Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Brasil, 21-26 abril 2007, INPE, p. 6541-6547.

Spectral pattern predicts floristic composition and forest structure of white sand forests in the upper Rio Negro

Marcia Barbosa Abraão [1]
Bruce Walker Nelson [2]
Glenn Harvey Shepard Jr. [2,3]

[1] Instituto Socioambiental – ISA
Caixa Postal 21
69750-000 -- São Gabriel da Cachoeira - AM
marciaabraao@yahoo.com.br

[2] Instituto Nacional de Pesquisas da Amazônia – INPA
Programa de Pós-Graduação em Ecologia
Av André Araújo 2936 – 69060-001 - Manaus- AM, Brasil
bnelsonbr@yahoo.com.br

[3] University of East Anglia
School of Biological Sciences
NR4 7TJ Norwich, UK
gshepardjr@gmail.com

Abstract Landsat Thematic Mapper band 5 brightness and average brightness of bands 3, 4, and 5, were good predictors of tree beta diversity across a set of 15 forest plots on white sand in the upper Negro River of the Brazilian Amazon. These two spectral attributes explained 80% and 81%, respectively of overall change in floristic composition. Beta diversity was represented by a non-metric multidimensional scaling ordination. Floristic distances between all plot pairs took into account the relative importance of each tree species. The same two spectral attributes were closely related to plot basal area (R^2 = 0.84 and 0.85, respectively), but less closely related to average canopy height (R^2 = 0.65 and 0.60). Tree names were indigenous (Baniwa) folk-taxonomy names provided by three different knowledgeable informants. When floristic distances were derived using presence/absence of each species, inconsistent knowledge between these informants for the many trees belonging to rare taxa appeared as a strong artifact in the ordination.

Keywords: Landsat, floristic ordination, beta diversity, white-sand forests, ordenamento florístico, diversidade beta, campinaranas.

1. Introduction

In the basins of the Içana, Tiquié and upper Uaupés Rivers, white-sand forests dominate the landscape and constitute homogeneous patches clearly differentiated by their distinct borders and colors on RGB composites of Landsat Thematic Mapper bands 3, 4 and 5 (Figure 1). To examine the usefulness of Landsat pixel brightness for predicting tree community composition and forest structure, these attributes were obtained for a set of standard forest plots in the middle Içana River basin. Changes in floristic composition were represented by the scores of a single-axis ordination. Floristic composition and forest structure were then related by simple linear regression to spectral attributes from a Landsat image.

2. Methods

Fifteen forest plots on white sand soils were selected away from disturbed areas near three Baniwa communities (Juvitera, Jandu, Aracu) along a 50 km linear stretch of the Içana River (1.3-1.6^o N, 68.5-68.9^o W). The 15 plots included replicates of six undisturbed white-sand vegetation types as recognized and distinctively named according to indigenous Baniwa forest classification (Abraão 2005). Diameter at breast height (DBH) was measured for all trees over 5 cm DBH to calculate basal area of each species, i.e., total stem cross-section per hectare. Canopy height for each plot was obtained by averaging 25 regularly spaced measurements of tree height.

