Received: 7 April 2021 | Revised: 6 August 2021 | Accepted: 5 September 2021

DOI: 10.1111/jvs.13081

RESEARCH ARTICLE

Journal of Vegetation Science IAVS

Revealing floristic variation and map uncertainties for different plant groups in western Amazonia

Gabriela Zuquim[1,2] | Hanna Tuomisto[2] | Pablo P. Chaves[2] | Thaise Emilio[3] | Gabriel M. Moulatlet[4] | Kalle Ruokolainen[2] | Jasper Van doninck[2,5] | Henrik Balslev[1]

[1]Ecoinformatics and Biodiversity, Department of Biology, Aarhus University, Aarhus, Denmark

[2]Amazon Research Team, Department of Biology, University of Turku, Turku, Finland

[3]Programa Nacional de Pós-Doutorado, Programa de Pós-Graduação em Ecologia, Institute of Biology, University of Campinas, Campinas, Brazil

[4]Facultad de Ciencias de la Tierra y Agua, Universidad Regional Amazónica Ikiam, Tena, Ecuador

[5]Spatial & Community Ecology Lab, Department of Integrative Biology, Michigan State University, East Lansing, Michigan, USA

Correspondence
Henrik Balslev, Ecoinformatics and Biodiversity, Department of Biology, Aarhus University, Aarhus, Denmark.
Email: henrik.balslev@bio.au.dk

Funding information
Natur og Univers, Det Frie Forskningsråd, Grant/Award Number: 4181–00158 and 9040-00136B; Academy of Finland, Grant/Award Number: 139959 and 273737

Co-ordinating Editor: Alessandra Fidelis

Abstract

Questions: Understanding spatial variation in floristic composition is crucial to quantify the extent, patchiness and connectivity of distinct habitats and their spatial relationships. Broad-scale variation in floristic composition and the degree of uniqueness of different regions remains poorly mapped and understood in several areas across the globe. We here aim to map vegetation heterogeneity in Amazonia.

Location: Middle Juruá river region, Amazonas State, Brazil.

Methods: We mapped four plant groups by applying machine learning to scale up locally observed community composition and using environmental and remotely sensed variables as predictors, which were obtained as GIS layers. To quantify how reliable our predictions were, we made an assessment of model transferability and spatial applicability. We also compared our floristic composition map to the official Brazilian national-level vegetation classification.

Results: The overall performance of our floristic models was high for all four plant groups, especially ferns, and the predictions were found to be spatially congruent and highly transferable in space. For some areas, the models were assessed not to be applicable, as the field sampling did not cover the spectral or environmental characteristics of those regions. Our maps show extensive habitat heterogeneity across the region. When compared to the Brazilian vegetation classification, floristic composition was relatively homogeneous within dense forests, while floristic heterogeneity in rainforests classified as open was high.

Conclusion: Our maps provide geoecological characterization of the regions and can be used to test biogeographical hypotheses, develop species distribution models and, ultimately, aid science-based conservation and land-use planning.

KEYWORDS
Amazonian biogeography, area of applicability, ferns, Juruá river, machine learning, Melastomataceae, niche, palms, plant community, remote sensing, species–environmental relationships, tropical forests, vegetation mapping, Zingiberales

1 | INTRODUCTION

Amazonian forests are highly heterogeneous, but the broad-scale variation in floristic composition and degree of uniqueness of different regions remains poorly mapped and understood. Improved mapping aids in identifying areas that share a unique common pool of species, which are integral to systematic conservation planning (Guisan & Zimmermann, 2000; Margules & Pressey, 2000). Vegetation types are central units in conservation planning (Capobianco et al., 2001) as one of the goals is to collectively represent proportions of each type of forest in the conservation units network. This is based on the assumption that the mapped and actual field-observed distributions of species are strongly linked (Ferrier & Watson, 1997; Emilio et al., 2010). The baseline vegetation data are often from few plant groups or structural components, such as canopy trees. Understorey plants, for example, are not included in forest mapping and are not directly observed by satellite images.

In biogeographical research, vegetation maps are an intuitive source of ecological information. They help to test biogeographical hypotheses and to quantify the extent of patchiness, connectivity of various habitats and their spatial relationships, which are relevant to understand species distributions and metapopulation dynamics (Hanski, 1998). For example, the riverine barrier hypothesis (Haffer, 1974; Boubli et al., 2015; Naka & Brumfield, 2018) has been influencing Amazonian biogeography studies since the 19th century (Wallace, 1852) and remains to some degree controversial, being frequently re-assessed as new data are accumulated (Fordham et al., 2020). For testing the riverine barrier hypothesis, it is important to consider the alternative hypothesis that a difference in species composition across the river emerges because the river banks represent different habitats (Tuomisto & Ruokolainen, 1997). An accurate map of vegetation can provide relevant information on habitat quality, as has already been shown in at least two Amazonian biogeographical studies (Maximiano et al., 2020).

Commonly used indicators of habitat variation include vegetation maps, which can vary greatly in quality and resolution. Vegetation maps may be periodically updated, often absorbing new products from the remote-sensing revolution (Kwok, 2018). Vegetation maps covering all of Brazilian Amazonia, for example, were initially based on radar images and field surveys of trees during the RADAMBRASIL project (RADAMBRASIL, 1978), but recent versions also incorporate data provided by the Landsat programme (IBGE, 2004, 2010; USGS, 2021). On the other hand, information from field data and vegetation inventories accumulate at a slower pace. Thus far, tree inventories cover an area of only 0.0013% of the total forest area of Brazilian Amazonia (Tejada et al., 2019) and the sampling suffers from spatial biases and large geographical gaps (Nelson et al., 1990; Hopkins, 2007; Schulman et al., 2007). The Instituto Brasileiro de Geografia e Estatística (IBGE) map is widely used in land-use planning (Roriz et al., 2017), and even though it has been updated using ancillary information, its information is still based on the same field inventories of canopy trees that were carried out in the 1970s (RADAMBRASIL, 1978). These limitations in vegetation mapping hamper our detailed understanding of habitat patchiness and connectivity, species distributions, metapopulation dynamics and other biogeographical and evolutionary processes.

To improve vegetation mapping and to understand its relation to remotely sensed spectral data and other GIS-based environmental descriptors, we here modelled relationships between compositional, environmental and spectral variation and made spatial predictions of floristic composition in western Amazonia. Species' responses to historical and environmental factors may vary among biological groups (Dambros et al., 2020) and can, therefore, be expected to vary also in relation to spectral data. We addressed this possibility by combining standardized data sets of inventories of phylogenetically distant plant groups with environmental field data and GIS layers from the middle Juruá River region. The Juruá River is one of the longest white-water tributaries of the Amazon River and the vegetation is highly heterogeneous, shaped by variation in flooding regimes (Hawes et al., 2012; Newton, Endo et al., 2012) and soil characteristics (Tuomisto et al., 2016, 2019). High plant species turnover in non-flooded forests is associated with geological history and soil heterogeneity (Schobbenhaus et al., 2004; Higgins et al., 2011; Tuomisto et al., 2016, 2019), which are reflected in a gradient of variation in floristic composition from habitats of clayey soils with higher concentrations of nutrients (especially exchangeable base cations) to sandier and relatively nutrient-poor soil areas (Tuomisto et al., 2016).

Besides the importance of describing the forests, detailed vegetation mapping may be of particular interest in the middle Juruá area due to the existence of community-based management programmes for sustainable production of non-timber forest products (Newton et al., 2011; Newton, Endo et al., 2012; Campos-Silva et al., 2018). Natural resources vary according to vegetation characteristics and are non-randomly distributed in the region (Newton, Peres et al., 2012). Patterns of density and spatial distribution of useful species and forest resources affect local community livelihood, hunting and harvesting strategies (Newton, Endo et al., 2012; Newton, Watkinson et al., 2012).

We aimed to map plant community composition in the Juruá River basin by scaling up locally defined relationships between community composition and remotely sensed spectral data and environmental descriptors. For areas beyond the sampled environmental gradient, such extrapolation may be problematic. Therefore, we made an assessment of the reliability and transferability of the spatial predictions of floristic composition for four plant groups with different morphological traits and dispersal strategies. To understand congruencies and variations with vegetation classification, we also compared how our floristic maps relate to the Brazilian map of vegetation classes provided by IBGE (2010).

