





**PREPRINT**

*Author-formatted, not peer-reviewed document posted on 28/01/2026*

DOI: <https://doi.org/10.3897/arphapreprints.e186656>

---

# **Restricted occupancy and environmental specialization of *Ragala ucuquirana-branca* (Sapotaceae) across the Amazon Basin**

 Kaio Cesar da Cunha,  Laura Lorena Rivera-Parada, Itamara Gonçalves da Gama, Glaucileide Ferreira,  Semirian Campos Amoêdo,  Caroline da Cruz Vasconcelos

1 **Title page**

2

3 Running title – Restricted occupancy and specialization in an Amazonian tree

4

5 **Restricted occupancy and environmental specialization of *Ragala ucuquirana-branca***  
6 **(Sapotaceae) across the Amazon Basin**

7

8 Kaio C. M. da Cunha<sup>1</sup>, Laura L. Rivera-Parada<sup>2</sup>, Itamara G. da Gama<sup>3</sup>, Glaucileide Ferreira<sup>4</sup>,  
9 Semirian C. Amoêdo<sup>5</sup>, Caroline C. Vasconcelos<sup>1</sup>

10

11 1 Herbário, Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus, AM, Brazil

12 2 Giunti Psychometrics, Madrid, Spain

13 3 Centro de Sementes Nativas do Amazonas, Universidade Federal do Amazonas (UFAM),  
14 Manaus, AM, Brazil

15 4 Programa de Pós-Graduação em Ciências Florestais, Universidade Federal do Espírito Santo  
16 (UFES), Jerônimo Monteiro, ES, Brazil

17 5 Ação Ecológica Guaporé (ECOPORE), Porto Velho, RO, Brazil

18

19 Corresponding author: Kaio C. M. da Cunha ([kaiocesar2013.kc19@gmail.com](mailto:kaiocesar2013.kc19@gmail.com))

20

21 **Abstract**

22

23 **Background and aims** – Understanding the distribution of Amazonian tree species is  
24 hindered by sparse occurrence data and strong environmental heterogeneity. Species with  
25 wide geographic ranges but restricted habitat occupancy challenge climate-only explanations  
26 of distribution patterns. Here, we investigate the environmental drivers shaping the  
27 distribution of *Ragala ucuquirana-branca* (Sapotaceae, Chrysophylloideae), a canopy tree  
28 native to the Amazon Basin, to disentangle the roles of environmental filters in constraining  
29 its realized distribution.

30 **Material and methods** – We compiled and spatially validated 72 occurrence records from  
31 verified herbarium specimens and long-term forest monitoring plots collected between 1941  
32 and 2025 across the Amazon Basin. Species distribution models were developed using an  
33 ensemble framework that combined multiple algorithms and integrated climatic, atmospheric,  
34 edaphic, topographic, and land cover variables. Model performance was evaluated using  
35 threshold-independent and threshold-dependent accuracy metrics, and habitat suitability was  
36 projected across the Amazon Basin.

37 **Key results** – Model performance was consistently high, indicating robust discrimination  
38 between suitable and unsuitable environments. Soil-related predictors collectively accounted  
39 for the largest share of model importance, exceeding climatic and other environmental  
40 variables. Soil pH in water (Phh2o) emerged as the most influential predictor of distribution,  
41 followed by atmospheric water vapor pressure (Vapr), minimum temperature of the coldest  
42 month (Bio6), and the cation exchange capacity (Cec). Although the species exhibits a broad  
43 Extent of Occurrence of approximately 1.48 million km<sup>2</sup>, its Area of Occupancy is highly  
44 restricted, revealing a spatially aggregated distribution. Suitable habitats are concentrated in  
45 the central and western Amazonia, with strong fragmentation toward the eastern and southern  
46 margins of the basin.

47 **Conclusion** – Our results demonstrate that the distribution of *R. ucuquirana-branca* is  
48 strongly constrained by edaphic conditions and atmospheric stability rather than by broad  
49 climatic gradients alone. This decoupling between geographic range size and effective habitat  
50 occupancy highlights pronounced ecological specialization and potential vulnerability. More

51 broadly, the study underscores the importance of incorporating soil and atmospheric filters  
 52 into assessments of species rarity, distribution, and conservation priorities in hyperdiverse  
 53 tropical forests.

54  
 55 **Keywords**

56 Amazonian tree species, Chrysophylloideae, conservation, edaphic specialization, ensemble  
 57 forecasting, spatial modeling, species distribution modeling, *terra-firme* forests

58  
 59 **INTRODUCTION**

60  
 61 The Amazon is the largest drainage basin and harbors one of the most complex and diverse  
 62 tree floras on Earth (Guayasamin et al. 2024), with the Sapotaceae family standing out as a  
 63 dominant and ecologically significant component of these forests (ter Steege et al. 2006,  
 64 2013). Within this family, *Ragala* Pierre (Chrysophylloideae subfamily) is a small genus that  
 65 comprises about four species and three subspecies concentrated in the Amazon Basin, but  
 66 extending to the coastal areas of the Guianas, occurring in lowland *terra-firme* forests (non-  
 67 flooded forests) on clayish and sandy soils (Pennington 1990; van Roosmalen et al. 2000;  
 68 Swenson et al. 2023).

69 Species currently placed in *Ragala* have historically been assigned to *Ecclinusa* Mart.,  
 70 and both genera were at times subsumed within *Chrysophyllum* L. sensu lato (Pennington  
 71 1990, 1991). *Ragala* was recently reinstated through phylogenetic analyses based on  
 72 molecular and morphological data (Swenson et al. 2023). These analyses supported the idea  
 73 that *Ragala* is a distinct genus from *Ecclinusa* and *Chrysophyllum* sensu lato, forming a  
 74 cohesive lineage sister to *Ecclinusa*, *Elaeoluma* Baill., and *Nemaluma* Baill. (Faria et al.  
 75 2017; Swenson et al. 2023).

76 Diagnostic characters for *Ragala* include oblique tertiary venation with areolate  
 77 higher-order venation, pentamerous flowers and fruits with an accrescent calyx—a key  
 78 feature identified since Pierre (1891)—a glabrous corolla, and the absence of staminodes  
 79 (Faria et al. 2017; Swenson et al. 2023). Furthermore, wood anatomy studies confirm that  
 80 *Ragala* possesses distinct structural patterns that separate it from both *Ecclinusa* and  
 81 *Chrysophyllum* sensu lato (Kukachka 1981).

82 A notorious member of this genus is *Ragala ucuquirana-branca* (Aubrév. & Pellegr.)  
 83 W.A.Rodrigues. This canopy tree, which can reach up to 30 m in height and 45 cm in  
 84 diameter, is characterized by dark brown scaling bark and a reddish-brown slash with  
 85 abundant sticky milky latex in the field (Pennington 1990, 1991). Taxonomically, it is  
 86 distinguished from its close relatives, *R. sanguinolenta* Pierre and *R. scalaris* (T.D.Penn.)  
 87 Swenson, by its densely ferruginous-tomentose indumentum of short, curly trichomes on the  
 88 young parts and lower leaf surfaces, as well as its villous reddish fruit (Pennington 1990,  
 89 1991). Despite its current well-established taxonomic identity, detailed knowledge regarding  
 90 the environmental drivers of this species remains necessary to fully understand its ecological  
 91 niche across its documented range in central and northern Amazonian Brazil and southern  
 92 Venezuela (Pennington 1990).

93 Species Distribution Models (SDMs) have become indispensable tools for evaluating  
 94 how broad-scale environmental shifts impact biological diversity (Hao et al. 2019; Zurell et  
 95 al. 2020). These correlative frameworks function by integrating georeferenced occurrence  
 96 data with spatially explicit environmental layers, most notably climatic variables (Miller  
 97 2010). By quantifying the relationship between a taxon and its habitat, SDMs allow  
 98 researchers to delineate geographical ranges and forecast shifts in distribution across space  
 99 and time (Guisan and Zimmermann 2000; Miller 2010), being useful for supporting policy  
 100 decisions in biodiversity conservation (Araújo and New 2007; Guisan et al. 2013).

101           Given the increasing pressure on Amazonian ecosystems (Flores et al. 2024), there is  
102 an urgent need for precise spatial data to guide management strategies for species  
103 conservation. Therefore, this study uses ensemble species distribution modeling to: (1)  
104 compile and validate available occurrence records of the *R. ucuquirana-branca*, (2) quantify  
105 the primary environmental factors shaping its potential distribution, and (3) produce spatially  
106 explicit predictions of its suitable habitat across the Amazon Basin.

107

## 108 **MATERIAL AND METHODS**

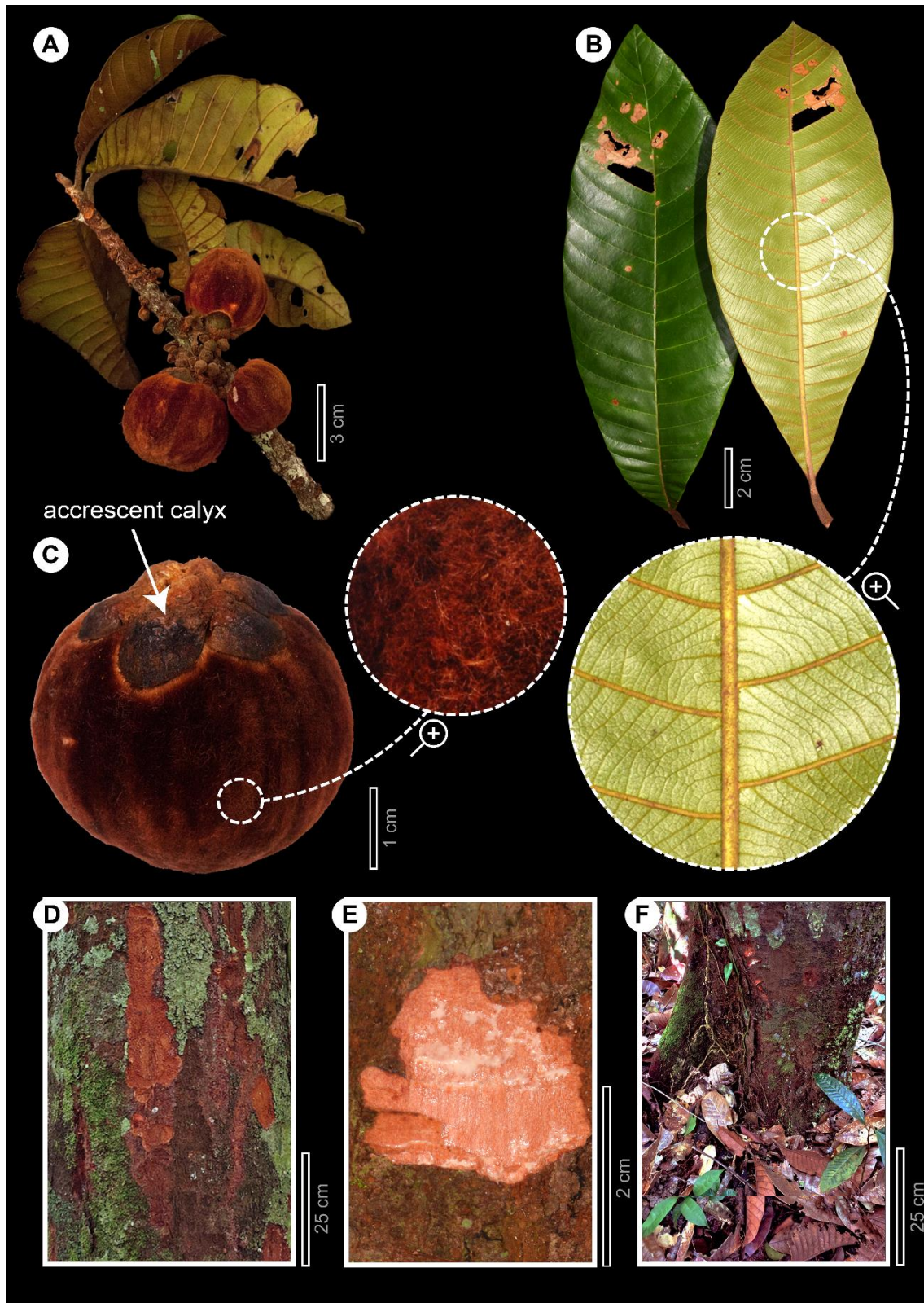
109

### 110 **Target taxon**

111

112 *Ragala ucuquirana-branca* (Fig. 1) is a canopy tree native to the Amazon Basin, typically  
113 occurring in lowland *terra-firme* forests. The species is locally known by several vernacular  
114 names, such as “ucuquirana-brava-da-folha-vermelha” (Rodrigues 1974), “balata-brava”,  
115 “coquirana-branca”, and “coquirana-rocha” (Pennington 1990; van Roosmalen and Garcia  
116 2000). Its distribution is primarily documented in central and northern Amazonia, extending  
117 into southern Venezuela (Pennington 1990). Given its restricted distribution, this species  
118 serves as an ideal model for investigating how environmental variables shape the distribution  
119 of high-diversity tropical tree families.

