Article:
Draper, FC, Baraloto, C, Brodrick, PG et al. (13 more authors) (2019) Imaging spectroscopy predicts variable distance decay across contrasting Amazonian tree communities. Journal of Ecology, 107 (2). pp. 696-710. ISSN 0022-0477

https://doi.org/10.1111/1365-2745.13067

Imaging spectroscopy predicts variable distance decay across contrasting Amazonian tree communities

Frederick C. Draper[1,2], Christopher Baraloto[2], Philip G. Brodrick[1], Oliver L. Phillips[3], Rodolfo Vasquez Martinez[4], Euridice N. Honorio Coronado[5], Timothy R. Baker[3], Ricardo Zárate Gómez[5], Carlos A. Amasifuen Guerra[6], Manuel Flores[6], Roosevelt Garcia Villacorta[7], Paul V. A. Fine[8], Luis Freitas[5], Abel Monteagudo-Mendoza [4,9], Roel J.W Brienen[3] and Gregory P. Asner[1]

[1] Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA

[2] International Center for Tropical Botany, Florida International University, Miami, FL, USA

[3] School of Geography, University of Leeds, Leeds, West Yorkshire, UK

[4] Jardín Botánico de Missouri, Oxapampa, Pasco, Perú

[5] Instituto de Investigaciones de la Amazonia Peruana, Iquitos, Loreto, Perú

[6] Facultad de Ciencias Forestales, Universidad Nacional de la Amazonía Peruana, Iquitos, Loreto, Perú

[7] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA

[8] Department of Integrative Biology, University of California, Berkeley, CA, USA

[9] Universidad Nacional de San Antonio Abad del Cusco, Cusco, Perú

Summary

1. The forests of Amazonia are among the most biodiverse on Earth, yet accurately quantifying how species composition varies through space (i.e. beta-diversity) remains a significant challenge. Here we use high-fidelity airborne imaging spectroscopy from the Carnegie Airborne Observatory to quantify a key component of beta-diversity, the distance decay in species similarity through space, across three landscapes in Northern Peru. We then compared our derived distance decay relationships to theoretical expectations obtained from a Poisson Cluster Process, known to match well with empirical distance decay relationships at local scales.

2. We used an unsupervised machine learning approach to estimate spatial turnover in species composition from the imaging spectroscopy data. We first validated this approach across two landscapes using an independent dataset of forest composition in 49 forest census plots (0.1-1.5 ha). We then applied our approach to three landscapes, which together represented terra firme clay forest, seasonally-flooded forest and white-sand forest. We finally used our approach to quantify landscape-scale distance decay relationships and compared these with theoretical distance decay relationships derived from a Poisson Cluster Process.

3. We found a significant correlation of similarity metrics between spectral data and forest plot data, suggesting that beta-diversity within and among forest types can be accurately estimated from airborne spectroscopic data using our unsupervised approach. We also found that estimated distance decay in species similarity varied among forest types, with seasonally-flooded forests showing stronger distance decay than white-sand and terra firme forests. Finally, we demonstrated that distance decay relationships derived from the theoretical Poisson Cluster Process compare poorly with our empirical relationships.

4. Synthesis: Our results demonstrate the efficacy of using high-fidelity imaging spectroscopy to estimate beta-diversity and continuous distance decay in lowland tropical forests. Furthermore, our findings suggest that distance decay relationships vary substantially among forest types, which has important implications for conserving these valuable ecosystems. Finally, we demonstrate that a theoretical Poisson Cluster Process poorly predicts distance decay in species similarity as conspecific aggregation occurs across a range of nested scales within larger landscapes.

Introduction

The forests of Amazonia are highly diverse, supporting as many as 16,000 tree species (ter Steege et al., 2013). The importance of this diversity, beyond its intrinsic value as a natural wonder, is increasingly well documented, for example, by underpinning key biogeochemical cycles and determining the resilience of Amazonian forests to climate change (Sakschewski et al. 2016). Despite this recognition of the importance of diversity, accurately quantifying how species composition varies through space (i.e., beta-diversity) in Amazonia remains a significant challenge given the remoteness of the largest tropical forest on Earth. Over recent years large networks of forest plots (e.g. RAINFOR, ATDN and CTFS) have provided invaluable insight into the spatial ecology of Amazon forests (Duque et al., 2017; Phillips et al., 2004; ter Steege et al., 2006). However, even summed together these networks represent only ~2000 ha of forest, with many plots in localised clusters. Therefore, using plot data alone to assess continuous spatial phenomena such as turnover in species composition represents a significant current limitation to understanding tropical biodiversity.

An alternative, yet complementary, approach to quantifying biodiversity is through the use of remotely sensed data integrated with existing plot data. Such an approach enables the acquisition of contiguous data over vast swaths of forests irrespective of accessibility, potentially transforming the power of an entirely ground-based approach. Multispectral data from satellite based remote sensing, in conjunction with plot data, has been used successfully to broadly classify different forest types (Draper et al., 2014; Salova et al., 2005) and to provide general assessment of species turnover in Amazonia (Thessler, 2008; Tuomisto et al., 2003). However, current satellite based multispectral sensors (e.g. Landsat) lack the spatial and spectral resolution required to sufficiently differentiate the high species-level diversity occurring within tropical forests (Rocchini, 2007a, 2007b; Rocchini et al., 2016). Recent advances in high-fidelity, laser guided imaging spectroscopy present a viable solution, and have been used successfully to estimate beta-diversity in Neotropical forests (Féret & Asner 2014a; b; Somers et al. 2015).

A key component of beta-diversity is the variation in species composition as a function of geographic distance (hereafter referred to as distance decay). Distance decay is a particularly useful concept as it allows for an understanding of the relative importance of different processes

that may determine patterns of beta-diversity, such as environmental filtering and dispersal limitation (Soininen, McDonald, & Hillebrand 2007; Tuomisto, Ruokolainen, & Yli-Halla 2003). Understanding variation in distance decay relationships among different landscapes and forest types also has important implications for designing effective conservation strategies (Socolar et al. 2016). For example, the gradient of distance decay can help to understand if conserving species in a given landscape or forest type will be maximised by many small or few large protected areas (Nekola & White 1999). Furthermore, distance decay relationships can be used to formally test theoretical predictions of community assembly, for example from neutral theory and sampling area theory (Chave & Leigh, 2002; Condit et al., 2002; Hubbell, 2001; Morlon et al., 2008).

One particularly significant theoretical model suggests that distance decay relationships are defined by the spatial aggregation of tree species, which can be characterized by a Poisson Cluster Process (hereafter PCP) (Plotkin et al. 2000; Morlon et al. 2008). This model is useful because it correctly recognises that tree species are spatially aggregated (Condit et al. 2000), but does not attempt to ascribe a particular community assembly mechanism. Furthermore, this model has accurately characterised species area curves, and distance decay relationships in number of tropical forests (Plotkin et al. 2000; Morlon et al. 2008). Importantly, while this model has found relatively good agreement at small scales (≤ 50 ha), it has not been possible to test this model at larger spatial scales. An key limitation of the PCP approach is that it assumes a single scale of aggregation, in this paper we test the validity of this assumption at larger spatial scales (>1000 ha).

Within western Amazonia, several plot based studies have examined distance decay relationships in tree communities, and most of these studies find an initial rapid decay in species similarity over the first few kilometres followed by a far more gradual decay over greater distances (Condit et al., 2002; Duque et al., 2009; Tuomisto et al., 2003). However, this relationship varies substantially with the spatial scale of study (Morlon et al., 2008; Phillips et al., 2003; Tuomisto et al., 2003), forest type (Draper et al., 2018), underlying geology (Phillips et al., 2003) and taxonomic group (Kristiansen et al., 2012; Tuomisto et al., 2003). Importantly, all of these plot based studies have been data limited, either using a relatively small number of plots (typically < 50 ha) to interpolate distance decay over tens to hundreds of kilometres (Condit et al., 2002;

Tuomisto et al., 2003), or using spatially continuous data to investigate distance decay over small spatial scales (≤ 50 ha) (May et al., 2016; Morlon et al., 2008).

Here we apply a sequence of unsupervised machine learning techniques (Féret & Asner 2014b) to continuous high-fidelity spectral datasets to quantify contiguous beta-diversity and associated distance decay relationships at a landscape scale (>1000 ha) across three lowland landscapes in Amazonian Peru. At each of these landscapes we apply our method to one of three distinct forest types: white-sand forest, seasonally-flooded forest and terra firme clay forest. We also use an extensive network of 49 forest census plots across two landscapes to thoroughly validate our approach and to answer the following questions.

1. Does high spatial resolution imaging spectroscopy accurately predict turnover in tree species composition across different forest types in lowland Amazonia?
2. How does distance decay in tree species composition vary across different forest types in lowland Amazonia?
3. How well does a theoretical PCP predict distance decay in tree species composition across a range of forest types in lowland Amazonia?

Methods

Study landscapes

Three distinct landscapes were used in this study: Allpahuayo Mishana, Jenaro Herrera, and Quebrada Braga. These landscapes are all located in the department of Loreto, Peru (Fig. 1), and were selected because they harbour at least one of the three most common forest types encountered across western Amazonia: terra firme clay forest, seasonally-flooded forest, and white-sand forests (Baraloto et al. 2011). The first landscape, Allpahuayo Mishana, is a national reserve located close to the city of Iquitos that contain a mosaic of terra firme clay and white-sand forest (Fine et al., 2010; García Villacorta et al., 2003). These white-sand forests have exceptionally nutrient poor sandy soils of cratonic origin, and harbour numerous endemic tree species (Fine et al. 2010). The second landscape, Jenaro Herrera, is a centre of research of the Instituto de Investigaciones de la Amazonía Peruana (IIAP). Jenaro Herrera is made up primarily of terra firme forest, although there are some small patches of white-sand forest, seasonally-flooded forest, and palm swamp forest (Honorio Coronado et al., 2008; Honorio Coronado et al., 2009). Finally, the Quebrada Braga landscape is located south of Jenaro Herrera, and is surrounded by the Ucayali river on three sides, these low-lying forests are inundated seasonally with nutrient-rich white water (Nebel et al. 2001).