2 | METHODS

2.1 | Study area and sampling strategy

Our study area in western Brazilian Amazonia comprised non-flooded forests in the middle Juruá area, including the lower Tarauacá River (Figure 1). The climate is mainly tropical and humid

(IBGE, 2004), with an average annual temperature of 27.1°C and average annual rainfall of approximately 3,700 mm (Hawes & Peres, 2016). The sampled area covered the Solimões and Içá geological formations (Schobbenhaus et al., 2004), with widely contrasting topography and soil characteristics (Tuomisto et al., 2016, 2019). We refer here to the 2004 geological map, as the more recent geological map (IBGE, 2010) does not recognize any boundary at this location.

We established 71 transects along the Juruá River and its largest tributary, the Tarauacá River (Figure 1). The field expedition consisted of two phases. In the first phase, palms (Arecaceae), ferns + lycophytes (pteridophytes), gingers (Zingiberales), and melastomes (Melastomataceae) were inventoried in 39 transects along the Juruá River over an air distance of 500 km. In the second phase, ferns + lycophytes and melastomes where inventoried in 32 transects along the same stretch of Juruá and along the Tarauacá River. To obtain comparable results for the four plant groups, we used the 39 first-phase transects to model the relationships between plant communities and several explanatory variables and the 32 second-phase transects separately as a validation data set (see Section 2.3 *Analysis*).

Each transect was 5 m wide and 500 m long and it crossed the local topography in order to obtain a representative sample of the hydrological variation at the site. The botanical inventories consisted of counting the number of individuals per species for each of the target plant groups. Representative voucher specimens were collected of each species and additional specimens were collected of individuals that could not be referred with certainty to an already collected species. Composite surface soil samples were taken at three points along each transect and analysed for several chemical variables, of which we here use the concentration of exchangeable base cations (Ca, Mg, K and Na). Further details on the field sampling protocol and soil sample analysis are available in Tuomisto et al. (2016).

2.2 | Predictors estimated from global maps and remote sensing (GIS layers)

We obtained GIS layers of height above the nearest drainage (HAND), cation exchange capacity of the surface soil (CEC), and reflectance values of the Landsat Thematic Mapper/Enhanced

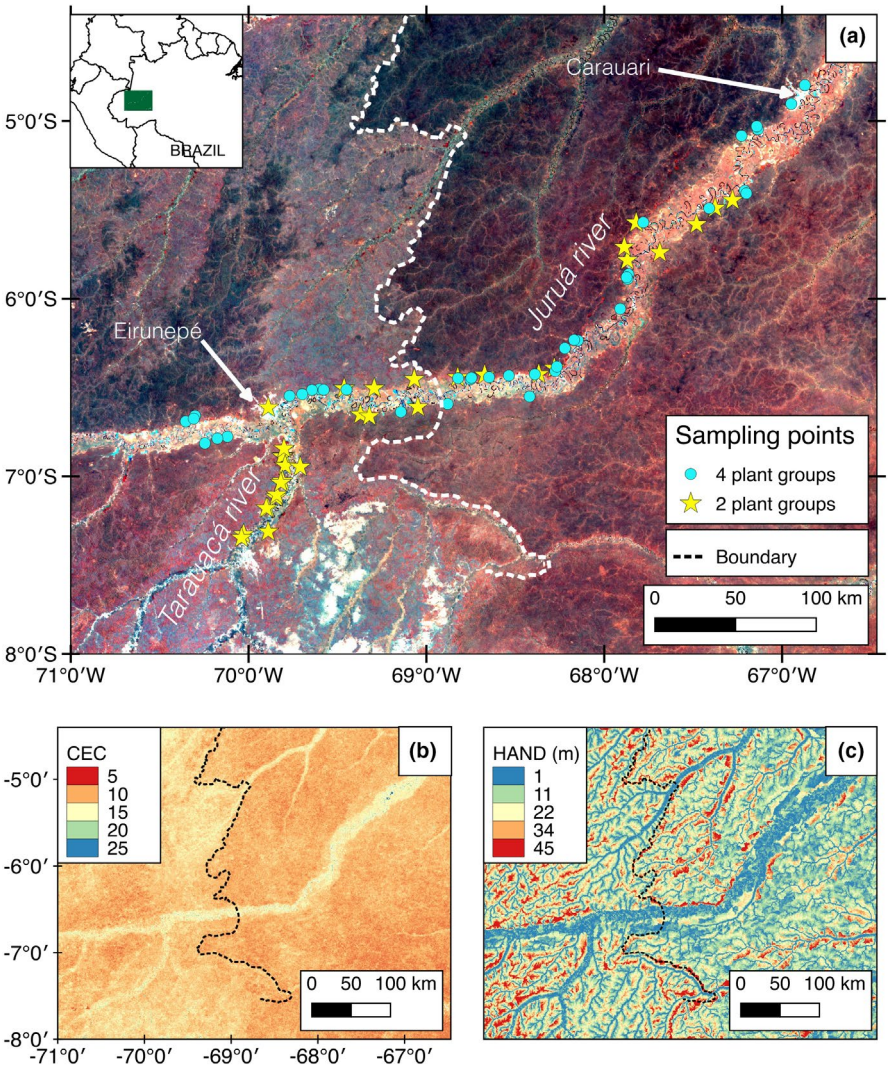

FIGURE 1 Maps of the study area in the middle reaches of the Juruá River. (a) Landsat TM/ETM+ reflectance variation. Bands 4, 5, and 7 were assigned to red, green, and blue colour channels, respectively. Blue circles represent locations in which all four plant groups (palms, ferns + lycophytes, melastomes, and gingers) were sampled; yellow stars represent sampling locations in which only ferns + lycophytes and melastomes were sampled. An inset map of northern South America shows the area of interest in a green rectangle. (b) Map of Cation Exchange Capacity (CEC) in cmol(+)/kg obtained from Soilgrids 250 m (soilgrids.org, Hengl. et al. 2017). (c) Height Above the Nearest Drainage (HAND; http://www.dpi.inpe.br/Ambdata/English/). In all maps, the dashed line corresponds to the boundary identified by Higgins et al. (2011) and interpreted by them as the limit between the Solimões (Pebas) and Içá geological formations

Thematic Mapper Plus (TM/ETM+) bands (Figure 1). HAND estimates the vertical distance of each transect to the nearest drainage, serving as an indicator of local hydrological conditions (Rennó et al., 2008; Nobre et al., 2011), and is available at 90-m resolution. We obtained the HAND model from http://www.dpi.inpe.br/Ambdata/English/download.php), based on a threshold of 50 pixels for the drainage area contribution to include smaller drainages (Rennó et al., 2008). We obtained soil CEC at pH = 7 as estimated for the surface soil (maximum depth 30 cm) from the SoilGrids project (Hengl et al., 2017; https://maps.isric.org) at 250-m resolution. To obtain a single value of CEC and HAND for each transect, we used the median value of the corresponding pixels that fell within a buffer of 200 m around the centre point of the transect. This was aimed at capturing most of the variation along the 500-m-long main axis of the transect without too much lateral coverage. The cation exchange capacity values were logarithmically transformed before analysis. Landsat TM/ETM+ reflectance values were obtained from a 30-m-resolution Landsat composite of images from 2000 to 2009 covering all Amazonia (Van doninck & Tuomisto, 2018). We obtained a single value for each band to represent the whole transect by calculating the median value from a 15 × 15 pixels window centred on the transects. This was done separately for each spectral band used in this study (bands 2, 3, 4, 5, and 7).

2.3 | Analysis

2.3.1 | Ordinations and floristic composition

Differences in floristic composition between transects were summarized in a three-dimensional ordination space using non-metric multidimensional scaling (NMDS). We used extended Sørensen dissimilarities (De'ath, 1999) based on presence/absence data as a dissimilarity measure. Separate dissimilarity matrices were calculated for each plant group using the 39 transects of the first fieldwork phase. A combined dissimilarity matrix was derived by summing the dissimilarity values of each of the four plant groups (hereafter called the "all" dissimilarity matrix). NMDS score values for the first three axes were used as response variables in a multiple regression model with values of locally collected soil nutrient concentration (sum of exchangeable base cations) and side of riverbank as predictors.

2.3.2 | Random Forest (RF) regression models, variable selection and spatial predictions

Non-metric multidimensional scaling ordination scores were interpreted as plot positions along floristic gradients and used as response variables in Random Forest (RF) regression models. RF is a powerful machine-learning algorithm that operates by constructing a multitude of regression trees and combining them to reach a mean prediction (Breiman, 2001). The RF regression analyses were performed separately for each plant group and for all groups together using the ordinations scores from the NMDS analyses based on 39 transects. As predictors, we used variables extracted from seven GIS layers for each transect. The variables were Landsat bands 2–5 and 7, HAND, and CEC.