120



121  
122 **Figure 1.** Field images of *Ragala ucuquirana-branca* (Sapotaceae). **A.** Fruiting branch. **B.**  
123 View of adaxial (upper) and abaxial (lower) leaf surface, with a close-up of the venation  
124 pattern. **C.** Mature fruit with a close-up of the reddish indument. **D.** Bark  
125 slash showing the milky latex. **E.** Bark slash showing the milky latex. **F.** Trunk base with small buttresses. Photographs: **A, C** by  
126 Isolde Ferraz; **B, D, E** by Francisco Farroñay; and **F** by José Edmilson Souza.  
127

128 **Species occurrence records**

129  
 130 We compiled all available occurrence (presence) records from two web-based biodiversity  
 131 repositories that integrate data from multiple herbaria worldwide: the Global Biodiversity  
 132 Information Facility (GBIF) (<https://www.gbif.org/>; GBIF.org 2024) and the *speciesLink*  
 133 network (<http://www.splink.org.br/>; Canhos et al. 2022). In addition, we incorporated records  
 134 from physical reference collections associated with long-term forest monitoring plots in  
 135 Central Amazonia, including the *Projeto Dinâmica Biológica de Fragmentos Florestais*  
 136 (PDBFF) and *Programa de Pesquisa em Biodiversidade* (PPBio). In all cases, only records of  
 137 preserved specimens were retained. During an initial data screening, we excluded specimens  
 138 from other genera through taxonomic verification of digital and/or physical specimens, as  
 139 well as duplicate records (i.e., specimens with the same collector and collection number, even  
 140 when deposited in different collections).

141 To minimize georeferencing errors, original coordinates (latitude and longitude in  
 142 decimal degrees) were manually verified to ensure consistency with the corresponding  
 143 locality descriptions. When coordinates were missing, they were estimated based on detailed  
 144 locality descriptions, allowing for reliable georeferencing using the Google Earth platform  
 145 (<https://earth.google.com/web/>). We also flagged and removed potentially erroneous records  
 146 by checking proximity to well-known geographic features (e.g., centroids for administrative  
 147 regions and institutions), identifying spatial outliers, and finding suspicious values or  
 148 formatting problems using the ‘CoordinateCleaner’ package (Zizka et al. 2019).

149 In addition to preparing the occurrence dataset for the SDM workflow, we quantified  
 150 the Extent of Occurrence (EOO) and Area of Occupancy (AOO) using the GEOCAT platform  
 151 (Bachman et al. 2011). These metrics summarize, respectively, the overall geographic spread  
 152 and the minimum occupied area represented by the verified records.

153  
 154 **Study region**

155  
 156 The species’ geographic range was cropped to the extent of the Amazon Basin (sensu Eva et  
 157 al. 2005), which includes areas of the Cerrado biome and montane regions that drain into the  
 158 Amazon River. This delimitation of the Amazon Basin (Amazonia sensu latissimo) covers  
 159 approximately 7,595,000 km<sup>2</sup> and includes areas of Brazil, Bolivia, Peru, Ecuador, Colombia,  
 160 Venezuela, Guyana, Suriname, and French Guiana. All subsequent modeling steps  
 161 (background selection, pseudo-absence generation, variable extraction, and ensemble  
 162 forecasting) were performed within this geographic extent to ensure ecological coherence and  
 163 minimize spatial bias in model performance and properly encompass the full known  
 164 distributional range of *Ragala ucuquirana-branca*.

165  
 166 **Environmental data**

167  
 168 Environmental variables were selected to capture the environmental gradients driving the  
 169 distribution of *Ragala ucuquirana-branca* in the Amazon Basin. The dataset included 37  
 170 predictors related to climate, topography, soil, land cover, and habitat heterogeneity (Table 1)  
 171 that are expected to influence plant physiology, morphology, distribution, and ecology  
 172 (Toledo et al. 2012; Stein et al. 2014; Lambers and Oliveira 2019).

173  
 174 **Table 1.** Variables selected for modeling the current potential distribution of *Ragala*  
 175 *ucuquirana-branca* (Sapotaceae) in the Amazon Basin. Variables 1–19 are related to climate,  
 176 20–22 to atmosphere; 23 to topography; 24–33 to soil; 34–36 to land cover; and 37 to habitat  
 177 heterogeneity. Data sources: WC – WorldClim 2.1; SG – SoilGrids; and EE – EarthEnv.

N	Label	Description	Unit	Source
1	Bio1	Annual mean temperature	°C	WC
2	Bio2	Mean diurnal range (mean of monthly (max temp - min temp))	°C	WC
3	Bio3	Isothermality (Bio2/Bio7) ( $\times 100$ )	%	WC
4	Bio4	Temperature seasonality (standard deviation $\times 100$ )	°C	WC
5	Bio5	Maximum temperature of warmest month	°C	WC
6	Bio6	Minimum temperature of coldest month	°C	WC
7	Bio7	Temperature annual range (Bio5–Bio6)	°C	WC
8	Bio8	Mean temperature of wettest quarter	°C	WC
9	Bio9	Mean temperature of driest quarter	°C	WC
10	Bio10	Mean temperature of warmest quarter	°C	WC
11	Bio11	Mean temperature of coldest quarter	°C	WC
12	Bio12	Annual precipitation	mm	WC
13	Bio13	Precipitation of wettest month	mm	WC
14	Bio14	Precipitation of driest month	mm	WC
15	Bio15	Precipitation seasonality (coefficient of variation)	fraction	WC
16	Bio16	Precipitation of wettest quarter	mm	WC
17	Bio17	Precipitation of driest quarter	mm	WC
18	Bio18	Precipitation of warmest quarter	mm	WC
19	Bio19	Precipitation of coldest quarter	mm	WC
20	Srad	Annual mean incident solar radiation	$\text{kJ m}^{-2} \text{day}^{-1}$	WC
21	Wind	Annual mean wind speed (2 m above the ground)	$\text{m s}^{-1}$	WC
22	Vapr	Annual mean water vapor pressure	kPa	WC
23	Elev	Elevation above sea level (from SRTM)	m	WC
24	Bdod	Bulk density of the fine earth fraction	$\text{kg dm}^{-3}$	SG
25	Cec	Cation exchange capacity of the soil	$\text{cmol}(+) \text{kg}^{-1}$	SG
26	Cfvo	Volume fraction of coarse fragments ( $> 2 \text{ mm}$ )	%	SG
27	Nitrogen	Total nitrogen (N)	$\text{g kg}^{-1}$	SG

28	Phh2o	pH in water (H <sub>2</sub> O)	-	SG
29	Sand	Sand (> 0.05 mm) in fine earth	%	SG
30	Silt	Silt (0.002–0.05 mm) in fine earth	%	SG
31	Clay	Clay (< 0.002 mm) in fine earth	%	SG
32	Soc	Soil organic carbon in fine earth	g kg <sup>-1</sup>	SG
33	Ocd	Organic carbon density	kg m <sup>-3</sup>	SG
34	Land2	Consensus land cover - Evergreen broadleaf trees (Class 2)	%	EE
35	Land5	Consensus land cover - Shrubs (Class 5)	%	EE
36	Land8	Consensus land cover - Regularly flooded vegetation (Class 8)	%	EE
37	Shannon	Diversity of Enhanced Vegetation Index (EVI)	-	EE

178

179

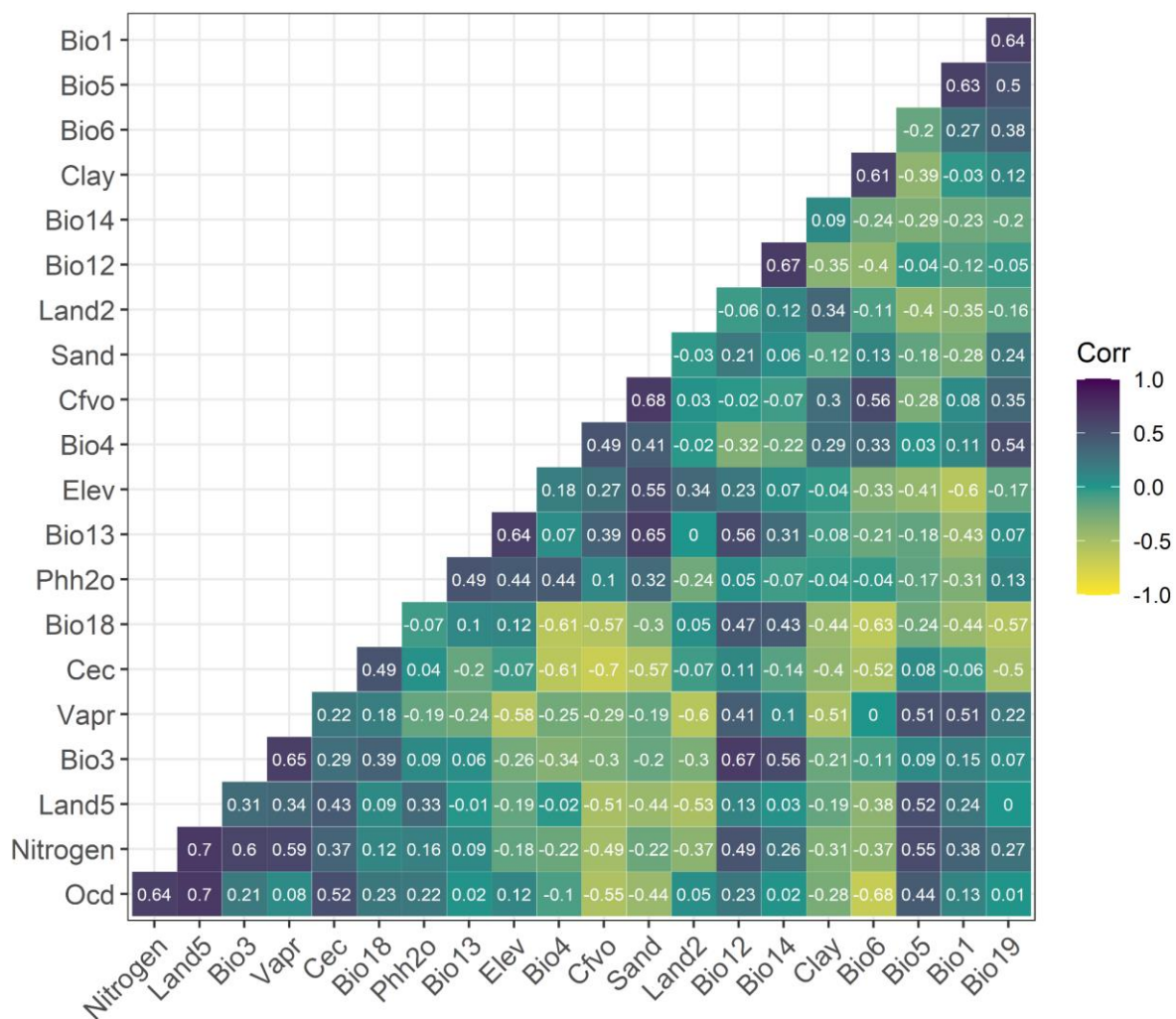
180 The near-global standardized environmental layers were downloaded as GeoTIFF  
 181 raster files from web-based repositories such as WorldClim 2.1 (<https://www.worldclim.org/>;  
 182 Fick and Hijmans 2017), SoilGrids (<https://isric.org/explore/soilgrids>; Poggio et al. 2021), and  
 183 EarthEnv (<https://www.earthenv.org/>; Tuanmu and Jetz 2014, 2015) at 30 seconds (~1 km)  
 184 spatial resolution and under historical (near current) environmental conditions (Table 1). Soil  
 185 data were downloaded using the ‘geodata’ package (Hijmans et al. 2024a) while other  
 186 variables were directly from the aforementioned web-based repositories.