Airborne data

We used the Carnegie Airborne Observatory (CAO) Airborne Taxonomic Mapping System (AToMS) to obtain fused high fidelity imaging spectroscopy and Light Detection and Ranging (LiDAR) data for all three of our landscapes (Asner et al. 2012). CAO flights took place between June and September 2012 at an altitude of approximately 2000 above ground level, with an average flight speed of 60 m s^{-1}, and a mapping swath of ~1.2 km. Spectral radiance data were collected between 380 and 2510 nm at 5 nm increments (Asner et al. 2012). These measurements were subsequently resampled to 10-nm resolution, resulting in 214 contiguous spectral bands at a ground-level resolution (pixel size) of 2 m. LiDAR data were obtained from a dual laser waveform scanner that was operated at 200 kHz, with a 17º scan half-angle from nadir, yielding a point density of 4 laser shots m^{-2} (up to 16 returns m^{-2}). Lidar data were used to produce maps of tree canopy height and ground surface at 1-m spatial resolution. Spectral and LiDAR data

were precisely geo-located using an embedded high resolution Global Positioning System-Inertial Measurement Unit (GPS-IMU).

The spectral radiance data were atmospherically corrected to apparent surface reflectance with the ACORN-5 model (Imspec LLC, Glendale, CA USA). Images were then processed to exclude pixels that were not fully sunlit (i.e. shaded by another tree), covered by cloud, or represented a non-forested land surface. Shade masks were built using LiDAR-derived ray tracing models (Asner et al. 2007), clouds were masked manually, and non-forested land surfaces were identified using a LiDAR derived map of tree canopy height where pixels with a canopy < 3 m were considered non-forested. In addition, spectral bands that contained sampling noise (wavelengths < 400 nm and > 2500 nm) or that were dominated by atmospheric water vapour (wavelengths 1350-1480 nm and 1780-2032 nm), were not used in this analysis.

Estimating beta-diversity from spectral data

To estimate beta-diversity from spectral data, we used the 'spectral species distribution' (SSD) approach, building on the previous work of Féret & Asner (2014 a, b) and more generally on the foundations of the spectral variation hypothesis (Palmer et al., 2002). Our approach assumes that the spectral properties of a landscape vary with species composition, and therefore we are able to use variation in spectral composition as a proxy for variation in species composition. At each of the three sites, we independently applied a seven-step analysis procedure to generate our mapped estimates of tree species compositional change as follows.

(1) We performed a principal component analysis (PCA) on our processed spectral image in order to reduce the high dimensionality of the spectral data and to isolate and remove sampling artefacts such as cross-track brightness gradients.
(2) We manually selected components associated with biological gradients by visually examining the first 35 components, and removing any that showed obvious artefacts, such as clear striping. This left 4-8 useful components that were used in steps 3-7. At all landscapes the first three components were always selected and the together the components represented >60% of the variance.
(3) We applied k-means clustering to the selected components, clustering each pixel into one of 50 possible 'spectral species'. Spectral species being simply clusters of pixels that have similar reflectance values, which may, but equally may not, trace onto actual

species. This process reduces the multi-layer image of PCs into a single layer image containing the spatial distribution of spectral species. Due to the large size of the dataset, k-means was applied using the 'mini-batch k-means' function in the Python package scikit learn, which provides near-equivalent performance at rapid computational speed (Pedregosa et al. 2012). Mini-batches of 10,000 pixels were used, each with 20 random starts.

(4) We then divided the resulting spectral species distribution image into 1 ha mapping kernels. Kernels in which > 66 % of pixels corresponded to either shade, non-vegetated ground, or were clouded were excluded from all further analysis. This led to a ~20 % loss of area from each landscape (Table 1).

(5) We then converted the image into a spectral species abundance matrix where each row corresponded to an individual kernel and each column to a spectral species, from which we calculated a Bray-Curtis distance matrix.

(6) We then applied Non-metric multi-dimensional scaling (NMDS) to the distance matrix in order to extract the most important compositional gradients in the spectral species data. The NMDS was optimized for three axes and run for 30 iterations.

(7) Finally, we re-projected the three NMDS axis scores into a raster format so that spatial variation in spectral species composition could be visualized.

The PCA and k-means analysis were undertaken using the Python package Sci-kit learn (Pedregosa et al. 2012). All beta-diversity analyses (steps 5 and 6) were performed in the R statistical environment using the Vegan package (Oksanen et al. 2013).

Plot inventory beta-diversity estimates

To validate our approach, we compared our estimates of beta-diversity derived from spectral data to measured beta-diversity obtained from inventory plot data at Allpahuayo Mishana and Jenaro Herrera. Our plot dataset consisted of 37 existing forest inventory plots distributed across Allpahuayo Mishana in white-sand and terra firme forest types, and 12 forest plots distributed across Jenaro Herrera in terra firme, white-sand and palm swamp forest types (Fig. 2). Plots varied in size from 0.1-1.5 ha, and five different sampling protocols were used as described below.

We used 12 large rectangular permanent sampling plots (0.5 to 1.5 ha), in which all tree stems with a diameter greater than 10 cm have been tagged and identified. Seven of these rectangular plots were one ha in size and belong to the RAINFOR Network, two of these plots were 1.5 ha in size (Peacock et al., 2007; Martinez & Phillips 2000). We also used three rectangular 0.5 ha plots in which all stems greater than 5 cm have been identified (Honorio Coronado et al., 2008). We further used 16 small 0.1 ha plots, in which all stems greater than 2.5 cm in diameter were identified. Six of these 0.1 ha plots were 'Gentry' plots consisting of ten 2 x 50 m intersecting transects (Gentry, 1982; Phillips et al., 2003). These six Gentry plots, alongside the seven 1 ha RAINFOR plots were downloaded from the ForestPlots.net online repository (Lopez-Gonzalez et al., 2009; Lopez-Gonzalez et al., 2011). The ten remaining 0.1 ha plots were rectangular 20 x 50 m plots (Zárate et al., 2006). We used four 0.5 ha modified Gentry plots, within which all stems greater than 2.5 cm in diameter were identified (Baraloto et al. 2011). The remaining 14 plots were circular plots in which all species greater than 10 cm dbh were identified (Baldeck et al., 2016); two of these circular plots were 0.25 ha and 12 were 0.14 ha. Summary details of the inventory plot dataset are given in Table 2, and full details of all plots are given in table S.1.

GPS coordinates were taken in the centre of each plot to determine its position within the landscape. There are significant uncertainties associated with using a GPS underneath a forest canopy, particularly for smaller inventory plots. Our approach partially mitigates these uncertainties as our aim is to align these plots with spectral species composition estimates at a 1 ha scale, and therefore, GPS locations need only be located in the correct 1 ha kernel. Ultimately, we removed five plots from this aggregate dataset in Allpahuayo Mishana (four 0.1 ha and one 0.5 ha), that were located < 10 m from a kernel boundary between white-sand forest and terra firme forest according to our spectrally derived map of estimated beta-diversity. As these plots were larger than 10 m in any dimension, there is a high likelihood that much of the area of these plots was situated in an incorrect kernel. These five boundary plots introduced additional variation in the relationship, as shown in Figure 3.

Because morpho-species were not standardised across datasets, it was necessary to exclude all individuals not identified to species level from the dataset before calculating beta-diversity. These exclusions led to a loss of 5-20% of individuals, which is likely to slightly increase the similarity among plots. However, patterns of beta-diversity among Amazonian tree census plots

have been shown to be generally robust to the exclusion of similar proportions of morpho-species (Pos et al. 2014).

Given that estimates of beta-diversity are sensitive to the number of individuals per plot, and that our dataset was made up of plots of different sizes (and different numbers of individuals), it was necessary to standardise our plot dataset by stem number before calculating beta-diversity. We did this by using a bootstrap resampling process. This process consisted of first establishing the minimum number of individuals in any plot, in this case 65, and then sampling (without replacement) 65 individuals from each plot at random. A Bray-Curtis distance matrix was then constructed using this subsample of 65 individuals per plot. Using this distance matrix, NMDS ordinations were performed. NMDS axis scores were then extracted for each plot. This process was then repeated 1000 times with a different set of 65 individuals per plot in order to develop confidence intervals for NMDS axis scores. Finally, we were able to compare NMDS axis scores derived from this plot inventory data with the corresponding NMDS axis scores derived from the spectral data.

Estimating spectral distance decay

To estimate the distance decay in species composition from spectral data within forest types, it was first necessary to isolate pixels that correspond to the forest type of interest. At Allpahuayo Mishana, the target forest type was white-sand forest. Using our validation data, we demonstrated that at this site white-sand forests can be readily separated from terra firme forests based on spectral composition (Fig. 2 and 3). Therefore, pixels with a value of greater than 0.3 on the first NMDS axis were classified as white-sand forest.

At Jenaro Herrera the target forest type was terra firme forest. We first used our spectral data to exclude small patches of white-sand forest from our analysis; to do this, we excluded all pixels with a value of greater than 0.2 on the second NMDS axis as this was shown to represent white sand forests in the validation data (Fig. 3). We then used the LiDAR derived DEM to separate pixels of seasonally-flooded forest from terra firme forest. Kernels with a mean elevation greater than 118 m were considered to be terra firme forest.

At Quebrada Braga the target forest type was seasonally-flooded forest. We used our LiDAR derived DEM to isolate those forests that are seasonally-flooded from those that are not. We

were able to use existing plot data to identify the elevation of seasonally-flooded forests (Kvist & Nebel 2001; Nebel et al. 2001). All kernels that had a mean elevation of 113-117 m a.s.l. were deemed to be seasonally-flooded. As this landscape is surrounded on three sides by a white-water river, we assume that seasonal flooding provides uniformly high nutrient deposition and that there are no further edaphic gradients.

