, B. (2000) *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations, and Causal Inference.* 1st edn. Cambridge University Press, Cambridge.

Shumacker, R. E. & Lomax, R. G. (2004) *A Beginner's Guide to Structural Equation Modeling*. 2nd edn. Lawrence Erlbaum Associates, New Jersey.

Vile, D., Shipley, B. & Garnier, E. (2006) A structural equation model to integrate changes in functional strategies during old-field succession. *Ecology*, **87**, 504-517.

Villasenor Alva, J. & Estrada, E. G. (2009) A Generalization of Shapiro-Wilk's Test for Multivariate Normality. *Communications in Statistics: Theories and Methods*, **38**, 1870-1883.

Wright, S. (1918) On the nature of size factors. *Genetics*, **3**, 367-374.

Wright, S. (1921) Correlation and causation. *Journal of Agricultural Research*, **10**, 557-587.

Wright, S. (1937) The method of path coefficients. *The Annals of Mathematics Statistics*, **5**, 161-215.

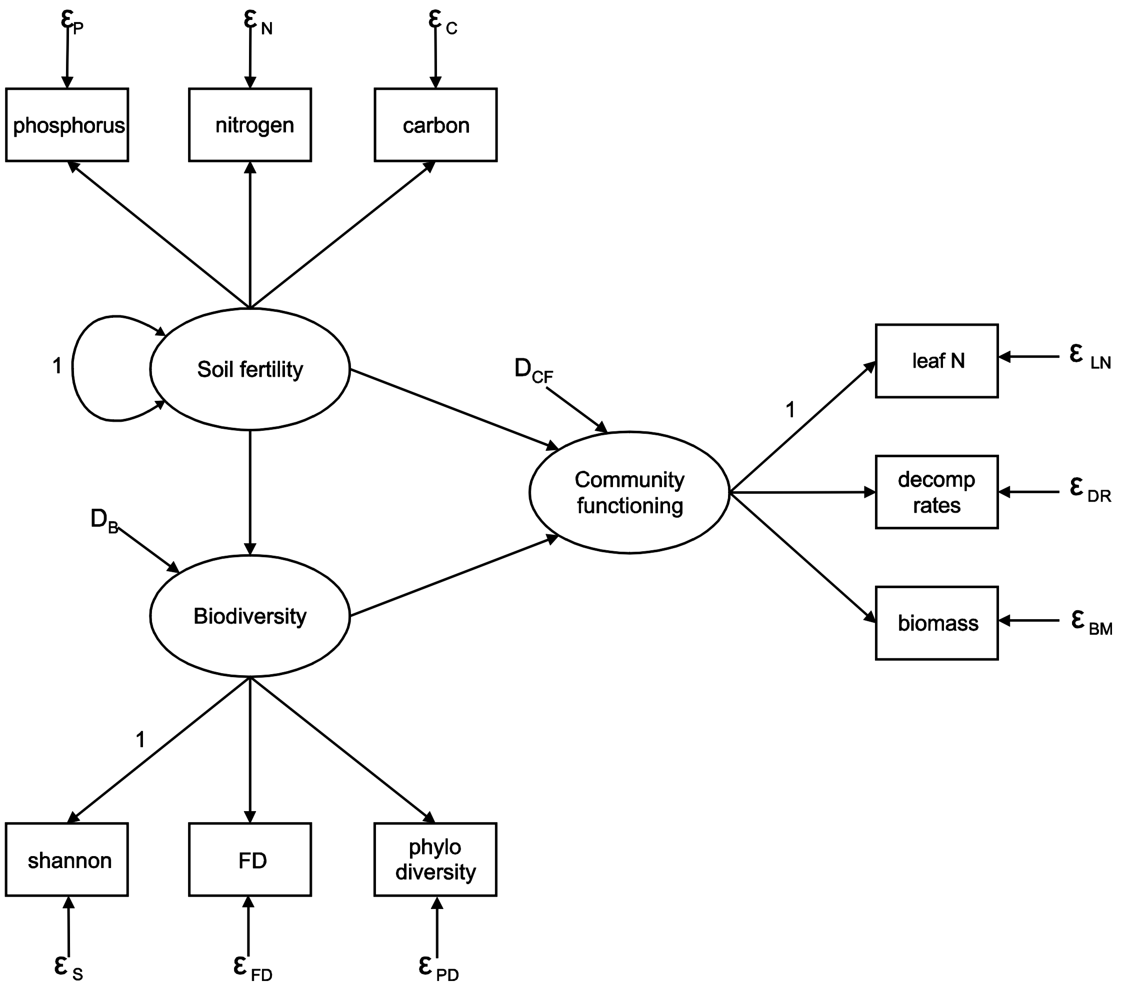
Zar, J. H. (2009) *Biostatistical Analysis*. 5th edn. Prentice Hall, New Jersey.

Table 1. Covariances (lower triangle), correlations (upper triangle, bold), and variances (diagonal) of the simulated observed variables of the structural equation model in Fig. 1. The dataset was normally distributed in multivariate space. phosphorus = soil phosphorus (mg kg^{-1}), nitrogen = soil nitrogen (mg kg^{-1}), carbon = soil carbon (mg kg^{-1}), shannon = Shannon's diversity index, FD = functional diversity index, phylodiversity = phylogenetic diversity index, leaf N = leaf nitrogen content (mg g^{-1}), decomp rates = decomposition rates of recently shed leaves, and biomass = live tissue biomass estimate (g m^{2}).

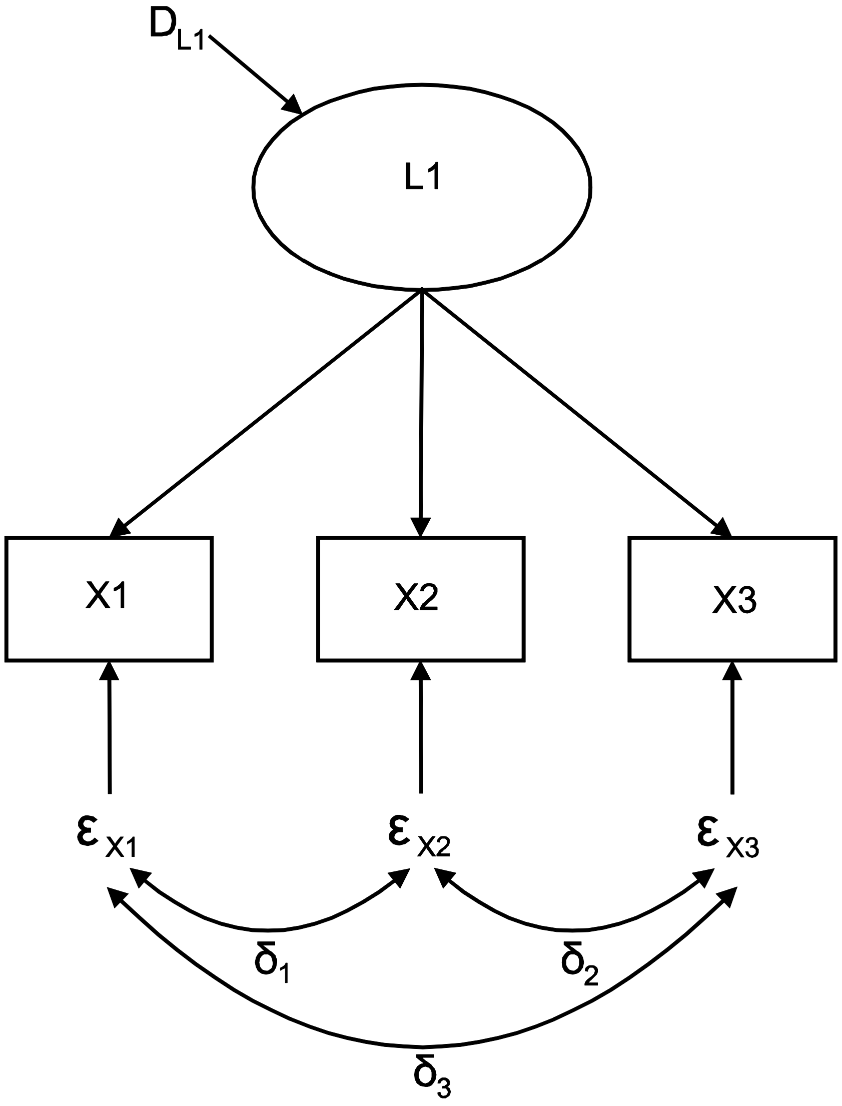
	phosphorus	**nitrogen**	**carbon**	**shannon**	**FD**	**phylo diversity**	**leaf N**	**decomp rates**	**biomass**
phosphorus	9.81	**0.57**	**0.55**	**0.08**	**0.12**	**0.03**	**0.00**	**0.03**	**0.01**
nitrogen	4.51	6.47	**0.49**	**0.12**	**0.12**	**0.05**	**0.10**	**0.04**	**0.11**
carbon	2.58	1.85	2.24	**0.20**	**0.20**	**0.13**	**0.13**	**0.14**	**0.09**
shannon	0.42	0.49	0.48	2.61	**0.66**	**0.62**	**0.35**	**0.42**	**0.42**
FD	1.14	0.97	0.92	3.27	9.50	**0.71**	**0.47**	**0.54**	**0.51**
phylodiversity	0.25	0.32	0.46	2.34	5.08	5.42	**0.49**	**0.55**	**0.52**
leaf N	-0.04	0.78	0.60	1.67	4.30	3.43	8.93	**0.61**	**0.59**
decomp rates	0.35	0.37	0.79	2.50	6.17	4.73	6.77	13.71	**0.66**
biomass	0.11	0.75	0.36	1.79	4.17	3.20	4.65	6.43	6.93

Table 2. Residual correlations (observed correlation matrix minus model-implied correlation matrix for the CFA model in Fig. 3). Values within the interval [-0.10, 0.10] indicate that the model adequately predicted all bivariate correlations.

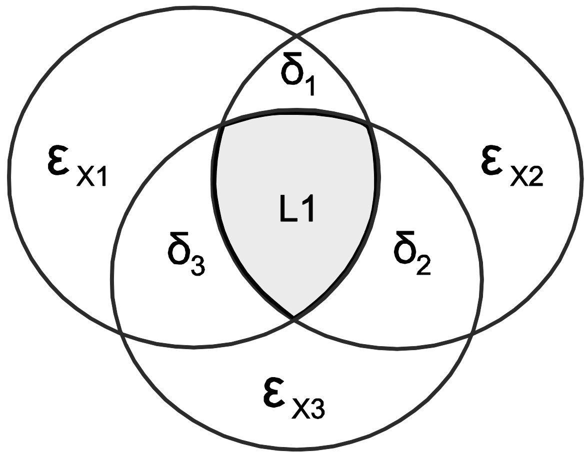
	phosphorus	**nitrogen**	**carbon**	**shannon**	**FD**	**phylo diversity**	**leaf N**	**decomp rates**	**biomass**
phosphorus	0.00								
nitrogen	0.01	0.00							
carbon	0.00	-0.01	0.00						
shannon	-0.02	0.02	0.10	0.00					
FD	-0.01	0.01	0.09	0.02	0.00				
phylodiversity	-0.09	-0.05	0.02	0.00	-0.01	0.00			
leaf N	-0.07	0.05	0.08	-0.06	-0.01	0.03	0.00		
decomp rates	-0.04	-0.02	0.08	-0.05	0.00	0.03	0.00	0.00	
biomass	-0.06	0.05	0.03	-0.02	0.00	0.02	0.00	0.00	0.00

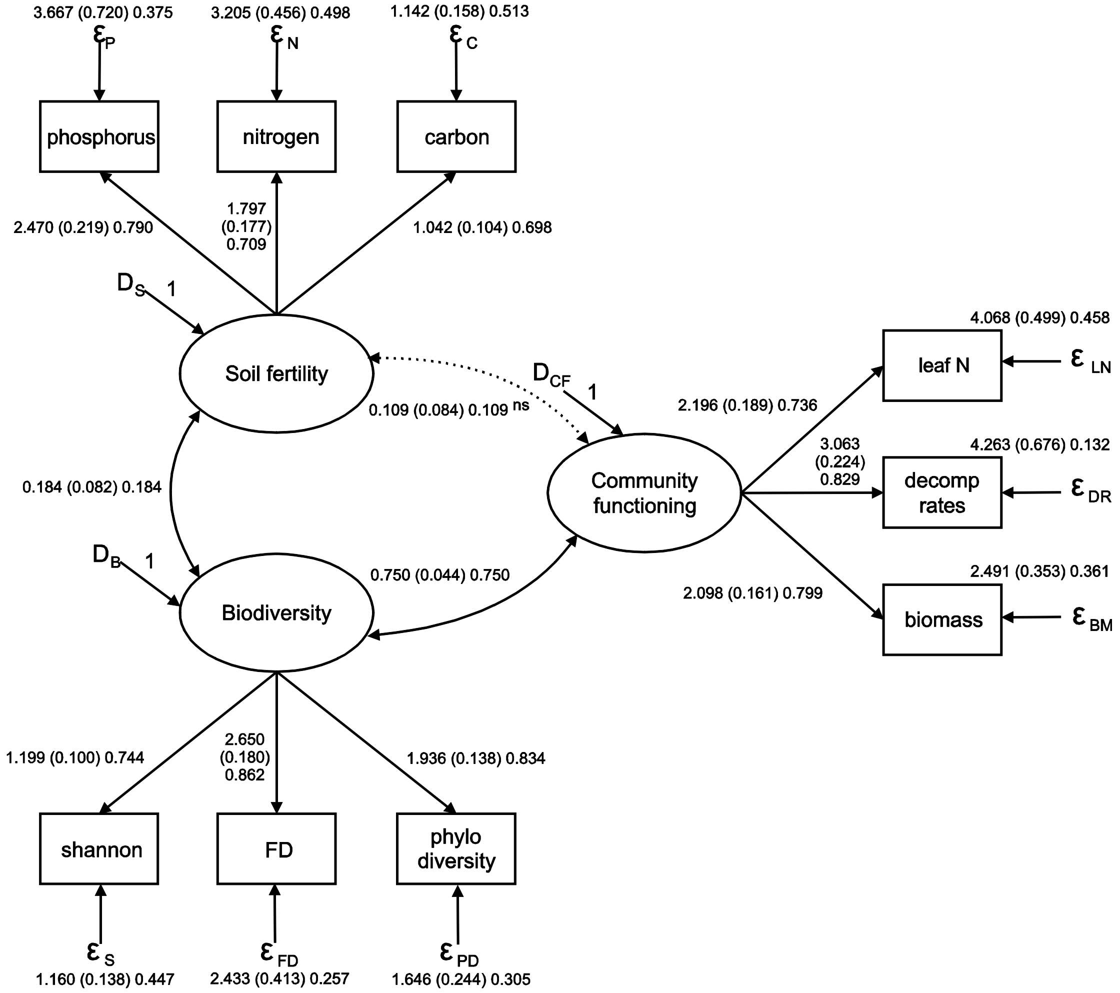

ε_P
ε_N
ε_C
phosphorus
nitrogen
carbon
1
Soil fertility
D_{CF}
leaf N
ε_{LN}
1
Community functioning
decomp rates
ε_{DR}
D_B
Biodiversity
biomass
ε_{BM}
1
shannon
FD
phylo diversity
ε_S
ε_{FD}
ε_{PD}

(a)


D_{L1}
L1
X1
X2
X3
ε_{X1}
ε_{X2}
ε_{X3}
δ_1
δ_2
δ_3

(b)


δ_1
ε_{X1}
ε_{X2}
L1
δ_3
δ_2
ε_{X3}


3.667 (0.720) 0.375
3.205 (0.456) 0.498
1.142 (0.158) 0.513
ε_P
ε_N
ε_C
phosphorus
nitrogen
carbon
2.470 (0.219) 0.790
1.797 (0.177) 0.709
1.042 (0.104) 0.698
D_S 1
Soil fertility
0.109 (0.084) 0.109 ns
D_{CF} 1
4.068 (0.499) 0.458
leaf N
ε_{LN}
2.196 (0.189) 0.736
3.063 (0.224) 0.829
4.263 (0.676) 0.132
decomp rates
ε_{DR}
Community functioning
0.184 (0.082) 0.184
D_B 1
0.750 (0.044) 0.750
2.098 (0.161) 0.799
2.491 (0.353) 0.361
biomass
ε_{BM}
Biodiversity
1.199 (0.100) 0.744
2.650 (0.180) 0.862
1.936 (0.138) 0.834
shannon
FD
phylo diversity
ε_S
ε_{FD}
ε_{PD}
1.160 (0.138) 0.447
2.433 (0.413) 0.257
1.646 (0.244) 0.305

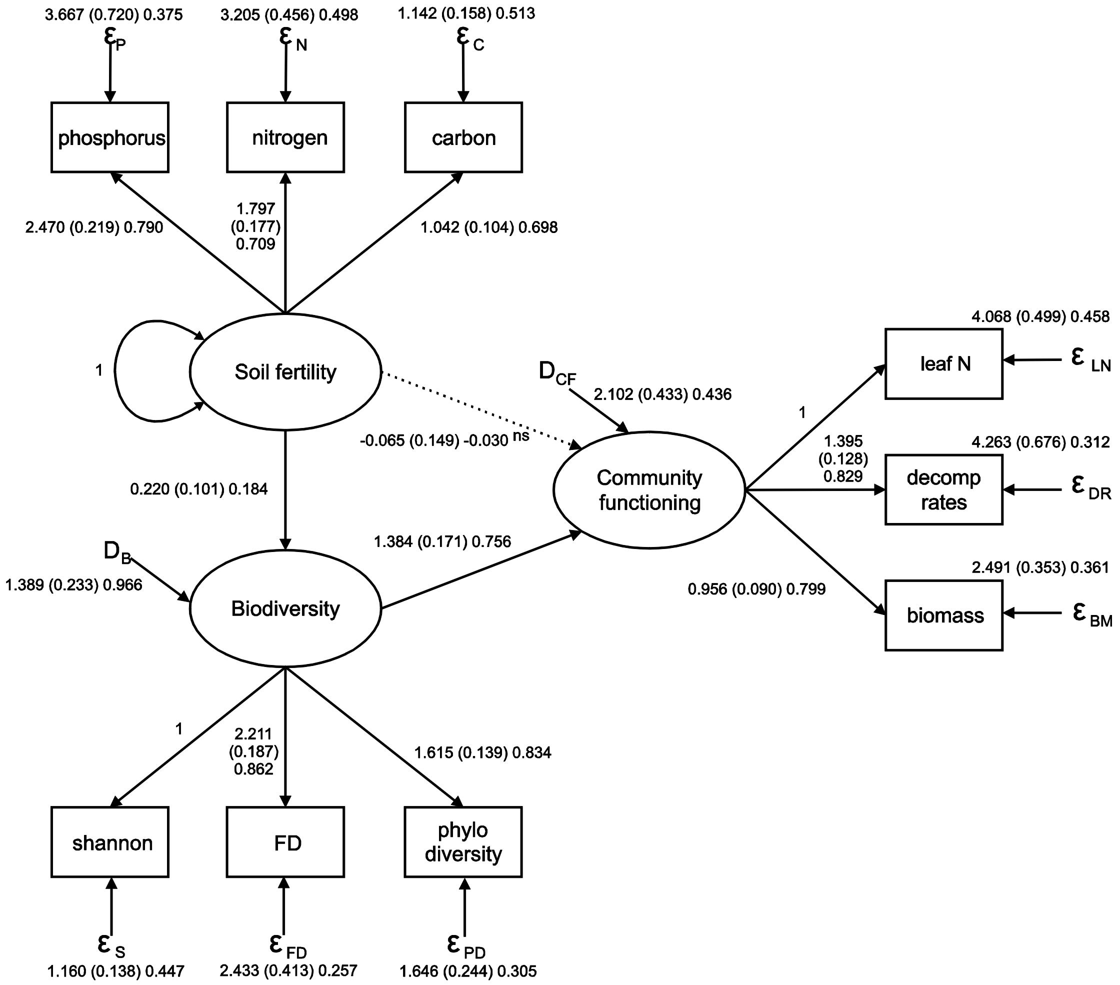

3.667 (0.720) 0.375
ε P
3.205 (0.456) 0.498
ε N
1.142 (0.158) 0.513
ε C
phosphorus
nitrogen
carbon
2.470 (0.219) 0.790
1.797 (0.177) 0.709
1.042 (0.104) 0.698
1
Soil fertility
D CF
2.102 (0.433) 0.436
-0.065 (0.149) -0.030 ns
0.220 (0.101) 0.184
Community functioning
4.068 (0.499) 0.458
leaf N
ε LN
1
1.395 (0.128) 0.829
4.263 (0.676) 0.312
decomp rates
ε DR
1.384 (0.171) 0.756
D B
1.389 (0.233) 0.966
Biodiversity
0.956 (0.090) 0.799
2.491 (0.353) 0.361
biomass
ε BM
1
2.211 (0.187) 0.862
1.615 (0.139) 0.834
shannon
FD
phylo diversity
ε S
1.160 (0.138) 0.447
ε FD
2.433 (0.413) 0.257
ε PD
1.646 (0.244) 0.305

Figure 1. Graphical representation of a structural equation model with hypothesised abiotic and biotic controls on community functioning. Parameters fixed at 1 set the scale of the latent variables. Variables in boxes are indicators of the latent variables in ellipses. phosphorus = soil phosphorus (mg kg^{-1}), nitrogen = soil nitrogen (mg kg^{-1}), carbon = soil carbon (mg kg^{-1}), shannon = Shannon's diversity index, FD = functional diversity index, phylodiversity = phylogenetic diversity index, leaf N = leaf nitrogen content (mg g^{-1}), decomp rates = decomposition rates of recently shed leaves, and biomass = live tissue biomass estimate (g m^{-2}).

Figure 2. Conceptual scheme of measurement error in structural equation modelling. (a) A latent variable (L1) with three indicators (X1, X2, and X3) and their respective error terms (ε_{X1}, ε_{X2}, and ε_{X3}). All errors are correlated (δ_1, δ_2, and δ_3); (b) circles represent the variances of each indicator. The intersection between all variances (grey area) defines the latent variable L1. All areas outside this intersection are the errors of the indicators, which consist of measurement error and unique variance due to the effects of omitted variables.

Figure 3. Confirmatory factor analysis to test the measurement model for the structural equation model proposed in Fig. 1. Variables in boxes are indicators (see Table 1 for variable names) of the latent variables in ellipses. Solid lines indicate significant paths. Dashed lines indicate not significant paths. $n = 210$; $\mathrm{ML}\chi^2 = 22.97$; $P = 0.52$; $df = 24$.

