Climate defined but not soil-restricted: the distribution of a Neotropical tree through space and time

Facundo Alvarez · Paulo S. Morandi · Ben Hur Marimon-Junior · Reginal Exavier · Igor Araújo · Lucas H. Mariano · Angélica O. Muller · Ted R. Feldpausch · Beatriz S. Marimon

Received: 5 April 2021 / Accepted: 21 October 2021 / Published online: 10 November 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract  Brosimum rubescens, a tree species with Neotropical distribution, can achieve local monodominance in Southern Amazonian forests. Understanding how and why this species varies across space and time is important because the monodominance of some species alters ecosystem complexity. Here we evaluated the fundamental ecological niche of B. rubescens by species distribution models (SDM), combining predictive environmental variables with occurrence points, and determined the temporal persistence and how the spatial distribution patterns of this species vary with different environmental predictive variables. To generate the SDMs, we incorporated predictive environmental variables as main components of climatic, hydric and edaphic variables. All algorithms showed higher performance in spatial predictions for hydric variables and for the combination of climatic, hydric and edaphic variables. We identified that the potential niches of B. rubescens seem to be defined by climatic fluctuations, with the edaphic conditions not limiting the presence of this species in the evaluated spatial scale. From the last glacial maximum to the present, this species seems to have increased its spatial amplitude; however, from the present to the future, predictions suggest that B. rubescens will experience a considerable loss of its range. Our findings showed independent and combined effects of different environmental variables, allowing us to identify which are limiting or facilitating the spatial distribution of B. rubescens. We corroborate the spatial persistence and geographical fidelity of the species’ distribution patterns over time.

Keywords  Species distribution models · Amazonia-Cerrado transition · Spatial hierarchy · Monodominant · Brosimum rubescens
Introduction

The global distribution of biodiversity follows a general pattern of spatial radiation from tropics to polar latitudes, and tropical forests are no exception to this pattern (Wiens and Donoghue 2004; Murphy and Bowman 2012). Within this spatial radiation, some tropical tree species dominate 50 to 100% of the total biomass forming the so-called monodominant forests (Hart et al. 1989). The monodominance of some species creates complex ecosystems, with no single or independent physical, chemical, or biological mechanism to explain the dominance over space and time (Torti et al. 2001; Peh et al. 2011; Marimon-Junior et al. 2019; ter Steege et al. 2019). The mechanisms promoting and maintaining the monodominant forests remain poorly understood and do not fit the spatio-temporal patterns of global biodiversity (Wiens and Donoghue 2004). One study suggested that these unique communities are associated with hydric, edaphic, and ecological traits occurring together (i.e. shade tolerance, slow decomposition rate, mast seeding, large seeds and poor dispersal, herbivore avoidance, ectomycorrhizal associations) under low exogenous disturbance over long time periods (Peh et al. 2011).

Neo- and paleo-tropical monodominant forests share some common edaphic characteristics, such as poor soil nutrient availability or an excess of certain elements (Marimon et al. 2001; Marimon-Junior 2007; Peh et al. 2011; Elias et al. 2018, 2019; Marimon-Junior et al. 2019). Nevertheless, there is no clear evidence that only soil properties are the main determinants creating or maintaining the monodominance (Brito-Rocha et al. 2016; Elias et al. 2018; Marimon-Junior et al. 2019). Modeling the spatial distribution of monodominant species, using different environmental variables and different temporal projections, allows a better understanding of their combined and independent spatio-temporal effects and a better approximation to their Grinnellian niche (Soberón 2010; Hijmans et al. 2015; Alvarez et al. 2020).

Brosimum rubescens Taub. (Moraceae), the target species of the present study has, a Neotropical distribution and is associated with non-seasonally flooded tropical forests, with the dispersion center in the Amazon Basin and Guyana, with some records in Central America and the Caribbean (Berg 1972; Palacios 2005). In the Amazonia-Cerrado transition, B. rubescens can form monodominant forests of up to 5000 ha and can represent up to 85% of all trees >10 cm diameter, one of the highest values recorded for a Neotropical monodominant species (Marimon et al. 2001). In addition, the high density of B. rubescens affects species diversity, unbalancing the uniformity of their abundance by acting as a biological barrier (Marimon-Junior 2007; Marimon et al. 2008; Marimon-Junior et al. 2019).

Although the occurrence and monodominance of some Neotropical forests are not conditioned by restrictive climatic conditions, differences in the distribution of the canopy understory light and species-specific capacities for the use of this resource is one of the factors of most significant repercussion for structuring such forests (Marimon et al. 2008; Peh et al. 2011). B. rubescens show development closely related to the amount and quality of sun radiation received by the young individuals (specially seedlings), and to the morphological and anatomical characteristics that provide phenotypic plasticity and periodical growth (Marimon et al. 2001, 2008, 2012). Also, soils affect the demography of B. rubescens in different phases of development, and the soil where this species is found has been reported to be acidic and nutrient poor, dystrophic (Quesada et al. 2010) and with Ca/Mg ration unbalanced (Marimon-Junior et al. 2019). Some studies have shown that B. rubescens invests in greater root biomass under more solar radiation as a response to seasonal hydric deficits (Butler et al. 2013), but the response of the species to climatic environmental dynamics is unknown (Marimon et al. 2008, 2012).

The temporal and spatial maintenance of B. rubescens monodominant forests is linked to episodic recruitment in response to extreme climatic interference, probably associated with the impacts of prolonged drought disturbances related to El Niño Southern Oscillation phenomenon (ENSO) (Marimon et al. 2014a, 2020; Marimon-Junior et al. 2019). Despite the morpho-anatomical and physiological attributes and environmental adaptations of B. rubescens, these drought events have caused a drastic reduction in the natural regeneration of the species, which may contribute to its population
decline in the Amazonia-Cerrado transition (Marimon et al. 2012, 2020). Therefore, future climatic variations may affect the sensitive balance between mortality and recruitment, affecting the local and regional distribution of *B. rubescens*. Despite several studies about the monodominant forest of *B. rubescens*, the patterns of the spatial distribution of this species, using different sets of variables and its spatial trends in future climate change scenarios, are not yet understood. In this work, we aimed to: 1) determine how the spatial distribution patterns of *B. rubescens* vary with different environmental predictive variables, and 2) evaluate the temporal and spatial persistence of *B. rubescens* in the Neotropics. Knowing the spatial and temporal patterns of this species would facilitate the planning of management strategies for conservation of this species and the unique monodominant tropical forest ecosystems.