To eliminate highly related variables and avoid model overfitting, we used forward feature selection (Meyer et al., 2018). We first produced all possible combinations of two variables and used each pair to build an RF regression model. The performance of the models was tested by 10-fold cross-validation, which consists of iteratively dividing the data set into 10 folds and using nine folds for modelling and one fold to test the model. The pair of predictors that produced the RF model with the highest R^2 in the test data set was kept and additional predictors were iteratively added until they no longer significantly improved the performance of the model. The best set of predictors of the ordination scores were determined separately for each of the first three NMDS ordination axes (1, 2, and 3) and each plant group. We generated 1,500 trees for each RF model and allowed different numbers of candidate predictors per node, but in order to make the results more comparable among models with different initial numbers of predictors, we used the results of models with two predictors per node.

To spatialize the predictions to the middle Juruá area, we applied the best GIS layer model of each plant group for predicting NMDS axes over the area of interest. The potential predictors of floristic composition were Landsat bands 2–5 and 7, HAND and CEC, but only the variables selected by RF were included in spatial predictions. For the spatial predictions, Landsat data were upscaled to 450-m resolution, which was obtained by assigning the median value from 15 × 15 original 30-m pixels to each new pixel.

2.3.3 | Map quality and applicability assessment

In addition to using 10-fold random cross-validation to calculate R^2 values for the RF models, we used two other strategies to assess the quality of the predicted NMDS maps. The second approach aimed to quantify model transferability by using a separate validation data set that consisted of 32 inventories of ferns + lycophytes and melastomes sampled in the second phase of the fieldwork (yellow stars in Figure 1). For this approach, we first produced new NMDS ordination axes for these two plant groups using all 71 transects. Then we used the NMDS scores of the 39 transects from the first phase of the fieldwork as training data to build the RF models and spatial predictions and the remaining 32 transects as test data.

The third quality-testing approach was to estimate the area of applicability (AOA) for each model. These are the areas for which the models are expected to provide reliable predictions, and we determined AOA using the method proposed by Meyer and Pebesma (2020). This is based on the environmental distance between a point to be predicted, k, and the most similar one of those points in the training data set that are not in the same cross-validation fold. This distance (d_k) is then divided by the mean distance between two points of the training data set (d_{mean}). The model is deemed applicable to

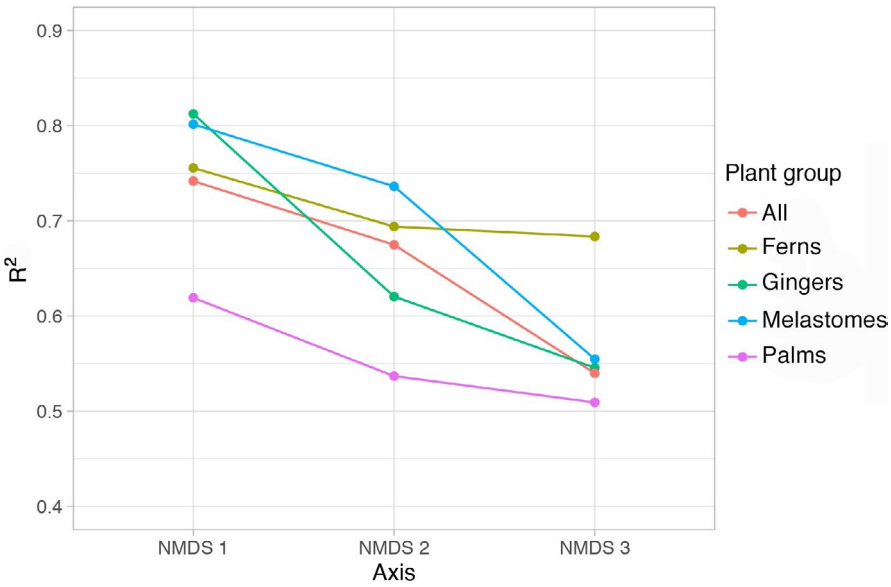

FIGURE 2 Comparison of cross-validation R^2 values of the best Random Forest (RF) regression models for predicting values of each non-metric multidimensional scaling (NMDS) ordination axis. Models were produced separately for each plant group and for all groups analysed together. The best models were defined as the ones with the highest R^2 using forward feature selection

those points within the area of interest for which the ratio d_k/d_{mean} is smaller than a threshold value (Meyer & Pebesma, 2020). The AOA maps are, therefore, binary maps that indicate where the predictions can be considered reliable. We generated an AOA map for each of the five predicted maps (four based on each plant group separately and one based on all plant groups combined).

2.4 | Cross-taxa comparisons of predicted floristic gradients

To quantify the correlation of floristic patterns among palms, ferns + lycophytes, melastomes, gingers, and all plant groups combined, we compared the predicted NMDS axis values for each group in 10,000 randomly sampled locations. We masked out the areas where the models for any of the plant groups were indicated not to be applicable and thereby restricted the random sampling to those areas where all models were applicable. These comparisons were made for each of the three NMDS axes separately.

2.5 | Comparison with Brazilian vegetation classes

Finally, we compared the floristic composition maps obtained with the predicted values of the first NMDS axis with a vegetation map of the Brazilian Amazonas state (IBGE, 2010). The map consists of polygons that depict the dominant vegetation type following the classification of Veloso et al. (1991), which uses as classification criteria, among other things, canopy characteristics (even, uneven, closed, open, with or without emergent trees or with dominance of palms, bamboos or lianas) and geomorphological characteristics (e.g. origin of sediments, elevation; IBGE, 2012). To quantify heterogeneity within IBGE polygons, we evaluated the variation in predicted floristic (NMDS axis) values in our maps for each IBGE vegetation class. The IBGE map was obtained from the Brazilian Institute of Geography and Statistics (IBGE) geoserver (https://geoservicos.ibge.gov.br/geoserver/web/; accessed in February 2021). When reporting the results, we will use the IBGE acronyms for the vegetation classes exactly as they are in the original publication, but translating and interpreting the Portuguese names into English with respect to flooding regime and geomorphology in the Juruá region. For example, the class "Dae", which is described as "floresta ombrófila densa aluvial com dossel emergente", corresponds to the forests growing in the terraces of the Juruá River.

All analyses were carried out using R software (R Core Team, 2020), and the final layouts of maps were designed in QGIS (QGIS Development Team, 2021).

3 | RESULTS

3.1 | Floristic variation in the Juruá area and explanatory power of the models

The NMDS ordinations had satisfactory stress values for three dimensions (0.12 — palms; 0.05 — ferns + lycophytes; 0.12 — melastomes and 0.11 — gingers) and showed that the species composition can be divided into two floristically distinct groups of transects (Appendix S1). For every plant group and all groups together, multiple linear regressions showed that the first NMDS axes were significantly related to the locally measured concentration of exchangeable base cations in the soil ($p < 0.01$) but not to riverbank ($p > 0.05$). Using the RF models, we found that Pearson R^2 for the best model predicting each axis consistently decreased from axis 1 to axis 3 for every plant group (Figure 2). The highest Pearson R^2 values were found for axis 1 and ranged from 0.62 (palms) to 0.81 (melastomes).

For all four plant groups, the maps of the predicted main gradient in community composition (NMDS axis 1) in middle Juruá revealed clear floristic turnover zones across the area (Figure 3). The maps produced with the four different plant groups were visually similar.

The eastern section was distinct from most of the western section. The floodplains of Juruá and other rivers and the southwestern region also appear to stand out from the rest of the area.

When the models for NMDS axis 2 were spatialized (Appendix S2), the visual differences between the eastern and western areas remained distinct, except for palms and all plants together. Similar NMDS axes values, indicating floristically similar plant communities, could be found on both sides of the Juruá River.

3.2 | Map evaluation, validation, and transferability of the predictions

In addition to cross-validation, we also validated the maps with a separate data set of 32 transects in which ferns + lycophytes and melastomes were inventoried. We found that the correlation between predicted and observed NMDS axis 1 scores was higher for ferns + lycophytes ($r = 0.70$) than for melastomes ($r = 0.55$). For axis 2 the correlation was moderate for both ferns + lycophytes and melastomes ($r = 0.51–0.57$, respectively). The correlations for the third axis were extremely low for both plant groups ($r = 0.08$ and -0.06 for ferns + lycophytes and melastomes, respectively).