Figure 4. Estimates of the structural equation model proposed in Fig. 1. Variables in boxes are indicators (see Table 1 for variable names) of the latent variables in ellipses. Solid lines indicate significant paths. Dashed lines indicate not significant paths. $n = 210$; ML$\chi2 = 22.97$; $P = 0.52$; $df = 24$.

III - CAPÍTULO 2

Artigo submetido ao periódico Oecologia

Are fire, soil fertility and toxicity, water availability, plant functional diversity, and litter decomposition causally related in a Neotropical savana?

Gustavo Henrique Carvalho[†1], Marco Antônio Batalha[1], Igor Aurélio Silva[2], Marcus Vinicius Cianciaruso[3], and Owen Petchey[4]

[1] Department of Botany, Federal University of São Carlos, PO Box 676, 13565-905, São Carlos, Brazil

[2] Department of Plant Biology, State University of Campinas, PO Box 6109, 13083-970, Campinas, Brazil

[3] Department of Ecology, Federal University of Goiás, PO Box 131, 74001-970, Goiânia, Brazil

[4] Institute for Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057, Zurich, Switzerland

Running title: Environment, diversity, and functioning

[†] Correspondence author. E-mail: gustavo.bio@gmail.com

Summary

1. Understanding how biodiversity and ecosystem functioning respond to changes in the environment is fundamental to the maintenance of ecosystem function. In realistic scenarios, the biodiversity-ecosystem functioning pathway may account for only a small share of all factors determining ecosystem function
2. Here, we investigated the strength to which variations in environmental characteristics in a Neotropical savanna affected functional diversity and decomposition. We sought an integrative approach, testing a number of pairwise hypotheses about how the environment, biodiversity, and functioning were linked.
3. We used structural equation modelling to connect fire frequency, soil fertility, exchangeable aluminium, water availability, functional diversity of woody plants, and litter decomposition rates in a causal chain
4. We found significant effects of soil nutrients, water availability, and aluminium on functional diversity and litter decomposition. Fire did not have a significant direct effect on functional diversity or litter decomposition. However, fire was connected to both variables through soil fertility.
5. The mediated effects that emerged from pairwise interactions (e.g., biodiversity-ecosystem functioning effects) are encouraging for predicting the functional consequences of changes in environmental variables and biodiversity, but also caution against predictions based on only environmental or only biodiversity change.
6. *Synthesis*. This study showed that soil fertility, aluminium, and water availability had significant direct effects on functional diversity and litter decomposition rates. Soil fertility mediated the indirect effects of fire on functional diversity and litter decomposition rates. By simultaneously considering the pairwise effects between our variables, we uncovered a number of indirect effects, which often acted in opposition to direct ones.

Key-words: Brazilian cerrado; community functioning; intraspecific variability; structural equation modeling

Introduction

Understanding how organisms respond to changing environmental conditions could help to develop more effective management policies, especially regarding how disturbances and other abiotic factors should be dealt with to reduce the loss of biodiversity and function (Loreau *et al.* 2001). A great amount of research on the interactions between the components of the triad of abiotic factors, biological diversity, and ecosystem functioning exist, with much of this research focusing on the relationships between two components at a time (Tilman, Lehman & Thomson 1997; Hooper *et al*. 2005). For instance, several studies on the relationships between biological diversity and ecosystem functioning (hereafter BEF) did not quantitatively consider the influence of abiotic factors in BEF components (Tilman, Lehman & Thomson 1997; Hooper & Vitousek 1997; Hector *et al*. 1999). Ecosystem functioning likely responds to changes in environmental conditions not only via effects on biological diversity (Chapin *et al.* 1997; Cardinale, Nelson & Palmer 2000), but also via more direct pathways of effect (Grace *et al.* 2007; Jonsson & Wardle 2010). For example, these more direct pathways can be modulated by abiotic controls, such as resource availability, which influence ecosystem properties (Hooper *et al.* 2005). In realistic scenarios, the BEF pathway may account for only a small share of all factors determining ecosystem properties and function (Srivastava & Vellend 2005; Grace *et al.* 2007; Jonsson & Wardle 2010). Incorporating the effects of the environment in BEF research in natural communities goes one step further with the unveiling of the relative importance of all factors that contribute to ecosystem functioning, not only biodiversity (Grace *et al.* 2007; Jonsson & Wardle 2010). Such knowledge of the strength of the direct and indirect effects of the environment on the interaction between organisms and ecosystem properties and functioning may aid efforts to protect biodiversity and services in fluctuating conditions and has been tagged as one of the areas that need attention in BEF research (Srivastava & Vellend 2005).

In niche-based models of assembly, several non-exclusive mechanisms of species sorting have been identified. One of these mechanisms is environmental filtering, where abiotic forces, such as nutrient availability, water availability, and fire, filter species with certain trait values that give them the ability to overcome the limitations imposed by the environment (Keddy 1992). Limiting similarity, another well-studied mechanism, is determined by the interactions between species, such as competition (MacArthur & Levins 1967, Fridley 2001). For instance, species with similar trait values are likely to have overlapping positions in niche space and may, thus, compete for the same resources (Fridley 2001; Kraft, Valencia & Ackerly 2008; Cornwell & Ackerly 2009). These interactions prevent the coexistence of species with too similar trait attributes and often operate with the environment to shape the distribution of traits in a community. Thus, the stronger the pressure of the environment on individuals the lower the expected diversity of functional traits (Díaz & Cabido 2001).

Species can influence ecosystem functioning in a variety of ways (Loreau *et al*. 2001; Hooper *et al*. 2005). Increasing the number of species increases the likelihood that key species for ecosystem functioning are present in the community in a process known as the selection effect (Loreau 2000). For instance, in boreal forests, plant richness and composition drive litter decomposition rates and net primary productivity, respectively (Jonsson & Wardle 2010). However, since some species may have similar roles or contribute little to ecosystem functioning, often the number of species is not a good predictor of ecosystem functioning (Wardle, Bonner & Nicholson 1997; Hooper *et al*. 2005). A meta-analysis including savannas, meadows, prairies, and grasslands has shown that richness seldom explains the variation in plant biomass (Grace *et al.* 2007). Similar results were reported for manipulated polycultures (Petchey, Hector & Gaston 2004). One alternative approach is to look into the diversity of functional traits. It has been suggested that communities with higher diversity of functional traits have a tendency to operate more efficiently due to higher niche complementarity, which leads to more efficient partitioning of resources (Díaz & Cabido 2001; Petchey & Gaston 2002; Hooper *et al*. 2005; Petchey & Gaston 2006). Also, functional diversity is expected to account for multiple ecosystem functions as it is measured from multiple functional traits (Cadotte 2011).

Home to at least 1,000 woody species (Castro *et al.* 1999), the Brazilian cerrado is under constant threat as it is being destroyed to give place to commercial cultures, such as soybean, sugarcane, and African grasses for cattle ranching (Ratter, Ribeiro & Bridgewater 1997; Durigan, Siqueira & Franco 2007). Cerrado soils are usually acidic Oxisols, with very low concentrations of plant nutrients (Motta, Curi & Franzmeier 2002). The concentrations of aluminium, on the other hand, are often very high (Goodland & Pollard 1973). Early investigators proposed that the higher the concentration of plant nutrients and the lower the concentration aluminium, the higher the density of woody individuals (Goodland & Pollard 1973). According to this classical theory, the cerrado is a fertility gradient, with different physiognomies occurring under soils with different concentrations of plant nutrients. However, even if nutrients play an important role in the establishment of woody individuals, it is unlikely that they are the sole responsible for the striking heterogeneity of the cerrado vegetation. Some studies on soil-vegetation relationships in cerrado areas failed to corroborate the fertility gradient theory (Ruggiero *et al.* 2002), indicating that other soil characteristics, such as water availability (Ferreira *et al*. 2007; Assis *et al.* 2011), may be of greater importance. Evidence shows that nutrients and plant available water can, thus, act as environmental filters, favouring a limited range of trait values. For example, nutrient-poor soil favours species with sclerophyllous leaves and drought favours deep-rooted trees (Coutinho 1990). Abiotic factors are also expected to affect decomposition. For instance, soil moisture can alter the dynamic of mass loss of litter (Gartner & Cardon 2004). Climate has also been shown to explain the variation in litter decomposition rates and carbon mineralisation (Anderson 1991; Berg *et al.* 1993; Madritch & Cardinale 2007).

In the cerrado, as in other savannas, fire plays an important role in determining the composition of species (Bond & Keeley 2005; Silva & Batalha 2010). Cerrado plant species evolved with fire and are adapted to it (Coutinho 1990). For example, woody species have thick barks and subterranean meristems that insulate internal tissues from the high temperatures and allow resprouting after burnings (Gottsberger & Silberbauer-Gottsberger 1996). Fire has been shown to act as an environmental filter, promoting clustering of functional traits in woody species (Silva & Batalha 2010). Also, areas where fires are more frequent have lower total biomass than areas with less frequent fires (Cianciaruso, Silva & Batalha 2010). Burning events have important impact on nutrient cycling and availability in the cerrado (Coutinho 1990; Silva & Batalha 2008). After a fire, nutrients that were in the vegetation are either deposited in the soil as ashes or lost by volatilisation (Coutinho 1990). Consequently, areas where fire events are frequent have more organic matter and nitrogen when compared to areas protected from fire (Silva & Batalha 2008).

Here, we investigated the strength to which variations in abiotic factors in the cerrado affect the functional diversity of woody individuals and ecosystem functioning, directly and indirectly. We sought an integrative approach, analysing how disturbance, water availability, and soil nutrients influence the association between the diversity of functional traits and decomposition. We proposed a structural equation model (Fig. 1) that we believed was a plausible representation of the current knowledge on how our variables of interest are connected. We answered the following questions: (1) is the variation in the frequency of burnings related to the variation in the fertility of the area? (2) Is the variation in the frequency of burnings related the variation in the functional diversity of woody individuals? (3) Is the variation in the functional diversity woody individuals related to the variation in litter decomposition? (4) Is the variation in soil aluminium contents related to the variation in the functional diversity of woody individuals? (5) Is the variation in slope and altitude, which are related to the water availability, related to the variations in nutrient availability, functional diversity of woody individuals, and litter decomposition?

Materials and methods

Site

Covering about 133,000 ha of the Brazilian Central Plateau (17°49’-18°28’S and 52°’39’-53°10’W), Emas National Park (ENP) is one of the largest cerrado reserves. ENP is under tropical and humid climate, with a wet summer (September to May) and dry winter (June to August). Annual rainfall and mean temperature lie around 1,745 mm and 24.6°, respectively. The cerrado vegetation in the park goes from open (68.1% of its area) to closed physiognomies (25.1%). Other vegetation types, such as wet grasslands, riparian forests, and semideciduous forests cover the remaining 6.8% of the area (Ramos-Neto & Pivello 2000). ENP is one of the most important sites containing the fauna, flora, and key habitats that characterise the cerrado.

Since the prohibition of cattle farming inside the park in 1984, a fire exclusion policy was instituted. However, the accumulation of dry biomass led to the occurrence of uncontrolled wildfires every 3-4 years, burning on average 80% of ENP's total area (França, Ramos-Neto & Setzer 2007). In 1994, a catastrophic fire burned about 95% of the park. Since then, precautions to avoid similar burnings have been taken. Preventive firebreaks are burned annually in the dry season and a fire brigade stays in ENP during this period to prevent anthropogenic fires (França, Ramos-Neto & Setzer 2007). Nevertheless, even with these precautions, in August 2010 an anthropogenic fire burned 93% of the park's area.

Sampling

We randomly placed 100 plots, each one with 25m^2, in all tangible areas of ENP. In each plot, we collected five soil samples at 5 cm depth to determine the following soil variables (see Silva & Batalha 2008 for details on chemical and physical analyses): pH, organic matter, available phosphorus, total nitrogen, exchangeable potassium, exchangeable calcium, exchangeable magnesium, exchangeable aluminium, sum of bases, base saturation, aluminium saturation, cation exchange capacity, sand content, silt content, and clay content. Since most variables presented high bivariate correlation, we kept exchangeable aluminium, total nitrogen, available phosphorus, and cation exchange capacity for subsequent analyses, reducing model complexity. These variables are commonly regarded as important for plant growth, whilst exchangeable aluminium is toxic to plants. We also measured altitude and slope as rough surrogates of water availability: in ENP, the higher the altitude and the lower the slope, the lower the water availability in deep soil layers. Even though we could have used variables that represent soil texture, like sand content, as surrogates of water availability, these variables were highly correlated with altitude. We decided to keep altitude as it is a continuous measure and, thus, easier to include in the model. We determined mean interval between fire events for each plot from 1984 to 2009 based on satellite images and in-field observations. We inverted this variable by multiplying it by minus one to make interpretation easier. Thus, larger values indicate more frequent burnings.

In each plot, we identified and tagged all individuals with at least 3 cm of stem diameter at the soil level. From September 2009 to January 2010, for each sampled individual, we collected the values of 10 functional traits that are relatively easy and inexpensive to measure (Cornelissen *et al.* 2003; Pausas & Paula 2005). These traits are surrogates of important responses of plants to environmental conditions, such as nutrient availability, water availability, and fire, and of impacts of plants on ecosystem functioning (Cornelissen *et al*. 2003; Pausas & Paula 2005). The traits we measured were: (1) basal area, related to space occupation, resource uptake, total biomass, and reproductive capability; (2) height, associated with competitive vigour, fecundity, and growth after disturbance; (3) bark thickness, related to resistance to disturbance; (4) wood density, related to the capacity to store carbon and growth after disturbance; (5) leaf toughness, associated with resistance to abiotic and biotic mechanical damage; (6) leaf size, related to resistance to environmental stress; (7) specific leaf area, associated with growth and maximum photosynthetic rate, and (8-10) leaf nitrogen, phosphorus, and potassium concentration, related to maximum photosynthetic rates and nutrient stress. We used trait values for each individual to calculate an individual-based measure of functional diversity (iFD; Cianciaruso *et al.* 2009). Instead of taking into account mean trait values for each species, iFD considers actual trait values for all individuals sampled. Despite the increasing awareness that trait variation within populations is important for ecological processes (Crutsinger *et al.* 2006; Cadotte, Carscadden & Mirotchnick 2011), few functional diversity indices can accommodate intraspecific trait variation (Schleuter *et al.* 2010). iFD can, thus, detect plastic responses of individuals to environmental conditions instead of assuming that all individuals of a given species are equal regarding their traits. In the cerrado, some traits can vary more within than between species (Dantas *et al.* 2012). Thus, selecting a measure of functional diversity that could account for this source of variability would give us a better representation of niche width and overlap. We calculated iFD values for each plot using the “treedive” function from the “vegan” package (Oksanen *et al*. 2010) for the R environment (R Development Core Team 2011). We standardised the trait matrix to zero mean and unit variance before all calculations.

As a surrogate of ecosystem functioning, we used litter decomposition, which is one of the key functions for the maintenance of communities (Sulkava & Huhta 1998) and regulates the cycle of matter (Clark *et al*. 2001). In cerrado sites, decomposition rates vary according to vegetation structure, increasing from open to closed cerrado physiognomies (Cianciaruso *et al.* 2006; Valenti, Cianciaruso & Batalha 2008). Overall, litter decomposition is very slow in the cerrado. In each plot, at the corners, we placed four sets of five decomposition bags made with 1.0 mm^2 mesh. So, we had 2,000 litter bags in total. To prepare the bags, we collected recently shed leaves from woody individuals of all species occurring nearby the plots, cleaned them with a soft brush, and oven-dried them at 80°C for 24 hours. After that, we placed all leaves in a bag and thoroughly mixed them. We put 5 g of mixed dried leaves in each decomposition bag, placed them in the plots in the middle of the rainy season (January 2010), collected them 6 months later, cleaned them with a soft brush, and weighed them. When we designed the litter decomposition experiment, we planned to collect one set of bags from each plot after 1, 3, 6, 9, and 12 months to calculate decomposition rates. However, a fire burned almost all the vegetation in the park about 8 months into the experiment and we lost two sets of litter bags from each plot. Them, we considered the differences between masses after 6 months and initial masses as an indicator of community functioning for each plot. Large differences between final and initial masses indicate faster decomposition and higher functioning.

Analysis

We proposed a structural equation model (Shipley 2000) with our causal assumptions regarding the following variables (Fig. 1): (1) inverted mean interval between fire events; (2)

exchangeable aluminium (3-5) total nitrogen, available phosphorus, and cation exchange capacity as indicators of soil fertility; (6) altitude; (7) slope; (8) iFD; and (9) litter decomposition times. Before testing our model, we screened our data for outliers and deviations from normality in uni- and multivariate space using robust Mahalanobis distances (Filmozer, Garret & Reimann 2005). When data are normally distributed in multivariate space, distances are expected to follow a chi-square distribution (Filmozer, Garret & Reimann 2005). Then, we log-transformed all variables except total nitrogen and available phosphorus to minimise the effects of deviations from uni- and multivariate normality. After this, we screened data for variables with high multivariate collinearity by fitting several multiple regression models, each with one of our measured variables as the response and all other variables as predictors. Models with a $R^2 > 0.85$ indicated multivariate collinearity. Then, we assured that the scales of our variances were uniform by multiplying them to constants, which helped avoiding ill-scaled covariance structures and consequent problems during estimation (Kline 2010). With data thus screened for possible sources of problems during estimation, we generated a covariance matrix from our data (Table 1). We assured that the covariance matrix was positive-definite and all variances and covariances were within bounds (Kline 2010).

To estimate the free parameters and access the fit of the structural equation model we proposed, we used maximum likelihood estimation (MLE) and maximum likelihood chi-square ($ML\chi^2$), respectively. Ultimately, MLE and $ML\chi^2$ provide estimates of effects, variances and disturbances, and of how well a proposed model explain the covariance structure generated from observed data. The covariance matrix, the sample size (in our case, 100 plots), and one or a few previously specified model are all that is needed to apply structural equation modelling using MLE. We used the covariance matrix and proposed model as input in the package 'lavaan' (Rossel 2011) for the R environment (R Development Core Team 2011). When the differences between observed and model-implied covariance structures are small, the model has good fit (small $ML\chi^2$). Otherwise, the model has poor fit (large $ML\chi^2$), which means that the model does not properly explain how variables interact in the system. After estimation and fit, we searched for Heywood cases in the results (Kline 2010), that is, inadmissible and not interpretable solutions (for example, negative variance estimates). Finally, we repeated the analysis passing different starting values to the estimator to assure that it would converge to the same set of parameter estimates and model fit statistic every time. We fixed the variances of all exogenous variables to their observed values. Previous screening indicated linear relationships between the variables in our model.

When considering parameter estimates that were not significantly different from zero, we looked at the corresponding bivariate residual correlation to determine whether to remove these parameters or not. We expected that some parameters would not be significant as the model we presented was complex and the hypotheses behind some of the parameters are not very well established yet.

Results

We sampled 531 woody individuals belonging to 55 species. Fabaceae and Myrtaceae were the richest families, with 10 and 9 species, respectively. The ranges of the traits were as follows: 1) basal area: 0.004-0.147 m^2; 2) height: 0.12-5.65 m; 3) bark thickness: 0.11-29.6 mm; 4) wood density: 0.006-0.943 mg mm^{-3}; 5) leaf toughness: 0.09-8.84 N; 6) leaf size: 542-13010 mm^2; 7) specific leaf area: 0.004-20.507 mm^2 mg^{-1}; 8) leaf nitrogen content: 7.32-44.85 mg g^{-1}; leaf phosphorus content: 0.21-2.60 mg g^{-1}; and 9) leaf potassium content: 1.53-24.74 mg g^{-1}. Please refer to Batalha *et al.* (2011) for trait means for each sampled species. Environmental variables ranged within the following values: 1) altitude: 709-884 m; 2) slope: 0.3-8.7 º; 3) available phosphorus: 1-13 mg kg^{-1}; 4) total nitrogen: 1019-2746 mg kg^{-1}; 5) cation exchange capacity: 33-387.8 mmol kg^{-1}; 6) exchangeable aluminium 5-38 mmol kg^{-1}; and 7) mean interval between fires: 1.18-8 years.

The initial model (Fig. 1) had good fit with data ($ML\chi^2 = 17.244$; $P = 0.243$; $df = 14$; Fig. 2). Even though this model had acceptable fit, not all parameters were significant (Fig. 2). For instance, the effects of fire on functional diversity and decomposition were not significantly different from zero. Likewise, altitude did not seem to have an effect on any of the variables of our proposed model. This model resulted in some high residuals (Table 2), so we went ahead and removed all the non-significant parameters from the model (Fig. 3). Removing these parameters yielded a model with much better fit (MLχ^2 = 10.289; P = 0.741; df = 14) and lower residuals (Table 3). In this final model, altitude was not significantly connected to any other variables, so we removed it (Fig. 3).

Fire had a positive effect on the latent variable fertility (fire → fertility = 0.404; Fig. 3). This result suggested that frequent burnings were related to higher nutrient availability. Fire did not directly affect iFD and decomposition; however, both variables were indirectly influenced by fire through fertility and iFD. The signals of these indirect effects were different though: larger intervals between fire events were related to less fertile soil, which increased iFD (fire → fertility → iFD = -0.172); thus, more fires resulted in lower iFD values. Furthermore, frequent burnings resulted in more fertile soils, which, in turn, resulted in quicker decomposition; thus, more fires resulted in faster decomposition (fire → fertility → decomposition = 0.225). Contrastingly, we found a negative effect of fire on decomposition through both fertility and iFD (fire → fertility → iFD → decomposition = -0.048).

The coefficients describing the direct relationships of aluminium with fertility (Al ←→ fertility = 0.706; Fig. 3), iFD (Al → iFD = 0.454; Fig. 3), and decomposition (Al → decomposition = -0.672; Fig. 3) were significant. The path coefficient connecting aluminium and fertility was positive (Al → fertility = 0.704; Fig. 3). Similarly, our model suggested that higher aluminium content increased iFD values when considering the direct path between them. There was also an indirect pathway connecting aluminium and decomposition through iFD (Al → iFD → decomposition = 0.129; Table 4).

The effects of altitude on iFD and decomposition were non-significant, so we removed altitude from the final model. Slope had a significant effect on iFD (slope → iFD = 0.212) and an indirect effect on decomposition through iFD (slope → iFD → decomposition = 0.065; Table 4). Plots with higher slopes had higher iFD values and faster decomposition.

Finally, iFD had a positive effect on decomposition (iFD → decomposition = 0.768). Plots with higher functional diversity presented quicker decomposition.