To visualise the distance decay across each landscape we calculated the mean similarity (inverse Bray Curtis) for all paired plots within bins of 100 m, (i.e. the mean similarity between plots located 0-100 m apart, 100-200 m apart etc.). We have presented the ensemble mean and standard deviation with each distance bin and do not assume independence among these pairwise distances. Additionally, we calculated the first order derivative of similarity every 100 m across each landscape. We used a LOESS smoothing function (span = 0.35), to demonstrate how the derivative varies with distance across each landscape.

Theoretical distance decay

To assess the extent to which our empirical spectral distance decay relationships could be reproduced by a PCP, we applied the theoretical framework outlined by Morlon et al. (2008). Because we applied this approach to 50 spectral species rather than hundreds or thousands of species, it was essential that our measure of similarity was calculated using abundance rather than occurrence data. Therefore, we did not fit the general formula supplied by Morlon et al. (2008) which had been developed to using the Sorensen index. Instead, we simulated maps of spectral species distributions with a PCP, which we parameterised using fits of Ripley's K curves to our spectral species maps. Subsequently, we were able to derive abundance-based distance decay relationships from these theoretically derived maps of spectral species distributions.

The PCP is a stochastic mathematical process of assigning clusters of objects (here spectral species) in space according to the following: 1. Cluster centres for each object are randomly distributed across a landscape assuming a constant cluster density. The number of individuals in each cluster is drawn from a Poisson distribution. 3. Individuals within each cluster are then distributed based on a radially symmetrical Gaussian distribution .

In this study, a PCP was produced for each of the 50 spectral species across each of the three landscapes according to the following process:

(1) Empirical Ripley's K curves were derived for each spectral species in each landscape using the R package Spatstat (Baddeley & Turner 2005).

(2) When a Ripley's K curve is calculated for a PCP, it can be shown to have the functional from presented in Equation 1 (Plotkin et al. 2000). Consequently, we use an inverse modelling framework to match each empirically derived Ripley's K curve with Equation 1 by adjusting ρ (the density of clusters across the landscape), and μ (the intensity of individuals within each cluster).

Equation 1: $K(d)^{PCP} = \pi d^2 + \rho^{-1}\left(1 - \exp\left(\frac{-d^2}{4\mu^2}\right)\right)$

(3) Species likelihood probabilities were then determined for each spectral species using the ρ and μ values in a PCP in concert with the radial Gaussian probability function defined in Equation 2. Probabilities from each clump were overlaid on top of one another and the maximum likelihood was used.

Equation 2: $h(x, y) = (2\pi\mu^2)^{-1} \exp\left(\frac{-(x^2+y^2)}{2\mu^2}\right)$

(4) The 50 species likelihood maps (one per spectral species) were then normalized based on the abundance of each spectral species in the empirical maps. These likelihoods were then used to weight a random draw that was used to condense the likelihoods into a single, theoretically-based spectral species map.

(5) A one ha grid was then fit over the simulated spectral species distribution map and the Bray Curtis distance among one ha kernels was calculated in exactly the same way as was done with the empirical data. From this grid, theoretical distance decay relationships were calculated in exactly the same manner as was done with the empirical spectral data (i.e. by calculating the mean similarity (inverse Bray Curtis) for all paired plots within bins of 100 m.

(6) Steps 2 to 5 were then repeated 20 times, to generate 20 distinct theoretical spectral species maps and associated distance decay curves. The final curves presented were the mean of means within each 100 m bin and the standard deviations of the means.

Results

Validation with forest plot data

At Allpahuayo Mishana, our estimates of species compositional turnover derived from spectral data were strongly correlated with field plot-based measures of beta-diversity ($R^2 = 0.85$; $P < 0.001$; Fig. 3). However, the residual variance was higher among only terra firme forest plots ($R^2 = 0.29$; $P = 0.05$) than among only white-sand forest plots ($R^2 = 0.76$; $P < 0.001$). At Jenaro Herrera, there was also a highly significant relationship between beta-diversity estimated with our spectral approach and field-measured beta-diversity ($P = <0.001$), although there was more residual variance at this site than at Allpahuayo Mishana ($R^2=0.68$). Most of the variation in the relationship between spectral and plot data came from palm swamp forests, which were poorly distinguished in the second NMDS axis; instead, the third NMDS axis was more useful at identifying areas of palm swamp (figure S.2). The relationship between spectral composition and species composition was consistent across two landscapes, and among different field plot datasets that were established using different sampling protocols with different stem diameter size limits.

Mapping beta-diversity

Our spectrally-derived maps of estimated tree species composition demonstrate clear gradients across the three study landscapes (Fig. 4). However, the underlying determinants of these floristic gradients appear to be different among the three sites. At Allpahuayo Mishana, the three NMDS axes show similar spatial patterns (Figs 4 and S.1), with NMDS axes 2 and 3 additionally containing a substantial element of sampling artefact (i.e. clear striping). This relative uniformity across NMDS axes suggests there is a single predominant floristic gradient at this site, because, if multiple important floristic gradients were present, we would expect them to be reflected in different NMDS axes. Combined with field validation data, our spectrally-derived maps indicate that the primary floristic gradient at this site reflects an underlying edaphic gradient from nutrient rich terra firme clay soils, to nutrient poor white-sand soils. These white-sand forests were always found at higher elevations (>145 m a.s.l.) at Allpahuayo Mishana.

Our estimates of tree species composition also suggest that there is a strong spatial gradient in floristic composition at Quebrada Braga. Similar to Allpahuayo Mishana, consistency among NMDS axes suggests there is a single primary floristic gradient at Quebrada Braga (Fig. 4 and S.2). Somewhat surprisingly, this floristic gradient did not correspond strongly with elevation. The Quebrada Braga landscape is seasonally flooded by the large and nutrient-rich Ucayali River, which surrounds this landscape on three sides. Therefore, elevation will primarily determine the intensity and duration of this seasonal flooding.

Jenaro Herrera appears to be a more complex landscape than the other two, as it contains three distinct floristic gradients, demonstrated by three distinctive NMDS axes (Fig. 4 and S.3). This landscape appears to contain two forms of flooded forest, one flooded by nutrient-rich white water from the large Ucayali River and another flooded by nutrient-poor black water. In addition, there are patches of white-sand forest as well as forests that have been significantly impacted by anthropogenic activities.

Empirical spectral distance decay

We observed a consistent pattern of a rapid decline in floristic similarity over distances of 500 m or less across all three forest types. Beyond this initial steep decay in similarity, three patterns distinguish these landscapes. In white-sand forests at Allpahuayo Mishana, after a rapid decay in similarity over the initial 800 m there was almost no discernible decrease in similarity with increasing distance (Fig. 5 panels A and D).

In seasonally-flooded forests at Quebrada Braga, we found a constant decay in floristic similarity with increasing distance. As with the other two landscapes, this decline was steepest over the initial 700 metres. However, the decline in compositional similarity persisted over the entirety of this landscape, as demonstrated by the consistently negative differential values (Fig. 5 panels B and D).

Finally, in terra firme forests at Jenaro Herrera we found a steep decay in compositional similarity over 500 m, followed by a more gradual decline up to distances of 3 km (Fig. 5 panels C and D). Beyond 3 km there was no discernible decrease in similarity with increasing distance up to 10 km. Additionally, at Jenaro Herrera there was greater overall variation in compositional similarity across all distances compared with the other two sites, as shown by the wider error

bars. We attribute this variation to the greater environmental variation at this site, as well as greater overall species diversity in terra firme forests as opposed to both white-sand forests and seasonally-flooded forests.

Theoretical distance decay model

Overall the theoretical models derived from our PCP approach poorly represented the three empirical (spectrally derived) distance decay relationships (Fig. 5). At Allpahuayo Mishana, although the form of the theoretical distance decay relationship was very similar to that derived from the empirical data, the theoretically derived distance decay generally overestimates similarity relative to the empirical data (Fig. 5 panel A). Similarly, Fig. 5 panel B shows that at Jenaro Herrera, the general pattern of the distance decay relationship was reasonably characterized relative to the empirical relationship, but the overall distance magnitude was not. At Quebrada Braga, we found a very different pattern, with the PCP models predicting a sustained sharp decrease in similarity over the first kilometre, which was not reflected in the empirical data (Fig. 5 panel C). However, the shallow but continuous decline in similarity beyond the first kilometre demonstrated by the PCP at Quebrada Braga showed reasonable agreement with the empirically-based relationship (Fig. 5 panel C).

Discussion

Our results demonstrate that distance decay relationships vary among forest types in lowland Amazonia at a landscape scale. This is significant, because in contrast with previous plot-based studies, we are able to investigate this distance decay relationship continuously across landscapes while simultaneously maintaining high resolution. Within terra firme forests, our estimated distance decay curves are broadly consistent with a number of previous studies in this region (Condit et al., 2002; Duque et al., 2009), showing both rapid decay in similarity over short distances, followed by almost no decay at distances greater than 4 km. The two other forest types that we investigated also demonstrate this initial rapid decline in similarity over the first kilometre, supporting the idea that canopy tree species across forest types are spatially aggregated over scales less than one kilometre (Condit et al. 2000). However, beyond this first kilometre, patterns of distance decay sharply differ among different forest types.

The variation in distance decay among forest types is particularly apparent in seasonally-flooded forest, which shows a strong and relatively continuous decline in similarity with increasing distance. There are few plot based estimates of distance decay relationships in seasonally-flooded forests with which to compare our data (but see Wittmann et al., 2006; Draper et al., 2018). Nevertheless, our broad pattern of continuous decline in similarity appears to be consistent with these plot-based analyses. Much of the variation in spectral species composition across the Quebrada Braga landscape appears to be broadly independent of elevation. As elevation here should be a reasonable proxy for flooding duration and intensity, our data suggest that flooding duration and intensity are not the most important determinant of species composition in this landscape. This contrasts with a number of previous studies that have found flooding depth and duration to be the most important determinants of species composition (Assis et al., 2015; Junk et al., 2011; Wittmann et al., 2004; Wittmann et al., 2006).