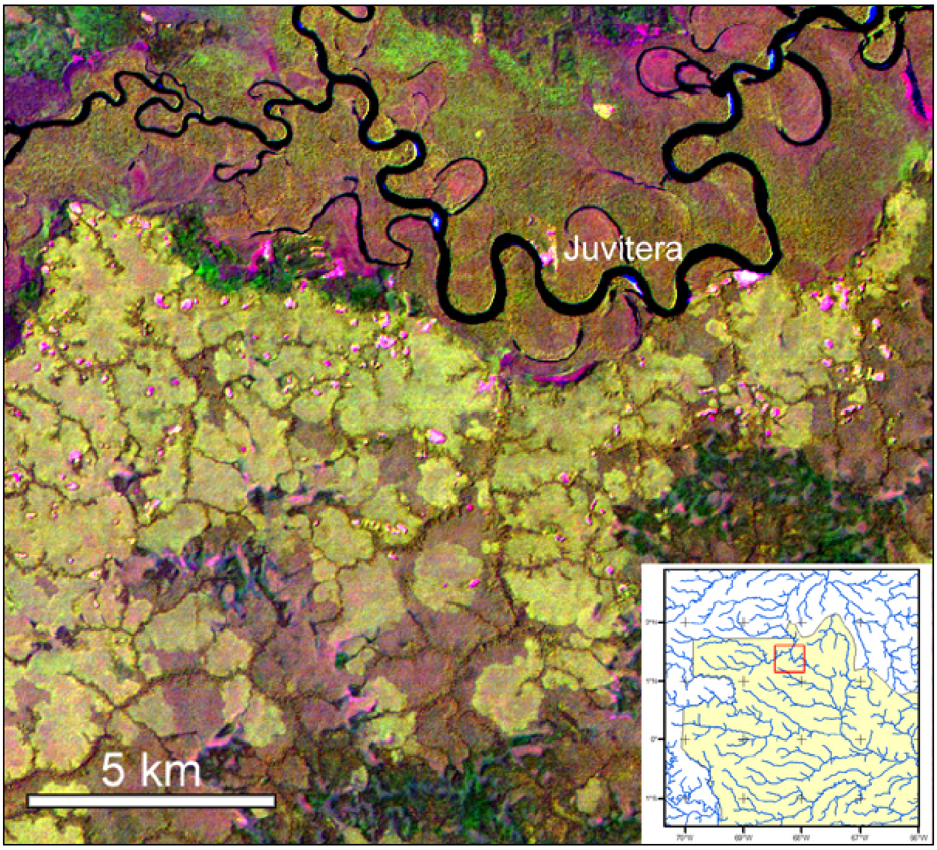

Figure 1. Distinct borders and colors of different forests on white sand south of the Içana River floodplain, near the Baniwa village of Juvitera. TM 543(RGB) false-color composite. Small pink patches are swidden fields. Inset shows image location in the "dog's head" region of NW Amazonas state.

To compensate for geometric inaccuracies of the Landsat image and of GPS readings at smaller scales (e.g., surface areas representing only a few pixels), each floristic inventory was divided among five sub-plots of 20x20 m (0.2 ha total inventory) spread over a larger 200 x 200m area representing a homogenous Baniwa vegetation type. GPS points were collected from each of the 5 sub-parcels, which were located on a geometrically corrected Landsat image, and polygons of about 8-10 pixels in size were digitized around each of these five points to identify 40-50 pixels per inventory plot. The average encoded radiance value of each plot's population of pixels was then extracted for each of three Landsat TM bands: 3, 4 and 5. Thus, for each band we used 15 average brightness values, one value for each inventory.

It was not possible to obtain authorization to make botanical collections in this region for later identification in the herbarium. Therefore, tree identifications (hereafter called "species") were based on names provided by three knowledgeable indigenous Baniwa informants, one from each study community. The three informants did not work together in the field so identifications were based on individual knowledge, not consensus. Though we could not compare proper scientific determinations from herbarium voucher specimens to the names given by the informants, we were nonetheless interested in evaluating the usefulness of folk-taxonomy names for detecting gradients in forest composition.

The basal area of each species in each inventory was used to detect linear gradients in floristic composition across the set of 15 plots. Basal area was used rather than abundance, because a few large trees of one species can occupy more canopy area and have a larger biomass than many small trees of another species. We also used presence/absence data, which emphasizes the contribution of less common species. Presence/absence analyses are important when detecting spatial patterns and indicators of beta diversity, because most tree species are locally rare. In forests on clay soils of the western Amazon Pitman *et al.* (2001) found that only 15-22% of all local tree species had densities ≥1 stem per hectare.