Discussion

Our model integrating the variables we assumed to be the major drivers of assembly and functioning in ENP and other cerrado areas offered support for most of our pairwise causal assumptions. Furthermore, combining these hypotheses in a structural equation model enabled us to reveal indirect effects between the variables representing the environment, biodiversity, and ecosystem functioning. Frequently, indirect effects acted in opposition to direct effects. Our results allowed us to quantify the strength at which alterations in one variable cause alterations in all other variables in the system, providing us with a more realistic mathematical translation of current theories of how this ecosystem works. The approach we chose also permitted us to include not only variables that are generally regarded as major drivers of assembly and functioning, but also those that are specific to the cerrado. According to our model, fire was positively related to soil fertility in ENP. Fire was also related to iFD and decomposition, although indirectly. Aluminium seemed to have an influence on several aspects of the ecosystem, including complementarity of functional traits and decomposition. The slope of the plots, which is rough estimate of water availability to plants, had a significant effect on iFD. Functional diversity, in turn, was directly related to the speed of decay in litter. Our whole hypothesized causal structure had good fit with data as indicated by the MLχ^2 statistic (Fig. 3). We were able to demonstrate that by simultaneously considering the pairwise hypotheses of how the variables in the system were directly connected, we were able to reveal the strength and sign of indirect relationships (Fig. 3).

The effect of fire on fertility met with our predictions. Sites that burned more frequently had more plant nutrients in the soil (fire → fertility in Fig. 3). Fire transfers nutrients from the burned vegetation to the upper soil layer as ash deposition (Coutinho 1990). Moreover, during burnings, woody individuals shed their leaves, including young ones, which are then deposited around trees as litter (Rodrígues *et al.* 2009). Young leaves have higher nutrient concentrations than mature leaves, so premature leaf shedding and decomposition are likely to increase nutrient availability. Several nutrients have fast turnover times in the cerrado (Pivello & Coutinho 1992) and, so, even though some chemical elements are lost by volatilisation or as particles in smoke, deposition usually compensates for this loss in 1-3 years (Coutinho 1979). Pivello & Coutinho (1992) estimated that 3-year intervals between burnings would be optimal to avoid impoverishment of the soils and maintain nutrient cycling. Slight increases in nutrients availability up to one year after moderate fires were also found in African savannas (Jensen, Michelsen & Gashaw 2001) and in the cerrado (Silva & Batalha 2008). Despite the fact that we found higher nutrient availability in the soil with increasing fire frequencies, the relationships between fire, nutrients, and the vegetation are still unclear, since contrasting results have been found, even in the cerrado (Kauffman, Cummings & Ward 1994; Moreira 2000; Pivello *et al.* 2010). Aluminium and plant nutrients are likely introduced in the soil by some of the same processes (e.g. weathering), thus the observed correlation between fertility and aluminium (fertility ← → Al in Fig. 3).

Contrarily to our expectations, the fire → iFD path was not significant. Low and moderate fire frequencies might have not been strong enough environmental filters to leave an imprint in functional diversity (Cianciaruso *et al.* 2012). Thus, the apparent absence of functional structuring in sites with few burnings might have decreased the strength of the path between fire and iFD, rendering it statistically non-significant. Also, fire might have caused a turnover of species without causing loss of functional diversity or ecosystem function in a process known as the insurance hypothesis (Yachi & Loreau 1999; Loreau 2001). Our model might have supported the path connecting fire and iFD if we had used a different set of traits (Cianciaruso *et al.* 2012). However, even in the absence of a direct effect, there was a small indirect effect of fire on iFD through soil fertility (Table 4). More frequent fires promoted faster nutrient cycling which, in turn, were associated with less trait complementarity and lower iFD. Absence of trait structuring in sites with different occurrence of burnings in ENP also suggests that fire might be filtering species at the regional level, so the species we sampled were already selected and local filtering by fire was not strong enough determine the distribution of trait-states (Dantas *et al.* 2012).

Variations in soil chemical elements triggered a response in iFD (fertility → iFD and Al → iFD paths in Fig. 3) and decomposition (fertility → decomposition and Al → decomposition in Fig. 3). We observed a negative relationship between fertility and iFD (fertility → iFD path in Fig. 3) and a positive one between aluminium and iFD (Al → iFD path in Fig. 3). These findings strongly support the hypotheses represented in the structural equation model, and highlight the importance of nutrients and aluminium for plant functional diversity and ecosystem decomposition rates. In the cerrado, soil characteristics have been regarded as promoters of trait clustering through environmental filtering (Batalha *et al.* 2011). Areas with low nutrient availability may promote competitive exclusion, limiting the similarity of individuals (Stubbs & Wilson 2004). This process decreases the overlap in niche occupation, which might lead to higher iFD. Although low nutrient availability can promote trait clustering, competition for limited resources contributes to the selection of specialised nutrient and nutrient uptake requirements, imposing a force in the opposite direction. Similarly, plants have several ways of circumventing the toxic effects of aluminium (Kochian 1995). Thus, high concentrations of exchangeable aluminium might increase trait diversity and iFD. Also, high concentrations of exchangeable aluminium are related to acidic soils, which are reported to have low bacterial diversity when compared to neutral soils (Fierer & Jackson 2006). Low bacterial biota might have led to slower decomposition. Moreover, ants can change soil properties near their nests (Wagner, Brown & Gordon 1997; Frouz, Michal & Kalčík 2003). Areas with ant nests nearby have more soil nutrients (Wagner, Brown & Gordon 1997; Frouz, Michal & Kalčík 2003) and higher pH (Frouz, Michal & Kalčík 2003), due to ant activity, so the positive link between fertility and decomposition and the negative link between aluminium and decomposition in our model might be related to the presence of ants. Our data did not include information on ant activity in plots, so we could not explicitly consider and this activity in the model. Including other trophic levels into future models will certainly help clarify some of the associations in BEF research (Hooper *et al*. 2005, Srivastava & Vellend 2005)

In ENP, hilly terrain seemed to increase functional diversity (slope → iFD path in Fig. 3), which supported the theory that water availability is one of the most important determinants of vegetation structure in the cerrado (Ferreira *et al*. 2007; Assis *et al.* 2011). In ENP's flatland, the water table is deep enough to prevent even shallow-rooted woody individuals to reach it. Hilly areas, however, have shallow water table that favours the establishment of those woody individuals incapable of reaching deep soil layers (Cole 1986; Franco 2002). Sites on flatlands, on the other hand, have less available water, limiting the occurrence of individuals to those with deep root systems. Thus, in sites where there is less water available, environmental filtering might be a much stronger force than competition on the distribution of trait attributes, leading to lower functional diversity. In spite of altitude also being related to the depth of the water table in ENP, it was not one of the variables in our model with best fit. In hilly areas, the water table might be shallow enough to allow some woody individuals to overcome hydraulic limitations.

Changes in iFD led to variations in decomposition (iFD → decomposition path in Fig. 3). Functional traits can have great influence on ecosystem fluxes, pools, and function (Hooper *et al*. 2005). Higher biodiversity, especially when considering the functional component, can increase complementarity in patterns of resource use (Tilman, Lehman & Thomson 1997; Petchey & Gaston 2002), which results in more efficient functioning. Moreover, soil biota may be influenced by plant functional diversity and, as a consequence, influence litter decomposition (Collins 1981; Chapman, Whittacker & Heal 1988). Fire, on the other hand, did not have a direct effect on decomposition. In the cerrado, arthropods are fundamental agents of litter breakdown. The communities of several leaf-litter arthropods slightly decrease in number of individuals following a fire, but they are able to quickly recover and resume litter breakdown (Vasconcelos *et al*. 2009). Thus, the litter-dwelling activities of these arthropods might not have been significantly affected by burnings, which supressed any possible direct effects of fire on decomposition. Even though fire was not directly related to functioning, there were two indirect pathways connecting these variables. There was a path linking fire to decomposition through soil fertility and another one through both soil fertility and iFD (please refer to Table 4 for net effects). Interestingly, the indirect effects of fire on decomposition were antagonistic: the relationship was positive when mediated only by soil fertility and negative when mediated by both soil fertility and iFD. We found the same pattern of opposite relationships in links between fertility and decomposition. The direct path between these variables was positive, but the indirect one mediated by iFD was negative. Functional diversity seemed to act as a buffer, controlling the strength of the links between environment and function. Evidence from other studies suggests similar patterns of ecosystem function buffering by biodiversity (Balvanera *et al.* 2006).

Several aspects of the environment might alter the strength of the relationship between iFD and decomposition, both directly and indirectly and through different pathways. We were able to identify and quantify some of the multiple paths that causally connect the environment, biodiversity, and ecosystem functioning. The causal links that are thought to be the most important for determining the BEF relationship can be complex. For instance, we could not identify a direct influence of fire on iFD. Fire did have, however, a strong link to soil fertility, which, in turn, helped shaping the distribution of iFD values in our plots. Similarly, we identified at least two ways in which fire affects decomposition in ENP, even though we did not find a significant direct connection. Having a better understanding of how abiotic factors interact with each other and with biodiversity and function can uncover critical paths for the conservation of biological diversity and ecosystem function (Grace *et al.* 2007; Jonsson & Wardle 2010; Srivastava & Vellend 2005). If we had analysed the effects of fire on biodiversity directly, we would probably have found a non-significant relationship. However, considering soil fertility as a mediator of the fire → iFD path, we were able to uncover a strong link between these variables. Interestingly, we also demonstrated that the sign of the associations between two variables might change depending on which variables are functioning as mediators (e.g. fire and decomposition through fertility and through both fertility and iFD). Moreover, the net effects of the multiple ways in which the effects of disturbances can propagate in a causal network can be large even without the presence of a direct connection between disturbance and biodiversity (Table 4) and the pathways connecting two variables can have opposite signs (e.g., aluminium and fire on decomposition).

Soil chemical elements and water availability were the most important direct causes of change to the iFD-decomposition relationship. We detected direct effects of fertility (fertility → iFD and fertility → decomposition in Fig. 3), aluminium (Al → iFD and Al → decomposition in Fig. 3), and slope (slope → iFD in Fig. 3) on iFD and decomposition. Our results also suggest that fire has an important role on the BEF relationship through soil nutrients. The absence of a significant path between fire and iFD indicate that fire might be selecting traits at the regional level. Interestingly, iFD seemed to regulate the influence of the environment on decomposition by controlling the strength of the direct links between environment and decomposition. It is important to note that this study was cross-sectional, so we did not try to model the feedback mechanisms of biodiversity and functioning on abiotic factors. Long-term studies with data on the fluctuations of biotic and abiotic factors in different seasons might shed light on these feedbacks. Also, modelling the multidimensionality of the biodiversity and functioning components by including more diversity indices and other ecosystem properties will yield even more adequate models. Structural equation models with latent variables are appropriate tools for modelling these multidimensional concepts (for instance, fertility in Fig. 1). Here we considered one aspect of biodiversity and functioning. Furthermore, different ecosystem fluxes and properties might be affected by different sets of trait attributes, possibly changing the strength of the BEF pathway as models become more realistic. Ecosystem functioning research must address all these uncertainties to propose BEF models that are more relevant to the conservation of biodiversity and services (Hooper *et al*. 2005; Srivastava & Vellend 2005).

In conclusion, we presented an initial framework of how the abiotic factors interact with the components of the BEF relationship in a neotropical savanna. We showed that several aspects of the environment had great influence on functional diversity and ecosystem properties through direct and indirect pathways. Different fire frequencies resulted in propagation of effects on iFD and decomposition through soil fertility. Fertility and aluminium were also major drivers of assembly and functioning in ENP. Thus, we were able to unveil the strength of some of the links connecting abiotic factors and how they interact with biodiversity and ecosystem functioning. It is, therefore, of major importance to conservation efforts that management policies take into account the several ways in which variables interact to regulate the functioning of ecosystems.

Acknowledgments

We are grateful to Fapesp, CNPq, and CAPES, for financial support; to Ibama and Emas National Park staff, for research permission; to Helena França, for providing the fire history; and to V.L. Dantas, N.A. Escobar, C.S. Gonçalves, P.P. Loiola, N.B. Rosatti, and D.M. Silva, for help in the field. O.L.P. was partly funded by the Royal Society and the University of Zurich.

References

Anderson, J.M. (1991) The effects of climate change on decomposition processes in grassland and coniferous forests. *Ecological Applications*, **1**, 326-347.

Assis, A.C.C., Coelho, R.M., Pinheiro, E.S. & Durigan, G. (2011). Water availability determines physiognomic gradient in an area of low-fertility soils under cerrado vegetation. *Plant Ecology*, **212**, 1135-1147.

Balvanera, P., Pfisterer, A.B., Buchmann, N., He, J.S., Nakashizuka, T., Raffaelli, D. & Schimid, B. (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters*, **9**, 1146-1156.

Batalha, M.A., Silva, I.A., Cianciaruso, M.V., França, H. & Carvalho, G.H. (2011) Phylogeny, traits, environment, and space in cerrado plant communities at Emas National Park (Brazil). *Flora*, **206**, 949-956.

Berg, B., Berg, M.P., Bottner, P., Box, E., Breymeyer, A., Calvo de Anta, R., Couteaux, M., Escudero, A., Gallardo, A., Kratz, W., Madeira, M., Mälkönen, E., McClaugherty, C., Meentemeyer, V., Muñoz, F., Piussi, P., Remacle, J. & Virzo de Santo, A. (1993) Litter mass-loss rates in pine forests of Europe and Eastern United States: some relationships with climate and litter quality. *Biogeochemestry*, **20**, 127-159.

Bond, W.J. & Keeley, J.E. (2005) Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems. *Trends in Ecology and Evolution,* **20**, 387-394.

Cadotte, M.W. (2011) The new diversity: management gains through insights into the functional diversity of communities. *Journal of Applied Ecology*, **48**, 1067-1069.

Cadotte, M.W., Carscadden, K. & Mirotchnick, N. (2011) Beyond species: functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology*, **48**, 1079-1087.

Cardinale, B.J., Nelson, K. & Palmer, M.A. (2000) Linking species diversity to the functioning of ecosystems: on the importance of environmental context. *Oikos*, **91**, 175-183.

Castro, A.A.J.F, Martins, F.R., Tamashiro, J.Y. & Shepherd, G.J. (1999) How rich is the flora of the Brazilian cerrados? *Annals of the Missouri Botanical Garden,* **86**, 192-224.

Chapin, F.S., III., Walker, B.W., Hobbs, R.J., Hooper, D.U., Lawton, J.H., Sala, O.E. & Tilman, D. (1997) Biotic control over the functioning of ecosystems. *Science*, **277**, 500-504.

Chapman, K., Whittacker, J.B. & Heal, O.W. (1988) Metabolical and faunal activity in litters of tree mixtures compared with pure stands. *Agriculture, Ecosystems, and Environment*, **24**, 33-40.

Cianciaruso, M.V., Batalha, M.A., Gaston, K.J. & Petchey, O.L. (2009) Including intraspecific variability in functional diversity. *Ecology*, **90**, 81-89.

Cianciaruso, M.V., Pires, J.S.R., Delitti, W.C.B. & Silva, E.F.P. (2006) Produção de serapilheira e decomposição do material foliar em um cerradão da Estação Ecológica de Jataí, Luiz Antônio, SP. *Acta Botanica Brasilica*, **20**, 49-59.

Cianciaruso, M.V., Silva, I.A. & Batalha, M.A. (2010) Above-ground biomass of functional groups in the herbaceous layer of savannas under different fire frequencies. *Australian Journal of Botany,* **58**, 169-174.

Cianciaruso, M.V., Silva, I.A., Batalha, M.A., Gaston, K.J. & Petchey, O.L. (2012) The role of fire on phylogenetic and functional structure of woody savannas: moving from species to individuals. *Perspectives in Plant Ecology, Evolution and Systematics*, **14**, 205-216.

Clark, D.A., Brown, S., Kichlighter, W.K., Charmbers, J.Q., Thomlinson. J.R., Ni, J. & Holland, E.A. (2001) Net primary production in tropical forests: an evaluation and synthesis of existing field data. *Ecological Applications*, **11**, 371–384.

Cole, M.M. (1986) *The savannas, biogeography and geobotany*. Academic Press, London.

Collins, N.M. (1981) The role of termites in the decomposition of wood and leaf litter in the southern Guinea savanna of Nigeria. *Oecologia,* **51**, 389-399.

Cornelissen, J.H.C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D.E., Reich, P.B., ter Steege, H., Morgan, H.D., var der Heijden, M.G.A., Pausas, J.G. & Poorter, H. (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany,* **51**, 335-380.

Cornwell, W.K. & Ackerly, D.D. (2009). Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs*, **79**, 109-126.

Coutinho, L.M. (1979). Aspectos ecológicos do fogo no cerrado: III. A precipitação atmosférica de nutrientes minerais. *Revista Brasileira de Botânica*, **1**, 93-97.

Coutinho, L.M. (1990). Fire in the ecology of the Brazilian cerrado. *Fire in the tropical biota* (eds J.G. Goldammer), pp 81-103. Springer, Berlin.

Crutsinger, G.M., Collins, M.D., Fordyce, J.A., Gompert, Z., Nice, C.C. & Sanders, N.J. (2006) Plant genotypic diversity predicts community structure and governs ecosystem process. *Science*, **313**, 966-968.

Dantas, V.L., Pausas, J.G., Batalha, M.A., Loiola, P.P. & Cianciaruso, M.V. (in press) The role of fire in structuring trait variability in neotropical savannas. *Oecologia*, **171**, 487-494.

Díaz, S. & Cabido, M. (2001) Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, **16**, 646-655.

Durigan, G., Siqueira, M.F. & Franco, G.A.D.C. (2007) Threats to the cerrado remnants of the state of São Paulo, Brazil. *Scientia Agricola*, **64**, 355-363.

Ferreira, J.N., Bustamante, M., Garcia-Montiel, D.C., Caylor, K.K. & Davidson, E.A. (2007) Spatial variation in vegetation structure coupled to plant available water determined by two-dimensional soil resistivity profiling in a Brazilian savanna. *Oecologia*, **153**, 417-430.

Fierer, N. & Jackson, R.B. (2006) The diversity and biogeography of soil microbial communities. *Proceedings of the National Academy of Sciences of The United States of America*, **103**, 626-631.

Filmozer, P., Garret, R.G. & Reimann, C. (2005) Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences*, **31**, 579-587.

França, H., Ramos-Neto, M.B. & Setzer, A. (2007) *O fogo no Parque Nacional das Emas*. Ministério do Meio Ambiente, Brasília.

Franco, A.C. (2002) Ecophysiology of woody plants. *The cerrados of Brazil: ecology and natural history of a neotropical savanna* (eds P.S. Oliveira & R.J. Marquis). pp 178-197. Columbia University, New York.

Fridley, J.D. (2001) The influence of species diversity on ecosystem productivity: how, where, and why? *Oikos*, **93**, 514-523.

Frouz, J., Michal, H. & Kalčík, J. (2003) The effect of *Lasius niger* (Hymenoptera, Formicidae) ant nest on selected soil chemical properties. *Pedobiologia*, **47**, 205-212.

Gartner, T.B. & Cardon, Z.G. (2004) Decomposition dynamics in mixed-species leaf litter. *Oikos*, **104**, 230-246.

Goodland, R. & Pollard, R. (1973) The Brazilian cerrado vegetation: a fertility gradient. *Journal of Ecology*, **61**, 219-224.

Gottsberger, G. & Silberbauer-Gottsberger, I. (2006). *Life in the cerrado: a South American tropical seasonal vegetation. Vol. 1. Origin, structure, dynamics and plant use*. Reta, Ulm.

Grace, J.B., Anderson, T.M., Smith, M.D., Seabloom, E., Andelman, S.J., Meche, G., Weiher, E., Allain, L.K., Jutila, H., Sankaran, M., Knops, J., Ritchie, M. & Willig, M.R. (2007) Does species diversity limit productivity in natural grassland communities? *Ecology Letters,* **10**, 680-689.

Hector, A., Schmid, B., Beierkuhnlein, C., Caldeira, M.C., Diemer, M., Dimitrakopoulos, P.G., Finn, J.A., Freitas, H., Giller, P.S., Good, J., Harris, R., Högberg, P., Huss-Danell, K., Joshi, J., Jumpponen, A., Körner, C., Leadley, P.W., Loreau, M., Minns, A., Mulder, C.P.H., O'Donovan, G., Otway, S.J., Pereira, J.S., Prinz, A., Read, D.J., Scherer-Lorenzen, M., Schulze, E.D., Siamantziouras, A.S.D., Spehn, E.M., Terry, A.C., Troumbis, A.Y., Woodward, F.I., Yachi, S. & Lawton, J.H. (1999) Plant diversity and productivity experiments in European grasslands. *Science*, **286**, 1123-1127.

Hooper, D.U., Chapin, F.S.III., Ewel, A.H., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setälä, H., Symstad, A.J., Vandermeer, J. & Wardle, D.A. (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, **75**, 3-35.

Hooper, D.U. & Vitousek, P.M. (1997) The effects of plant composition and diversity on ecosystem processes. *Science*, **277**, 1302-1305.

Jensen, M., Michelsen, A. & Gashaw, M. (2001) Responses in plant, soil inorganic and microbial nutrient pools to experimental fire, ash and biomass addition in a woodland savanna. *Oecologia*, **128**, 85-93.

Jonsson, M. & Wardle, D. A. (2010) Structural equation modelling reveals plant-community drivers of carbon storage in boreal forest ecosystems. *Biology Letters,* **6**, 116-119.

Kauffman, J.B., Cummings, D.L. & Ward, D.E. (1994) Relationships of fire, biomass and nutrient dynamics along a vegetation gradient in the Brazilian cerrado. *Journal of Ecology*, **82**, 519-531.

Keddy, P.A. (1992) Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science*, **3**, 157-164.

Kline, R.B. (2010) *Principles and practice of structural equation modeling*. Guilford, New York.

Kochian, L.V. (1995) Cellular mechanisms of aluminum toxicity and resistance in plants. *Annual Review of Plant Physiology and Molecular Biology,* **46**, 237-260.

Kraft, N.J.B., Valencia, R. & Ackerly, D.D. (2008) Functional traits and niche-based tree community assembly in an Amazonian forest. *Science*, **322**, 580-582.

Loreau, M. (2000) Biodiversity and ecosystem functioning: recent theoretical advances. *Oikos*, **91**, 3-17.

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J.P., Hector, A., Hooper, D.U., Huston, M.A., Raffaelli, D., Schmid, B., Tilman, D. & Wardle, D.A. (2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, **294**, 804-808.

Macarthur, R. & Levins, R. (1967) The limiting similarity, convergence, and divergence of coexisting species. *The American Naturalist*, **101**, 377-385.

Madritch, M.D. & Cardinale, B.J. (2007) Impacts of tree species diversity on litter decomposition in northern temperate forests of Wisconsin, USA: a multi-site experiment along a latitudinal gradient. *Plant and Soil*, **292**, 147-159.

Moreira, A.G. (2000) Effects of fire protection on savanna structure in Central Brazil. *Journal of Biogeography*, **27**, 1021-1029.