**Materials and methods**

**Study area**

The results of the models were restricted to the northern portion of the South American continent (Fig. 1 and Supplementary Material S1), but to generate the models, the Neotropical region was chosen (Leroy et al. 2018). *B. rubescens* is a canopy species belonging to Moraceae family (Berg 1972) that can reach 90 cm diameter at breast height with heights of up to 35 m (Laurance et al. 2004). In Brazil, the occurrence of this species is reported in forests in the Amazonia,
a) OCCs: Obtaining and pre-processing

Spatial filter
spThin filter

Final database
Background points
70% Training 30% Test

b) Environmental predictors: pre-processing

Standardization (z ~ Normal distribution)
Spatial resolution adjustment (10 km ± 5 arc-minutes)
Cut (Neotropic)

Independent treatments

PCA 1
24 Climate variables

PCA 2
22 Edaphic variables

PCA 3
28 Hydric variables

PCA 4
All variables (69)

PCA 5
CCSM4 MPI-ESM-P MIROC-ESM

C
E
H
CEH-1
LMG-MH-F

Modeling methods: BIO, GWR, ME, GLM, RF and SVM

c) SDMs algorithms applied

d) SDM configuration and evaluation

- 70% for test and 30% for training the SDMs
- Adjustment of SDMs parameters
- SDMs executed from R
- 10,000 pseudo-absences
- 30 randomizations
- 1,260 final models

True Skill Statistic (TSS >0.4)
Area Under the Curve (AUC >0.7)
Cerrado, and Atlantic Forest Biomes, with spatial interruptions coinciding with the Caatinga Biome and monodominant forests registered only in Amazonia-Cerrado transition zone (Marimón et al. 2001, 2014a, 2020). In Amazonia and Atlantic Forest Biomes the soils are dystrophic but shallow for the former and humid and deep for the latter (Quesada et al. 2010; Joly et al. 2012). Cerrado soils are highly weathered and deep, acidic, dystrophic, and with high levels of iron and aluminum (Ribeiro and Walter 2008). The predominant climate types in Cerrado, Amazonia and Atlantic Forest Biomes, especially in the Brazilian areas where B. rubescens was registered are Af, Am and Aw (Dubreuil et al. 2018).

OCCs: Obtaining and pre-processing

We obtained the occurrence data (OCCs: Fig. 2a) for B. rubescens from the virtual platforms GBIF (http://www.gbif.org) and speciesLink (http://splink.cria.org.br), complemented with data from the Laboratório de Ecologia Vegetal (Universidade do Estado de Mato Grosso, Nova Xavantina). We incorporate all OCCs data regardless of the sampling date. The initial database had 256 records that we filtered spatially to identify and eliminate data outside the Neotropics. Subsequently, we eliminated all repeated, null, or dubious taxonomic data. A common problem in data obtained from virtual platforms is the spatial biases associated with the frequency, intensity and heterogeneity of the sampling efforts (Pearson et al. 2007; de Oliveira et al. 2014; De Marco and Nóbrega 2018). Using Kernel density estimation (KDE) spatial heterogeneity (Silverman 2018), was observed in the OCCs of B. rubescens (Fig. 1). To obtain the KDE, in the QGIS software (version 3.14; QGIS.org 2020), we used the interpolation extension ‘Heatmap’ with Epanechnikov Kernel function and 10,000 km as bandwidth value. We evaluated the spatial autocorrelation (SAC) of the OCCs using Moran index (Fig. 2a). As the data presented SAC (values greater than 0.2!) and spatial heterogeneity, we applied the “spThin” package with 10 km as the cutoff value (Aiello-Lammens et al. 2015) from the R platform (R Core Team 2020).

Preparation of environmental predictors

To better approximate the fundamental niche of B. rubescens we used predictive environmental variables (Supplementary Material S2) that show an effect at different scales and magnitudes (Soberón 2010; Domisch et al. 2015). Environmental variables act with spatial connections that follow a spatial hierarchical dependence pattern, with climate acting on the global scale, geomorphology on the regional scale and hydrological variables on the local scale (Wiens and Donoghue 2004; Domisch et al. 2015). In this sense, despite the system emergence properties, evaluating each set of variables in blocks (climate, edaphic and hydric) may promote a better understanding of the effect of each set on the spatial distribution patterns of different species (Velazco et al. 2017; Alvarez et al. 2020). We established three subsets of environmental predictors: climatic, edaphic, and hydric variables. In the climatic subset, we used 19 predictor variables obtained from the WorldClim platform (http://www.worldclim.org, Hijmans et al. 2005). For the edaphic characterization, we used 17 variables (Supplementary Material S2 and Fig. 1) obtained from the SoilGrids project (http://www.soilgrids.org).

As described in different studies, life cycles, biomass dynamics, growth, stature, and species distribution are conditioned by the seasonality of climatic variables related to hydric stress (Bruijnizeel 2004; Feldpausch et al. 2011, 2016). In this sense, we incorporated 23 variables in the hydric subset (Supplementary Material S2 and Fig. 1) obtained from different platforms: six variables from the potential evapotranspiration (http://www.cgiar-CSI.org), one from actual evapotranspiration and three from soil water stress (http://www.cgiar-CSI.org/data), being six variables of relative humidity (https://www.climond.org), three of cloud cover (http://www.cgiar-CSI.org/data), three of water vapor pressure (http://worldclim.org/version2) and one topographic wetness index variable (http://envirem.github.io). Considering the relative temporal homogeneity and importance for the characterization and physical description of the landscape of the five
following variables: global relief model (http://geodata.grid.unep.ch), elevation, slope-aspect (http://srtm.csi.cgiar.org) and terrain roughness index (http://envirm.github.io), we included them to each one of the three subsets described above.

To generate the species distribution models (SDMs) of past and future projections, we selected three scenarios; last glacial maximum (LGM: 22,000 years ago), middle Holocene (MH: 6000 years ago), and future (F: 2041–2060). We also used three models of global circulation for the selected periods: CCSM4, MPI-ESM-P, and MIROC-ESM, available in the WorldClim platform and selected because they generate robust temporal projections (Hijmans et al. 2005, 2015; Carnicer et al. 2020). Due to the increase in greenhouse gas emissions and trends expected for the future (Riahi et al. 2011; IPCC 2013), we considered the most extreme emission scenario [representative concentration paths (RCPs)=8.5]. Due to the global climate context and intensified anthropic actions, we selected RCP 8.5 which, according to the Intergovernmental Panel on Climate Change in 2004, represents the most pessimistic and probable scenario for projections until the year 2100 (Schwalm et al. 2020).

