

ORIGINAL RESEARCH

Mapping woody plant community turnover with space-borne hyperspectral data – a case study in the Cerrado

Pedro J. Leitão^{1,2} , Marcel Schwieder¹ , Fernando Pedroni³, Maryland Sanchez³, José R. R. Pinto⁴ , Leandro Maracahipes⁵ , Mercedes Bustamante⁶  & Patrick Hostert^{1,7} 

¹Geography Department, Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

²Department Landscape Ecology and Environmental System Analysis, Technische Universität Braunschweig, Langer Kamp 19c, 38106 Braunschweig, Germany

³Instituto de Ciências Biológicas e da Saúde, Universidade Federal de Mato Grosso, Pontal do Araguaia MT, 78698000, Brazil

⁴Departamento de Engenharia Florestal, Universidade de Brasília, Brasília 70919-970, Brazil

⁵Instituto de Pesquisa Ambiental da Amazônia, Canarana MT, 78640-000, Brazil

⁶Departamento de Ecologia, Universidade de Brasília, Brasília DF, 70919-970, Brazil

⁷Integrated Research Institute on Transformations of Human-Environment Systems (IRI THESys), Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

Keywords

Beta diversity, community modelling, GDM, hyperspectral remote sensing, SGDM, species turnover

Correspondence

Pedro J. Leitão, Geography Department, Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany.
Tel: +49-531-3915945;
Fax: +49-531-3918170;
E-mails: p.leitao@geo.hu-berlin.de;
p.leitao@tu-bs.de

Funding Information

This study is part of the research activities of the EnMAP Scientific Advisory Group (EnSAG), and was funded by the German Aerospace Centre (DLR) – Project Management Agency, granted by the Ministry of Economics and Technology (BMW grant 50EE1309). The field data were collected within the project CNPq 457497/2012-2 and Edital Sisbiota – CNPq n° 47/2010 (CNPq 563134/2010-0, Projeto Diversidade biológica do Cerrado: estrutura e padrões). J.R.R.Pinto benefitted from a Productivity fellowship (CNPq 307701/2014-0)

Editor: Harini Nagendra

Associate Editor: Ned Horning

Received: 15 February 2018; Revised: 8 June 2018; Accepted: 6 July 2018

doi: 10.1002/rse2.91

Remote Sensing in Ecology and Conservation 2019; **5** (1):107–115

Abstract

Effective conservation measures require the knowledge on the spatial patterns of species communities and their turnover. This knowledge is, however, many times lacking, particularly so for complex systems. On the other hand, recent developments have resulted in tools that enable the mapping of these patterns from remote sensing data, such as Sparse Generalized Dissimilarity Modelling (SGDM). SGDM is a two-stage approach, which combines a Sparse Canonical Component analysis and a Generalized Dissimilarity Modelling (GDM), thus being designed to deal with high-dimensional data to predict community turnover in GDM. In this study, we use space-borne hyperspectral data to map woody plant community patterns collected in two study sites in the Cerrado (Brazilian savannah), namely, the Parque Estadual da Serra Azul (PESA) in Mato Grosso state and Parque Nacional da Chapada dos Veadeiros (PNCV) in Goiás state. Field data were collected in both study sites, following a systematic sampling scheme adapted for the Cerrado. The Cerrado is the most diverse of all the world's savannahs, and while holding a high diversity and endemism of species, this biome is mostly unprotected and understudied. We used Hyperion data acquired over the two study sites, which were subject to data pre-processing (including radiometric and geometric corrections, as well as correction for sensor errors) and quality screening before analysis. Our models were used to map woody plant community patterns and turnover for the study areas. We also inspected the Hyperion spectral bands which most contributed in the SGDM, for each site. Furthermore, the modelled patterns were interpreted with respect to the ecological characteristics of the respective species, this way further enhancing our understanding of this complex system. This study has demonstrated that this approach is suitable for mapping woody plant communities in heterogeneous systems, based on combined field and space-borne hyperspectral data.

Introduction

Effective conservation measures require a deeper knowledge about the spatial patterns of the biotic communities (Gaston 2000). Beta diversity, which can be interpreted as the turnover of species (Legendre et al. 2005), is particularly important as it allows to predict changes in overall species richness (gamma diversity) based on changes in local species assemblages (alpha diversity), as well as provides useful insights on species interactions and community dynamics (Cornell and Lawton 1992; Socolar et al. 2016). Despite their importance, spatial patterns of beta diversity are mostly unknown, and few studies analyse these patterns over large extents (McKnight et al. 2007). On the other hand, the use of remote sensing data together with field-collected compositional data has great potential for directly mapping species assemblages and turnover at larger extents (Turner et al. 2003; Rocchini et al. 2018). Indeed, advances in remote sensing technology, such as the advent of forthcoming space-borne hyperspectral systems, will enable the detailed characterisation of complex natural systems with great potential for ecological applications (Leitão et al. 2015a).

In this study we use space-borne hyperspectral data for mapping woody plant community patterns based on the turnover of shrub-tree species in two study sites in the Brazilian savannah biome, commonly known as the Cerrado. The knowledge of these patterns will increase the understanding of the ecology of this complex system with high species richness, thus supporting efficient and sustainable conservation actions.

To this purpose we use a Sparse Generalized Dissimilarity Modelling (SGDM) approach to fit the field-collected woody plant inventory data to the hyperspectral data (Leitão et al. 2015b). The SGDM is particularly suitable for our study, as it was designed for modelling (and mapping) community turnover with high-dimensional remote sensing data. For getting a better understanding of our models, we inspect the contributions of the individual spectral bands, as well as the canonical scores of each individual species, for each site. Finally, we illustrate the interpretation of the main axes of turnover in the modelled communities in terms of their ecological characteristics.

