SCALING SPECTRAL DIVERSITY TO MAP TROPICAL FOREST BIODIVERSITY: A COMPARATIVE ANALYSIS OF COMPLETE AND RANDOMIZATION-BASED METHODS IN BRAZIL'S SOUTHEASTERN RESERVES

Markus Gastauer[1], Priscila Sanjuan Medeiros de Sarmento[2] e Larissa Ranielle Silva Parente[3]

[1]Instituto Tecnológico Vale, R. Boaventura da Silva, 955, Belém, Brazil, markus.gastauer@itv.org; [2] Instituto Tecnológico Vale, R. Boaventura da Silva, 955, Belém, Brazil, priscila.sanjuan.medeiros@gmail.com e [3] Instituto Tecnológico Vale, R. Boaventura da Silva, 955, Belém, Brazil, larissarsparente@gmail.com

ABSTRACT

This study compares traditional and randomization-based methods to calculate spectral diversity across two conservation units in Southeastern Brazil, Vale's Natural Reserve and Sooretama Biological Reserve. Leveraging Sentinel-2 imagery, we examined spectral community richness and Shannon diversity to assess tropical forest biodiversity. Both methods yielded consistent spatial patterns and high correlations, validating the randomization approach for areas beyond standard computational capacities. This scalable technique offers a robust pathway for monitoring biodiversity across large ecosystems.

Key words **—** Biodiversity indicators, tropical ecosystems, Sentinel-2 imagery, conservation units, spatial analysis.

RESUMO

Este estudo compara métodos tradicionais e baseados em randomização para calcular a diversidade espectral em duas unidades de conservação no sudeste do Brasil: a Reserva Natural da Vale e a Reserva Biológica de Sooretama. Usando imagens Sentinel-2, examinamos a riqueza e a diversidade de Shannon de comunidades espectrais para avaliar a biodiversidade em florestas tropicais. Ambos os métodos apresentaram padrões espaciais consistentes e alta correlação, validando a abordagem de randomização para áreas além da capacidade computacional padrão. Esta técnica escalável oferece um caminho robusto para monitorar a biodiversidade em grandes ecossistemas.

Palavras-chave **—** *Indicadores de biodiversidade, ecossistemas tropicais, imagens Sentinel-2, unidades de conservação, análise espacial.*

1. INTRODUCTION

The hypothesis of spectral diversity suggests that variations in the spectral properties (e.g., the range of wavelengths reflected from vegetation) correlate with biological diversity, particularly in plant communities [1]. In ecosystems like forests, higher spectral diversity, often visible as a greater number of green shades or distinct spectral signatures, could indicate a more diverse plant communities, species variety, or functional diversity [2] This hypothesis leverages remote sensing approaches, which can detect subtle spectral differences, allowing researchers to assess and monitor biodiversity non-invasively in a cost-effective way.

The availability of high-resolution satellite images from the Sentinel mission in combination with computation tools offers promising avenues to asses tropical forest biodiversity [3]. Recent studies revealed the accuracy of this technique to map successional advances in minelands rehabilitation [4], but limitations were recognized, as the approach requires computational capacities increasing disproportionally with the size of the study area.

Objective of this study was (i) to reveal the spectral diversity of two important conservation unites from Southeastern Brazil, Vale's Natural Reserve and Sooretama Biological Reserve from Linhares, Espírito Santo, and (ii) to develop a method to compute spectral diversity for large areas unable to deal with a standard desktop computer using a randomization technique. We compared the outcomes of conventional and randomization technique and present the first spectral diversity map of the study area.

2. MATERIAL E METHODS

Study area was Vale's Natural Reserve and Sooretama Biological Reserve from Linhares, Espirito Santo, Brazil (Fig. 1). Both conservation units form a huge old-growth forest remnant in a region characterized by intense agricultural activities, and provide hundreds of different tree species [5], some of them such as *Dinizia jueirana-facao* (Fabaceae) with spatial distribution restricted to the reserves.

To carry out our study, we downloaded the atmospheric corrected T24KUE and T24KUD Sentinel tiles (L2A) from July 15th, 2019. Single bands and tiles were stacked together in QGIS Desktop 3.32.2 (Lima version) and clipped to the limits of the reserves.

We used the biodivMapR package [3] to compute spectral diversity for the entire raster in a first step (complete computation). We used radiometric filtering to separate

pixels covered by vegetation from vegetation-free or cloud-covered pixels. For that, we used the following threshold values: Blue Band (B02) of 1700, Near Infrared (NIR, B08) of 3700, and Normalized Difference Vegetation Index (NDVI, [B08-B03]/[B08+B03]) of 0.4. Based on pixel reflectance patterns, we build the spectral community map clustering the vegetated pixels within 100 clusters after dimensionality reduction by principal component analysis. Then, we retrieved the number of different clusters, i.e., spectral community richness, and the Shannon index of spectral diversity for each cell of 100 x 100 m. Manually, we carried out 20 iterations of the spectral community map and following indices.