We incorporated all the predictor variables with a resolution of 5 arc minutes (~10 km in Ecuador), and we delimited them geographically using a Neotropical cut mask. We standardized all environmental predictor variables (μ = 0; SD = 1) to avoid different weights (Fig. 2). Subsequently, we performed a principal component analysis (PCA) to reduce environmental collinearity (De Marco and Nóbrega 2018); PCA1 (C: 24 climatic variables), PCA2 (E: 22 edaphic variables), PCA3 (H: 28 hydric variables), PCA4 (CEH-1: all variables together, 69). Before performing the PCAs, we projected the linear coefficients obtained for the current scenario in each past and future climate scenario. In each climate scenario, we used the first eight axes of the PCA that explained more than 95% of the original environmental variation of the original predictor variables as the new predictors for species distribution. We also obtained the independent contributions of each variable considered in the axes of all PCA (Supplementary Material S2 and Fig. 1).

SDM methods configuration and evaluation

We applied the SDMs that present greater predictive performances and flexibility with presence-only data (Fig. 2c), using the following algorithms: bioclim (BIO), Gower (GWR), maxent (ME), generalized linear models (GLM), random forest (RF) and support vector machine (SVM). To meet the requirements of the algorithms, we generated 10,000 background points randomly distributed within the study area, without spatially coinciding with known occurrence points. Then, we divided the data into 70% for test and 30% for training and validation. We generated the first models for the present scenario (PCA1, PCA2, PCA3, and PCA4) and later for the LGM, MH, and future scenarios. To obtain the potential distribution of B. rubescens, we used the most straightforward configurations of the models (Pearson et al. 2007; Varela et al. 2011), leaving the results to depend on the interactions between the algorithms and the input data (Fig. 2d).

We configured BIO and GWR algorithms by default using the “bioclim” (Hijmans and Graham 2006) and “gower” (Gower 1971) packages, respectively, and ME algorithm (“maxnet” R package: Phillips and Dudík 2008) in a predetermined way using linear characteristics. Where the mean and variance of the variables remained close to the observed values (Elith et al. 2011), the output was logistic type, with 1000 iterations and 10,000 background points. For GLM (“glmnet” R package: Hastie et al. 2016), the response data assume values of zero (absence) and one (presence), we used a logistic link function, considering all single predictor variables and without interaction between them. We optimized the estimation of the smoothing parameter using the proximal Newton method without selection of covariates backwards or forwards (Velazco et al. 2017). We adjusted the RF algorithm (“randomForest” R package: Liaw and Wiener 2002) automatically, used 500 trees in the adjustment step, and considered only those models with a minimum out-of-bag error, maintaining standard values in the means of the variables (Hastie et al. 2016; Velazco et al. 2017). Finally, we adjusted the SVM (“libsvm” R package: Chang and Lin 2011) algorithm in a predetermined manner with a linear kernel function type c-SVC with probabilistic output (Velazco et al. 2017).

We executed and built all SDMs using the R “dismo” package (Hijmans et al. 2015) with specific functions of each algorithm. With some exceptions, we configured the SVM models from the “ksvm” function of the “kernlab” package (Karatzoglou et al. 2004) and RF from the “randomForest” package.
Moreover, we defined the randomization process to 30 iterations for each algorithm (using a self-initialization analysis) to decrease the spatial structure between the training/test data (subsets of the same OCC points), to reduce possible correlation errors, and to reduce biased evaluations. Therefore, we generated 180 models for each set of predictive variables, totaling 1260 models.

The performance of the SDMs generated was evaluated by two families of statistical metrics: dependent and independent thresholds (Liu et al. 2005). In this sense, we applied the true skill statistic estimator (TSS) in which the response values vary between −1 to 1 where SDMs >0.5 are considered acceptable or adequate (Allouche et al. 2006). To complement, we applied the independent threshold area under the receiver-operator curve (AUC: Fielding and Bell 1997), which discriminates areas of omission from areas with known occurrence. Values range from 0 to 1, where the closer to 1 the better is the model’s ability to represent reality (Elith 2000). To verify the differences in AUC and TSS values among the PCAs and the applied algorithms, we performed Kruskal-Wallis non-parametric tests since the data did not meet the assumption of normality.

We used as a threshold the lowest adequacy value within the incorporated OCCs [(lowest presence training (LPT)], allowing the generation of presence/absence matrices (binary models) from the adequacy of continuous matrices projected for South America (environmental models). Applying the consensus of the weighted average we assembled all the suitability matrices obtained for the different temporal scenarios. In this way, to generate a final distribution matrix for B. rubescens in the temporal scenarios, we applied a consensus method, assembling all the predictions considering only those models with TSS values greater than the mean TSS value.

Results

Model evaluation

All algorithms showed high performance and we registered statistical differences among them (Fig. 3). Moreover, all models showed TSS and AUC adequate values (Table 1), but the best performances were associated with the SVM, ME, and RF algorithms, with TSS values >0.5 and AUC>0.8 (Table 1, Fig. 3). The models generated exclusively from hydric variables and combining all the variables performed better, regardless of the algorithms (Table 1). For both PCAs and algorithms, the null hypothesis was rejected (p value <0.001) (Fig. 3).

Present projections

The spatial complexity of the generated predictions was different depending on the built-in predictor variables. The suitability matrices generated from PCA1 (C: Climatic variables) and PCA3 (H: Hydric variables) showed wide continuity and spatial aggregation (Fig. 4). The matrices generated from the PCA3 (E: Edaphic variables) showed diffuse spatial distribution patterns of wide spatial amplitude (Fig. 4). When considering all the variables together (CEH-1), two situations can be observed: high connectivity in spatial distribution patterns and/or low spatial complexity concerning the other PCAs used. Regardless of the variables incorporated, all models showed centers of greater suitability in the central region of the Amazon Basin, the Atlantic coast, and the North region of South America (Fig. 4). As a particular case, CEH-1 showed predictions of suitability associated with the Central America and the Caribbean islands, although we have not incorporated occurrence records for these areas. Regardless of the set of built-in predictor variables, the predictions of environmental suitability were superimposed on the occurrence records obtained in the field.

When we combined all the variables to assess the independent contribution of each within a 200,000 km² hexagon, we observed in more detail the environmental structure (C: Climatic; E: Edaphic and H: Hydric) of the study area (CEH-2: Fig. 4 and Supplementary Material S3). In this sense, in the Amazon Basin and the Atlantic Coast, the balance between the three sets of variables makes it possible to recognize, geographically, the higher probabilities of environmental suitability and they coincide spatially with the reports of occurrence of B. rubescens.