186

187

188 To reduce multicollinearity and prevent model overfitting, we selected a subset of  
 189 weakly correlated bioclimatic predictors (Dormann et al. 2013). Automated variable selection  
 190 was performed using the *select07()* function implemented in the ‘mecofun’ package (Zurell  
 191 2024), which identifies correlated pairs (Spearman’s rank correlation coefficient  $|r| > 0.7$ )  
 192 based on environmental values (i.e., the full set of 37 variables) extracted at species  
 193 occurrence points. Within each correlated pair, we retained the variable with the highest  
 194 explanatory power, assessed by its univariate importance in a quadratic generalized linear  
 195 model (GLM) ranked by Akaike Information Criterion (AIC). This procedure resulted in a  
 final set of 21 variables, which were used in all subsequent analyses (Fig. 2).

195



196  
 197 **Figure 2.** Correlogram plot showing pairwise Spearman’s rank correlation between variables  
 198 selected as predictors ( $|r| \leq 0.7$ ) for further analysis. Darker colors indicate stronger  
 199 correlations. Dark purple indicates positive and yellow indicates negative correlations.

200  
 201 **Ensemble modeling**

202  
 203 We modeled species’ potential distribution using five algorithms: (1) Boosted Regression  
 204 Trees (Elith et al. 2008); (2) Classification and Regression Trees (Breiman et al. 1984; De’ath  
 205 and Fabricius 2000); (3) Generalized Additive Model (Hastie and Tibshirani 1986; Yee and  
 206 Mitchell 1991); (4) Maximum Entropy (Phillips et al. 2006); and (5) Random Forest (Breiman  
 207 2001) (Table 2). We used the ‘dismo’ package (Hijmans et al. 2024b) to perform Boosted  
 208 Regression Trees and Maximum Entropy; ‘rpart’ (Therneau and Atkinson 2023) to perform  
 209 Classification and Regression Trees; ‘mgcv’ (Wood 2011) to perform Generalized Additive  
 210 Model; and ‘randomForest’ (Liaw and Wiener 2002) for the Random Forest models. We used  
 211 an ensemble forecasting approach (i.e., combining multiple models for better predictions) to  
 212 minimize errors (Araújo and New 2007).

213  
 214 **Table 2.** Summary of the algorithms used in the ensemble species distribution model (SDM).  
 215 Statistical approach: ML – Machine Learning; R – Regression.

Label	Algorithm	Data type	Approach
BRT	Boosted Regression Trees	Pseudo-absence	ML
CART	Classification and Regression Trees	Pseudo-absence	ML
GAM	Generalized Additive Model	Pseudo-absence	R
MaxEnt	Maximum Entropy	Presence-only (background sample)	ML
RF	Random Forest	Pseudo-absence	ML

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

We sampled pseudo-absence points across the Amazon Basin. We matched and randomly generated pseudo-absence points, 10 times the number of occurrence points, for all algorithms (Barbet-Massin et al. 2012). Moreover, we randomly split the combined dataset (comprising occurrence records and pseudo-absence data) so that 70% was used for model training and the remaining 30% for testing.

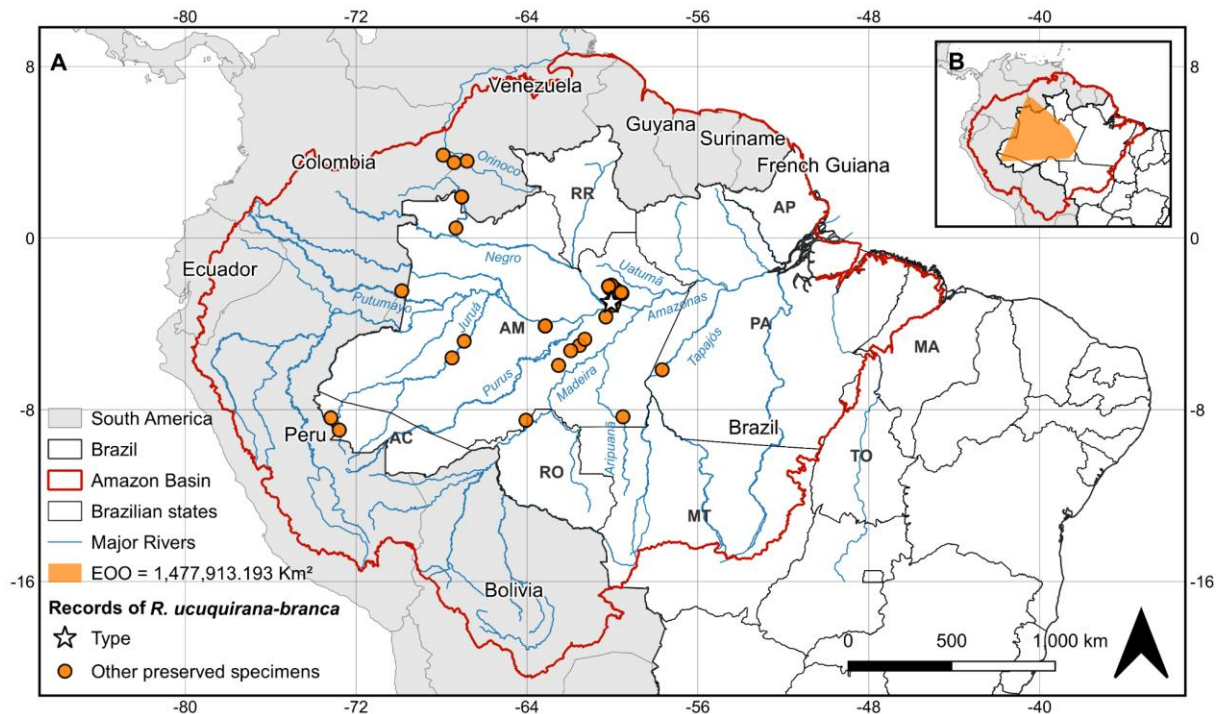
We evaluated the model’s performance by calculating the True Skill Statistic (TSS; [(sensitivity + specificity) - 1]) (Allouche et al. 2006), and the Area Under the Curve (AUC) of receiver operator characteristics (ROC; Fielding and Bell 1997). We generated 10 replications for each algorithm and selected the replicates presenting accuracy values of AUC  $\geq 0.75$  and TSS  $\geq 0.55$  to construct the ensemble rasters. Then, we ensembled the models based on the average of the selected rasters weighted by AUC values  $\geq 0.75$  (Hao et al. 2020).

To create binary distribution (presence-absence) of the predicted habitat suitability, we applied two threshold-selection methods: minimum (or lowest) training presence (MPT), which consists of the lowest predicted suitability value for an occurrence point; and 10th percentile training presence (P10), which omits all regions with habitat suitability lower than the suitability values for the lowest 10% of occurrence records (Morrow 2019).

All SDM workflow was performed in the open-source statistical software R v.4.4.2 (R Core Team 2024) using custom-made R scripts and add-on libraries. Final maps were prepared using QGIS v3.28.1 (QGIS Development Team 2022).

## RESULTS

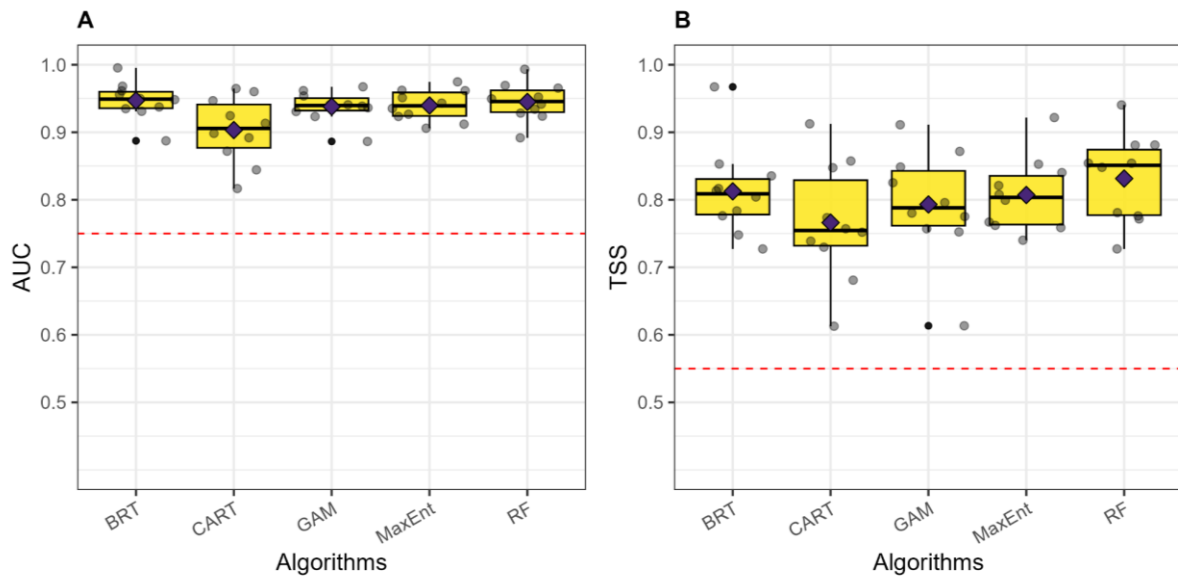
After data cleaning and spatial verification, a total of 72 unique occurrence records of *Ragala ucuquirana-branca* (Fig. 1) were retained for subsequent analyses (Fig. 3A). These records span collections made between 1941 and 2025 and represent the most comprehensive and spatially vetted dataset currently available for the species. Occurrences are unevenly distributed across the Amazon Basin, with a pronounced concentration in central Amazonia, particularly in lowland *terra-firme* forests of northern Brazil, and a smaller number of records extending into southern Venezuela and southern Colombia (Fig. 3A). In Brazil, records are restricted to the states of Acre, Amazonas, and Pará. No records were confirmed for large portions of the southwestern and southeastern Amazon Basin.