The overall aim of this study is to demonstrate the usage of space-borne hyperspectral data in a SGDM for mapping woody plant community turnover patterns in heterogeneous landscapes.

Materials and Methods

Study area

The Cerrado covers ca. 2 million km² and is the most biodiversity rich of all the world's savannahs (Myers et al.

2000). Indeed, it is estimated to hold ca. 12,000 species of vascular plants (Mendonça et al. 2008), of which almost half are endemic (Silva and Bates 2002). Despite its richness, this highly complex system is mostly understudied and with little legal protection area and limited conservation incentives (Françoso et al. 2015; Strassburg et al. 2017). The Cerrado biome constitutes a global biodiversity hotspot, as it is under great pressure for conversion into agriculture, with almost half of its area already converted (Sano et al. 2010). Also, up to 40% of the remaining natural vegetation can be legally converted according to Brazil's Forest Code (Soares-Filho et al. 2014), and projections of land use change indicate the extinction of about 500 endemic plant species (Ferreira et al. 2013; Lahsen et al. 2016).

Our study was focused on two sites of natural Cerrado vegetation, and both within protected areas: Parque Estadual da Serra Azul (PESA), in Mato Grosso state; and Parque Nacional da Chapada dos Veadeiros (PNCV), in Goiás state (Fig. 1). These two sites represent only a small fraction of the great variability of the Cerrado biome (Felfili et al. 2004; Mendonça et al. 2008), and serve as demonstration sites to test our methodology.

Data collection and processing

Woody plant inventory data were collected in the field, following a sampling protocol based on the PPBio programme, adapted for the Cerrado (Magnusson et al. 2005; Schwieder et al. 2018). This protocol consists on a grid of 10 plots systematically distributed along two parallel 5-km tracks, 1 km apart. Each plot consisted of a 250 m × 40 m sampling area (the longer dimension following the contour of the terrain), divided into subsections (sampling units) of varied size (10 m long and 5–20 m wide). Within each sampling unit, every woody plant with a minimum trunk diameter of 10 cm, collected at 30 cm above the ground, was identified at the species level. The field campaigns were conducted between February 2012 (one single plot) and May 2014 in PESA and between August 2014 and May 2015 in PNCV. By overlaying the sampling units with the co-registered remote sensing image pixels, we assigned to each pixel all species that potentially occurred on it. Whenever a sampling unit (partially) covered more than one pixel, the respective species list was assigned to all overlaying pixels. In order to minimize undesired community differences due to undersampling, only pixels with at least 75% sample cover were considered. This was done for both study sites for all available plot samples (8 for PESA and 6 for PNCV), which resulted in a total of 65 observations (185 species) for PESA and 29 observations (94 species) for PNCV, totalling 241 species overall.

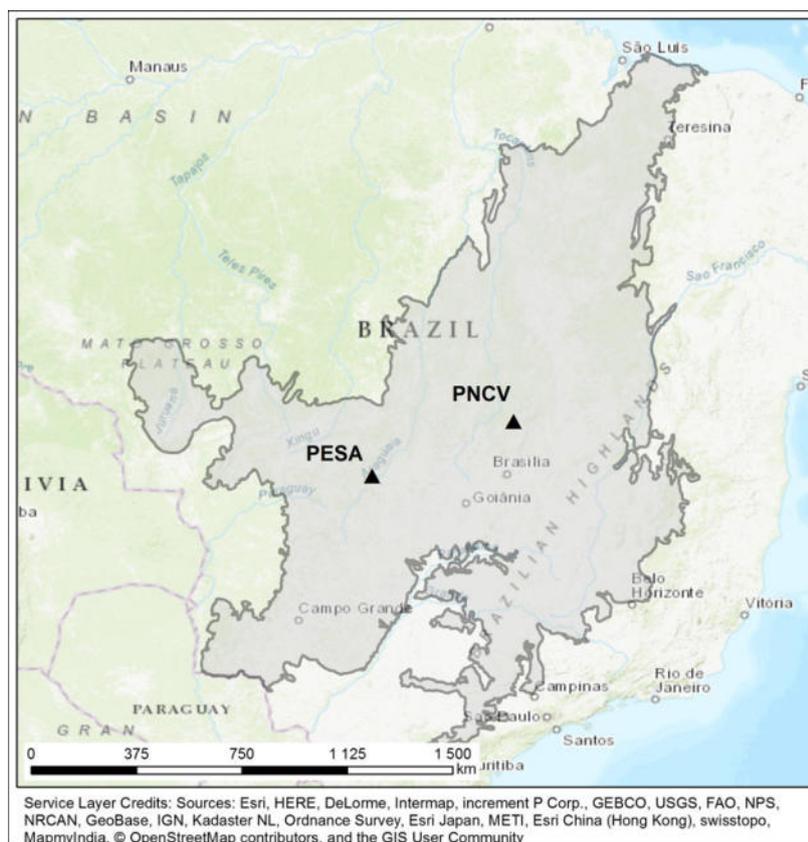


Figure 1. Distribution of Cerrado in Brazil (in grey) and the locations of the two study sites: PESA, Parque Estadual da Serra Azul, Mato Grosso state; and PNCV, Parque Nacional da Chapada dos Veadeiros, Goiás state.