Motta, P.E.F., Curi, N. & Franzmeier, D.P. (2002). Relation of soils and geomorphic surfaces in the Brazilian cerrado. *The cerrados of Brazil: ecology and natural history of a Neotropical savanna* (eds P.S Oliveira & R.J. Marquis). pp. 13-32. Columbia University, New York.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H. & Wagner, H. (2010) Vegan: community ecology package. URL: http://CRAN.R-project.org/package=vegan.

Pausas, J.G. & Paula, S. (2005) Plant functional traits database for Euro–Mediterranean ecosystems. Eufirelab Deliverable D–04–06. URL: http//www.eufirelab.org.

Petchey, O.L. & Gaston, K.J. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, **5**, 402-411.

Petchey, O.L., Hector, A. & Gaston, K.J. (2004) How do different measures of functional diversity perform? *Ecology*, **85**, 847-857.

Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters*, **9**, 741-758.

Pivello, V.R. & Coutinho, L.M. (1992) Transfer of macro-nutrients to the atmosphere during experimental burnings in an open cerrado (Brazilian savanna). *Journal of Tropical Ecology*, **8**, 487-497.

Pivello, V.R., Oliveras, I., Miranda, H.S., Haridasan, M., Sato, M.N. & Meirelles, S.T. (2010) Effects of fire on soil nutrient availability in an open savanna in Central Brazil. *Plant and Soil*, **337**, 111-123.

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. URL: http://www.r-project.org.

Ramos-Neto, M.B. & Pivello, V.R. (2000) Lightning fires in a Brazilian savanna national park: rethinking management strategies. *Environmental Management*, **26**, 675–684.

Ratter, J.A., Ribeiro, J.F. & Bridgewater, S. (1997) The Brazilian cerrado vegetation and threats to its biodiversity. *Annals of Botany*, **80**, 223-230.

Rodrígues, A., Duran, J., Fernández-Palacios, J.M. & Gallardo, A. (2009) Short-term wildfire effects on the spatial pattern and scale of labile organic-N and inorganic-N and P pools. *Forest Ecology and Management*, **257**, 739-746

Rosseel, Y. (2011) lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, **48**, 1-35.

Ruggiero, P.G.C., Batalha, M.A., Pivello, V.R. & Meirelles, S.T. (2002) Soil-vegetation relationships in cerrado (Brazilian savanna) and semideciduous forest, Southeastern Brazil. *Plant Ecology*, **160**, 1-16.

Schleuter, D., Daufresne, M., Massol, F. & Argillier, C. (2010) A user's guide to functional diversity indices. *Ecological Monographs*, **80**, 469-484.

Shipley, B. (2002) *Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference.* Cambridge University Press, Cambridge.

Silva, D.M. & Batalha, M.A. (2008) Soil–vegetation relationships in cerrados under different fire frequencies. *Plant and Soil*, **311**, 87–96.

Silva, I.A. & Batalha, M.A. (2010) Woody plant species co-occurrence in Brazilian savannas under different fire frequencies. *Acta Oecologica*, **36**, 85-91.

Srivastava, D.S & Vellend, M. (2005) Biodiversity-ecosystem functioning research: is it relevant to conservation? *Annual Review of Ecology, Evolution, and Systematics*, **36**, 267-294.

Stubbs, W.J. & Wilson, J.B. (2004) Evidence for limiting similarity in a sand dune community. *Journal of Ecology*, **92**, 557-567.

Sulkava, P. & Huhta, V. (1998) Habitat patchiness affects decomposition and faunal diversity: a microcosm experiment on forest floor. *Oecologia*, **116**, 390-396.

Tilman, D., Lehman, C.L. & Thomson, K.T. (1997) Plant diversity and ecosystem productivity: theoretical considerations. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 1857-1861.

Valenti, M.W., Cianciaruso, M.V. & Batalha, M.A. (2008) Seasonality of litterfall and leaf decomposition in a cerrado site. *Brazilian Journal of Biology*, **68**, 459-465.

Vasconcelos, H.L., Pacheco, R., Silva, R.C., Vasconcelos, P.B., Lopes, C.T., Costa, A.N. & Bruna, E.M. (2009) Dynamics of the leaf-litter arthropod fauna following fire in a neotropical woodland savanna. *PLoS ONE*, **11**, e7762.

Wagner, D., Brown, M.J.F., Gordon, D.M. (1997) Harvester ant nests, soil biota and soil chemestry. *Oecologia*, **112**, 232-236.

Wardle, D.A., Bonner, K.I. & Nicholson, K.S. (1997) Biodiversity and plant litter: experimental evidence which does not support the view that enhanced species richness improves ecosystem function, *Oikos*, **79**, 247-258.

Yachi, S. & Loreau, M. (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proceedings of the National Academy of Sciences of The United States of America*, **96**, 1463-1468.

Table 1. Variances (diagonal), covariances (upper triangle; boldface), and correlations (lower triangle) of the variables in our model (Fig. 1). We log-transformed all variables except N and P. N = total nitrogen (mg kg^{-1}), P = available phosphorus (mg kg^{-1}), CEC = cation exchange capacity (mmol kg^{-1}), Al = exchangeable aluminium (mmol kg^{-1}), slope (degrees), altitude (m), fire = inverted mean time between burnings (years), iFD = individual-based functional diversity, decomp = 6-month litter mass loss (g).

	N	**P**	**CEC**	**Al**	**slope**	**altitude**	**fire**	**iFD**	**decomp**
N	0.11	**0.27**	**0.05**	**0.05**	**-0.03**	**0.00**	**0.03**	**-0.06**	**0.01**
P	0.47	3.02	**0.14**	**0.13**	**-0.09**	**0.03**	**0.17**	**0.00**	**0.01**
CEC	0.57	0.31	0.07	**0.04**	**-0.01**	**0.00**	**0.01**	**-0.02**	**0.00**
Al	0.59	0.28	0.51	0.07	**-0.02**	**0.00**	**-0.01**	**0.04**	**-0.02**
slope	-0.15	-0.08	-0.03	-0.12	0.44	**-0.01**	**-0.03**	**0.17**	**0.01**
altitude	0.29	0.32	0.05	0.23	-0.22	0.00	**0.01**	**-0.01**	**0.00**
fire	0.31	0.30	0.11	-0.09	-0.15	0.40	-0.11	**-0.08**	**0.02**
iFD	-0.16	0.00	-0.05	0.14	0.24	-0.18	0.22	1.22	**0.03**
decomp	0.06	0.02	-0.03	-0.26	0.04	0.00	-0.24	0.10	0.07

Table 2. Standardised residuals for the model in Fig. 3. N = total nitrogen (mg kg^{-1}), P = available phosphorus (mg kg^{-1}), CEC = cation exchange capacity (mmol kg^{-1}), Al = exchangeable aluminium (mmol kg^{-1}), slope (degrees), fire = inverted mean time between burnings (years), iFD = individual-based functional diversity, decomp = 6-month litter mass loss (g).

	N	**P**	**CEC**	**Al**	**slope**	**fire**	**iFD**	**decomp**
N	0.00							
P	0.02	0.00						
CEC	0.00	-0.01	0.00					
Al	-0.01	-0.06	0.08	0.00				
slope	-0.02	-0.01	0.06	0.00	0.00			
fire	0.00	0.13	-0.10	0.00	0.00	0.00		
iFD	-0.02	0.08	0.04	0.00	0.00	-0.01	0.00	
decomp	0.01	-0.01	-0.07	0.00	-0.02	-0.05	0.00	0.00

Table 3. Predicted covariances for the model in Fig. 3. N = total nitrogen (mg kg^{-1}), P = available phosphorus (mg kg^{-1}), CEC = cation exchange capacity (mmol kg^{-1}), Al = exchangeable aluminium (mmol kg^{-1}), slope (degrees), fire = inverse of the mean time between burnings (years), iFD = individual-based functional diversity, decomp = 6-month litter mass loss (g).

	N	**P**	**CEC**	**Al**	**slope**	**fire**	**iFD**	**decomp**
N	0.11							
P	0.26	2.99						
CEC	0.05	0.14	0.07					
Al	0.05	0.15	0.03	0.07				
slope	-0.03	-0.09	-0.02	-0.02	0.44			
fire	0.03	0.10	0.02	-0.01	-0.03	-0.11		
iFD	-0.05	-0.15	-0.03	0.04	0.17	0.08	1.20	
decomp	0.00	0.01	0.00	-0.02	0.01	-0.02	0.03	0.07

Table 4. Standardised net effects (combined estimates of direct and indirect effects) between the variables in Fig. 3. fire = inverse of the mean time between burnings (years), fertility = latent variable representing soil fertility, Al = exchangeable aluminium (mmol kg^{-1}), slope (degrees), iFD = individual-based functional diversity, and decomp = 6-month litter mass loss (g).

effect	**direct**	**indirect**	**total**
fire → iFD	0.000	-0.172	-0.172
fire → decomp	0.000	0.176	0.176
fertility → iFD	-0.422	0.000	-0.422
fertility → decomp	0.555	-0.120	0.435
Al → iFD	0.454	0.000	0.454
Al → decomp	-0.672	0.129	-0.543
slope → iFD	0.229	0.000	0.229
slope → decomp	0.000	0.065	0.065

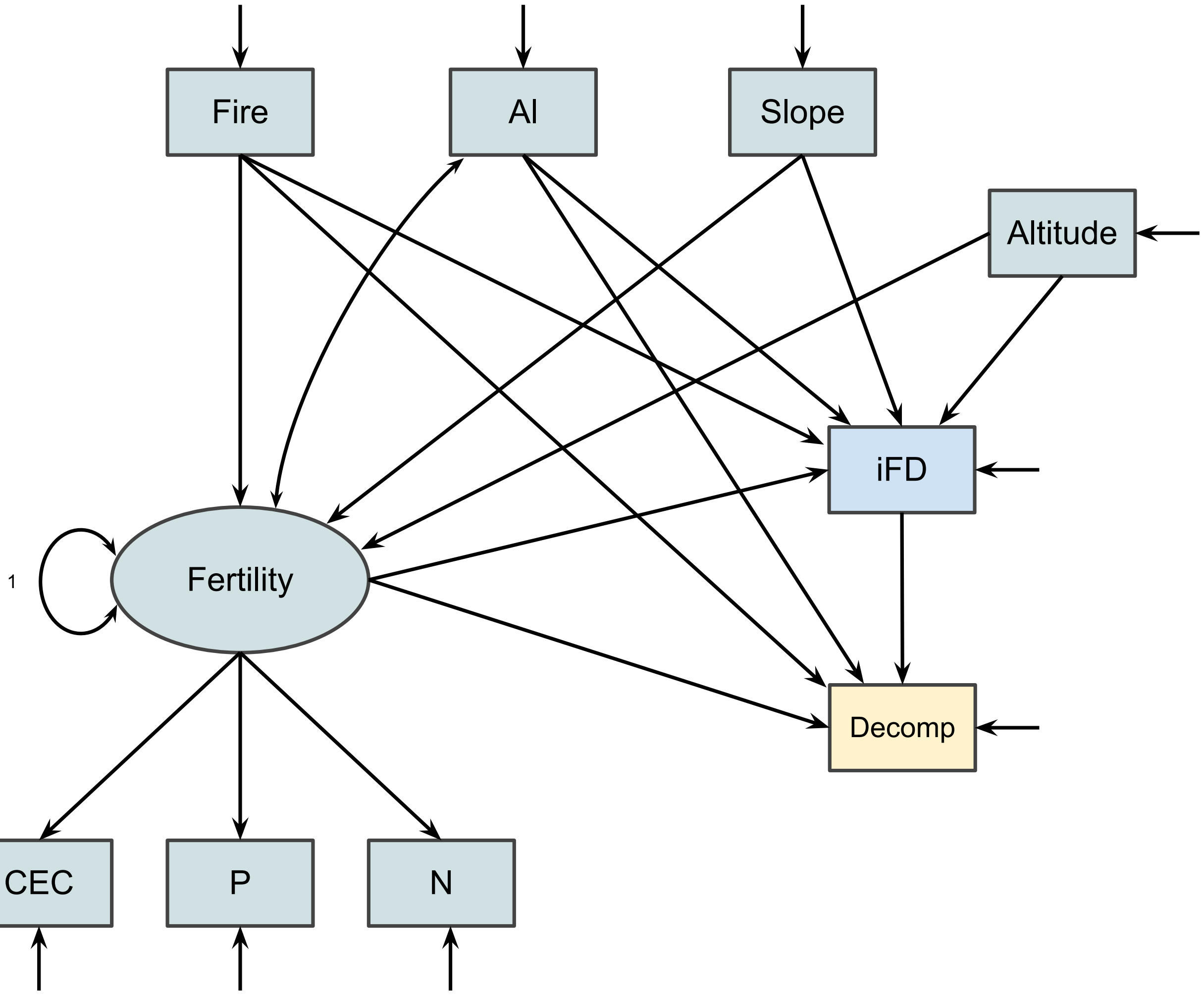


Fig 1. Structural equation model connecting fire, soil fertility, aluminium, water availability, functional diversity, and decomposition in Emas National Park, Brazil. Each arrow without a fixed value represents a parameter to be estimated. Rectangles represent directly measurable variables (for example, fire). Ovals represent theoretical concepts called latent variables (for example, fertility. Variables in green represent the environment, iFD represents biodiversity (blue), and decomposition (yellow) represents ecosystem functioning. N = total nitrogen (mg kg^{-1}), P = available phosphorus (mg kg^{-1}), CEC = cation exchange capacity (mmol kg^{-1}), Al = exchangeable aluminium (mmol kg^{-1}), slope (degrees), altitude (m), fire = inverse of the mean time between burnings (years), iFD = individual-based functional diversity, decomp = 6-month litter mass loss. We fixed the disturbance of the latent variable fertility to set the scale. We also fixed the variance of all exogenous variables to their observed values.

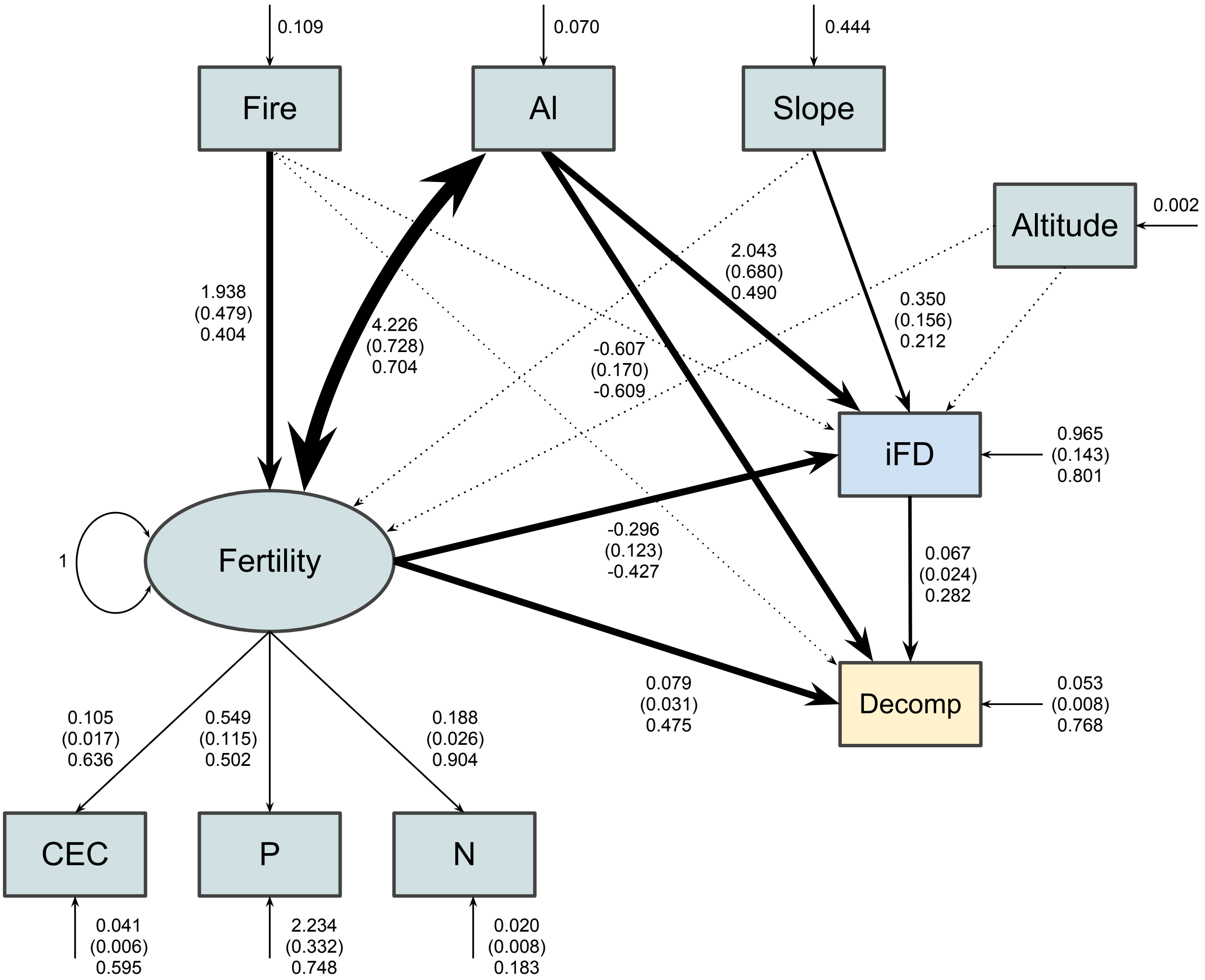


Fig 2. Unstandardised parameter estimates, followed by standard errors in parenthesis and standardised estimates of the model in Fig. 1. Solid lines indicate significant parameters ($\alpha = 0.05$), whilst dashed lines indicate parameters not significantly different from 0. Except for the indicators of fertility, line weights indicate the strength of the relationship. Variables in green represent the environment, iFD represents biodiversity (blue), and decomposition (yellow) represents ecosystem functioning. We fixed the disturbance of fertility to set the scale of the latent variable. We also fixed the variance of all exogenous variables to their observed values. The model had good fit with data ($ML\chi^2 = 17.244$; $P = 0.243$; $df = 14$). N = total nitrogen (mg kg^{-1}), P = available phosphorus (mg kg^{-1}), CEC = cation exchange capacity (mmol kg^{-1}), Al = exchangeable aluminium (mmol kg^{-1}), slope (degrees), altitude (m), fire = inverse of the mean time between burnings (years), iFD = individual-based functional diversity, decomp = 6-month litter mass loss.

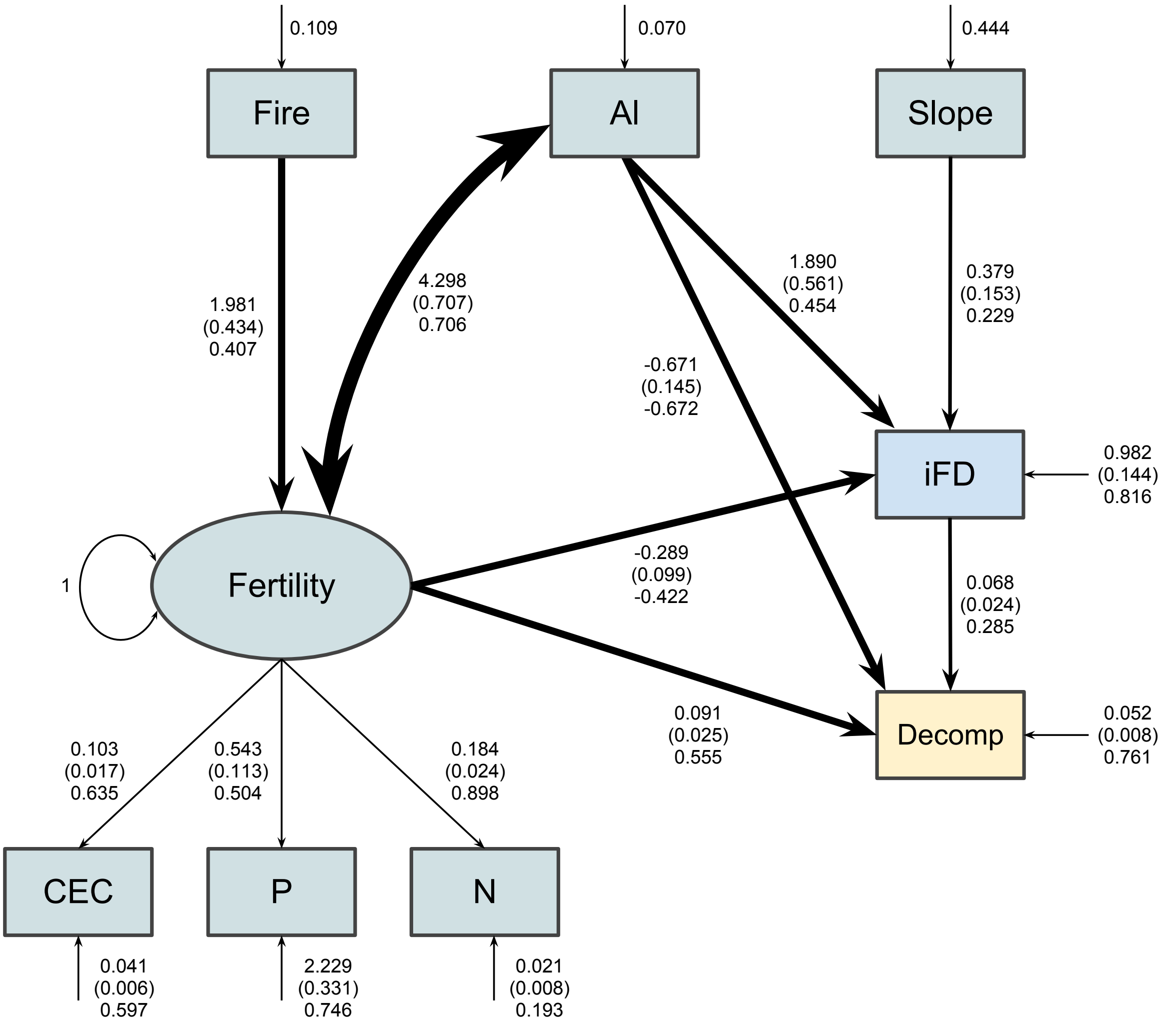


Fig 3. Unstandardised parameter estimates, followed by standard errors in parenthesis and standardised estimates of the model in Fig. 2 refitted without the non-significant parameters. Except for the indicators of fertility, line weights indicate the strength of the relationship. Variables in green represent the environment, iFD represents biodiversity (blue), and decomposition (yellow) represents ecosystem functioning. We fixed the disturbance of fertility to set the scale of the latent variable. We also fixed the variance of all exogenous variables to their observed values. The model had good fit with data (MLχ^2 = 10.289; *P* = 0.741; *df* = 14). N = total nitrogen (mg kg^{-1}), P = available phosphorus (mg kg^{-1}), CEC = cation exchange capacity (mmol kg^{-1}), Al = exchangeable aluminium (mmol kg^{-1}), slope (degrees), fire = inverse of the mean time between burnings (years), iFD = individual-based functional diversity, decomp = 6-month litter mass loss.

