Global Ecology and Biogeography

WILEY

Global Ecology and Biogeography

RESEARCH ARTICLE OPEN ACCESS

Climate Predicts Tree Dispersal Strategies Across Brazil's Ecosystems

Fernanda C. Souza[1,2] | Vinícius Andrade Maia[3] | Renata Françoso[3] | Matheus Coutinho Freitas de Oliveira[3] | Ana Beatriz Ligo[1] | Bruno X. Pinho[4] | Jesús Aguirre-Gutiérrez[5,6,7,8] | Martin Sullivan[9] | Larissa Rocha-Santos[10] | Laura de Brito Évora[11] | Leonardo de Sousa Miranda[12] | María Natalia Umaña[13] | Pablo R. Stevenson[14] | Paulo Amador Tavares[15] | Adriane Esquivel-Muelbert[16] | Alan Bernardes da Silveira[17] | Ana Magalhães Cordeiro Teixeira[18] | Ângelo Gilberto Manzatto[19] | Átila Oliveira[20] | Aurora Levesley[21] | Beatriz Schwantes Marimon[22,23] | Ben Hur Marimon Junior[22,23] | Carlos Joly[24,25] | Carlos A. Peres[26,27] | Carolina V. Castilho[28] | Celice Alexandre Silva[29] | Danielle Storck-Tonon[30] | Domingos Cardoso[31,32] | Domingos de Jesus Rodrigues[33] | Dora M. Villela[34] | Douglas C. Morton[35] | Eddie Lenza de Oliveira[22,23] | Edmar Almeida de Oliveira[36] | Eduardo Hase[20] | Eliana Ramos[37,38] | Elmar Veenendaal[39] | Erika Berenguer[5,12] | Estelle Dugachard[40] | Everton Almeida[41] | Fabrício Alvim Carvalho[42] | Fernanda Coelho de Souza[21,43,44] | Fernando Elias[45] | Filipe M. França[2,46] | Flávia R. C. Costa[47] | Flávia Delgado Santana[20] | Flávia Rodrigues Barbosa[33] | Foster Brown[48] | Françoise Yoko Ishida[49] | Frederico Augusto Guimarães Guilherme[50] | Gabriel Damasco[51] | Georgia Pickavance[21] | Geraldo Wilson Fernandes[52] | Haroldo Cavalcante de Lima[31] | Helena Lemos Rocha[53,54] | Henrique A. Mews[55] | Iêda Leão do Amaral[56] | Igor Broggio[34] | Jarcilene Silva de Almeida[57] | John E. B. L. Cunha[58] | Jhonathan Oliveira Silva[59] | Joice Ferreira[2,60] | Jorcely G. Barroso[61] | Jose Ataliba Mantelli Aboim Gomes[62] | José Roberto Rodrigues Pinto[63] | Juliana Schietti[64] | Karina Melgaço[21,43] | Laura B.Vedovato[65,66] | Leandro V. Ferreira[67] | Letícia Fernandes da Silva[68] | Lilian Blanc[69] | Lilian Gomes da Silva Rocha[63] | Lola da Costa[70] | Lucas José Mazzei de Freitas[60,63] | Luciana F. Alves[71] | Luciano Paganucci de Queiroz[72] | Luiz E. O. C. Aragão[66] | Magna Soelma Beserra de Moura[73] | Maíra Benchimol[10] | Marcelo Brilhante de Medeiros[74] | Marcelo Fragomeni Simon[74] | Marcelo T. Nascimento[34] | Marcos José Salgado Vital[75] | Marcos Silveira[76] | Maria A. Carniello[77] | Maria das Dores Magalhães Veloso[78] | Mário Marcos do Espírito Santo[78] | Maryane B. T. Andrade[20] | Milton Serpa de Meira Junior[79] | Nayane C. C. S. Prestes[22] | Oliver L. Phillips[21] | Patrícia Corrêa Guedes de Souza[54] | Paulo Eduardo Barni[80] | Paulo S. Morandi[22,23] | Peter Moonlight[81] | Philip Fearnside[20] | Plinio Sist[69] | Priscyla Maria Silva Rodrigues[59] | Queila S. Garcia[82] | Rafael F. Ramos[83] | Rainiellen de Sá Carpanedo[33] | Reinaldo Imbrozio Barbosa[84] | Ricardo de Oliveira Perdiz[75,85] | Richarlly da Costa Silva[86] | Robson dos Santos Alves da Silva[87] | Rodolfo Souza[88] | Rubens Manoel dos Santos[3] | Sabina Cerruto Ribeiro[89] | Samuel de Padua Chaves e Carvalho[90] | Sílvia da Luz Lima Mota[91] | Simone Aparecida Vieira[92] | Simone Matias de Almeida Reis[22,89] | Ted R. Feldpausch[93] | Tereza Sposito[94] | Thaiane Rodrigues de Sousa[95] | Thiago Metzker[94] | Tiina Särkinen[96] | Toby A. Garder[97] | Toby Pennington[96,98,99,100] | Tomas F. Domingues[101] | Tony César de Sousa Oliveira[102,103] | Vitor de Andrade Kamimura[104] | Wendeson Castro[76] | Wesley Jonatar Alves da Cruz[105] | William E. Magnusson[20] | Yadvinder Malhi[5,6] | Yvonne V. Bakker[106] | Jos Barlow[12]

Correspondence: Fernanda C. Souza (fesouzasjdr@hotmail.com)

Received: 1 April 2026 | **Revised:** 7 July 2026 | **Accepted:** 9 July 2026

Handling Editor: Petr Keil

Keywords: climate change | dispersal agents | dispersal modes | evapotranspiration | precipitation | primates | seed mass | temperature seasonality | tropical ecosystems | wind speed

For affiliations refer to page 13.

ABSTRACT

Aim: We investigate the distribution of dispersal modes—endozoochory (via animal ingestion), synzoochory (via animal seed accumulation) and anemochory (via wind)—and seed mass of woody plants across Brazilian ecosystems, testing whether their relative abundance in local plant communities could be predicted by climate, edaphic factors and dispersal agents (e.g., the main vectors moving seeds). We expected climate to be the main predictor of dispersal strategy abundance.

Location: Brazilian ecosystems (Amazon, Atlantic Forest, Cerrado and Caatinga).

Time Period: 2004–2020.

Major Taxa Studied: Woody angiosperms, trees and shrubs.

Methods: We assigned seed mass to 1984 species and dispersal mode to 2203 species within 300 old-growth and non-flooded plots. We employed Random Forest regressions to assess the influence of water availability, temperature seasonality, soil fertility, wind and frugivorous primates on the community-weighted means of seed mass and dispersal modes. We mapped predictions and estimated the area of applicability to identify regions where predictions were reliable.

Results: Climate was the main factor predicting tree dispersal strategies, followed by dispersal agents or soil properties. Endozoochory was prevalent in wetter regions with high water availability and greater diurnal temperature range. Anemochory was more common in drier regions with low water availability, stronger winds and greater diurnal temperature range, whereas synzoochory was more abundant in areas with low temperature seasonality. Heavier seeds were more prevalent in regions with less seasonal temperatures, smaller diurnal temperature ranges and weaker winds.

Main Conclusions: Climate is a key predictor of the distribution and prevalence of plants with contrasting dispersal modes and seed masses. We thus expect functional responses to environmental change associated with shifts in the relative abundance of dispersal strategies, with implications for ecosystem services. Our results also highlight the critical role of plot-based datasets in improving our understanding of how environmental factors influence seed dispersal strategies across tropical ecosystems.

1 | Introduction

Identifying the biotic and abiotic factors that drive species assembly in ecological communities, and the mediating functional traits, is a critical goal in functional biogeography (Violle et al. 2014). Morphological seed traits, such as size, are essential to characterize two key dimensions of the regeneration niche underpinning dispersal and establishment functions (Saatkamp et al. 2018). For example, seed mass influences several aspects of plant life history, including seed production (Moles 2018), dormancy (Rubio de Casas et al. 2017) and seedling survival and competition (Westoby et al. 1996; Walters and Reich 2000; Moles et al. 2007). Furthermore, dispersal mode can indicate affinity to abiotic and biotic dispersal vectors, influencing how seeds reach sites for germination, seedling establishment and survival (van der Pijl 1982; Pinho et al. 2025). Despite increasing knowledge on the functional significance of seed mass and dispersal modes, the distribution of dispersal strategies across tropical ecosystems and their ecological predictors remains poorly understood. Assessing the predictors underlying plant dispersal modes and seed mass across broad-scale ecological gradients can help clarify the current patterns and anticipate potential shifts in plant communities under environmental change.

Seed mass and the proportion of zoochory—animal-dispersed seeds—increase from temperate to tropical regions, and from dry to humid forests (Westoby et al. 2002; Moles et al. 2007; Chen et al. 2017). These empirical patterns have supported the development of three non-mutually exclusive hypotheses that try to explain the mechanisms underpinning patterns in dispersal mode and seed mass. The resource-availability hypothesis proposes that high metabolic costs are involved in constructing and maintaining fleshy fruits and large seeds—typically zoochorous species—compared to non-fleshy fruits and small seeds (Moles et al. 2007; Chen et al. 2017). Thus, the proportion of zoochorous plants is expected to increase in areas with high resource availability, characterized by high and constant air temperature and water availability from rainfall, and fertile soils (Tabarelli et al. 2003; Butler et al. 2007; Correa et al. 2015, 2022). However, this hypothesis has received limited support at community-level scales in neotropical forests, indicating inconsistent relationships between air temperature, climate seasonality, soil fertility and the dominance of dispersal modes (Correa et al. 2015, 2022).

A second hypothesis, the disperser-availability hypothesis, predicts that the distribution of dispersal modes is associated with the availability of dispersal agents (Griffiths and Lawes 2006; Umaña et al. 2018; Galetti et al. 2013). This hypothesis has received more support for wind and water-dispersed plants (Umaña et al. 2018; Correa et al. 2015, 2022) while results have been mixed for animal-dispersed plants when related to the availability of frugivorous primates (Correa et al. 2015, 2022).

Last, the recruitment hypothesis predicts that large seeds have a greater competitive advantage than small seeds by promoting higher seedling performance in shaded conditions associated with closed-canopy forests (Westoby et al. 1996; Eriksson et al. 2000; Walters and Reich 2000). This hypothesis has been supported when associated with physical factors typically found in deeply shaded conditions such as warmer, wetter and less seasonal climates (ter Steege et al. 2006; Malhado et al. 2015). While higher seedling survival from large seeds is expected, their dispersal potential without biotic or water vectors tends to be low (Thomson et al. 2011). Thus, dispersal by animals and water plays a crucial role for heavy-seeded species, allowing transport far from the parental plant and avoiding negative density-dependent effects (Janzen 1970; Connell 1971; Comita et al. 2014).

Despite advances in our understanding of the distribution patterns of seed mass and dispersal modes and their ecological correlates, the varied empirical patterns and lack of clear support for existing hypotheses indicate there are many gaps that require further investigation. Some of these relate to the nature and scale of the data typically used. Large-scale patterns in seed mass and dispersal modes are often derived from species occurrence data aggregated across spatial grid cells or latitudinal bands (Moles et al. 2007; Malhado et al. 2015), which typically fail to capture the ecological processes operating at community-level scales. Furthermore, most studies have focused on the distribution of dispersal modes within specific regions (Almeida-Neto et al. 2008; Correa et al. 2015, 2022), leaving the ecological gradients related to plant dispersal strategies at broader scales relatively unexplored. For instance, Brazil, which spans over half of South America, encompasses a remarkable diversity of geological formations, climates, soils, fauna and vegetation types, delivering ecosystem services across major tropical biomes recognized globally, including rainforests, dry forests and savannas (Ab'Saber 2003; IBGE (Instituto Brasileiro de Geografia e Estatística) 2012; MMA (Ministério do Meio Ambiente) 2024). Examining these issues across the major Brazilian ecosystems, which span a wide range of humid to arid conditions, and incorporating plot-based data could provide more robust insights into the mechanisms underlying the dispersal strategies of plant assemblages in tropical ecosystems.

To advance our understanding of variation in dispersal strategies at continental scales, we assessed the local abundance of seed mass and dispersal modes of 1984 species in 300 old-growth vegetation plots across major Brazilian ecosystems, including Amazon, Atlantic Forest, Cerrado and Caatinga. We investigated the distribution of the most common dispersal modes—endozoochory, anemochory and synzoochory—and seed mass, expressed as community-weighted means, to examine whether and how climate, soil properties and dispersal agents (e.g., the main vectors moving seeds) influence their abundance in local plant communities. For vertebrate dispersers, we focus exclusively on primates, as they are a key taxonomic group for seed dispersal in the Neotropics and encompass functional groups relevant to both endozoochory and synzoochory (van Roosmalen 2013; Fuzessy et al. 2016).

We examine how seed mass and dispersal modes vary among regions, and which explanatory variables are most important in predicting these dispersal strategies. Our overarching expectation is that climate influences the relative abundance of dispersal modes and seed mass, reflecting the key role of temperature and precipitation in driving dissimilarities in vegetation composition across tropical ecosystems (Dexter et al. 2018; Feeley et al. 2020). Specifically, we expect a higher prevalence of zoochorous and heavy-seeded plants in regions with lower temperature seasonality and wetter climates because (i) lower seasonality in temperature would create conditions favouring large-seeded plants, whose seeds require longer growing seasons to develop and are less likely to exhibit dormancy under variable climates (Rubio de Casas et al. 2017); and (ii) higher water availability is critical for the development and maintenance of fleshy fruits, whereas non-fleshy fruits are less dependent on water availability (Tabarelli et al. 2003). Conversely, more seasonal and drier environments should favor anemochory and light-seeded plants. In addition, we expected differences within the animal dispersed plants. Although zoochorous plants require high water availability for fruit development (Tabarelli et al. 2003; Almeida-Neto et al. 2008; Correa et al. 2015), this dispersal strategy encompasses a wide variation in seed size (Souza 2023), which may be influenced by contrasting temperature ranges (Correa et al. 2022). We predict that endozoochory would be more prevalent in regions with greater temperature seasonality, while synzoochory, characterized by the largest seed sizes among dispersal modes (Correa et al. 2022; Souza 2023), would be more strongly associated with regions that experience lower temperature seasonality. We discuss our results in terms of the role each variable group plays in predicting seed mass and dispersal modes, and how informative they are for understanding both current patterns and potential responses of tropical tree floras to climate change and anthropogenic disturbances.

2 | Methods

2.1 | Study Area and Vegetation Plot Data

We used 300 plots (0.1 ha–1.6 ha; ~32–126 m per side; totalling 247.24 ha) distributed across seven Brazilian ecosystems, including four Amazonian biogeographical regions: Central ($n = 62$), Eastern ($n = 45$), Southern ($n = 54$) and Western ($n = 45$), as well as the Atlantic Forest ($n = 29$), Cerrado ($n = 36$) and Caatinga ($n = 29$) (Figure 1). These regions encompass their major vegetation types: the Amazon and Atlantic Forest regions are dominated by moist forests, the Cerrado by Brazilian savanna and the Caatinga by dry forest. For analytical purposes, the Amazon was divided into four regions to reduce sampling imbalance associated with its large spatial extent and the resulting disproportionate availability of plots. All plots were located on non-flooded soils and in old-growth vegetation. These inventories, carried out by multiple botanical investigators over the past two decades (2004–2020), were curated at ForestPlots.net (Lopez-Gonzalez et al. 2011; ForestPlots.net et al. 2021). Although other floristic datasets are available, the use of forest-plot inventories allows the standardization of taxonomy and methodological criteria. Unlike floristic datasets, plot-level data provide community-level data including species abundances. These inventories enable us to establish consistent criteria for plant size inclusion (e.g., minimum diameter) and to consider the disturbance history of sites, allowing us to prioritize undisturbed or minimally disturbed forests. The area and total number of individuals and species sampled by region are summarized in Table S1.

The difference in sampling methods (e.g., plot size) should not affect our results as we focus on the relative abundance of dispersal strategies within local plant communities and the resulting variation across the Brazilian ecosystems. To further address this concern, we tested whether any systematic patterns or effects not captured by the model were associated with variation in plot size. No significant correlation was found between model residuals and plot area, indicating that any unexplained effects are unlikely to be associated with variation in plot size (Figure S1). In addition, to ensure standardized species, genus and family nomenclature, we verified taxon names against the Flora e Funga do Brasil database (BFG (Brazil Flora Group) et al., 2022; https://floradobrasil.jbrj.gov.br/consulta/).

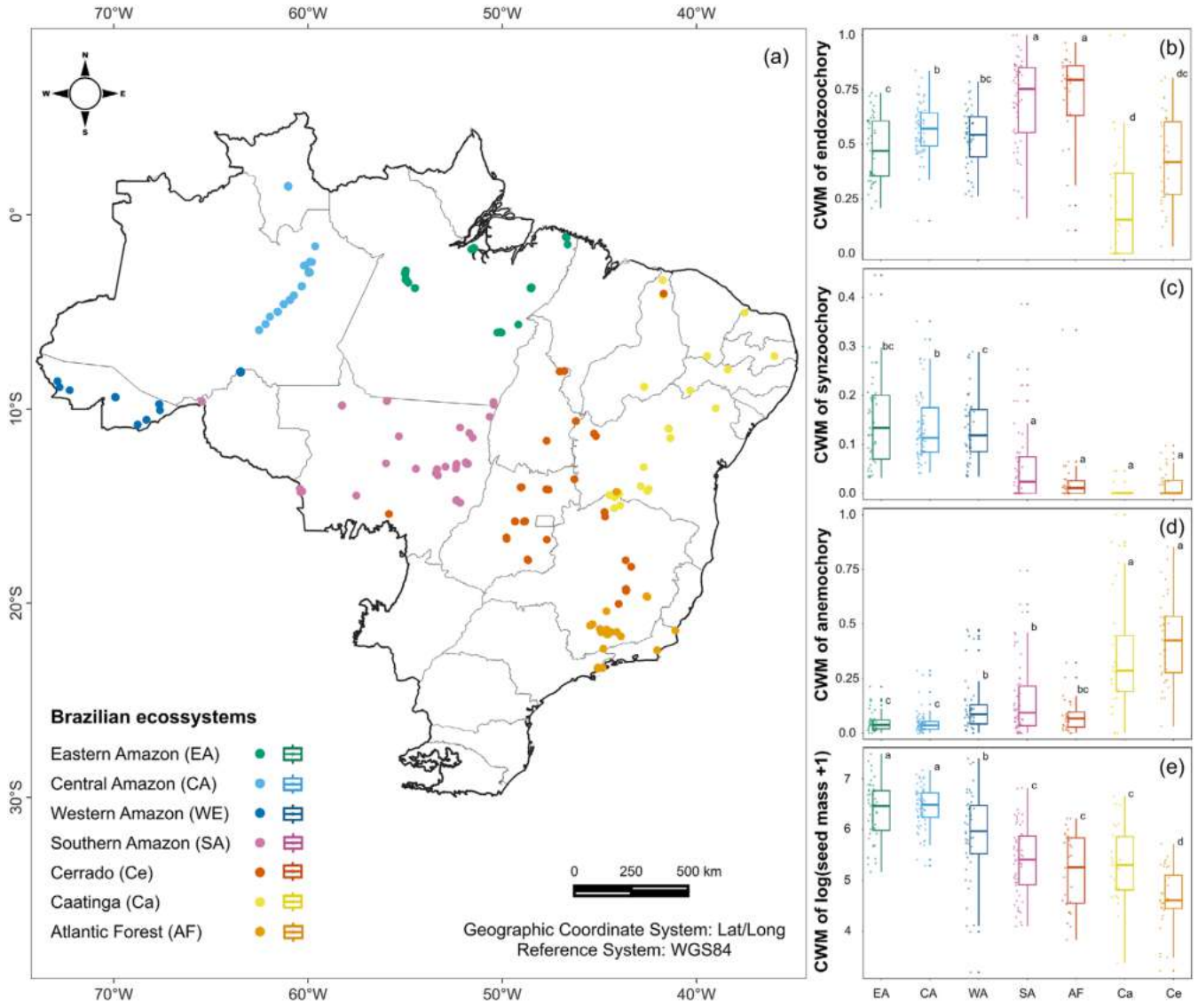


FIGURE 1 | Location of forest inventory plots where tree dispersal strategies were quantified (a). Variation in community-weighted means (CWM) of endozoochory (b), synzoochory (c), anemochory (c) and seed mass across Brazilian ecosystems. Colours in (a) correspond to regions in (b–e). Different letters represent statistical differences between regions. Jittered points are the actual values, boxplot (centre line; median), 25%–75% quartiles (box edges), < 1.5 times the interquartile range (whiskers) and extreme values (dots). $N = 300$ plots (Central Amazon = 62, Eastern Amazon = 45, Southern Amazon = 54, Western Amazon = 45, Atlantic Forest = 29, Cerrado = 36, Caatinga = 29).

2.2 | Vegetation Data

We used data from 172,185 individuals representing 541 genera and 2203 species (Table S1). To include woody stems, we established the following criteria. First, within each plot, species were ranked by stem abundance in descending order, and only those contributing to 80% of the cumulative number of individuals were included. Second, we focused on stems from tree (including palms) and shrub species, excluding herbaceous species, lianas and bamboos. Next, we applied distinct diameter thresholds for woody stems in each ecosystem: DBH ≥ 10 cm for the Amazon regions, ≥ 5 cm for the Atlantic Forest and Cerrado, and ≥ 3 cm for the Caatinga. These thresholds were based on the minimum diameter commonly reported in forest plot inventories for different ecosystems. For instance, drier systems typically use lower diameter thresholds due to the smaller average plant size (Moonlight et al. 2021). The effect of employing different thresholds for the two moist forest regions (Amazon and Atlantic Forest) was negligible in practice, as only 0.02% of individuals included from the Atlantic Forest were ≤ 10 cm diameter after applying the initial criteria. By employing all these steps, we sought to ensure that the most representative woody species from each region were represented in our datasets while avoiding potential biases that would arise by including non-woody plants.

2.3 | Dispersal Modes and Seed Mass

We gathered information on the dispersal modes from a mix of sources, including scientific literature, such as books, papers and theses, as well as online databases (e.g., Seed Information Database

[SID]; Royal Botanic Gardens, Kew 2020). Our classification divided species and morphospecies into two categories: zoochorous and non-zoochorous. The former included endozoochorous and non-endozoochorous (synzoochory and myrmecochory) dispersal modes, while non-zoochorous species were classified according to anemochory, barochory, autochory and hydrochory (see Table 1).

Due to the lack of information denoting primary dispersal modes, we also applied the classification scheme proposed by Hawes et al. (2020), separating modes into *sensu stricto* and *sensu lato* subcategories. *Sensu stricto* referred to species or morphospecies with definitive dispersal modes, while *sensu lato* referred to those with two or more possible dispersal modes (see Table S2). For the main analyses, we focused exclusively on the *sensu stricto* category because it is less ambiguous, including only cases in which species have a single, well-defined dispersal mode. Results incorporating both *sensu stricto* and *sensu lato* classifications are presented in the Supporting Information (Table S9; Figures S13–S18).

Dispersal modes were predominantly imputed at the genus level, accounting for ~93% of all individuals and 85% of species. Species for which dispersal modes could not be assigned (i.e., completely unknown) were negligible, representing only 0.58% of individuals and 2.77% of species (Table S3).

We collected and compiled seed-mass data from scientific literature, global databases such as TRY Plant Trait Database (Kattge et al. 2020), BIEN Botanical Information and Ecology Network (Enquist et al. 2026), and SID Seed Information Database (Royal Botanic Gardens, Kew 2020) and personal communications. We prioritized the data compilation from dry mass. However, in cases where dry materials were unavailable (28% of individuals), we converted fresh mass to dry mass using the following correction factor: dry measurements = $[0.92 \times \text{fresh mass}]^{0.94}$ (Moles et al. 2005).

We focused on collecting seed-mass data from endozoochorous, anemochorous and synzoochorous species (*sensu stricto* and *sensu lato*), which together span a wide range of seed masses and represent 88.8% of the total number of individuals in our database ($n = 152{,}966$) (Table S3). For each species and genus, a single value for seed mass was calculated as the median of all available records, ensuring a consistent representation across sources. Within these records, 647 species had complete seed-mass measurements at the species level (approximately 32.6% of the total). Since seed size traits tend to be conserved along the phylogenetic tree (Fuzessy et al. 2022), for species lacking species-level data, we assigned median genus-level values (58% of species). For the remaining 9.4% of species without genus-level information, missing seed-mass values were estimated via multivariate imputation with chained equations, with each value imputed five times using predictive mean matching in the *mice* package (van Buuren and Groothuis-Oudshoorn 2011). To further improve the accuracy of imputed seed-mass values, we incorporated seed width and length, which are moderately to strongly correlated with seed mass (Souza 2023). Seed width was available for 33% of species ($n = 663$) and seed length for 38% of species ($n = 767$). When species-level measurements were unavailable, genus-level values were used, covering 49% of species for width ($n = 986$) and 44% of species for length ($n = 882$).

The imputed seed mass data accounted for only 4.41% of all individuals across species and plots ($n = 152{,}966$) and largely overlapped with the distribution of observed values (Figure S2). To assess the reliability of the imputation procedure, we artificially introduced gaps in species with complete seed-mass data by randomly removing 10% of observed values, approximately matching the proportion of missing data, and imputed them using the same framework (five multiple imputations). Imputation accuracy was quantified using the root mean square error (RMSE), which averaged 13,780 ± 7,537 mg (~13.7 ± 7.5 g). Finally, to evaluate whether imputation influenced model outcomes, we fitted models using only complete species-level data and compared their performance with models including both complete and imputed values.

TABLE 1 | Definition of dispersal modes, following van Roosmalen (2013), Kuhlmann and Ribeiro (2016), Hawes et al. (2020) and Correa et al. (2022).

Dispersal mode	Description
Endozoochory	Fleshy diaspores with small to large seeds dispersed through the gut of animals
Synzoochory	Fleshy or dry diaspores with large seeds carried externally through the body of bats, monkeys (primary dispersal) or scatter hoarders (secondary dispersal)
Myrmecochory	Dry diaspores with small seeds associated with a lipid-rich food body (elaiosome) carried by ants
Anemochory	Diaspores without fleshy structures carried by the wind
Barochory	Diaspores carried by gravity
Autochory	Dry diaspores carried by 'itself' through explosive opening or elastic dehiscence (active autochory), or without any evident adaptation to specific dispersal agents (passive autochory)
Hydrochory	Diaspores without fleshy structures and floating abilities carried by water

2.4 | Response Variables

Seed-mass data were scaled-up from the species level to the plot level by calculating the Community-Weighted Mean (CWM). For each plot p, CWM was calculated as $\text{CWM}_{tp} = \sum^{s}_{i-1} \text{AB}_{ip} \times t_i$, where AB_{ip} is the relative abundance of species i in plot p, calculated as the number of stems of species i divided by the total number of stems in the plot, and t_i is the median seed mass value of species i. Median values were used as a single representative species-level trait value in the CWM calculation, based on all available trait records, to provide a robust estimate that is less sensitive to outliers. Trait values were then log-transformed $\left[\log(\text{seed mass} + 1)\right]$ to reduce skewness in trait distributions. CWM values for seed mass were calculated using the *FD* package (Laliberté and Legendre 2010). This approach captures the

dominance of trait values within the community and provides a meaningful summary of community functional composition (Muscarella et al. 2017).