We used space-borne hyperspectral data collected from the Hyperion sensor on board the Earth Observing-1 platform. These data were collected during the dry season, as during this period, there is little to no photosynthetically active herbaceous vegetation (Ferreira et al. 2003), and therefore the vegetation spectral signal is predominantly coming from woody vegetation thus optimising its discrimination (Sano et al. 2010). In order to match the dates of the field campaigns, the Hyperion data used were, respectively, acquired on the 27th June 2014 and on the 27th of July 2015 for both the PESA and PNCV study sites. The Hyperion data were radiometrically corrected, including correction for pixel shifts, striping, keystone and smile, and atmospheric effects (Leitão et al., in press; Rogass et al. 2014). Data from the two individual sensor detectors (visible and near infrared, and shortwave infrared) were co-registered using precision terrain-corrected (L1T) Landsat Operational Land Imager (OLI) scenes for spatial consistency. Erroneous or noisy spectral bands were interactively screened and excluded, resulting in a total of 81 spectral bands per image, covering the visible, near and shortwave infrared portions of the

electromagnetic spectrum. The final band stacks were spectrally smoothed with a Savitzky-Golay filter, with a filter width of 3 in combination with a first-order polynomial (Savitzky and Golay 1964; Miglani et al. 2011). The spectral smoothing was performed in the EnMAP-Box software (van der Linden et al. 2015).

Woody plant community modelling and mapping

In this study, we used a SGDM approach for modelling and mapping the woody plant communities in our study sites (Leitão et al. 2015b). This method is a two-step approach, consisting of initially reducing the environmental space (in this case, the Hyperion data) by means of a sparse canonical correlation analysis (SCCA; Witten et al. 2013), and then fitting the resulting components with Generalized Dissimilarity Modelling (GDM; Ferrier et al. 2007). The SCCA is a form of penalized canonical correlation analysis based on the Lasso regularisation, and is thus designed to deal with high-dimensional data, such as hyperspectral data. The

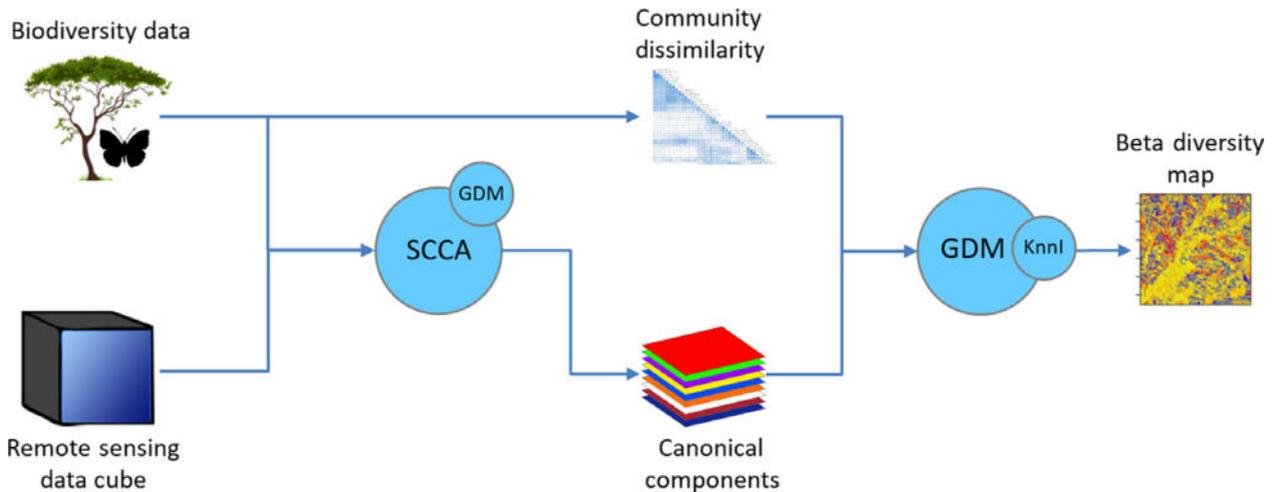


Figure 2. The SGDM workflow is designed to be used with high-dimensional data (such as remote sensing data cubes), for mapping beta diversity: The Sparse Canonical Component Analysis (SCCA) coupled with the Generalized Dissimilarity Model (GDM) reduce the dimension of these data (in canonical components) while keeping their relation with the biodiversity data; the final GDM followed by the knn-imputation (Knnl) allows predicting the main axis of community variation throughout the study area. SGDM, Sparse Generalized Dissimilarity Modelling.

GDM is an established method for modelling species turnover, by using monotonic I-splines to fit dissimilarities between pairs of community data samples with environmental predictors (Ferrier et al. 2007). This way, SGDM is highly suited for modelling beta diversity (community turnover) based on high-dimensional remote sensing data (Fig. 2; Leitão et al. 2015b). The SGDM procedure includes the parameterization of the SCCA (based on the optimization of a GDM model, in a fivefold cross-validation), the reduction (canonical transformation) in the hyperspectral data, model reduction (i.e. exclusion of non-significant components) and fitting of the reduced data (significant sparse components) with the final GDM.

Subsequently, and following the approach suggested by Leitão et al. (2015b), a non-metric dimensional scaling (NMDS) was applied on the GDM model predictions (dissimilarities) in order to extract the main axes of variation in the turnover of the modelled communities. NMDS is a non-parametric ordination approach designed to analyse dissimilarity matrices. The number of NMDS axes to extract was determined automatic, based on the respective stress values (Clarke 1993). The result of the NMDS analysis was extrapolated to the whole study area by means of a knn-imputation (Leitão et al. 2017), an nearest-neighbour imputation approach designed for dealing with missing values and which has been previously used for estimating the floristic ordination scores based on spectrally nearest neighbours (Thessler et al. 2005). This procedure resulted in a map of the spatial turnover (beta diversity) of woody plants along the study sites.