Projections of temporal climate scenarios

The spatial distribution patterns of the models for different scenarios showed some overlap in the historical spatial distribution patterns of B. rubescens (Fig. 5).
Fig. 3 Kruskal-Wallis (KW) distributions of means and standard deviation of AUC and TSS values between different environmental scenarios (Climate: C; Edaphic: E; Hydric: H and CEH-1: all in one) and SDM algorithms (BIO, GLM, GWR, ME, RF, and SVM). The red lines indicate the values from which the models have high performance.

Table 1 Mean values of AUC (Area Under the receiver-operator Curve) and TSS (True Skill Statistic) of the different SDMs applied to assess the distribution of *Brosimum rubescens* in the treatment of climate (C), edaphic (E), hydric (H), climate + edaphic + hydric (CEH-I) and the last glacial maximum (LGM), middle Holocene (MH) and future (F) scenarios.

| Algorithms | C (PCA1) | E (PCA2) | H (PCA3) | CEH-I (PCA4) | Total (AUC) | Total (TSS) |
|------------|----------|----------|----------|--------------|-------------|-------------|
|            | AUC   | TSS   | AUC   | TSS   | AUC   | TSS   | AUC   | TSS   | AUC   | TSS   |
| BIO        | 0.741 | 0.417 | 0.710 | 0.359 | 0.753 | 0.436 | 0.757 | 0.458 | 0.740 | 0.417 |
| GLM        | 0.805 | 0.566 | 0.756 | 0.489 | 0.797 | 0.528 | 0.808 | 0.598 | 0.792 | 0.545 |
| GWR        | 0.726 | 0.376 | 0.715 | 0.394 | 0.781 | 0.504 | 0.784 | 0.501 | 0.752 | 0.444 |
| ME         | 0.812 | 0.560 | 0.769 | 0.481 | 0.810 | 0.560 | 0.823 | 0.620 | 0.803 | 0.555 |
| RF         | 0.832 | 0.598 | 0.784 | 0.458 | 0.832 | 0.542 | 0.837 | 0.621 | 0.821** | 0.555** |
| SVM        | 0.798 | 0.544 | 0.768 | 0.460 | 0.823 | 0.557 | 0.818 | 0.588 | 0.801 | 0.537 |
| Total (PCAs) | 0.786 | 0.510 | 0.750 | 0.440 | 0.799 | 0.521 | 0.804* | 0.564* | 0.785 | 0.509 |

*Higher values of AUC and TSS comparing the different sets of variables

**Higher AUC and TSS values of the different SDMs applied
Despite the “temporal distance” of the compared scenarios, the environmental suitability matrices (Fig. 5: Species model distribution) show high spatial overlap, with good connectivity, being more significant and evident in the environmental suitability centers associated with the Amazon Basin, North of South America and Atlantic coast. Although the matrices obtained from the projections have diffuse regions, both for the past (LGM and MH) and future (F) scenarios, the highest suitability values are superimposed on the present spatial distribution patterns (CEH-1). The projections were made using only climatic variables and not considered them as predictors of the landscape, such as topographic or edaphic variables.

While we did not incorporate occurrence records for Central America or the Caribbean islands, projections for both past and future scenarios showed environmental suitability associated with these regions (Fig. 5: Species model distribution). The suitability matrices for all temporal projection scenarios also show values associated with the Patagonian region, with wide spatial amplitude of predictions in the LGM and future scenarios. In the matrices of all projections two large areas can be seen, one between the Caribbean islands and Amazonia and the other between Amazonia and Atlantic Forest biomes (Fig. 5: Species model distribution).
The pattern of spatial overlap between different scenarios can be highlighted in higher resolution in the binary matrices (Fig. 5: Binary models). The present, past and future (PPF) map is the result of overlapping all of them in a single matrix (Fig. 5: Binary models). The values associated with each binary matrix represent the area (km²) occupied by *B. rubescens* within the study area. The central triangle shows losses and gains areas from one scenario to another and a gain of 932,659 km² can be noticed from the LGM to the CEH-1 (present). Also, it shows a substantial loss of 568,118 km² from the present to the future projection. However, this loss is not total: when comparing the surface calculated for the LGM with the amount of surface that was projected for the future, the result remains positive (Fig. 5: Binary models). Despite these variations in the total area values over time, in the PPF matrix, *B. rubescens* has a high and recurring presence associated with the coastal region of the South American Caribbean, Amazon Basin, and the Atlantic Forest, where we found highest adequacy values for all projections (Fig. 5: Binary models).

**Discussion**

The spatial distribution patterns predicted for *B. rubescens*, regardless of the variables, allowed a great spatial and historical description of the potential niches in the Neotropics (Fig. 5). All algorithms showed high performance, with greater responses for hydric
variables and all predictors combined (Table 1 and Fig. 3). The spatial distribution patterns predicted using edaphic variables showed higher amplitude and spatial connectivity than those predicted in a similar way for climatic and hydric variables (Fig. 4). In the evaluated spatial scale, the potential niches of *B. rubescens* calculated from climatic variables presented smaller spatial amplitude (spatially restrictive) compared to those obtained from edaphic variables (not spatially restricted). Yet, historical patterns showed high spatial overlap and geographic fidelity, with increases in spatial amplitude from the past to the present and decrease from the present to the future (Fig. 5).

*Brosimum rubescens* and climate

The SDMs predictions obtained from climatic variables showed high performance, strong relationship, and dependence between climate and the spatial distributions of *B. rubescens* (Fig. 4), confirming results supported and found in multiple works of vegetation SDM with climatic variables (e.g. Peh et al. 2011; Velazco et al. 2017; Soong et al. 2020). Our models showed suitable areas highly connected and spatially continuous, however, with two spatial segregations, one in Amazonia and other associated with the Atlantic Forest (Fig. 5). Climatic conditions are very similar in the Central Amazonia and the *Hileia Baiana* region within the Atlantic Forest in Brazil (Andrade-Lima 1966; Gentry 1982; Sobral-Souza et al. 2015). Following the Köppen-Geiger (Beck et al. 2018) climate classification, these two regions share similar climatic conditions, corresponding to Af (equatorial), Am (monsoon), and Aw (savanna). Mainly Af and Am are characterized as tropical without (or short) dry season. In fact, the low prediction associated with the Amazonia-Cerrado transition (Aw type) corresponds to a tropical climate with more extended dry season (Beck et al. 2018).