249  
 250 **Figure 3. A.** Known geographic distribution of *Ragala ucuquirana-branca* (Sapotaceae)  
 251 across the Amazon Basin; the red outline represents the basin boundary, blue lines represent  
 252 major rivers, and markers indicate georeferenced records (type represented by a white star and  
 253 other preserved specimens by orange circles). **B.** Inset map showing the Extent of Occurrence  
 254 (EOO) of *Ragala ucuquirana-branca* in the study region.  
 255

256 The spatial configuration of the occurrence records defines a broad Extent of  
 257 Occurrence (EOO) of approximately 1.48 million km<sup>2</sup>, encompassing a substantial portion of  
 258 the Amazon Basin (Fig. 3B). However, this large EOO contrasts sharply with the sparse and  
 259 discontinuous distribution of records within it. Occurrences form geographically isolated  
 260 clusters separated by vast areas without confirmed records, resulting in a fragmented spatial  
 261 pattern. Even within the EOO polygon, large regions lack occurrences, indicating that the  
 262 species' distribution is not continuous across its geographic envelope. The Area of Occupancy  
 263 (AOO) was estimated at 148 km<sup>2</sup>, reflecting the limited number of occupied grid cells within  
 264 the broader EOO. Together, the EOO and AOO maps highlight the spatial configuration of  
 265 known records within the Amazon Basin.

266 Model evaluation revealed consistently high predictive performance across all five  
 267 algorithms included in the ensemble framework (Table 3). Mean Area Under the Curve  
 268 (AUC) values were generally high, exceeding 0.80 for most model replicates (Fig. 4A), while  
 269 True Skill Statistic (TSS) values were predominantly above 0.60 (Fig. 4B), indicating strong  
 270 discriminatory power between suitable and unsuitable conditions.  
 271



272  
 273 **Figure 4.** Evaluation of the algorithm performance in the ensemble species distribution model  
 274 for *Ragala ucuquirana-branca* (Sapotaceae) in the Amazon Basin using (A) AUC and (B)  
 275 TSS metrics. Each black transparent point represents one model replicate; yellow boxes show  
 276 the median and interquartile range, and dark purple diamonds mark the mean value for each  
 277 algorithm. The red dashed line indicates the commonly used threshold for acceptable model  
 278 performance.  
 279

280 Random Forest (RF) and Boosted Regression Trees (BRT) showed the highest median  
 281 values for both metrics and the narrowest interquartile ranges, reflecting high stability among  
 282 replicates. Maximum Entropy also performed robustly, whereas CART and GAM exhibited  
 283 greater variability, particularly in TSS values, although still meeting minimum performance  
 284 criteria (Table 3).  
 285

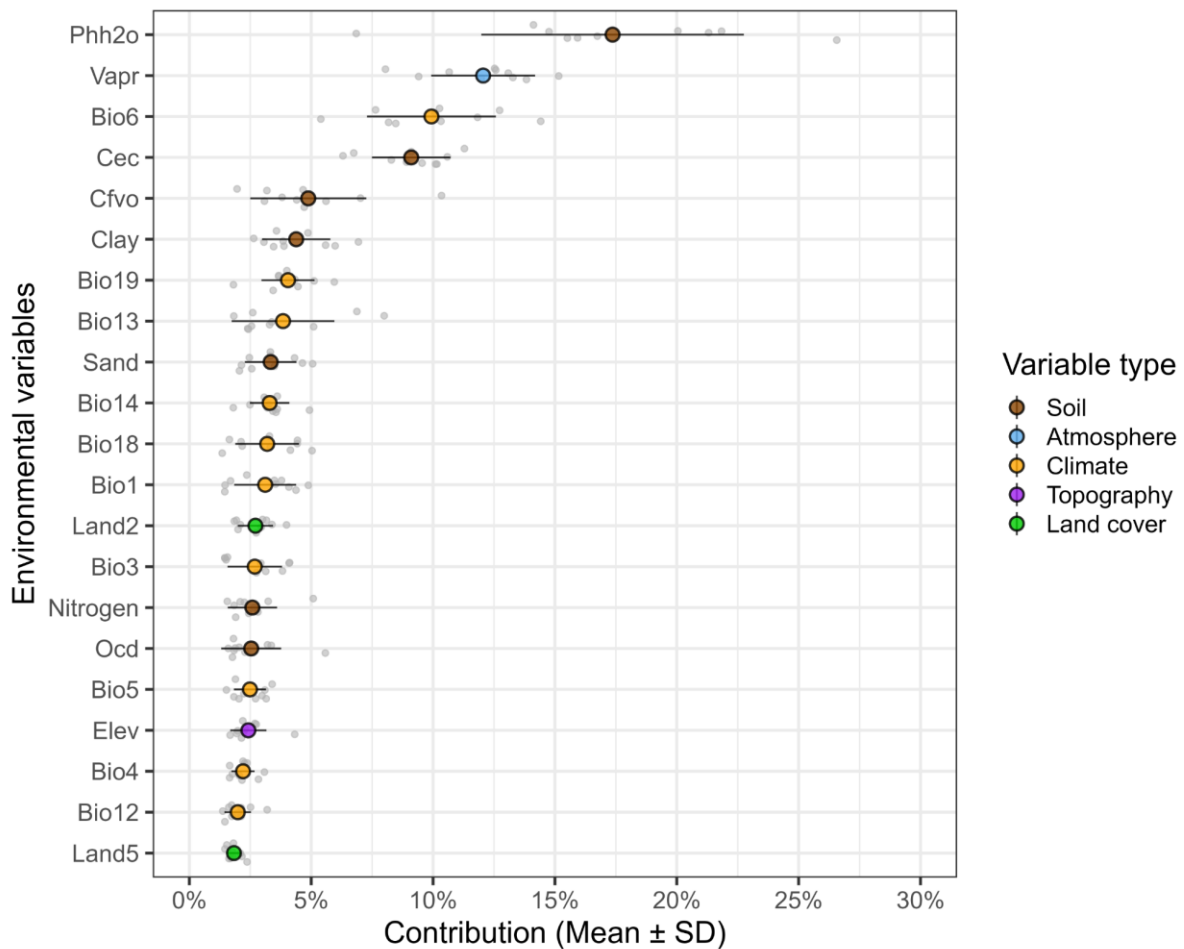
286 **Table 3.** Mean and standard deviation ( $\pm$ ) of the evaluation metrics for each algorithm in the  
 287 ensemble species distribution model for *Ragala ucuquirana-branca* (Sapotaceae) in the  
 288 Amazon Basin.

Algorithm	Area Under the Curve	True Skill Statistic
	Mean $\pm$ SD	Mean $\pm$ SD
Boosted Regression Trees (BRT)	0.95 $\pm$ 0.03	0.81 $\pm$ 0.07
Classification and Regression Trees (CART)	0.90 $\pm$ 0.05	0.77 $\pm$ 0.09
Generalized Additive Model (GAM)	0.94 $\pm$ 0.02	0.79 $\pm$ 0.08
Maximum Entropy (MaxEnt)	0.94 $\pm$ 0.02	0.81 $\pm$ 0.06
Random Forest (RF)	0.95 $\pm$ 0.03	0.83 $\pm$ 0.07

289  
 290 Only model replicates meeting the predefined thresholds (AUC  $\geq$  0.75 and TSS  $\geq$   
 291 0.55) were retained for ensemble construction. This filtering step resulted in a final ensemble  
 292 characterized by strong internal agreement among algorithms, as indicated by overlapping  
 293 distributions of performance metrics (Fig. 4). Weighting ensemble predictions by AUC values  
 294 further reduced the influence of lower-performing models and enhanced the robustness of  
 295 spatial predictions across the Amazon Basin.

296 The relative importance of environmental predictors retained in the ensemble model  
 297 (see Fig. 2 for correlation analysis) shows a markedly uneven pattern (Fig. 5), with a small  
 298 subset of variables explaining a disproportionate share of the total contribution. Overall, soil-

299 related predictors accounted for the largest fraction of model importance (44.2%), followed  
 300 by climate (36.8%), atmosphere (12.1%), land cover (4.5%), and topography (2.4%).  
 301



302 **Figure 5.** Relative contribution of environmental predictors retained in the ensemble species  
 303 distribution model for *Ragala ucuquirana-branca* (Sapotaceae) in the Amazon Basin.  
 304 Predictor importance was weighted by algorithm performance and summarized across model  
 305 replicates (transparent grey circles). Variables are grouped and color-coded according to  
 306 environmental category (soil, atmosphere, climate, topography, and land cover). Higher  
 307 values indicate a greater contribution of each predictor to overall model performance.  
 308  
 309

310 Among individual predictors, soil pH measured in water (Phh2o) emerged as the  
 311 single most influential variable, contributing 17.4% to the ensemble model, and also  
 312 displaying the highest variability across model runs (SD = 5.4%). This was followed by vapor  
 313 pressure (Vapr; 12.1%), minimum temperature of the coldest month (Bio6; 9.9%), and cation  
 314 exchange capacity (Cec; 9.1%). Together, these four predictors accounted for approximately  
 315 48.5% of the total explanatory power of the model, indicating that habitat suitability for *R.*  
 316 *ucuquirana-branca* is primarily structured along a limited number of key environmental  
 317 gradients.

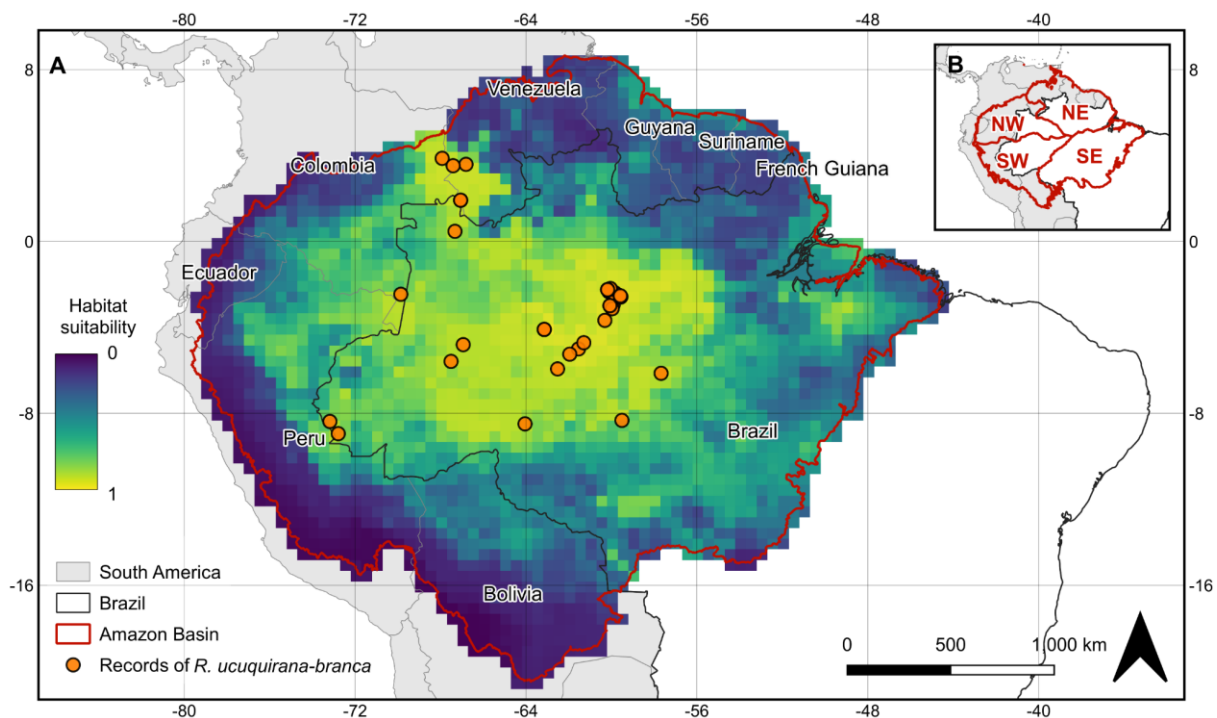
318 Climate variables collectively showed moderate but consistent contributions, with  
 319 Bio6 standing out as the dominant climatic predictor, while other temperature- and  
 320 precipitation-related variables (e.g. Bio19, Bio13, Bio14, Bio18) each contributed between  
 321 ~3–4%. Soil texture and fertility proxies (e.g., clay, sand, Cfvo, nitrogen, and organic carbon  
 322 density) exhibited intermediate importance values, suggesting a secondary but non-negligible  
 323 role in shaping suitability patterns. In contrast, land cover variables and topographic elevation

324 contributed comparatively little to the ensemble model, each individually accounting for less  
 325 than 3% of total importance.