IV - CAPÍTULO 3

Artigo preparado para ser submetido ao periódico Community Ecology

The relative importance of traits, dispersal limitation, and demographic stochasticity to community assembly in a Neotropical savanna

Gustavo Henrique Carvalho‡[1] and Marco Antônio Batalha[1]

[1] Department of Botany, Federal University of São Carlos, PO Box 676, 13565-905, São Carlos, Brazil

---

‡ Correspondence author. E-mail: gustavo.bio@gmail.com

Abstract

Identifying the relative importance of deterministic and stochastic factors to the assembly of local communities has been tagged as one of the fundamental questions ecologists need to answer. Niche-based theories predict that the subset of species from the regional pool that occur in a local community presents a trait distribution that has been shaped by biotic and abiotic factors. The neutral theory, on the other hand, predicts that the species that compose the regional pool are functionally equivalent. Local communities formed from stochastic processes would, thus, present abundances not significantly different from those in the species pool. We collected traits on all woody individuals in 100 25m^2 quadrats to test if we could predict local abundances using a pool of species and traits. We combined plots into large scale, intermediate scale, and fine scale samples. We used random sampling, spatial distance, soil characteristics, and fire to assemble species in different scales and environmental gradients. In the Brazilian cerrado, fire and nutrient-poor soils are likely promoters of habitat filtering. To test if traits improved the predictions generated by the information present in the pool, we used maximum entropy models coupled with permutation tests. We could accurately predict local abundances of the 73 species in the pool. Dispersal limitation was the main factor assembling communities at all the scales we studied, but the effects of stochasticity became more important as the scale became local. Traits explained little of the uncertainty present in local abundances, but coupled with pool frequencies they yielded large coefficients of determination.

Keywords: maxent, savanna, species pool, cerrado, traits

Introduction

Unveiling the processes that drive community composition is one of the most pursued goals in ecology. Niche-based theories assume that the attributes of the species in the regional pool are key to determine which species are more likely to coexist (Silvertown 2004). Contrastingly, the neutral theory (Bell 2000, Hubbell 2001) carries a null assumption of functional equivalence between the species in the pool. Where there is strong biotic or environmental pressure, trait filtering will likely influence the assembly of plant communities (Kraft et al. 2008). However, in neutral communities, dispersal limitation and demographic stochasticity determine species occurrence, as all species in the pool are assumed functionally equivalent. Thus, species with higher frequencies in the regional the pool have more chance of dispersal and establishment into local assemblages, even though the per capita probabilities of immigration are assumed equal (Hubbell 2001, Shipley et al. 2012). Local species abundances should, therefore, closely match the abundances in the regional pool, with deviances being caused by ecological drift and trait differences playing a minimal role in the coexistence of species.

If deterministic forces like environmental filtering or competition are shaping the abundance distributions of local communities, co-occurring plants are expected to be either more similar or more different in their attributes than expected by chance (Kraft et al. 2008, Mayfield and Levine 2010). Regardless of the scenario, including traits in models of assembly where habitat filtering is expected to occur should allow us to identify whether deterministic or stochastic forces are contributing the most to species establishment and increase predictive power (Shipley et al. 2012). This is a fundamental step towards the goal of predicting species abundances based on their characteristics and on the environment and will render more accurate models on the impacts of a changing world on the distribution of plants. In addition, identifying the relative importance of deterministic and stochastic factors to assembly has been tagged as one of the fundamental questions ecologists still need to tackle (Sutherland et al. 2013).

Initially, studies of community assembly assumed both niche and neutral theories to be mutually exclusive, especially because the resistance against neutral assembly has always been strong (Wennekes et al. 2012). However, efforts towards unified frameworks have been appearing in the literature (Tilman 2004, Chase 2005, Gravel et al. 2006, Stokes and Archer 2010). In these frameworks, niche structuring, dispersal limitation, and demographic stochasticity are assumed either to act simultaneously or subsequently to shape communities. The importance of each driver of assembly, either deterministic or stochastic, is expected to differ given the strength of the environmental filter and interactions. For instance, according to the continuum hypothesis (Gravel et al. 2006), niche and neutral sorting form ends from deterministic to stochastic exclusion in a continuum. Where interactions are strong, competition may pull species apart, creating a community with higher complementarity of functional traits (Cavender-Bares et al. 2004, Kraft et al. 2007). At the other end of the continuum, where competition is less intense, dispersal limitation and random demographic events will be the main forces driving assembly and local communities will be neutral (Graves et al. 2006).

Besides biotic and abiotic factors, the scale of the sampling area is also crucial when investigating assembly rules (Harrison and Cornell 2008, Kraft and Ackerly 2010, Chase and Myers 2011, Yuan et al. 2011, Götzenberger et al. 2012). Whereas environmental filters are expected to be the dominant deterministic force sorting species at broader scales, local competitive interactions also influence colonisation

(Laliberté et al. 2009, Kooyman et al. 2010, Mokany and Roxburgh 2010, Lambers et al. 2012). Environmental factors acting as habitat filters at multiple scales will likely narrow the range of trait values in local communities. Competition between neighbours, on the other hand, limits trait similarity, acting as a disruptive force selecting traits.

In the Brazilian cerrado, low nutrient availability, high aluminium content, low pH, seasonal drought, and fire set the stage for trait-structured communities (Eiten 1972, Motta et al. 2002). Thus, habitat filtering is expected to be a major force governing the co-occurrence of plant species in local cerrado assemblages. Indeed, classical theories proposed by early investigators predicted that the higher the concentration of plant nutrients and the lower the concentration aluminium, the higher the density of woody individuals (Goodland & Pollard 1973). Therefore, according to these theories, the cerrado is a fertility gradient, with distinct physiognomies and subsets of species occurring under soils with distinct nutrient and aluminium concentrations. Using different mixtures of traits, Cianciaruso et al. (2012) observed fire-induced phenotypic structuring in some of the assemblages they studied. Silva & Batalha (2010) found phylogenetic overdispersion across a high number of woody cerrado species, whilst Silva et al. (2010) observed random structuring and trait clustering along gradients of fire frequency in a cerrado area. Although there is substantial evidence of trait clustering in cerrado communities, a question that remains largely unanswered: can we use traits to predict which subset of species from the pool are more likely to occur in areas under certain environmental conditions?

In this study, we built a pool with some of the species occurring within the boundaries of a conservation unit in Brazil, measured a few of their traits, and tried to answer the following questions: 1) how well can we predict species abundances over different spatial scales and environmental gradients using a species pool and community weighed traits?; 2) if traits are important for predicting local assemblages, which traits are the most important? These questions are important because, to our knowledge, this is the first study attempting to predict local species-specific abundances in the Brazilian cerrado. To assess abundances as well as the shared influence of deterministic (traits) and stochastic (dispersal limitation and demographic stochasticity) processes in shaping local assemblages, we used a maximum entropy framework (Shipley 2009, 2012). This framework has shown potential to be a powerful tool for predicting specific abundances under a variety of spatial scales, in current (Sonnier et al. 2010, Laughlin et al. 2011, Shipley 2012, Frenette-Dussalt et al. 2012) and future scenarios (Frenette-Dussalt et al. 2012).

Material and methods

Study area

We conducted this study in Emas National Park (ENP), Central Brazil. ENP is one of the largest and most important cerrado reserves in Brazil, with an area of around 133,000 ha and containing key elements that characterise the fauna, flora, and habitats of the cerrado. The cerrado vegetation in the park goes from open (68.1% of its area) to closed physiognomies (25.1% of its area). Other vegetation types, such as wet grasslands, riparian forests, and semideciduous forests cover the remaining 6.8% of the area (Ramos-Neto and Pivello 2000). The climate in ENP is tropical and humid, with a wet season between September and May and a dry season extending from June to late August. Annual rainfall and mean temperature lie around 1,745 mm and 24°C, respectively.

Sampling

We randomly placed 100 25m^2 quadrats in alongside the maintenance roads that cross ENP. In each quadrat, from September 2009 to January 2010, we identified all woody individuals with at least 3 cm of stem diameter at the soil level. On all identified individuals, we measured the values of six functional traits that are surrogates of important responses of plants to environmental conditions, such as nutrient availability, water availability, and fire (Cornelissen et al. 2003; Pausas and Paula 2005): (1) basal area (m^2), related to space occupation, resource uptake, total biomass, and reproductive capability; (2) height (m), associated with competitive vigour, fecundity, and growth after disturbance; (3) bark thickness (mm), related to resistance to disturbance; (4) wood density (mg mm^{-3}), related to potential carbon storage and growth after disturbance; (5) leaf size (mm^2), related to resistance to environmental stress and (6) specific leaf area (mm^2 mg^{-1}), associated with growth and maximum photosynthetic rate. Next, we calculated mean trait values per species and standardised the trait matrix to unit variance. Even though we had data on each plant we sampled in all plots, individual-based values were not appropriate for the statistical analyses we used. Thus, with the species-based trait matrix, we calculated community-weighted trait means (Shipley et al. 2012), $\bar{t}_{jk} = \sum_{i=1}^{S} ra_{ik}t_{ij}$, for each trait (*j*) in each local community (*k*) using the local abundances (*ra*) for the *i*th species present the local community. Community-weighted trait means give more weight to the traits of the most abundant species and are related to the biomass ratio hypothesis, according to which the traits of the dominant species will contribute the most to ecosystem properties (Grime 1998).

Species pool

To construct the pool of species, we combined the frequencies we found in 2009 with the frequencies from another study conducted in ENP in 2006 (see Silva and Batalha 2009 for details). Our pool consisted of 3,529 individuals of 73 species, around half of all cerrado woody species that occur in ENP (Batalha and Martins 2002).

Environmental variables

We collected five soil samples from each quadrat at 0-5 cm depth to organise plots in a gradient of soil fertility, toxicity, and water availability. We measured the following environmental variables from the soil samples: pH, organic matter, available phosphorus, total nitrogen, exchangeable potassium, exchangeable calcium, exchangeable magnesium, exchangeable aluminium, sum of bases, base saturation, aluminium saturation, cation exchange capacity, sand content, silt content, and clay content (see Silva and Batalha 2008 for details on chemical and physical analyses). These variables are commonly regarded as important for plant establishment and development. We also measured the altitude and slope of each plot as rough surrogates of water availability: in ENP, the higher the altitude and the lower the slope, the lower the availability of water in deep soil layers. To position our plots in a gradient of soil fertility, we used soil nutrient availability, aluminium, and the surrogates of water availability to construct a matrix of plots per environmental variables. Next, we standardised all variables to zero mean and unit variance and calculated a distance matrix using Euclidean distances. We used the environmental distance matrix as input in a principal coordinate analysis (PCoA). Finally, we ordered our plots according to their scores in the first PCoA axis to position them in the gradient.

Since the prohibition of cattle farming inside the park in 1984, a fire exclusion policy was instituted. However, the accumulation of dry biomass led to the occurrence of uncontrolled wildfires every 3-4 years, burning on average 80% of ENP's total area (França et al. 2007). In 1994, a catastrophic fire burned about 95% of the park. Since then, precautions to avoid similar burnings have been taken. Managers burn preventive firebreaks annually in the dry season and a fire brigade stays in ENP during this period to prevent anthropogenic fires (França et al. 2007). Nevertheless, even with these precautions, in August 2010, an anthropogenic fire burned 93% of the park's area. Preventive firebreaks, maintenance roads, and heterogeneous fuel availability make the fire map of ENP resemble a mosaic, with different plots possibly having very distinct fire histories, even if they are not distant from each other. Using in-field observations and satellite images, we could determine the fire history of each of our plots tracing back to 1984. We built a binary fire matrix with plots as rows and years as columns. We did a PCoA using the fire matrix and placed plots in the fire gradient following the order of their scores in the first PCoA axis.

Statistical analyses

To determine the relative importance of traits, dispersal limitation, and demographic stochasticity for the assembly of local communities, we used community assembly through trait selection (CATS) models developed by Shipley (2006, 2009, 2010). These models use the maximum entropy formalism (maxent) proposed by Shipley (2006, 2009) coupled with a permutation test (Shipley 2010) to estimate the relative importance of traits, dispersal limitation, and both dispersal limitation and traits to the assembly of communities. In sum, we first used the CATS models to predict local abundances and assess model fit through coefficients of determination. We used pool abundances and traits in this step. Next, we decomposed the coefficients of determination to remove model bias and determine how much of the explained variance in predictions was due to traits, dispersal limitation, and ecological drift, excluding biologically irrelevant uncertainty.

The first step to fit the CATS models was to calculate local predicted abundances with the 'maxent' function in the 'FD' library (Laliberté and Shipley 2011) for the R environment (R Core Team 2012). Maxent models are Bayesian models that use the frequencies in the pool to predict local abundances. Then, the permutation tests assess the amount of information gained when the traits of the species in the pool are included in the model. If traits do not explain a significant amount of the uncertainty in predicted abundances, local communities are a sample of the pool and traits do not influence these frequencies. This is an indication that the environment and interactions between the regional and local scales are not enough to leave patterns in the distribution of traits. However, if traits do provide significantly better predictions, both dispersal limitation and niche-based assembly can contribute to the formation of communities. As input in the 'maxent' function, we used community-weighted trait means as the macroscopic constraints, species-level trait means as the states, and regional abundances as the prior frequencies. We assessed the significance of model predictions using the function 'maxent.test' from the 'FD' library (Laliberté and Shipley 2011) with 999 randomisations. Hence, we could test if observed probabilities of each species in the pool and predicted local probabilities, estimated from the CATS models, were significantly similar.

We tested if traits significantly improved the information present in the pool by fitting CATS models for 2 groups of plots (large scale, 50 plots in each group), 5 groups of plots (intermediate scale, 20 plots in each group), and 10 groups of plots (fine scale, 10 plots in each group). We combined plots randomly, spatially, to account for possible significant factors we did not measure, and by their position in the soil fertility and fire gradients. We considered each of these groups as a local community. We had, thus, four CATS models for each local community in each grouping factor. So, we were able to test the joint effects of abiotic filters and spatial scale on frequency distributions. The coefficients of determination of CATS models included model bias, that is, biologically irrelevant uncertainty. To determine the amount of biologically relevant variation in local species abundances predicted by each CATS model, we followed the method developed by Shipley et al. (2012). The first CATS model we fitted for each group had permuted community-weighted trait means and a maximally uninformative prior, which consisted of a uniform pool where all species had the same abundances. By doing this, we removed any association between frequencies of the species in the pool, traits, and local abundances. All information present in the coefficient of determination given by this model [R^2(uniform)] was due to model structure and was removed from subsequent models. The second CATS model we fitted for each group of plots had observed pool abundances as the prior and permuted community-weighted trait means. So, this model returned a coefficient of determination [R^2(dispersal)] with only the contribution of dispersal limitation, that is, the frequencies in the pool, to local abundances. In this second CATS model, any contribution of traits to assembly was removed after we permuted the trait matrix. We fitted the third CATS model using observed community-weighted trait means and a maximally uninformative prior. By doing this, we removed the contribution of regional abundances to the coefficient of determination and retained only the information contributed by species traits [R^2(uniform, traits)]. Finally, we estimated the variation jointly explained by the pool and traits using the observed prior and observed community-weighted trait means [R^2(neutral, traits)].

To remove model bias, that is, to decompose the coefficients of determination to leave only biologically relevant information, we followed Shipley et al. (2012). This decomposition can be further divided into proportions explained uniquely by dispersion limitation, uniquely by traits, jointly by dispersal and traits, and unexplained but relevant biological information.

The proportion of biologically relevant information explained uniquely by the neutral prior (the abundances in the pool of species) is:

$$\frac{R^2(\text{neutral}, \text{traits}) - \hat{R}^2(\text{uniform}, \text{traits})}{1 - \hat{R}^2(\text{uniform})} \tag{1}$$

With this equation, we discounted the effects of traits and bias to local communities.

The proportion of biologically relevant information explained by local traits only is:

$$\frac{R^2(\text{neutral}, \text{traits}) - \hat{R}^2(\text{neutral})}{1 - \hat{R}^2(\text{uniform})} \tag{2}$$

This decomposition removes the effects of dispersal limitation and model bias from the local uncertainty explained by traits. The proportion of biologically relevant information explained by both local traits and dispersal limitation is as follows:

$$\frac{\hat{R}^2(\text{neutral}) + R^2(\text{uniform}, \text{traits}) - \hat{R}^2(\text{neutral}, \text{traits}) - \hat{R}^2(\text{uniform})}{1 - \hat{R}^2(\text{uniform})} \tag{3}$$

Finally, the proportion of biologically relevant uncertainty (demographic stochasticity) not explained by either dispersal or traits is as follows:

$$\frac{1 - R^2(\text{neutral}, \text{traits})}{1 - \hat{R}^2(\text{uniform})} \tag{4}$$

To determine the direction and strength of trait selection, we followed the procedures proposed by Sonnier et al. (2012). The method consists of fitting several multiple regression models with predicted probabilities as the response variables and standardised trait means as the explanatory variables. These coefficients are calculated during estimation in maxent models and are known as Lagrange multipliers (λ_{jk}). Lagrange multipliers quantify the strength of the relationship between each trait and predicted local abundances. $\lambda_{jk} > 0$ indicates that species with larger values for trait j in community k will become more abundant if all other traits are held constant. $\lambda_{jk} < 0$ indicates the opposite, species with larger values for trait j will become less abundant if all other traits are held constant. Finally, $\lambda_{jk} = 0$ indicates that trait j does not influence predicted abundances. We calculated Lagrange multipliers only for scenarios of pure trait selection as they provided us with sufficient information to answer our questions regarding the importance of traits for abundance predictions.

Results

Models with observed pool abundances and traits combined had the highest coefficients of determination when considering the total information present in local abundances, prior to bias removal (Table 1). At the larger scale, traits added significantly better information to that already present in the pool for all models (Fig. 1). At the intermediate scale, all but one model including traits and regional abundances were significant. Finally, at the finer scale, two joint-effect models were significant and two were not.

At the large scale, prior to removal of model bias, trait-only models explained around 28-29% of the uncertainty in species abundances for all grouping factors (‘traits’ column in Table 1). All coefficients of determination were significant, indicating that traits added valuable information to the maximally uninformative species pool. Using observed abundances in the pool and randomised traits in models explained about 62-65% for all grouping factors (‘dispersal’ column in Table 1). Including both observed traits and pool frequencies in models resulted in coefficients of determination ranging between 70-72% of explained uncertainty for all grouping factors (‘both’ column in

Table 1). We saw similar results at the intermediate scale, with traits explaining 24-26% of uncertainty in predictions with the maximally uninformative species pool ('traits' column in Table 1). Dispersal limitation explained 55-58% ('dispersal column in Table 1) and both traits and dispersal explained 60-64% ('both' column in Table 1). Traits added significant information to the frequencies in the observed species pool in all but spatially grouped assemblages. Finally, at the fine scale, traits explained 23-24% of the uncertainty for all grouping factors, dispersal limitation accounted for 47-52% of this uncertainty, and both traits and dispersal limitation explained 53-56% of the variation in predictions. Traits did not add significant information to observed pool frequencies in assemblages constructed randomly and by fire occurrence. Thus, traits and pool frequencies had lower explanatory power as the scale became local.

By decomposing the coefficients of determination to remove biologically irrelevant information, we observed that dispersal limitation and demographic stochasticity explained most of the uncertainty in local communities, regardless of how we grouped our plots (Table 2). Again, how we grouped plots had very little impact on how well traits, dispersal, both, and stochasticity were able to explain local observed abundances (Table 2). However, as we further divided plots into more groups, narrowing the scale of local communities, the role of chance events (demographic stochasticity) assembling these communities became more important. For instance, for all four ways we used to sort plots into groups, demographic stochasticity explained more than 30% of the uncertainty when we arranged plots into two groups and about 50% when we separated plots into 10 groups (Table. 2). Thus, biologically relevant uncertainty became higher as the scale of the communities became local.

Bark thickness was the trait that had the greatest influence on most local abundances (Appendix 3). We observed positive effects on frequency predictions in all communities, despite of spatial scale and biotic factor we used to group plots. The effects of wood density on abundances were also mainly positive, although much smaller than the effects of bark thickness (Appendix 3). Thus, species with thicker barks and higher wood densities were favoured. Oppositely, basal area and height had negative effects on abundances in most assemblages. Thinner and shorter individuals had larger predicted frequencies. Leaf size and specific leaf area had negative effects on most communities, with individuals with smaller and thinner leaves being more abundant.

Discussion

Using maximum entropy models with broad scale species frequencies and a few traits, we could accurately predict local abundances of the 73 species in the pool. Dispersal limitation was the main factor assembling communities at all the scales we studied. Pure trait models explained some of the uncertainty in species abundances, but most of it was explained when we added trait information to pool frequencies in maxent models. When we decomposed the coefficients of determination to account for model bias, an interesting pattern emerged: the importance of traits and dispersal limitation and traits became smaller at finer scales, whereas the importance of demographic stochasticity became greater. We also quantified the importance of some traits (Lagrange multipliers) to species abundances. For instance, species co-occurrence in this Neotropical savanna was greatly influenced by bark thickness. Species with thicker barks had higher local frequencies.

In all maxent models where we used a maximally uninformative prior, the inclusion of observed community-weighed trait means rendered significantly better predictions of local abundances. This was expected since there is empirical evidence of niche-based assembly in tropical communities (Karst et al. 2005, Kraft et al. 2008, Shipley et al. 2012), including the cerrado (Silva and Batalha 2010, Batalha et al. 2011a, Cianciaruso et al. 2012). Even though most pure-trait models were significant, they accounted for a relatively low amount of the uncertainty present in predictions for all scales and environmental gradients, as shown in their small coefficients of determination, especially after partitioning. It appears that we either could not detect the environmental cause of trait structuring since different grouping factors rendered models with similar predictions or did not include important response traits. For instance, water is a key resource in the cerrado and likely an important source of niche partitioning in its plant communities (Oliveira et al. 2005). Additionally, fire occurrence influence soil characteristics in the cerrado, so fire and soil gradients should have some degree of overlapping. Including more accurate measurements of soil water availability might shed some light on the factors causing local trait structuring in this Neotropical savanna. Using the same framework we used in this study and meta-community trait means, Shipley et al. (2012) found results similar to ours, with pure-trait models accounting for a small proportion of the assembly of tropical communities in French Guiana. An attempt to disentangle the importance of niche and spatial processes in a Canadian temperate forest yielded contrasting results, though (Laliberté et al. 2009). In this temperate forest, niche partitioning was the most important cause of the distribution of tree seedling abundances.