For dispersal modes, CWM represents the relative abundance of species belonging to each dispersal mode category (endozoochory, synzoochory and anemochory) within each plot.

2.5 | Explanatory Variables

For each plot, we extracted long-term mean values (1970–2000) of eight WorldClim v2.0 bioclimatic variables (1-km resolution), capturing climatic seasonality and intra-annual variability in temperature and rainfall-driven water availability (Table S5) (Fick and Hijmans 2017). Additionally, we used potential evapotranspiration (PET), calculated from a set of WorldClim variables which involve minimum, maximum and average temperature, solar radiation, wind speed and water vapour pressure, as a proxy for water availability (Trabucco and Zomer 2018). PET reflects the amount of water expected to be removed from the land surface through evapotranspiration processes annually and provides a more ecologically meaningful estimate of plant-available water compared with precipitation, as it integrates both atmospheric demand and energy balance (Moles et al. 2014; Mod et al. 2016).

We also extracted mean values of five soil physical and chemical properties that are known to influence nutrient availability and, therefore, soil fertility (Table S5). These variables were obtained from SoilGrids version 2.0 (250-m resolution) and represent average conditions over the 0–200 cm soil depth layer (Poggio et al. 2021).

To assess the role of biotic dispersal agents on the CWM values of seed mass and zoochorous dispersal modes, we used the body-size breadth of frugivorous primates as a functional index. Primates are among the most important arboreal vertebrates, constituting 25%–40% of the frugivorous biomass in South American tropical forests (Haugaasen and Peres 2005). A greater variation in body mass across frugivorous primate assemblages may indicate complementary roles in plant-frugivore interactions, increasing the chances that a fruit-consuming vertebrate acts as an effective seed disperser (Fuzessy et al. 2022). Seeds dispersed by primates often show higher germination rates and faster germination compared to non-dispersed seeds, reflecting enhanced dispersal quality (Fuzessy et al. 2016). Moreover, primates play a key role as arboreal dispersers of large-seeded tree species, either through endozoochory or synzoochory, a role that is poorly represented in avian assemblages (van Roosmalen 2013; Fuzessy et al. 2018). Their richness is also strongly congruent with other taxa, including plants (Stevenson 2001), and their biogeography and association with areas of endemism align with birds (da Silva et al. 2005). While birds are also important seed dispersers, we did not include them in the analyses because deriving a bird functional metric directly comparable to our primate-based index would require substantial additional data processing across multiple species and guilds to integrate these components into a spatially explicit functional index at a continental scale, which was not trivial within the scope of this study.

To calculate body-size breadth within each assemblage, we first obtained primate species lists for each georeferenced location using species distribution polygons available from the International Union for Conservation of Nature (IUCN (International Union for Conservation of Nature) 2023). We then extracted diet information from the *EltonTraits* database, classifying species according to their level of frugivory (Wilman et al. 2014). We also compiled the body mass of each frugivorous primate species and calculated the interquartile range (IQR) of this trait per plot (Table S6). The IQR, representing the middle 50% of the data distribution (Q3–Q1), is robust to outliers and skewed distributions, making it a reliable measure of functional breadth of frugivorous primates in each plot.

To assess the role of abiotic dispersal agent on CWM values of anemochory, we obtained annual mean wind speeds (at 10 m–200 m above the ground) for each plot from the Global Wind Atlas maps at 250-m resolution (Davis et al. 2023).

These three sets of explanatory variables (climate, soil and dispersal agents) are known to influence large-scale plant trait distributions in tropical regions (Moles 2018), including seed size and dormancy, the prevalence of fleshy fruits and the occurrence of major dispersal modes (Rubio de Casas et al. 2017; Huang et al. 2021; Correa et al. 2022). Since climate and soil variables were strongly correlated ($|r| > 0.7$) (Figure S3), we performed a variance inflation factor (VIF) analysis to avoid multicollinearity in the models for each response variable. Only variables with VIF < 4 were retained for model construction.

The final models included four climate variables: PET, precipitation of the wettest quarter, mean diurnal temperature range and temperature seasonality; three soil variables (clay content, cation exchange capacity and organic carbon stock) and one variable associated with dispersal agents, which depended on the dispersal mode (e.g., the IQR of primate body mass for zoochory and wind speed for anemochory). For seed mass, both dispersal-related variables were included.

2.6 | Data Analyses

To assess variation in the community-weighted mean (CWM) of seed mass and dispersal modes across regions, we used generalized linear models (GLMs) implemented in the *glmmTMB* package (Brooks et al. 2017) with beta-binomial distribution for proportional data and Gaussian distribution for continuous data. Post hoc tests were conducted using the *emmeans* package (Lenth 2023) to assess differences in CWM values across Brazilian ecosystems.

We trained Random Forest (RF) models to investigate the influence of climate, soil and dispersal agents in predicting CWM values of seed mass and dispersal modes (endozoochory, anemochory and synzoochory). RF is a robust machine-learning algorithm that captures complex, non-linear relationships between response and explanatory variables by combining multiple decision trees. For each tree, the algorithm applies bootstrap aggregating (bagging), first selecting a random sample from the original dataset to create a bootstrap dataset and then choosing a random subset of variables for each tree. Predictions are

obtained by averaging the outcomes of all individual trees, which helps avoid overfitting—a common issue in a single decision tree (Breiman 2001).

To optimize the performance of our models and minimize overfitting, we used a hyperparameter tuning process with five-fold cross validation using the *caret* package (Kuhn 2008). To find the best hyperparameters, we explored a set combination of hyperparameters, including number of trees (*ntree*: 50 to 300 by 20), number of variables as candidates at each split (*mtry*: 3 to 9, by 1) and size of terminal nodes (*nodesize*: 5 to 20 by 2). For each combination of hyperparameters, we trained RF models using four folds as the training set, while the remaining one was allocated as the test set. This iterative process was repeated until each fold had served as a test set. The combination of hyperparameters that yielded the lowest root mean square error (RMSE) was selected as optimal. The final RF model, with the best hyperparameters, was trained on the full dataset using the *caret* package (Kuhn 2008) and the *ranger* method (Wright and Ziegler 2017) to generate predictions and assess the influence of each predictor. Variable importance was quantified for each RF model using the model's internal importance scores. To express predictor contributions, we normalized each variable's importance by the sum of all importance values and multiplied it by the cross-validated R^2 (obtained from the 5-fold CV described below). The sum of the contributions of all variables within a predictor group represents their relative contribution to the predictive power of the model.

To assess the influence of each predictor on model outcomes, we generated partial dependence plots (PDPs) for all explanatory variables retained in each RF model. PDPs show the mean marginal influence of each predictor on the predicted response, while keeping all other predictors constant at their sample mean (Friedman 2001).

We assessed spatial autocorrelation in model residuals using Moran's index (Moran I) using the *spdep* package (Bivand et al. 2015). To select the optimal number of nearest neighbours (k), we tested values from one to 50 and identified the smallest k that produced a fully connected network while monitoring Moran's I for each k. Using this optimal k, Moran's I was calculated for residuals with asymptotic, randomization and Monte Carlo tests (999 permutations) to evaluate the presence and significance of spatial autocorrelation. Spatial patterns were visualized with Moran correlograms using the *ncf* package (Bjørnstad 2022) for both the observed response and models' residuals to examine the scale of spatial dependence. Moran's I values for all model residuals were close to zero (ranging from −0.047 to −0.019) and non-significant across all methods ($p > 0.05$; Table S8), indicating no evidence of spatial autocorrelation.

Although low Moran's I values indicate that the models successfully captured the spatial structure in the training data, they do not necessarily imply that the inferred relationships are transferable beyond the sampled locations. Therefore, we evaluated model performance using both non-spatial and spatial cross-validation. Non-spatial cross-validation assessed predictive performance under random data partitioning, whereas spatial cross-validation provided a more stringent assessment of transferability across spatially independent locations(Roberts et al. 2017).

For non-spatial cross-validation, we applied two complementary approaches. First, we divided the data into five equally sized folds, with each fold iteratively left out of model training and used as independent test data. Folds were stratified by region to ensure proportional representation of each region across the folds. This procedure was applied to the full (unbalanced) dataset. In the second approach, a balanced stratified procedure was implemented. Specifically, for each of 100 iterations, 29 plots per region (matching the smallest region) were randomly sampled without replacement within each iteration, allowing plots to be reused across iterations. The same five-fold stratified cross-validation was then applied to each subset, resulting in a total of 500 model evaluations.

For spatial cross-validation, we performed a spatial blocking procedure in which sampling locations were grouped into spatial blocks defined at three spatial scales (300, 500 and 800 km). The spatial blocks were assigned to five folds. For each spatial scale, folds were iteratively left out of model training and used as independent test data. These spatial scales represent increasing levels of spatial separation between training and validation data, allowing us to evaluate model generalizability under progressively more challenging spatial extrapolation scenarios.

Model performance was quantified in all cross-validation procedures using three metrics, averaged across folds and repetitions: the root mean square error (RMSE) between predicted and observed values, the correlation coefficient between predicted and observed values, and the coefficient of determination ($R^2 = 1 -$ residual sum of squares/total sum of squares).

2.7 | Mapping Seed Dispersal Strategies

RF was used to predict seed mass and dispersal modes based on climate, soil and dispersal agents across Brazilian ecosystems. To create prediction maps for each response variable, we constructed a raster stack of all model variables at a grid resolution of 0.0417° × 0.0417° (~5-km^2 grid-cell) using the *terra* and *raster* packages (R. Hijmans 2025a, 2025b). We obtained raster layers for soil, climate variables and wind speed from their respective spatial sources. For attributes of primate communities, we converted species-distribution polygons into a raster and calculated the interquartile range (IQR) of primate body mass values for each grid cell. Model predictions were masked to areas classified as natural vegetation using land-cover data from MapBiomas collection 10 (Souza et al. 2020; MapBiomas Project 2025).

We evaluated the spatial applicability of model predictions using the method proposed by Meyer and Pebesma (2021) implemented in the *cast* package (Meyer et al. 2025). The area of applicability (AOA) represents the region where the model's relationships, based on training data, are valid and where the model's cross-validation performance can be reliably applied. To define the AOA, we first calculated the dissimilarity index (DI), which quantifies the minimum Euclidean distance between

a new prediction point and the nearest training data point in the space of explanatory variables, with predictors weighted by their respective importance in the model. Values greater than one indicate that a point is more dissimilar to its nearest training data point than the average dissimilarity among training data points. Therefore, the AOA is defined by applying a binary DI threshold, where values above the 75th percentile plus 1.5 times the interquartile range of the DI distribution are removed, and the maximum remaining DI is used as the threshold. Prediction points with DI values below this threshold are considered within the AOA, whereas those exceeding it are considered outside the AOA. All spatial and modelling analyses were conducted in R version 4.3.1 (R Core Team 2023).

3 | Results

3.1 | Variation in Tree Dispersal Strategies

Endozoochory was the most prevalent dispersal mode across all studied plots, followed by anemochory and synzoochory, with all three showing significant differences among regions (Tables S3, S4). CWM values of endozoochory were highest in the Southern Amazon and Atlantic Forest compared to other regions (Figure 1b), while anemochory peaked in the Cerrado and Caatinga (Figure 1d). Synzoochory had the highest CWM values in the Eastern, Central and Western Amazon compared to Southern Amazon, Atlantic Forest, Caatinga and Cerrado (Figure 1c). CWM of seed mass was significantly higher in the Eastern and Central Amazon than in the five other regions (Figure 1e; Table S7).

3.2 | Explanatory Variables Influencing Tree Dispersal Strategies

To identify the main ecological factors influencing variation in seed mass and dispersal modes, we constructed RF models using three sets of explanatory variables: climate, soil and dispersal agents. For endozoochory, the most important explanatory variables were associated with water availability and seasonality and variability in climate (Figure 2). Endozoochory increased with precipitation in the wettest quarter and mean diurnal temperature range and declined with higher potential evapotranspiration (Figure S5). For synzoochory, the most important explanatory variable was related to seasonality in climate, with a negative relationship between synzoochory and temperature seasonality. The next most important variables were mean diurnal temperature range and precipitation in the wettest quarter, with the former showing a negative relationship with synzoochory and the latter mostly constant along its gradient but showing a sharp increase in synzoochory around 900 mm–1000 mm (Figure 2; Figure S6). Other climate variables (e.g., seasonality of temperature for endozoochory and potential evapotranspiration for synzoochory), along with dispersal agents and soil fertility, had little influence on both dispersal modes (Figure S5,S6). For anemochory, the most important explanatory variables were mainly associated with water availability and dispersal agents. Anemochory was positively associated with potential evapotranspiration and wind speed. Soil variables, together with climatic seasonality and variability, had little influence on the dominance of anemochory (Figure 2; Figure S7).

For seed mass, variability and seasonality in climate and dispersal agents were the most influential explanatory variables (Figure 2; Figure S10). Seed mass decreased with increasing mean diurnal temperature, seasonality of temperature and wind speed. Variables related to soil, biotic dispersal agents and water availability had little effect on CWM values of seed mass (Figure S8,S11).

3.3 | Performance of Models

Using non-spatial cross-validation, we found positive correlations between predicted and observed values across all dispersal modes and seed mass under both balanced and unbalanced approaches ($r = 0.63–0.78$; $R^2 = 0.36–0.55$). Root mean squared prediction errors ranged from 0.05 to 0.17 for dispersal modes. For seed mass, RMSE values ranged from 215.11 to 217.42 mg when using complete plus imputed data, and from 248.04 to 253.04 mg when only complete cases were considered (Table 2).

When model performance was evaluated using independent spatial blocks, correlations remained positive across all models and block sizes ($r = 0.21–0.67$). For dispersal modes, prediction errors remained broadly comparable to those obtained under non-spatial validation (RMSE = 0.07–0.21; Table 2). For seed mass, RMSE ranged from 260.76 to 289.81 mg using the complete plus imputed data and from 303.20 to 323.57 mg using only complete cases. Coefficients of determination decreased relative to non-spatial cross-validation, remaining positive for the 300- and 500-km block sizes (except for anemochory) and becoming negative for the 800-km block size (except for seed mass) (Table S10).

3.4 | Spatial Mapping

Spatial extrapolation of community-weighted mean seed mass (CWM) and dispersal modes (endozoochory, anemochory and synzoochory) showed a strong association with Brazilian ecosystems. For example, forested and moist regions were associated with high levels of endozoochory (Figure 3a), while open and drier regions were related to a high dominance of anemochory (Figure 3c). Within moist ecosystems, however, the highest CWM values of endozoochory were concentrated in specific regions, particularly in the Southern Amazon and Atlantic Forest (Figure 3a). Synzoochory was most prevalent across Amazonian regions, especially in the Eastern, Central and Western Amazon (Figure 3b), where the largest seed masses were also predicted (Figure 3d). Uncertainty in predictions varied among models, being higher for endozoochory (Figure S9a) and synzoochory (Figure S9b). These models were applicable only in limited regions, mainly corresponding to sampled areas, and showed low applicability in the Caatinga, large parts of the Amazon and the coastal Atlantic Forest. Predictions for anemochory and seed mass were generally more consistent, except in the northwestern Amazon, large portions of the Caatinga, the northern Cerrado and the coastal Atlantic Forest (Figure S9c,d). Predictions for seed mass based on complete data only were slightly more restrictive than those based on complete plus imputed data, although the AOA was very similar (Figure S12).

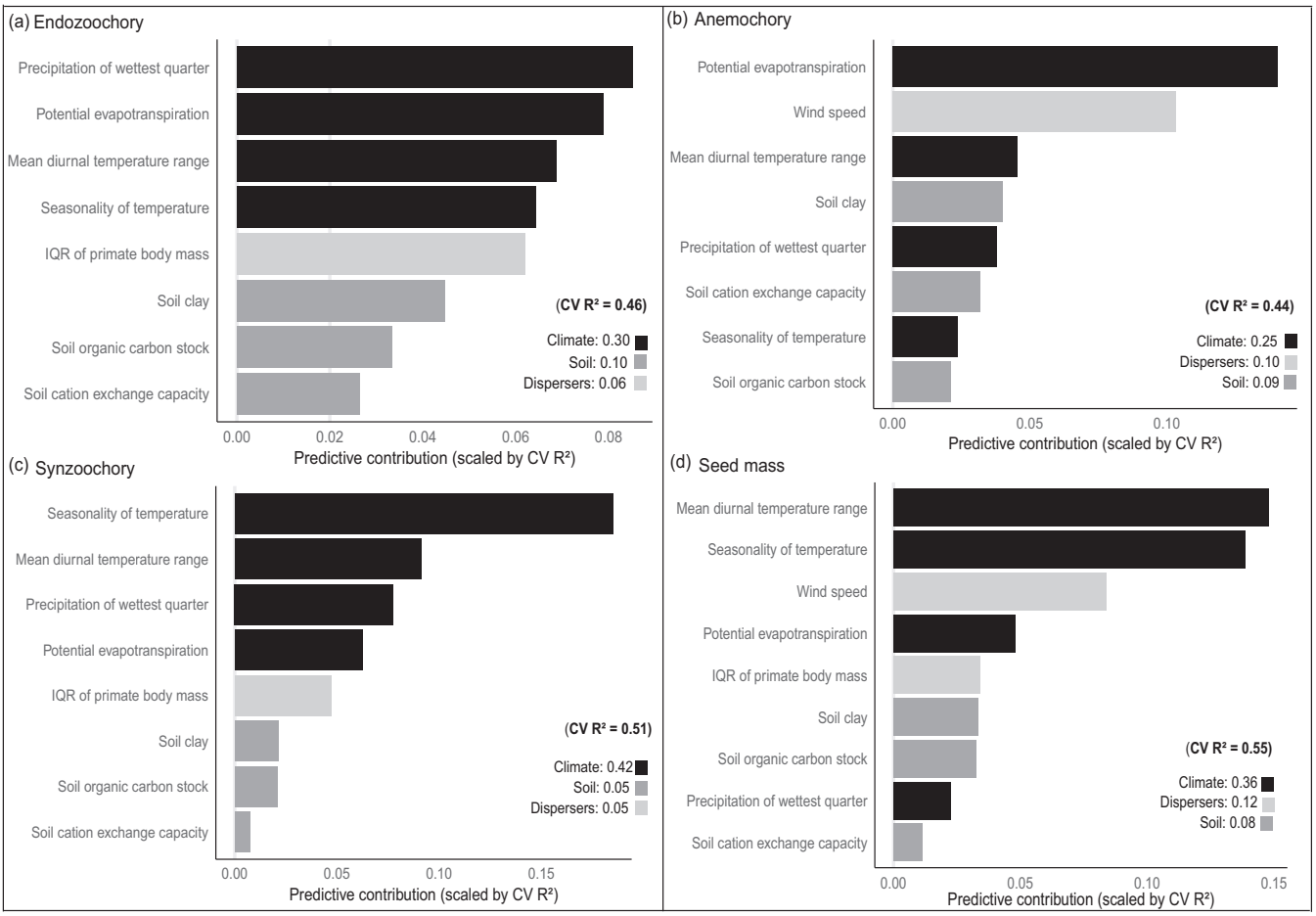


FIGURE 2 | Proportional contribution of each predictor to the predictive performance of the models for endozoochory (a), anemochory (b), synzoochory (c) and seed mass (d). Contributions were calculated from the models' internal importance scores and normalized so that their sum equals the cross-validated R^2. Bar lengths represent the proportion of explained variance, and colours indicate predictor groups: Climate (black), dispersers (light grey) and soil (dark grey).

TABLE 2 | Performance of models when applied to data not used for model fitting.

	Cross-validation					
	Balanced			Unbalanced		
Model RF	**RMSE (SD)**	**Correlation (SD)**	**R^2 (SD)**	**RMSE**	**Correlation**	**R^2**
Endozoochory	0.17 (0.02)	0.68 (0.09)	0.45 (0.14)	0.16	0.68	0.46
Anemochory	0.16 (0.02)	0.63 (0.09)	0.36 (0.14)	0.14	0.68	0.44
Synzoochory	0.05 (0.01)	0.74 (0.09)	0.52 (0.14)	0.06	0.73	0.51
Seed mass (complete + imputed data)	215.11 (45.35)	0.76 (0.07)	0.53 (0.12)	217.42	0.78	0.55
Seed mass (only complete data)	248.04 (46.21)	0.75 (0.07)	0.53 (0.12)	253.04	0.77	0.55

Note: Model performance for dispersal modes and seed mass was evaluated using stratified k-fold cross-validation under both balanced and unbalanced approaches (see methods). Performance metrics include: (1) Root mean square error (RMSE), indicating the average prediction error (proportion for dispersal modes and milligrammes for seed mass), (2) the correlation coefficient between observed and predicted values and (3) the coefficient of determination [1−(residual sum of squares/total sum of squares)]. SD = standard deviation.

4 | Discussion

Our large-scale assessment shows considerable variation in the dominance of dispersal strategies across Brazilian regions, and the important role of climate in predicting these strategies. While seasonality and variability in temperature and water availability were the most influential predictors of dominance of seed mass and dispersal modes, specific dispersal agents had

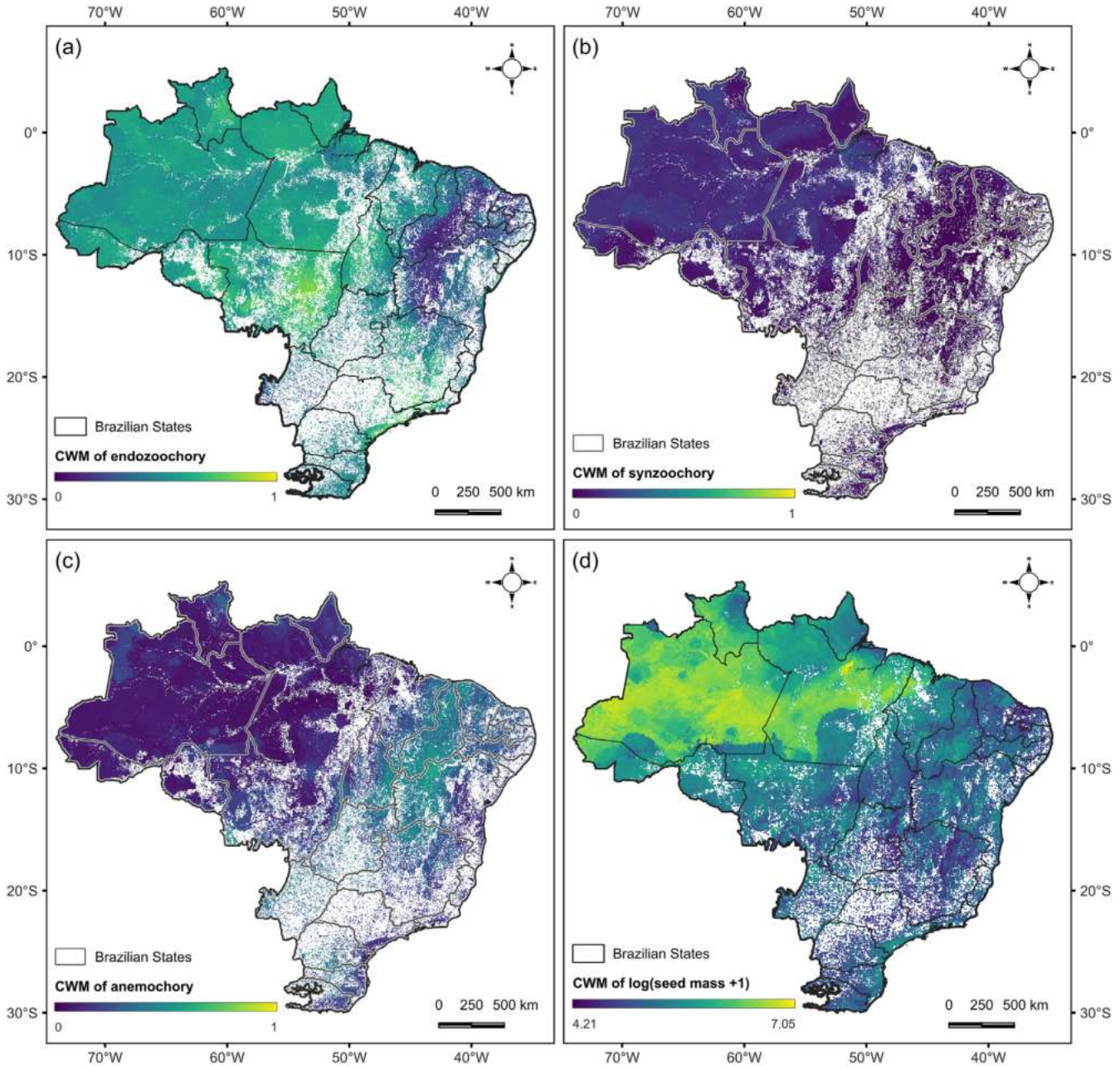


FIGURE 3 | Prediction map of community-weighted means (CWM) of endozoochory (a), synzoochory (b), anemochory (c) and seed mass (complete plus imputed data) (d) across Brazilian ecosystems. Dark purple and blue colours represent low CWM values, while ligth green and yellow colours represent high CWM values. Brazilian state boundaries are indicated by black lines in (a) and (d), and grey lines in (b) and (c). Non-natural vegetation areas are masked in white.

important roles within their respective dispersal modes, such as wind contributing to anemochorous groups and seed mass and primates linking to endozoochorous groups.

4.1 | The Role of Climate in Predicting Plant Dispersal Modes and Seed Mass

Climate was the main predictor of plant dispersal modes and seed mass across Brazilian ecosystems. Endozoochory and anemochory were dominant in regions with contrasting levels of evapotranspiration, with endozoochorous plants were more frequent in wetter environments (lower evapotranspiration), whereas anemochorous plants occurred more often in drier environments (higher evapotranspiration). These findings support the expectations that animal-dispersed fruits/seeds require high and consistently available water to develop and maintain their fleshy fruits, whereas wind-dispersed fruits/seeds are adapted for efficient wind transport, with light and dry structures making them less dependent on water availability (Tabarelli et al. 2003; Almeida-Neto et al. 2008; Correa et al. 2015). In contrast to previous studies in the Amazon (Correa et al. 2022), water availability showed a weaker overall influence on synzoochory, but its effects were spatially

structured and stronger in regions where synzoochory is more prevalent, particularly along precipitation gradients across Amazonian regions.