Topographic variables, global relief model (GRM) and elevation were also selected within PCA1: these variables together with pluviometry, conditioned the temporal and spatial connectivity of hydrological networks (Marimon-Junior 2007; Domisch et al. 2015; Parreira et al. 2019) and species distribution (Elias et al. 2019). For some monodominant species, precipitation is important due to its impacts on water table fluctuations and flood pulses (Nunes Da Cunha and Junk 2001). However, there is no relationship between *B. rubescens* and soil moisture, but a set of other ecological conditions acting together on the mechanisms of monodominance origin and maintenance (Marimon-Junior et al. 2019). For *B. rubescens*, fruiting and seed dispersal processes are associated with the beginning of rainy season (Rivera et al. 2014) or end of dry season (Marimon and Felfili 2006). Also, ENSO-related extreme drought events affect negatively the recruitment and survival rates of *B. rubescens* (Laurance 2003; Phillips et al. 2010; Marimon et al. 2020), probably compromising the monodominance maintenance. This species is light-demanding with episodic recruitment highly dependent on the temporal variations of light and temperature (Marimon et al. 2012). Although variables that could characterize this dynamic were not incorporated into our analysis, some variables such as those of cloud cover could act as proxies for this spatial variation of the environment (Clark and Clark 1994).

*Brosimum rubescens* and edaphic variables

Edaphic properties have been considered the best predictors of woody species spatial distribution (Velazco et al. 2017; Elias et al. 2018). The models obtained from edaphic variables enabled to correctly predicting the distribution of *B. rubescens* (Fig. 4). However, their performances were lower than those obtained with other variables, since the spatial distribution of tropical species is driven by the interaction of multiple variables in a set of ecological conditions acting together, in which soil attributes can predominate as a key determinant (e.g., soil texture, exchangeable cations) (Marimon-Junior et al. 2019; Elias et al. 2019). The spatial distribution patterns showed association with the central region of Amazonia and the Atlantic coast that have similar edaphic conditions, such as deep soils, predominantly Oxisols, and Argisols (Quesada et al. 2010; Marimon-Junior et al. 2019). The wide spatial amplitude obtained with edaphic variables suggests that the species could have a greater amplitude of niches, even so, it is subject to other environmental conditions acting together, such as climate and hydrology (i.e., soil texture, drainage/depth of water table), that limit its presence in some places (Fig. 4 and Supplementary Material S3).

The models resulting from edaphic variables show high spatial overlap with the areas of highest
concentration of interchangeable cations (Supplementary Material S1: Cation exchange capacity). Thus, high concentrations of exchangeable cations in the top soil could act as a competitive advantage allowing the species to reach a wide range of niches (Elias et al. 2019) and, occasionally, reaching the monodominance with other ecological factors acting together (Marimon-Junior et al. 2019). However, some studies suggested that edaphic properties affect *B. rubescens* according to ontogenetic stage, thus the dependency relationship of this species with the edaphic properties increases in complexity (Brito-Rocha et al. 2016; Elias et al. 2018). In this sense, models generated from edaphic variables, for species with dissimilar edaphic dependencies in their development, could be conditioning the performance of the predictions.

**Brosimum rubescens** and hydric variables

The spatial distributions of the predictions of suitability generated from hydric variables showed high spatial connectivity, few diffuse areas, and spatial overlap with the Amazonas-Tocantins-Araguaia hydrographic networks and the East Atlantic and East-Northeast Atlantic Basins (Fig. 4). Water resources predictors are primarily responsible for defining hydraulic vegetation conditions (Borchert 1994; Anderegg et al. 2018; Bittencourt et al. 2020). Fluctuations in hydric properties act as a strong environmental filter, spatially segregating phytosociognomies according to hydric and thermal stress tolerance thresholds (Borchert 1994; Bittencourt et al. 2020; Araújo et al. 2021). In the geographical extension (Leroy et al. 2018) delimited in our case, Neotropics, climate and edaphic properties may affect the predictions on macro-ecological spatial scales (regional or continental) and the hydric variables could be acting on smaller spatial scales (Domisch et al. 2015; Velazco et al. 2017). Since no relationship between soil moisture and *B. rubescens* monodominance has been recorded (Marimon-Junior et al. 2019), other soil hydraulic conditions, such as water table fluctuations, could be acting as one of the soil hydraulic factors influencing the distribution of the species. However, considering the difficulty of determining the water table depth on a large scale, further studies must be carried out to corroborate this hypothesis.

We observed that soil water stress (SWS) and topographic humidity index (TWI) were related to *B. rubescens* occurrence. Drought is an important factor related to the hyperdynamic of some tropical forests, affecting the structure, dynamics, and composition of tree species (Marimon et al. 2014b). The regeneration of this species may be conditioned by decreases in the soil moisture derived from droughts during and after ENSO events (Marimon et al. 2020). However, *B. rubescens* also produce larger amounts of seeds during drought periods (Marimon and Felfili 2006), and this strategy may be advantageous for maintaining its dominance (Rivera et al. 2014), specially due to the chance that some seeds survive the drought.

**Brosimum rubescens**, CEH-1 and CEH-2

Each set of variables can generate different predictions for a species, but the combination of all these variables in a single subset allows a better representation of ecosystems by improving the approximations of the fundamental niche (Soberón 2010; Velazco et al. 2017). When the different subsets of variables were combined in a single block (CEH-1: Fig. 4 and Supplementary Material S3), the result for *B. rubescens* showed a balance in the values of AUC and TSS (De Marco and Nóbrega 2018). When the variables are incorporated into subsets, they respond at their thresholds, but when they are complemented, they respond with spatial hierarchical dependence (Rivera et al. 2014; Domisch et al. 2015; Elias et al. 2018).

The environmental suitability matrices obtained from the CEH-1 scenario were adjusted to the spatial distribution patterns of occurrence reported for *B. rubescens* within the Neotropics (Fig. 4). The spatial distribution pattern of the Brosimeae tribe extends from Mexico to southern Brazil, with records in some Antilles islands (Cuba and Jamaica) and many species associated with the Amazon Basin (Berg 1972). According to this, *B. rubescens* are spatially discontinuous in the Neotropics, but the highest frequency of records is associated with the Amazon Basin in Brazil, Peru, Colombia, and Guianas. In the Brazilian Amazonia-Cerrado transition, *B. rubescens* form extensive monodominant forests (Marimon et al. 2001, 2020), and in Central Amazonia, this species had wide spatial connectivity in the environmental suitability.

As evidenced in the CEH-2 map (Fig. 4 and Supplementary Material S3), the regions where the
balance between climatic, edaphic, and hydraulic factors were greater for this species and had a high spatial overlap with the matrices of environmental suitability obtained from the different algorithms. Other areas where the proportions of C, E, and H, were not in balance, showed the possible limiting factors for the absence. An example of this would be the Patagonia region where environmental suitability was predicted, but it would only be the edaphic variables (and climate in a smaller proportion) that would facilitate the presence of *B. rubescens*, as the hydric restrictions do not allow its development.