Instead, our results appear to emphasize the importance of disturbance in determining species composition at this site. Disturbance has been recognised as an important driver of beta-diversity in West Amazonian floodplain forests (Puhakka et al., 1992; Salo et al., 1986). This may be especially true in Quebrada Braga as it is surrounded by the large and dynamic Ucayali River, which migrates laterally over decadal timescales (Salo et al., 1986; Schwenk et al., 2017).

Therefore, while some areas might have experienced large-scale disturbance relatively recently, other areas may not have been disturbed for many decades or centuries. Such disturbance patterns would also be spatially auto-correlated, and therefore consistent with the distance decay patterns we observe. The discrepancy between our study and previous plot-based studies (e.g. Assis et al., 2015; Junk et al., 2011; Wittmann et al., 2004; Wittmann et al., 2006) may arise from plot-based studies sampling predominantly mature seasonally-flooded forests over disturbed forests, whilst our study samples the whole landscape without this apparent bias. Fluvial disturbance is not the only form of large-scale spatially auto-correlated disturbance that may be driving beta-diversity patterns in Amazonian forests; for example, in central Amazonia large blow-down events have an important role in driving turnover in species composition (Marra et al. 2014).

In white-sand forests, the initial rapid decline in similarity with increasing distance is even more pronounced than in the other forest types and does not persist beyond the initial 800 m. This initial rapid decay may reflect the patchiness of white-sand forests at Allpahuayo Mishana. Patches of white-sand forests at this site are frequently smaller than 800 metres across, and ecological similarity is likely to be higher within a patch than between patches. In this way, white-sand forest tree communities may be functioning as meta-communities, separated by terra firme forests (Adeney et al., 2016; Palacios et al., 2016). The lack of declining similarity with increasing distance beyond 800 m is consistent with some published distance decay curves for white-sand forests in this region (Draper et al., 2018), whist others that have been developed for much broader spatial scales appear to show a more constant decay (García-Villacorta et al., 2016; Guevara et al. 2016), presumably because they include several compositionally distinct floras.

Jenaro Herrera presents a different, and perhaps more complex pattern than in the other landscapes, indicated by the three NMDS axes showing distinct spatial patterns that reflect different underlying gradients. For example, patches of white-sand forests and terra firme forests are clearly distinct in NMDS axis 2, whilst palm swamp forests appear more strongly in the third NMDS axis. Furthermore, unlike the other two landscapes, Jenaro Herrera appears to show a strong anthropogenic disturbance gradient, which can be seen in high values in NMDS axis 1 that cluster near the town (fig. S.1). This apparently high level of anthropogenic disturbance is in

some ways unsurprising as Jenaro Herrera supports a larger population than the other two sites and is surrounded by forests that are accessible and without formal legal protection. This contrasts with the other two landscapes, with Allpahuayo Mishana being accessible but protected and Quebrada Braga being unprotected but further from human development and due to seasonal flooding, relatively inaccessible.

A clear feature revealed by our LiDAR-derived DEM at Jenaro Herrera is the sharp increase in elevation that bisects the landscape from West to East (Fig. 4). This geological feature appears to be a boundary between the upland Tertiary Iquitos geanticline and Pleistocene alluvial terraces (Dumont et al., 1990; Dumont et al., 1991; Rasanen et al., 1992). Interestingly, this boundary appears to have little impact on floristic composition unlike other geological features in this region (Higgins et al., 2011, 2012). While field data will be required to confirm that there is little floristic turnover across this boundary, the boundary does not appear in local floristic classifications nor in maps of forest types (López Parodi & Freitas 1990; Honorio et al. 2008).

We were able to validate our approach by comparing our spectrally-derived estimates of beta-diversity with an extensive network of 53 forest plots distributed across two sites. Overall, this comparison provides compelling evidence that high fidelity imaging spectroscopy can be used to understand the spatial organisation of biodiversity in hyper-diverse tropical forests. Our results show highly significant linear relationship between spectrally-derived and plot-based estimates of beta-diversity consistent with previous studies that have used similar unsupervised approaches (Baldeck & Asner, 2013; Féret & Asner, 2014a, 2014b; Somers et al., 2015). Importantly, this strong relationship is preserved across plots using both 2 cm and 10 cm diameter cut-offs. As the spectral signal is derived entirely from the uppermost canopy layer, our results suggest that canopy level species composition may an excellent proxy for species composition in understory strata in these landscapes. The weaker relationship between spectral similarity and floristic similarity in terra firme forests may reflect the fact that fewer canopy species were recorded in this forest type. This is because the majority of stems recorded in the 0.1 ha plots are < 10 cm dbh, which will not reach the forest canopy in these tall forests. In the shorter stature white-sand forests, a larger proportion of small-stemmed trees will reach the canopy and therefore will be included in the spectral data.

Across all forest types, the distance decay relationships derived from the theoretical PCP compared poorly with the comparable empirical data. This mismatch suggests that the decay in community composition cannot be easily predicted by the clustering of conspecific individuals following a PCP. Major limitations of the PCP approach include the assumption that conspecific individuals are aggregated at a single scale, and the assumption that each clump of individuals throughout the landscape has the same Gaussian dispersal pattern (Morlon et al. 2008). The single scale of aggregation assumption may be largely correct at small spatial scales (≤ 50 ha) in relatively homogenous environments (Morlon et al. 2008), where trees are aggregated mainly at small scales < 50 m (Condit et al. 2000). However, at larger spatial scales (> 500 ha), conspecific individuals aggregate at a range of different scales due to dispersal limitation, environmental specificity, Janzen-Connell effects, and competition among individuals (Levin, 1992; Wiegand et al., 2007). Similarly, a species is unlikely to have constant density across a 50 ha plot, however across a landscape > 1000 ha, assuming a constant density becomes an even less plausible assumption. Our results demonstrate that these assumptions would need to be relaxed in order to reasonably predict distance decay relationships a landscape scales from theoretical spatial point process models such as the PCP.

Furthermore, our theoretical approach calculates PCP distributions for each spectral species independently; these distributions are then combined into a single map using random draws weighted by the landscape abundance of each spectral species in the empirical spectral species map. Our approach does not include interactions among species and between species and the environment, instead assuming the landscape is a homogeneous plane. Incorporating these biotic and abiotic interactions in future models could provide a way to further explore the relative influence of neutral and niche processes at landscape scales. Finally, our PCP was parameterized with by spectral species distributions. It is possible that parameterization based on actual species distribution data, which would be extremely difficult to collect at such large scales, may lead to different results. An approach integrating field and spectral species distributions could provide further insight.

A more general limitation of our approach is that we cluster the spectral signal of the entire landscape into just 50 spectral species and assume they are representative of hyper-diverse tropical forest landscapes that will contain hundreds (if not thousands) of tree species. While this

approach is well supported by both our comparisons with field data, and previous work that has shown 40 spectral species to be optimal (Feret & Asner 2014a), there are limitations. In general, it is likely that common canopy species will dominate the spectral signal as they make up a far greater proportion of the sunlit canopy, whilst rare and or understory species will be under-represented. Rare species are thought to have more localised and environmentally specific distributions (Hubbell, 2013), and therefore, the extent to which common species can be used to investigate spatial patterns of beta-diversity merits further investigation.

Additionally, using 50 spectral species elevates the similarity among plots within each forest type. This is especially evident in white-sand forests, where overall similarity is far higher in our spectral based analysis than has been found previously in plot based studies (Fine et al., 2010; García-Villacorta et al., 2016; Guevara et al., 2016, Draper et al., 2018). Many white-sand specialist tree species share functional characteristics that are likely to make them spectrally similar, such as increased leaf thickness and toughness, as well as lower concentrations of foliar N and P (Asner et al., 2016; Fortunel et al., 2014; Fyllas et al., 2009). Therefore, the diversity within white-sand forests may be poorly represented by our approach, resulting in an artificial increase in similarity between plots. However, the tight correlation between spectral and plot-based estimates of species composition in white-sand forests suggests that despite the overall increase in similarity among plots, our approach is still able to capture the main correlates of plot diversity.

The strength of our approach is that we can apply this method continuously to much larger areas than would be impossible using field data alone. Therefore, there is great potential for using our method to quantify beta-diversity and distance decay relationships continuously over far greater spatial extents. Furthermore, our approach is not only able to quantify beta-diversity, but also to precisely geo-locate where turnover occurs and therefore to suggest which environmental features may be important. We suggest that unsupervised spectral-based approaches, such as ours, can be used to actively guide field efforts to areas containing floristic assemblages that are poorly represented by current plot networks. We advocate for closer collaboration among ecologists using field-based data and those using imaging spectroscopy data.

In summary, this study demonstrates that distance decay relationships vary substantially among landscapes and forest types in lowland Amazonia, consistent with much of what has been found

previously using field plot-based data. Nevertheless, we also present findings that challenge previous hypotheses regarding the environmental drivers of tree species composition. In particular, we suggest that edaphic properties and topography may not always be the most important determinants of floristic composition, and in dynamic floodplain landscapes, disturbance may be a more important driver of tree species composition. Comparing estimates derived from our spectral data with a large dataset of forest plots, we provide compelling evidence for the validity of our approach, not only in classifying broad forest types, but also in describing subtle changes in floristic composition. Finally, our results demonstrate that distance decay relationships are driven by conspecific individuals aggregating at a range of nested scales across landscapes. Reproducing these patterns from theory will require the assumptions of PCP models to be relaxed.