326 Overall, the distribution of predictor contributions highlights a strong dependence of  
 327 the species' modeled distribution on edaphic conditions and atmospheric moisture, modulated  
 328 by temperature extremes, while broad-scale land cover classes and topography appear to play  
 329 a minor role at the spatial resolution considered (Fig. 5).

330 The continuous habitat suitability map derived from the ensemble model revealed a  
 331 highly heterogeneous and spatially restricted distribution of suitable environments across the  
 332 Amazon Basin (Fig. 6A). Areas of high suitability (values > 0.9) are not continuous but  
 333 instead occur as discrete and spatially isolated patches embedded within a largely unsuitable  
 334 matrix. The most extensive and cohesive suitable areas are concentrated in central Amazonia,  
 335 whereas smaller and more fragmented patches are predicted in the northwestern and  
 336 northeastern portions of the basin.

337



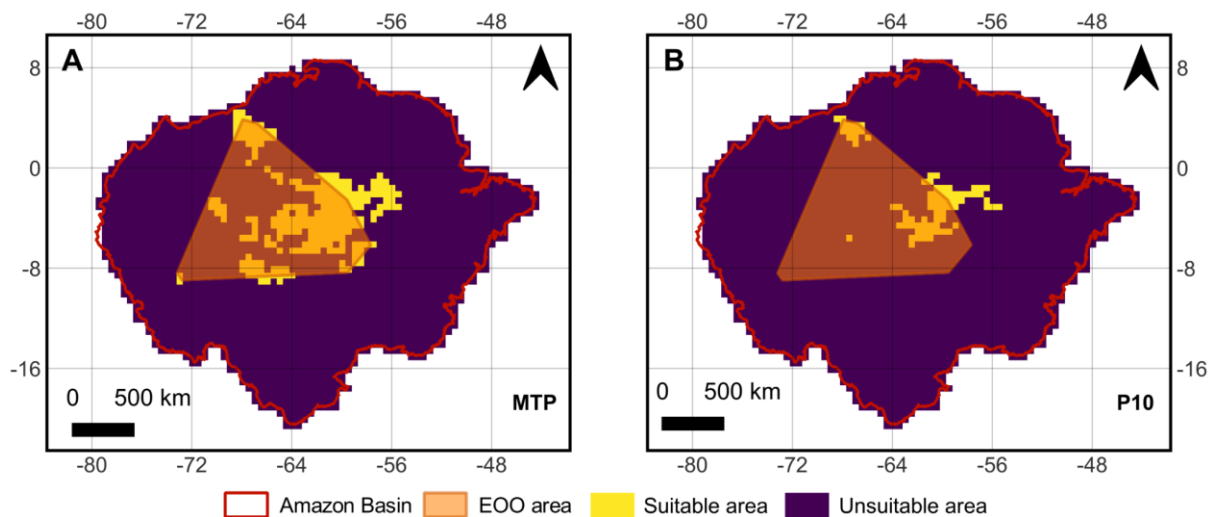
338 **Figure 6. A.** Current habitat suitability map derived from an ensemble model using  
 339 occurrence records of *Ragala ucuquirana-branca* (Sapotaceae) in the Amazon Basin; the red  
 340 outline represents the Amazon Basin, and the color gradient indicates habitat suitability  
 341 values, ranging from highly suitable areas (yellow, > 0.9) to extremely unsuitable areas (dark  
 342 purple, < 0.1). **B.** Inset map showing the geographic subdivisions of the Amazon Basin used  
 343 for regional interpretation (NW = Northwest, SW = Southwest, NE = Northeast, SE =  
 344 Southeast).  
 345

346  
 347 Despite the species' broad extent of occurrence (EOO), large regions of southwestern and  
 348 southeastern Amazonia were consistently predicted as having low or negligible suitability.  
 349 This pronounced mismatch between geographic extent and environmentally suitable area  
 350 underscores the strong environmental structuring of habitat availability across the basin.  
 351 Sharp spatial transitions between suitable and unsuitable areas further reflect the high  
 352 environmental heterogeneity characterizing Amazonian landscapes.

353 Comparison between continuous suitability predictions and observed occurrence  
 354 records indicates that most known records fall within areas predicted as moderately to highly  
 355 suitable, supporting good spatial correspondence between empirical data and model

356 predictions. At the same time, several areas with high predicted suitability lack confirmed  
 357 occurrences, suggesting the presence of potentially suitable but undersampled regions within  
 358 the basin.

359 Binary projections reinforced the restricted potential occupancy of *Ragala*  
 360 *ucuquirana-branca* within the Amazon Basin (Fig. 7). Application of the Minimum Training  
 361 Presence (MTP) threshold resulted in a relatively inclusive prediction, encompassing all  
 362 known occurrences and extending into adjacent areas of marginal suitability, particularly in  
 363 central Amazonia (Fig. 7A).  
 364



365 **Figure 7.** Binary distribution maps of *Ragala ucuquirana-branca* (Sapotaceae) based on two  
 366 threshold criteria (MTP and P10). **A.** Minimum Training Presence (MTP). **B.** 10th Percentile  
 367 Training Presence. Yellow areas indicate predicted presence (suitable), and dark purple areas  
 368 indicate predicted absence (unsuitable) under each threshold criterion. The red outline  
 369 represents the Amazon Basin, and the orange polygon represents the Extent of Occurrence  
 370 (EOO).  
 371  
 372

373 In contrast, application of the 10th Percentile Training Presence (P10) threshold produced  
 374 a markedly more conservative projection, restricting suitable habitat to a limited number of  
 375 core areas characterized by the highest suitability values (Fig. 7B). Under this criterion, large  
 376 portions of the EOO identified as suitable under the MTP threshold were excluded, and the  
 377 total area predicted as suitable represented only a small fraction of the Amazon Basin.  
 378 Notably, both thresholding approaches consistently identified the same core regions, differing  
 379 primarily in the extent rather than the spatial configuration of suitable areas.

380 Taken together, occurrence data, variable contributions, continuous suitability maps, and  
 381 binary projections converge on a consistent empirical pattern: *Ragala ucuquirana-branca*  
 382 exhibits a broad basin-wide geographic range coupled with environmentally constrained,  
 383 spatially fragmented, and geographically restricted habitat occupancy.  
 384

## 385 DISCUSSION

### 387 Occurrence data and known geographic distribution

388  
 389 The compiled records from preserved herbarium specimens and long-term forest monitoring  
 390 plots substantially expand the empirical basis previously available for the species, whose  
 391 distribution had been documented primarily through sparse records in the last taxonomic

392 treatments for the genus (e.g., Pennington 1990, 2006). Large portions of the basin that lack  
 393 confirmed occurrences, including regions that have been historically subject to botanical  
 394 surveys (Fig. 3A) (Hopkins 2007; Stropp et al. 2020; Carvalho et al. 2023), suggest that the  
 395 absence of records cannot be attributed solely to sampling gaps. Instead, this pattern points to  
 396 underlying ecological constraints that limit the establishment of the species across large areas  
 397 of apparently continuous forest cover.

398 The pronounced discrepancy between EOO and AOO highlights a strongly aggregated  
 399 spatial pattern, with occurrences restricted to discrete localities (Gaston 1996; Clavel et al.  
 400 2011) rather than being evenly distributed across Amazonian forest cover. This indicates that  
 401 the species occupies a limited fraction of its overall geographic range.

402 Taken together, the occurrence data indicate that *R. ucuquirana-branca* is  
 403 geographically widespread but spatially sparse, with confirmed records forming distinct  
 404 clusters across the basin. Similar patterns of broad geographic extent combined with highly  
 405 aggregated occurrence records have been reported for other Amazonian Sapotaceae, including  
 406 *Ecclinusa guianensis* Eyma (Vasconcelos et al. 2020), *Pradosia ptychandra* (Eyma)  
 407 T.D.Penn. (Vasconcelos et al. 2024) and *Pouteria decorticans* T.D.Penn. (Gama et al. 2025).  
 408 This pattern provides an essential empirical baseline for evaluating the environmental drivers  
 409 and modeled habitat suitability explored in the subsequent sections.

#### 411 **Model performance and ensemble robustness**

412  
 413 AUC and TSS metrics indicate a superior predictive capability (Fig. 4), effectively  
 414 distinguishing between suitable and unsuitable habitats with minimal omission errors  
 415 (Allouche et al. 2006). The consistently high performance observed across algorithms is in  
 416 agreement with previous studies showing that machine learning methods are more adept at  
 417 capturing complex, non-linear interactions between abiotic variables and species presence  
 418 than traditional frequentist models (Elith et al. 2008; Elith and Leathwick 2009).

419 The strength of the ensemble approach lies in its ability to mitigate the individual  
 420 biases of different algorithms, providing a consensus projection that is generally more robust  
 421 and biologically realistic than single-model outputs (Araújo and New 2007; Marmion et al.  
 422 2009). By combining predictions across algorithms, ensemble models tend to reduce  
 423 overfitting commonly observed in high-performance methods such as MaxEnt or Random  
 424 Forest, resulting in more conservative and reliable suitability estimates. This methodological  
 425 rigor increases confidence that the identified suitability hotspots are not merely artifacts of  
 426 sampling bias but instead reflect areas where the environmental conditions fall within the core  
 427 ecological envelope of *Ragala ucuquirana-branca* (Phillips et al. 2009).

#### 429 **Environmental drivers of distribution**

430  
 431 In the Amazon, soil chemistry acts as a powerful filter; species adapted to specific pH levels  
 432 are often associated with physiological trade-offs that prevent them from competing in soils  
 433 with different nutrient availabilities or aluminum toxicity levels (Phillips et al. 2003; John et  
 434 al. 2007; Quesada et al. 2012; Condit et al. 2013; Figueiredo et al. 2018).

435 Furthermore, the influence of vapor pressure and Bio6 suggests a narrow  
 436 physiological tolerance for atmospheric water demand and thermal stress. Vapor pressure,  
 437 closely related to atmospheric water demand and vapor pressure deficit, influences stomatal  
 438 regulation in canopy trees, limiting carbon assimilation and growth (Esquivel-Muelbert et al.  
 439 2017; Grossiord et al. 2020). The reliance on stable, humid conditions explains the species'  
 440 absence from the more seasonal fringes of the southern and eastern Amazon. This suggests  
 441 that the species' distribution is governed by a dual-filter system: a chemical filter at the

442 rhizosphere (pH) (e.g., Dubuis et al. 2013) and a physical filter at the leaf-atmosphere  
 443 interface (Vapor pressure) (e.g., Salm et al. 2007), both of which must be met simultaneously  
 444 for the species to persist.