*Brosimum rubescens* and historical occurrence

In our results, we observed variation in the spatial amplitudes of the environmental suitability projected in different temporal scenarios (Fig. 5). Other studies suggested that decreases in population densities over time in conjunction with predominantly random spatial distribution pattern may be responding to density-dependent mortality, so when evaluating different dominant species of tropical forests, we observed that those with lower spatial densities tend to easily disappear (Felfili et al. 2000; Arieira et al. 2016). After recent drought events, *B. rubescens* showed variations in population density and demography, mainly in the regenerating stratum, but the robustness of the intraspecific spatial distribution pattern was maintained for adult trees, demonstrating resistance and resilience of the species to exogenous disturbances (Elias et al. 2018). In this sense, the extreme climate events may not change the spatial distribution pattern of the Neotropical distribution of *B. rubescens*, as was observed on a smaller spatial scale (Elias et al. 2018; Marimon et al. 2020). In this context, the strategies for spatial occupation and temporal maintenance of the occurrence areas are well characterized in the SDMs (Marimon et al. 2001; Soberón 2010; Elias et al. 2018).

Conclusion

Our SDMs results show the independent and combined effects of different environmental predictors for *B. rubescens* distribution, allowing the identification of variables that can increase or decrease the spatial amplitudes of its occurrence (Fig. 4). At the same time, we corroborate the spatial persistence and geographic fidelity of the spatial distribution patterns of the species over time (Fig. 5). The SDMs generated for *B. rubescens* from edaphic variables showed wide spatial amplitudes, greater than when considering just climatic or soil hydric variables, for which the predictions were spatially restrictive. Knowing the responses of species to environmental variables that condition their spatial and historical patterns are important in the global context of human actions and intensification of climate change.

Acknowledgments A special thanks for the contributions and suggestions of Maria Eduarda Maldaner (Dudita), as well as Angélica Faria de Resende and Narain Loureiro Benone, which helped to substantially improve this work. This study was financed in part by the Coordenação de Aperfeiçoamento
Code availability  Available upon request.

Authors' contributions  Conceived and designed the investigation: FA, PSM, BHM-J, RE, IA, LHM, AOM, BSM. Performed field and laboratory work: FA, PSM, IA, LHM, AOM. Analyzed the data: FA, RE. Contributed materials and analysis tools: FA, RE, BHM-J, TRF, BSM. Wrote the paper: FA, PSM, BHM-J, RE, IA, LHM, AOM, TRF, BSM.

Funding  This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brazil (CAPES) - Finance Code 001. We also thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico/Projetos Ecológicos de Longa Duração-CNPq/PELD (441244/2016-5) and FAPEMAT/ReFlor (0589267/2016), for financial support. B.S. Marimon and B.H. Marimon-Junior acknowledge CNPq for their PQ-1D productivity grants (301153/2018-3 and 311027/2019-9), and P.S. Morandi acknowledges CAPES for his post-doc grant (88887.185186/2018-00). T.R. Feldpausch was supported by the UK Natural Environment Research Council (NERC, NE/N011570/1).

Data availability  Species occurrences are available in the supplementary material.

Declarations

Conflict of interest  The authors declare they have no conflict of interests to declare.

Ethics approval  Not applicable.

Consent to participate  Not applicable.

Consent for publication  All authors agree to the publication of the present work.

References

Aiello-Lammens ME, Boria RA, Radosavljevic A et al (2015) spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography (Cop) 38:541–545. https://doi.org/10.1111/eco.01132

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol 43:1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Alvarez F, Gerhard P, Paiva Silva D, Spacek Godoy B, Montag LFDA (2020) Effects of different variable sets on the potential distribution of fish species in the Amazon Basin. Ecol Freshw Fish 29(4):764–778. https://doi.org/10.1111/eff.12552

Anderegg WRL, Konings AG, Trugman AT et al (2018) Hydraulic diversity of forests regulates ecosystem resilience during drought. Nature 561:538–541. https://doi.org/10.1038/s41586-018-0539-7

Andrade-Lima AD (1966) Contribuição ao estudo do paralelismo da flora Amazônico-Nordestina (1st edn). Secretaria de Agricultura, Indústria e Comércio, Recife

Araújo I, Marimon BS, Scalon MC, Fauset S, Junior BHM, Tiwari R, Gloor MU (2021) Trees at the Amazonia-Cerrado transition are approaching high temperature thresholds. Environ Res Lett 16(3):034047. https://doi.org/10.1088/1748-9326/abeb39

Arieara J, Penha J, Nunes da Cunha C, Couto EG (2016) Ontogenetic shifts in habitat-association of tree species in a neotropical wetland. Plant Soil 404:219–236. https://doi.org/10.1007/s11104-016-2844-y

Beck HE, Zimmermann NE, McVicar TR et al (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci Data 5:180214. https://doi.org/10.1038/sdata.2018.214

Berg CC (1972) Flora Neotropica - monograph no. 7: Olmecidea, Brosimeae (moraceae), reprinted. Hafner Publishing Company, Inc.: New York

Bittencourt PRL, Oliveira RS, da Costa ACL et al (2020) Amazonia trees have limited capacity to acclimate plant hydraulic properties in response to long-term drought. Glob Chang Biol 26:3569–3584. https://doi.org/10.1111/gcb.15040

Borchert R (1994) Soil and stem water storage determine phenology and distribution of tropical dry Forest trees. Ecology 75:1437–1449. https://doi.org/10.2307/1937467

Brito-Rocha E, Schilling AC, Dos Anjos L et al (2016) Regression models for estimating leaf area of seedlings and adult individuals of Neotropical rainforest tree species. Braz J Biol 76:983–989. https://doi.org/10.1590/1519-6984.05515

Bruijnzeel LA (2004) Hydrological functions of tropical forests: not seeing the soil for the trees? Agric Ecosyst Environ 104:185–228. https://doi.org/10.1016/j.agee.2004.01.015

Butler A, Marimon-Junior BH, Maracahipes L, Marimon BS, Silvério DV, Oliveira EA, Lenza E, Feldpausch TR, Meir P, Grace J (2013) Absorbing roots areas and transpiring leaf areas at the tropical Forest and savanna boundary in Brazil. 2013. In: Perrault C, Bellamy (eds) Savannas: climate, biodiversity and ecological significance. Nova Science Publishers, Exeter, pp 107–126