Table 1. Fifteen most important Hyperion spectral band variables (average wavelengths in nm), and their absolute contributions in the SGDM models, for both study sites (PESA – Parque Estadual da Serra Azul, Mato Grosso state; and PNCV – Parque Nacional da Chapada dos Veadeiros, Goiás state). For interpretation, the spectral bands were grouped into the following spectral regions: blue (B: 450–520 nm); green (G: 520–600 nm); red (R: 600–690 nm); red edge (RE: 690–730 nm); near infrared (NIR: 730–1400 nm) and shortwave infrared (SWIR: 1400–2600 nm).

PESA		PNCV	
Spectral band	Relative contribution	Spectral band	Relative contribution
620.15 (R)	4.229	569.27 (G)	4.039
701.55 (RE)	4.137	1517.83 (SWIR)	3.377
630.32 (R)	3.804	1507.73 (SWIR)	3.237
599.80 (G)	3.679	609.97 (R)	3.138
640.50 (R)	3.637	477.69 (B)	2.979
609.97 (G)	3.498	1568.22 (SWIR)	2.911
1285.76 (NIR)	3.442	1598.51 (SWIR)	2.821
732.07 (RE)	3.107	1235.27 (NIR)	2.711
650.67 (R)	2.939	579.45 (G)	2.664
589.62 (G)	2.642	467.52 (B)	2.629
1719.60 (SWIR)	2.544	1225.17 (NIR)	2.615
518.39 (G)	2.529	1003.30 (NIR)	2.570
498.04 (B)	2.503	1527.92 (SWIR)	2.509
579.45 (G)	2.376	589.62 (G)	2.042
1537.92 (SWIR)	2.343	874.53 (NIR)	1.928

The performance of the SGDM models was assessed by means of the root mean square error (RMSE) and the coefficient of determination (r^2), following a 10-fold cross-validation. The final model variable contributions

were also extracted, based on the loss in model deviance explained when dropping each variable at a time. By multiplying the variable (component) contributions of the GDM model, with the absolute value of the respective canonical vectors used to derive them, in the SCCA, from the original spectral bands, it is possible to derive the relative model contribution of each spectral band in the final model. For easier interpretation, these values were scaled so that their sum corresponds to 100. Finally, the individual species scores for each sparse component can be calculated by multiplying the species matrix with the one with the component scores (output of the NMDS). All analyses were done in R, using code from the *sgdm* package (Leitão et al. 2017; R Development Core Team, 2017).

The ecological and biogeographical characteristics of the species can be used to better interpret the modelled patterns. In this case, and as an illustrative example, we considered the biome of origin of each species, their phytogeographic domain (biomes where they occur), typical formation (forest or savannah) and fire resistance

properties as based on their bark thickness (Hoffmann et al. 2012; Flora do Brasil 2020).

Results

The SGDM models were able to model the species turnover in the woody plant communities for both study sites. The model performances consisted in a RMSE of 0.09762 and 0.15255, and a R^2 of 70.972 and 39.193, for the PESA and PNCV study sites respectively.

These models included 8 and 7 significant sparse components, respectively, for the PESA and PNCV study sites. By inspecting the 15 most important spectral band variables for each model (Table 1), we could observe that they used information from all spectral regions, ranging from the visible to the microwave infrared. In the PESA model, there is a certain predominance (10 of 15) of spectral bands within the visible range (blue to red), whereas in the PNCV model, the infrared showed to be more important to characterise the observed turnover patterns (9 of 15 bands). Also, information extracted