As recommended by McCune & Grace (2002) we used the Sorensen quantitative index as a measure of similarity between all pairs of plots and non-metric multi-dimensional scaling (MDS) as the ordination method. In a matrix of plots (rows) by species (columns), the cells contained the basal area of each species. These values were then normalized to the row sum, so that plots with more total basal area had the same weight in the ordination as plots with less basal area. For presence/absence data, each cell was filled with a value of 1 or zero and the same similarity index was used.

In a triangular matrix of 15 plots x 15 plots, the cells were filled with the Sorensen index for each plot pair. An MDS ordination attempts to reproduce the rank order of all Sorensen distances in this matrix when determining each plot's coordinates in a space of reduced dimensions. Note that in a one-dimensional solution, there will be no difference in the output of paired distances if the order of the plots is inverted. In a two-dimensional solution the data cloud can flip horizontally, vertically or rotate and still maintain all paired distances. In order to maintain consistent axis scores for all plots when repeating a two-dimensional analysis, the point cloud was always rotated so that its long axis was parallel to axis 1. This reduced the number of possible outcomes to the four combinations of two horizontal and two vertical flips. The analysis was repeated just a few times until the desired score signs were output.

Gradual change in community composition across all plots was represented by one or two indirect floristic gradients, which were the scores of the MDS ordination axes. Each forest plot's score from a single-axis MDS that explained a large portion of the original variation in floristic composition was then related by simple linear regression to the following spectral attributes from each plot: NDVI, brightness of each Landsat Thematic Mapper band (TM bands 3, 4, and 5) and the average of the three standardized band brightness values, which is a proxy for canopy shade content. These simple regressions allowed us to determine how well each spectral attribute "predicted" change in species composition from one place to another in the upper Rio Negro white-sand forests. Plot basal area and plot height were also regressed against the same five spectral attributes, to see how well these two forest structure attributes were predicted by the orbital sensor. Because the five spectral attributes were correlated, no multiple regressions were employed.

The Landsat image used for extracting spectral attributes was for the scene at WRS2 row 4, path 59, acquired 29 December 1993. The image was uncorrected, so it was geometrically adjusted to the Landsat Geocover tile for the region, dated ~1990. This base image has an RMS error of just 50 meters (https://zulu.ssc.nasa.gov/mrsid/docs/). The geometric quality of the cartographic base image was checked by collecting GPS points from a few distinct landmarks, such as the mouth of a narrow entrenched tributary of the Içana River. Forty ground control points were located on the cartographic base image and on a contrast-enhanced color composite of the Landsat image. The set of 40 ground control points had an average RMS error of 0.53 pixels. The individual bands were then geometrically corrected using these same control points feeding a cubic polynomial mapping function and the nearest-neighbor resampling method.

3. Floristic Composition

A total of 7541 trees were inventoried and 353 plant names were provided by the three Baniwa informants. After excluding species found only once in the entire study area, 7409 trees and 223 different tree names were retained for the community analysis. Most species in this study were uncommon or rare. The 30 most abundant species constituted 80% of all trees in the study, while the 120 rarest species contributed only 5% of all stems, even after single occurrences were excluded.

An ordination was also conducted with all 353 species to examine the effect of the less common plants. The single occurrences had no effect on the ordination using relative ecological importance of each species, probably because the more abundant species have an overwhelming influence on the Sorensen distances.

For the analysis using relative ecological importance, distances between all pairs of 15 plots arrayed along a single ordination axis accurately emulated the ranking of the original Sorensen distances of all plot pairs. Specifically, 87% of the variability in the Sorensen distances between all plot pairs was accounted for by a single-axis MDS. A two-dimensional MDS increased the explained variance only slightly, to 92%. Plot scores from the single MDS axis were therefore regressed against the five spectral attributes (Table 1). "Total explanatory power" (TEP) was obtained by multiplying the R^2 of each simple regression by the amount of variance in original Sorensen distances accounted for by the single MDS axis. For example, the R^2 value of 0.914 from regressing TM band 5 against the single-axis MDS score was multiplied by 0.87. Band 5 therefore predicted 80% of the change in floristic composition across the 15 forest plots.