Carnicer C, Eisenlohr PV, Jácomo A et al (2020) Running to the mountains: mammal species will find potentially
suitable areas on the Andes. Biodivers Conserv 29:1855–1869. https://doi.org/10.1007/s10531-020-01951-5

Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol (TIST) 2(3):1–27

Clark DA, Clark DB (1994) Climate-induced annual variation in canopy tree growth in a Costa Rican tropical RainForest. J Ecol 82:865. https://doi.org/10.2307/2261450

De Marco P, Nóbrega CC (2018) Evaluating collinearity effects on species distribution models: an approach based on virtual species simulation. PLoS One 13:202403. https://doi.org/10.1371/journal.pone.020243

De Oliveira G, Rangel TF, Lima-Ribeiro MS et al (2014) Evaluating, partitioning, and mapping the spatial autocorrelation component in ecological niche modeling: a new approach based on environmentally equidistant records. Ecography (Cope) 37:637–647. https://doi.org/10.1111/j.1600-0587.2013.00564.x

Domisch S, Jähnig SC, Simaika JP et al (2015) Application of species distribution models in stream ecosystems: the challenges of spatial and temporal scale, environmental predictors and species occurrence data. Fundam Appl Limnol 186:45–61. https://doi.org/10.1139/fal.2015/0627

Dubreuil V, Fante KP, Planchon O, Neto JLSA (2018) Os tipos de climas anuais no Brasil: uma aplicação da classificação de Köppen de 1961 a 2015. Confinis. Revue Franco-brésilienne de géographie/Revista franco-brasileira de geografia, (37). https://doi.org/10.4000/confins.15738

Elias F, Marimón BS, Marimón-Junior BH et al (2018) Idiosyncratic soil-tree species associations and their relationships with drought in a monodominant Amazon forest. Acta Oecol 91:127–136. https://doi.org/10.1016/j.actao.2018.07.004

Elias F, Marimón Junior BH, de Oliveira FJM et al (2019) Soil and topographic variation as a key factor driving the distribution of tree flora in the Amazonia/Cerrado transition. Acta Oecol 100:103467. https://doi.org/10.1016/j.actao.2019.103467

Elith J (2000) Quantitative methods for modelling species habitat: comparative performance and an application to Australian plants. In: Ferson S, Burgman M (eds) Quantitative methods for conservation biology, 1st edn. Springer, New York, pp 39–58

Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Divers Distrib 17(1):43–57

Feldpausch TR, Banin L, Phillips OL et al (2011) Height-diameter allometry of tropical forest trees. Biogeosciences 8:1081–1106. https://doi.org/10.5194/bg-8-1081-2011

Feldpausch TR, Phillips OL, Brienen RJW et al (2016) Amazon forest response to repeated droughts. Glob Biogeochem Cycles 30:964–982. https://doi.org/10.1002/2015GC005133

Felfili JM, Rezende AV, Da Silva J, Silva MA (2000) Changes in the floristic composition of cerrado sensu stricto in Brazil over a nine-year period. J Trop Ecol 16:579–590. https://doi.org/10.1017/S0266467400001589

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49. https://doi.org/10.1017/S037689297000088

Gentry AH (1982) Neotropical floristic diversity: phytogeographical connections between central and South America, Pleistocene climatic fluctuations, or an accident of the Andean orogeny? Ann - Missouri Bot Gard 69:557–593. https://doi.org/10.2307/2399084

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:623–637

Hallé F, Oldeman RAA, Tomlinson PB (1978) Tropical trees and forests, 1st edn. Springer, Berlin

Hart TB, Hart JA, Murphy PG (1989) Monodominant and species-rich forests of the humid tropics: causes for their co-occurrence. Am Nat 133:613–633. https://doi.org/10.1086/284941

Hastie T, Qian J, Tay K (2016) An introduction to glmmnet

Hijmans RJ, Graham CH (2006) Testing the ability of climate envelope models to predict the effect of climate change on species distributions. Glob Chang Biol 12:2272–2281. https://doi.org/10.1111/j.1365-2486.2006.01256.x

Hijmans RJ, Cameron SE, Parra JL et al (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965–1978. https://doi.org/10.1002/joc.1276

Hijmans RJ, Phillips S, Leathwick J, Elith J. (2015) dismo: species distribution modeling. R package

IPCC (2013) Summary for Policymakers. In: Stocker T.F., Qin G-K D., Plattner MT, et. al. (eds) Climate Change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, pp. 3–29

Joly CA, Assis MA, Bernacci LC et al (2012) Florística e fitossociologia em parcelas permanentes da Mata Atlântica do sudeste do Brasil ao longo de um gradiente altitudinal. Biota Neotrop 12:123–145. https://doi.org/10.1590/s1676-06032012000100012

Karatzioglou A, Hornik K, Smola A, Zeileis A (2004) Kernlab - An S4 package for kernel methods in R. J Stat Softw 11:1–20. https://doi.org/10.18637/jss.v011.i09

Laurance WF (2003) Slow burn: the insidious effects of surface fires on tropical forests. Trends Ecol Evol 18:209–212. https://doi.org/10.1016/S0169-5347(03)00064-8

Laurence WF, Nascimento HEM, Laurance SG et al (2004) Inferred longevity of Amazonian rainforest trees based on a long-term demographic study. For Ecol Manag 190:131–143. https://doi.org/10.1016/j.foreco.2003.09.011

Leroy B, Delsol R, Hugueny B et al (2018) Without quality presence-absence data, discrimination metrics such as TSS can be misleading measures of model performance. J Biogeogr 45:1994–2002. https://doi.org/10.1111/jbi.13402

Liaw A, Wiener M (2002) Classification and regression by random forest. R News 2:18–22

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting predictors and species occurrence data. Fundam Appl Environ Conserv 20:33–53. https://doi.org/10.1007/s10376-005-00395-x

Springer
Marimon BS, Felfili JM (2006) Chhua de sementes em uma floresta monodominate de *Brosimum rubescens* Taub. e em uma floresta mista adjacente no Vale do Araguaia, MT, Brasil. Acta Bot Bras 20:423–432. https://doi.org/10.1590/S0102-3366200600020017

Marimon BS, Felfili JM, Haridasan M (2001) Studies in monodominant forests in eastern Mato Grosso, Brazil: I. A forest of *Brosimum rubescens* taub. Edinb J Bot 58:123–137. https://doi.org/10.1017/S096042860100049X