Wind-dispersed seeds have also been linked to environments with a high annual temperature range, whereas animal-dispersed seeds have been associated with both low and high annual temperature range (Correa et al. 2022). Extending these findings across a different temporal scale, our study shows that anemochory and endozoochory were dominant in environments with a high mean diurnal tempertaure range, whereas synzoochory was more frequent in areas with a lower mean diurnal temperature range. These trends are best explained from the perspective of seed size. Greater temperature variation could favor the dominance of small-seeded species (many of them anemochorous and endozoochorous), whose seeds are more likely to exhibit dormancy and survive adverse seasonal conditions (Rubio de Casas et al. 2017; Pinho et al. 2021). Conversely, synzoochorous species have typically large seeds that are known for their greater recruitment success in shady environments under stable climates (Eriksson et al. 2000), near the equator (Pinho et al. 2021). This adaptive advantage is due to the greater reserves that help seedlings tolerate low light conditions and establish successfully (Walters and Reich 2000). Indeed, heavy-seeded plants exhibited higher dominance under physical factors typically found in deeply shaded conditions, such as low temperature variation (seasonal and diurnal), which were most prevalent in the core regions of the Amazon.

4.2 | The Role of Dispersal Agents in Predicting Plant Dispersal Modes and Seed Mass

Beyond climate, our cross-regions assessment supports the association between dispersal agents and seed mass and dispersal modes (c.f. Correa et al. 2015, 2022). Anemochorous and light-seeded plants were more dominant in communities under high wind speed. Anemochorous and other light-seeded species were more dominant in communities under high wind speed, and primates were important for predicting the dominance of endozoochory; however, they were less important for predicting the dominance of synzoochory and heavy-seeded species. One potential reason explaining this last result may be due to defaunation processes that have affected the primates' populations. However, to test this hypothesis, would require local information on the relative abundance of primates and other dispersers across communities, both currently and over the past decades or centuries during which present tree communities assembled. For instance, large-bodied frugivores, which are key dispersers of large-seeded species, have been disproportionately lost across many tropical forest regions due to hunting and habitat loss (Dirzo et al. 2014). Furthermore, the legacy of Pleistocene megafaunal extinctions (e.g., in savannas) may have further contributed to the erosion of dispersal functions. Several large-seeded species in these systems still show morphological traits associated with dispersal by extinct megafauna, some of which may be partly explained by the influence of humans as disperser agents (van Zonneveld et al. 2018); such historical events suggest a long-term functional mismatch between plant traits and current disperser assemblages (Janzen and Martin 1982; Guimarães et al. 2008). Therefore, the arrival of humans and selective and historical loss of large-bodied dispersers can therefore lead to a functional decoupling between heavy-seeded plants (many of which rely on synzoochory) and the dispersal agents that originally dispersed them (Fricke et al. 2025).

Another reason for the weak relationship between the dominance of zoochory and seed mass, and the IQR of primate body mass is that, although primates are important dispersers, they are not the only ones. Birds and bats are key dispersers in open habitats (Eriksson et al. 2000), and scatter-hoarders play an important role in dispersing large, heavy seeds (Mittelman et al. 2021). Finally, the directionality of the relationship between zoochory and the IQR of primate body mass remains unclear, as animal-dispersed plants may also be important predictors of the biomass of frugivorous primates (Stevenson 2015). These effects could introduce positive feedback, where changes in the availability of animal-dispersed plants would positively impact the biomass of primates, and the high biomass of primates would increase seed dispersal and lead to greater recruitment of zoochorous plants. Despite these challenges, unraveling the role that frugivorous animals play in determining tree species composition is important given that the loss of large-frugivore biomass could affect the recruitment of large-seeded trees, especially those dispersed by primates (Terborgh et al. 2008), with potential impacts on other variables such as forest above-ground carbon stocks (Bello et al. 2015; Peres et al. 2016; Gardner et al. 2019).

4.3 | The Role of Soil in Predicting Plant Dispersal Modes and Seed Mass

Although soil fertility is known to affect forest dynamics by influencing productivity and turnover rates in tropical ecosystems (ter Steege et al. 2006), it showed only a weak influence on predicting dispersal modes and seed mass. There are some factors which may potentially influence this response. First, soil properties are expected to exert limited influence on size-related traits such as seed mass compared to plant economics traits (e.g., leaf area). Indeed, along the global spectrum of plant trait variation, seed mass is more closely associated with latitudinal gradients driven by water and energy availability rather than by soil nutrient conditions (Joswig et al. 2022). Second, it is important to point out that local-scale variation in soil properties is generally large. Thus, a coarse-scale soil database could not effectively capture the fine-grained variations in soil properties, such as nutrient supply, which is known to influence fruit resource construction at local scales (Carvalheiro et al. 2021).

4.4 | Performance and Applicability of Model's Predictions

Although our dataset spans most Brazilian ecosystems, it still represents only a small fraction of each region. Consequently, spatial predictions should be interpreted with caution, as some environmental conditions may fall outside the range represented by the training data (Meyer and Pebesma 2021). Areas of lower applicability were particularly evident in parts of the Amazon and Caatinga. In the Amazon, these areas overlap with

regions of low ecological research probability, which have been historically neglected due to limited accessibility and research infrastructure (Carvalho et al. 2023). In the Caatinga, however, lower applicability may reflect the distinctive climatic conditions, which occupy a relatively unique portion of the environmental space represented in our analyses.

The extent of model applicability also differed substantially among dispersal strategies. The endozoochory and synzoochory models showed substantially more restricted areas of applicability than other response variables, suggesting lower transferability beyond the environmental conditions represented in the training data. This pattern likely reflects the nature of the disperser predictors used in these models. For example, primate distributions were derived from species range polygons (IUCN (International Union for Conservation of Nature) 2023), which provide relatively coarse representations of species occurrence and may not capture fine-scale spatial variation in plant–frugivore interactions that influence seed dispersal. Consequently, these predictors may be less effective at representing ecological variation across environmentally distinct regions. In contrast, the anemochory and seed mass models were more broadly applicable, likely because wind speed, a key predictor in these models, is available as a spatially continuous, high-resolution variable (Davis et al. 2023), providing a more consistent representation of environmental gradients across the study region.

Model performance under spatial cross-validation further reflects these constraints. Although model residuals did not show evidence of spatial autocorrelation, model performance decreased under spatially structured evaluation, as expected under spatial cross-validation, with lower and occasionally negative R^2 values. In contrast, prediction errors remained comparable to those obtained under non-spatial cross-validation. Importantly, the relative importance of climatic predictors remained consistent across validation schemes, suggesting that the main environmental drivers of the models are robust, even when predictive performance decreases under spatial assessment.

4.5 | Data Shortfalls

While additional plot data would provide better knowledge of plant communities across different Brazilian regions, our understanding of the mechanisms underlying variation in seed mass and dispersal modes would also benefit from improved information on key predictors. Most notably, local-scale variation in soil properties and faunal composition could be important determinants of tree assemblages at smaller scales. Moreover, given the potential influence of historical and ongoing defaunation on large-seeded trees (many of them synzoochorous), future studies could test these effects on seed mass using defaunation indices (Benítez-López et al. 2026) or proxies such as the human footprint (Sanderson et al. 2002). Finally, seed mass and dispersal modes could be viewed as part of broader plant economic spectra, and it would be interesting to relate them to other traits, including those recently mapped, such as wood density (Sullivan et al. 2025) and leaf traits mapped from space (Aguirre-Gutiérrez et al. 2025).

4.6 | Insights From Plot-Based Data on Environmental–Plant Relationships

Although large-scale studies have reported a higher prevalence of zoochory and larger seeds in wetter and less seasonal environments, our plot-based approach provides new insights into these relationships. Zoochory showed contrasting responses to climatic gradients when analysed separately: synzoochorous plants, which typically bear the heaviest seeds (Souza 2023), showed the opposite trend compared to endozoochorous plants. This result highlights the importance of accounting for seed mass along with dispersal modes when predicting or explaining dispersal mechanisms. Furthermore, although precipitation has been shown to strongly influence seed mass across broad latitudinal gradients (Moles et al. 2014), this variable was a weaker driver at finer spatial scales. Such findings suggest that ecological hypotheses derived from global patterns, such as resource availability, may not always hold at community-level scales. In this context, plot-based datasets play a critical role in refining existing ecological theory by improving our understanding of how environmental factors influence seed-dispersal strategies across tropical ecosystems.

Importantly, these plot-based relationships between dispersal strategies and environmental gradients can help inform and anticipate potential shifts in plant communities under future climatic or anthropogenic pressures (e.g., habitat loss and fragmentation). For instance, it is expected that global warming will increase climate instability through the increase of seasonality of precipitation and intensification of the annual hottest day temperature and wettest day precipitation (IPCC (Intergovernmental Panel on Climate Change) 2023). As a result, we may expect (i) reduced recruitment success of larger-seeded species that are typical of deeply shaded rainforests due to increased temperature seasonality; (ii) increases in wind-dispersed and small-seeded species, such as those commonly found in dry forests, that may be better able to cope with increasing climatic instability; and (iii) reductions of fleshy endozoochorous fruits (which strongly depend on water availability) under higher seasonal rainfall and drought predicted with global warming. However, all these predictions could be modified if climate change triggers adaptive phenotypic plasticity in fruit and seed traits (van Kleunen and Fischer 2005; Matesanz et al. 2010).

5 | Conclusions

We found that climate, particularly moisture and thermal variables, is the main driver of the distribution and dominance of dispersal modes and seed masses. Ongoing environmental changes, including climate change and anthropogenic disturbances, may disrupt these patterns, potentially affecting seed dispersal and recruitment processes, with consequences for ecosystem services including vegetation carbon storage (Bello et al. 2015). These disruptions could also affect faunal communities and people reliant on forest fruits and seeds, as illustrated by climate-driven yield losses in iconic species such as the Brazil nut tree *Bertholletia excelsa*, a large-seeded synzoochorous species that is important for local livelihoods (Anjos et al. 2024). Incorporating seed mass and dispersal modes, along with other seed traits related to the regeneration niche (Saatkamp

et al. 2018; Oliveira et al. 2026) into vegetation models could complement existing work on leaf and stem traits (e.g., Dwyer et al. 2014; Hernández-Calderón et al. 2014), helping to improve our understanding of the impacts of environmental change on tropical ecosystems.

Affiliations

[1]Departamento de Ecologia e Conservação, Instituto de Ciências Naturais, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil | [2]Programa de Pós-Graduação Em Ecologia, Universidade Federal do Pará, Belém, Pará, Brazil | [3]Departamento de Ciências Florestais, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil | [4]Institute of Plant Sciences, University of Bern, Bern, Switzerland | [5]Environmental Change Institute, University of Oxford, Oxford, UK | [6]Leverhulme Centre for Nature Recovery, University of Oxford, Oxford, UK | [7]Georgina Mace Centre for the Living Planet, Imperial College London, Ascot, UK | [8]Facultad de Ciencias Forestales y Ambientales, Universidad Juárez del Estado de Durango, Durango, Mexico | [9]Department of Natural Sciences, Manchester Metropolitan University, Manchester, UK | [10]Applied Ecology and Conservation Lab, Universidade Estadual de Santa Cruz, Ilhéus, Brazil | [11]Herbário, Instituto Agronômico de Campinas (IAC), São Paulo, Brazil | [12]Lancaster Environment Centre, Lancaster University, Lancaster, UK | [13]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA | [14]Laboratorio de Ecología de Bosques Tropicales y Primatología, Universidad de los Andes, Bogotá, Colombia | [15]Universidade Do Estado Do Amapá, Amapá, Brazil | [16]School of Geography, University of Birmingham, Birmingham, UK | [17]Departamento de Engenharia Florestal, Universidade Federal de Mato Grosso, Cuiabá, Brazil | [18]School of Forestry, University of Canterbury, Christchurch, New Zealand | [19]Departamento de Biologia, Universidade Federal de Rondônia, Porto Velho, Rondônia, Brazil | [20]Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Brazil | [21]School of Geography, University of Leeds, Leeds, UK | [22]Programa de Pós-Graduação Em Ecologia e Conservação, Universidade Do Estado de Mato Grosso, Nova Xavantina, Brazil | [23]Departamento de Ciências Biológicas, Universidade Do Estado de Mato Grosso, Mato Grosso, Brazil | [24]Departamento de Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas/UNICAMP, Campinas, Brazil | [25]Plataforma Brasileira de Biodiversidade e Serviços Ecossistêmicos (BPBES), Campinas, Brazil | [26]School of Environmental Sciences, University of East Anglia, Norwich, UK | [27]Instituto Juruá, Manaus, Brazil | [28]Núcleo de Apoio à Pesquisa Em Roraima, Instituto Nacional de Pesquisas da Amazônia, Boa Vista, Brazil | [29]Departamento de Agronomia, Universidade Do Estado de Mato Grosso, Nova Xavantina, Brazil | [30]Programa de Pós-Graduação Em Ambiente e Sistemas de Produção Agrícola, Universidade Do Estado de Mato Grosso, Tangará da Serra, Brazil | [31]Instituto de Pesquisas Jardim Botânico Do Rio de Janeiro, Rio de Janeiro, Brazil | [32]Instituto de Biologia, Universidade Federal da Bahia, Salvador, Bahia, Brazil | [33]Instituto de Ciências Naturais, Humanas e Sociais (ICNHS), Universidade Federal de Mato Grosso (UFMT), Sinop, Brazil | [34]Laboratório de Ciências Ambientais, Universidade Estadual Do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, RJ, Brazil | [35]Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA | [36]Laboratório de Ecologia Vegetal, Universidade Do Estado de Mato Grosso, Nova Xavantina, Brazil | [37]Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil | [38]Instituto Nacional da Mata Atlântica (INMA), Santa Teresa, Brazil | [39]Plant Ecology and Nature Conservation Group, Wageningen University and Research, Wageningen, the Netherlands | [40]Office National Des Forêts, Paris, France | [41]Programa de Pós-graduação Em Ciência, Tecnologia e Inovação Florestal (PPGCTIF) do Instituto de Biodiversidade e Floresta, Universidade Federal Do Oeste Do Pará, Santarém, Brazil | [42]Departamento de Botânica, Universidade Federal de Juiz de Fora, Juiz de Fora, Brazil | [43]BeZero, London, UK | [44]Coordenação de Pesquisas Em Ecologia, Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Brazil | [45]Museu Paraense Emílio Goeldi, Coordenação de Botânica, Belém, Brazil | [46]School of Biological Sciences, University of Bristol, Bristol, UK | [47]Coordenação de Dinâmica Ambiental, Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Brazil | [48]Woodwell Climate Research Center, Falmouth, Massachusetts, USA | [49]Centre for Tropical Environmental and Sustainability Science, College of Science and Engineering, James Cook University, Cairns, Australia | [50]Departamento de Biodiversidade, Instituto de Biociências, Universidade Federal de Jataí, Goiás, Brazil | [51]Gothenburg Global Biodiversity Centre, University of Gothenburg, Gothenburg, Sweden | [52]Departamento Genética, Ecologia e Evolução, ICB/Universidade Federal de Minas Gerais, Belo Horizonte, Brazil | [53]Tikré Brazil Soluções Ambientais Eireli, Goiás, Brazil | [54]Universidade de Brasília, Brasília, Distrito Federal, Brazil | [55]Instituto de Ciências Exatas e Naturais, Universidade Federal de Rondonópolis, Rondonópolis, Brazil | [56]Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Brazil | [57]Departamento de Botânica, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil | [58]Universidade Federal de Campina Grande, Unidade Acadêmica de Tecnologia Do Desenvolvimento, Sumé, Paraíba, Brazil | [59]Colegiado de Ecologia, Universidade Federal Do Vale Do São Francisco, Senhor do Bonfim, Bahia, Brazil | [60]Empresa Brasileira de Pesquisa Agropecuária, Embrapa Amazônia Oriental, Belém, Pará, Brazil | [61]Centro Multidisciplinar, Universidade Federal Do Acre, Rio Branco, Brazil | [62]Centro de Recursos Genéticos Vegetais, Instituto Agronômico de Campinas, São Paulo, Brazil | [63]Departamento de Engenharia Florestal, Universidade de Brasília, Brasília, Distrito Federal, Brazil | [64]Departamento de Biologia, Universidade Federal Do Amazonas, Manaus, Amazonas, Brazil | [65]University of Exeter, Exeter, UK | [66]Instituto Nacional de Pesquisas Espaciais (INPE), São Paulo, Brazil | [67]Museu Paraense Emílio Goeldi, Coordenação de Ciências da Terra e Ecologia, Belém, Pará, Brazil | [68]Programa de Pós-Graduação Em Sanidade e Produção Animal Sustentável Na Amazônia Ocidental da Universidade Federal Do Acre, Acre, Brazil | [69]Centre de Coopération International en Recherche Agronomique Pour le Développement (CIRAD), Montpellier, France | [70]Departamento de Geociências, Universidade Federal do Pará, Belém, Pará, Brazil | [71]Center for Tropical Research, Institute of the Environment and Sustainability, University of California, Los Angeles, USA | [72]Departamento de Ciências Biológicas, Universidade Estadual de Feira de Santana, Feira de Santana, Bahia, Brazil | [73]Empresa Brasileira de Pesquisa Agropecuária, Embrapa Semiárido, Fortaleza, Ceará, Brazil | [74]Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Brasília, Brazil | [75]Programa de Pós-graduação Em Recursos Naturais (PRONAT), Universidade Federal de Roraima, Boa Vista, Brazil | [76]Laboratório de Botânica e Ecologia Vegetal, Centro de Ciências Biológicas e da Natureza, Universidade Federal Do Acre, Rio Branco, Acre, Brazil | [77]Programa de Pós-Graduação Em Ciências Ambientais (PPGCA), Universidade Do Estado de Mato Grosso, Mato Grosso, Brazil | [78]Departamento de Biologia Geral, Universidade Estadual de Montes Claros, Montes Claros, Minas Gerais, Brazil | [79]Instituto de Ciências Agrárias, Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brazil | [80]Universidade Estadual de Roraima, Rorainópolis, Roraima, Brazil | [81]Botany, School of Natural Sciences, Trinity College Dublin, Dublin, Ireland | [82]Departamento de Botânica, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil | [83]Laboratório de Ecologia e Manejo de Ecossistemas, Universidade Estadual de Campinas, Campinas, Brazil | [84]Departamento de Dinâmica Ambiental, Instituto Nacional de Pesquisa da Amazônia, Boa Vista, Roraima, Brazil | [85]Centro de Estudos da Biodiversidade (CBIO), Universidade Federal de Roraima, Boa Vista, Roraima, Brazil | [86]Instituto Federal de Educação, Ciência e Tecnologia Do Acre (IFAC), Rio Branco, Acre, Brazil | [87]Universidade Do Estado de Mato Grosso, Nova Xavantina, Brazil | [88]Environmental Modeling Program, Texas A&M Transportation Institute, Bryan, Texas, USA | [89]Centro de Ciências Biológicas e da Natureza, Universidade Federal Do Acre, Rio

Branco, Acre, Brazil | [90]Departamento de Silvicultura, Universidade Federal Rural Do Rio de Janeiro, Instituto de Florestas, Seropédica, Brazil | [91]Faculdade de Engenharia Florestal, Universidade Federal de Mato Grosso, Cuiabá, Brazil | [92]Núcleo de Estudos e Pesquisas Ambientais (NEPAM), Universidade Estadual de Campinas (UNICAMP), São Paulo, Brazil | [93]Geography, Faculty of Environment, Science and Economy, University of Exeter, Exeter, UK | [94]Instituto Bem Ambiental (IBAM), Belo Horizonte, Minas Gerais, Brazil | [95]Programa de Pós-Graduação Em Ecologia, Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, Brazil | [96]Royal Botanic Gardens Edinburgh, Edinburgh, UK | [97]Stockholm Environment Institute, Stockholm, Sweden | [98]School of Geography, University of Exeter, Exeter, UK | [99]Department of Biology, Washington University in St. Louis, St. Louis, Missouri, USA | [100]Missouri Botanical Garden, St. Louis, Missouri, USA | [101]Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil | [102]Faculty of Communication and Environment, Hochschule Rhein-Waal, Kamp-Lintfort, Germany | [103]Institute of Biogeosciences, IBG2: Plant Sciences, Forschungszentrum Jülich GmbH, Jülich, Germany | [104]Programa de Pós-Graduação Em Biologia Vegetal, Departamento de Biodiversidade, Instituto de Biociências, Universidade Estadual Paulista (UNESP), São Paulo, Brazil | [105]AMAP (botAnique et Modélisation de L'architecture Des Plantes et Des Végétations), CIRAD, CNRS, INRAE, IRD, Montpellier, France | [106]Departamento de Ecologia, Instituto de Biologia, Universidade Estadual de Campinas, São Paulo, Brazil

Acknowledgements

We are deeply thankful for everyone involved in field collection, identification and data curation from vegetation inventories. We also thank the Brazilian and international herbaria for curating and making their collections and specimen images available online, which enabled trait measurements for this study. We are thank the NERC/FAPESP funded project ARBOLES 'A trait-based Understanding of LATAM Forest Biodiversity and Resilience' (NE/S011811/1) for supporting data collection, academic exchanges and JB's time. We acknowledge the contributions of ForestPlots.net, a meta-network and cyber-initiative to promote collaborative forest science, as well as the ForestPlots.net Collaboration and Data Request Committee (B.S.M., Euridice Honoria., O.P., Tim Baker., A.L.) for facilitating this project and associated data management; this work was enabled by ForestPlots.net approved Research Project number 98 "*Mapping reproductive traits across South American biomes*". F.C.S. acknowledges the Coordination for the Improvement of Higher Education Personnel—Brazil (CAPES) – Finance Code 001, for providing a PhD scholarship and the Brazilian National Research Council (CNPq) for the postdoctoral fellowship (PDJ Process no. 152806/2024-5). F.C.S. also acknowledges the Centro Avançado em Pesquisas Socioecológicas para a Recuperação Ambiental da Amazônia (CAPOEIRA, CNPq Grant no. 443849/2024-2) for providing a collaborative research environment and opportunities for scientific exchange and professional development. F.C.S. is grateful to Fabricio Beggiato Baccaro for reading the manuscript and providing valuable feedback, and to Marc van Roosmalen for his guidance on the literature on Amazonian fruits. G.W.F., M.M.E.-S. and M.D.M.V. thanks CNPq and FAPEMIG (APQ 03020-22) for grant support. M.N.U. was funded by a Seed Grant from Universidad de Los Andes, Colombia. B.S.M., B.H.M.J.R., F.A.G.G. (Process no. 302200/2025-8) and J.R.R.P. (Process no. 312571/2021-6) thank CNPq for productivity grant support. F.A.G.G. thanks Research Support Foundation of Goiás State (FAPEG—Process no 2025 10267001634) and CNPq (Process no. 445997/2024-9) for the financial support through the PELD CEMA. J.A.G. was funded by the Natural Environment Research Council (NERC) (NE/T011084/1 and NE/Z504191/1), the Royal Society (RG\R1\251370) the Leverhulme Trust (RPG-2024-342) and European Union (ERC, ADAPTA, 101231095). F.E. is supported by the Serrapilheira Institute (Grant no. R-2401–357 46863) and FAPESPA (TO no. 158/2024). F.M.F. acknowledges funding from CNPq (406767/2022-0; 445970/2024-3, 443860/2024-6, 446514/2024-1, 316833/2025-8, 441257/2023-2 and 444350/2024-1) and the UKRI FLF (MR/X032949/1). L.S.M. acknowledges funding from DEFRA's Global Centre on Biodiversity for Climate (RG2-009854) and the Ramboll Foundation. T.F.D. acknowledges the CNPq for productivity grant support (Grant no. 312589/2022-0) and FAPESP (Grant no. 2015/50488-5). V.A.K. acknowledges the FAPESP (Grant no. 2022/09041-0). DC's research in plant biodiversity is supported by grants from CNPq (Research Productivity Fellowship no. 420895/2025-6 and PPBIO Semiárido no. 441271/2023-5) and FAPERJ (Programa Jovem Cientista do Nosso Estado—2022, Grant no. 200.153/2023). W.E.M. and the surveyed plots were supported by PELD-PSAM (CNPq/MCTI/CONFAP-FAPs/PELD call no. 21/2020, processes no. CNPq 441366/2020-1, 445988/2024-0, FAPEAM), PPBio (CNPq/MCTI process no. 441260/2023-3 and 441228/2023-2); and the INCT for Amazonian Biodiversity (CNPq/MCTI processes no. 406474/2022-2 and, FAPEAM 01.02.016301.02394/2024-50 and 01.02.016301.02637/2022-98). T.R.F. was supported by the NERC (Grant no. NE/N011570/1). Plot data collection from the Brazilian Atlantic Forest was supported by the State of São Paulo Research Foundation/FAPESP (BIOTA Grants no. 2003/12595-7, 2010/20811-7, 2012/51509-8 and 2012/51872-5); CNPq (PELD Grant no. 403710/2012-0 and Universal Grant no. 459941/2014-3); British Natural Environment Research Council/NERC (Grant no. K016431/1) and Project Sustainable Landscapes (US Forest Service, USAID, EMBRAPA & NASA-Goddard), under COTEC/IF 41.065/2005 and IBAMA/CGEN 093/2005 permits. 'Nordeste Project: New Science for a Neglected Biome' was funded by the São Paulo Research Foundation (FAPESP, Grant no. 2015/50488-5) and the NERC (Newton Fund/NERC, Grants no. NE/N012488/1, NE/N012550/1, NE/N012526/1). The RAINFOR forest monitoring network in Brazil has been supported principally by the NERC (Grants no. NE/B503384/1, NE/D01025X/1, NE/I02982X/1, NE/F005806/1, NE/D005590/1 and NE/I028122/1), the European Research Council (ERC) (Advanced Grant no. 291585 – 'TFORCES'), the Gordon and Betty Moore Foundation (#1656 'RAINFOR'), and the European Union's Fifth, Sixth and Seventh Framework Programme (EVK2-CT-1999-00023 – 'CARBONSINK-LBA', 283080 – 'GEOCARBON', 282664 – 'AMAZALERT'), as well as by NERC/State of Sao Paulo Research Foundation (FAPESP) consortium grants 'BIO-RED' (NE/N012542/1, 2012/51872-5), 'ECOFOR' (NE/K016431/1, 2012/51509-8), 'ARBOLES' (NE/S011811/1), the Royal Society (University Research Fellowships and Global Challenges Awards) ('FORAMA', ICA/R1/180100) and CNPq, including several PVE and Productivity Grants. Data from RAINFOR and other networks are curated at ForestPlots.net, uniting permanent plot records and supporting scientists from the world's tropical forests. The monitoring of plots in the Amazon-Cerrado transition was financed by CNPq/PELD-TRAN (441572/2020-0) and FAPEMAT/PELD (0346321/2021). The development of ForestPlots.net and its data curation has been funded by numerous grants including NE/B503384/1, NE/N012542/1, ERC AdG 291585, NE/F005806/1, NE/N004655/1 and the Gordon and Betty Moore Foundation. This is paper number 137 of the Rede Amazonia Sustentavel, who recognize support from CNPq PELD-RAS 3: N°445994/2024-0; UKRI/National Environment Research Council NE/K016431/1, NE/P004512/1, NE/X019039/1, NE/S01084X/1 and NE/N01250X/1; and BNP Paribas Foundation's Climate and Biodiversity Initiative and DEFRA's Global Centre of Biodiversity for Climate. The Article Processing Charge for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614).