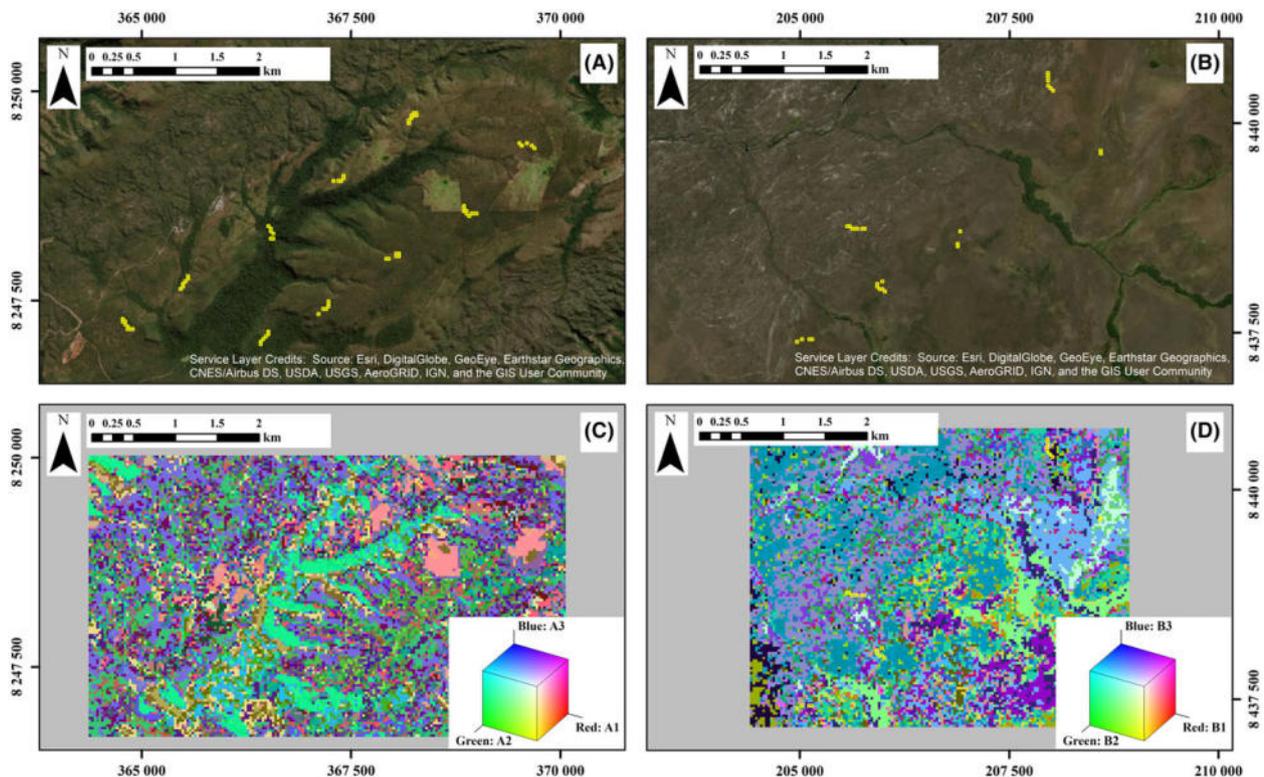


Figure 3. Above (in yellow) the location of the sampled pixels with a sample cover of at least 75% in PESA, Parque Estadual da Serra Azul (A) and PNCV, Parque Nacional da Chapada dos Veadeiros (B). Below, RGB maps of the main axes of variation in species turnover in PESA (C) and PNCV (D) woody plant communities. Each of the first three NMDS axes (A1 to A3 for PESA and B1 to B3 for PNCV) were, respectively, plotted in the R, G and B channels, so that, for example, a bright turquoise pixel would mean high loadings on axes 2 and 3 (high values on both the G and B channels).

from adjacent spectral bands (e.g. 620.15 nm and 630.32 nm for PESA or 1507.73 and 1517.83 for PNCV) showed to be important in both models.

The NMDS analysis resulted of 11 axes for PESA and 8 for PNCV. The patterns of variation in species spatial turnover in the study sites can be visualized, for example, by plotting the first 3 NMDS axes in RGB (Fig. 3).

The careful inspection of the species loadings in the NMDS axes, together with information on the species ecological and biogeographical characteristics, allow us to interpret the modelled patterns (Fig. 4). The first NMDS axis extracted from the PESA model depicts the turnover between typical Savannah species from those which occur

both in Forests and Savannah formations, whereas the second axis shows the turnover between species originated from the Amazon biome and non-fire resistant from those original from the Cerrado and resistant to fire. In the PNCV model, on the other hand, the first and second axes show the turnover between species which occur both in the Cerrado and the Caatinga biomes from those which occur in the Cerrado only.

Discussion

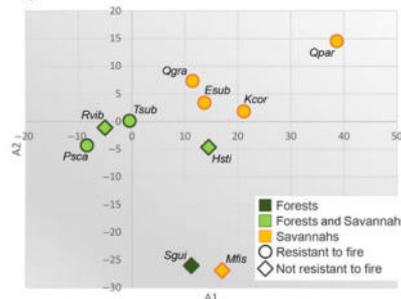
The SGDM models used information contained in every single spectral band, although reduced in few significant

Main axes of woody plant community turnover, their representative species and respective ecological and biogeographical characteristics, for both study sites

- Parque Estadual da Serra Azul (PESA):

Species	BIO	Dom	For	FR	A1	A2	A3
<i>Erythroxylum suberosum</i> (Esub)	C	AC	S	R	13.70	3.40	11.54
<i>Hymenaea stigonocarpa</i> (Hstig)	C	ACMPT	FS	N	14.47	-4.61	-13.77
<i>Kielmeyera coriacea</i> (Kcor)	C	AC	S	R	21.17	1.83	-8.51
<i>Mabea fistulifera</i> (Mfis)	A	ACMT	S	N	17.02	-26.87	-8.60
<i>Physocalymma scaberrimum</i> (Pscab)	C	AC	FS	R	-8.44	-4.31	-8.55
<i>Qualea grandiflora</i> (Qgra)	C	ACMT	S	R	11.49	7.33	-1.52
<i>Qualea parviflora</i> (Qpar)	C	ACMT	S	R	38.69	14.52	-15.58
<i>Rudgea viburnoides</i> (Rvib)	C	ACT	FS	N	-5.08	-1.08	0.96
<i>Sloanea guianensis</i> (Sgui)	A	ACM	F	N	11.20	-25.99	9.88
<i>Tachigali subulatina</i> (Tsub)	C	C	FS	R	-0.41	0.13	-1.22

Two highest and lowest scoring species for each of the first 3 NDMS axes (A1 to A3), and their characteristics



Scaterplot of the first two NMDS axes, including the formation on which the species occur and their fire resistance properties

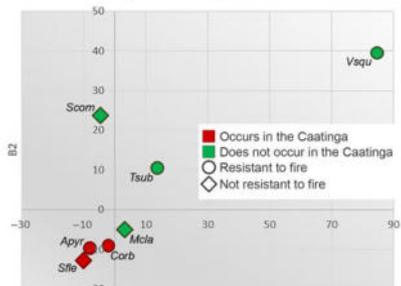


© Leandro Maracahipes
Qualea parviflora has a thick bark as an adaptation to fire