445 Cec is a useful indicator of soil fertility because it reflects the soil's ability to retain and  
 446 exchange essential nutrients and is closely associated with clay mineralogy and organic matter  
 447 content. Its relevance suggests that *R. ucuquirana-branca* may preferentially occur in soils  
 448 with greater nutrient-retention capacity, which can buffer nutrient limitation in highly  
 449 weathered Amazonian substrates. Similar associations between Cec and tree species  
 450 distributions have been documented for several Amazonian forest communities (Phillips et al.  
 451 2003; Quesada et al. 2012; Toledo et al. 2017). Together, these results indicate that the  
 452 realized niche of *R. ucuquirana-branca* is shaped by a tight coupling between belowground  
 453 chemical constraints and aboveground atmospheric conditions, reinforcing the view that  
 454 multiple, interacting environmental filters determine tree species distributions in the Amazon.  
 455

### 456 **Predicted habitat suitability across the Amazon Basin**

457  
 458 The fact that the ensemble model predicts low suitability in regions with different soil pH and  
 459 atmospheric moisture, even those historically subject to botanical surveys, strongly supports  
 460 the hypothesis of edaphic specialization over a simple sampling artifact. This suggests that *R.*  
 461 *ucuquirana-branca* is a true habitat specialist whose occupancy is physically constrained by  
 462 the weathered substrates and stable climates of the central-western core.

463 The concentration of highly suitable areas in the central-western Amazonia suggests  
 464 the existence of a core ecological region for the species, where environmental conditions  
 465 closely match its physiological and edaphic requirements. In contrast, marginal areas toward  
 466 the eastern and southern basin likely represent environments beyond the species' tolerance  
 467 limits, rather than zones of transient occupancy. The isolated "islands" of habitat (Fig. 7)  
 468 suggest that much of the eastern basin is environmentally marginal or entirely unsuitable,  
 469 likely due to the combined effects of lower soil fertility and increased climatic seasonality  
 470 (Toledo et al. 2017).

471 This projected fragmentation has deep evolutionary and ecological implications.  
 472 Small, isolated patches of suitable habitat can lead to reduced gene flow and increased  
 473 vulnerability to stochastic events (Rosas et al. 2011). As climate change alters rainfall patterns  
 474 and vapor pressure across the basin (Esquivel-Muelbert et al. 2019; Flores et al. 2024; Lima et  
 475 al. 2026), these core areas may shrink further under projected increases in climatic seasonality  
 476 and atmospheric water demand. The suitability gap observed in the southern Amazonia acts as  
 477 a contemporary barrier, suggesting that the species may be unable to migrate southward in  
 478 response to changing climates, effectively trapping populations in the central basin (Araújo  
 479 and Pearson 2005; Bush et al. 2011).

### 481 **Taxonomic and ecological implications**

482  
 483 Our results indicate that *Ragala ucuquirana-branca* is characterized by ecological  
 484 specialization associated with soil chemical properties and atmospheric stability, rather than  
 485 broad environmental tolerance. This finding has important implications for taxonomy, as  
 486 environmentally structured distributions may contribute to morphological and genetic  
 487 differentiation across the species' range, a pattern frequently observed in Amazonian lineages  
 488 (Gentry 1981; Tuomisto et al. 2003; Fine et al. 2005, 2014). The environmentally structured  
 489 distribution revealed by the SDM provides testable hypotheses for future taxonomic and  
 490 phylogeographic studies. If populations occupying distinct soil or climatic domains exhibit

491 consistent morphological or genetic differentiation, this would support the role of ecological  
 492 specialization as a driver of diversification within the lineage.

493 From a conservation standpoint, the restricted AOO and the concentration of highly  
 494 suitable habitats in environmentally stable regions highlight potential vulnerability to ongoing  
 495 climatic change and land-use intensification, particularly in southeastern Amazonia.  
 496 Integrating SDM outputs with field-based inventories and phylogeographic data will be  
 497 essential for refining conservation assessments and understanding the evolutionary and  
 498 ecological processes shaping the distribution of *R. ucuquirana-branca*.

499

## 500 CONCLUSIONS

501

502 Our study provides the first assessment of the environmental drivers and potential distribution  
 503 of *Ragala ucuquirana-branca* across the Amazon Basin, illustrating how edaphic and  
 504 atmospheric filters can decouple geographic range size from effective habitat occupancy in  
 505 Amazonian trees. Our results demonstrate that the species' distribution is not determined  
 506 solely by broad climatic gradients, but is strongly filtered by specific edaphic conditions,  
 507 particularly soil pH, in combination with atmospheric moisture stability.

508 The high predictive performance of the ensemble model supports the identification of  
 509 the western and central Amazonia as the main biophysical strongholds for the species. The  
 510 marked discrepancy between the large Extent of Occurrence (EOO) and the highly restricted  
 511 Area of Occupancy (AOO) highlights pronounced ecological specialization and potential  
 512 vulnerability, consistent with a narrow niche breadth and limited tolerance to environmental  
 513 heterogeneity. As an edaphic specialist, *R. ucuquirana-branca* occupies fragmented and  
 514 environmentally stable habitats, which may increasingly constrain gene flow and long-term  
 515 resilience under ongoing environmental change.

516 From a conservation perspective, the species' dependence on specific soil chemistry  
 517 and stable atmospheric moisture regimes makes it particularly sensitive to the combined  
 518 effects of deforestation and climate change. More broadly, our findings emphasize the need to  
 519 move beyond climate-only frameworks when assessing species rarity, vulnerability, and  
 520 conservation priorities in hyperdiverse tropical forests.

521

## 522 ACKNOWLEDGEMENTS

523

524 We are grateful to the teams of the *Projeto Dinâmica Biológica de Fragmentos Florestais*  
 525 (PDBFF) and the *Centro de Estudos Integrados da Biodiversidade Amazônica* (CENBAM)  
 526 for providing access to and loans of specimens from the PDBFF and PPBio collections, with  
 527 special thanks to Ana Andrade, Alberto Vicentini, and Flavia Costa. We also thank Isolde  
 528 Ferraz, Francisco Farroñay, and José Edmilson Souza for providing the high-quality images  
 529 used in Fig. 1. Finally, we acknowledge all institutions and collections responsible for  
 530 generating, curating, and sharing the specimen data used in this study.

531

## 532 REFERENCES

533

- 534 Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution  
 535 models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*  
 536 43(6): 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>  
 537 Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology*  
 538 & *Evolution* 22(1): 42–47. <https://doi.org/10.1016/j.tree.2006.09.010>  
 539 Araújo MB, Pearson RG (2005) Equilibrium of species' distributions with climate. *Ecography*  
 540 28(5): 693–695. <https://doi.org/10.1111/j.2005.0906-7590.04253.x>

- 541 Bachman S, Moat J, Hill AW, Torre J, Scott B (2011) Supporting Red List threat assessments  
542 with GeoCAT: geospatial conservation assessment tool. *ZooKeys* 150: 117–126.  
543 <https://doi.org/10.3897/zookeys.150.2109>
- 544 Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for  
545 species distribution models: How, where and how many? *Methods in Ecology and*  
546 *Evolution* 3(2): 327–338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- 547 Breiman L (2001) Random Forests. *Machine Learning* 45(1): 5–32.  
548 <https://doi.org/10.1023/A:1010933404324>
- 549 Breiman L, Friedman J, Olshen RA, Stone CJ (1984) *Classification and Regression Trees*.  
550 Chapman and Hall/CRC, New York, 368. <https://doi.org/10.1201/9781315139470>
- 551 Bush M, Flenley J, Gosling W (2011) *Tropical rainforest responses to climatic change*, 2nd  
552 ed. Springer Berlin, Heidelberg, 454.
- 553 Canhos DAL, Almeida EA, Assad AL, Bustamante MMC, Canhos VP, Chapman AD,  
554 Giovanni R, Imperatriz-Fonseca VL, Lohmann LG, Maia LC, Miller JT, Nelson G,  
555 Peterson AT, Pirani JR, Souza S de, Stehmann JR, Thiers B (2022) *speciesLink*: Rich  
556 data and novel tools for digital assessments of biodiversity. *Biota Neotropica* 22:  
557 e20221394. <https://doi.org/10.1590/1676-0611-BN-2022-1394>
- 558 Carvalho RL, Resende AF, Barlow J, França F, Moura MR, Maciel R, Alves-Martins F, Shutt  
559 J, Nunes CA, Elias F, Juliana M. Silveira, Stegmann L, Baccaro FB, Juen L, Schiatti J,  
560 Aragão L, Berenguer E, Castello L, Costa FRC, Guedes ML, Leal CG, Lees AC, Isaac  
561 V, Nascimento RO, Phillips OL, Schmidt FA, Steege H ter, Vaz-de-Mello F,  
562 Venticinque EM, Vieira ICG, Zuanon J, Ferreira J, Filho AN dos SG, Ruschel A, Calor  
563 AR, Alves A de L, Muelbert AE, Quaresma A, Vicentini A, Piedade AR da, Oliveira  
564 AA de, Aleixo A, Casadei-Ferreira A, Gontijo A, Hercos A, Andriolo A, Lopes A,  
565 Pontes-Lopes A, Santos APM dos, Oliveira AB da S de, Mortati AF, Salcedo AKM,  
566 Albernaz AL, Fares AL, Luiza-Andrade A, Pes AMO, Faria APJ, Batista APB, Puker  
567 A, Bueno AS, Junqueira AB, Andrade ALRH de, Ghidini AR, Galuch A V., Menezes  
568 ASO de, Manzatto AG, Correa ASAS, Queiroz ACM, Zanzini AC da S, Neto AMO,  
569 Melo AWF de, Guimaraes AF, Castro AB, Borges A, Ferreira AB, Marimon BS,  
570 Marimon-Junior BH, Flores BM, Resende BO de, Albuquerque BW, Villa B, Davis B,  
571 Nelson B, Williamson B, Melo BSB de, Cintra BBL, Santos BB, Prudente B da S,  
572 Luize BG, Godoy BS, Rutt CL, Ritter CD, Silva CVJ, Ribas CR, Peres CA, Azevêdo  
573 CAS de, Freitas C, Cordeiro CL, Brocardo CR, Castilho C, Levis C, Doria CR da C,  
574 Arantes CC, Santos CA dos, Jakovac CC, Silva CA, Benetti CJ, Lasmar C, Marsh CJ,  
575 Andretti CB, Oliveira CP de, Cornelius C, Rosa CA da, Baider C, Gualberto CG, Deus  
576 CP de, Jr. C da SM, Neto CR dos S, Lobato CMC, Santos CRM dos, Penagos CCM,  
577 Costa D da S, Vieira DLM, Aguiar DPP de, Veras DS, Pauletto D, Braga D de L,  
578 Storck-Tonon D, Almeida D da F, Douglas D, Amaral DD do, Gris D, Luther D,  
579 Edwards DP, Guimarães DP, Santos DC dos, Campana DR de S, Nogueira DS, Silva  
580 DR da, Dutra DB de S, Rosa DCP, Silva DAS da, Pedroza D, Anjos D V., Lima DVM,  
581 Silvério D V., Rodrigues D de J, Bastos D, Daly D, Barbosa EM, Arenas ERC, Oliveira  
582 EA de, Santos EA dos, Santana ECC de, Guilherme E, Vidal E, Campos-Filho EM,  
583 Berg E van den, Morato EF, Silva ER da, Marques EE, Pringle EG, Nichols E,  
584 Andresen E, Farias E de S, Siqueira ELS de, Albuquerque EZ de, Görgens EB, Cunha  
585 EJ da, Householder E, Novo EMM de L, Oliveira FF de, Roque F de O, Coletti F,  
586 Reis F, Moreira FFF, Todeschini F, Carvalho FA, Souza FC de, Silva FAB, Carvalho  
587 FG, Cabeceira FG, D’Horta FM, Mendonça FP, Florêncio FP, Carvalho FR de, Arruda  
588 FV de, Nonato FA da S, Santana FD, Durgante F, Souza FKS de, Obermuller FA,  
589 Castro FS de, Wittmann F, Sales FM da S, Valente-Neto F, Salles FF, Borba GC,  
590 Damasco G, Barros GG, Brejão GL, Jardim GA, Prance GT, Lima GR, Desidério GR,