Funding

This research was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico: 152806/2024-5, 302200/2025-8, 312571/2021-6, 445997/2024-9, 406767/2022-0, 445970/2024-3, 443860/2024-6, 446514/2024-1, 316833/2025-8, 441257/2023-2, 444350/2024-1, 312589/2022-0, 420895/2025-6, 441271/2023-5, 441366/2020-1, 445988/2024-0, 441260/2023-3, 441228/2023-2, 406474/2022-2, 403710/2012-0, 459941/2014-3, 441572/2020-0, 445994/2024-0; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior: Finance Code 001; European Research Council: 291585, 101231095; Fundação Amazônia

Paraense de Amparo à Pesquisa: 158/2024; Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro: 200.153/2023; Fundação de Amparo à Pesquisa do Estado de Goiás: 2025/10267001634; Fundação de Amparo à Pesquisa do Estado de Mato Grosso: 0346321/2021; Fundação de Amparo à Pesquisa do Estado de Minas Gerais: APQ 03020-22; Fundação de Amparo à Pesquisa do Estado do Amazonas: 01.02.016301.02394/2024-50, 01.02.016301.02637/2022-98; Fundação de Amparo à Pesquisa do Estado de São Paulo: 2003/12595-7, 2010/20811-7, 2012/51509-8, 2012/51872-5, 2015/50488-5, 2022/09041-0; Natural Environment Research Council: NE/S011811/1, NE/T011084/1, NE/Z504191/1, NE/N011570/1, NE/K016431/1, NE/N012488/1, NE/N012550/1, NE/N012526/1, NE/B503384/1, NE/D01025X/1, NE/I02982X/1, NE/F005806/1, NE/D005590/1, NE/I028122/1, NE/N012542/1, NE/N004655/1, NE/P004512/1, NE/X019039/1, NE/S01084X/1, NE/N01250X/1; UK Research and Innovation: MR/X032949/1; Royal Society: RG/R1/251370, ICA/R1/180100; Leverhulme Trust: RPG-2024-342; Department for Environment, Food & Rural Affairs: RG2-009854; Instituto Serrapilheira: R-2401-35746863; and Gordon and Betty Moore Foundation: #1656.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data and R code used for the analyses are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.r2280gbpg.

References

Ab'Saber, A. N. 2003. *Os domínios de natureza no Brasil: Potencialidades paisagísticas*. Ateliê Editorial.

Aguirre-Gutiérrez, J., S. W. Rifai, X. Deng, et al. 2025. "Canopy Functional Trait Variation Across Earth's Tropical Forests." *Nature* 641: 129–136.

Almeida-Neto, M., F. Campassi, M. Galetti, P. Jordano, and A. Oliveira-Filho. 2008. "Vertebrate Dispersal Syndromes Along the Atlantic Forest: Broad-Scale Patterns and Macroecological Correlates." *Global Ecology and Biogeography* 17: 503–513.

Anjos, L. J., G. S. Gonçalves, V. A. Dutra, et al. 2024. "Brazil Nut Journey Under Future Climate Change in Amazon." *PLoS One* 19, no. 11: e0312308.

Bello, C., M. Galetti, M. A. Pizo, et al. 2015. "Defaunation Affects Carbon Storage in Tropical Forests." *Science Advances* 1: e1501105.

Benítez-López, A., L. Santini, C. A. Peres, et al. 2026. "Defaunation Maps for Tropical Mammals and Birds." *Global Ecology and Biogeography* 35, no. 7: e70280.

BFG (Brazil Flora Group), J. Gomes-da-Silva, F. L. R. Filardi, et al. 2022. "Brazilian Flora 2020: Leveraging the Power of a Collaborative Scientific Network." *Taxon* 71: 178–198.

Bivand, R., T. Keitt, and B. Rowlingson. 2015. "rgdal: Bindings for the 'Geospatial' Data Abstraction Library." https://CRAN.R-project.org/package=rgdal.

Bjornstad, O. N. 2022. "ncf: Spatial Covariance Functions." R Package Version 1.3-2. https://CRAN.R-project.org/package=ncf.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32.

Brooks, M. E., K. Kristensen, K. J. van Benthem, et al. 2017. "*glmmTMB* Balances Speed and Flexibility Among Packages for Zero-Inflated Generalized Linear Mixed Modeling." *R Journal* 9, no. 2: 378–400.

Butler, D. W., R. Green, D. Lamb, W. McDonald, and P. Forster. 2007. "Biogeography of Seed-Dispersal Syndromes, Life-Forms and Seed Sizes Among Woody Rain-Forest Plants in Australia's Subtropics." *Journal of Biogeography* 34: 1736–1750.

Carvalheiro, L. G., I. Bartomeus, O. Rollin, S. Timóteo, and C. F. Tinoco. 2021. "The Role of Soils on Pollination and Seed Dispersal." *Philosophical Transactions of the Royal Society, B: Biological Sciences* 376, no. 1834: 20200171.

Carvalho, R. L., A. F. Resende, J. Barlow, et al. 2023. "Pervasive Gaps in Amazonian Ecological Research." *Current Biology* 33, no. 16: 3495–3504.

Chen, S. C., W. K. Cornwell, H. X. Zhang, and A. T. Moles. 2017. "Plants Show More Flesh in the Tropics: Variation in Fruit Type Along Latitudinal and Climatic Gradients." *Ecography* 40: 531–538.

Comita, L. S., S. A. Queenborough, S. J. Murphy, et al. 2014. "Testing Predictions of the Janzen–Connell Hypothesis: A Meta-Analysis of Experimental Evidence for Distance-and Density-Dependent Seed and Seedling Survival." *Journal of Ecology* 102: 845–856.

Connell, J. H. 1971. "On the Role of Natural Enemies in Preventing Competitive Exclusion in Some Marine Animals and in Rain Forest Trees." *Dynamics of Populations* 298, no. 312: 1–610.

Correa, D. F., E. Álvarez, and P. R. Stevenson. 2015. "Plant Dispersal Systems in Neotropical Forests: Availability of Dispersal Agents or Availability of Resources for Constructing Zoochorous Fruits?" *Global Ecology and Biogeography* 24: 203–214.

Correa, D. F., P. R. Stevenson, M. N. Umaña, et al. 2022. "Geographic Patterns of Tree Dispersal Modes in Amazonia and Their Ecological Correlates." *Global Ecology and Biogeography* 32: 49–69.

da Silva, J. M. C., A. B. Rylands, and G. A. da Fonseca. 2005. "The Fate of the Amazonian Areas of Endemism." *Conservation Biology* 19, no. 3: 689–694.

Davis, N. N., J. Badger, A. N. Hahmann, et al. 2023. "The Global Wind Atlas: A High-Resolution Dataset of Climatologies and Associated Web-Based Application." *Bulletin of the American Meteorological Society* 104, no. 8: 1507–1525.

Dexter, K. G., R. T. Pennington, A. T. Oliveira-Filho, M. L. Bueno, P. L. Silva de Miranda, and D. M. Neves. 2018. "Inserting Tropical Dry Forests Into the Discussion on Biome Transitions in the Tropics." *Frontiers in Ecology and Evolution* 6: 104.

Dirzo, R., H. S. Young, M. Galetti, et al. 2014. "Defaunation in the Anthropocene." *Science* 345, no. 6195: 401–406.

Dwyer, J. M., R. J. Hobbs, and M. M. Mayfield. 2014. "Specific Leaf Area Responses to Environmental Gradients through Space and Time." *Ecology* 95, no. 2: 341–399.

Enquist, B. J., B. Boyle, B. S. Maitner, et al. 2026. "BIEN: A Biodiversity Informatics Ecosystem Advancing Open and Reproducible Workflows for Plant Observation, Plot and Trait Data." *Methods in Ecology and Evolution* 17, no. 5: 1556–1584.

Eriksson, O., E. M. Friis, and P. Löfgren. 2000. "Seed Size, Fruit Size, and Dispersal Systems in Angiosperms From the Early Cretaceous to the Late Tertiary." *American Naturalist* 156: 47–58.

Feeley, K. J., C. Bravo-Avila, B. Fadrique, T. M. Perez, and D. Zuleta. 2020. "Climate-Driven Changes in the Composition of New World Plant Communities." *Nature Climate Change* 10, no. 10: 965–970.

Fick, S. E., and R. J. Hijmans. 2017. "Worldclim 2: New 1-Km Spatial Resolution Climate Surface for Global Land Areas." *International Journal of Climatology* 37: 4302–4315.

ForestPlots.net, C. Blundo, J. Carilla, et al. 2021. "Taking the Pulse of Earth's Tropical Forests Using Networks of Highly Distributed Plots." *Biological Conservation* 260: 108849.

Fricke, E. C., C. Bello, R. Chaplin-Kramer, et al. 2025. "Drivers and Impacts of Global Seed Disperser Decline." *Nature Reviews Biodiversity* 1, no. 6: 386–400.

Friedman, J. H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *Annals of Statistics* 29, no. 5: 1189–1232.

Fuzessy, L., F. A. O. Silveira, L. Culot, P. Jordano, and M. Verdu. 2022. "Phylogenetic Congruence Between Neotropical Primates and Plants Is Driven by Frugivory." *Ecology Letters* 25: 320–325.

Fuzessy, L. F., T. G. Cornelissen, C. Janson, and F. A. Silveira. 2016. "How Do Primates Affect Seed Germination? A Meta-Analysis of Gut Passage Effects on Neotropical Plants." *Oikos* 125, no. 8: 1069–1080.

Fuzessy, L. F., C. Janson, and F. A. Silveira. 2018. "Effects of Seed Size and Frugivory Degree on Dispersal by Neotropical Frugivores." *Acta Oecologica* 93: 41–47.

Galetti, M., R. Guevara, M. C. Cortes, et al. 2013. "Functional Extinction of Birds Drives Rapid Evolutionary Changes in Seed Size." *Science* 340, no. 6136: 1086–1090.

Gardner, C. J., J. E. Bicknell, W. Baldwin-Cantello, M. J. Struebig, and Z. G. Davies. 2019. "Quantifying the Impacts of Defaunation on Natural Forest Regeneration in a Global Meta-Analysis." *Nature Communications* 10: 4590.

Griffiths, M. E., and M. J. Lawes. 2006. "Biogeographic, Environmental, and Phylogenetic Influences on Reproductive Traits in Subtropical Forest Trees, South Africa." *Ecography* 29: 614–622.

Guimarães, P. R., Jr., M. Galetti, and P. Jordano. 2008. "Seed Dispersal Anachronisms: Rethinking the Fruits Extinct Megafauna Ate." *PLoS One* 3, no. 3: e1745.

Haugaasen, T., and C. A. Peres. 2005. "Mammal Assemblage Structure in Amazonian Flooded and Unflooded Forests." *Journal of Tropical Ecology* 21: 133–145.

Hawes, J., I. Vieira, L. Magnago, et al. 2020. "A Large-Scale Assessment of Plant Dispersal Mode and Seed Traits Across Human-Modified Amazonian Forests." *Journal of Ecology* 108: 1373–1385.

Hernández-Calderón, E., R. Méndez-Alonzo, J. Martínez-Cruz, et al. 2014. "Altitudinal Changes in Tree Leaf and Stem Functional Diversity in a Semi-Tropical Mountain." *Journal of Vegetation Science* 25, no. 4: 955–966.

Hijmans, R. 2025a. "terra: Spatial Data Analysis. R Package Version 1.8 29." https://github.com/rspatial/terra.

Hijmans, R. J. 2025b. "raster: Geographic Data Analysis and Modeling. R Package Version 3.6–32." http://cran.stat.unipd.it/web/packages/raster.

Huang, E., Y. Chen, M. Fang, Y. Zheng, and S. Yu. 2021. "Environmental Drivers of Plant Distributions at Global and Regional Scales." *Global Ecology and Biogeography* 30, no. 3: 697–709.

IBGE (Instituto Brasileiro de Geografia e Estatística). 2012. *Manual técnico da vegetação brasileira*, Série Manuais técnicos em Geociências. 2nd ed. IBGE.

IPCC (Intergovernmental Panel on Climate Change). 2023. *Climate Change 2023: Synthesis Report. A Report of the Intergovernmental Panel on Climate Change. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team*, edited by H. Lee and J. Romero. IPCC.

IUCN (International Union for Conservation of Nature). 2023. "The IUCN Red List of Threatened Species. Version 2023–1." https://www.iucnredlist.org.

Janzen, D. H. 1970. "Herbivores and the Number of Tree Species in Tropical Forests." *American Naturalist* 104, no. 940: 501–528.

Janzen, D. H., and P. S. Martin. 1982. "Neotropical Anachronisms: The Fruits the Gomphotheres Ate." *Science* 215, no. 4528: 19–27.

Joswig, J. S., C. Wirth, M. C. Schuman, et al. 2022. "Climatic and Soil Factors Explain the Two-Dimensional Spectrum of Global Plant Trait Variation." *Nature Ecology & Evolution* 6: 36–50.

Kattge, J., G. Bönisch, S. Díaz, et al. 2020. "TRY Plant Trait Database - Enhanced Coverage and Open Access." *Global Change Ecology* 26: 119–188.

Kuhlmann, M., and J. F. Ribeiro. 2016. "Evolution of Seed Dispersal in the Cerrado Biome: Ecological and Phylogenetic Considerations." *Acta Botânica Brasílica* 30: 271–282.

Kuhn, M. 2008. "Building Predictive Models in R Using the Caret Package." *Journal of Statistical Software* 28, no. 5: 1–26.

Laliberté, E., and P. Legendre. 2010. "A Distance-Based Framework for Measuring Functional Diversity From Multiple Traits." *Ecology* 91: 2099–3005.

Lenth, R. 2023. "emmeans: Estimated Marginal Means, Aka Least-Squares Means." https://cran.r-project.org/web/packages/emmeans/emmeans.pdf.

Lopez-Gonzalez, G., S. L. Lewis, M. Burkitt, T. R. Baker, and O. L. Phillips. 2011. "ForestPlots.Net: A Web Application and Research Tool to Manage and Analyse Tropical Forest Plot Data." *Journal of Vegetation Science* 22: 610–613.

Malhado, A., J. A. Oliveira-Neto, J. Stropp, et al. 2015. "Climatological Correlates of Seed Size in Amazonian Forest Trees." *Journal of Vegetation Science* 26: 956–963.

MapBiomas Project. 2025. "Collection 10 of the Annual Land Use and Land Cover Maps of Brazil." https://plataforma.brasil.mapbiomas.org/.

Matesanz, S., E. Gianoli, and F. Valladares. 2010. "Global Change and the Evolution of Phenotypic Plasticity in Plants." *Annals of the New York Academy of Sciences* 1206, no. 1: 35–55.

Meyer, H., C. Milà, M. Ludwig, J. Linnenbrink, and F. Schumacher. 2025. "CAST: 'Caret' Applications for Spatial-Temporal Models." R Package Version 1.0.3. https://github.com/HannaMeyer/CAST.

Meyer, H., and E. Pebesma. 2021. "Predicting Into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models." *Methods in Ecology and Evolution* 12, no. 9: 1620–1633.

Mittelman, P., C. M. Dracxler, P. R. O. Santos-Coutinho, and A. S. Pires. 2021. "Sowing Forests: A Synthesis of Seed Dispersal and Predation by Agoutis and Their Influence on Plant Communities." *Biological Reviews* 96: 2425–2445.

MMA (Ministério do Meio Ambiente). 2024. "Biomas." https://antigo.mma.gov.br/biomas.html.

Mod, H. K., D. Scherrer, M. Luoto, and A. Guisan. 2016. "What We Use Is Not What We Know: Environmental Predictors in Plant Distribution Models." *Journal of Vegetation Science* 27, no. 6: 1308–1322.

Moles, A. T., D. D. Ackerly, J. C. Tweddle, et al. 2007. "Global Patterns in Seed Size." *Global Ecology and Biogeography* 16: 109–116.

Moles, A. T., D. D. Ackerly, C. O. Webb, et al. 2005. "Factors That Shape Seed Mass Evolution." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 30: 10540–10544.

Moles, A. T., S. E. Perkins, S. W. Laffan, et al. 2014. "Which Is a Better Predictor of Plant Traits: Temperature or Precipitation?" *Journal of Vegetation Science* 25: 1167–1180.

Moles, A. T. 2018. "Being John Harper: Using Evolutionary Ideas to Improve Understanding of Global Patterns in Plant Traits." *Journal of Ecology* 106, no. 1: 1–18.

Moonlight, P. W., K. Banda-r, O. L. Phillips, et al. 2021. "Expanding Tropical Forest Monitoring Into Dry Forests: The DRYFLOR Protocol for Permanent Plots." *Plants, People, Planet* 3, no. 3: 295–300.

Muscarella, R., M. Lohbeck, M. Martínez-Ramos, et al. 2017. "Demographic Drivers of Functional Composition Dynamics." *Ecology* 98: 2743–2750.

Oliveira, D. N., C. A. Ordóñez-Parra, S. C. Chen, et al. 2026. "Tropical Seed Trait Database: Advancing Seed Functional Ecology in the World's Most Biodiverse Region." *New Phytologist* 251, no. 3: 945–958.

Peres, C. A., T. Emilio, J. Schietti, S. J. M. Desmouliere, and T. Levi. 2016. "Dispersal Limitation Induces Long-Term Biomass Collapse in Overhunted Amazonian Forests." *Proceedings of the National Academy of Sciences of the United States of America* 113, no. 4: 892–897.

Pinho, B. X., F. P. Melo, C. J. Ter Braak, et al. 2025. "Winner–Loser Plant Trait Replacements in Human-Modified Tropical Forests." *Nature Ecology & Evolution* 9, no. 2: 282–295.

Pinho, B. X., M. Tabarelli, C. J. F. ter Braak, et al. 2021. "Functional Biogeography of Neotropical Moist Forests: Trait–Climate Relationships and Assembly Patterns of Tree Communities." *Global Ecology and Biogeography* 30: 1430–1446.

Poggio, L., L. M. de Sousa, N. H. Batjes, et al. 2021. "SoilGrids 2.0: Producing Soil Information for the Globe With Quantified Spatial Uncertainty." *Soil* 7: 217–240.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Roberts, D. R., V. Bahn, S. Ciuti, et al. 2017. "Cross-Validation Strategies for Data With Temporal, Spatial, Hierarchical, or Phylogenetic Structure." *Ecography* 40, no. 8: 913–929.

Royal Botanic Gardens, Kew. 2020. "Seed Information Database (SID) Version 7.1." http://data.kew.org/sid.

Rubio de Casas, R., C. G. Willis, W. D. Pearse, C. C. Baskin, J. M. Baskin, and J. Cavender-Bares. 2017. "Global Biogeography of Seed Dormancy Is Determined by Seasonality and Seed Size: A Case Study in the Legumes." *New Phytologist* 214: 1527–1536.

Saatkamp, A., A. Cochrane, L. Commander, et al. 2018. "A Research Agenda for Seed-Trait Functional Ecology." *New Phytologist* 221: 1764–1775.

Sanderson, E. W., M. Jaiteh, M. A. Levy, et al. 2002. "The Human Footprint and the Last of the Wild: The Human Footprint is a Global Map of Human Influence on the Land Surface, Which Suggests That Human Beings are Stewards of Nature, Whether We Like It or Not." *Bioscience* 52, no. 10: 891–904.

Souza, C. M., Jr., J. Z. Shimbo, M. R. Rosa, et al. 2020. "Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine." *Remote Sensing* 12, no. 17: 2735.

Souza, F. C. 2023. *Características de dispersão de sementes e frutos ao longo dos biomas brasileiros: explorando tendências, predizendo e mapeando correlatos ecológicos. Tese de doutorado*. Universidade Federal de Lavras.

Stevenson, P. R. 2001. "The Relationship Between Fruit Production and Primate Abundance in Neotropical Communities." *Biological Journal of the Linnean Society* 72, no. 1: 161–178.

Stevenson, P. R. 2015. "Neotropical Primate Communities: Effects of Disturbance, Resource Production and Forest Type Heterogeneity." *American Journal of Primatology* 78: 391–401.

Sullivan, M. J., O. L. Phillips, D. Galbraith, et al. 2025. "Variation in Wood Density Across South American Tropical Forests." *Nature Communications* 16, no. 1: 2351.

Tabarelli, M., A. Vicente, and D. Barbosa. 2003. "Variation of Seed Dispersal Spectrum of Woody Plants Across a Rainfall Gradient in North-Eastern Brazil." *Journal of Arid Environments* 53: 197–210.

ter Steege, H., N. C. Pitman, O. L. Phillips, et al. 2006. "Continental-Scale Patterns of Canopy Tree Composition and Function Across Amazonia." *Nature* 443: 444–447.

Terborgh, J., G. Nuñez-Iturri, N. C. Pitman, et al. 2008. "Tree Recruitment in an Empty Forest." *Ecology* 89, no. 6: 1757–1768.

Thomson, F. J., A. T. Moles, T. D. Auld, and R. T. Kingsford. 2011. "Seed Dispersal Distance Is More Strongly Correlated With Plant Height Than With Seed Mass." *Journal of Ecology* 99: 1299–1307.

Trabucco, A., and R. J. Zomer. 2018. Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2.

Umaña, M. N., D. F. Correa, Á. Cano, et al. 2018. "Diversity of Dispersal Systems in Igapó Forests: An Analysis of Local Tree Diversity, Species Turnover, and Dispersal Systems." In *Igapó (Black-Water Flooded Forests) of the Amazon Basin*, edited by R. W. Myster, 23–35. Springer International Publishing.

van Buuren, S., and K. Groothuis-Oudshoorn. 2011. "Mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45: 1–67.

van der Pijl, L. 1982. *Principles of Dispersal*. Springer-Verlag.

van Kleunen, M., and M. Fischer. 2005. "Constraints on the Evolution of Adaptive Phenotypic Plasticity in Plants." *New Phytologist* 166, no. 1: 49–60.

van Roosmalen, M. 2013. *Wild Fruits of Amazon*. Utrecht Herbarium.

van Zonneveld, M., N. Larranaga, B. Blonder, L. Coradin, J. I. Hormaza, and D. Hunter. 2018. "Human Diets Drive Range Expansion of Megafauna-Dispersed Fruit Species." *Proceedings of the National Academy of Sciences of the United States of America* 115, no. 13: 3326–3331.

Violle, C., P. B. Reich, S. W. Pacala, B. J. Enquist, and J. Kattge. 2014. "The Emergence and Promise of Functional Biogeography." *Proceedings of the National Academy of Sciences* 111, no. 38: 13690–13696.

Walters, M. B., and P. B. Reich. 2000. "Seed Size, Nitrogen Supply, and Growth Rate Affect Tree Seedling Survival in Deep Shade." *Ecology* 81: 1887–1901.

Westoby, M., D. S. Falster, A. T. Moles, P. A. Vesk, and I. J. Wright. 2002. "Plant Ecological Strategies: Some Leading Dimensions of Variation Between Species." *Annual Review of Ecology and Systematics* 33: 125–159.

Westoby, M., M. Leishman, and J. Lord. 1996. "Comparative Ecology of Seed Size and Dispersal." *Philosophical Transactions of the Royal Society of London B* 351: 1309–1318.

Wilman, W., J. Belmaker, J. Simpson, C. de la Rosa, M. M. Rivadeneira, and W. Jetz. 2014. "EltonTraits 1.0: Species-Level Foraging Attributes of the World's Birds and Mammals." *Ecology* 95: 2027.