- Parque Nacional da Chapada dos Veadeiros (PNCV):

Species	BIO	Dom	For	FR	B1	B2	B3
<i>Aspidosperma pyriforme</i> (Apyr)	T	CT	S	R	-7.77	-9.85	-9.04
<i>Chamaecrista orbiculata</i> (Corb)	C	CT	S	R	-1.90	-9.22	-73.16
<i>Mimosa catantoni</i> (Mcat)	C	C	S	N	3.14	-4.91	9.06
<i>Syagrus comosa</i> (Scom)	C	AC	S	N	-4.76	23.74	-12.70
<i>Syagrus flexuosa</i> (Sflx)	C	CT	FS	N	-10.11	-12.81	-22.30
<i>Tachigali subulatina</i> (Tsub)	C	C	FS	R	13.98	10.35	10.82
<i>Vellutia squamata</i> (Vsqm)	C	C	S	R	85.02	39.41	-95.71

Two highest and lowest scoring species for each of the first 3 NDMS axes (B1 to B3), and their characteristics



Scaterplot of the first two NMDS axes, including the phytogeographic domain of each species and their fire resistance properties



© Fernando Pedroni
Syagrus flexuosa occurs in the dry Caatinga

Figure 4. Illustrative example of the NMDS scores for individual species, and their ecological and biogeographical characteristics. The two highest and lowest scoring species for each of the first three axes for both PESA, Parque Estadual da Serra Azul (A1 to A3, above left) and PNCV, Parque Nacional da Chapada dos Veadeiros (B1 to B3, below left) models are presented (respective values in bold). In the middle, the two-first axes are plotted (PESA above and PNCV below), including labels for some of the species characteristics. On the right, two species are illustrated as examples. Legend: BIO, Biome of Origin; Dom, Phytogeographic Domain; For, Formation; FR, Fire resistance; A, Amazon; C, Cerrado; M, Atlantic Forest; P, Pantanal; T, Caatinga; F, Forest; S, Savannah; R, Resistant; N, Non-resistant.

components, which agrees with Leitão et al. (2015b). We found, however, differences in the models between the two study sites. First, the model performance for PESA was much higher than that for PNCV, which could be due to several factors. First, the larger sample size in PESA allows the extraction of a larger number of sparse components, thus optimizing the information extraction capability of the SGDM model (Leitão et al. 2015b). Second, this study site has a greater variability of woody plant communities than PNCV, including the transitions from savannah to forest and thus a greater species richness, also beneficial for the discriminating power of the models. Also, the presence of rock outcrops in PNCV might serve as noise in the species spectral discrimination. And finally, it might be that the species present in PESA are spectrally more distinct from those in PNCV. Also, the differences in the number of NMDS axes extracted are most probably due to the greater complexity (higher species richness) of PESA.

Despite the model differences between sites, the achieved model performances are high for both sites, for example, when comparing to the performances achieved by Baldeck and Asner (2013). In this study, the authors use an unsupervised clustering approach to map beta diversity of woody plants in an African Savannah with airborne hyperspectral imagery (Baldeck and Asner 2013). Considering the high model performances achieved, we consider that our approach was successful in modelling the woody plant species turnover in our study sites. Our results also allowed the interpretation of the turnover in terms of some ecological and biogeographical characteristics of the species, such as biome of origin, phytogeographic domain, formation habitat and fire resistance.

It is worth noting that the herbaceous and small shrub vegetation, although also contributing to the surface reflectance as captured by the satellite imagery, was not considered in this. This is due to the lack of field data at a sufficient spatial extent to be representatively allocated to the image pixel size. Our study showed, however, that it is possible to map the species turnover of the woody layer based on the surface reflectance values, in such complex landscapes.

The use of the RGB colour space for mapping species communities, by plotting their axes' loadings, is not new (Schmidtlein et al. 2007; Féret and Asner 2014), although not widely used. Although this requires a good understanding of the colour space (additive colour space defined by the three red, green and blue additive primary colours), using the RGB colour space for mapping community distributions and turnover has great potential as it allows the visualisation of three ordination axes at a time, thus allowing to easily interpret the spatial patterns of species community variation.

Our results further emphasize the potential of space-borne hyperspectral data for mapping species turnover in heterogeneous landscapes. It is thus expected that forthcoming hyperspectral missions, such as EnMAP (Guanter et al. 2015) or HypSIIRI (Lee et al. 2015), will have an important role in monitoring complex natural ecosystems and this way provide further insights into the ecology and functioning of these systems (Jetz et al. 2016; Pettorelli et al. 2018).

Finally, this study is a further step into the understanding of the ecology of the Cerrado and enables follow-up studies, for example, focused on the patterns of species traits, their interactions, trade-offs and trait diversity (Batalha et al. 2011; Schneider et al. 2017). Indeed, only by better understanding these patterns and processes, it is possible to implement efficient conservation measures for the long-term sustainability of this threatened biodiversity hotspot (Klink and Machado 2005).