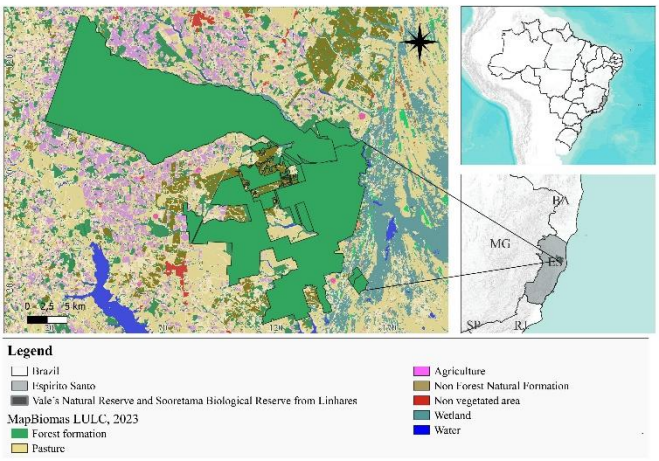


Fig. 1. Localization of the study site.

Then, we divided the entire raster in in 692 cells of 1 x 1 km and carried out different randomization runs. During each run, we randomly selected 200 1 x 1 km cells and pooled them in a single raster. Using the same computation parameters as in the previous section, we computed spectral community richness and the Shannon index of spectral diversity for each 100 x 100 m cell from the pooled raster. Number of randomizations was set to 70, so that each 1 x 1 km cell is expected to enter in around 20 distinct randomizations.

For evaluation, all cells with vegetation cover (recognized by radiometric filtering) small than 90% were excluded to reduce the effects of low vegetation cover on spectral diversity outcomes. Then, we compared the mean values across iterations (complete approach) with the mean values across randomizations using linear regression. Furthermore, we selected a couple of the 100 x 100 m cells to show how iterations and randomizations stabilize the spectral diversity values.

Finally, we plotted the spatial distribution of spectral diversity within the two reserves using QGIS.

3. RESULTS

Computation time of spectral diversity for two reserves with a 64 GB RAM virtual machine was similar for the complete and the randomization approach, about 1.5 days in both cases. Tests for larger areas showed that the computation of randomization trials can be run in up to three parallel R sessions, reducing computation time considerably compared to the complete version.

Outcomes for spectral community richness and Shannon diversity from both computation methods correlate significantly with R^2 values larger than 0.95 (Fig. 2). Single values originating from different computation approaches do not differ significantly (Fig. 3).

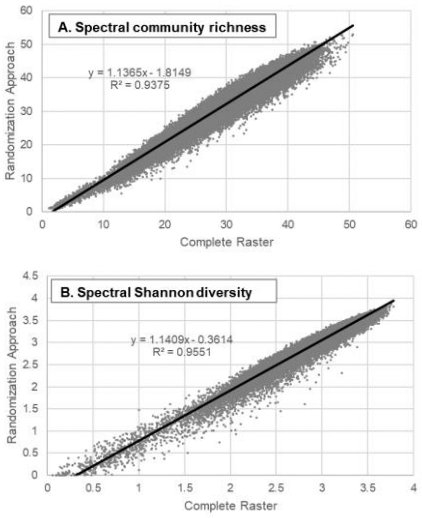


Fig. 2. Correlation between spectral species richness (A) and spectral Shannon diversity (B) computed by the complete and the randomization method. See Methods for details.

Consequently, the spatial distribution of spectral diversity computed from both methods reveals similar distributions, with lower spectral diversity identified for the northwestern and the southern tip of the study area. Lower spectral diversity than in the rest of the study was furthermore detected in the flooded area of the Barra Seca River and its tributaries at the northern border of the study area (Fig. 4).