Wright, M. N., and A. Ziegler. 2017. "*Ranger*: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R." *Journal of Statistical Software* 77, no. 1: 1–17.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Summary of sample sizes for each Brazilian regions studied. **Table S2:** Definitions of seed dispersal modes categories. **Table S3:** Proportion of dispersal modes (*sensu stricto* and *sensu lato*) based on 2203 species and 172,185 individuals within 300 old-growth plots. **Table S4:** Statistical results from Generalized Linear Models (GLMs) testing differences in community weighted-means (CWM) of dispersal modes and seed mass across Brazilian ecosystems. Test values (χ^2), degrees of freedom (df) and p-values are shown for each model. All models indicate significant differences among ecosystems ($p < 0.001$). **Table S5:** Descriptions of the climate and soil predictor variables used in this study. **Table S6:** List of frugivorous primate species including their body mass (g) and degree of frugivory (%) according to Elton traits database. **Table S7:** Proportion of dispersal modes (*sensu stricto* and *sensu lato*) (%) by Brazilian ecosystems, based on 2203 species and 172,185 individuals from 300 old-growth plots. Dispersal modes analysed in this

study are highlighted in bold. **Table S8:** Moran's I tests for spatial autocorrelation of residuals from dispersal mode and seed mass models. Results are reported for normality, randomization and Monte Carlo methods. Columns show Moran's I, expectation, variance, standardized deviate (Z) and p-value. **Table S9:** Model performance for dispersal mode models, including both *sensu stricto* and *sensu lato* classifications, assessed using stratified k-fold cross-validation under balanced and unbalanced sampling schemes (see Methods). Performance metrics include: (1) root mean square error (RMSE), indicating the average prediction error (proportion for dispersal modes and milligrammes for seed mass), (2) the correlation coefficient between observed and predicted values and (3) the coefficient of determination [1 − (residual sum of squares/total sum of squares)]. **Table S10:** Model performance assessed using spatial block cross-validation under different block sizes (300, 500 and 800 km). Performance metrics include: (1) root mean square error (RMSE), indicating the average prediction error (proportion for dispersal modes and milligrammes for seed mass), (2) the correlation coefficient between observed and predicted values, and (3) the coefficient of determination [1 − (residual sum of squares/total sum of squares)]. **Table S11:** Cross-validated model performance across Brazilian ecosystems for the main models of seed mass and dispersal modes (endozoochory, anemochory and synzoochory). Model performance was assessed using stratified 5-fold cross-validation and predictive metrics were calculated separately for each biome. Performance metrics include: (1) root mean square error (RMSE), indicating the average prediction error (proportion for dispersal modes and milligrammes for seed mass), (2) the correlation coefficient (Cor) between observed and predicted values and (3) the coefficient of determination (1 − [residual sum of squares/total sum of squares]). **Figure S1:** Model residuals as a function of plot area for endozoochory (a), anemochory (b), synzoochory (c), seed mass with complete and imputed data (d) and seed mass using only complete data (e). Blue points represent individual plots, and the red line corresponds to zero residuals, separating positive residuals (above the line) from negative residuals (below the line). **Figure S2:** Plots show the density of seed-size distribution at species level for the original data (blue line) and five imputed datasets (red) on log scale. Trait imputation was performed through chained equations by predictive mean matching, using R package *mice*. **Figure S3:** Correlation structure among selected climatic, soil and dispersal variables. The correlograms illustrate the Pearson correlations between predictors used in the models by the colour and size of disk. Red colour indicates negative correlations, and blue colour indicates negative correlations. Disk size indicates the strength of the correlations from weak (small) to strong (large). BDOD, Bulk density; CEC, Cation exchange capacity; CLAY, Proportion of clay particles; et0_yr, mean annual evapotranspiration; ISO, Isothermality; MAP, Annual Precipitation; MeanDiuRan, Mean Temperature Diurnal Range; OCS, Organic carbon stocks; PH, pH of water; PrecDriQuar, Precipitation of Driest Quarter; PrecWarQuar, Precipitation of Warmest Quarter; PrecWetQuar, Precipitation of Wettest Quarter; pri_bIQR, IQR of primate body mass; SP, Precipitation Seasonality; ST, Temperature Seasonality; TempAnnRan, Temperature Annual Range; WIND SPEED; Wind speed. **Figure S4:** Spatial correlograms across distance classes, based on raw data (left) and model residuals (right), for endozoochory (a, b), anemochory (c, d), synzoochory (e, f), seed mass with complete and imputed data (g, h) and seed mass using only complete data (i, j). Filled symbols indicate statistically significant correlations, whereas open symbols indicate non-significant correlations. The dashed horizontal line indicates zero correlation. **Figure S5:** Partial dependence of community weighted-mean (CWM) of endozoochory on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other variables constant at their sample mean. **Figure S6:** Partial dependence of community weighted-mean (CWM) of synzoochory on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other predictors constant at their sample mean. **Figure S7:** Partial dependence of community weighted-mean (CWM) of anemochory on each predictor variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other predictors constant at their sample mean. **Figure S8:** Partial dependence of community weighted-mean (CWM) of log (seed mass + 1) on each predictor variable, using both complete and imputed seed mass data. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other predictors constant at their sample mean. **Figure S9:** Area of Applicability (AOA) of the spatial prediction model for endozoochory (a), synzoochory (b), anemochory (c) and seed mass (d) across Brazilian ecosystems. Areas shown in grey fall outside the AOA (AOA = 0) and were excluded from predictions. Colours represent the spatial predictions of the models based on community-weighted mean values, with dark purple and blue indicating lower values and light green and yellow indicating higher values. **Figure S10:** Proportional contribution of each predictor to the predictive power of the seed mass model using only complete seed mass data (i.e., without imputation). Contributions were derived from the model's internal importance scores and normalized so that their sum equals the cross-validated R^2. Bar lengths represent the proportion of explained variance, and colours indicate predictor groups: climate (black), dispersers (light grey) and soil (dark grey). **Figure S11:** Partial dependence of community weighted-mean (CWM) of seed mass (log scale) on each predictor variable, using only complete seed mass data (i.e., without imputation). We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other predictors constant at their sample mean. **Figure S12:** Area of applicability (AOA) of the spatial prediction models (a, b) and predicted spatial variation in seed mass (c, d), using complete and imputed data (a, c) and only complete data (b, d). Areas shown in grey fall outside the AOA (AOA = 0) and were excluded from predictions in panels a and b. Colours represent spatial predictions based on community-weighted mean values, with dark purple and blue indicating lower values and light green and yellow indicating higher values. **Figure S13:** Proportional contribution of each predictor to the predictive power of the models for endozoochory (a), anemochory (b), synzoochory (c) and seed mass (d), including both *sensu stricto* and *lato* classifications. Contributions were calculated from the models' internal importance scores and normalized so that their sum equals the cross-validated R^2. Bar lengths represent the proportion of explained variance, and colours indicate predictor groups: climate (black), dispersers (light grey) and soil (dark grey). **Figure S14:** Partial dependence of community weighted-mean (CWM) of endozoochory, including both *sensu stricto* and *sensu lato* classifications, on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other variables constant at their sample mean. **Figure S15:** Partial dependence of community weighted-mean (CWM) of anemochory, including both *sensu stricto* and *sensu lato* classifications, on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other variables constant at their sample mean. **Figure S16:** Partial dependence of community weighted-mean (CWM) of synzoochory, including both *sensu stricto* and *sensu lato* classifications, on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other variables constant at their sample mean. **Figure S17:** Prediction map of community-weighted means (CWM) of endozoochory (a), synzoochory (b) and anemochory (c), including both *sensu stricto* and *sensu lato* classifications, across Brazilian ecosystems. Dark purple and blue colours represent low CWM values, while light green and yellow colours represent high CWM values. **Figure S18:** Area of Applicability (AOA) of the spatial prediction model for endozoochory (a), synzoochory (b) and anemochory (c), including both *sensu stricto* and *sensu lato* classifications, across Brazilian ecosystems. Areas shown in grey fall outside the AOA (AOA = 0) and were excluded from predictions. Colours represent the spatial predictions of the models based on community-weighted mean (CWM) values, with dark purple and blue indicating lower values and light green and yellow indicating higher values.

Supporting information – Souza et al. *Global Ecology and Biogeography*

Table S1: Summary of sample sizes for each Brazilian regions studied.

Region	**N plots**	**area (ha)**	**N individuals**	**N species**
Atlantic forest	29	27.26	39,349	431
Amazon	206	169.35	71,500	2626
Eastern	45	19.51	10,906	688
Central	62	61.65	24,751	1,047
Western	45	44.51	12,266	432
Southern	54	43.68	23,577	459
Caatinga (dry forest)	29	16.68	27,702	147
Cerrado (Brazilian savanna)	36	33.42	33,634	280
Brazilian ecosystems	300	246.71	172,185	2,203

Table S2: Definitions of seed dispersal modes categories

Dispersal categories	Description
1) Endozoochory *sensu stricto*	endozoochorous; endozoochorous AND synzoochorous/ hydrochorous
2) Endozoochory *sensu lato*	endozoochorous OR (AND/OR) autochorous/synzoochorous /hydrochorous
3) Synzoochory *sensu stricto*	synzoochorous; synzoochorous AND autochorous/barochorous/ hydrochorous
4) Synzoochory *sensu lato*	synzoochorous OR autochorous/hydrochorous/barochorous AND/OR hydrochorous
5) Mymercochory *sensu stricto*	myrmecochorous; myrmecochorous and autochorous/barochorous
6) Mymercochory *sensu lato*	myrmecochorous AND/OR autochorous
7) Anemochory *sensu stricto*	anemochorous; anemochorous AND hydrochorous
8) Anemochory *sensu lato*	anemochorous OR autochorous/hydrochorous AND/OR hydrochorous
9) Autochory *sensu stricto*	autochorous
10) Autochory *sensu lato*	autochorous OR barochorous/hydrochorous AND/OR hydrochorous

Table S3: Proportion of dispersal modes (*sensu stricto* and *sensu lato*) based on 2,203 species and 172,220 individuals within 300 old-growth plots.

Dispersal mode	N species	%	N individuals	%
Anemochory *sensu lato*	29	1.32	2,131	1.24
Anemochory *sensu stricto*	301	13.66	37,110	21.55
Autochory *sensu lato*	18	0.82	1,632	0.95
Autochory *sensu stricto*	91	4.13	12,490	7.25
Barochory	1	0.05	81	0.05
Endozoochory *sensu lato*	151	6.85	9,047	5.25
Endozoochory *sensu stricto*	1,308	59.37	90,641	52.64
Hydrochory	18	0.82	383	0.22
Mymercochory *sensu lato*	2	0.09	169	0.10
Mymercochory *sensu stricto*	28	1.27	3,469	2.01
Synzoochory *sensu lato*	57	2.59	4,912	2.85
Synzoochory *sensu stricto*	138	6.26	9,125	5.30
unknown	61	2.77	995	0.58

Table S4: Statistical results from Generalized Linear Models (GLMs) testing differences in community weighted-means (CWM) of dispersal modes and seed mass across Brazilian ecosystems. Test values (χ^2), degrees of freedom (df), and p-values are shown for each model. All models indicate significant differences among ecosystems ($p < 0.001$).

Models	χ^2	df	p value
Anemochory	324.19	6	< 0.001
Endozoochory	85.41	6	<0.001
Synzoochory	73.43	6	<0.001
Seed mass	382.36	6	<0.001

Table S5: Descriptions of the climate and soil predictor variables used in this study.

Code	Description	Unit	Min-Max
Climate			
MeanDiuRan	Mean Temperature Diurnal Range (Mean of monthly (max temp - min temp))	º C	7.91-14.80
ISO	Isothermality (MeanDiuRan/TempAnnRan) *100	%	55.9-85.93
ST	Temperature Seasonality (standard deviation *100)	º C	28.95-260.51
TempAnnRan	Temperature Annual Range (MaxTempWarMon-MinTempColMon)	º C	9.80-21.29
MAP	Annual Precipitation	mm	507-2569
SP	Precipitation Seasonality (Coefficient of Variation)	mm	39.21-109.89
PrecWetQuar	Precipitation of Wettest Quarter	mm	219-1200
PrecDriQuar	Precipitation of Driest Quarter	mm	3-335
PET	Potential evapotranpiration	mm	1,121-1,877
Soil			
Chemical			
OCS	Organic carbon stocks	kg/m²	1.9-8.7
CEC	Cation exchange capacity	cmol(c)/kg	5.48-21.16

PH	pH of water	pH	4.13-6.83
Physical			
BDOD	Bulk density of the fine earth fraction	kg/dm³	1.08-1.44
CLAY	Proportion of clay particles (< 0.002 mm) in the fine earth fraction	g/100g (%)	21.25-68.13

Table S6: List of frugivorous primate species including their body mass (g) and degree of frugivory (%) according to Elton traits database.

Primate species*	**Body mass (g)**	**Degree of frugivory (%)**
Alouatta belzebul	6400	40
Alouatta caraya	5862.5	40
Alouatta discolor	6262.8	40
Alouatta guariba	5589.2	40
Alouatta juara	6573.34	40
Alouatta macconnelli	7554.8	40
Alouatta nigerrima	7662	40
Alouatta puruensis	6573.34	40
Alouatta ululata	6682.1	40
Aotus azarae	963	20
Aotus nigriceps	1060	20
Aotus trivirgatus	900	20
Ateles belzebuth	5000	60
Ateles chamek	6000	60
Ateles marginatus	6000	60
Ateles paniscus	7900.1	60
Brachyteles arachnoides	13499.9	20
Brachyteles hypoxanthus	8865	20
Callimico goeldii	480	50
Cebus albifrons	2629	20
Cebus kaapori	2600	70
Cebus olivaceus	2614.5	20
Cebus unicolor	2614.5	30
Lagothrix lagotricha	846	70

Leontocebus fuscicollis	387	30
Leontocebus weddelli	598	30
Leontopithecus rosalia	535.5	50
Saguinus bicolor	465	30
Saguinus imperator	400	30
Saguinus labiatus	575	30
Saguinus melanoleucus	491	30
Saguinus midas	540	30
Saguinus mystax	618	30
Saguinus niger	591.2	30
Saguinus ursulus	711	30
Saimiri boliviensis	615	20
Saimiri cassiquiarensis	759.8	20
Saimiri collinsi	681	20
Saimiri sciureus	743.2	20
Saimiri ustus	1000	20
Sapajus apella	2500	20
Sapajus libidinosus	3200	40
Sapajus nigritus	3200	40
Sapajus robustus	3200	40
Sapajus xanthosternos	3200	40

* To avoid occasional frugivores (<20% of fruits in the diet), we considered only those species of primates with a primary (>50%) or secondary (20%<50%) frugivory degree. In addition, we excluded records of *Pitheciid* genera (*Callicebus*, *Cheracebus*, and *Pithecia*), which are medium-to-large primates that primarily act as seed predators rather than seed dispersers (Hawes and Peres, 2014; Fuzessy et al. 2021).

Table S7: Proportion of dispersal modes (*sensu stricto* and *sensu lato*) (%) by Brazilian ecosystems, based on 2,203 species and 172,220 individuals from 300 old-growth plots. Dispersal modes analysed in this study are highlighted in bold.

Dispersal mode	**Atlantic forest**	**Amazon**				**Caatinga**	**Cerrado**
		Eastern	*Central*	*Southern*	*Western*		
Anemochory *sensu lato*	0.81	0.31	0.19	0.25	0.31	3.40	0.28
Anemochory *sensu stricto*	**8.69**	**4.60**	**4.37**	**23.43**	**4.60**	**34.01**	**40.15**
Autochory *sensu lato*	0.02	0.81	0.00	0.40	0.81	3.40	0.88
Autochory *sensu stricto*	7.52	10.10	2.89	0.74	10.10	12.93	8.80
Endozoochory *sensu lato*	3.22	6.42	9.90	7.02	6.42	5.44	2.07
Endozoochory *sensu stricto*	**75.42**	**47.02**	**57.40**	**60.96**	**47.02**	**29.93**	**45.35**
Hydrochory	0.06	0.69	0.97	0.03	0.69	1.34	0.00
Mymercochory *sensu lato*	0.00	0.00	0.00	0.00	0.00	0.61	0.00
Mymercochory *sensu str*	2.69	1.22	0.09	0.47	1.22	0.48	0.57
Synzoochory *sensu lato*	0.00	11.11	9.87	5.24	11.11	4.49	0.00
Synzoochory *sensu stricto*	**1.50**	**16.38**	**12.68**	**13.01**	**16.38**	**9.55**	**1.80**
unknown	0.08	0.60	1.63	0.76	0.60	3.06	0.11

Table S8: Moran's I tests for spatial autocorrelation of residuals from dispersal mode and seed mass models. Results are reported for normality, randomisation, and Monte Carlo methods. Columns show Moran's I, expectation, variance, standardized deviate (Z), and p-value.

Model	Method	Moran I	Expectation	Variance	Z	p-value
	Normality	-0.02	-0.00	0.00	-1.59	0.94
	Randomisation	-0.02	-0.00	0.00	-1.60	0.94
Endozoochory	Monte-Carlo	-0.02	—	—	—	0.97
	Normality	-0.01	-0.00	0.00	-0.95	0.83
	Randomisation	-0.02	-0.00	0.00	-0.96	0.83
Anemochory	Monte-Carlo	-0.01	—	—	—	0.82
	Normality	-0.02	-0.00	0.00	-1.24	0.89
	Randomisation	-0.02	-0.00	0.00	-1.26	0.89
Synzoochory	Monte-Carlo	-0.02	—	—	—	0.91
Seed mass	Normality	-0.04	-0.00	0.00	-2.72	0.99
(complete and	Randomisation	-0.01	-0.00	0.00	-2.73	0.99
imputed data)	Monte-Carlo	-0.04	—	—	—	0.99
	Normality	-0.04	-0.00	0.00	-2.72	0.99
Seed mass (only	Randomisation	-0.04	-0.00	0.00	0.00	0.99
complete data)	Monte-Carlo	-0.04	—	—	—	0.99

Table S9: Model performance for dispersal mode models, including both *sensu stricto* and *sensu lato* classifications, assessed using stratified k-fold cross-validation under balanced and unbalanced sampling schemes (see Methods). Performance metrics include: (1) root mean square error (RMSE), indicating the average prediction error (proportion for dispersal modes and milligrams for seed mass), (2) the correlation coefficient between observed and predicted values, and (3) the coefficient of determination [1 − (residual sum of squares / total sum of squares)].

Model RF	**Cross-validation**					
	Balanced			**Unbalanced**		
	RMSE (SD)	**Correlation (SD)**	**R^2 (SD)**	**RMSE**	**Correlation**	**R^2**
Endozoochory	0.16 (0.02)	0.72 (0.09)	0.49 (0.14)	0.15	0.71	0.50
Anemochory	0.16 (0.02)	0.69 (0.09)	0.44 (0.14)	0.14	0.71	0.47
Synzoochory	0.08 (0.02)	0.77 (0.12)	0.59 (0.17)	0.08	0.79	0.61

Table S10: Model performance assessed using spatial block cross-validation under different block sizes (300, 500, and 800 km). Performance metrics include: (1) root mean square error (RMSE), indicating the average prediction error (proportion for dispersal modes and milligrams for seed mass), (2) the correlation coefficient between observed and predicted values, and (3) the coefficient of determination [1 − (residual sum of squares / total sum of squares)].

Model RF	**Spatial cross-validation**		
Endozoochory	**RMSE**	**Correlation**	**R^2**
300 km	0.19	0.46	0.17
500 km	0.21	0.31	0.02
800 km	0.20	0.21	-0.03
Anemochory			
300 km	0.16	0.51	0.11
500 km	0.18	0.45	-0.02
800 km	0.17	0.36	-0.13
Synzoochory			
300 km	0.07	0.50	0.11
500 km	0.07	0.60	0.24
800 km	0.07	0.39	0.00
Seed mass (complete + imputed data)			
300 km	271.74	0.66	0.19
500 km	289.81	0.58	0.14
800 km	260.76	0.59	0.14
Seed mass (only complete data)			

300km	313.13	0.67	0.19
500km	323.58	0.62	0.18
800km	303.21	0.55	0.09

47

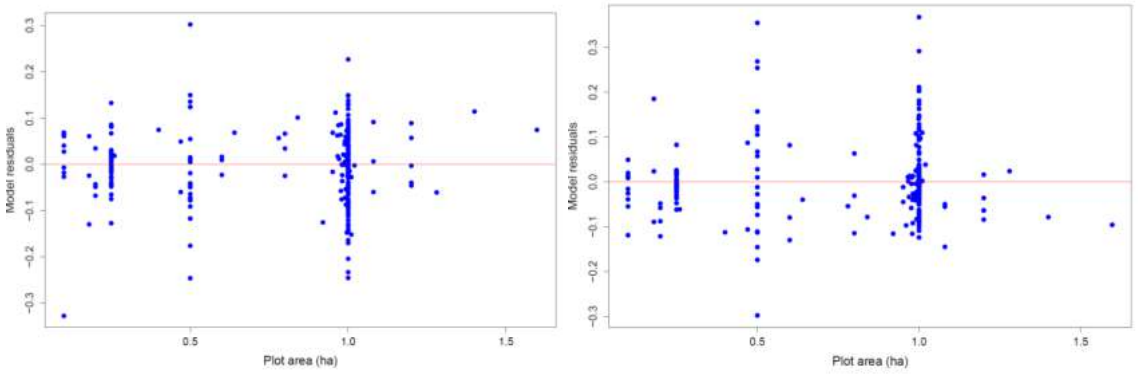
48

Table S11. Cross-validated model performance across Brazilian ecosystems for the main models of seed mass and dispersal modes (endozoochory, anemochory, and synzoochory). Model performance was assessed using stratified 5-fold cross-validation, and predictive metrics were calculated separately for each biome. Performance metrics include: (1) root mean square error (RMSE), indicating the average prediction error (proportion for dispersal modes and milligrams for seed mass), (2) the correlation coefficient (Cor) between observed and predicted values, and (3) the coefficient of determination [1 − (residual sum of squares / total sum of squares)].

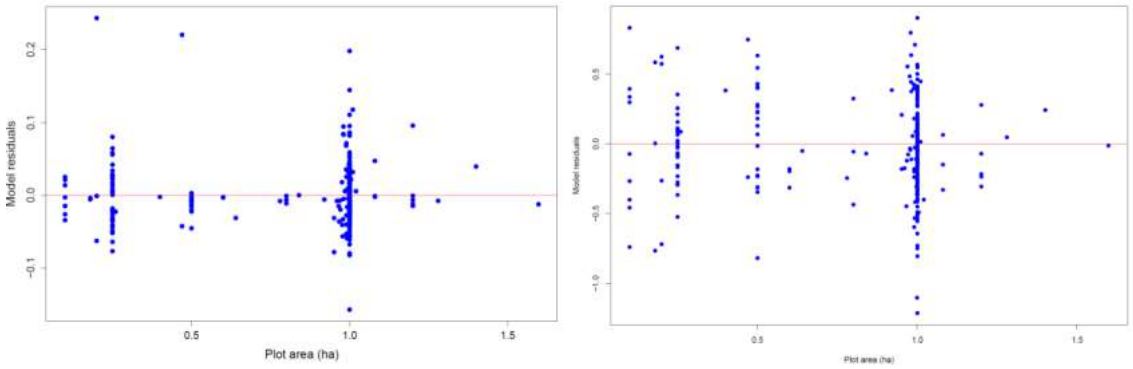
Model	Region	RMSE	R^2	Cor
Endozoochory	Atlantic Forest	0.14	0.51	0.73
	Caatinga	0.24	-0.00	0.21
	Central Amazon	0.09	0.25	0.52
	Cerrado	0.21	-0.21	0.03
	Eastern Amazon	0.12	0.33	0.58
	Southern Amazon	0.18	0.14	0.4
	Western Amazon	0.11	0.24	0.49
Anemochory	Atlantic Forest	0.1	-1.21	0.51
	Caatinga	0.25	0.07	0.28
	Central Amazon	0.06	-0.41	-0.09
	Cerrado	0.22	-0.40	-0.02
	Eastern Amazon	0.04	0.16	0.6
	Southern Amazon	0.14	0.28	0.53
	Western Amazon	0.09	0.4	0.66

	Atlantic Forest	0.06	-0.15	0.03
	Caatinga	0.01	-0.73	0.59
	Central Amazon	0.05	0.32	0.57
Synzoochory	Cerrado	0.02	0.00	0.36
	Eastern Amazon	0.07	0.29	0.55
	Southern Amazon	0.07	0.01	0.23
	Western Amazon	0.05	0.37	0.62
	Atlantic Forest	92.2	0.57	0.76
	Caatinga	198	-0.09	0.21
	Central Amazon	198	0.36	0.65
Seed mass	Cerrado	95.7	-1.19	0.14
	Eastern Amazon	268	0.46	0.69
	Southern Amazon	193	0.00	0.29
	Western Amazon	341	0.16	0.46

(a) (b)



(c) (d)



(e)

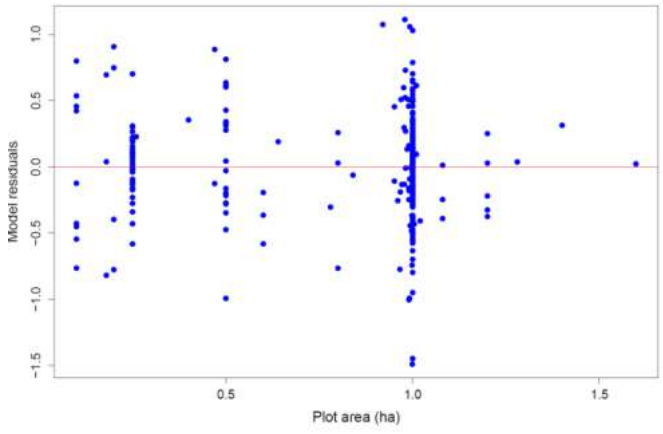


Figure S1: Model residuals as a function of plot area for endozoochory (a), anemochory (b), synzoochory (c), seed mass with complete and imputed data (d), and seed mass using only complete data (e). Blue points represent individual plots, and the red line corresponds to zero

residuals, separating positive residuals (above the line) from negative residuals (below the line).

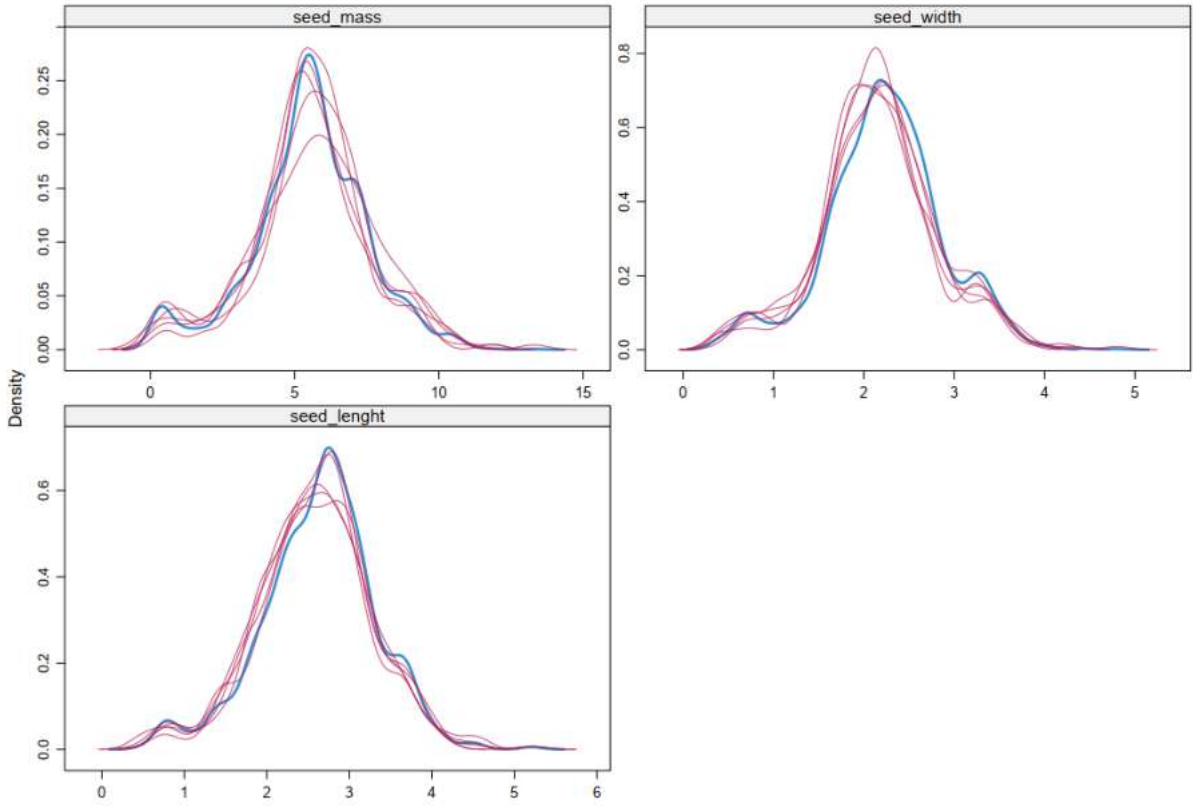


Figure S2: Plots show the density of seed-size distribution at species level for the original data (blue line) and five imputed datasets (red) on log scale. Trait imputation was performed through chained equations by predictive mean matching, using R package *mice*.