When evaluating the areas of applicability (Figure 4), we found that all models of the individual plant groups were applicable to a large part of the middle Juruá. As could be expected given the distribution of the field data, all models were more commonly applicable over the non-inundated areas than over the floodplains. The spatial distribution of applicability areas differed among plant groups, as their models used different sets of explanatory variables, with palms having the largest AOA and gingers the smallest. The area over which all four models were applicable (the intersection map in Figure 4) was considerably smaller than the applicability areas for each individual plant group. The model obtained for the combined data set (all four plant groups put together), in turn, had an intermediate size between those of the individual plant groups.

3.3 | Correlations of predicted floristic variation between plant groups

We assessed whether the maps generated for the four plant groups separately and combined were correlated with each other. For NMDS axis 1, the correlations were high, varying from $r = 0.96$ between ferns + lycophytes and all plant groups combined to $r = 0.64$ between palms and gingers. All correlations between ferns + lycophytes, melastomes, and gingers for NMDS axis 1 were above 0.85, while the correlations involving palms were slightly lower (between 0.76 and 0.64; Figure 5). The correlations involving axes 2 and 3 were moderate to very low (r between 0.50 and 0.03) except between gingers and ferns + lycophytes for NMDS axis 2 ($r = 0.72$; Figure 5).

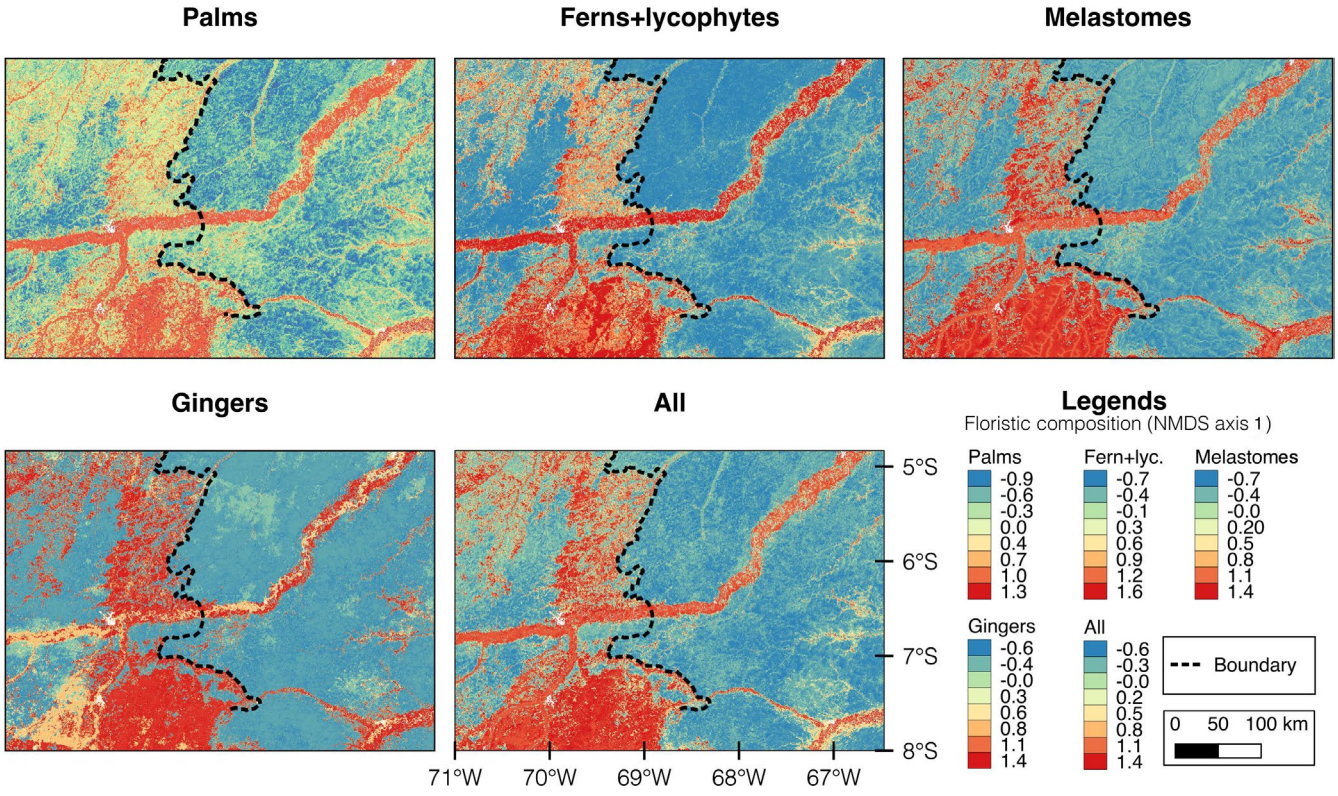

FIGURE 3 Main spatial patterns in the community composition of four different plant groups in the middle Juruá area as represented by the predicted non-metric multidimensional scaling (NMDS) axis 1 scores for presence/absence data. Human settlements were masked out and are shown in white; black dashed line corresponds to the boundary identified by Higgins et al. (2011) and interpreted by them as the limit between the Solimões (Pebas) and Içá geological formations

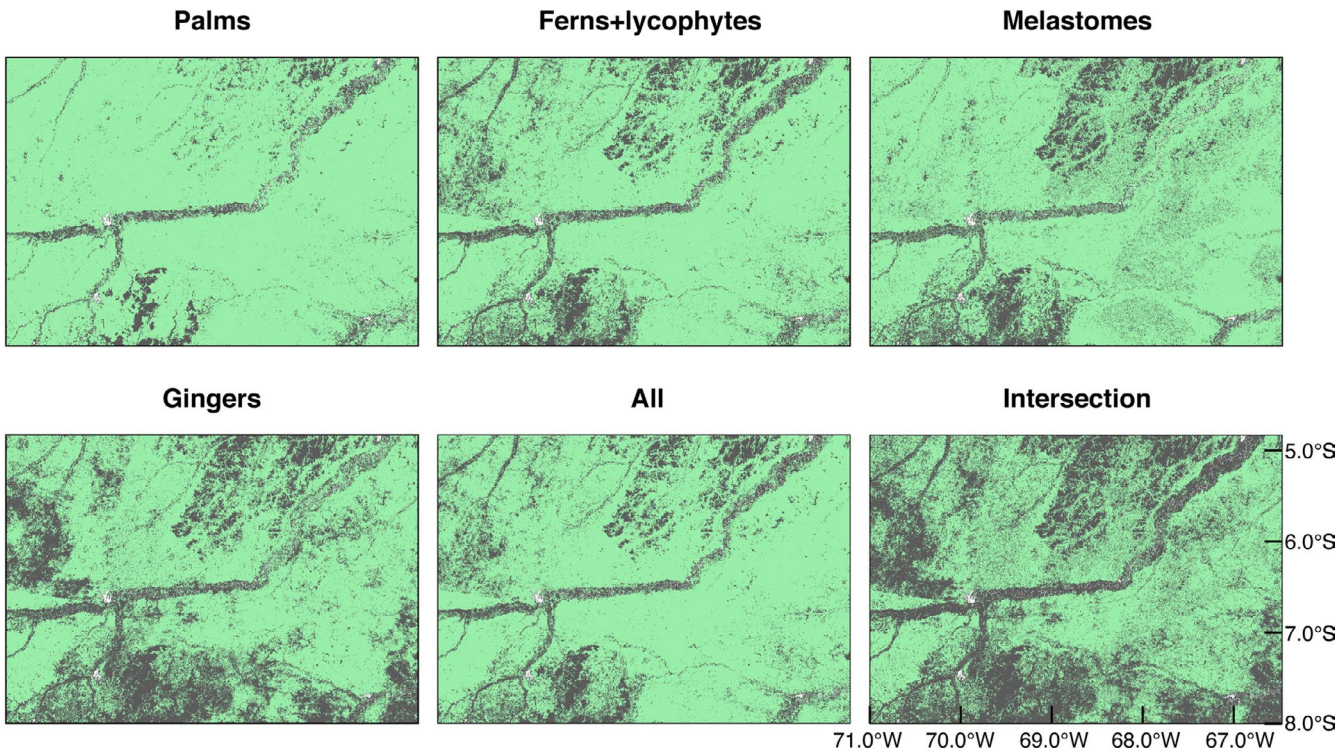

FIGURE 4 Area of applicability (AOA) of the predictions of the main floristic gradient (non-metric multidimensional scaling [NMDS] axis 1 values presented in Figure 3) in the middle Juruá area for palms, ferns + lycophytes, melastomes, gingers and all plant groups combined. The intersection map highlights those areas over which all five models were applicable. Grey corresponds to areas where predictions are not applicable, and green to areas where the predictions are applicable