Acknowledgements

This study was supported through a joint project between the Carnegie Institution for Science and the International Center for Tropical Botany at Florida International University. CAO data collection was supported by the John D. and Catherine T. MacArthur Foundation and the Avatar Alliance Foundation. We thank N Vaughn and D Knapp for processing the CAO data used in this analysis. Plot installation, fieldwork and botanical identification by the authors and colleagues has been supported by several grants including a Gordon and Betty Moore Foundation grant to RAINFOR, the EU's Seventh Framework Programme (283080, 'GEOCARBON') and NERC Grants to OLP (Grants NER/A/S/2000/0053, NE/B503384/1, NE/F005806/1, and a NERC Postdoctoral Fellowship), and a National Geographic Society for supporting forest dynamics research in Amazonian Peru (grant #5472-95). OLP is supported by an ERC Advanced Grant and is a Royal Society-Wolfson Research Merit Award holder. The Carnegie Airborne Observatory is made possible by grants and donations to GP Asner from the Avatar Alliance Foundation, Grantham Foundation for the Protection of the Environment, Gordon and Betty Moore Foundation, the John D. and Catherine T. MacArthur Foundation, W. M. Keck Foundation, the Margaret A. Cargill Foundation, Mary Anne Nyburg Baker and G. Leonard Baker, Jr., and William R. Hearst III.

Authors contributions

FCD, GPA and CB conceived the ideas and designed the methodology. GPA collected the imaging spectroscopy data. CB, OLP, RVM, RZG, CAAG, MF, ENHC, TRB, RGV, PVAF, LF, AMM, and RJWB collected the field validation data. FCD, PB and GPA analysed the data. FCD wrote the manuscript with input from GPA, CB and PB. All authors contributed to drafts and gave final approval for publication.

References

Adeney, J.M., Christensen, N.L., Vicentini, A. & Cohn-Haft, M. (2016) White-sand Ecosystems in Amazonia. Biotropica, **48**, 7–23.

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E. & Vaughn, N. (2016) Large-scale climatic and geophysical controls on the leaf economics spectrum. Proceedings of the National Academy of Sciences of the United States of America, **113**, E4043-51.

Asner, G.P., Knapp, D.E., Boardman, J., Green, R.O., Kennedy-Bowdoin, T., Eastwood, M., … Field, C.B. (2012) Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. Remote Sensing of Environment, **124**, 454–465.

Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J. & Field, C.B. (2007) Carnegie Airborne Observatory: in-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. Journal of Applied Remote Sensing, **1**, 013536.

Assis, R.L., Wittmann, F., Piedade, M.T.F. & Haugaasen, T. (2015) Effects of hydroperiod and substrate properties on tree alpha diversity and composition in Amazonian floodplain forests. Plant Ecology, **216**, 41–54.

Baddeley, A. & Turner, R. (2005) Spatstat: an R package for analyzing spatial point patterns. Journal of statistical software, **12**, 1–42.

Baldeck, C. & Asner, G. (2013) Estimating Vegetation Beta Diversity from Airborne Imaging Spectroscopy and Unsupervised Clustering. Remote Sensing, **5**, 2057–2071.

Baldeck, C.A., Tupayachi, R., Sinca, F., Jaramillo, N. & Asner, G.P. (2016) Environmental drivers of tree community turnover in western Amazonian forests. Ecography, **39**, 1089–1099.

Baraloto, C., Rabaud, S., Molto, Q., Blanc, L., Fortunel, C., Hérault, B., … Fine, P.V.A. (2011) Disentangling stand and environmental correlates of aboveground biomass in Amazonian forests. Global Change Biology, **17**, 2677–2688.

Chave, J. & Leigh, E.G. (2002) A spatially explicit neutral model of beta-diversity in tropical forests. Theoretical population biology, **62**, 153–68.

Condit, R., Ashton, P.S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., … Yamakura, T. (2000) Spatial Patterns in the Distribution of Tropical Tree Species. Science, **288**.

Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R.B., … Hubbell, S.P. (2002) Beta-diversity in tropical forest trees. Science (New York, N.Y.), **295**, 666–9.

Draper, F.C., Honorio Coronado, E.N., Roucoux, K.H., Lawson, I.T., Pitman, N.C.A., Fine, P.V.A., … Baker, T.R. (2018) Peatland forests are the least diverse tree communities documented in Amazonia, but contribute to high regional beta-diversity. Ecography, **41**, 1–14.

Draper, F.C., Roucoux, K.H., Lawson, I.T., Mitchard, E.T.A., Honorio Coronado, E.N., Lähteenoja, O., Montenegro, L.T., Sandoval, E. V, Zaráte, R. & Baker, T.R. (2014) The distribution and amount of carbon in the largest peatland complex in Amazonia. Environmental Research Letters, **9**.

Dumont, J.F., Deza, E. & Garcia, F. (1991) Morphostructural provinces and neotectonics in the Amazonian lowlands of Peru. Journal of South American Earth Sciences, **4**, 373–381.

Dumont, J.F., Lamotte, S. & Kahn, F. (1990) Wetland and upland forest ecosystems in Peruvian Amazonia: Plant species diversity in the light of some geological and botanical evidence. Forest Ecology and Management, **33**–**34**, 125–139.

Duque, A., Muller-Landau, H.C., Valencia, R., Cardenas, D., Davies, S., … Vicentini, A. (2017) Insights into regional patterns of Amazonian forest structure, diversity, and dominance from three large terra-firme forest dynamics plots. Biodiversity and Conservation, **26**, 669–686.

Duque, Á., Phillips, J.F., von Hildebrand, P., Posada, C.A., Prieto, A., Rudas, A., Suescún, M. & Stevenson, P. (2009) Distance Decay of Tree Species Similarity in Protected Areas on Terra Firme Forests in Colombian Amazonia. Biotropica, **41**, 599–607.

Féret, J.-B. & Asner, G.P. (2014a) Microtopographic controls on lowland Amazonian canopy diversity from imaging spectroscopy. Ecological Applications, **24**, 1297–1310.

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

Fine, P.V.A., García-Villacorta, R., Pitman, N.C.A., Mesones, I. & Kembel, S.W. (2010) A Floristic Study of the White-Sand Forests of Peru. Annals of the Missouri Botanical Garden, **97**, 283–305.

Fortunel, C., Paine, C.E.T., Fine, P.V. a., Kraft, N.J.B. & Baraloto, C. (2014) Environmental factors predict community functional composition in Amazonian forests (ed G De Deyn). Journal of Ecology, **102**, 145–155.

Fyllas, N.M., Patiño, S., Baker, T.R., Bielefeld Nardoto, G., Martinelli, L.A., Quesada, C.A., … Lloyd, J. (2009) Basin-wide variations in foliar properties of Amazonian forest: phylogeny, soils and climate. Biogeosciences, **6**, 2677–2708.

García-Villacorta, R., Dexter, K.G. & Pennington, T. (2016) Amazonian White-Sand Forests Show Strong Floristic Links with Surrounding Oligotrophic Habitats and the Guiana Shield. Biotropica, **48**, 47–57.

García Villacorta, R., Ahuite Reátegui, M. & Olórtegui Zumaeta, M. (2003) Clasificación de bosques sobre arena blanca de la Zona Reservada Allpahuayo-Mishana. Folia Amazonica, **14**, 17–33.

Gentry, A.H. (1982) Patterns of Neotropical Plant Species Diversity. Evolutionary Biology, pp. 1–84. Springer US, Boston, MA.

Guevara, J.E., Damasco, G., Baraloto, C., Fine, P.V.A., Peñuela, M.C., Castilho, C., … ter Steege, H. (2016) Low Phylogenetic Beta Diversity and Geographic Neo-endemism in Amazonian White-sand Forests. Biotropica, **48**, 34–46.

Higgins, M.A., Asner, G.P., Perez, E., Elespuru, N., Tuomisto, H., Ruokolainen, K. & Alonso, A. (2012) Use of Landsat and SRTM Data to Detect Broad-Scale Biodiversity Patterns in Northwestern Amazonia. Remote Sensing, **4**, 2401–2418.

Higgins, M.A., Ruokolainen, K., Tuomisto, H., Llerena, N., Cardenas, G., Phillips, O.L., … Räsänen, M. (2011) Geological control of floristic composition in Amazonian forests. Journal of Biogeography, **38**, 2136–2149.

Honorio Coronado, E.N., Baker, T.R., Phillips, O.L., Pitman, N.C.A., Pennington, R.T., Vásquez Martínez, R., … Freitas Alvarado, L. (2009) Multi-scale comparisons of tree composition in Amazonian terra firme forests. Biogeosciences, **6**, 2719–2731.

Honorio, E.N., Pennington, T.R., Freitas, L.A., Nebel, G. & Baker, T.R. (2008) Análisis de la composición florística de los bosques de Jenaro Herrera ,. , **15**, 53–60.

Hubbell, S.P. (2001) The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32). Princeton University Press.

Hubbell, S.P. (2013) Tropical rain forest conservation and the twin challenges of diversity and rarity. Ecology and Evolution, **3**, 3263–3274.

Junk, W.J., Piedade, M.T.F., Schöngart, J., Cohn-Haft, M., Adeney, J.M. & Wittmann, F. (2011) A classification of major naturally-occurring Amazonian lowland wetlands. Wetlands, **31**, 623–640.

Kristiansen, T., Svenning, J.C., Eiserhardt, W.L., Pedersen, D., Brix, H., Munch Kristiansen, S., … Balslev, H. (2012) Environment versus dispersal in the assembly of western Amazonian palm communities. Journal of Biogeography, **39**, 1318–1332.

Kvist, L.P. & Nebel, G. (2001) A review of Peruvian flood plain forests: ecosystems, inhabitants and resource use. Forest Ecology and Management, **150**, 3–26.

Levin, S.A. (1992) The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. Ecology, **73**, 1943–1967.