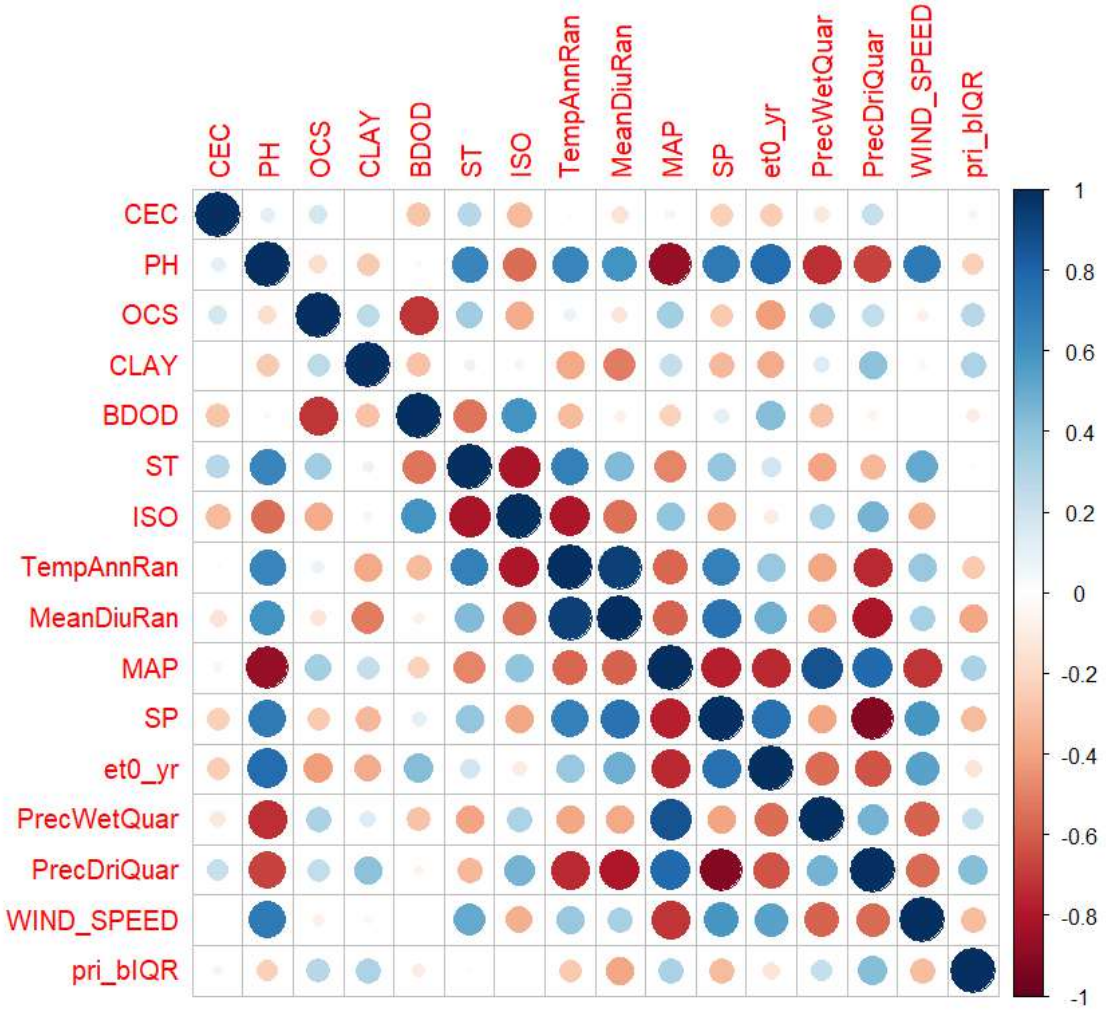
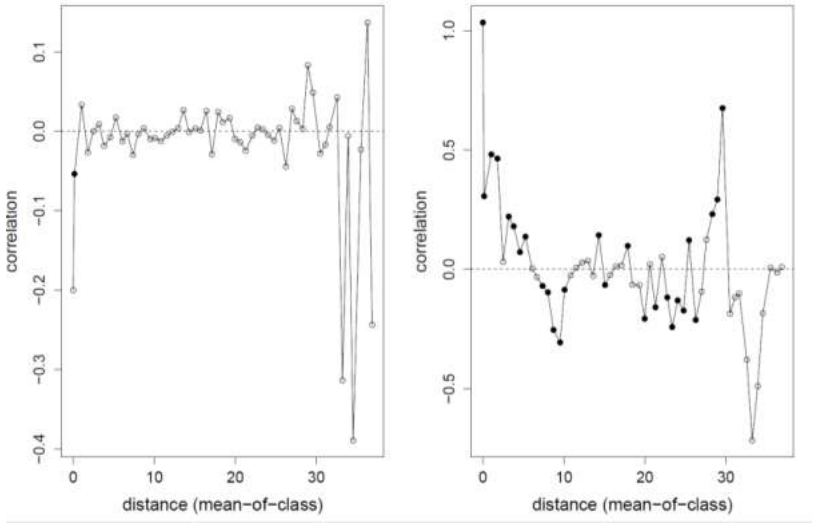
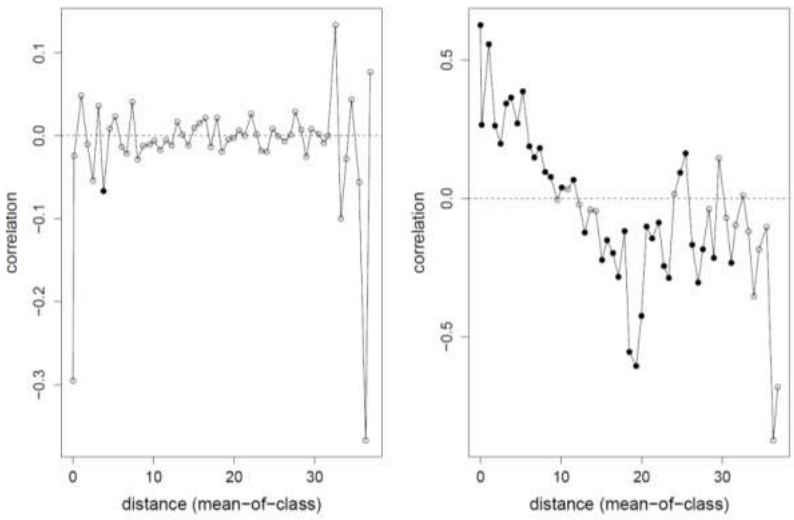


Figure S3: Correlation structure among selected climatic, soil, and dispersal variables. The correlograms illustrate the Pearson correlations between predictors used in the models by the colour and size of disk. Red colour indicates negative correlations, and blue colour indicates negative correlations. Disk size indicates the strength of the correlations from weak (small) to strong (large). **MeanDiuRan:** Mean Temperature Diurnal Range; **ISO:** Isothermality; **ST:** Temperature Seasonality; **TempAnnRan:** Temperature Annual Range; **MAP:** Annual Precipitation; **SP**: Precipitation Seasonality; **PrecWetQuar:** Precipitation of Wettest Quarter, **PrecDriQuar:** Precipitation of Driest Quarter, **PrecWarQuar:** Precipitation of Warmest Quarter; **OCS:** Organic carbon stocks, **CEC:** Cation exchange capacity, **PH:** pH of water; **BDOD:** Bulk density; **CLAY:** Proportion of clay particles; pri_bIQR: **IQR** of primate body mass; **WIND SPEED:** Wind speed; **et0_yr:** mean annual evapotranspiration.

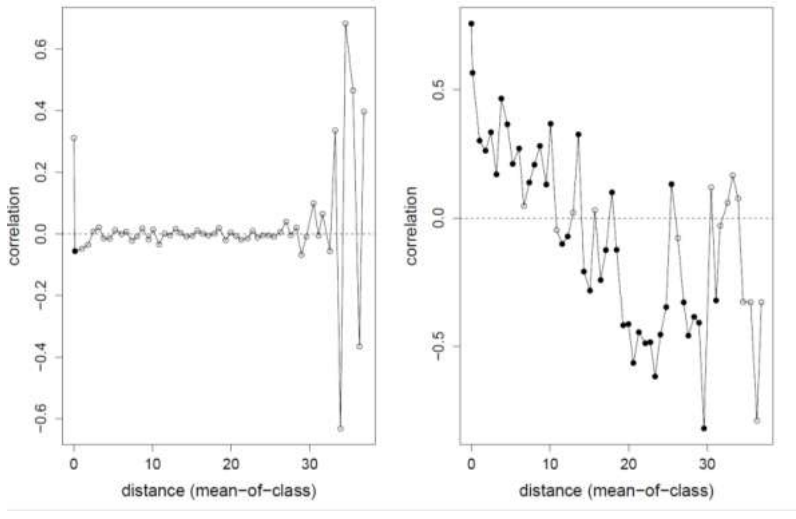
(a) (b)

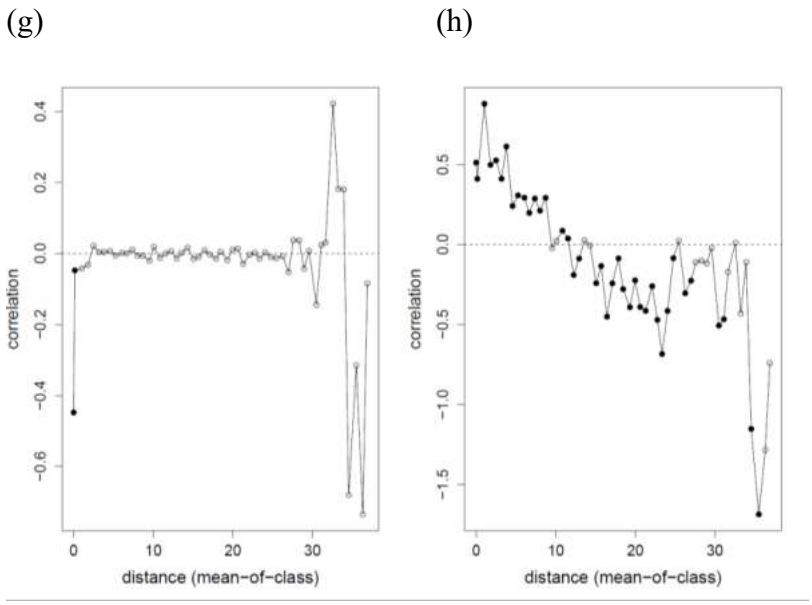

correlation
distance (mean−of−class)
correlation
distance (mean−of−class)

(c) (d)


correlation
distance (mean−of−class)
correlation
distance (mean−of−class)

(e) (f)


correlation
distance (mean−of−class)
correlation
distance (mean−of−class)



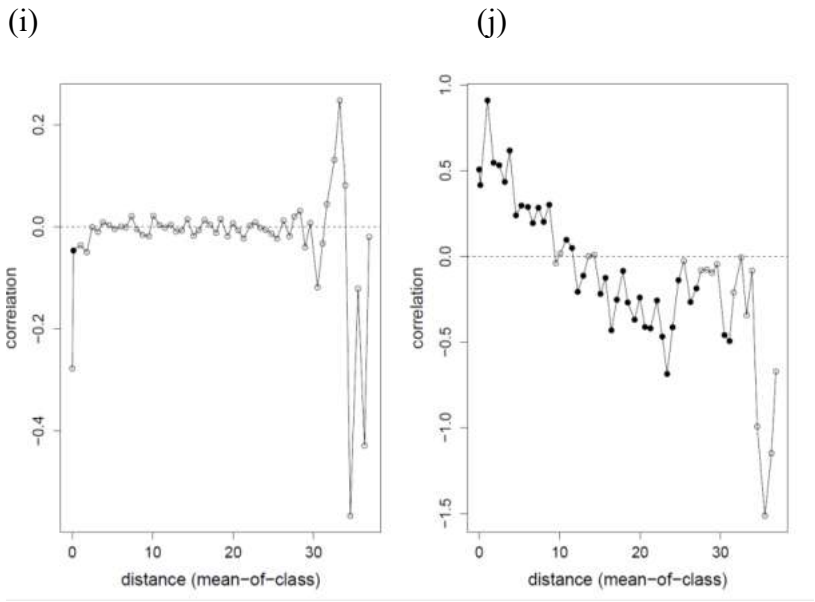


Figure S4: Spatial correlograms across distance classes, based on raw data (left) and model residuals (right), for endozoochory (a, b), anemochory (c, d), synzoochory (e, f), seed mass with complete and imputed data (g, h), and seed mass using only complete data (i, j). Filled symbols indicate statistically significant correlations, whereas open symbols indicate non-significant correlations. The dashed horizontal line indicates zero correlation.

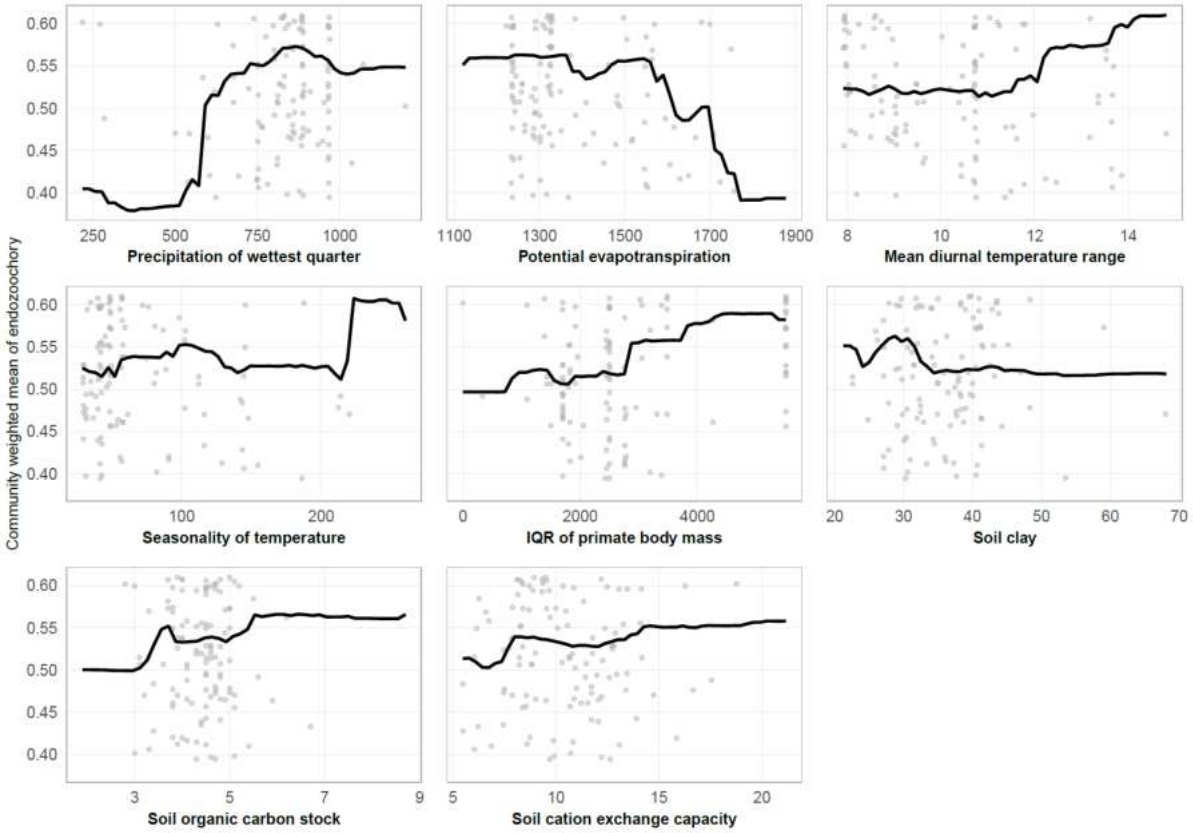


Figure S5: Partial dependence of community weighted-mean (CWM) of endozoochory on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other variables constant at their sample mean.

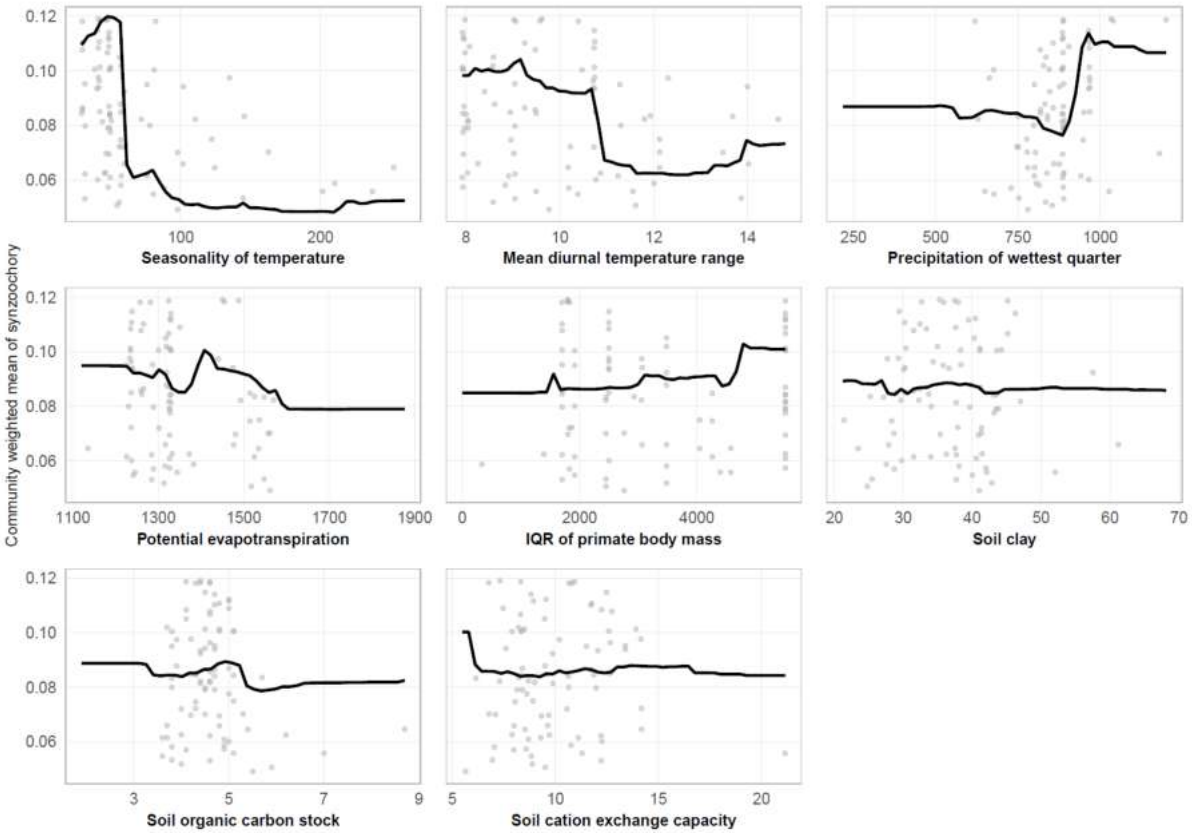


Figure S6: Partial dependence of community weighted-mean (CWM) of synzoochory on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other predictors constant at their sample mean.

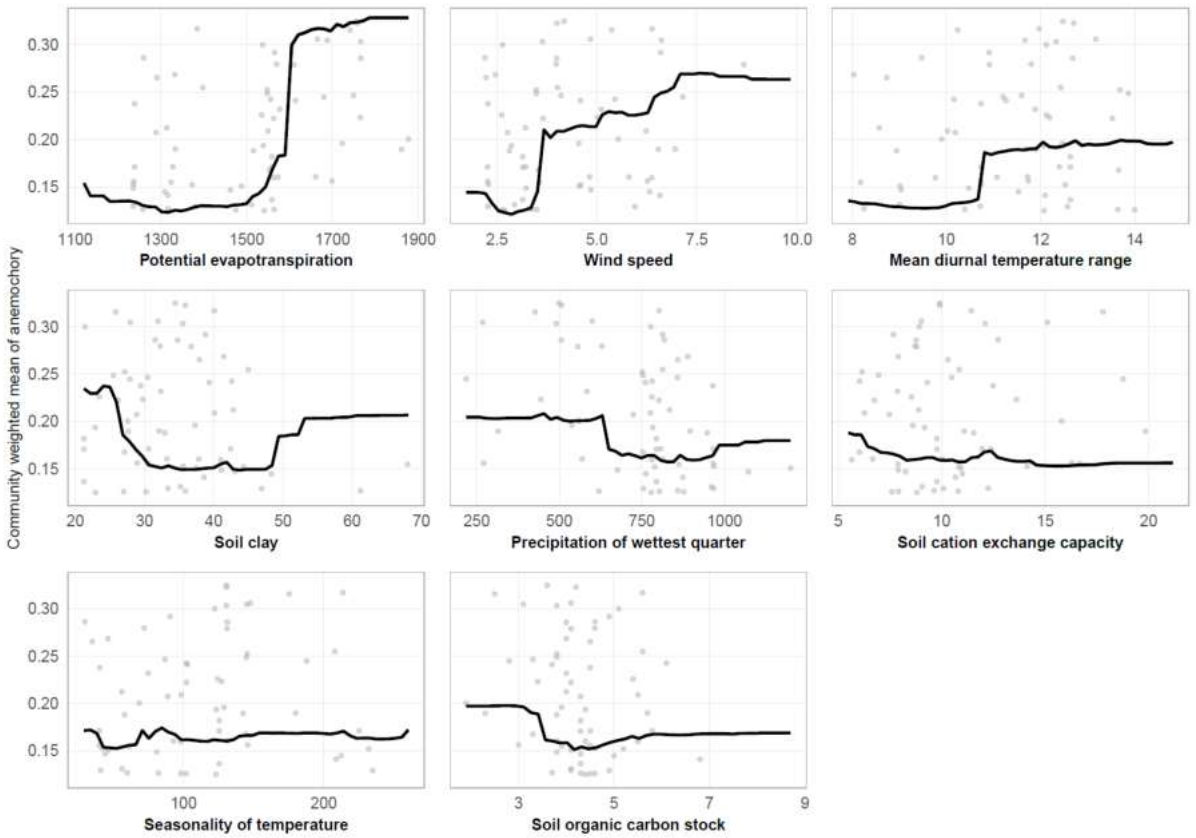
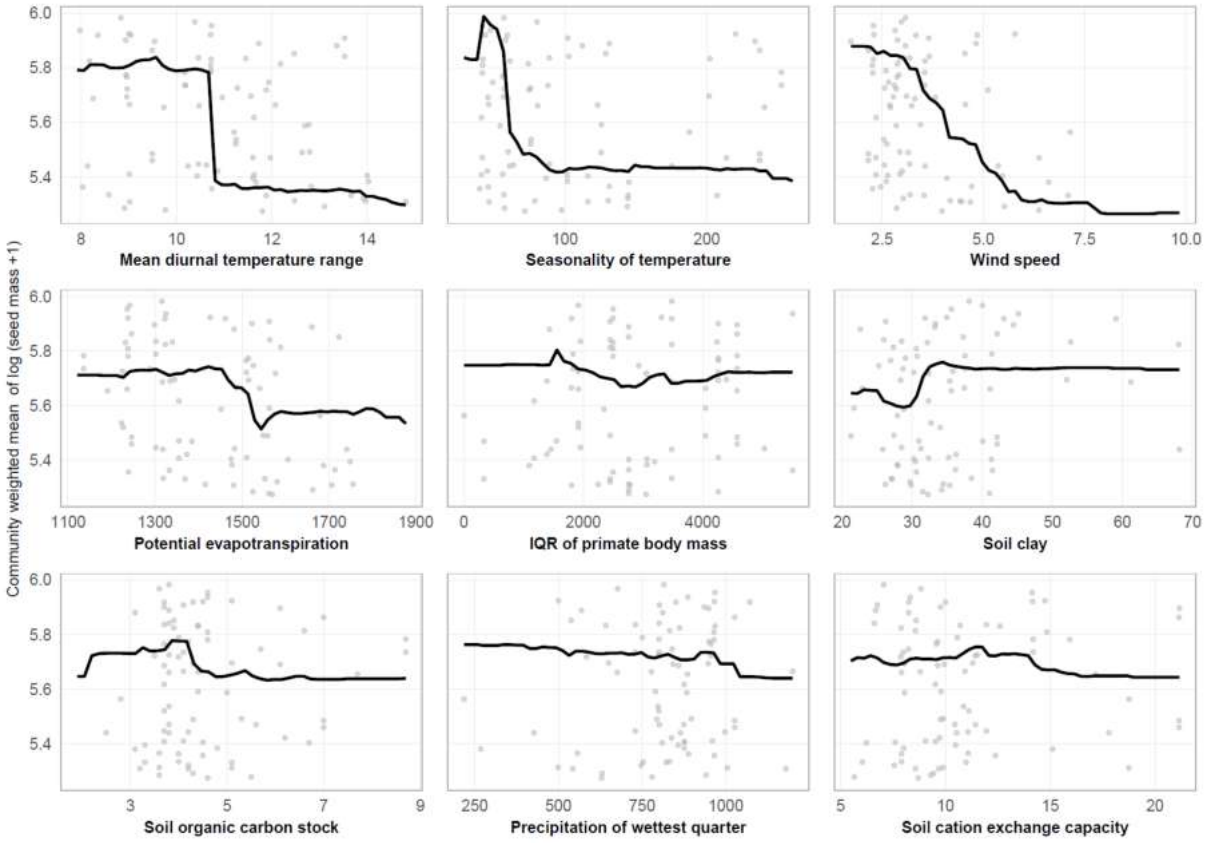


Figure S7: Partial dependence of community weighted-mean (CWM) of anemochory on each predictor variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other predictors constant at their sample mean.



117

118 **Figure S8:** Partial dependence of community weighted-mean (CWM) of log (seed mass+1)
119 on each predictor variable, using both complete and imputed seed mass data. We draw a
120 partial dependence plot showing the mean marginal influence of climate, dispersal agents and
121 soil variables based on the CWM values while keeping all other predictors constant at their
122 sample mean.