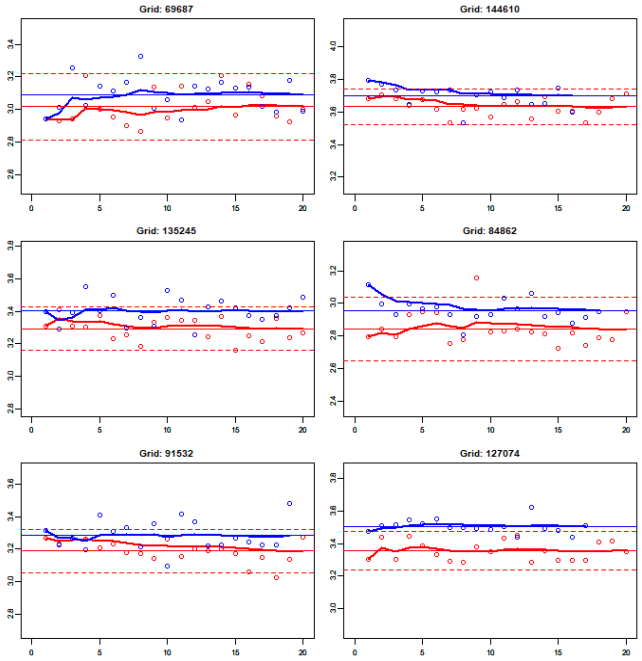


Fig. 3. Influence of the number of iterations (red dots) and randomizations (blue dots) on the computation of spectral Shannon diversity for six randomly chosen 100 x 100 m cells. Continuous lines represent mean values, and dashed lines (for complete computation only) are mean values ±standard deviation.

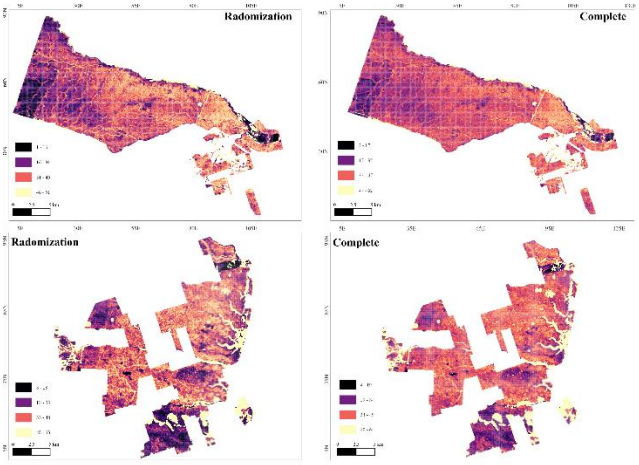


Fig. 4. Spatial distribution of spectral diversity in Vale's Natural Reserve and Sooretama Biological Reserve from Linhares, Espirito Santo, Brazil.

4. DISCUSSION

The computation of spectral diversity for two important conservation units from Espírito Santo State was straightforward. Both methods applied here retrieved a consistent pattern of the spatial distribution of spectral diversity in the study areas. This highlights that the randomization method presented here enables the use of spectral diversity to analyze larger areas not able to be processed by the complete approach due to memory limitations, and requires further investigations of the factors that predict the distribution of spectral diversity in space.

High determination coefficients and no statistical differences between the outcomes of spectral diversity computed from both methods encourages the usage of the randomization approach to compute spectral diversity for larger areas unable to be processed by machines with limited memory. Follow-up studies should focus in revealing the effects of the number of cells per randomization as well as the influence of cell size on the outcomes.

Within the two reserves, areas of lower spectral diversity were detected. More homogeneous spectral reflectance from non-forest vegetation at river margins may contribute to lower spectral diversity at the margins of the Barra Seca River. Furthermore, terrain roughness is expected to increase spectral diversity. These and further investigations are ongoing to reveal the predictors of spectral diversity to make this method a tool to monitor forest ecosystems.

5. CONCLUSION

In this study, we made important progress to upscale the usage of spectral diversity for the evaluation of tropical forest ecosystems. We show that the randomization returns results consistent with traditional approaches, which enables researchers to (1) test the hypothesis of spectral diversity in unprecedent scales, (2) upscale the applications of spectral diversity for the monitoring of changes in tropical forest ecosystems in response of global changes, and (3) outline predictors of spectral diversity in large or even very large areas.

8. REFERÊNCIAS

[1] S. L. Ustin, and J. A. Gamon . Remote sensing of plant functional types. *New Phytologist*, 186(4), 795–816. 2010.

[2] J. B. Féret and G. P. Asner. Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications*, 24(6), 1289–1296. 2014.

[3] J. B. Féret, and F. Boissieu. BiodivMapR: An R Package for α- and β-diversity Mapping Using Remotely Sensed Images. Edited by Kylie Scales Methods in *Ecology and Evolution/British Ecological Society*, 11 (1), pp. 64-70. 2020.

[4] M. Gastauer, W. R. Nascimento Jr, C. F. Caldeira, S. J. Ramos, P. W. M. Souza-Filho and J. B. Féret. Epectral diversity allows remote detection of the rehabilitation status in an Amazonian iron

mining complex. *International Journal of Applied Earth Observation and Geoinformation*, 106, p.102653. 2022.

[5] R. Pellens, I. Garay, and P. Grandcolas. Biodiversity conservation and management in the Brazilian Atlantic forest: every fragment must be considered. *Biodiversity Hotspots*. New York, Nova Science Publishers Inc, pp.1-45. 2009.