Lopez-Gonzalez, G., Lewis, S.L., Burkitt, M., Baker, T.R. & Phillips, O.L. (2009) Forestplots. net database. University of Leeds, Leeds, UK. Available at: http://www. forestplots. net/(accessed 13 June 2015).

Lopez-Gonzalez, G., Lewis, S.L., Burkitt, M. & Phillips, O.L. (2011) ForestPlots.net: a web application and research tool to manage and analyse tropical forest plot data. Journal of Vegetation Science, **22**, 610–613.

López Parodi, J. & Freitas, D. (1990) Geographical aspects of forested wetlands in the lower Ucayali, Peruvian Amazonia. Forest Ecology and Management, **33**–**34**, 157–168.

Marra, D.M., Chambers, J.Q., Higuchi, N., Trumbore, S.E., Ribeiro, G.H.P.M., dos Santos, J., … Wirth, C. (2014) Large-Scale Wind Disturbances Promote Tree Diversity in a Central Amazon Forest. PLoS ONE, **9**, e103711.

Martinez, R.V. & Phillips, O.L. (2000) Allpahuayo: Floristics, Structure, and Dynamics of a High-Diversity Forest in Amazonian Peru. Annals of the Missouri Botanical Garden, **87**, 499.

May, F., Wiegand, T., Lehmann, S. & Huth, A. (2016) Do abundance distributions and species aggregation correctly predict macroecological biodiversity patterns in tropical forests? Global Ecology and Biogeography, **25**, 575–585.

Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., … Green, J.L. (2008) A general framework for the distance-decay of similarity in ecological communities. Ecology letters, **11**, 904–17.

Nebel, G., Kvist, L.P., Vanclay, J.K., Christensen, H., Freitas, L. & Ruiz, J. (2001) Structure and floristic composition of flood plain forests in the Peruvian Amazon I. Overstorey. Forest Ecology and Management, **150**, 27–57.

Nekola, J.C. & White, P.S. (1999) The distance decay of similarity in biogeography and ecology. Journal of Biogeography, **26**, 867–878.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., … Wagner, H. (2013) Package 'vegan.' Community ecology package, version, **2**.

Palacios, J., Zarate, R., Torres, G., Denux, J.P., Maco, J.T., Gallardo, G.P., … Cuadros, A. (2016) Mapeo de los bosques tipo varillal utilizando Imagenes de satelite rapideye en la provinica Maynas, Loreto, Peru. Folia Amazónica, **25**, 25–36.

Palmer, M.W., Earls, P.G., Hoagland, B.W., White, P.S. & Wohlgemuth, T. (2002) Quantitative tools for perfecting species lists. Environmetrics, **13**, 121–137.

Peacock, J., Baker, T.R., Lewis, S.L., Lopez-Gonzalez, G. & Phillips, O.L. (2007) The RAINFOR database: monitoring forest biomass and dynamics. Journal of Vegetation Science, **18**, 535–542.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, É. (2012) Scikit-learn: Machine Learning in Python.

Phillips, O.L., Baker, T.R., Arroyo, L., Higuchi, N., Killeen, T.J., Laurance, W.F., … Vinceti, B. (2004) Pattern and process in Amazon tree turnover, 1976-2001. Philosophical Transactions of the Royal Society B: Biological Sciences, **359**, 381–407.

Phillips, O.L., Martínez, R.V., Vargas, P.N., Monteagudo, A.L., Zans, M.-E.C., Sánchez, W.G., … Rose, S. (2003a) Efficient plot-based floristic assessment of tropical forests. Journal of Tropical Ecology, **19**, 629–645.

Phillips, O.L., Vargas, P.N., Lorenzo, A., Cruz, A.P., Chuspe, M., Sánchez, W.G., … Rose, S. (2003b) Habitat association among Amazonian tree species : a landscape-scale approach. , 757–775.

Plotkin, J.B., Potts, M.D., Leslie, N., Manokaran, N., Lafrankie, J. & Ashton, P.S. (2000) Species-area Curves, Spatial Aggregation, and Habitat Specialization in Tropical Forests. Journal of Theoretical Biology, **207**, 81–99.

Pos, E.T., Guevara Andino, J.E., Sabatier, D., Molino, J.F., Pitman, N., Mogollón, H., … ter Steege, H. (2014) Are all species necessary to reveal ecologically important patterns? Ecology and Evolution, **4**, 4626–4636.

Puhakka, M., Kalliola, R., Rajasilta, M. & Salo, J. (1992) River Types, Site Evolution and Successional Vegetation Patterns in Peruvian Amazonia. Journal of Biogeography, **19**, 651–665.

Rasanen, M., Nellerf, R.O.N., Saloj, J. & Jungner, H. (1992) Recent and ancient fluvial deposition systems in the Amazonian foreland basin , Peru. , **129**, 293–306.

Rocchini, D. (2007a) Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. Remote Sensing of Environment, **111**, 423–434.

Rocchini, D. (2007b) Distance decay in spectral space in analysing ecosystem β-diversity. International Journal of Remote Sensing, **28**, 2635–2644.

Rocchini, D., Boyd, D.S., Féret, J.-B., Foody, G.M., He, K.S., Lausch, A., … Pettorelli, N.

(2016) Satellite remote sensing to monitor species diversity: potential and pitfalls (eds A Skidmore and A Chauvenet). Remote Sensing in Ecology and Conservation, **2**, 25–36.

Sakschewski, B., von Bloh, W., Boit, A., Poorter, L., Peña-Claros, M., Heinke, J., … Thonicke, K. (2016) Resilience of Amazon forests emerges from plant trait diversity. Nature Climate Change, **6**, 1032–1036.

Salo, J., Kalliola, R., Häkkinen, I., Mäkinen, Y., Niemelä, P., Puhakka, M. & Coley, P.D. (1986) River dynamics and the diversity of Amazon lowland forest. Nature, **322**, 254–258.

Salovaara, K.J., Thessler, S., Malik, R.N. & Tuomisto, H. (2005) Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery. Remote Sensing of Environment, **97**, 39–51.

Schwenk, J., Khandelwal, A., Fratkin, M., Kumar, V. & Foufoula-Georgiou, E. (2017) High spatiotemporal resolution of river planform dynamics from Landsat: The RivMAP toolbox and results from the Ucayali River. Earth and Space Science, **4**, 46–75.

Socolar, J.B., Gilroy, J.J., Kunin, W.E. & Edwards, D.P. (2016) How Should Beta-Diversity Inform Biodiversity Conservation? Trends in Ecology & Evolution, **31**, 67–80.

Soininen, J., McDonald, R. & Hillebrand, H. (2007) The distance decay of similarity in ecological communities. Ecography, **30**, 3–12.

Somers, B., Asner, G.P., Martin, R.E., Anderson, C.B., Knapp, D.E., Wright, S.J. & Van De Kerchove, R. (2015) Mesoscale assessment of changes in tropical tree species richness across a bioclimatic gradient in Panama using airborne imaging spectroscopy. Remote Sensing of Environment, **167**, 111–120.

Ter Steege, H., Pitman, N.C.A., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., … Castellanos, H. (2006) Continental-scale patterns of canopy tree composition and function across Amazonia. Nature, **443**, 444–447.

Ter Steege, H., Pitman, N.C.A., Sabatier, D., Baraloto, C., Salomão, R.P., Guevara, J.E., … Silman, M.R. (2013) Hyperdominance in the Amazonian tree flora. Science, **342**.

Thessler, S. (2008) Remote Sensing of Floristic Patterns in the Lowland Rain Forest Landscape.

Tuomisto, H., Poulsen, A.D., Ruokolainen, K., Moran, R.C., Quintana, C., Celi, J. & Cañas, G. (2003a) Linking floristic patterns with soil heterogeneity and satellite imagery in Ecuadorian Amazonia. Ecological Applications, **13**, 352–371.

Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. (2003b) Dispersal, environment, and floristic variation of Western Amazonian forests. Science, **299**, 241–244.

Wiegand, T., Gunatilleke, S., Gunatilleke, N. & Okuda, T. (2007) Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. Ecology, **88**, 3088–3102.

Wittmann, F., Junk, W.J. & Piedade, M.T.F. (2004) The varzea forests in Amazonia: flooding and the highly dynamic geomorphology interact with natural forest succession. Forest Ecology and Management, **196**, 199–212.

Wittmann, F., Schongart, J., Montero, J.C., Motzer, T., Junk, W.J., Piedade, ... & Worbes, M. (2006) Tree species composition and diversity gradients in white-water forests across the Amazon Basin. Journal of Biogeography, **33**, 1334–1347.

Zárate, R., Amasifuen, C. & Flores, M. (2006) Floración y Fructificación de plantas leñosas en bosques de arena blanca y de suelo arcilloso en la Amazonía Peruana. Revista Peruana de Biología, **13**.