591 Melo G da C de, Carmo GHP do, Cabral GS, Rousseau GX, Silva GC da, Schwartz G,  
592 Griffiths H, Queiroz HL de, Espírito-Santo HVM, Cabette HSR, Nascimento HEM,  
593 Vasconcelos HL, Medeiros H, Aguiar HJAC de, Leão H, Wilker I, Gonçalves IC,  
594 Gorayeb I de S, Miranda IP de A, Brown IF, Santos ICS, Fernandes IO, Fernandes I,  
595 Delabie JHC, Abreu JC de, Neto J de LG, Costa JBP, Noronha JC, Brito JG de, Wolfe  
596 J, Santos JC, Ferreira-Ferreira J, Gomes JO, Lasky JR, Falcão JC de F, Costa JG, Cravo  
597 JS, Guerrero JEB, Gutiérrez JAM, Carreiras J, Lanna J, Brito JS, Schöngart J, Aguiar  
598 JJM, Lima J, Barroso JG, Noriega JA, Pereira JL da S, Nessimian JL, Souza JLP de,  
599 Toledo JJ de, Magalhães JLL, Camargo JL, Jr. JMBO, Ribeiro JMF, Silva JO de A,  
600 Guimarães JR da S, Hawes JE, Andrade-Silva J, Revilla JDC, Silva JS da, Menger J da  
601 S, Rechetelo J, Stropp J, Barbosa JF, Vale JD do, Louzada J, Silva JCC, Silva KD da,  
602 Melgaço K, Carvalho KS, Yamamoto KC, Mendes KR, Vulinec K, Maia LF,  
603 Cavalheiro L, Vedovato LB, Demarchi LO, Giacomini L, Dumas LL, Maracahipes L,  
604 Brasil LS, Ferreira LV, Calvão LB, Maracahipes-Santos L, Reis LP, Silva LF da, Melo  
605 L de O, Carvalho LC da S, Casatti L, Amado LL, Matos LS de, Vieira L, Prado LP do,  
606 Alencar L, Fontenele L, Mazzei L, Paolucci LN, Zanzini LP, Carvalho LN, Crema LC,  
607 Brulinger LFB, Montag LF de A, Naka LN, Azara L, Silveira LF, Nunes LG de O,  
608 Rosalino LM do C, Mestre LAM, Bonates LC de M, Coelho L de S, Borges LHM,  
609 Lourenço L da S, Freitas MAB, Brito MT da S, Pombo MM, Rocha M da, Cardoso MR,  
610 Guedes MC, Raseira MB, Medeiros MB de, Carim M de JV, Simon MF, Pansonato  
611 MP, Anjos MR dos, Nascimento MT, Souza MR de, Monteiro MGT, Silva MJ da,  
612 Uehara-Prado M, Oliveira MA de, Callisto M, Vital MJS, Santos MPD, Silveira M,  
613 Oliveira MVND, Pérez-Mayorga MA, Carniello MA, Lopes MA, Silveira MAP de A,  
614 Esposito MC, Maldaner ME, Passos MIS, Anacléto MJP, Costa MKS, Martins MP,  
615 Piedade MTF, Irumé MV, Costa MMS da, Maximiano MF de A, Freitas MG, Cochrane  
616 MA, Gastauer M, Almeida MRN, Souza MF de, Catarino M, Batista MC, Massam MR,  
617 Martins MF de O, Holmgren M, Almeida M, Dias MS, Santo NBE, Benone NL,  
618 Ivanauskas NM, Medeiros N, Targhetta N, Félix NS, Jr. NF, Hamada N, Campos N,  
619 Giehl NF da S, Metcalf OC, Silva OGM da, Cerqueira PV, Moser P, Miranda PN,  
620 Peruquetti PSF, Alverga PP de P, Prist P, Souto P, Brando P, Pompeu P dos S, Barni  
621 PE, Graça PM de A, Morandi PS, Cruz PV, Silva PG da, Bispo PC, Camargo PB de,  
622 Sarmiento PS de M, Souza P, Andrade RB de, Braga RB, Boldrini R, Bastos RC, Assis  
623 RL de, Salomão RP, Leitão RP, Raimundo N.G. Mendes J, Carpanedo R de S, Melinski  
624 RD, Ligeiro R, Pérez REP, Barbosa RI, Cajaiba RL, Silvano RAM, Salomão RP,  
625 Hilário RR, Martins RT, Perdiz R de O, Vicente RE, Silva RJ da, Koroiva R, Solar R,  
626 Silva R da C, Lima RB de, Silva R dos SA da, Mariano R, Ribeiro RAB, Fadini RF,  
627 Oliveira RLC de, Feitosa RM, Matavelli R, Mormul RP, Silva RR da, Zanetti R,  
628 Barthem R, Almeida RPS, Ribeiro SC, Neto SV da C, Nienow S, Oliveira SAV de,  
629 Borges SH, Milheiras S, Ribeiro SP, Couceiro SRM, Sousa SA de, Rodrigues SB, Dutra  
630 SL, Mahood S, Vieira SA, Arrolho S, Silva SS da, Triana SP, Laurance S, Kunz SH,  
631 Alvarado ST, Rodrigues THA, Santos TF dos, Machado TL da S, Feldpausch TR,  
632 Sousa T, Michelin TS, Emilio T, Brito T de F, André T, Barbosa TAP, Miguel TB, Izzo  
633 TJ, Laranjeiras TO, Mendes TP, Silva TSF, Krolow TK, Begot TO, Baker TR,  
634 Domingues TF, Giarrizzo T, Bentos TV, Haugaasen T, Peixoto U, Pozzobom UM,  
635 Korasaki V, Ribeiro VS, Scudeller VV, Oliveira VHF, Landeiro VL, Ferreira VRS,  
636 Silva V de NG, Gomes VHF, Oliveira VC de, Firmino V, Santiago WTV, Beiroz W,  
637 Almeida WR de, Oliveira WL de, Silva WC da, Castro W, Dáttilo W, Cruz WJA da,  
638 Silva WFM da, Magnusson WE, Laurance W, Milliken W, Paula WS de, Malhi Y,  
639 Shimabukuro YE, Lima YG de, Shimano Y, Feitosa Y (2023) Pervasive gaps in  
640 Amazonian ecological research. *Current Biology* 33(16): 3495–3504.

- 641 <https://doi.org/10.1016/j.cub.2023.06.077>
- 642 Clavel J, Julliard R, Devictor V (2011) Worldwide decline of specialist species: toward a  
643 global functional homogenization? *Frontiers in Ecology and the Environment* 9(4):  
644 222–228. <https://doi.org/10.1890/080216>
- 645 Condit R, Engelbrecht BMJ, Pino D, Pérez R, Turner BL (2013) Species distributions in  
646 response to individual soil nutrients and seasonal drought across a community of  
647 tropical trees. *Proceedings of the National Academy of Sciences* 110(13): 5064–5068.  
648 <https://doi.org/10.1073/pnas.1218042110>
- 649 De'ath G, Fabricius KE (2000) Classification and Regression Trees: A powerful yet simple  
650 technique for ecological data analysis. *Ecology* 81(11): 3178–3192.  
651 [https://doi.org/10.1890/0012-9658\(2000\)081\[3178:CARTAP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2)
- 652 Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B,  
653 Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B,  
654 Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of  
655 methods to deal with it and a simulation study evaluating their performance. *Ecography*  
656 36(1): 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- 657 Dubuis A, Giovanettina S, Pellissier L, Pottier J, Vittoz P, Guisan A (2013) Improving the  
658 prediction of plant species distribution and community composition by adding edaphic  
659 to topo-climatic variables. *Journal of Vegetation Science* 24(4): 593–606.  
660 <https://doi.org/10.1111/jvs.12002>
- 661 Elith J, Leathwick JR (2009) Species Distribution Models: Ecological explanation and  
662 prediction across space and time. *Annual Review of Ecology, Evolution, and*  
663 *Systematics* 40(1): 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- 664 Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *Journal*  
665 *of Animal Ecology* 77(4): 802–813. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>
- 666 Esquivel-Muelbert A, Baker TR, Dexter KG, Lewis SL, Brienen RJW, Feldpausch TR, Lloyd  
667 J, Monteagudo-Mendoza A, Arroyo L, Álvarez-Dávila E, Higuchi N, Marimon BS,  
668 Marimon-Junior BH, Silveira M, Vilanova E, Gloor E, Malhi Y, Chave J, Barlow J,  
669 Bonal D, Davila Cardozo N, Erwin T, Fauset S, Hérault B, Laurance S, Poorter L, Qie  
670 L, Stahl C, Sullivan MJP, Steege H ter, Vos VA, Zuidema PA, Almeida E, Almeida de  
671 Oliveira E, Andrade A, Vieira SA, Aragão L, Araujo-Murakami A, Arets E, Aymard C  
672 GA, Baraloto C, Camargo PB, Barroso JG, Bongers F, Boot R, Camargo JL, Castro W,  
673 Chama Moscoso V, Comiskey J, Cornejo Valverde F, Lola da Costa AC, Aguila  
674 Pasquel J del, Fiore A Di, Fernanda Duque L, Elias F, Engel J, Flores Llampazo G,  
675 Galbraith D, Herrera Fernández R, Honorio Coronado E, Hubau W, Jimenez-Rojas E,  
676 Lima AJN, Umetsu RK, Laurance W, Lopez-Gonzalez G, Lovejoy T, Aurelio Melo  
677 Cruz O, Morandi PS, Neill D, Núñez Vargas P, Pallqui Camacho NC, Parada Gutierrez  
678 A, Pardo G, Peacock J, Peña-Claros M, Peñuela-Mora MC, Petronelli P, Pickavance  
679 GC, Pitman N, Prieto A, Quesada C, Ramírez-Angulo H, Réjou-Méchain M, Restrepo  
680 Correa Z, Roopsind A, Rudas A, Salomão R, Silva N, Silva Espejo J, Singh J, Stropp J,  
681 Terborgh J, Thomas R, Toledo M, Torres-Lezama A, Valenzuela Gamarra L, Meer PJ  
682 van de, Heijden G van der, Hout P van der, Vasquez Martinez R, Vela C, Vieira ICG,  
683 Phillips OL (2019) Compositional response of Amazon forests to climate change.  
684 *Global Change Biology* 25(1): 39–56. <https://doi.org/10.1111/gcb.14413>
- 685 Esquivel-Muelbert A, Baker TR, Dexter KG, Lewis SL, Steege H ter, Lopez-Gonzalez G,  
686 Monteagudo Mendoza A, Brienen R, Feldpausch TR, Pitman N, Alonso A, Heijden G  
687 van der, Peña-Claros M, Ahuite M, Alexiades M, Álvarez Dávila E, Murakami AA,  
688 Arroyo L, Aulestia M, Balslev H, Barroso J, Boot R, Cano A, Chama Moscoso V,  
689 Comiskey JA, Cornejo F, Dallmeier F, Daly DC, Dávila N, Duivenvoorden JF, Duque  
690 Montoya AJ, Erwin T, Fiore A Di, Fredericksen T, Fuentes A, García-Villacorta R,