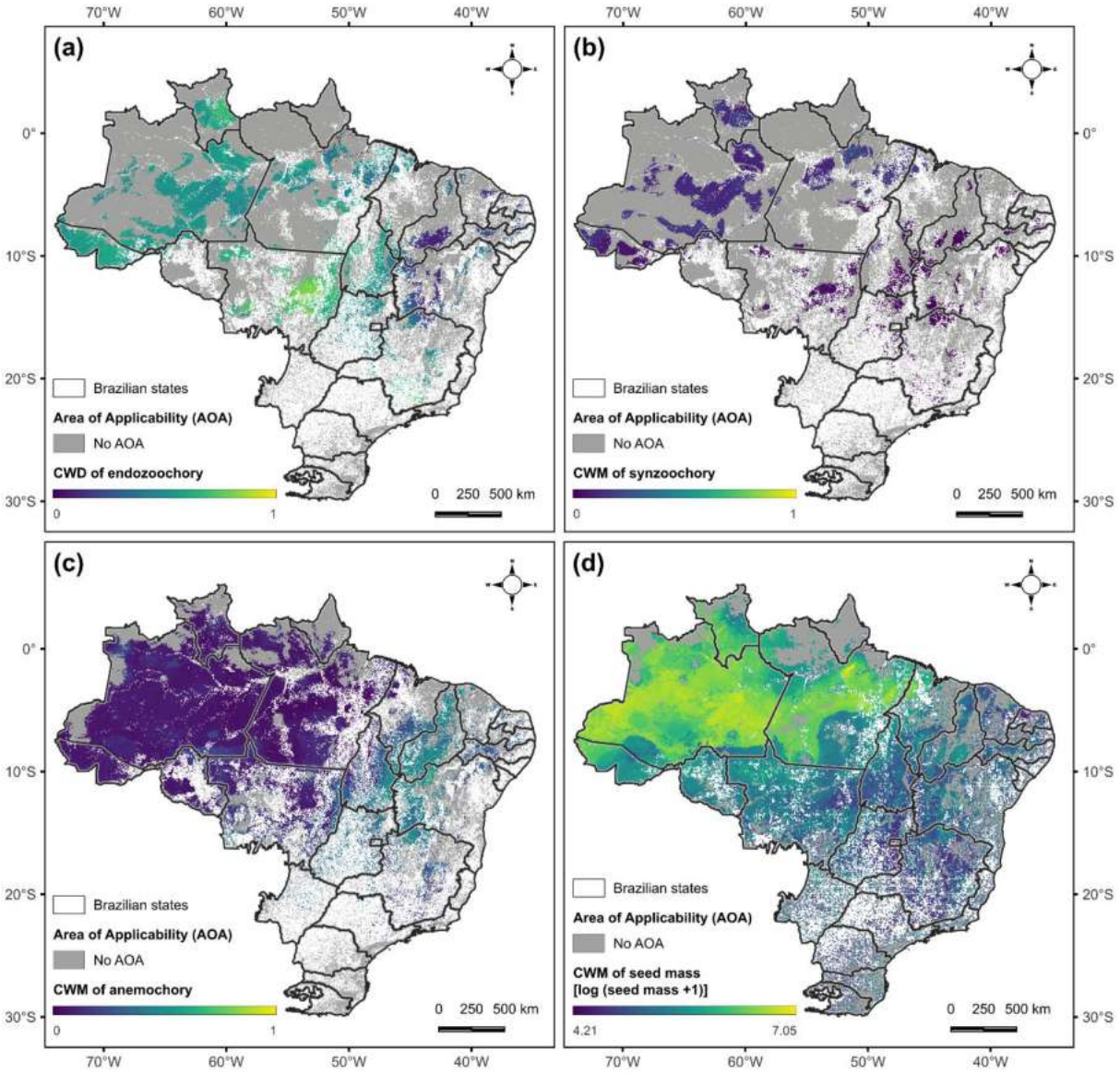
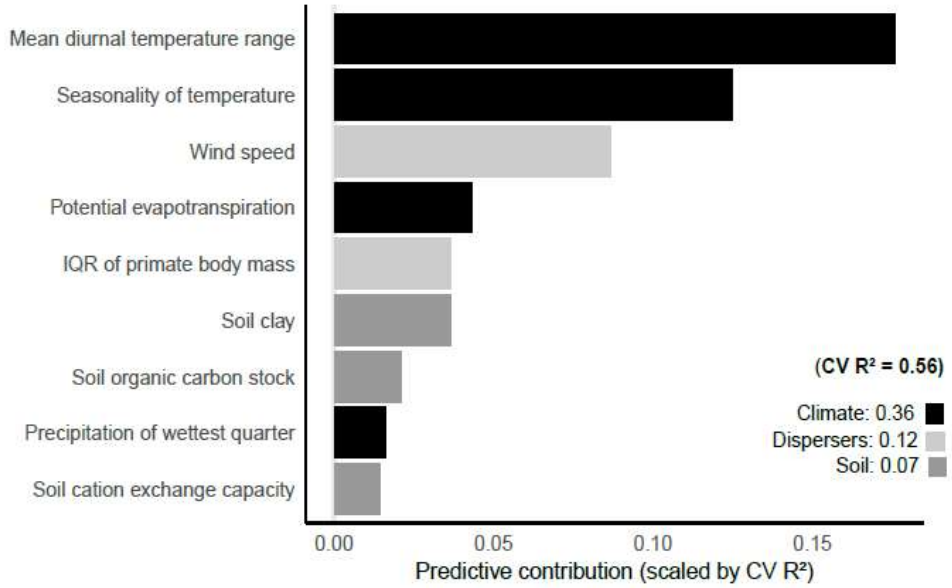


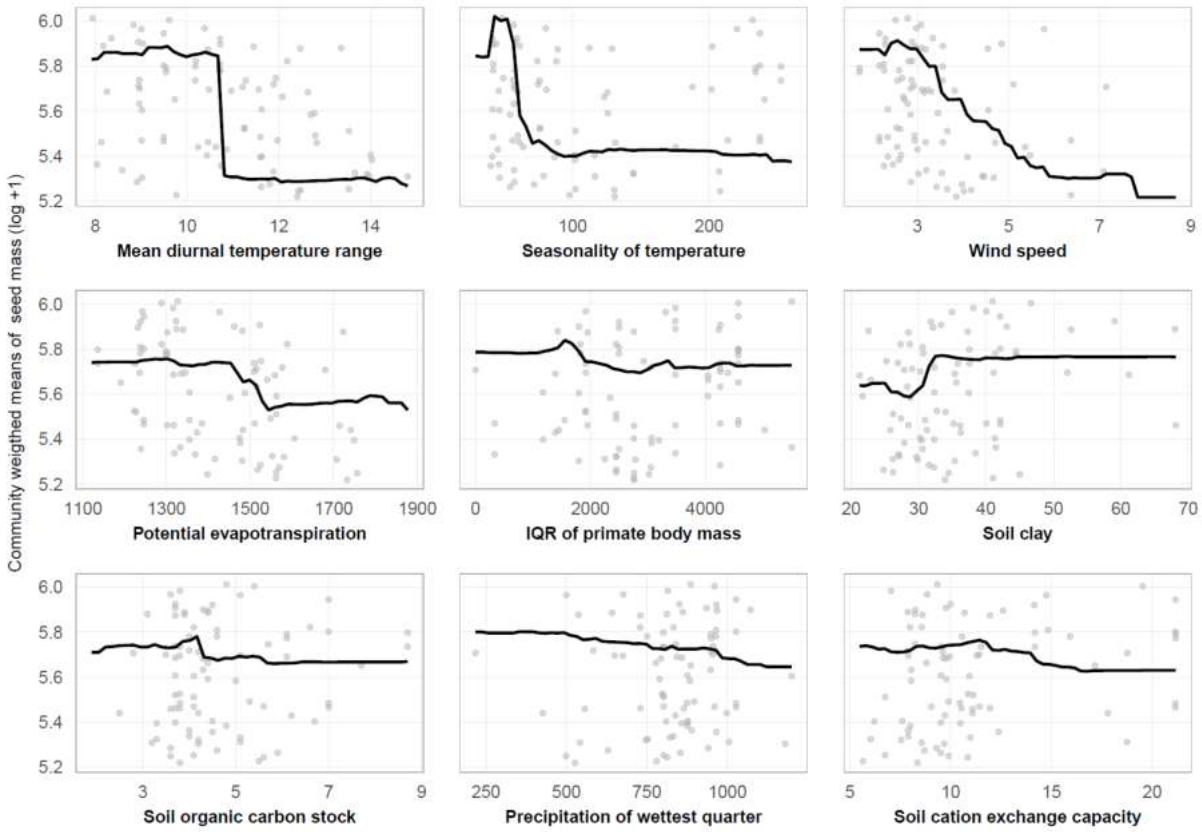
Figure S9: Area of Applicability (AOA) of the spatial prediction model for endozoochory (a), synzoochory (b), anemochory (c), and seed mass (d) across Brazilian ecosystems. Areas shown in grey fall outside the AOA (AOA = 0) and were excluded from predictions. Colours represent the spatial predictions of the models based on community-weighted mean values, with cooler tones (blue and green) indicating lower values and warmer tones (yellow and red) indicating higher values.



130

131 **Figure S10:** Proportional contribution of each predictor to the predictive power of the seed
132 mass model using only complete seed mass data (i.e., without imputation). Contributions were
133 derived from the model's internal importance scores and normalised so that their sum equals
134 the cross-validated R^2. Bar lengths represent the proportion of explained variance, and colours
135 indicate predictor groups: climate (black), dispersers (light grey), and soil (dark grey).

136



137

138 **Figure S11:** Partial dependence of community weighted-mean (CWM) of seed mass (log
139 scale) on each predictor variable, using only complete seed mass data (i.e., without
140 imputation).We draw a partial dependence plot showing the mean marginal influence of
141 climate, dispersal agents and soil variables based on the CWM values while keeping all other
142 predictors constant at their sample mean.

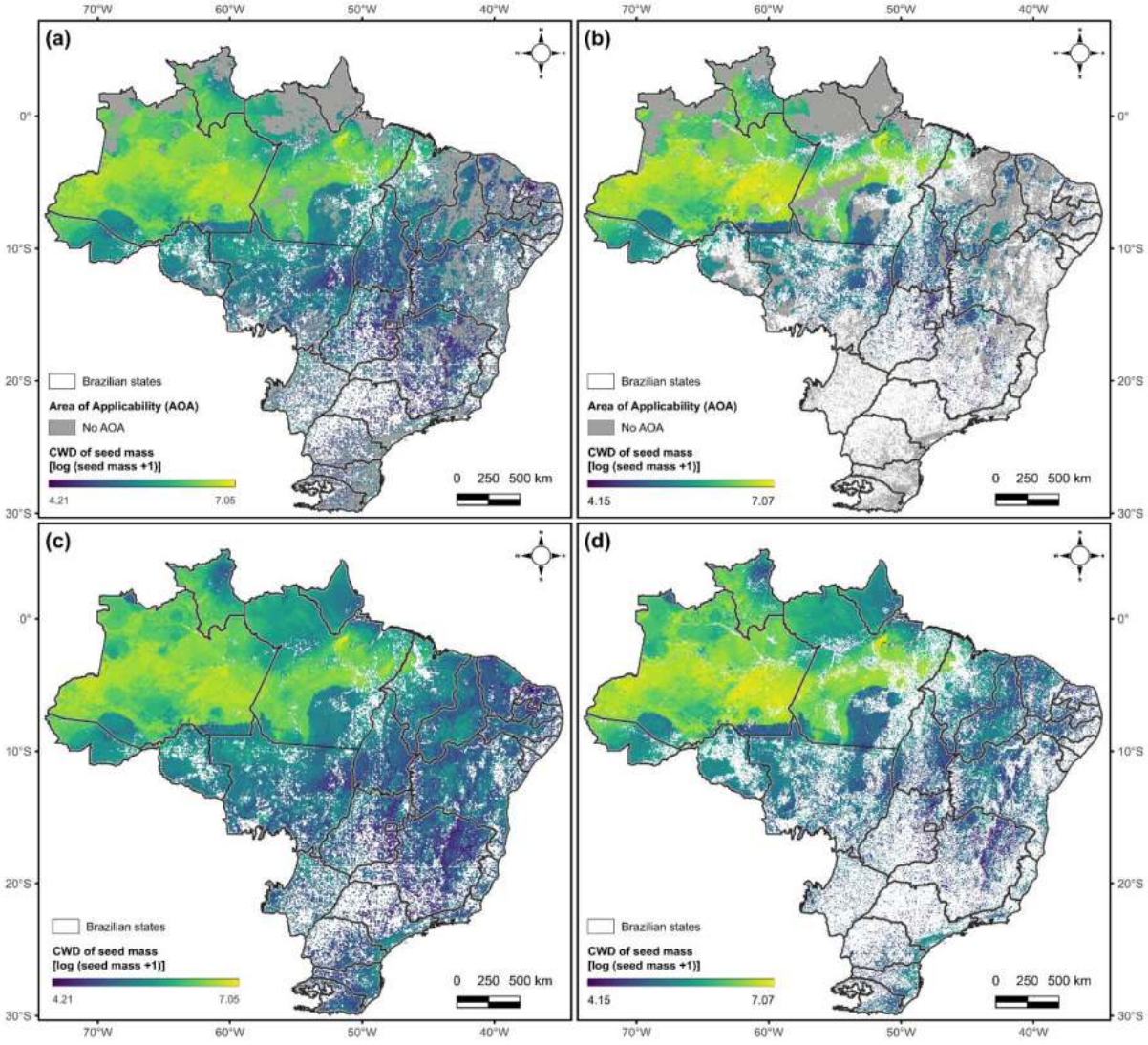


Figure S12: Area of applicability (AOA) of the spatial prediction models (a, b) and predicted spatial variation in seed mass (c, d), using complete and imputed data (a, c) and only complete data (b, d). Areas shown in grey fall outside the AOA (AOA = 0) and were excluded from predictions in panels a and b. Colours represent spatial predictions based on community-weighted mean values, with cooler tones (blue and green) indicating lower values and warmer tones (yellow and red) indicating higher values.

Sensitivity analyses – Dispersal modes (*sensu stricto* and *sensu lato*)

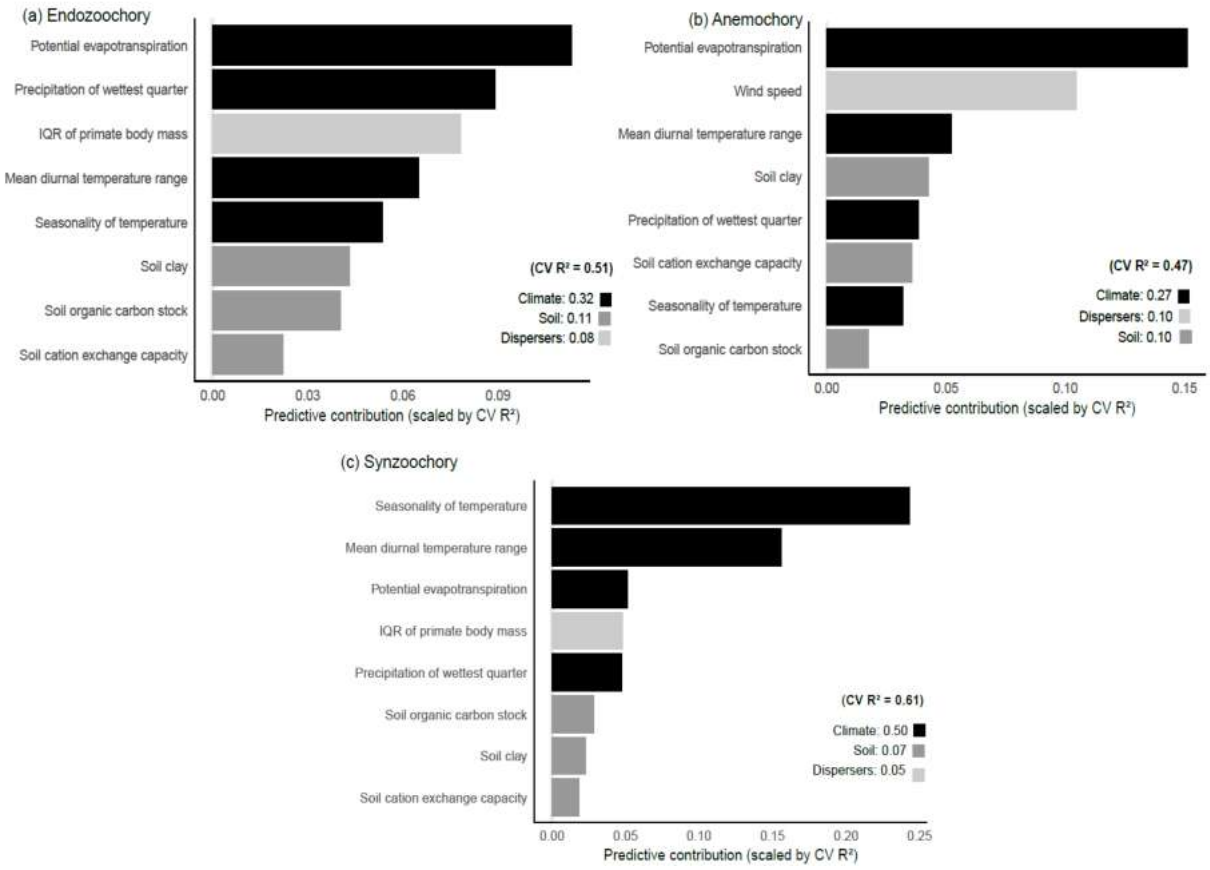
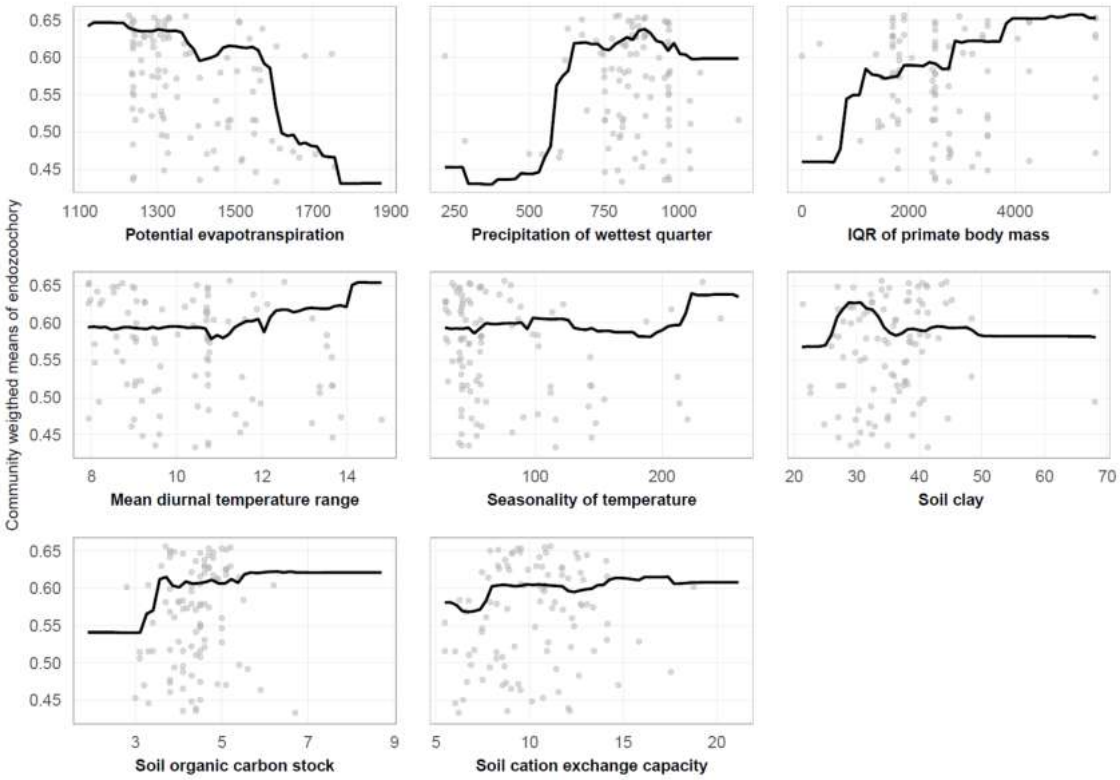
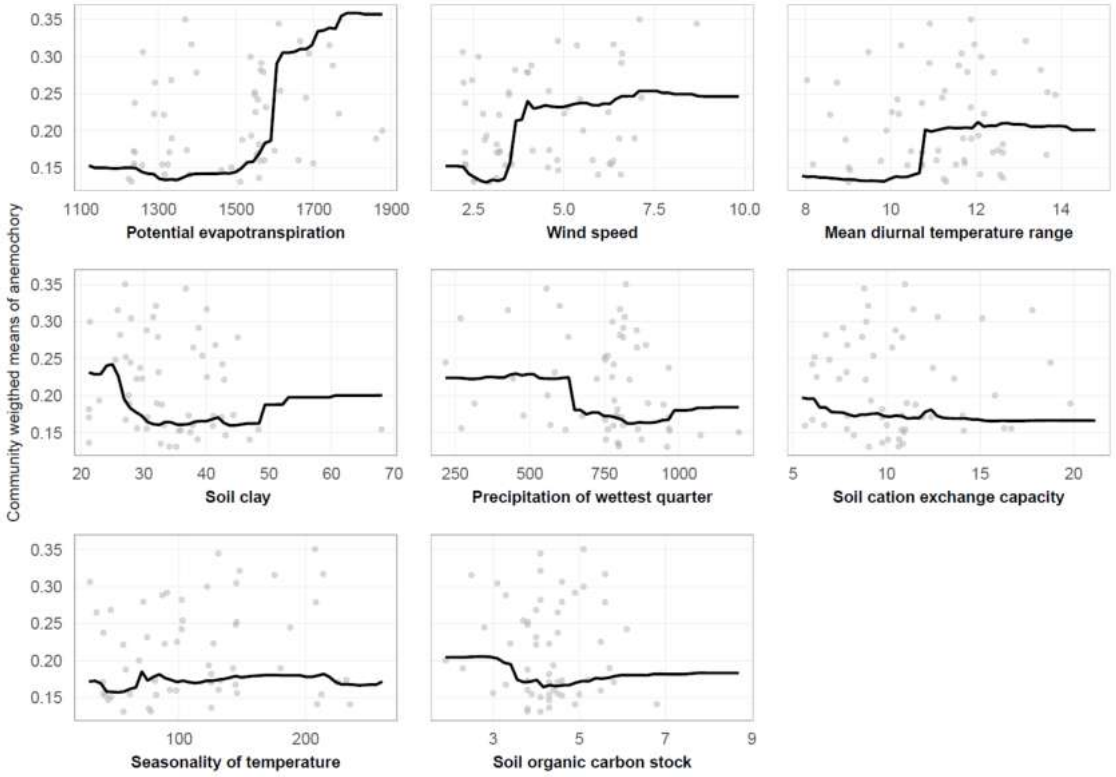


Figure S13: Proportional contribution of each predictor to the predictive power of the models for endozoochory (a), anemochory (b), synzoochory (c), and seed mass (d), including both *sensu stricto* and *lato* classifications. Contributions were calculated from the models' internal importance scores and normalised so that their sum equals the cross-validated R^2. Bar lengths represent the proportion of explained variance, and colours indicate predictor groups: climate (black), dispersers (light grey), and soil (dark grey).



158

159 **Figure S14**: Partial dependence of community weighted-mean (CWM) of endozoochory,
160 including both *sensu stricto* and *sensu lato* classifications, on each explanatory variable. We
161 draw a partial dependence plot showing the mean marginal influence of climate, dispersal
162 agents and soil variables based on the CWM values while keeping all other variables constant
163 at their sample mean.



164

165 **Figure S15:** Partial dependence of community weighted-mean (CWM) of anemochory,
166 including both *sensu stricto* and *sensu lato* classifications, on each explanatory variable. We
167 draw a partial dependence plot showing the mean marginal influence of climate, dispersal
168 agents and soil variables based on the CWM values while keeping all other variables constant
169 at their sample mean.