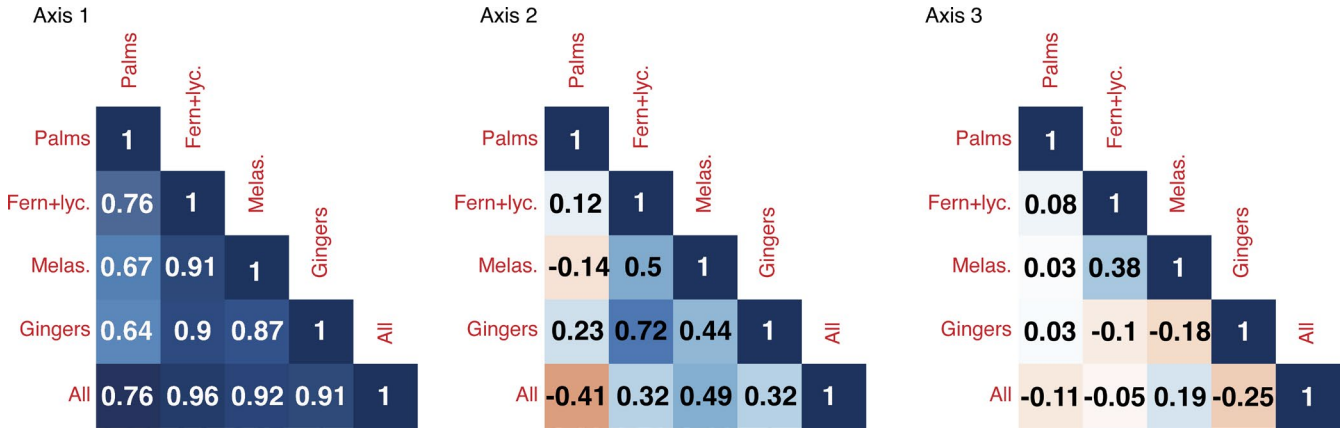

FIGURE 5 Correlations between floristic gradients of four plant groups as represented with scores of non-metric multidimensional scaling (NMDS) axes 1–3 predicted using Random Forest models over the middle Juruá area of Brazilian Amazonia. Predicted values of each axis for each plant group were extracted for 10,000 randomly sampled points from the predictive maps (shown for axis 1 and 2 in Figure 3, and Appendix S2, respectively)

3.4 | Comparison with Brazilian vegetation map

We observed that the degree of floristic heterogeneity predicted by our models within the IBGE dominant vegetation classes varied among the classes (Figures 6a and 7). Open lowland forest with palms (Abp) and open floodplain forest with palms (Aap) were highly heterogeneous classes, each encompassing virtually all the floristic variation observed in the entire area. The remaining classes appeared more homogeneous in floristic composition. Dense lowland forest with an emergent canopy (Dbe — depicted in pink in Figure 6a) overlaps to a large degree with one extreme of the floristic gradient (negative NMDS axis 1 values of our floristic composition map — depicted in blue in Figure 6b). On the western side, parts of the open lowland forest with palms (Abp) overlap with the opposite extreme of the floristic gradient (positive NMDS axis

1 values — depicted in red in Figure 6b). The biggest differences in average floristic composition values were observed between open lowland forest with bamboo (Abb; average NMDS axis 1 score 1.23) and all the dense forest classes (Dae, Dau and Dbe; average NMDS axis 1 scores 0.04, 0.01 and −0.38, respectively; Figure 7).

4 | DISCUSSION

Landscapes are viewed as networks of habitat patches in which species occur as discrete local populations connected by migration (Hanski, 1998). Understanding the spatial distribution of habitats allows predictions of species occurrences and population densities, which are relevant for testing biogeographical hypotheses and can influence strategies for conservation and the use of natural resources. Below we discuss the maps of floristic gradients that were presented above and how they can advance vegetation mapping and biogeographical understanding.

4.1 | Congruence among plant groups

Strong congruence in community composition among plant groups is an indication that there is a common driver of their spatial patterns, which can be understood as habitat variation. We interpreted the strong congruent patterns among the maps for each plant group as a consequence of common species responses to variation in edaphic conditions in the region (Tuomisto et al., 2016, 2019). Maps for palms had less explanatory power than those for the other groups, probably because palm species tend to be more generalists than species of the other groups, and compositional differences would be more apparent with abundance data than with the presence/absence data

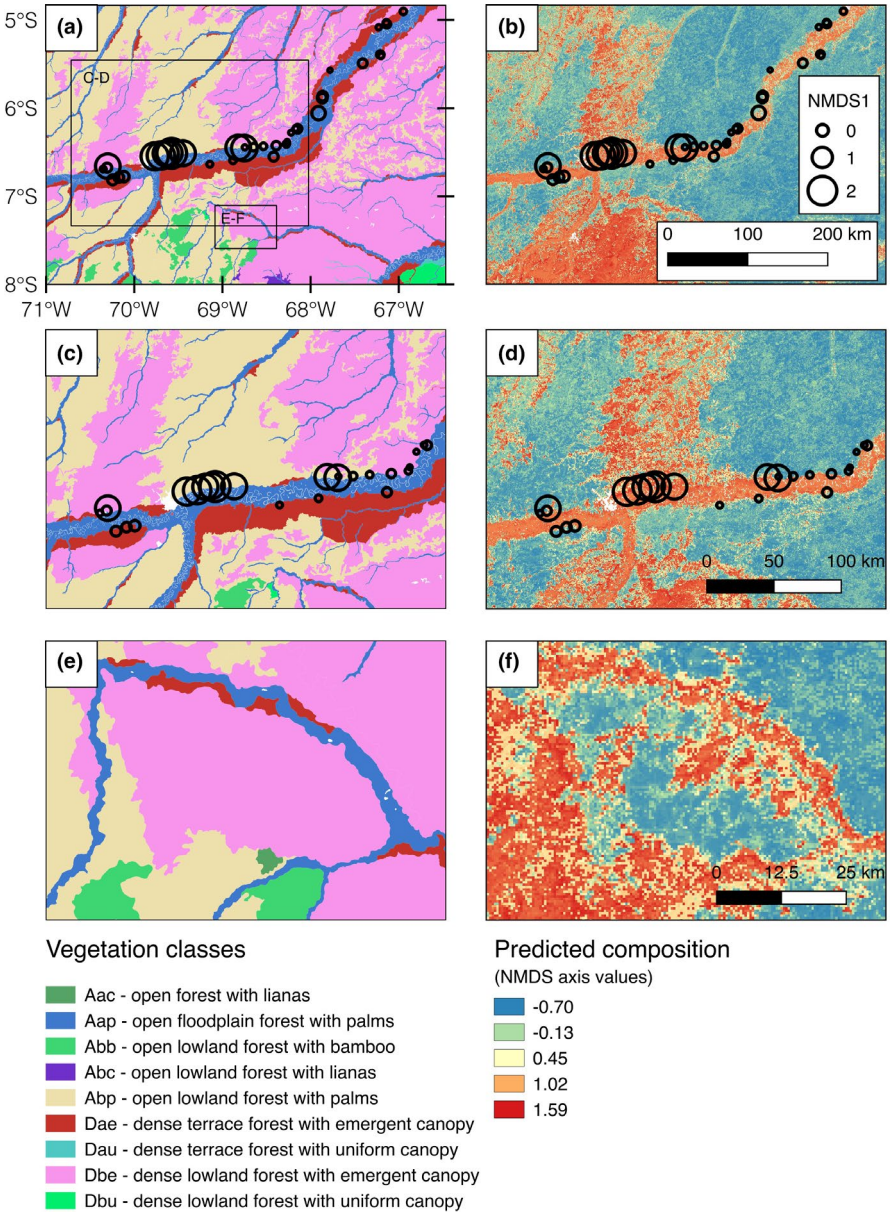

FIGURE 6 Comparison between the Instituto Brasileiro de Geografia e Estatística (IBGE) vegetation map (a, c, e) and the predicted non-metric multidimensional scaling (NMDS) axis 1 scores obtained with all four plant groups together (b, d, f). Each circle corresponds to one floristic inventory site. Circle sizes are scaled according to the observed NMDS axis 1 values, which represent position along the main observed floristic gradient. Acronyms for dominant vegetation classes in the legend are as in the original publication (IBGE, 2012), but the names have been translated on the basis of interpretation with respect to flooding regime and geomorphology in the Juruá area

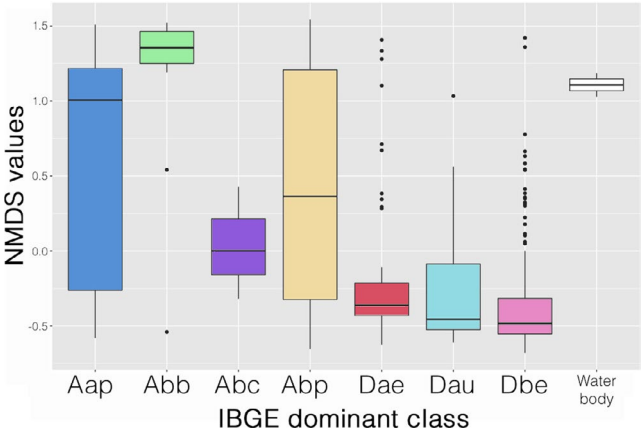

FIGURE 7 Boxplots of the variation in predicted floristic composition values (non-metric multidimensional scaling [NMDS] axis 1) based on all four plant groups together within each dominant vegetation class of IBGE (2012). For full names of the vegetation classes, see Figure 6

used here (Ruokolainen & Vormisto, 2000; Kristiansen et al., 2012; Tuomisto et al., 2016; Cámara-Leret et al., 2017).