Table 1. Simple regressions relating change in florístic composition to spectral attributes.

MDS score = ax + b + error				
a	*x*	*b*	R^2	*TEP*
-25.89	NDVI	14.00	0.31	0.27
-1.17	TM band 3	22.40	0.81	0.70
-0.15	TM band 4	9.76	0.68	0.59
-0.20	TM band 5	9.62	0.91	0.80
-1.07	Average brightness	0.0000	0.93	0.81

4. Forest Structure

Basal area of the fifteen plots ranged from 14.2 to 39.5 m^2 ha^{-1}. Average canopy height of a plot ranged from 7.3 m to 19.4 m. As expected, these two structural attributes were highly correlated (0.81 Pearson correlation). There was a strong correlation between floristic and structural changes. The single-axis MDS score from the quantitative floristic ordination had a Pearson correlation coefficient of 0.79 with canopy height, and 0.94 with plot basal area. Plots with low MDS scores were those with low canopy and less stem biomass, while plots with high MDS scores were tall and dense with more stem biomass.

The linear models for predicting basal area and canopy height from the five spectral attributes are shown in two tables below. Basal area had a stronger relationship with each independent variable than did canopy height. As with floristic composition, the best predictors of forest structure were TM band 5 and average brightness of all three bands (TM 3, 4, and 5). NDVI was again the worst predictor.

Table 2. Simple regressions relating forest structure to spectral attributes.

Basal Area = ax + b + error					Canopy Height = ax + b + error			
a	*x*	*b*	R^2		*a*	*x*	*b*	R^2
-193.13	NDVI	132.37	0.28		-54.76	NDVI	44.00	0.08
-8.88	TM band 3	197.49	0.75		-4.27	TM band 3	95.95	0.60
-1.14	TM band 4	101.34	0.62		-0.45	TM band 4	43.34	0.33
-1.50	TM band 5	100.44	0.84		-0.71	TM band 5	48.69	0.65
-8.09	Average brightness	27.95	0.85		-3.65	Average brightness	14.40	0.60

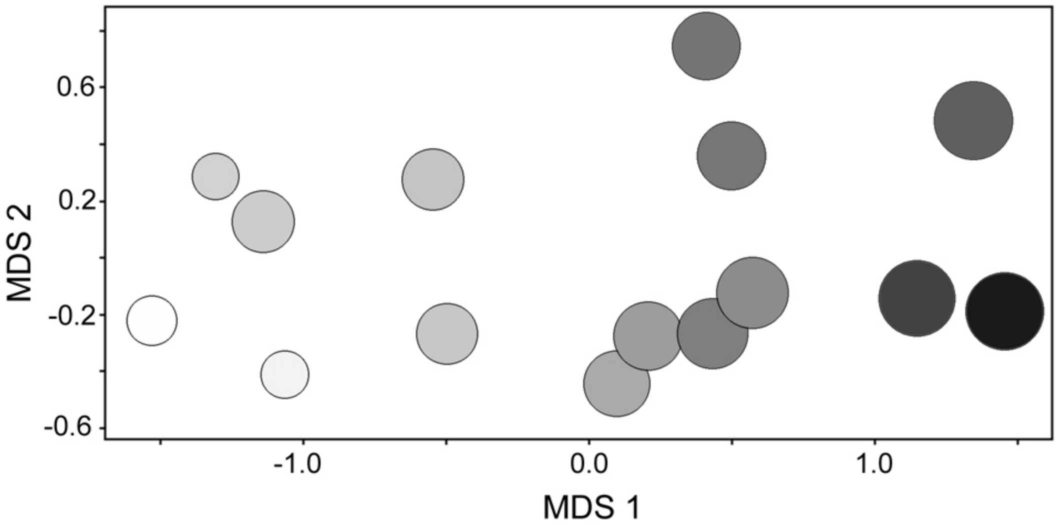

Figure 2. Congruent changes in floristic composition (based on quantitative data for 223 tree names), forest structure and spectral characteristics. Each dot's area is proportional to total basal area of the trees in that plot. Grey tones are scaled to TM band 5 brightness of each plot.