Acknowledgments

This study is part of the research activities of the EnMAP Scientific Advisory Group (EnSAG), and was funded by the German Aerospace Centre (DLR) – Project Management Agency, granted by the Ministry of Economics and Technology (BMW grant 50EE1309). The field data were collected within the project CNPq 457497/2012-2 and Edital Sisbiota – CNPq n° 47/2010 (CNPq 563134/2010-0, Projeto Diversidade biológica do Cerrado: estrutura e padrões). J.R.R.Pinto benefitted from a Productivity fellowship (CNPq 307701/2014-0). We are grateful to Christian Rogass for pre-processing the Hyperion data. Finally, we are also thankful to Prof. Dr. Duccio Rocchini and a further anonymous referee for the useful comments during the review process, which further improved the manuscript.

References

- Baldeck, C. A., and G. P. Asner. 2013. Estimating vegetation beta diversity from airborne imaging spectroscopy and unsupervised clustering. *Remote Sens.* **5**, 2057–2071.
- Batalha, M. A., I. A. Silva, M. V. Cianciaruso, and G. H. de Carvalho. 2011. Trait diversity on the phylogeny of cerrado woody species. *Oikos* **120**, 1741–1751.
- Brasil, F.D. (2020) under construction. In Jardim Botânico do Rio de Janeiro. Available at: <http://floradobrasil.jbrj.gov.br/>.
- Clarke, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* **18**, 117–143.
- Cornell, H. V., and J. H. Lawton. 1992. Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. *J. Anim. Ecol.* **61**, 1–12.

- Felfili, J. M., M. C. D. Junior, A. C. Sevilha, C. W. Fagg, B. M. T. Walter, P. E. Nogueira, et al. 2004. Diversity, floristic and structural patterns of cerrado vegetation in Central Brazil. *Plant Ecol.* **175**, 37–46.
- Féret, J.-B., and G. P. Asner. 2014. Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecol. Appl.* **24**, 1289–1296.
- Ferreira, L. G., H. Yoshioka, A. Huete, and E. E. Sano. 2003. Seasonal landscape and spectral vegetation index dynamics in the Brazilian Cerrado: an analysis within the Large-Scale Biosphere-Atmosphere Experiment in Amazonia (LBA). *Remote Sens. Environ.* **87**, 534–550.
- Ferreira, M. E., L. G. Ferreira, F. Miziara, and B. S. Soares-Filho. 2013. Modeling landscape dynamics in the central Brazilian savanna biome: future scenarios and perspectives for conservation. *J. Land Use Sci.* **8**, 403–421.
- Ferrier, S., G. Manion, J. Elith, and K. Richardson. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* **13**, 252–264.
- Françoso, R. D., R. Brandão, C. C. Nogueira, Y. B. Salmona, R. B. Machado, and G. R. Colli. 2015. Habitat loss and the effectiveness of protected areas in the Cerrado Biodiversity Hotspot. *Nat. Conserv.* **13**, 35–40.
- Gaston, K. J. 2000. Global patterns in biodiversity. *Nature* **405**, 220–227.
- Guanter, L., H. Kaufmann, K. Segl, S. Chabrillat, S. Förster, C. Rogass, et al. 2015. The EnMAP spaceborne imaging spectroscopy mission for Earth Observation. *Remote Sens.* **7**, 8830–8857.
- Hoffmann, W. A., E. L. Geiger, S. G. Gotsch, D. R. Rossatto, L. C. R. Silva, O. L. Lau, et al. 2012. Ecological thresholds at the savanna-forest boundary: how plant traits, resources and fire govern the distribution of tropical biomes. *Ecol. Lett.* **15**, 759–768.
- Jetz, W., J. Cavender-Bares, R. Pavlick, D. Schimel, F. W. Davis, G. P. Asner, et al. 2016. Monitoring plant functional diversity from space. *Nat. Plants* **2**. Article number 16024.
- Klink, C. A., and R. B. Machado. 2005. Conservation of the Brazilian Cerrado. *Conserv. Biol.* **19**, 707–713.
- Lahsen, M., M. M. C. Bustamante, and E. L. Dalla-Nora. 2016. Undervaluing and overexploiting the Brazilian Cerrado at our Peril. *Environment* **58**, 4–15.
- Lee, C. M., M. L. Cable, S. J. Hook, R. O. Green, S. L. Ustin, D. J. Mandl, et al. 2015. An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sens. Environ.* **167**, 6–19.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecol. Monogr.* **75**, 435–450.
- Leitão, P. J., M. Schwieder, F. Pötzschner, J. R. R. Pinto, A. Teixeira, F. Pedroni, et al. In press. From sample to pixel: using high-resolution remote sensing data for the spatial allocation and extrapolation of field collected data in heterogeneous landscapes. *Ecosphere* <https://doi.org/10.1002/ecs1002.2298>.
- Leitão, P. J., M. Schwieder, S. Suess, A. Okujeni, L. S. Galvão, S. van der Linden, et al. 2015a. Monitoring natural ecosystem and ecological gradients: perspectives with EnMAP. *Remote Sens.* **7**, 13098–13119.
- Leitão, P. J., M. Schwieder, S. Suess, I. Catry, E. J. Milton, F. Moreira, et al. 2015b. Mapping beta diversity from space: sparse generalised dissimilarity modelling (SGDM) for analysing high-dimensional data. *Methods Ecol. Evol.* **6**, 764–771.
- Leitão, P. J., M. Schwieder, and C. Senf. 2017. sgdm: an R package for performing sparse generalized dissimilarity modelling with tools for gdm. *ISPRS Int. J. Geo-Inform.* **6**, 23.
- van der Linden, S., A. Rabe, M. Held, B. Jakimow, P. J. Leitão, A. Okujeni, et al. 2015. The EnMAP-Box – a toolbox and application programming interface for EnMAP data processing. *Remote Sens.* **7**, 11249–11266.
- Magnusson, W. E., A. P. Lima, R. Luizão, F. Luizão, F. R. C. Costa, C. V. D. Castilho, et al. 2005. RAPELD: a modification of the Gentry method for biodiversity surveys in long-term ecological research sites. *Biota. Neotrop.* **5**, 19–24.
- McKnight, M. W., P. S. White, R. I. McDonald, J. F. Lamoreux, W. Sechrest, R. S. Ridgely, et al. 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. *PLoS Biol.* **5**, 2424–2432.
- Mendonça, R. C., J. M. Felfili, B. M. T. Walter, M. C. Silva Júnior, A. V. Rezende, T. S. Filgueiras, et al. 2008. Flora vascular do bioma cerrado checklist com 12.356 espécies. Pp. 422–442 in S. M. Sano, S. P. Almeida and J. F. Ribeiro, eds. *Cerrado: ecologia e flora*. Embrapa Cerrados, Brasília, DF.
- Migliani, A., S. S. Ray, D. P. Vashishta, and J. S. Parihar. 2011. Comparison of two data smoothing techniques for vegetation spectra derived from EO-1 hyperion. *J. Ind. Soc. Remote Sens.* **39**, 443–453.
- Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. D. Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.
- Pettorelli, N., H. Schulte to Bühne, A. Tulloch, G. Dubois, C. Macinnis-Ng, A. M. Queirós, et al. 2018. Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward. *Remote Sens. Ecol. Conserv.* **4**, 71–93.
- R Development Core Team. 2017. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rocchini, D., S. Luque, N. Pettorelli, L. Bastin, D. Doktor, N. Faedi, et al. 2018. Measuring beta-diversity by remote sensing: a challenge for biodiversity monitoring. *Methods Ecol. Evol.* **9**, 1787–1798.
- Rogass, C., L. Guanter, C. Mielke, D. Scheffler, N. K. Boesche, C. Lubitz, et al. 2014. An automated processing chain for