4.2 | Assessment of map transferability

Recently, predictive spatial modelling for mapping exercises has become more common (Hengl et al., 2017), but the degree of uncertainty of these maps has rarely been assessed, leading to a loss of confidence in the accuracy of global maps (Meyer & Pebesma, 2020). Explicit documentation of locations where predictions are less certain is crucial, especially in applied fields, such as when prioritizing conservation efforts (Pelletier et al., 2018). The AOA maps revealed that our models were not transferable to the southwestern part of the study area in the Tarauacá river drainage. This could be related to geological or climatic factors or a combination of both. The area has a higher prevalence of the Solimões formation than the main Juruá River drainage and is also in a transition zone towards the more seasonal bamboo forests further south (Carvalho et al., 2013; Tuomisto et al., 2019). The presence of bamboo in the forest clearly changes the reflectance of the canopy (Carvalho et al., 2013; Van doninck et al., 2020), and the gradient of reflectance was probably longer in the SW part of the area of interest than in the training data set.

The distinctness of the Tarauacá River drainage was also reflected in that we found lower predictive power for floristic variation when using a separate 32-transect test data set than when using cross-validation within the 39 transects used for model building. The 39-transect set included no samples from the Tarauacá River, so the predictions for that area were based on field sites with different characteristics. Nevertheless, the relationship between predicted and observed floristic composition for ferns + lycophytes remained relatively high, suggesting that our results were spatially transferable to another subbasin in the larger Juruá region and possibly to other areas with similar environmental conditions.

4.3 | Floristic patterns and biogeographical hypothesis

Our maps were able to reproduce a sharp turnover zone in floristic composition, which has been previously associated with geologically driven changes in soil characteristics (Higgins et al., 2011; Tuomisto et al., 2016, 2019). Our maps also indicate that the boundary is not absolute, but rather that both sides of it are floristically heterogeneous. Consequently, the Juruá River itself runs across a heterogeneous landscape, and knowledge of the configuration of the different habitats is important to understand species distributions and population structures in the area. For example, our maps provide guidance for effective geographical placement of field sampling for the testing of the riverine barrier hypothesis along the Juruá River. As can be seen in Figures 3 and 5, the northern and southern sides of the Juruá differ floristically in the stretch between the mouth of the Tarauacá and the Solimões/Içá geological boundary, which reflects environmental (especially soil) differences (Tuomisto et al., 2016). If this environmental background is overlooked, across-river differences in the species composition or genetic structure of other organisms might erroneously be interpreted as evidence for the riverine barrier hypothesis. To ensure robust biogeographical conclusions, sampling should be planned using appropriate vegetation or floristic maps such that similar habitats are sampled and compared on both sides of the river (Tuomisto & Ruokolainen, 1997).

Earlier studies on small mammals and plants found no evidence that the Juruá River would be a distribution barrier (Gascon et al., 2000; Tuomisto et al., 2016). This might be expected, because the Juruá river channel is narrow, extensively meandering and with slow streamflow compared with rivers that have more often been cited as barriers (Wallace, 1852; Ayres & Clutton-Brock, 1992; Aleixo, 2004; Ribas et al., 2012). However, the Juruá has an extensive floodplain, and recent analyses of taxonomically updated data have revealed more pronounced differences in primate species composition (Fordham et al., 2020) and plumage colour of some understorey birds (Del-Rio et al., 2021) from the non-inundated areas across the river. To what degree these observations are related to habitat variation vs the presence of the river and its floodplain remains an open question. Disentangling potential environmental and historical effects is crucial to clarify the evolutionary history of Amazonian biota, and this requires a proper control of environmental variation (Tuomisto & Ruokolainen, 1997; Maximiano et al., 2020). Here we provide floristic maps that help to address this challenge.

4.4 | Comparison with Brazilian Vegetation map from IBGE

We observed that the degree of floristic heterogeneity predicted by our models within the IBGE dominant vegetation classes varied among the classes. We found that the IBGE open-forest classes were floristically more heterogeneous, which was also previously observed for canopy trees (Emilio et al., 2010). Open lowland forest

with palms (Abp) contained a large variation in floristic composition. This class overlapped to some degree with the forests depicted in red on our maps which are forests growing over soils with a relatively high concentration of exchangeable base cations. Forests on nutrient-rich soils in Amazonia tend to be more dynamic and to have lower trees and more canopy gaps than forests on poorer soils (Higgins et al., 2015). Canopy palms are an important structural component of the Abp forest class, and taller palms seem to prefer habitats with higher soil nutrient concentration (Cámara-Leret et al., 2017). In contrast to the Abp class, the Dbe class (dense lowland forest with emergent canopy) was found to have low tree beta diversity (Emilio et al., 2010) and was also floristically more homogeneous in our maps. This class overlapped to a large degree with areas that in our maps corresponded to negative values in the main floristic composition gradient (NMDS axis 1). This area also corresponds to the Içа formation (Schobbenhaus et al., 2004) with relatively nutrient-poor soils (Higgins et al., 2011; Tuomisto et al., 2016). The Aap class as defined by IBGE (2012) corresponds to forests over floodplains. Because the Juruá is an actively meandering river, its floodplain is a highly dynamic and spatially heterogeneous mosaic of different flooding conditions and sedimentation regimes that support successional vegetation patches of different ages (Salo et al., 1986; Toivonen et al., 2007). Although our floristic sampling did not cover the floodplains, their high heterogeneity was revealed by the satellite imagery.

The congruencies and incongruences between the floristic map and the vegetation classes suggest some degree of correspondence between canopy and understorey plant species composition patterns with dominant vegetation classes but also reveals a high degree of heterogeneity within some of those classes.

5 | CONCLUSION

Our analyses in the middle Juruá area revealed habitat distributions, patch sizes and connectivity using RF models that were assessed to be accurate and transferable to areas beyond the localities of floristic inventories. The mapped floristic patterns were based on GIS predictors that captured variation associated with environmental gradients. To what degree the identified patterns can also reflect turnover of species or genotypes in animal groups remains to be investigated. Some studies have suggested that distribution patterns of plants and animals in Amazonia are, to some degree, congruent (Pomara et al., 2012, 2014; Dambros et al., 2020). Therefore, our floristic maps can be used to formulate expectations for other taxonomic groups on species turnover and distribution patterns related to current environmental variation. These can be compared to predictions based on dispersal barriers or other historical explanations to disentangle their relative roles in driving the current biogeographical patterns in the Juruá region.

Because densities of plant and animal populations with economic importance are also likely to be driven by habitat preferences, habitat mapping has direct implications for community-based resource use and management planning (Newton, Peres et al., 2012; Newton, Watkinson et al., 2012; Campos-Silva et al., 2018) as well as livelihood strategies (Newton, Endo et al., 2012). A spatially explicit approach such as ours is also needed to understand metapopulation dynamics and to support science-based conservation and land-use planning.

ACKNOWLEDGEMENTS

We are thankful to several field assistants and the boat staff who made the fieldwork possible. Research permits were granted by CNPq, CEUC/SDS, and SISBIO. Fernando G. Figueiredo kindly allowed the use of ginger inventory data. Funding was received from the Academy of Finland (grants #139959 and #273737 to HT) and the Danish Council for Independent Research - Natural Sciences (grants #4181–00158 and #9040-00136B to HB). The authors declare to have no conflicts of interest.