Figure 2 clearly shows parallel trends in floristic, structural and spectral attributes of the white sand forests. A two-axis MDS was employed here so as to form a two-dimensional graph, but axis 1 explained most (84%) of the variability in floristic composition.

5. Ordinations of presence/absence data

For the ordination based on presence/absence data, a single-axis MDS explained 70%, while a two-axis MDS explained 91% of the original variability in composition. Visual inspection of the two-axis solution (Figure 3) indicated that the person identifying the trees influenced apparent floristic composition. Plots identified by informants from Aracu and Jandu (both belonging to the same Baniwa clan) overlapped completely, but those identified by the informant from Juvitera (belonging to a different, multilingual Baniwa clan) formed a separate group. The three geographic regions of field work where each of the informants provided tree names were spaced at regular intervals along the Içana River. So distance between plots was not a likely factor. Rather, cultural and linguistic "distance" among informants may be at play.

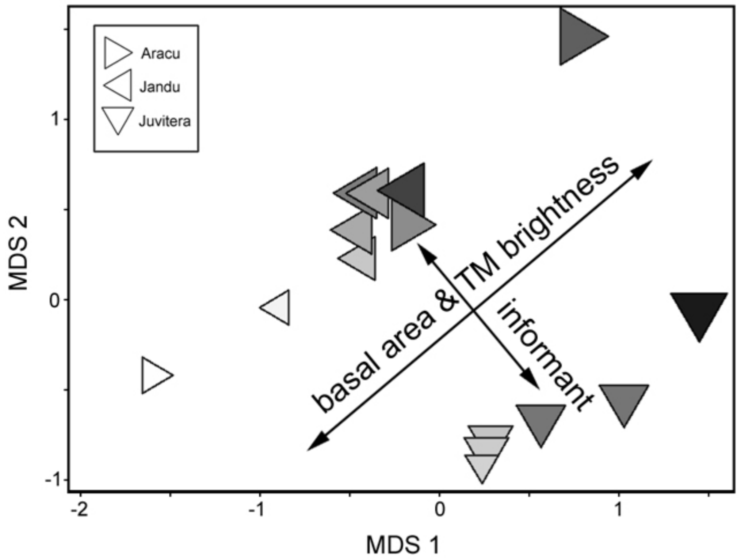

Figure 3. Two-dimensional MDS based on presence/absence data for 223 tree species. Basal area of each plot is proportional to triangle size; grey tones are scaled to TM 5 brightness. The 3 different informants are indicated by the names of their respective indigenous communities.

Along each elongate group of forest plots in Figure 3, toward the upper right, basal area increases and TM band 5 brightness decreases. There was a predictive relationship between canopy brightness and floristic composition that could be extracted in a multiple regression with three variables: TM 5, informant category and the interaction between these two variables. However, the magnitude of the interaction between canopy brightness in TM band 5 and the informant category would be different if other informants were used. Some informants agree, others have more idiosyncratic knowledge, know fewer trees or make more guesses. So even if interaction were accounted for with one pair of informants, this cannot be applied to compensate for differences in another pair of informants. This is clear from Figure 3, where one pair of informants is in greater agreement. Consequently, useful linear predictions of forest composition provided by other informants, using spectral attributes, cannot be extracted from the presence/absence data in this study.

6. Discussion and Conclusions

For the single axis MDS ordination that used relative ecological importance of each species, average brightness was a good predictor of change in floristic composition (R^2 = 0.93, Total Explanatory Power = 81%). Band 5 had a similarly high predictive power. Bands 3 and 4 had a poorer predictive ability with R^2 values of 0.81 and 0.68, respectively. NDVI was the worst predictor of composition and of structure. The reason for this is clear when NDVI is plotted against the MDS score. The relationship is non-linear. Inventory sites with very low canopy and probably some sandy patches, dry grass, dry lichens and illuminated bare branches are those at the low end of the MDS ordination. Moving up the MDS scale, NDVI increases as leafy canopy closure increases, covering the dry material. But then as canopy texture increases in the tallest white sand forests, NDVI decreases.