Dispersal limitation was the main factor influencing community structuring in our study area. Prior to bias removal, pool abundances explained at least half of the variation in local abundances and this percentage was higher at broader scales. This is an indication that traits are being weakly filtered within ENP, especially at narrower scales. A review of community assembly publications estimated that 18% niche-based assembly studies did not find evidence of trait distribution different from random

(Götzenberger et al. 2012), so the low importance of traits for assembly we found here was not surprising. Since our pool included only species found in the cerrado vegetation in ENP, they already had gone through regional environmental filtering and had the appropriate attributes to establish and grow in the park (Batalha et al. 2011a). Local abundances should, thus, reflect abundances in the environmental conditions in the landscape (Shipley et al. 2012). However, models with dispersal limitation and community-weighted trait means rendered significantly better predictions, accounting for as much as 73.5% of local abundances. Therefore, even if environmental filtering is not greatly sorting attributes at finer scales, there is some deterministic mechanism shaping local trait distributions. This needs further investigation since spatial distance, soil characteristics, and fire as grouping factors yielded similar coefficients of determination. Alternatively, we may have overlooked important response traits and some of their contribution to community sorting could be emerging through their correlations with the traits we did measure.

Decomposing coefficients of determination to leave only biologically relevant information showed that cerrado assemblages in ENP were sensitive to stochastic processes. The importance of demographic stochasticity to assembly increased with increasing spatial scales. Indeed, these patterns were in accordance with theoretical predictions (Chase and Myers 2011). Moreover, the patterns that emerged from information decomposition in this study were in agreement with those from a previous one (Shipley et al. 2012). Niche and scale partitioning in a temperate forest also resulted from sharing effects of both deterministic and stochastic events (Laliberté et al. 2009). Similarly, stochastic events were the main process determining the establishment of species and, afterwards, interactions among individuals further shaped a parkland community (Stokes and Archer 2010). Thus, growing evidence indicates that both types of processes simultaneously shape communities from different vegetation types, including this Neotropical savanna, at different spatial scales.

Bark thickness was the trait that most contributed to abundance predictions according to Lagrange multipliers. The cerrado vegetation in ENP suffers from recurring burning events, thus it is expected that species with thicker barks would have more individuals. In general, basal area, height, leaf size, and specific leaf area were negatively related to predicted abundances. Smaller and thinner species were more abundant, which may be caused by recurrent fires imposing a difficulty for individuals to grow (Batalha et al. 2011b). Species with smaller leaves and lower specific leaf area were also more abundant. Since Lagrange multipliers are partial regression coefficients, it is hard to evaluate their unique impact on predictions in the presence of other traits. High correlation between them could be an important confounding factor (Sonnier et al. 2012). Investigating the causal relations between traits with structural equation modelling prior to maxent fitting could provide guidance to accommodate trait covariance (Sonnier et al. 2012). However, since maxent models heavily rely on Lagrange multipliers, including structural equation models in their structure would require careful consideration and is beyond the scope of this article.

In conclusion, we showed that it is possible to predict cerrado species in local assemblages with great accuracy combining pool abundances and trait data. Also, the environmental variables we assumed to be possible environmental filters yielded very similar predictions. Researchers can easily extend the framework we used to predict the outcome of local and broad environmental changes to species occurrence, making it a powerful tool for conservationists. Moreover, we were able to unveil the traits that contribute the most to species abundances. Future studies could identify which species from the regional pool can become invasive and predict the outcome of this invasion. Since researchers can use maxent models with virtually any measure of biological diversity, future studies could map invasive grasses occurrence and predict the fate of native species, especially rare ones, under different management scenarios. Finally, as previously noted, including more species in the pool, measuring more traits and, and more accurately assessing other possible sources of environmental variation, such as water availability, may render even better predictions and aid conservation efforts.

Acknowledgements

We are grateful to Fapesp, CNPq, and CAPES, for financial support; to Ibama and Emas National Park staff, for research permission; to Helena França, for providing the fire history.

References

Algar, AC, Kerr, JT, Currie, DJ (2011) Quantifying the importance of regional and local filters for community trait structure in tropical and temperate zones. Ecology 92: 903-914.

Batalha, MA, Martins, FR (2002) The vascular flora of the cerrado in Emas National Park (Goiás, Central Brazil). Sida 20: 295-311.

Batalha, MA, Silva, IA, Cianciaruso, MV, Carvalho, GH (2011a) Trait diversity on the phylogeny of cerrado woody species. Oikos 120: 1741-1751.

Batalha, MA, Silva, IA, Cianciaruso, MV, França, H, Carvalho, GH (2011b) Phylogeny, traits, environment, and space in cerrado plant communities at Emas National Park (Brazil). Flora 206: 949-956.

Bell, G (2000) The distribution of abundance in neutral communities. The American Naturalist 155: 606-617.

Cavender-Bares, J, Ackerly, DD, Baum, DA, Bazzaz, FA (2004) Phylogenetic overdispersion in Floridian oak communities. American Naturalist 163: 823-843.

Chase, JM (2005) Towards a really unified theory for metacommunities. Functional Ecology 19: 182-186.

Chase, JM, Myers, JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. Philosophical Transactions of the Royal Society 366: 2351-2363.

Cianciaruso, MV, Silva, IA, Batalha, MA, Gaston, KJ, Petchey, OL (2012) The role of fire on phylogenetic and functional structure of woody savannas: moving from species to individuals. Perspectives in Plant Ecology, Evolution and Systematics 14: 205-216.

Cornelissen, JHC, Lavorel, S, Garnier, E, Díaz, S, Buchmann, N, Gurvich, DE, Reich, PB, ter Steege, H, Morgan, HD, var der Heijden, MGA, Pausas, JG, Poorter, H (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany 51: 335-380.

Eiten, G (1972) The cerrado vegetation of Brazil. The Botanical Review 38: 201-341.

França, H, Ramos-Neto, MB, Setzer, A (2007) O fogo no Parque Nacional das Emas. Ministério do Meio Ambiente, Brasília.

Frenette-Dussalt, C, Shipley, B, Meziane, D, Hingrat, Y (2012) Trait-based climate change predictions of plant community structure in arid steppes. Journal of Ecology *in press*.

Goodland, R, Pollard, R (1973) The Brazilian cerrado vegetation: a fertility gradient. Journal of Ecology 61: 219-224.

Götzenberger, L, de Bello, F, Bråthen, KA, Davison, J, Dubuis, A, Guisan, A, Leps, J, Lindborg, R, Moora, M, Pärtel, M, Loic, P, Pottier, J, Vittoz, P, Zobel, K, Zobel, M (2012) Ecological assembly rules in plant communities - approaches, patterns, and prospects. Biological Reviews 87: 111-127.

Gravel, D, Canham, CD, Beaudet, M, Messier, C (2006) Reconciling niche and neutrality: the continuum hypothesis. Ecology Letters 9: 399-409.

Grime, JP (1998) Benefits of plant diversity to ecosystems: immediate filter and founder effects. Journal of Ecology, 86, 902-910.

Harrison, S, Cornell, H (2008) Toward a better understanding of the regional causes of local community richness. Ecology Letters 11: 969-979.

Hubbell, SP (2000) The unified neutral theory of biodiversity and biogeography. In Monographs in Population Biology (eds Simon A. Levin and Henry S. Horn). Princeton University Press.

Karst, J, Gilbert, B, Lechowicz, MJ (2005) Fern community assembly: the roles of chance and the environment at local and intermediate scales. Ecology 86: 2473-2486.

Kooyman, R, Cornwell, W, Westoby, M (2010) Plant functional traits in Australian subtropical rain forest: partitioning within-community from cross-landscape variation. Journal of Ecology 98: 517-525.

Kraft, NJB, Cornwell, WK, Webb, CO, Ackerly, DD (2007) Trait evolution, community assembly, and the phylogenetic structure of ecological communities. American Naturalist 170: 271-283.

Kraft, NJB, Valencia, R, Ackerly, DD (2008) Functional traits and niche-based tree community assembly in an Amazonian forest. Science 322: 580-582.

Kraft, NJB, Ackerly, DD (2010) Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. Ecological Monographs 80: 401-422.

Laliberté, E, Paquette, A, Legendre, P, Bouchard, A (2009) Assessing the scale-specific importance of niche and other spatial processes on beta diversity: a case study from a temperate forest. Oecologia 159: 377-388.

Lambers, JHR, Adler, PB, Harpole, WS, Levine, JM, Mayfield, MM (2012) Rethinking community assembly through the lens of coexistence. Annual Review of Ecology, Evolution, and Systematics 43: 227-248.

Laughlin, DC, Fulé, PZ, Huffman, DW, Crouse, J, Laliberté, E (2011) Climatic constrains on trait-based forest assembly. Journal of Ecology 99: 1489-1499.

Mayfield, M, Levine, JM (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. Ecology Letters 13: 1085-1093.

Mokany, K, Roxburgh, SH (2010) The importance of spatial scale for trait-abundance relations. Oikos 119: 1504-1514.

Motta, PEF, Curi, N, Franzmeier, DP (2002) Relation of soils and geomorphic surfaces in the Brazilian cerrado. The cerrados of Brazil: ecology and natural history of a neotropical savanna (eds P.S Oliveira & R.J. Marquis). pp. 13-32. Columbia University, New York.

Oliveira, RS, Bezerra, L, Davidson, EA, Pinto, F, Klink, CA, Nesptad, DC, Moreira, A (2005) Deep root function in soil water dynamics in cerrado savannas of central Brazil. Functional Ecology 19: 574-581.

Pausas, JG, Paula, S (2005) Plant functional traits database for Euro–Mediterranean ecosystems. Eufirelab Deliverable D–04–06. URL: http//www.eufirelab.org.

Punchi-Manage, R, Getzin, S, Wiegand, T, Kanagaraj, R, Gunatilleke, CVS, Gunatilleke, IAUN, Wiegand, K, Huth, A (2013) Effects of topography on structuring local species assemblages in a Sri Lankan mixed dipterocarp forest. Journal of Ecology 101: 149-160.

R Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Ramos-Neto, MB, Pivello, VR (2000) Lightning fires in a Brazilian savanna national park: rethinking management strategies. Environmental Management 26: 675–684.

Rosindell, J, Hubbell, SP, & Etienne, RS (2011) The unified neutral theory of biodiversity and biogeography at age ten. Trends in Ecology and Evolution 26: 340-348.

Shipley, B (2009) From plant traits to vegetation structure. Cambridge University Press, Cambridge.

Shipley, B, Paine, CET, Baraloto, C (2012) Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly. Ecology 93: 760-769.

Silva, IA, Batalha, MA (2010) Woody plant species co-occurrence in Brazilian savannas under different fire frequencies. Acta Oecologica 36: 85-91.

Silva, IA, Carvalho, GH, Loiola, PP, Cianciaruso, MV, Batalha, MA (2010) Herbaceous and shrubby species co-occurrences in Brazilian savannas: the roles of fire and chance. Community Ecology 11: 97-104.

Silvertown, J (2004) Plant coexistence and the niche. Trends in Ecology and Evolution 19: 605-610.

Sonnier, G, Shipley, B, Navas, M (2010) Plant traits, species pools and the prediction of relative abundance in plant communitites: a maximum entropy approach. Journal of Vegetation Science 21: 318-331.

Sonnier, G, Navas, M, Fayolle, A, Shipley, B (2012) Quantifying trait selection driving community assembly: a test in herbaceous plant communities under contrasted land use regimes. Oikos 121: 1103-1111.

Stokes, CJ, Archer, SR (2010) Niche differentiation and neutral theory: an integrated perspective on shrub assemblages in a parkland savanna. Ecology 91: 1152-1162.

Sutherland, WJ, Freckleton, RP, Godfray, HCJ, Beissinger, SR, Benton, T, Cameron, D, Carmel, Y, Coomes, DA, Coulson, T, Emmerson, MC, Hails, RS, Hays, GC, Hodgson, DJ, Hutchings, MJ, Johnson, D, Jones, JPG, Keeling, MJ, Kokko, H, Kunin, WE, Lambin, X, Lewis, OT, Malhi, Y, Mieszkowska, N, Milner-Gulland, EJ, Norris, K, Phillimore, AB, Purves, DW, Reid, JM, Reuman, DC, Thompson, K, Travis, JMJ, Turnbull, LA, Wardle, DA, Wiegand, T (2013) Identification of 100 fundamental ecological questions. Journal of Ecology 101: 58-67.

Tilman, D (2004) Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences 101: 10854-10861.

Wennekes, PL, Rosindell, J, Etienne, RS (2012) The neutral-niche debate: a philosophical perspective. Acta Biotheorica 60: 257-271.

Yuan, Z, Gazol, A, Wang, X, Lin, F, Ye, J, Bai, X, Li, B, Hao, Z (2011) Scale specific determinants of tree diversity in an old growth temperate forest in China. Basic and Applied Ecology 12: 488-495.

Table 1. Estimated coefficients of determination (R^2) of the total information in assemblages of woody species in Emas National Park, Brazil. * indicates R^2 from models where adding trait data significantly improved predictions. 'traits' refer to pure-trait models, 'dispersal' to pure-neutral models, 'both' to the information explained jointly by traits and dispersal. We divided the sampling area into groups of 50 (large scale), 20 (intermediate scale), and 10 (fine scale) plots. We arranged plots into groups randomly, according to spatial proximity, to a gradient of soil nutrient, aluminium, and water availability, and to a fire frequency gradient. Boldface indicates significant coefficients.

	Large scale			**Intermediate scale**			**Fine scale**		
Random	traits	dispersal	both	traits	dispersal	both	traits	dispersal	both
	0.294	0.624	**0.709**	**0.253**	0.581	**0.642**	**0.230**	0.501	0.564
Space	traits	dispersal	both	traits	dispersal	both	traits	dispersal	both
	0.280	0.656	**0.725**	**0.248**	0.553	0.603	**0.237**	0.475	**0.538**
Soil	traits	dispersal	both	traits	dispersal	both	traits	dispersal	both
	0.286	0.624	**0.706**	**0.263**	0.551	**0.630**	**0.246**	0.497	**0.568**
Fire	traits	dispersal	both	traits	dispersal	both	traits	dispersal	both
	0.288	0.623	**0.706**	**0.253**	0.562	**0.627**	**0.234**	0.520	0.564

Table 2. Decomposition of the biologically relevant information from coefficients of determination. 'traits' refer to pure-trait models, 'dispersal' to pure-neutral models, 'both' to the information explained jointly by traits and dispersal, and 'stochasticity' to the biologically relevant uncertainty that could not be explained by traits, dispersal limitation, or both. We divided the sampling area into groups of 50 (large scale), 20 (intermediate scale), and 10 (fine scale) plots. We arranged plots into groups randomly, according to spatial proximity, to a gradient of soil nutrient, aluminium, and water availability, and to a fire frequency gradient.

	Large scale				**Intermediate scale**				**Fine scale**			
Random	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity
	0.077	0.484	0.099	0.339	0.051	0.451	0.082	0.416	0.029	0.400	0.051	0.520
Space	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity
	0.078	0.522	0.077	0.322	0.040	0.420	0.069	0.471	0.039	0.355	0.058	0.547
Soil	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity
	0.074	0.487	0.097	0.341	0.065	0.428	0.075	0.433	0.039	0.382	0.068	0.511
Fire	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity	traits	dispersal	both	stochasticity
	0.072	0.486	0.100	0.342	0.046	0.436	0.082	0.435	0.032	0.399	0.040	0.529

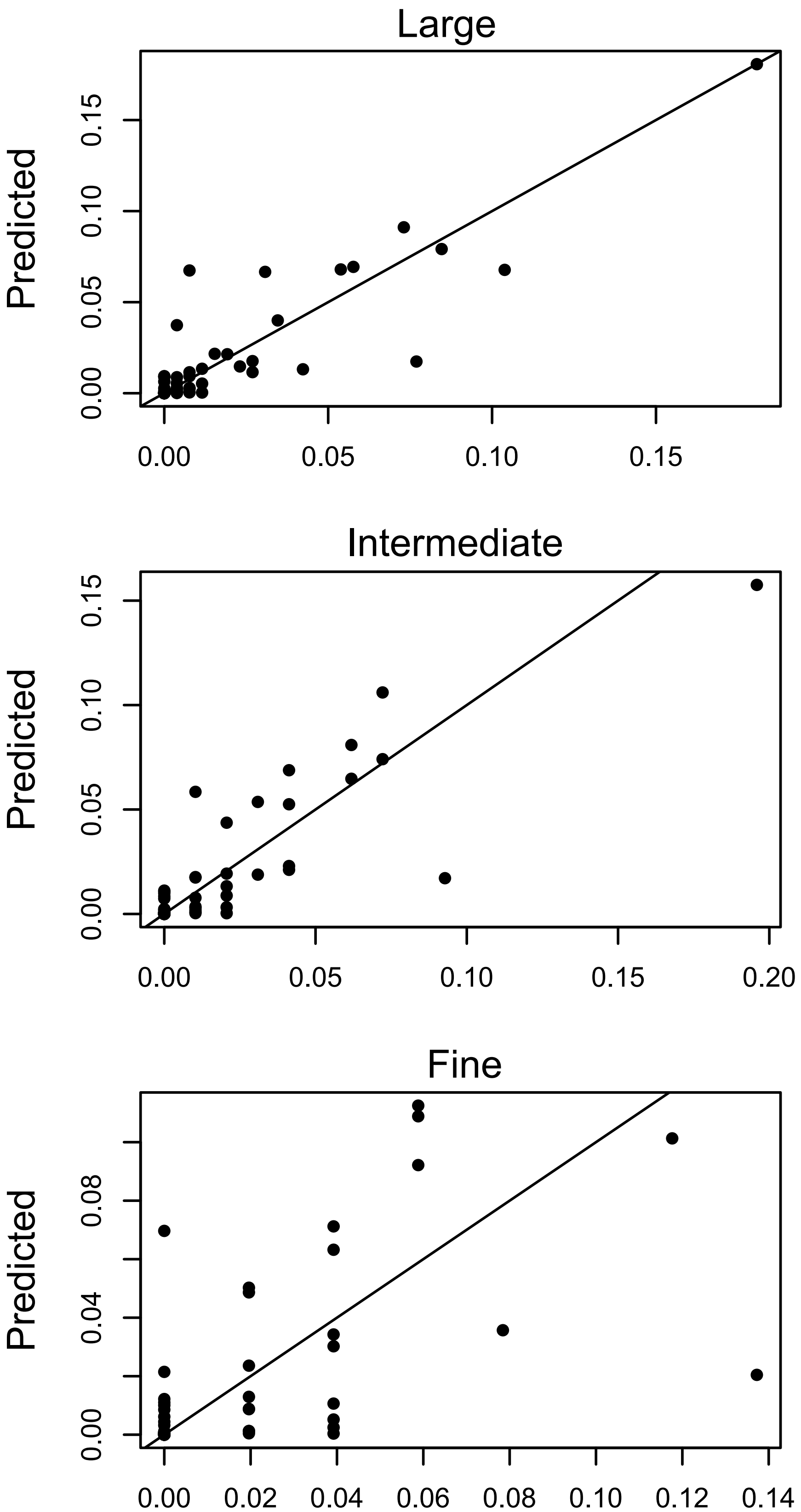

Large
Predicted
0.00
0.05
0.10
0.15
0.00
0.05
0.10
0.15
Intermediate
Predicted
0.00
0.05
0.10
0.15
0.00
0.05
0.10
0.15
0.20
Fine
Predicted
0.00
0.04
0.08
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
Observed

Figure 1. Predicted vs. observed abundances in communities along a fire gradient over different scales. The solid line represents the 1:1 line. Large scale: R^2=0.70, $P < 0.05$. Intermediate scale: R^2=0.63, $P < 0.05$. Fine scale: R^2=0.56, $P > 0.05$.

Appendix 1. Pool species with abundances and average trait values. BA = basal area (m^2), H = height (m), Brk = bark thickness (mm), Woo = wood density (mg mm^{-3}), LSz = leaf size (mm^2), SLA = specific leaf area (mm^2 mg^{-1}).