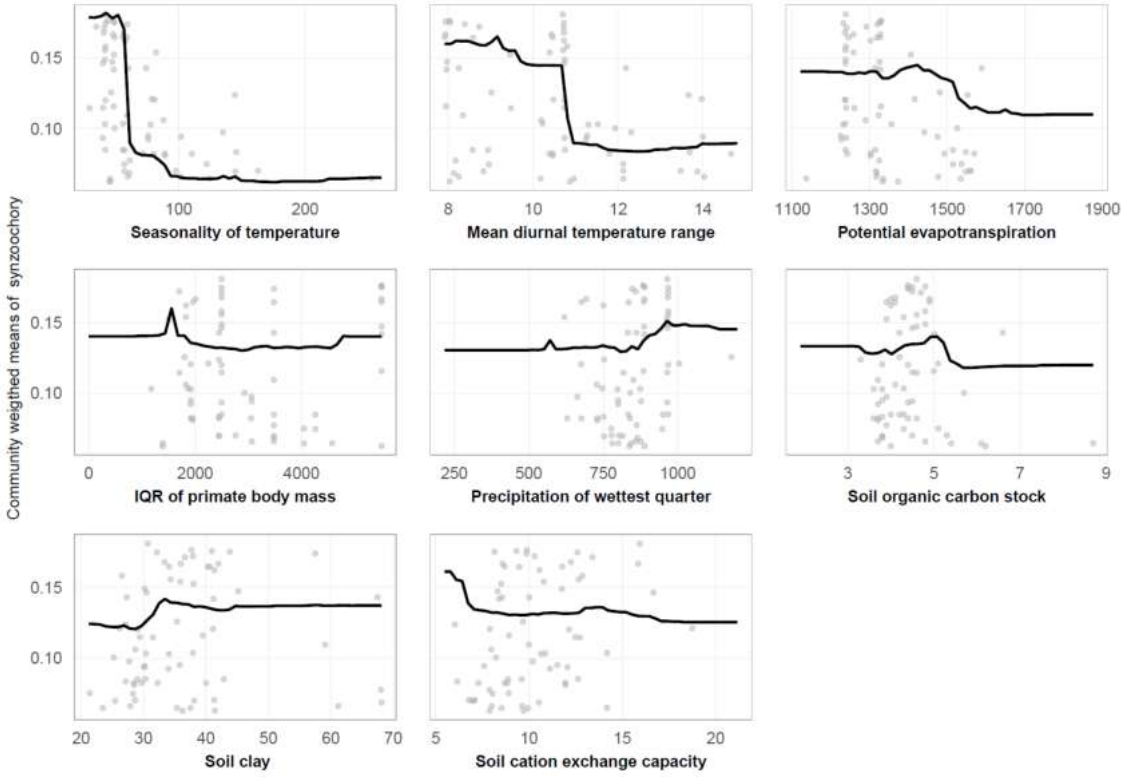
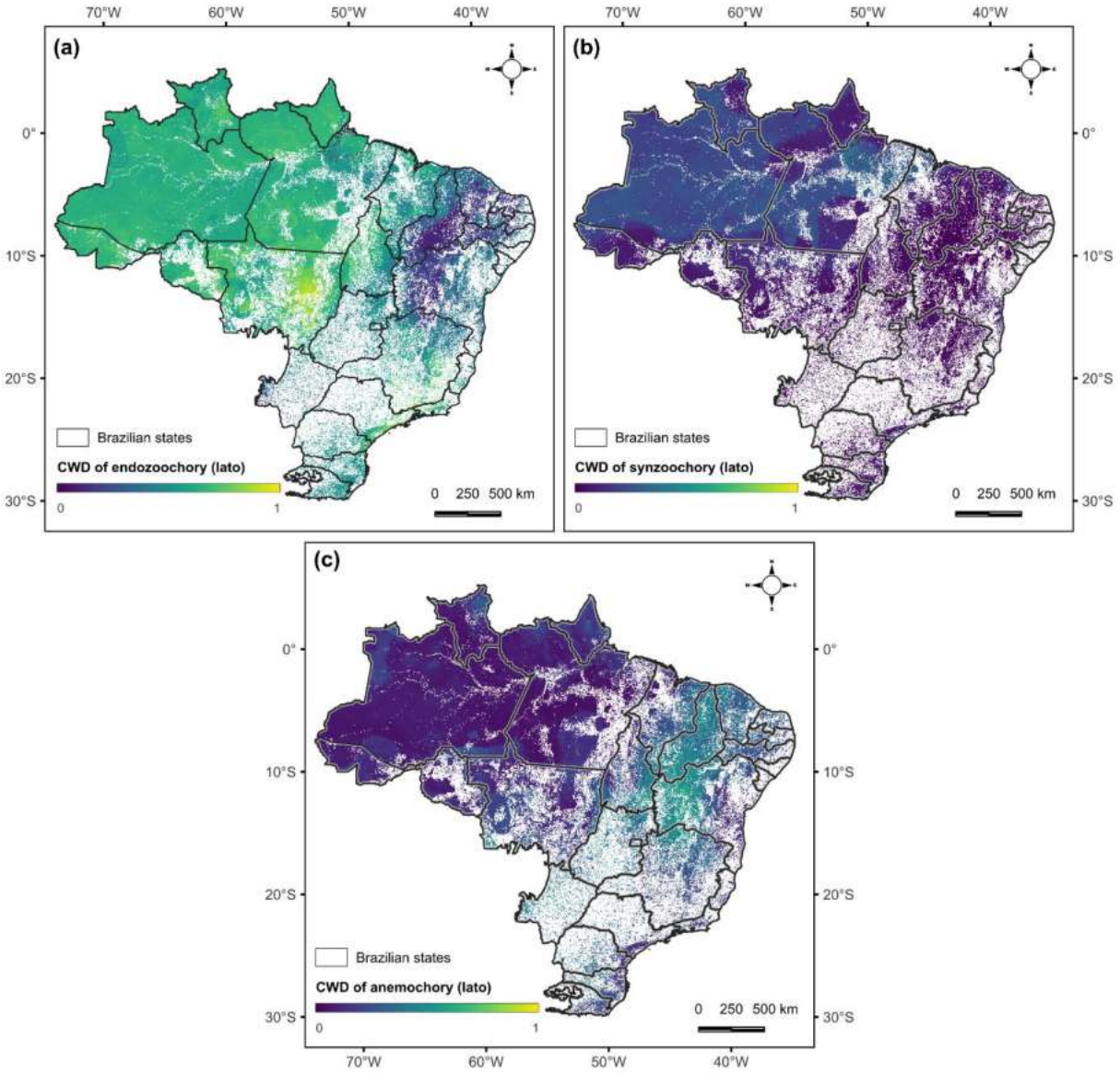
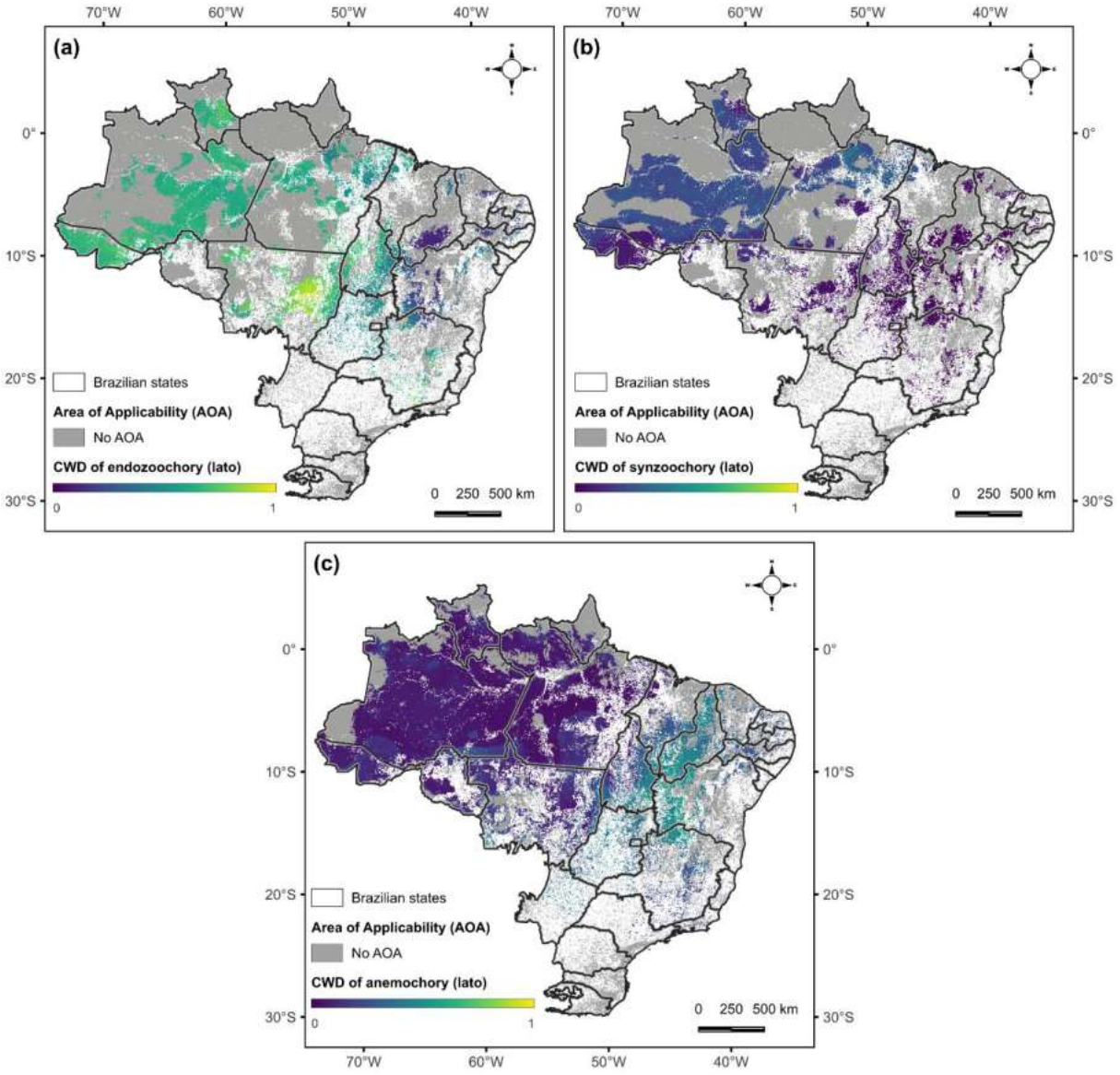


Figure S16: Partial dependence of community weighted-mean (CWM) of synzoochory, including both *sensu stricto* and *sensu lato* classifications, on each explanatory variable. We draw a partial dependence plot showing the mean marginal influence of climate, dispersal agents and soil variables based on the CWM values while keeping all other variables constant at their sample mean.



176

177 **Figure S17**: Prediction map of community-weighted means (CWM) of endozoochory (a),
178 synzoochory (b), and anemochory (c), including both *sensu stricto* and *sensu lato*
179 classifications, across Brazilian ecosystems. Cool colours (blue and green) represent low
180 CWM values, while warmer colours (yellow and red) represent high CWM values.



181

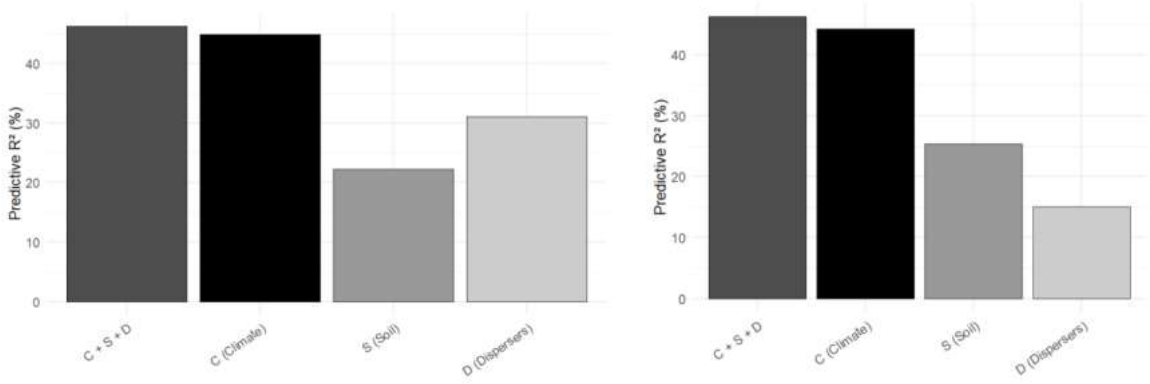
182 **Figure S18**: Area of Applicability (AOA) of the spatial prediction model for endozoochory (a),
183 synzoochory (b), and anemochory (c), including both *sensu stricto* and *sensu lato* classifications,
184 across Brazilian ecosystems. Areas shown in grey fall outside the AOA (AOA = 0) and were
185 excluded from predictions. Colours represent the spatial predictions of the models based on
186 community-weighted mean (CWM) values, with cooler tones (blue and green) indicating lower
187 values and warmer tones (yellow and red) indicating higher values.

Variance partitioning

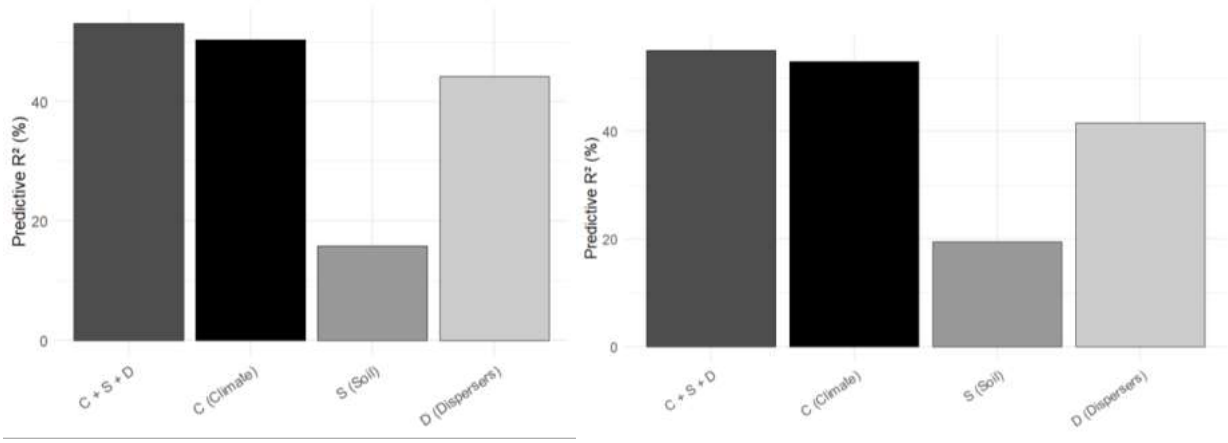
We performed a variation partitioning analysis based on Random Forest models to quantify the predictive contributions of climate (C), soil (S), and disperser-related (D) predictors. Models were fitted using each group separately (C, S, and D), as well as using all predictors combined (C + S + D). Predictive performance was assessed using a stratified 5-fold cross-validation scheme. In each fold, Random Forest models were trained using the final hyperparameters selected for the main analysis and evaluated on held-out data. Out-of-fold predictions were pooled across folds, and predictive R^2 was calculated as $1 - \mathrm{SSE}/\mathrm{SST}$, where SSE is the sum of squared prediction errors and SST is the total sum of squares of the observed response variable.

Predictor set	Endozoochory	Anemochory	Synzoochory	Seed mass
C (Climate)	0.45	0.44	0.50	0.53
S (Soil)	0.22	0.25	0.16	0.19
D (Dispersers)	0.31	0.15	0.44	0.42
C + S + D	0.46	0.46	0.53	0.55

(a) (b)



201 (c) (d)



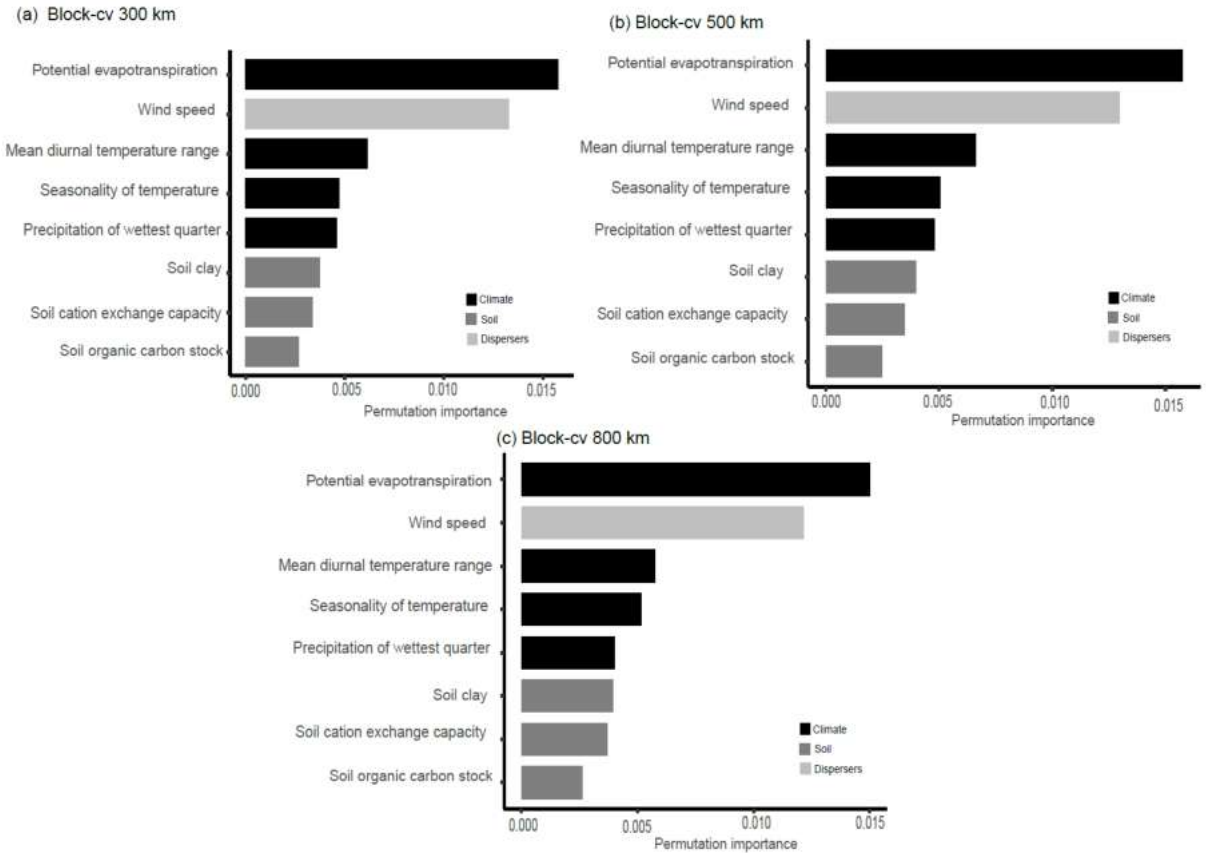
202

203

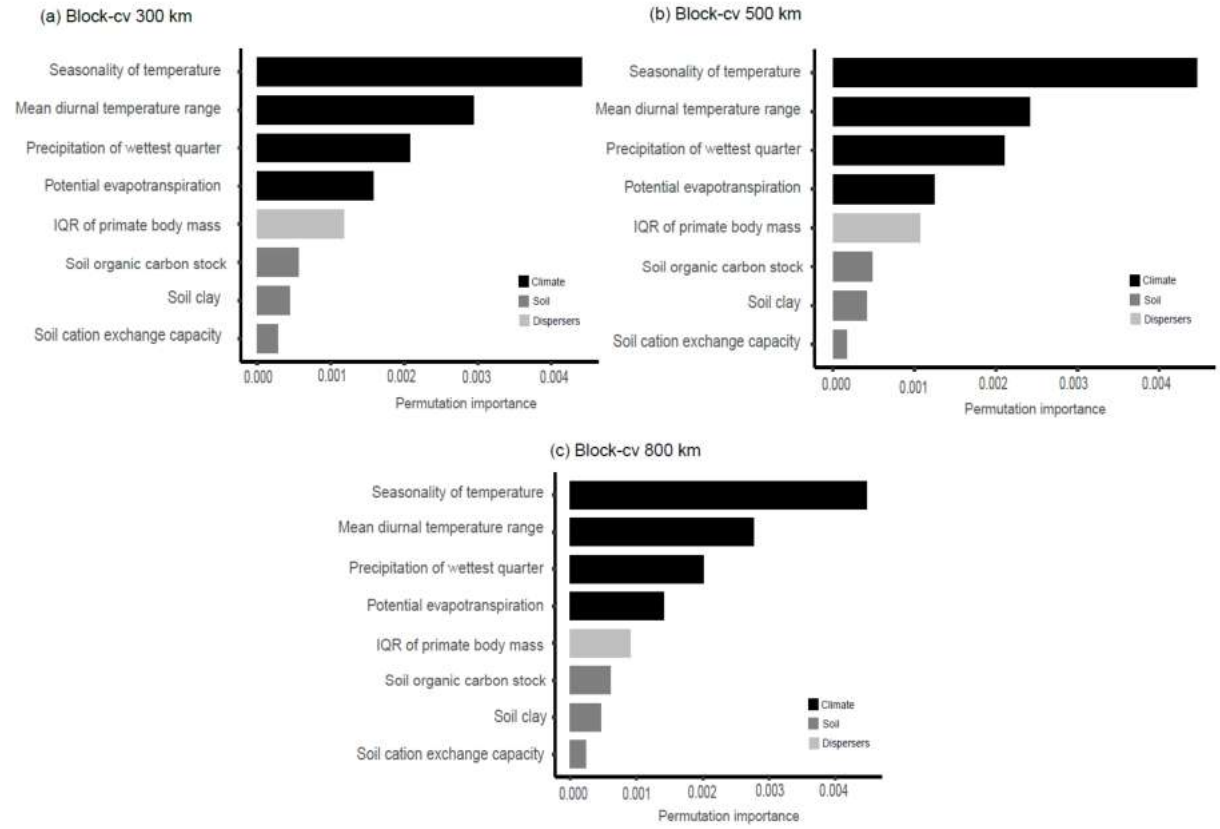
Relative contribution of each predictor under spatial block cross-validation

Variable importance was quantified using permutation importance from Random Forest models evaluated using spatially independent 5-fold cross-validation across three spatial scales: 300 km (a), 500 km (b), and 800 km (c) blocks. Importance values represent the average decrease in model performance after permuting each predictor across spatial folds. Predictors are grouped into climate (black), soil (dark grey), and disperser-related variables (light grey). The ranking of predictors is shown independently for each spatial scale.

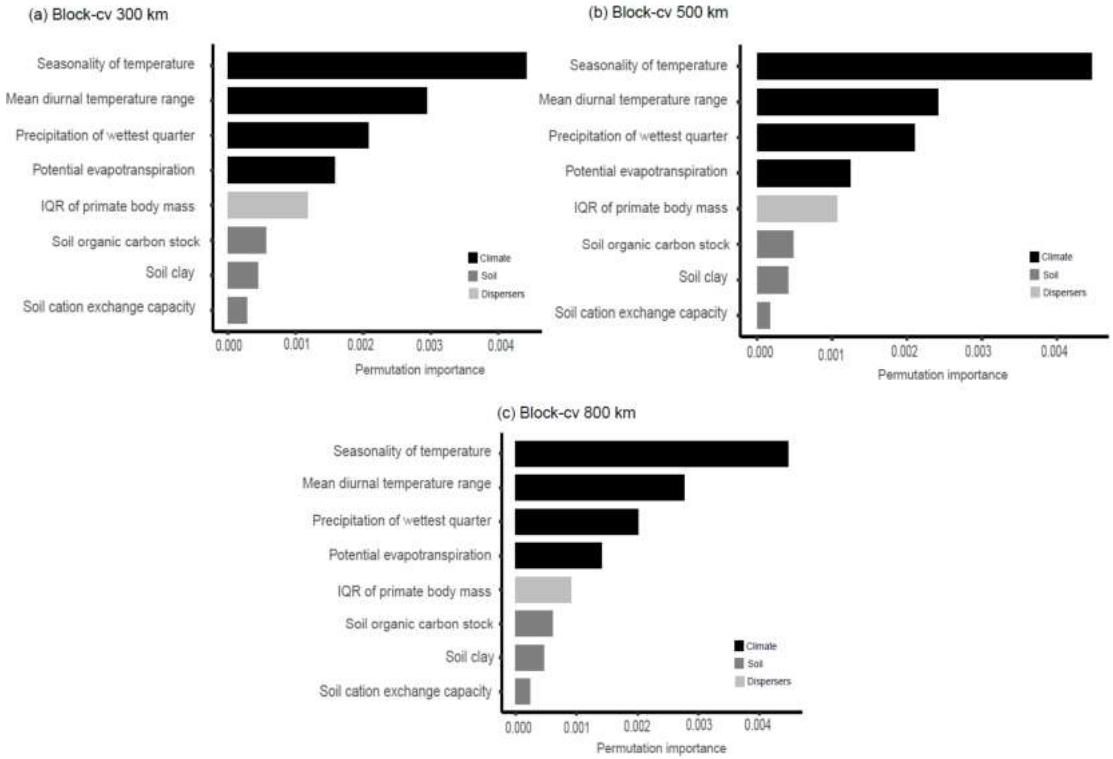
I) ENDOZOOCHORY



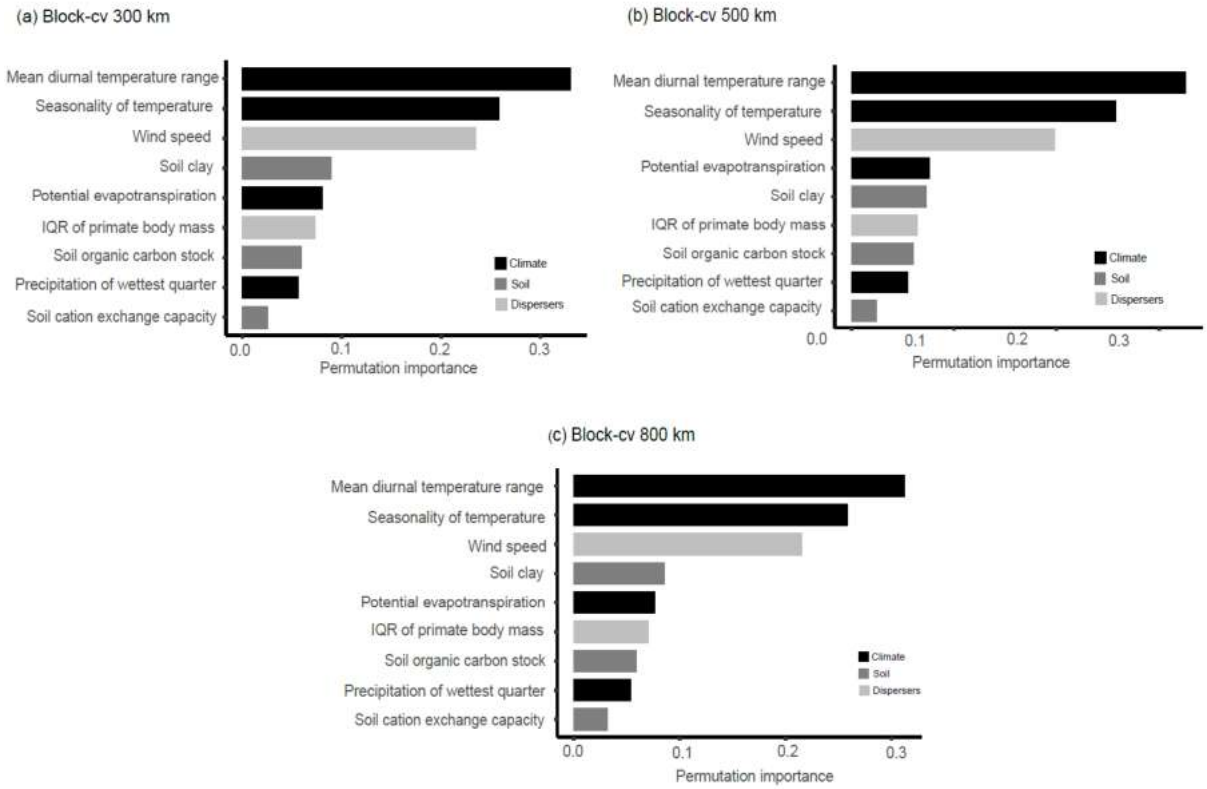
ANEMOCHORY



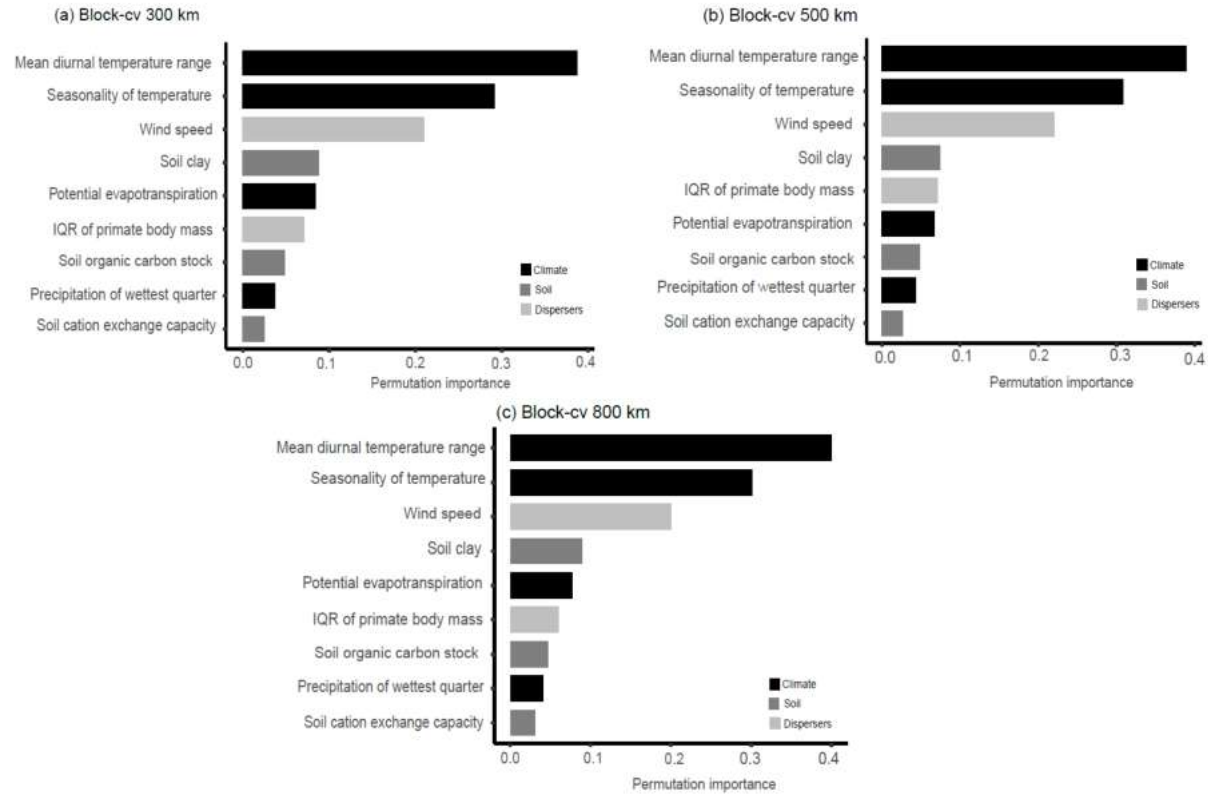
II) SYNZOOCHORY



III) SEED MASS (complete +imputed data)



IV) SEED MASS (only complete data)



References

Hawes, J.E., & Peres, C.A. (2014). Ecological correlates of trophic status and frugivory in neotropical primates. *Oikos*, *123*, 365–377.

Fuzessy, L., Silveira, F.A.O., Culot, L., Jordano, P., & Verdu, M. (2021). Phylogenetic congruence between Neotropical primates and plants is driven by frugivory. *Ecology Letters*, *25*, 320–325.