- the retrieval of georeferenced reflectance data from hyperspectral EO-1 Hyperion acquisitions. Pp. 3.1–3.7 in Z. B. M. Kycko and R. Reuter, eds. *34th EARSeL symposium*. EARSeL, Warsaw, Poland.
- Sano, E. E., R. Rosa, J. L. S. Brito, and L. G. Ferreira. 2010. Land cover mapping of the tropical savanna region in Brazil. *Environ. Monit. Assess.* **166**, 113–124.
- Savitzky, A., and M. J. E. Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639.
- Schmidtlein, S., P. Zimmermann, R. Schüpferling, and C. Weiß. 2007. Mapping the floristic continuum: ordination space position estimated from imaging spectroscopy. *J. Veg. Sci.* **18**, 131–140.
- Schneider, F. D., F. Morsdorf, B. Schmid, O. L. Petchey, A. Hueni, D. S. Schimel, et al. 2017. Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nat. Commun.* **8**, 1441.
- Schwieder, M., P. J. Leitão, J. R. R. Pinto, A. M. C. Teixeira, F. Pedroni, M. Sanchez, et al. 2018. Landsat phenological metrics and their relation to aboveground carbon in the Brazilian Savanna. *Carbon Balance Manage.* **13**, 7.
- Silva, J. M. C., and J. M. Bates. 2002. Biogeographic patterns and conservation in the South American Cerrado: a Tropical Savanna Hotspot The Cerrado, which includes both forest and savanna habitats, is the second largest South American biome, and among the most threatened on the continent. *Bioscience* **52**, 225–234.
- Soares-Filho, B., R. Rajão, M. Macedo, A. Carneiro, W. Costa, M. Coe, et al. 2014. Cracking Brazil's Forest Code. *Science* **344**, 363–364.
- Socolar, J. B., J. J. Gilroy, W. E. Kunin, and D. P. Edwards. 2016. How Should Beta-Diversity Inform Biodiversity Conservation? *Trends Ecol. Evol.* **31**, 67–80.
- Strassburg, B. B. N., T. Brooks, R. Feltran-Barbieri, A. Iribarrem, R. Crouzeilles, R. Loyola, et al. 2017. Moment of truth for the Cerrado hotspot. *Nat. Ecol. Evol.* **1**, 0099.
- Thessler, S., K. Ruokolainen, H. Tuomisto, and E. Tomppo. 2005. Mapping gradual landscape-scale floristic changes in Amazonian primary rain forests by combining ordination and remote sensing. *Glob. Ecol. Biogeogr.* **14**, 315–325.
- Turner, W., S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger. 2003. Remote sensing for biodiversity science and conservation. *Trends Ecol. Evol.* **18**, 306–314.
- Witten, D., R. Tibshirani, S. Gross, and B. Narasimhan. 2013. PMA: penalized multivariate analysis.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Contribution of the individual Hyperion spectral bands (average wavelengths in nm) in the SGDM models for the Parque Estadual da Serra Azul (PESA) study site.

Table S2. Contribution of the individual Hyperion spectral bands (average wavelengths in nm) in the SGDM models for the Parque Nacional da Chapada dos Veadeiros (PNCV) study site.

Table S3. Individual species loadings in each of the NMDS axes for the Parque Estadual da Serra Azul (PESA) study site.

Table S4. Individual species loadings in each of the NMDS axes for the Parque Nacional da Chapada dos Veadeiros (PNCV) study site.

Table S5. R code for mapping species community turnover using sgdm.