Family	**Species**	**Abundance**	**BA**	**H**	**Brk**	**Woo**	**LSz**	**SLA**
Fabaceae	*Mimosa amnis-atri* Barneby	486	0.002	1.074	4.315	0.583	7690.353	10.623
Sapotaceae	*Pouteria ramiflora* (Mart.) Radlk.	454	0.004	1.950	11.138	0.435	9964.381	7.308
Sapotaceae	*Pouteria torta* (Mart.) Radlk.	303	0.003	1.343	7.497	0.449	8709.128	6.798
Fabaceae	*Stryphnodendron adstringens* (Mart.) Coville	255	0.004	1.842	7.170	0.471	51890.161	7.169
Myrtaceae	*Psidium laruotteanum* Cambess.	234	0.002	1.105	10.586	0.560	3290.758	5.621
Asteraceae	*Eremanthus erythropappus* (DC.) MacLeish	218	0.004	1.209	12.527	0.466	3012.500	9.048
Ochnaceae	*Ouratea acuminata* (DC.) Engl.	210	0.003	1.513	11.622	0.462	3232.800	6.388
Bignoniaceae	*Tabebuia ochracea* A.H. Gentry	149	0.002	1.003	8.873	0.441	24108.000	5.486
Ochnaceae	*Ouratea spectabilis* (Mart. ex Engl.) Engl.	148	0.007	2.616	12.120	0.494	5070.800	5.018
Connaraceae	*Rourea induta* Planch.	126	0.001	0.750	9.520	0.223	1997.000	6.200
Asteraceae	*Piptocarpha rotundifolia* (Less.) Baker	102	0.004	1.574	7.099	0.508	9672.211	5.889
Ebenaceae	*Diospyros hispida* A.DC.	100	0.003	1.430	6.529	0.348	19711.375	4.683
Annonaceae	*Annona crassiflora* Mart.	84	0.006	1.314	5.861	0.258	7618.196	6.127
Fabaceae	*Anadenanthera falcata* (Benth.) Speg.	77	0.008	1.676	9.452	0.475	23616.427	6.681

Family	**Species**	**Abundance**	**BA**	**H**	**Brk**	**Woo**	**LSz**	**SLA**
Fabaceae	*Acosmium dasycarpum* (Vogel) Yakovlev	52	0.002	0.821	8.871	0.436	10697.083	6.983
Erythroxylaceae	*Erythroxylum su*berosum A.St.-Hil.	49	0.003	1.398	11.524	0.525	2864.882	8.164
Dilleniaceae	*Davilla elliptica* A.St.-Hil.	45	0.002	1.185	7.218	0.560	3663.833	7.734
Connaraceae	*Connarus suberosus* Planch.	42	0.004	1.390	12.059	0.429	13160.595	5.972
Arecaceae	*Allagoptera leucocalyx* (Drude) Kuntze	36	0.002	0.839	6.453	0.248	191166.886	4.114
Fabaceae	*Dimorphandra mollis* Benth.	32	0.007	1.949	10.700	0.438	33899.000	8.504
Malvaceae	*Eriotheca gracilipes* (K.Schum.) A.Robyns	26	0.006	2.312	10.820	0.359	30491.500	4.508
Erythroxylaceae	*Erythroxylum campestre* A.St.-Hil.	24	0.001	1.135	12.365	0.493	2555.000	6.919
Calophyllaceae	*Kielmeyera coriacea* Mart. & Zucc.	21	0.006	1.785	15.136	0.283	8908.438	5.898
Solanaceae	*Solanum lycocarpum* A. St.-Hil.	20	0.007	1.380	12.880	0.486	6925.000	8.875
Rubiaceae	*Palicourea rigida* Kunth	19	0.004	1.380	9.465	0.228	16986.500	5.201
Myrtaceae	*Eugenia piauhiensis* O. Berg	17	0.006	2.042	9.195	0.357	3265.805	7.133
Melastomataceae	*Miconia albicans* (Sw.) Steud.	17	0.001	1.742	5.870	0.607	4985.750	4.922
Fabaceae	*Anadenanthera peregrina* (L.) Speg.	15	0.010	2.603	16.021	0.579	16813.000	7.390
Malpighiaceae	*Byrsonima coccolobifolia* Kunth	15	0.008	2.768	9.304	0.441	7494.400	8.802
Caryocaraceae	*Caryocar brasiliense* A.St.-Hil.	15	0.007	2.900	11.930	0.400	51227.000	9.081
Bignoniaceae	*Tabebuia aurea* (Silva Manso) Benth. & Hook.f. ex S.Moore	15	0.002	1.267	13.600	0.235	45682.000	3.851

Family	**Species**	**Abundance**	**BA**	**H**	**Brk**	**Woo**	**LSz**	**SLA**
Anacardiaceae	*Anacardium humile* A.St.-Hil.	11	0.003	0.445	0.308	0.680	8822.463	4.369
Salicaceae	*Casearia sylvestris* Sw.	11	0.002	1.725	7.365	0.482	1200.750	8.667
Malvaceae	*Eriotheca pubescens* (Mart. & Zucc.) Schott & Endl.	8	0.001	1.050	6.340	0.257	59792.000	4.943
Myrtaceae	*Myrcia bella* Cambess.	7	0.004	1.301	12.701	0.498	908.714	7.330
Asteraceae	*Vernonia bardanoides* Less.	7	0.001	0.783	0.473	0.328	8657.146	7.875
Proteaceae	*Roupala montana* Aubl.	6	0.002	2.065	6.270	0.541	7738.500	4.577
Myrtaceae	*Eugenia punicifolia* (Kunth) DC.	5	0.003	1.880	14.105	0.544	874.500	10.183
Lythraceae	*Lafoensia pacari* A. St.-Hil.	5	0.006	2.200	6.370	0.551	2718.000	13.230
Fabaceae	*Machaerium acutifolium* Vogel	5	0.010	4.080	17.110	0.563	16743.500	7.280
Vochysiaceae	*Qualea grandiflora* Mart.	5	0.019	3.337	10.601	0.416	8875.862	4.468
Myrtaceae	*Eugenia aurata* O.Berg	4	0.002	1.130	11.030	0.435	2087.500	10.200
Euphorbiaceae	*Manihot tripartita* (Spreng.) Müll.Arg.	4	0.001	0.407	0.173	0.613	1414.595	7.700
Styracaceae	*Styrax ferrugineus* Nees & Mart.	4	0.023	4.025	12.390	0.484	2229.250	5.186
Malpighiaceae	*Byrsonima basiloba* A.Juss.	3	0.050	2.083	4.923	0.551	8648.000	5.066
Asteraceae	*Chromolaena squalida* (DC.) R.M.King & H.Rob.	3	0.0007	0.816	0.101	0.608	1010.143	8.741
Fabaceae	*Diptychandra aurantiaca* Tul.	3	0.002	1.560	11.420	0.563	12617.667	14.715
Fabaceae	*Hymenaea stigonocarpa* Hayne	3	0.003	1.603	2.423	0.523	22160.667	7.937

Family	**Species**	**Abundance**	**BA**	**H**	**Brk**	**Woo**	**LSz**	**SLA**
Annonaceae	*Duguetia furfuracea* (A.St.-Hil.) Saff.	2	0.0005	0.745	0.125	0.352	4418.803	4.821
Myrtaceae	*Myrcia camapuanensis* N.Silveira	2	0.002	0.980	6.610	0.704	6382.500	4.072
Vochysiaceae	*Qualea parviflora* Mart.	2	0.008	5.150	9.190	0.399	2888.000	12.127
Fabaceae	*Sclerolobium aureum* (Tul.) Baill.	2	0.003	1.915	5.710	0.414	31410.500	7.874
Lamiaceae	*Aegiphila lhotzkiana* Cham.	1	0.009	2.822	7.861	0.279	7046.156	6.431
Fabaceae	*Albizia niopoides* (Benth.) Burkart	1	0.001	1.790	6.470	0.553	11511.000	13.330
Apocynaceae	*Aspidosperma tomentosum* Mart.	1	0.002	0.830	11.170	0.413	5681.000	7.228
Fabaceae	*Bauhinia rufa* (Bong.) Steud.	1	0.0007	1.366	0.618	0.734	6661.036	6.739
Malpighiaceae	*Byrsonima verbascifolia* (L.) Rich. ex Juss.	1	0.008	3.210	14.660	0.463	10873.000	7.615
Myrtaceae	*Campomanesia pubescens* (Mart. ex DC.) O.Berg	1	0.0005	0.681	0.396	1.074	1611.874	5.925
Asteraceae	*Chresta sphaerocephala* DC.	1	0.0002	0.447	0.079	0.493	2721.754	5.849
Araliaceae	*Didymopanax macrocarpus* (Cham. & Schltdl.) Seem.	1	0.003	0.594	1.244	0.321	50857.153	3.757
Erythroxylaceae	*Erythroxylum tortuosum* Mart.	1	0.002	0.890	7.670	0.487	5180.000	6.199
Myrtaceae	*Eugenia bimarginata* DC.	1	0.005	1.560	11.630	0.445	4026.000	8.492
Nyctaginaceae	*Guapira noxia* (Netto) Lundell	1	0.007	0.680	5.700	0.205	12436.000	10.518
Apocynaceae	*Hancornia speciosa* Gomes	1	0.011	2.760	2.600	0.361	3150.000	10.468
Apocynaceae	*Himatanthus obovatus* (Müll.Arg.) Woodson	1	0.001	0.690	2.014	0.211	14465.307	5.726

Family	**Species**	**Abundance**	**BA**	**H**	**Brk**	**Woo**	**LSz**	**SLA**
Melastomataceae	*Miconia ferruginata* DC.	1	0.010	1.754	5.924	0.474	31426.604	3.635
Melastomataceae	*Mouriri elliptica* Mart.	1	0.002	0.970	9.870	0.543	5567.000	6.504
Myrtaceae	*Myrcia guianensis* (Aubl.) DC.	1	0.001	0.710	7.910	0.743	4691.000	3.874
Myrtaceae	*Myrcia lasiantha* DC.	1	0.005	1.540	15.230	0.261	936.000	6.375
Myrtaceae	*Myrcia obovata* (O.Berg) Nied.	1	0.004	1.130	6.610	0.560	4087.000	9.135
Celastraceae	*Plenckia populnea* Reissek	1	0.011	5.370	11.210	0.449	4079.000	15.755
Araliaceae	*Schefflera malmei* (Harms) Frodin	1	0.002	3.930	3.010	0.401	12199.000	3.809
Fabaceae	*Sclerolobium* sp1. Vogel	1	0.013	3.361	2.655	0.523	28092.070	5.500

Appendix 2. Correlations between average trait values. BA = basal area (m^2), H = height (m), Brk = bark thickness (mm), Woo = wood density (mg mm^{-3}), LSz = leaf size (mm^2), SLA = specific leaf area (mm^2 mg^{-1}).

	BA	**H**	**Brk**	**Woo**	**LSz**	**SLA**
BA	1.000					
H	0.184	1.000				
BRK	-0.132	0.347	1.000			
Woo	-0.099	-0.064	-0.192	1.000		
LSz	0.099	-0.068	-0.048	-0.309	1.000	
SLA	-0.249	0.277	0.148	0.081	-0.242	1.000

Appendix 3. Lagrange multipliers of pure-trait models. Grey shading represents significant values. BA = basal area (m^2), H = height (m), Brk = bark thickness (mm), Woo = wood density (mg mm^{-3}), LSz = leaf size (mm^2), SLA = specific leaf area (mm^2 mg^{-1}).

Grouping	Scale	BA	H	Brk	Woo	LSz	SLA
Random	**Large**	-0.0037	-0.0025	0.0074	0.0016	-0.0014	-0.0017
		-0.0038	-0.0034	0.0077	0.0002	-0.0003	-0.0022
	Intermediate	-0.0042	-0.002	0.0071	0	-0.0011	-0.0027
		-0.0036	-0.0037	0.0083	0.001	-0.0007	-0.0021
		-0.0036	-0.0023	0.0082	0.0004	-0.0005	-0.001
		-0.0037	-0.0033	0.0076	0.002	-0.0009	-0.0023
		-0.0036	-0.0035	0.0066	0.001	-0.0012	-0.0017
	Fine	-0.0037	-0.004	0.0066	0.0001	-0.0017	-0.0022
		-0.0037	-0.0024	0.0095	0.0025	-0.0012	-0.003
		-0.0026	-0.0062	0.0109	-0.0003	-0.0002	-0.0006
		-0.0026	-0.0057	0.0064	0.0033	0.0009	0.0008
		-0.0038	-0.0025	0.0068	-0.0001	-0.001	-0.0011
		-0.0048	0.0007	0.0087	0.0011	-0.0017	-0.0045
		-0.0041	-0.003	0.0082	0.0001	-0.002	-0.0037
		-0.0047	-0.0009	0.0067	0.0011	0.0005	-0.0035
		-0.0043	0.0004	0.0026	0.0025	-0.0008	0.0005
		-0.0036	-0.0039	0.007	-0.0008	-0.003	-0.0019
Space	**Large**	-0.0043	-0.0019	0.0059	0.0017	-0.0017	-0.003
		-0.0034	-0.0037	0.0089	0.0001	-0.0016	-0.0017
	Intermediate	-0.005	-0.0003	0.0049	0.0011	-0.001	-0.0038
		-0.0035	-0.0036	0.0062	0.0025	-0.0016	-0.0015
		-0.003	-0.0047	0.0086	0.0015	0.0008	-0.0004
		-0.0034	-0.0029	0.0081	0.0001	-0.0017	-0.0011
		-0.0038	-0.0029	0.0094	0.0001	-0.0015	-0.003
	Finer	-0.0027	-0.0037	0.0104	0.0026	-0.0028	-0.0014
		-0.0037	-0.003	0.0102	0.0012	0.0001	-0.0029
		-0.0045	-0.0026	0.007	-0.0012	-0.0035	-0.0047
		-0.0038	-0.0019	0.0089	0.0008	-0.0007	-0.0022
		-0.0044	-0.0035	0.0056	-0.0002	-0.0002	-0.0031
		-0.0049	-0.0017	0.0028	-0.002	0.0003	-0.0016
		-0.004	-0.0008	0.0055	0.0018	0	-0.0001
		-0.0033	-0.0047	0.006	0.0002	0.0018	0.0007
		-0.0048	0.0004	0.0042	0.002	-0.0018	-0.0027
		-0.0021	-0.0064	0.0106	0.0033	-0.0001	-0.0001
Soil	**Large**	-0.0037	-0.0029	0.0079	0.0004	-0.0006	-0.0017
		-0.0038	-0.0028	0.0072	0.0015	-0.0012	-0.0022

	Intermediate	-0.0041	-0.0011	0.0078	0.0005	-0.0009	-0.0018
		-0.0035	-0.005	0.0083	-0.0004	0.0001	-0.0019
		-0.0031	-0.0046	0.0053	0.002	-0.0007	0.0003
		-0.0046	-0.0017	0.0059	0.0003	-0.0017	-0.0036
		-0.0032	-0.003	0.01	0.0028	-0.001	-0.002
	Fine	-0.0043	-0.0003	0.008	0.0003	-0.0022	-0.0029
		-0.0034	-0.0032	0.0073	0.0009	0.0022	0.0008
		-0.0034	-0.0053	0.009	-0.0017	-0.0019	-0.0025
		-0.0036	-0.0046	0.0075	0.001	0.0022	-0.0012
		-0.0032	-0.0037	0.0071	0.0023	-0.0013	-0.0009
		-0.0029	-0.0057	0.0033	0.0017	-0.0001	0.0017
		-0.0045	-0.0013	0.007	0.0003	-0.0028	-0.0041
		-0.0046	-0.0022	0.0046	0.0002	-0.0005	-0.003
		-0.0028	-0.0045	0.0104	0.0037	-0.0025	-0.0027
		-0.0035	-0.0021	0.0097	0.0022	0	-0.0015
Fire	**Large**	-0.0038	-0.003	0.0073	0.0002	-0.0012	-0.0019
		-0.0037	-0.0028	0.0078	0.0016	-0.0006	-0.002
	Intermediate	-0.00353	-0.00458	0.00786	-0.00096	-0.00145	-0.00195
		-0.00401	-0.00231	0.00754	0.00102	-0.00008	-0.00226
		-0.00332	-0.00341	0.00782	0.00094	-0.00204	-0.00118
		-0.00365	-0.00204	0.00817	0.00296	0.00035	-0.00154
		-0.00427	-0.00233	0.00625	0.00012	-0.00148	-0.00284
	Fine	-0.0027	-0.0051	0.0096	-0.0013	-0.0033	-0.0005
		-0.0043	-0.0041	0.0063	-0.0006	0.0002	-0.0033
		-0.0038	-0.0015	0.0061	0.002	-0.0003	-0.0007
		-0.0042	-0.0033	0.0092	-0.0002	0.0002	-0.0041
		-0.0039	-0.0015	0.0059	0.0007	-0.0026	-0.0012
		-0.0026	-0.0058	0.0103	0.0012	-0.0013	-0.0011
		-0.0028	-0.003	0.0086	0.0049	0.0013	0.0005
		-0.0047	-0.0008	0.0076	0.0005	-0.0009	-0.0042
		-0.0034	-0.0036	0.0095	-0.0002	-0.0018	-0.0021
		-0.005	-0.0012	0.0033	0.0004	-0.0011	-0.0035

V - CAPÍTULO 4

Artigo preparado para ser submetido ao periódico Biology Letters (máximo de 2500 palavras)

Revisiting the drivers of woody species richness and density in a Neotropical savanna using structural equation modelling

Gustavo Henrique Carvalho§[1] and Marco Antônio Batalha[1]

[1] Department of Botany, Federal University of São Carlos, PO Box 676, 13565-905, São Carlos, Brazil

---

§ Correspondence author. E-mail: gustavo.bio@gmail.com

Abstract

Environmental filtering prevents species without certain attributes to occur in local communities. Traits respond differently to different abiotic factors, assembling communities with varying composition along environmental gradients. Here, we measured proxies of soil fertility, disturbance by fire, response and physiological traits to assess how these variables interact to determine woody species richness and density in a Neotropical savanna. We explicitly incorporated our assumptions about how different abiotic filters influence different subsets of traits into a statistical test using structural equation modelling, yielding a more accurate representation of assembly process. Fire had an effect on resistance traits, whereas soil fertility influenced physiological traits. Resistance traits explained both the richness and density of plots, whereas physiological traits explained only the density. Fewer fire events lead to richer and denser plots. Similarly, areas with lower cation exchange capacity assembled less dense communities. Furthermore, we showed that structural equation modelling allowed us to better represent the interactions of distinct environmental filters with different subsets of traits.

Keywords: cerrado, community assembly, savanna, structural equation modelling, traits.

Introduction

Environmental filtering prevents species without certain trait values to occur in local communities. Distinct abiotic factors can filter species at different spatial scales and points in time (Lambers et al. 2012). Thus, assemblages are likely formed by multiple sources of trait filtering (Lambers et al. 2012). Accordingly, subsets of traits respond differently to different abiotic factors, assembling communities with varying composition along environmental gradients (Keddy 1992, Lavorel and Garnier 2002). Since some traits are more relevant to certain ecosystem processes than others (Petchey and Gaston 2006), incorporating their functional relevance in models of assembly will render more realistic translations of how the environment and organisms interact to shape communities and rates of ecosystem processes.

In savannas, soil fertility and fire are important promoters of trait filtering (Gignoux et al. 1997, Batalha et al. 2010), determining which species from the regional pool are able to co-occur locally (Keddy 1992, Lambers et al. 2012). In the Brazilian cerrado, studies have indicated that soil characteristics (Goodland and Pollard 1973) and disturbance caused by fire (Moreira 2000, Silva and Batalha 2010, Batalha et al. 2010) play a majour role in the composition of communities, influencing the distribution of traits, richness, and density of woody individuals. Indeed, patches of forest formations within the cerrado domain indicate that both vegetations share the same climate requirements. However, not all functional traits respond similarly to soil fertility and fire in the cerrado. For instance, only half of the functional traits measured by Dantas et al. (2013) in a cerrado area had distributions correlated with fire frequency. Likewise, the effects of fire on functional diversity are dependent on the subset of traits used to calculate the index (Cianciaruso et al. 2012). Even though there is a building amount of research relating fire and the distribution of traits in the

Brazilian cerrado, it is still uncertain how traits mediate the interaction between abiotic factors and richness and density of plants. The role of different subsets of functional traits in this mediation also remains to be addressed.

Here, we measured proxies of soil fertility, disturbance by fire, response and physiological traits to assess how these variables interact to determine woody species richness and density, while describing how structural equation modelling with latent variables can improve assembly models. To our knowledge, this is the first study to quantify the joint effects of soil fertility and fire on traits and, consequently, the effects of traits on community composition in a Neotropical savanna.

Material and methods

We conducted this study in Emas National Park, Central Brazil. Emas is one of the most important cerrado reserves in Brazil, with an area of around 133,000 ha. The cerrado vegetation in the park presents a striking variation in the density of woody individuals. We randomly placed 100 25 m^2 quadrats in the park and, on each of the 531 woody individuals that occurred in the quadrats, measured the values of six functional traits that represent responses of plants to environmental conditions, such as nutrient availability and fire (Cornelissen et al. 2003): (1) basal area (m^2), related to space occupation, resource uptake, total biomass, and reproductive capability; (2) height (m), associated with competitive vigour, fecundity, and growth after disturbance; (3) bark thickness (mm), related to resistance to disturbance; (4) leaf nitrogen (mg g^{-1}), related to maximum photosynthetic rates and nutrient stress; (5) specific leaf area (mm^2 mg^{-1}), associated with growth and maximum photosynthetic rate; and (6) leaf toughness (N), associated with resistance to herbivore and leaf lifespan.

Fire is a major source of disturbance in Emas (França et al. 2007). Preventive firebreaks, maintenance roads, and heterogeneous fuel availability make the fire map of Emas resemble a mosaic, with nearby plots possibly having very distinct fire histories. Using in-field observations and satellite images, we counted the number of fire events between 1984 and 2010 as a proxy for the disturbance caused by burnings. As an indicator of soil fertility, we used cation exchange capacity. To measure this variable, we collected five soil subsamples from each plot, combined them into one sample, and sent them for chemical analysis.

We used structural equation modelling with robust estimators (Shipley 2000) to test two competing models of the causal connections between abiotic factors, traits, and community composition. Structural equation models provide the means to test hypotheses that represent alternative causal structures of any level of complexity, allowing researchers to analyse their data from a system perspective. If appropriate, researchers can include theoretical concepts in structural models as latent variables, which are expected to express themselves in the shared covariance between observed variables called indicators (for instance, leaf nitrogen content and basal area in Fig. 1a). We proposed two *a priori* structural equation models (Fig. 1). In the first model (Fig. 1a), all traits are caused by only one latent variable. In the second model (Fig 1b), observed traits are further divided into resistance and physiological traits. Each trait value in our proposed models was the average of that trait for all individuals in a quadrat. We had, thus, 100 observations for each trait. We determined density as the number of woody individuals per square meter. Prior to testing the full structural equation models, we did two confirmatory factor analyses (Shipley 2000) to assess the validity of the latent variables and their indicators in both structural equation models. Poor fit in these confirmatory models indicated that our choice of latent variables was not appropriate, that is, the shared variance between subsets of traits did not accurately represent a theoretical variable, in this case, the role of the traits in the interaction between plants and the environment. We did all analyses using the 'lavaan' package (Rosseel 2012) for R (R Core Team 2012).

Results

The causal structure with all traits represent by one latent variable had poor fit with data ($\chi^2 = 79.296$, $df = 26$, $p = 0$). A confirmatory factor analysis of the latent variable in this model also showed poor fit ($\chi^2 = 53.249$, $df = 9$, $p = 0$). The factor analysis indicated that all the traits we measured were not caused by a single latent variable, since bark thickness, basal area, and height had non-significant path coefficients, whereas leaf nitrogen content, specific leaf area, and leaf toughness had significant coefficients. Separating traits into two latent variables yielded a confirmatory factor model with a much better fit ($\chi^2 = 7.304$, $df = 8$, $p = 0.50$). All indicators of the two factors were significant.

The structural equation model with two sets of traits represented by two latent variables quickly converged to a solution and fitted the data well (Fig. 2, $\chi^2 = 38.284$, $df = 28$, $p = 0.093$). The loadings of both latent variables were significant. Of all paths we initially considered plausible, the estimates of the effects of fire on physiological traits, cation exchange capacity on resistance traits, physiological traits on richness, and the covariance between physiological and resistance traits were non-significant. Fire had a high influence on resistance traits and a non-significant one on physiological traits. Similarly, the path coefficient connecting cation exchange capacity and physiological traits was significant, whilst the estimate of the effect of cation exchange capacity on resistance traits was not. Richness was explained mainly by resistance traits, whereas the density of woody individuals was explained by both resistance and physiological traits.

Discussion

Our results offered support to our expectation that separating functional traits into latent variables better representing their functional roles would render more accurate models of how the environment and traits interact. The structural equation model with one latent variable representing all traits had poor fit with data, indicating that even if all measured traits are influenced by habitat filtering, different subsets of traits are subjected to different levels of trait filtering. Furthermore, including subsets of traits as latent factors according to their main role in the interaction between plants and the environment allowed us to more acuratelly represent the importance of different traits for community assembly. Finally, models with latent variables account for the covariance between traits and our imprecision in measuring them, leading to more reliable estimates (Shipley 2000).