Table 1. Summary of the spectral data used to estimate species composition for the three study landscapes

	Allpahuayo Mishana	**Jenaro Herrera**	**Quebrada Braga**
Forest type	White-sand forest	Terra firme forest	Seasonally-flooded forest
Total landscape area (ha)	4540	4910	3107
Area of forest type (ha)	794	2309	2522
No. pairwise comparisons	315,218	2,665,740	2,412,585

Table 2 Summary of field plot inventory data used to calibrate spectral data at Allpahuayo Mishana (AM) and Jenaro Herrera (JH)

Plot Type	Large rectangular	Small rectangular	Small circular	Large 'Gentry'	Small 'Gentry'
Reference	Vasquez & Phillips 2000; Honorio Coronado et al. 2008	Zarate et al. 2006	Baldeck et al. 2016	Baraloto et al. 2011	(Phillips et al. 2003a)
Site	AM & JH	AM	AM	AM & JH	AM
No. plots	15	10	14	4	6
Plot area (ha)	0.5-1.5	0.1	0.1-0.25	0.5	0.1
Min. dap (cm)	5/10	2.5	10	2.5	2.5
Mean individuals (per plot)	663	358	88	242	260
Mean identified species (per plot)	139	89	33	78	79

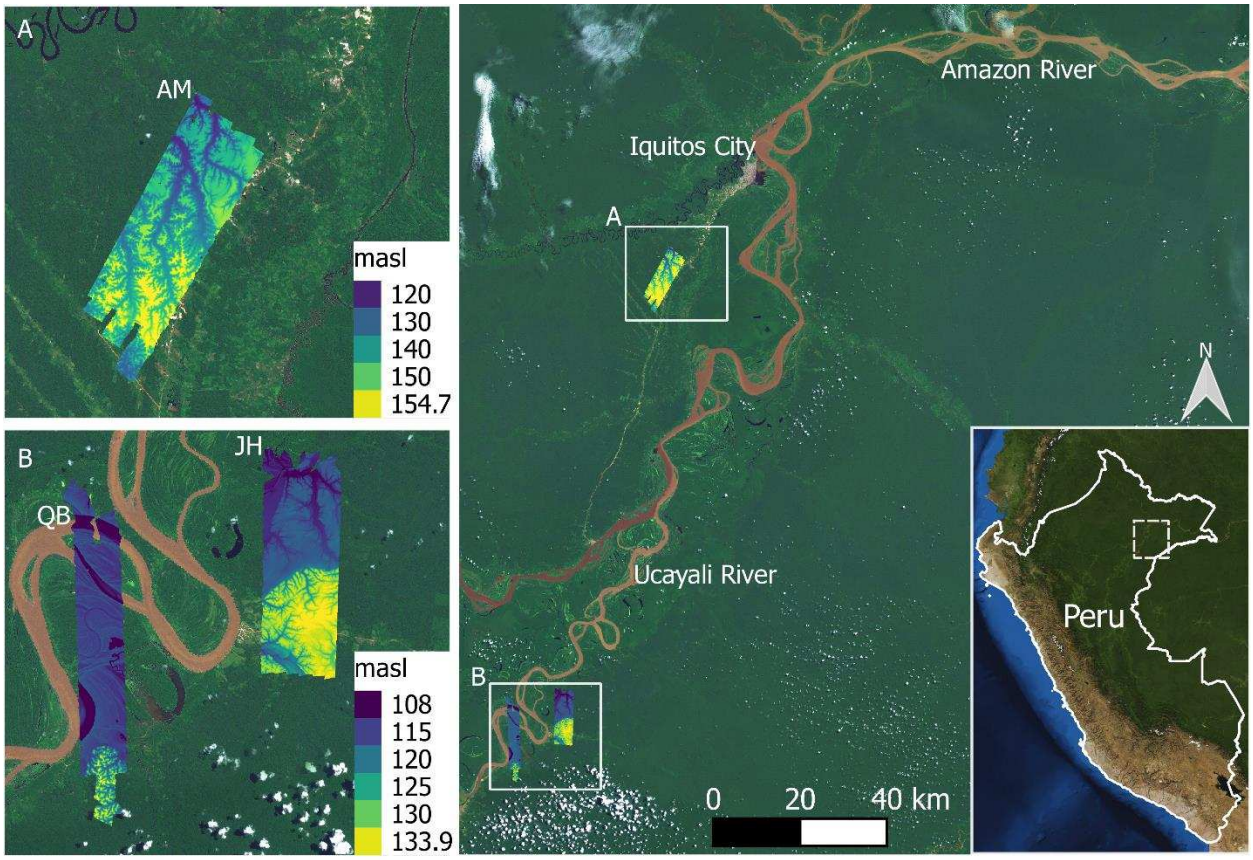

Figure 1. Maps of the three study landscapes: Allpahuayo Mishana (AM), Quebrada Braga (QB) and Jenaro Herrera (JH). Inset maps A and B show the immediate surroundings of the study landscapes as well as the CAO LiDAR-derived digital terrain models for each landscape. The third inset map shows the wider study region (dashed white box) in the context of Peru.

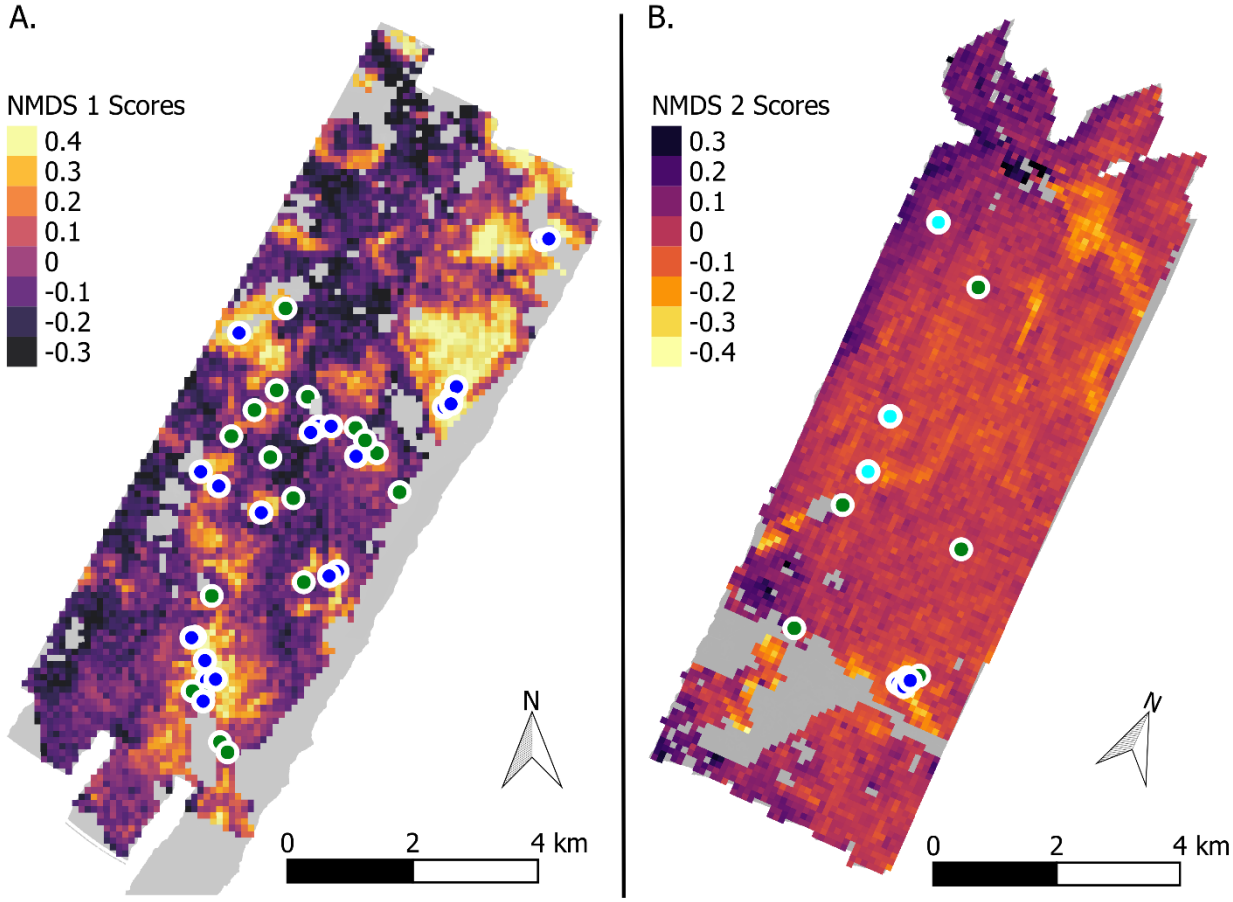

Figure 2 Distribution of field plots across the Allpahuayo Mishana landscape (panel A) and the Jenaro Herrera landscape (panel B). Blue circles represent plots in white-sand forest, green circles represent terra firme forest plots and cyan represent palm swamp forest plots. The backdrop of the map shows the first NMDS axis of the estimated species composition of Allpahuayo Mishana and the second NMDS axis of the estimated species composition of Jenaro Herrera, derived from airborne imaging spectroscopy.