Brightness of TM band 5 does not suffer a reversal with increasing stand height and basal area. It first decreases as dry material is covered by green leafy canopies and continues to drop as the canopy becomes taller and more textured. Consequently TM band 5 and average brightness of three bands were the best predictors of forest structure. NDVI was a poor predictor for the same reasons given above.

The strong floristic gradient detected using vegetation types and individual trees identified by Baniwa names, and this gradient's strong correlation with canopy spectral attributes, plot basal area and canopy height, indicate that indigenous knowledge can be used for floristic analysis of white sand forests, and perhaps other Amazonian forests (see Shepard, Yu & Nelson 2004; Halme & Bodmer in press). The three informants will have agreed on the tree names when identifying the most common species in each plot, since the Baniwa name for each white sand formation is based on an abundant indicator species. For example, six plots, including four from a vegetation type named "Anelima", had very high densities of the tree "Ane" (apparently *Cassia spruceana*). This abundant species contributed 24-44% of the basal area in each of these six plots. It should be recognized consistently by all three informants.

When using presence/absence data of species' names provided even by highly knowledgeable indigenous informants, the influence of canopy spectral characteristics on the ordination was still detectable, but the ordination scores would probably change using other informants. Idiosyncratic knowledge will affect the ordination in an unpredictable manner. Differences in knowledge repertoires are likely to be much greater for less common species. With quantitative similarity indices, the identifications need only be reliable for the few abundant tree species, since these have an overwhelming influence on the similarity index. Presence/absence procedure gives equal weight to all species. Because there are many more rare species than common ones in the list of 223 names, the rare species have a greater effect on the presence/absence ordination. For conservation planning purposes, the many rare species are of greater interest when seeking proxies for spatial changes in species composition. For these many rare species, proper identification requires collecting voucher specimens for comparison with reference collections, before predictive relationships can be developed between canopy spectral properties and spatial change in species composition.

7. Acknowledgements

This research was supported by Brazil’s National Science Council (CNPq) and Amazonas state’s research foundation (FAPEAM). Special thanks go to João Claudio, our Baniwa field research collaborator, and to our informants Custódio Gaudêncio, Alberto Antônio Lourenço and Roberto Paiva. We thank the Baniwa Indigenous Organization of the Içana River (OIBI), the Rio Negro Indigenous Federation (FOIRN), and Instituto Socioambiental (ISA) for logistical support.

References

Abraão, M. B. **Conhecimento indígena, atributos florísticos, estruturais e espectrais como subsídio para inventariar diferentes tipos de florestas de Campinarana no rio Içana, alto rio Negro**. Thesis (Masters in Ecology), Instituto Nacional de Pesquisas da Amazônia/Univ. Federal do Amazonas, Manaus, 2005. 108 p.

Halme, K. J.; Bodmer, R. E. Correspondence between scientific and traditional ecological knowledge: rain forest classification by the non-indigenous ribereños in Peruvian Amazonia. **Biodiversity and Conservation**, In Press (accepted May 19, 2006).

Landsat Geocover specifications. Available at: <https://zulu.ssc.nasa.gov/mrsid/docs/>. Accessed 10 November 2006.

McCune, B.; Grace, J. B. **Analysis of ecological communities**. Gleneden Beach, Oregon: MJM Software Design, 2002. 300 p.

Pitman, N. C. A.; Terborgh, J. W.; Silman, M. R.; Nuñez, P.; Neill, D. A.; Cerón, C. E.; Palacios, W. A.; Aulestia, M. **Ecology**, v. 82, n. 8, p. 2101–2117, 2001.

Shepard, G. H. Jr.; Yu, D. W.; Nelson, B. W. Ethnobotanical ground-truthing and forest diversity in the Western Amazon. **Advances in Economic Botany** v. 15, p. 133-171, 2004.