AUTHOR CONTRIBUTIONS

GZ, HT, and HB conceived the ideas; HT, KR, HB, TE, and GMM collected the field data; JVd produced the Landsat TM/ETM+ imagery; GZ analysed the data with THE support of PPC; and GZ led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

The codes to reproduce the ordinations, Random Forest models using GIS layers as predictors, spatial floristic predictions, and Area of Applicability maps are available in github (https://github.com/gabizuquim/Floristic-map) accompanied by palm inventory data and post-processed predictor layers as supporting data for the codes.

ORCID

Gabriela Zuquim https://orcid.org/0000-0003-0932-2308
Hanna Tuomisto https://orcid.org/0000-0003-1640-490X
Pablo P. Chaves https://orcid.org/0000-0002-1025-9505
Thaise Emilio https://orcid.org/0000-0001-5415-1822
Gabriel M. Moulatlet https://orcid.org/0000-0003-2571-1207
Kalle Ruokolainen https://orcid.org/0000-0002-7494-9417
Jasper Van doninck https://orcid.org/0000-0003-2177-7882
Henrik Balslev https://orcid.org/0000-0002-7101-7120

REFERENCES

Aleixo, A. (2004) Historical diversification of a terra-firme forest bird superspecies: A phylogeographic perspective on the role of different hypotheses of Amazonian diversification. *Evolution*, 58(6), 1303–1317.

Ayres, J.M. & Clutton-Brock, T.H. (1992) River boundaries and species range size in Amazonian primates. *The American Naturalist*, 140(3), 531–537. https://doi.org/10.2307/2462782

Boubli, J.P., Ribas, C., Lynch Alfaro, J.W., Alfaro, M.E., da Silva, M.N.F., Pinho, G.M. et al. (2015) Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. *Molecular Phylogenetics and Evolution*, 82, 400–412. https://doi.org/10.1016/j.ympev.2014.09.005

Breiman, L. (2001) Random forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Cámara-Leret, R., Tuomisto, H., Ruokolainen, K., Balslev, H. & Munch Kristiansen, S. (2017) Modelling responses of western Amazonian palms to soil nutrients. *Journal of Ecology*, 105(2), 367–381. https://doi.org/10.1111/1365-2745.12708

Campos-Silva, J.V., Hawes, J.E., Andrade, P.C.M. & Peres, C.A. (2018) Unintended multispecies co-benefits of an Amazonian community-based conservation programme. *Nature Sustainability*, 1(11), 650–656. https://doi.org/10.1038/s41893-018-0170-5

Capobianco, J.P.R., Veríssimo, A., Moreira, A., Sawer, D., Ikeda, S. & Pinto, L.P. (2001) *Biodiversidade na Amazônia brasileira: Avaliação e ações prioritárias para a conservação, uso sustentável e repartição de benefícios*. Estação Liberdade.

de Carvalho, A.L., Nelson, B.W., Bianchini, M.C., Plagnol, D., Kuplich, T.M. & Daly, D.C. (2013) Bamboo-dominated forests of the Southwest Amazon: Detection, spatial extent, life cycle length and flowering waves. *PLoS One*, 8(1), e54852. https://doi.org/10.1371/journal.pone.0054852

Dambros, C., Zuquim, G., Moulatlet, G.M., Costa, F.R.C., Tuomisto, H., Ribas, C.C. et al. (2020) The role of environmental filtering, geographic distance and dispersal barriers in shaping the turnover of plant and animal species in Amazonia. *Biodiversity and Conservation*, 29(13), 3609–3634. https://doi.org/10.1007/s10531-020-02040-3

De'ath, G. (1999) Extended dissimilarity: A method of robust estimation of ecological distances from high beta diversity data. *Plant Ecology*, 144(2), 191–199. https://doi.org/10.1023/A:1009763730207

Del-Rio, G., Mutchler, M.J., Costa, B, Hiller, A.E, Lima, G, & Matinata, B. et al. (2021). Birds of the Juruá River: extensive várzea forest as a barrier to terra firme birds. *J Ornithol*–162, 565–577. https://doi.org/10.1007/s10336-020-01850-0

Emilio, T., Nelson, B.W., Schietti, J., Desmoulière, S.-J.-M., Espírito Santo, H.M.V. & Costa, F.R.C. (2010) Assessing the relationship between forest types and canopy tree beta diversity in Amazonia. *Ecography*, 33(4), 738–747. https://doi.org/10.1111/j.1600-0587.2009.06139.x

Ferrier, S. & Watson, G. (1997) *An evaluation of the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity*. Environment Australia. Canberra: NSW National Parks and Wildlife Services.

Fordham, G., Shanee, S. & Peck, M. (2020) Effect of river size on Amazonian primate community structure: A biogeographic analysis using updated taxonomic assessments. *American Journal of Primatology*, 82(7). e23136. https://doi.org/10.1002/ajp.23136

Gascon, C., Malcolm, J.R., Patton, J.L., da Silva, M.N., Bogart, J.P., Lougheed, S.C. et al. (2000) Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences*, 97(25), 13672–13677.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2), 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Haffer, J. (1974) *Avian speciation in tropical South America, with a systematic survey of the toucans (Ramphastidae) and jacamars (Galbulidae).* , 1st edition. Cambridge: Nuttall Ornithological Club.

Hanski, I. (1998) Metapopulation dynamics. *Nature*, 396, 41–49.

Hawes, J.E., & Peres, C.A. (2016). Patterns of plant phenology in Amazonian seasonally flooded and unflooded forests. *Biotropica*, 48, 465–475. https://doi.org/10.1111/btp.12315

Hawes, J.E., Peres, C.A., Riley, L.B. & Hess, L.L. (2012) Landscape-scale variation in structure and biomass of Amazonian seasonally flooded and unflooded forests. *Forest Ecology and Management*, 281, 163–176. https://doi.org/10.1016/j.foreco.2012.06.023

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., & Blagotić, A. et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, 12(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Higgins, M.A, Asner, G.P, Anderson, C.B, Martin, R.E, Knapp, D.E, & Tupayachiet, R. et al. (2015). Regional-Scale Drivers of Forest Structure and Function in Northwestern Amazonia. *PLOS ONE*, 10(3), e0119887. https://doi.org/10.1371/journal.pone.0119887

Higgins, M.A., Ruokolainen, K., Tuomisto, H., Llerena, N., Cardenas, G., Phillips, O.L. et al. (2011) Geological control of floristic composition in Amazonian forests. *Journal of Biogeography*, 38(11), 2136–2149. https://doi.org/10.1111/j.1365-2699.2011.02585.x

Hopkins, M.J.G. (2007) Modelling the known and unknown plant biodiversity of the Amazon Basin. *Journal of Biogeography*, 34(8), 1400–1411. https://doi.org/10.1111/j.1365-2699.2007.01737.x

IBGE (2004) *Mapa de Vegetação do Brasil 1:5,000,000.* , 3rd edition. IBGE - Instituto Brasileiro de Geografia e Estatística. fttp://geoftp.ibge.gov.br/mapas/tematicos/mapas_murais

IBGE (2010) *Estado do Amazonas—Geologia [Map]*. Instituto Brasileiro de Geografia e Estatística. https://mapas.ibge.gov.br/tematicos/geologia

IBGE (2012) *Manual técnico da vegetação Brasileira* (2a edição). Instituto Brasileiro de Geografia e Estatística - IBGE.

Kristiansen, T., Svenning, J.-C., Eiserhardt, W.L., Pedersen, D., Brix, H., Munch Kristiansen, S. et al. (2012) Environment versus dispersal in the assembly of western Amazonian palm communities. *Journal of Biogeography*, 39(7), 1318–1332. https://doi.org/10.1111/j.1365-2699.2012.02689.x

Kwok, R. (2018) Ecology's remote-sensing revolution [News]. *Nature*, 556(7699), 137–138. https://doi.org/10.1038/d41586-018-03924-9

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, 405(6783), 243–253. https://doi.org/10.1038/35012251

de Maximiano, M.F.D.A., d'Horta, F.M., Tuomisto, H., Zuquim, G., Van doninck, J. & Ribas, C.C. (2020) The relative role of rivers, environmental heterogeneity and species traits in driving compositional changes in southeastern Amazonian bird assemblages. *Biotropica*, 52(5), 946–962. https://doi.org/10.1111/btp.12793

Meyer, H. & Pebesma, E. (2020). *Predicting into unknown space? Estimating the area of applicability of spatial prediction models*. ArXiv:2005.07939 [Cs, Stat]. http://arxiv.org/abs/2005.07939

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M. & Nauss, T. (2018) Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, 101, 1–9. https://doi.org/10.1016/j.envsoft.2017.12.001