Marimon BS, Felfili JM, Marimon-Junior BH et al (2008) Desenvolvimento inicial e partição de biomassa de *Brosimum rubescens* Taub. (Moraceae) sob diferentes níveis de sombreamento. Acta Bot Bras 22:941–953. https://doi.org/10.1590/S0102-33662008000400005

Marimon BS, Felfili JM, Fagg CW et al (2012) Monodominance in a forest of *Brosimum rubescens* Taub. (Moraceae): structure and dynamics of natural regeneration. Acta Oecol 43:134–139. https://doi.org/10.1016/j.actao.2012.07.001

Marimon BS, Felfili JM, Marimon BH et al (2014a) Leaf herbivory and monodominance in a Cerrado–Amazonia transitional forest, Mato Grosso, Brazil. Plant Biodivers Int J Deal Asp Plant Biol 150:124–130. https://doi.org/10.1080/11263504.2014.983577

Marimon BS, Marimon-Junior BH, Feldpausch TR, Oliveira-Santos C, Mews HA, Lopez-Gonzalez G et al (2014b) Disequilibrium and hyperdynamic tree turnover at the forest-cerrado transition zone in southern Amazonia. Plant Ecol Divers 7(1–2):281–292. https://doi.org/10.1080/17550874.2013.818072

Marimon BS, Oliveira-Santos C, Marimon-Junior BH et al (2020) Drought generates large, long-term changes in tree and liana regeneration in a monodominant Amazon forest. Plant Ecol 221:733–747. https://doi.org/10.1007/s11258-020-01478-8

Marimon-Junior BH (2007) Relação entre diversidade arbórea e aspectos do ciclo biogeoclimático de uma floresta monodominante de *Brosimum rubescens* Taub. e uma floresta mista adjacente no Leste Mato-grossense. PhD Thesis, Universidade de Brasília

Marimon-Junior BH, Hay JDV, Olivares I et al (2019) Soil water-holding capacity and monodominance in southern Amazon tropical forests. Plant Soil 450:65–79. https://doi.org/10.1007/s11104-019-04257-w

Murphy BP, Bowman DMJS (2012) What controls the distribution of tropical forest and savanna? Ecol Lett 15:748–758. https://doi.org/10.1111/j.1461-0248.2012.01771.x

Nunes Da Cunha C, Junk WJ (2001) Distribution of Woody plant communities along the flood gradient in the Pantanal of Poconé, Mato Grosso, Brazil. Int J Ecol Environ Sci 27:63–70

Palacios PA (2005) Patrones estructurales y distribución espacial de poblaciones de *Brosimum rubescens* taub. En relación con la variabilidad fisiográfica en la ribera colombiana del río amazonas. Universidad Nacional de Colombia-Sede Amazonia

Parreira MR, Nabout JC, Tessarollo G et al (2019) Disentangling uncertainties from niche modeling in freshwater ecosystems. Ecol Model 391:1–8. https://doi.org/10.1016/j.ecolmodel.2018.10.024

Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr 34:102–117. https://doi.org/10.1111/j.1365-2699.2006.01594.x

Peh KSH, Lewis SL, Lloyd J (2011) Mechanisms of monodominance in diverse tropical tree-dominated systems. J Ecol 99:891–898. https://doi.org/10.1111/j.1365-7354.2011.01827.x

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31(2):161–175

Phillips OL, van der Heijden GM, Lewis SL et al (2010) Drought-mortality relationships for tropical forests. New Phytol 187:631–646. https://doi.org/10.1111/j.1469-8137.2010.03359.x

QGIS.org (2020) QGIS geographic information system. Open Source Geospatial Foundation Project. http://qgis.org

Quesada CA, Lloyd J, Schwarz M et al (2010) Variations in chemical and physical properties of Amazon forest soils in relation to their genesis. Biogeosciences 7:1515–1541. https://doi.org/10.5194/bg-7-1515-2010

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3–900051–07–0. http://www.R-project.org/

Riahi K, Rao S, Krey V et al (2011) RCP 8.5—a scenario of comparatively high greenhouse gas emissions. Clim Chang 109:33–57. https://doi.org/10.1007/s10584-011-0149-y

Ribeiro JF, Walter BMT (2008) As principais fitofisionomias do bioma Cerrado. In: Sano SM, de Almeida SP, Ribeiro JF (eds) Cerrado: Ecologia e flora, 1st edn. Embrapa Cerrados/Embrapa Informação Tecnológica, Brazil, pp 152–212

Rivera LE, Peaúela MC, Moreno F (2014) Intra annual seed production and availability of two morphotypes of *Brosimum rubescens* taubert in forests of the Colombian Amazon. Biota Neotrop 14:e20130073. https://doi.org/10.1590/0109-3125140007313

Schwalm CR, Glendon S, Duffy PB (2020) RCP 8.5 tracks cumulative CO2 emissions. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.2007117117

Silverman BW (2018) Density estimation: for statistics and data analysis, 1st edn. Routledge, New York

Soberón JM (2010) Niche and area of distribution modeling: a population ecology perspective. Ecography (Cop) 33:159–167. https://doi.org/10.1111/j.1600-0587.2009.06074.x

Soobral-Souza T, Lima-Ribeiro MS, Solferini VN (2015) Biogeography of Neotropical rainforests: past connections between Amazon and Atlantic Forest detected by ecological niche modeling. Evol Ecol 29:643–655. https://doi.org/10.1007/s10682-015-0786-z

Soong JL, Janssens IA, Grau O et al (2020) Soil properties explain tree growth and mortality, but not biomass, across phosphorus-depleted tropical forests. Sci Rep 10:1–13. https://doi.org/10.1038/s41598-020-58913-8

Steeger H, Henkel TW, Helal N et al (2019) Rarity of monodominance in hyperdiverse Amazonian
forests. Sci Rep 9:13822. https://doi.org/10.1038/s41598-019-50323-9
Torti SD, Coley PD, Kursar TA (2001) Causes and consequences of monodominance in tropical lowland forests. Am Nat 157:141–153. https://doi.org/10.1086/318629
Varela S, Lobo JM, Hortal J (2011) Using species distribution models in paleobiogeography: a matter of data, predictors and concepts. Palaeogeogr Palaeoclimatol Palaeoecol 310:451–463. https://doi.org/10.1016/j.palaeo.2011.07.021
Velazco SJE, Galvão F, Villalobos F, De Marco P (2017) Using worldwide edaphic data to model plant species niches: an assessment at a continental extent. PLoS One 12:1–24. https://doi.org/10.1371/journal.pone.0186025
Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. Trends Ecol Evol 19:639–644. https://doi.org/10.1016/j.tree.2004.09.011

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.