The use of confirmatory factor analyses prior to full model estimation provided evidence for our assumptions about the concepts each subset of traits represented. Indeed, as we expected, resistance and physiological traits were indeed better represented by distinct latent variables. This approach could be useful even in studies where structural equation modelling is not used, as it provides a statistical test of the assumptions researchers make in studies involving functional traits. Ecologists often rely on methods that do not involve statistical analyses to choose which traits are important for a given system (Petchey and Gaston 2006). However, by using confirmatory factor analysis and structural equation modelling, one can test whether subsets of traits indeed represent a given concept. Moreover, this method can incorporate causal connections between traits, helping with the identification of redundancy.

Our model corroborated the observation that careful selection of traits will have an impact on models of community assembly (Petchey et al. 2004, Cianciaruso et al. 2012). Instead of simply removing or doing individual statistical analysis for each subset of traits based on their functional roles, we incorporated these roles in the analysis by using latent variables. Indeed, we were able to corroborate that fire did influence resistance traits, as previously described (Silva and Batalha 2010, Cianciaruso et al. 2012), but not physiological traits. Physiological and resistance traits were correlated in the confirmatory factor analysis, although they were not correlated in the structural equation model, indicating that the inclusion of fire and cation exchange capacity accounted for most of the variation between the latent variables.

We showed that plots that burn less have, on average, taller and thicker individuals, with thicker barks, which is in line with previous findings (Batalha et al. 2011, Dantas et al. 2013). High number of fires led, via the indirect effects of fire on richness and density through resistance traits, to plots with fewer species and individuals, probably due to recurrent top-killing (Higgins et al. 2007). Fire suppression has been shown to assemble richer and denser communities (Moreira 2000) and our results corroborated this.

The path coefficient connecting cation exchange capacity and resistance traits was not significant. Physiological traits, on the other hand, were significantly influenced by cation exchange capacity. Plots with higher fertility had individuals with higher leaf nitrogen content and specific leaf area and softer leaves. Physiological traits did not have a significant effect on richness. They did have, however, a significant negative effect on plant density. Higher cation exchange capacity led to less dense plots via physiological traits. Thus, our results unveiled a significant negative path between soil fertility and density. Previous findings either described positive (Goodland and Pollard 1973) or not significant (Moreira 2000, Ruggiero et al. 2002) relationships between soil nutrient availability and the density of woody plants.

In conclusion, we showed that explicitly incorporating theoretical concepts about the functional roles of plant traits into assembly models could indeed yield more realist representations of the how communities respond to abiotic filters and the consequences of these responses to patterns of diversity.

Acknowledgements

We are grateful to Fapesp, CNPq, and CAPES, for financial support; to Ibama and Emas National Park staff, for research permission; to Helena França, for providing the fire history.

References

Batalha, MA, Silva, IA, Cianciaruso, MV, França, H, Carvalho, GH (2011) Phylogeny, traits, environment, and space in cerrado plant communities at Emas National Park (Brazil). Flora 206: 949-956.

Cianciaruso, MV, Silva, IA, Batalha, MA, Gaston, KJ, Petchey, OL (2012) The role of fire on phylogenetic and functional structure of woody savannas: moving from species to individuals. Perspectives in Plant Ecology, Evolution and Systematics 14: 205-216.

Cornelissen, JHC, Lavorel, S, Garnier, E, Díaz, S, Buchmann, N, Gurvich, DE, Reich, PB, ter Steege, H, Morgan, HD, var der Heijden, MGA, Pausas, JG, Poorter, H

(2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. Australian Journal of Botany 51: 335-380.

Dantas, VL, Pausas, JG, Batalha, MA, Loiola, PP, Cianciaruso, MV (2013) The role of fire in structuring trait variability in Neotropical savannas. Oecologia 171: 487-494.

França, H., Ramos-Neto, M.B., Setzer, A. (2007) O fogo no Parque Nacional das Emas. Ministério do Meio Ambiente, Brasília.

Gignoux, J, Clobert, J, Menaut, JC (1997) Alternative fire resistance strategies in savanna trees. Oecologia 110: 576-583.

Goodland, R, Pollard, R (1973) The Brazilian cerrado vegetation: a fertility gradient. Journal of Ecology 61: 219-224.

Higgins, SI, Bond, WJ, February, EC, Bronn, A, Euston-Brown, DI, Enslin, B, Govender, N, Rademan, L, O'Reagan, S, Potgieter, AL, Scheiter, S, Sowry, R, Trollope, L, Trollope, WS (2007) Effects of four decades of fire manipulation on woody vegetation structure in savanna. Ecology 88: 1119-1125.

Keddy, P.A. (1992) Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3: 157-164.

Lambers, JHR, Adler, PB, Harpole, WS, Levine, JM, Mayfield, MM (2012) Rethinking community assembly through the lens of coexistence. Annual Review of Ecology, Evolution, and Systematics 43: 227-248.

Lavorel, S, Garnier, E (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. Functional Ecology 16: 545-556.

Petchey, OL, Hector, A, Gaston, KJ (2004) How do different measures of functional diversity perform? Ecology 85: 847-857.

Petchy, OL, Gaston, KJ (2006) Functional diversity: back to basics and looking forward. Ecology Letters 9: 741-758.

R Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

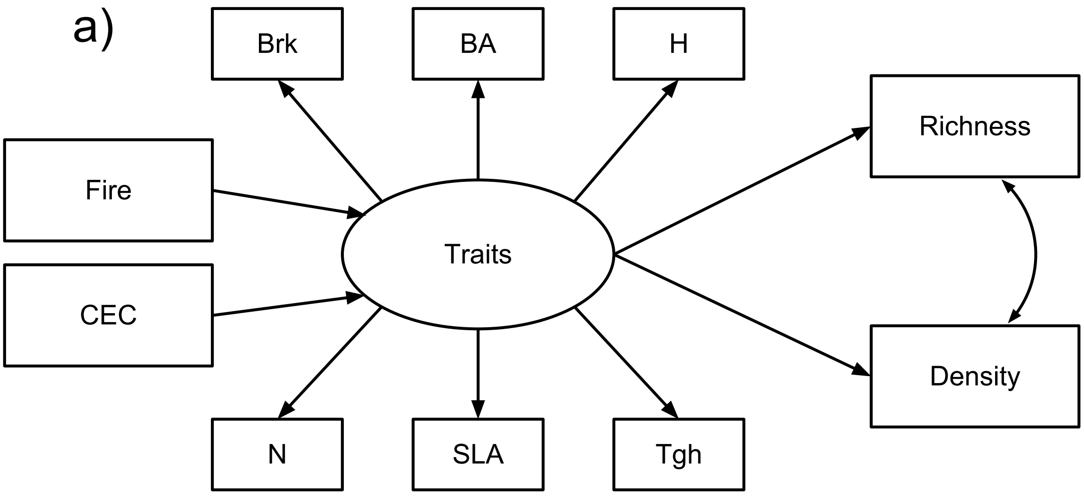
Rosseel, Y (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software 48: 1-36.

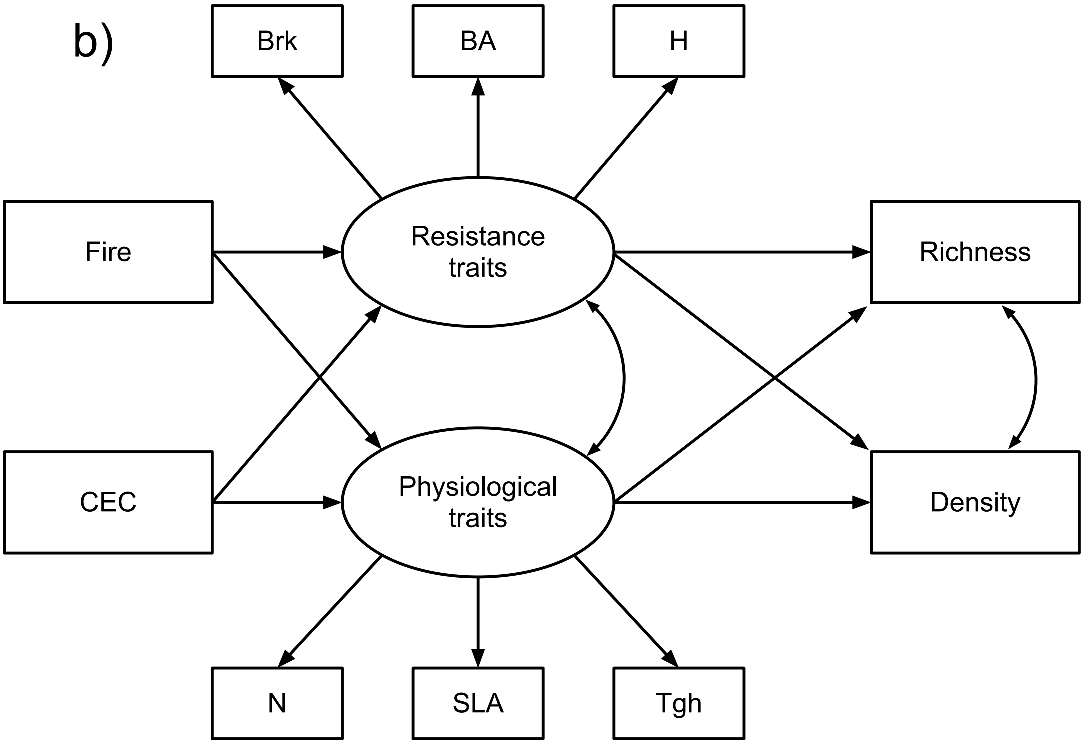
Ruggiero, PGC, Batalha, MA, Pivello, VR, Meirelles, ST (2002) Soil-vegetation relationships in cerrado (Brazilian savanna) and semideciduous forest, Southeastern Brazil. Plant Ecology 160: 1-16.

Silva, I.A. & Batalha, M.A. (2010) Woody plant species co-occurrence in Brazilian savannas under different fire frequencies. Acta Oecologica 36: 85-91.

Silva, IA, Carvalho, GH, Loiola, PP, Cianciaruso, MV, Batalha, MA (2010) Herbaceous and shrubby species co-occurrences in Brazilian savannas: the roles of fire and chance. Community Ecology 11: 97-104.

Shipley, B. (2000) Cause and Correlation in Biology: A User’s Guide to Path Analysis, Structural Equations, and Causal Inference. 1st edn. Cambridge University Press, Cambridge.


a)
Brk
BA
H
Fire
Traits
Richness
CEC
Density
N
SLA
Tgh


b)
Brk
BA
H
Fire
Resistance traits
Richness
CEC
Physiological traits
Density
N
SLA
Tgh

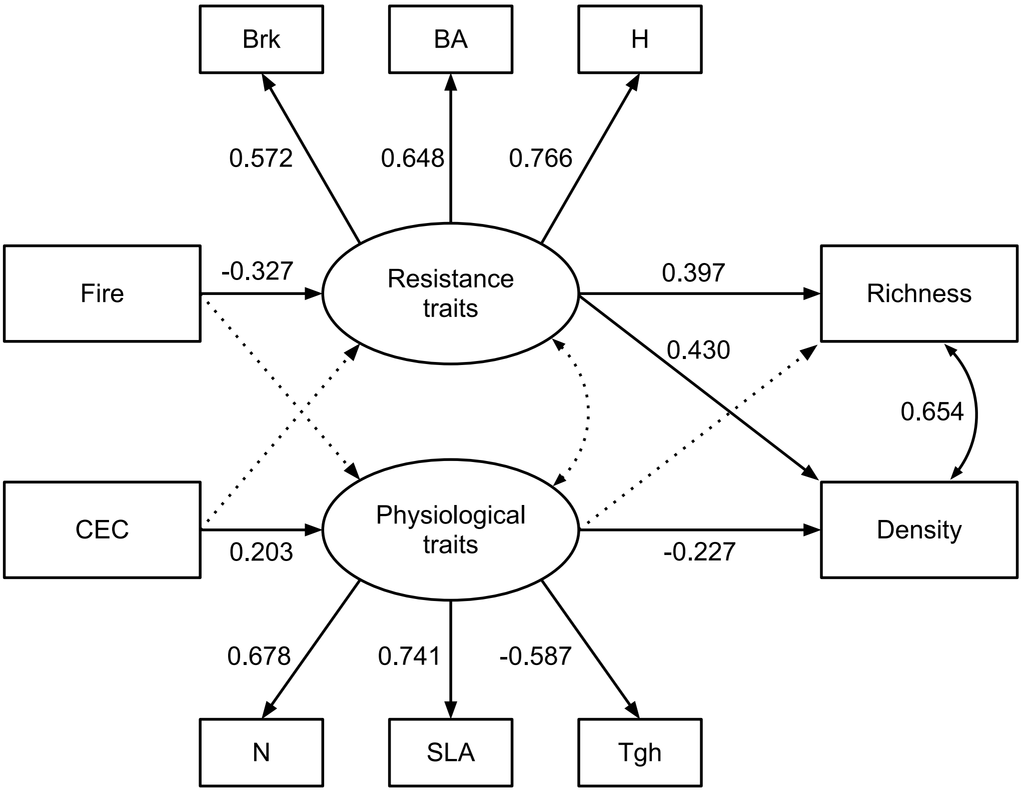
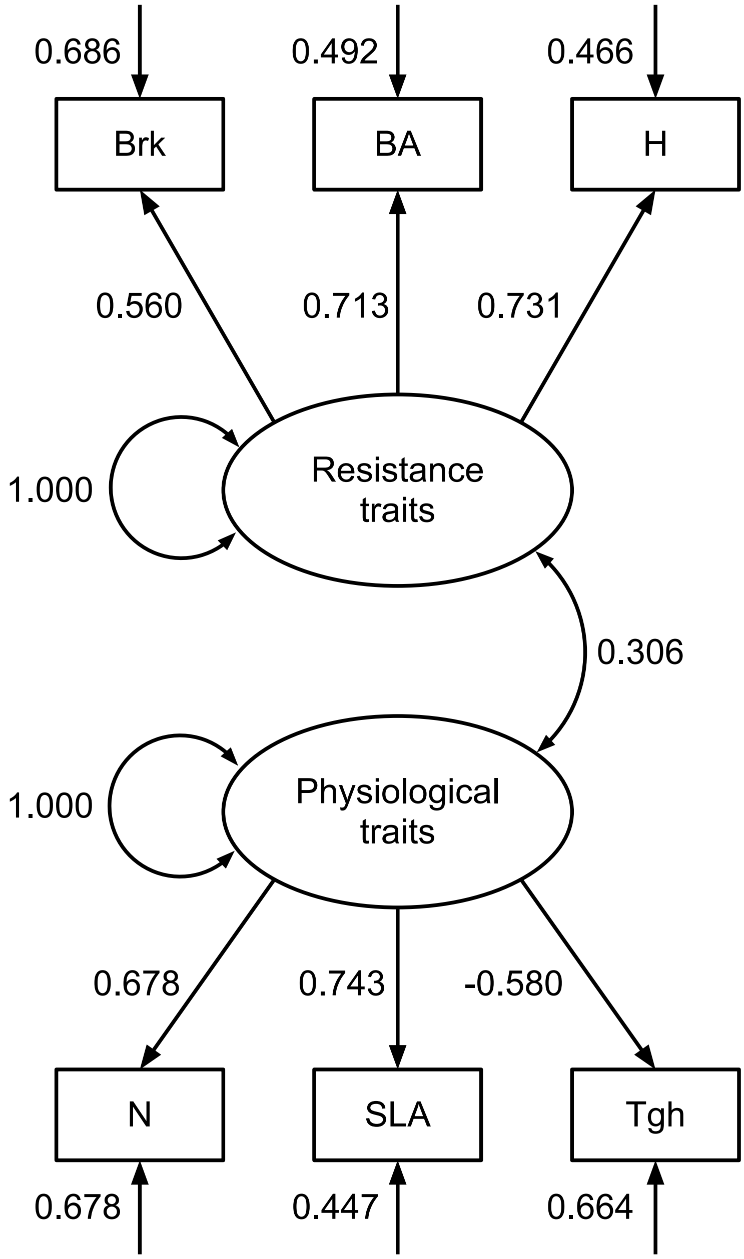

Brk
BA
H
0.572
0.648
0.766
Fire
-0.327
Resistance traits
0.397
Richness
0.430
0.654
CEC
0.203
Physiological traits
-0.227
Density
0.678
0.741
-0.587
N
SLA
Tgh

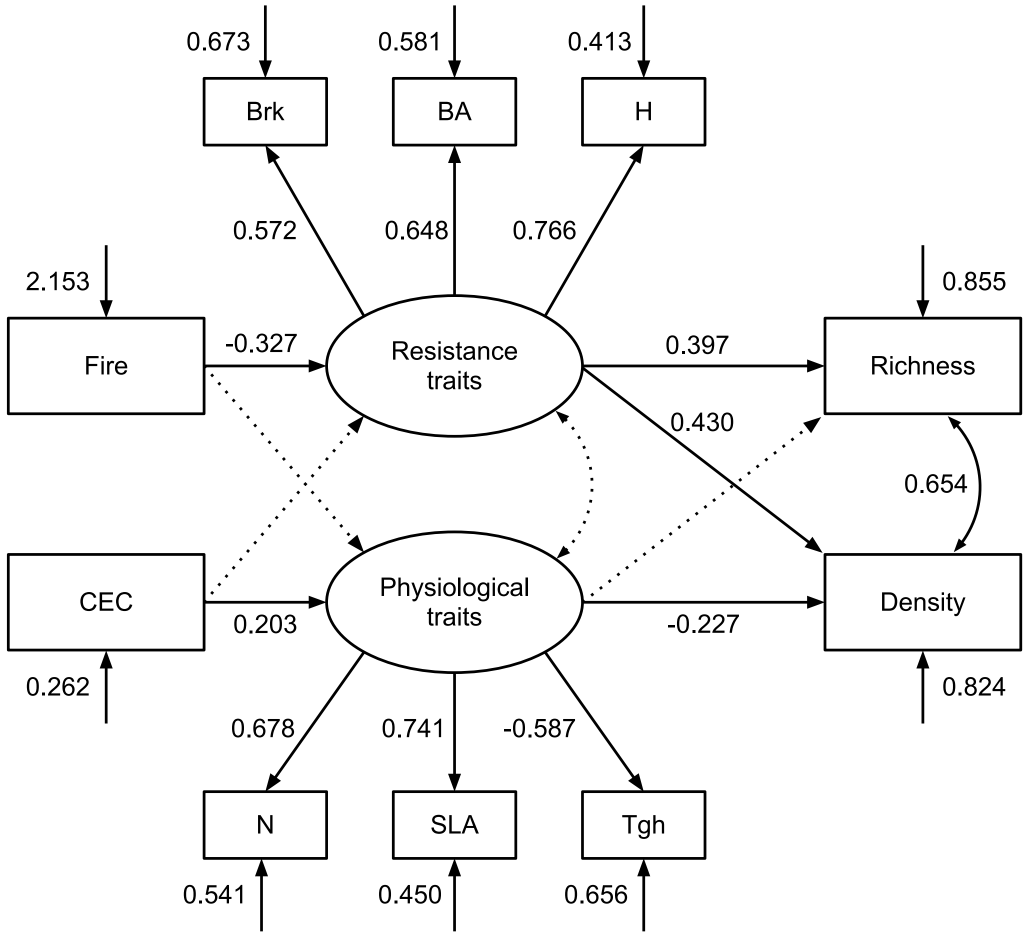
Figure 1. Competing causal models relating disturbance, traits, richness, and density of woody individuals in a Neotropical savanna. In the first model (a), we consider all measured traits to be caused by a unique latent variable, whereas in the second model (b), we incorporate trait multidimensionality by further dividing traits into two latent variables. Fire = number of fire events between 1984 and 2010. CEC = cation exchange capacity (mmol kg^{-1}), Brk = bark thickness (mm), BA = basal area (m^2), H = height (m), N = leaf nitrogen content (mg g^{-1}), SLA = specific leaf area (mm^2 mg^{-1}), Tgh = leaf toughness (N).

Figure 2. Standardised estimates of the model with best fit with data ($\chi^2 = 38.284$, $df = 28$, $p = 0.093$). Solid arrows indicate significant paths. Dotted arrows indicate paths not significantly different from 0. See the electronic supplementary material for error estimates and robust standard errors and Fig. 1 for details on the variables.

Appendix 1. Confirmatory factor analysis with standardised estimates of the model with traits separated into two latent variables ($\chi^2 = 7.304$, $df = 8$, $p = 0.50$). All estimates were significant. Brk = bark thickness (mm), BA = basal area (m^2), H = height (m), N = leaf nitrogen content (mg g^{-1}), SLA = specific leaf area (mm^2 mg^{-1}), Tgh = leaf toughness (N).



Appendix 2. Standardised estimates of the model with best fit with data ($\chi^2 = 38.284$, $df = 28$, $p = 0.093$). Solid arrows indicate significant paths. Dotted arrows indicate paths not significantly different from 0. Fire = number of fire events between 1984 and 2010. CEC = cation exchange capacity (mmol kg^{-1}), Brk = bark thickness (mm), BA = basal area (m^2), H = height (m), N = leaf nitrogen content (mg g^{-1}), SLA = specific leaf area (mm^2 mg^{-1}), Tgh = leaf toughness (N). We estimated the errors of endogenous variables and fixed the errors of fire and CEC to their observed standard deviations.



VI - CONCLUSÃO GERAL

CONCLUSÃO GERAL

Neste trabalho, nós mostramos que diversos aspectos do ambiente tiveram grande influência na diversidade funcional de espécies arbóreas de cerrado e no funcionamento da comunidade, tanto por meio de caminhos causais diretos quanto indiretos. Vimos também que o efeito do fogo na diversidade funcional e no funcionamento foi indireto, via seus efeitos nas características do solo. A fertilidade do solo e o alumínio também foram fatores importantes para a formação e funcionamento das comunidades no Parque Nacional das Emas. Portanto, pudemos identificar e quantificar as formas com que o ambiente interage com a biodiversidade e como a relação entre esses dois componentes influencia o funcionamento da comunidade.

Ainda, conseguimos prever as abundâncias locais das espécies utilizando frequências observadas em escalas maiores e alguns traços. Concluimos que os traços possuem importância significativa na formação das comunidades locais, mas que a maioria das espécies é filtrada na escala regional. Mostramos também que diferentes fatores ambientais agem de forma distinta sobre diferentes conjuntos de traços ambientais. O fogo influencia traços de resistência, enquanto fertilidade do solo influencia traços fisiológicos. Assim, maior frequência de queimadas leva à áreas com menor riqueza e densidade de indivíduos lenhosos, enquanto solos mais férteis levam à comunidades com menor densidade de indivíduos lenhosos.