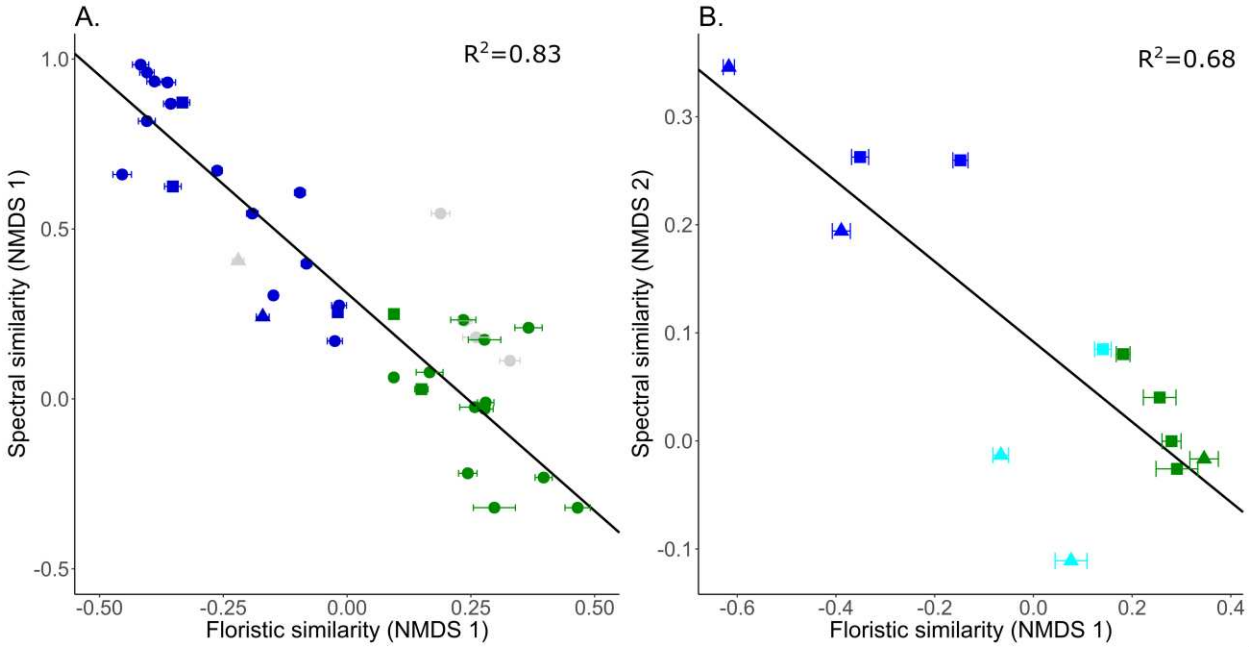

Figure 3 The relationship between spectrally derived estimates of tree species compositional turnover (represented by the first axis of the NMDS ordination of spectral species) and measured tree species compositional turnover (represented by the first axis of the NMDS ordination of tree species) at Allpahuayo Mishana (panel A.) and Jenaro Herrera (panel B.). Colours represent different forest types: Dark blue (white-sand forests); green (terra firme clay forests); cyan (palm swamp forests, grey symbols were those excluded from the analysis are they were < 10 m from a border between forest types. Error bars signify 95 % confidence intervals around floristic NMDS axis scores. Symbol shape corresponds to size of forest census plots, square (1 – 1.5 ha), triangle (0.5 ha), and circle (0.1 – 0.25 ha). Black lines represent linear regressions, both regressions were highly significant ($P \leq 0.001$).

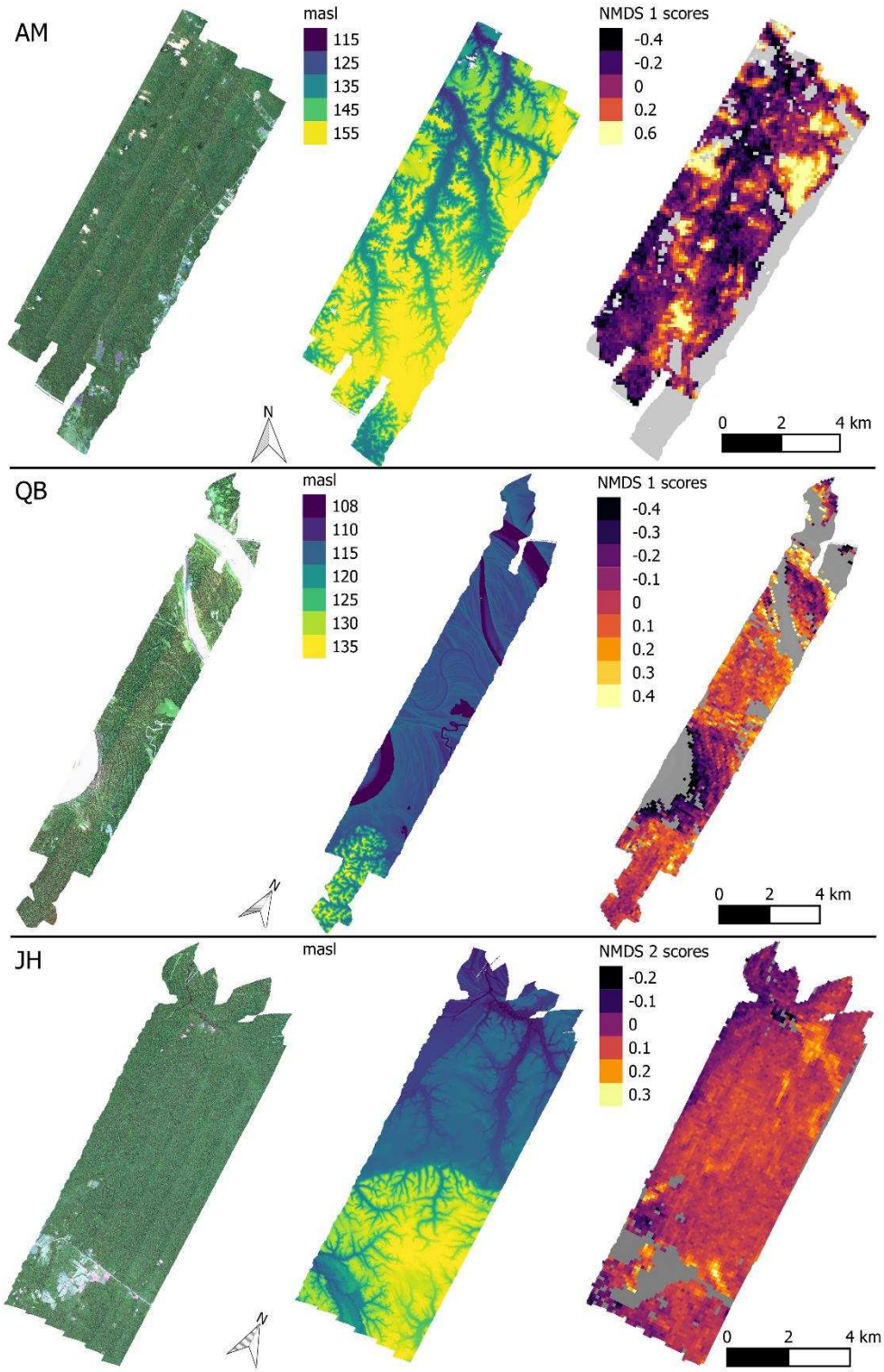

Figure 4 Maps of the three study landscapes, Allpahuayo Mishana (AM), Jenaro Herrera (JH) and Quebrada Braga (QB), The maps show RGB true colour (column 1), LiDAR-derived elevation (column 2) and spectrally-derived estimates of tree species composition, summarised by a single NMDS axis (column 3).

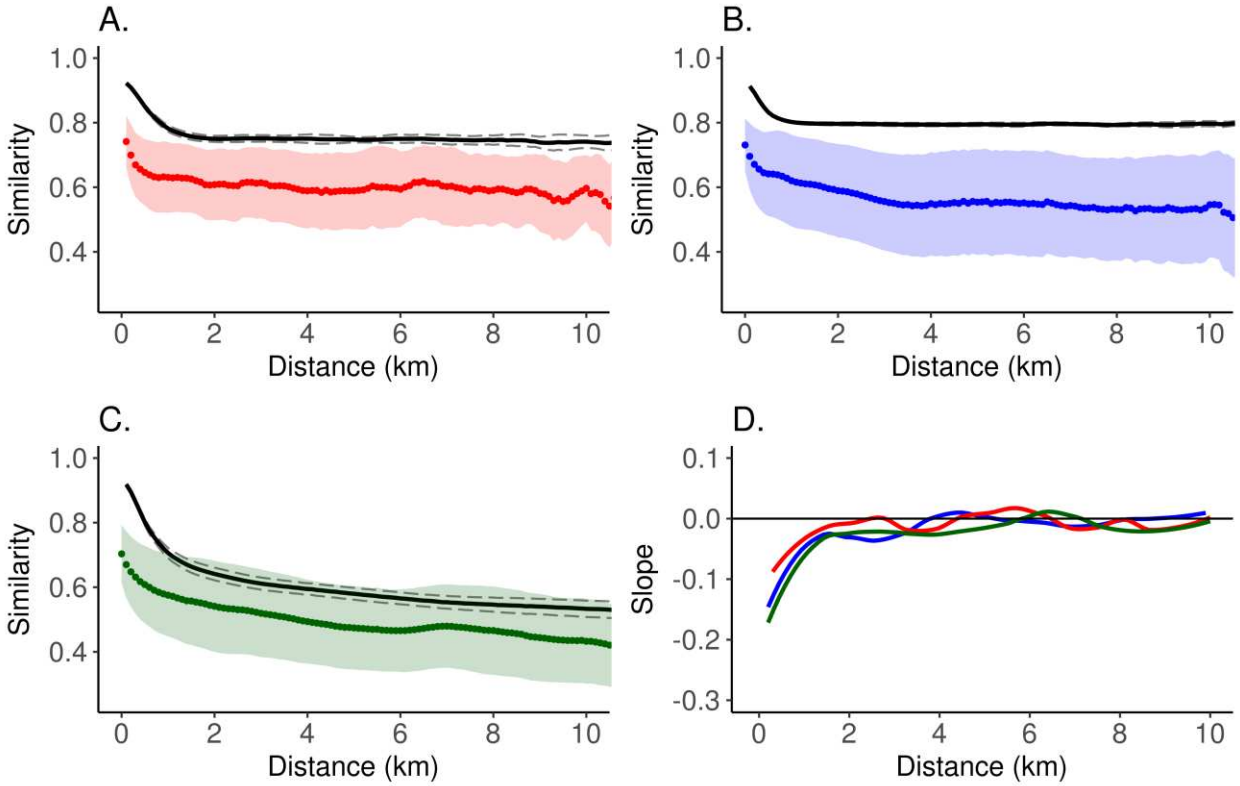

Figure 5. Distance decay relationships in three examples of forest types in the three different landscapes: white-sand forests at Allpahuayo Mishana (panel A); terra firme forest at Jenaro Herrera (panel B); seasonally-flooded forest at Quebrada Braga (Panel C). Points indicate mean Bray-Curtis indices of similarity every 100 meters, and shaded areas are the standard deviations surrounding each 100m point. Panel D shows the loess smoothed line (span=0.35) through the first order derivative, calculated every 100m at each site. Colours correspond to different landscapes/forest types: red = Allpahuayo Mishana white-sand, blue = Jenaro Herrera terra firme clay, green = Quebrada Braga seasonally-flooded. Solid black lines indicate the mean PCP theoretical predicted distance decays, and dashed black lines the standard deviations surrounding these means.