Naka, L.N. & Brumfield, R.T. (2018) The dual role of Amazonian rivers in the generation and maintenance of avian diversity. *Science. Advances*, 4(8), eaar8575. https://doi.org/10.1126/sciadv.aar8575

Nelson, B.W., Ferreira, C.A.C., da Silva, M.F. & Kawasaki, M.L. (1990) Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature*, 345(6277), 714–716. https://doi.org/10.1038/345714a0

Newton, P., Endo, W. & Peres, C.A. (2012) Determinants of livelihood strategy variation in two extractive reserves in Amazonian flooded and unflooded forests. *Environmental Conservation*, 39(2), 97–110. https://doi.org/10.1017/S0376892911000580

Newton, P., Peres, C.A., Desmoulière, S.J.M. & Watkinson, A.R. (2012) Cross-scale variation in the density and spatial distribution of an Amazonian non-timber forest resource. *Forest Ecology and Management*, 276, 41–51. https://doi.org/10.1016/j.foreco.2012.03.020

Newton, P., Watkinson, A.R. & Peres, C.A. (2011) Determinants of yield in a non-timber forest product: Copaifera oleoresin in Amazonian extractive reserves. *Forest Ecology and Management*, 261(2), 255–264. https://doi.org/10.1016/j.foreco.2010.10.014

Newton, P., Watkinson, A.R. & Peres, C.A. (2012) Spatial, temporal, and economic constraints to the commercial extraction of a non–timber forest product: Copaíba (Copaifera spp.) oleoresin in Amazonian reserves. *Economic Botany*, 66(2), 165–177. https://doi.org/10.1007/s12231-012-9198-z

Nobre, A.D., Cuartas, L.A., Hodnett, M., Rennó, C.D., Rodrigues, G., Silveira, A. et al. (2011) Height Above the Nearest Drainage – a hydrologically relevant new terrain model. *Journal of Hydrology*, 404(1–2), 13–29. https://doi.org/10.1016/j.jhydrol.2011.03.051

Pelletier, T.A., Carstens, B.C., Tank, D.C., Sullivan, J. & Espíndola, A. (2018) Predicting plant conservation priorities on a global scale. *Proceedings of the National Academy of Sciences*, 115(51), 13027–13032. https://doi.org/10.1073/pnas.1804098115

Pomara, L.Y., Ruokolainen, K., Tuomisto, H. & Young, K.R. (2012) Avian composition co-varies with floristic composition and soil nutrient concentration in Amazonian upland forests. *Biotropica*, 44(4), 545–553. https://doi.org/10.1111/j.1744-7429.2011.00851.x

Pomara, L.Y., Ruokolainen, K. & Young, K.R. (2014) Avian species composition across the Amazon River: The roles of dispersal limitation and environmental heterogeneity. *Journal of Biogeography*, 41(4), 784–796. https://doi.org/10.1111/jbi.12247

QGIS Development Team (2021) *QGIS Geographic Information System. Open Source Geospatial Foundation Project*. http://qgis.osgeo.org

R Core Team (2020) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL: http://www.R-project.org/

RADAMBRASIL. (1978). In *Folha SB.20 Purus; geologia, geomorfologia, pedologia, vegetacão e uso potencial da terra* (Vol. 17, p. 556). Rio de Janeiro: Projeto RADAMBRASIL/DNPM.

Rennó, C.D., Nobre, A.D., Cuartas, L.A., Soares, J.V., Hodnett, M.G., Tomasella, J. et al. (2008) HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. *Remote Sensing of Environment*, 112(9), 3469–3481. https://doi.org/10.1016/j.rse.2008.03.018

Ribas, C.C., Aleixo, A., Nogueira, A.C.R., Miyaki, C.Y. & Cracraft, J. (2012) A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729), 681–689. https://doi.org/10.1098/rspb.2011.1120

Roriz, P.A.C., Yanai, A.M. & Fearnside, P.M. (2017) Deforestation and carbon loss in Southwest Amazonia: Impact of Brazil's revised forest code. *Environmental Management*, 60(3), 367–382. https://doi.org/10.1007/s00267-017-0879-3

Ruokolainen, K. & Vormisto, J. (2000) The most widespread Amazonian palms tend to be tall and habitat generalists. *Basic and Applied Ecology*, 1(2), 97–108. https://doi.org/10.1078/1439-1791-00020

Salo, J., Kalliola, R., Häkkinen, I., Mäkinen, Y., Niemelä, P., Puhakka, M. et al. (1986) River dynamics and the diversity of Amazon lowland forest. *Nature*, 322(6076), 254–258. https://doi.org/10.1038/322254a0

Schobbenhaus, C., Gonçalves, J.H., Santos, J.O.S., Abram, M.B., Leao Neto, R., Matos, G.M.M. et al. (2004) *Carta geológica do Brasil ao milionésimo, sistemas de informações geográficas-SIG / Geological map of Brazil, 1:1,000,000 scale, geographic information system-GIS. CD-ROM*. CPRM, Geological Survey of Brazil, Brasilia.

Schulman, L., Toivonen, T. & Ruokolainen, K. (2007) Analysing botanical collecting effort in Amazonia and correcting for it in species range estimation. *Journal of Biogeography*, 34(8), 1388–1399. https://doi.org/10.1111/j.1365-2699.2007.01716.x

Tejada, G., Görgens, E.B., Espírito-Santo, F.D.B., Cantinho, R.Z. & Ometto, J.P. (2019) Evaluating spatial coverage of data on the aboveground biomass in undisturbed forests in the Brazilian Amazon. *Carbon Balance and Management*, 14(1), 11. https://doi.org/10.1186/s13021-019-0126-8

Toivonen, T., Mäki, S. & Kalliola, R. (2007) The riverscape of Western Amazonia - a quantitative approach to the fluvial biogeography of the region: The riverscape of Western Amazonia. *Journal of Biogeography*, 34(8), 1374–1387. https://doi.org/10.1111/j.1365-2699.2007.01741.x

Tuomisto, H., Moulatlet, G.M., Balslev, H., Emilio, T., Figueiredo, F.O.G., Pedersen, D. et al. (2016) A compositional turnover zone of biogeographical magnitude within lowland Amazonia. *Journal of Biogeography*, 43(12), 2400–2411. https://doi.org/10.1111/jbi.12864

Tuomisto, H. & Ruokolainen, K. (1997) The role of ecological knowledge in explaining biogeography and biodiversity in Amazonia. *Biodiversity & Conservation*, 6(3), 347–357. https://doi.org/10.1023/A:1018308623229

Tuomisto, H., Van doninck, J., Ruokolainen, K., Moulatlet, G.M., Figueiredo, F.O.G., Sirén, A. et al.(2019) Discovering floristic and geoecological gradients across Amazonia. *Journal of Biogeography*, 46(8), 1734–1748. https://doi.org/10.1111/jbi.13627

USGS. (2021). https://www.usgs.gov/core-science-systems/nli/landsat. accessed in February 2021.

Van doninck, J. & Tuomisto, H. (2018) A Landsat composite covering all Amazonia for applications in ecology and conservation. *Remote Sensing in Ecology and Conservation*, 4(3), 197–210. https://doi.org/10.1002/rse2.77

Van doninck, J., Westerholm, J., Ruokolainen, K., Tuomisto, H. & Kalliola, R. (2020) Dating flowering cycles of Amazonian bamboo-dominated forests by supervised Landsat time series segmentation. *International Journal of Applied Earth Observation and Geoinformation*, 93, 102196. https://doi.org/10.1016/j.jag.2020.102196

Veloso, H.P., Rangel Filho, A.L.R. & Lima, J.C.A., Fundação Instituto Brasileiro de Geografia e Estatística, & Departamento de Recursos Naturais e Estudos Ambientais. (1991). *Classificação da vegetação brasileira, adaptada a um sistema universal*. Ministério da Economia, Fazenda e Planejamento, Fundação Instituto Brasileiro de Geografia e Estatística, Diretoria de Geociências, Departamento de Recursos Naturais e Estudos Ambientais.

Wallace, A.R. (1852) On the monkeys of the Amazon. *Proceedings of the Zoological Society of London*, 20, 107–110.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Appendix S1. Ordination plots.

Appendix S2. Secondary spatial patterns of community composition.

How to cite this article: Zuquim, G., Tuomisto, H., Chaves, P.P., Emilio, T., Moulatlet, G.M., Ruokolainen, K., et al. (2021) Revealing floristic variation and map uncertainties for different plant groups in western Amazonia. *Journal of Vegetation Science*, 32:e13081. https://doi.org/10.1111/jvs.13081