- 691 Gonzales T, Guevara Andino JE, Honorio Coronado EN, Huamantupa-Chuquimaco I,  
692 Eliana Maria Jiménez R, Killeen TJ, Malhi Y, Mendoza C, Mogollón H, Jørgensen PM,  
693 Montero JC, Mostacedo B, Nauray W, Neill D, Vargas PN, Palacios S, Palacios Cuenca  
694 W, Pallqui Camacho NC, Peacock J, Phillips JF, Pickavance G, Quesada CA, Ramírez-  
695 Angulo H, Restrepo Z, Reynel Rodriguez C, Paredes MR, Peñuela-Mora MC, Sierra R,  
696 Silveira M, Stevenson P, Stropp J, Terborgh J, Tirado M, Toledo M, Torres-Lezama A,  
697 Umaña MN, Urrego LE, Vasquez Martinez R, Gamarra LV, Vela CIA, Vilanova Torre  
698 E, Vos V, Hildebrand P von, Vriesendorp C, Wang O, Young KR, Zartman CE, Phillips  
699 OL (2017) Seasonal drought limits tree species across the Neotropics. *Ecography* 40(5):  
700 618–629. <https://doi.org/10.1111/ecog.01904>
- 701 Eva HD, Huber O, Achard F, Balslev H, Beck S, Behling H, Belward AS, Beuchle R, Cleef  
702 AM, Colchester M, Duivenvoorden J, Hoogmoed M, Junk W, Kabat P, Kruijt B, Malhi  
703 Y, Müller JM, Pereira JM, Peres C, Prance GT, Roberts J, Salo J (2005) A proposal for  
704 defining the geographical boundaries of Amazonia. Office for Official Publications of  
705 the European Communities, Luxembourg, 52.
- 706 Faria AD, Pirani JR, Ribeiro JELS, Nylinder S, Terra-Araujo MH, Vieira PP, Swenson U  
707 (2017) Towards a natural classification of subfamily Chrysophylloideae (Sapotaceae) in  
708 the Neotropics. *Botanical Journal of the Linnean Society* 185(1): 27–55.  
709 <https://doi.org/10.1093/botlinnean/box042>
- 710 Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for  
711 global land areas. *International Journal of Climatology* 37(12): 4302–4315.  
712 <https://doi.org/10.1002/joc.5086>
- 713 Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in  
714 conservation presence/absence models. *Environmental Conservation* 24(1): 38–49.  
715 <https://doi.org/10.1017/S0376892997000088>
- 716 Figueiredo FOG, Zuquim G, Tuomisto H, Moulatlet GM, Balslev H, Costa FRC (2018)  
717 Beyond climate control on species range: The importance of soil data to predict  
718 distribution of Amazonian plant species. *Journal of Biogeography* 45(1): 190–200.  
719 <https://doi.org/10.1111/jbi.13104>
- 720 Fine PVA, Daly DC, Muñoz GV, Mesones I, Cameron KM (2005) The contribution of  
721 edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the western  
722 Amazon. *Evolution* 59(7): 1464–1478. <https://doi.org/10.1111/j.0014-3820.2005.tb01796.x>
- 724 Fine PVA, Zapata F, Daly DC (2014) Investigating processes of Neotropical rain forest tree  
725 diversification by examining the evolution and historical biogeography of the Proteae  
726 (Burseraceae). *Evolution* 68(7): 1988–2004. <https://doi.org/10.1111/evo.12414>
- 727 Flores BM, Montoya E, Sakschewski B, Nascimento N, Staal A, Betts RA, Levis C, Lapola  
728 DM, Esquivel-Muelbert A, Jakovac C, Nobre CA, Oliveira RS, Borma LS, Nian D,  
729 Boers N, Hecht SB, Steege H ter, Arieira J, Lucas IL, Berenguer E, Marengo JA, Gatti  
730 L V, Mattos CRC, Hirota M (2024) Critical transitions in the Amazon forest system.  
731 *Nature* 626(7999): 555–564. <https://doi.org/10.1038/s41586-023-06970-0>
- 732 Gama IG, Cunha KCM, Ferreira G, Amoêdo SC, Vasconcelos CC (2025) Modeling the  
733 current distribution of *Pouteria decorticans* (Sapotaceae) in the Amazon Basin.  
734 *Observatório de La Economía Latinoamericana* 23(12): e12751.  
735 <https://doi.org/10.55905/oelv23n12-158>
- 736 Gaston KJ (1996) Species-range-size distributions: patterns, mechanisms and implications.  
737 *Trends in Ecology & Evolution* 11(5): 197–201. [https://doi.org/10.1016/0169-5347\(96\)10027-6](https://doi.org/10.1016/0169-5347(96)10027-6)
- 739 GBIF.org (2024) Occurrence Download. <https://doi.org/10.15468/dl.pwe8bu> [accessed  
740 11.18.2024]

- 741 Gentry AH (1981) Distributional patterns and an additional species of the *Passiflora vitifolia*  
742 complex: Amazonian species diversity due to edaphically differentiated communities.  
743 Plant Systematics and Evolution 137(1/2): 95–105
- 744 Grossiord C, Buckley TN, Cernusak LA, Novick KA, Poulter B, Siegwolf RTW, Sperry JS,  
745 McDowell NG (2020) Plant responses to rising vapor pressure deficit. New Phytologist  
746 226(6): 1550–1566. <https://doi.org/10.1111/nph.16485>
- 747 Guayasamin JM, Ribas CC, Carnaval AC, Carrillo JD, Hoorn C, Loham LG, Riff D, Ulloa-  
748 Ulloa C, Albert JS (2024) Evolution of Amazonian biodiversity: A review. Acta  
749 Amazonica 54: e54bc21360. <https://doi.org/10.1590/1809-4392202103601>
- 750 Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, Tulloch AIT,  
751 Regan TJ, Brotons L, McDonald-Madden E, Mantyka-Pringle C, Martin TG, Rhodes  
752 JR, Maggini R, Setterfield SA, Elith J, Schwartz MW, Wintle BA, Broennimann O,  
753 Austin M, Ferrier S, Kearney MR, Possingham HP, Buckley YM (2013) Predicting  
754 species distributions for conservation decisions. Ecology Letters 16(12): 1424–1435.  
755 <https://doi.org/10.1111/ele.12189>
- 756 Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology.  
757 Ecological Modelling 135(2): 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- 758
- 759 Hao T, Elith J, Guillera-Arroita G, Lahoz-Monfort JJ (2019) A review of evidence about use  
760 and performance of species distribution modelling ensembles like BIOMOD. Diversity  
761 and Distributions 25(5): 839–852. <https://doi.org/10.1111/ddi.12892>
- 762 Hao T, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2020) Testing whether ensemble  
763 modelling is advantageous for maximising predictive performance of species  
764 distribution models. Ecography 43(4): 549–558. <https://doi.org/10.1111/ecog.04890>
- 765 Hastie T, Tibshirani R (1986) Generalized Additive Models. Statistical Science 1(3): 297–  
766 310.
- 767 Hijmans RJ, Barbosa M, Aniruddha G, Mandel A (2024a) Geodata: Download Geographic  
768 Data. Version 0.6-2. <https://cran.r-project.org/package=geodata> [accessed 12.04.2025]
- 769 Hijmans RJ, Etten J, Sumner M, Cheng J, Baston D, Bevan A, Bivand R, Busetto L, Canty M,  
770 Fasoli B, Forrest D, Ghosh A, Golicher D, Gray J, Greenberg JA, Hiemstra P, Hingee  
771 K, Ilich A, Karney C, Mattiuzzi M, Mosher S, Naimi B, Nowosad J, Pebesma E,  
772 Lamigueiro OP, Racine EB, Rowlingson B, Shortridge A, Venables B, Wueest R  
773 (2024b) Dismo: Species Distribution Modeling. Version 1.3-16.  
774 <https://doi.org/10.32614/CRAN.package.dismo> [accessed 11.03.2025]
- 775 Hopkins MJG (2007) Modelling the known and unknown plant biodiversity of the Amazon  
776 Basin. Journal of Biogeography 8(34): 1400–1411. <https://doi.org/10.1111/j.1365-2699.2007.01737.x>
- 777
- 778 John R, Dalling JW, Harms KE, Yavitt JB, Stallard RF, Mirabello M, Hubbell SP, Valencia  
779 R, Navarrete H, Vallejo M (2007) Soil nutrients influence spatial distributions of  
780 tropical tree species. Proceedings of the National Academy of Sciences 104(3): 864–  
781 869. <https://doi.org/10.1073/pnas.0604666104>
- 782 Kukachka BF (1981) Wood anatomy of the Neotropical Sapotaceae. XXV. *Ragala*. USDA  
783 Forest Service Research Paper FPL 396: 1–6.
- 784 Lambers H, Oliveira RS (2019) Plant Physiological Ecology, 3rd ed. Springer, Switzerland,  
785 736. <https://doi.org/10.1007/978-3-030-29639-1>
- 786 Liaw A, Wiener M (2002) Classification and Regression by randomForest. R News 2/3: 18–  
787 22.
- 788 Lima RB, Oliveira CP, Silva DAS, Granato-Souza D, Gonçalves LA, Vasconcelos CC,  
789 Batista APB, Matos-Filho JR, Silva JPG, Aparício PS, Souza CSC, Abreu JC, Lopes  
790 IJC, Ometto JP, Görgens EB (2026) Warmer climate threatens the occurrence of giant

- 791 trees in the Amazon basin. *Journal of Environmental Management* 399: 128616.  
792 <https://doi.org/10.1016/j.jenvman.2026.128616>
- 793 Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of  
794 consensus methods in predictive species distribution modelling. *Diversity and*  
795 *Distributions* 15(1): 59–69. <https://doi.org/10.1111/j.1472-4642.2008.00491.x>
- 796 Miller J (2010) Species Distribution Modeling. *Geography Compass* 4(6): 490–509.  
797 <https://doi.org/10.1111/j.1749-8198.2010.00351.x>
- 798 Morrow CB (2019) Thresholding Species Distribution Models.  
799 <https://babichmorrowc.github.io/post/2019-04-12-sdm-threshold/> [accessed 11.10.2025]
- 800 Pennington TD (1990) *Flora Neotropica Monograph 52: Sapotaceae*. New York Botanical  
801 Garden, New York, NY, 770.
- 802 Pennington TD (1991) The genera of Sapotaceae. Royal Botanical Gardens, Kew, UK, 307.
- 803 Pennington TD (2006) *Flora da Reserva Ducke, Amazonas, Brasil: Sapotaceae*. *Rodriguésia*  
804 57(2): 252–366. <https://doi.org/10.1590/2175-7860200657210>
- 805 Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species  
806 geographic distributions. *Ecological Modelling* 190(3): 231–259.  
807 <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- 808 Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample  
809 selection bias and presence-only distribution models: implications for background and  
810 pseudo-absence data. *Ecological Applications* 19(1): 181–197.  
811 <https://doi.org/10.1890/07-2153.1>
- 812 Phillips OL, Vargas PN, Monteagudo AL, Cruz AP, Zans MEC, Sánchez WG, Yli-Halla M,  
813 Rose S (2003) Habitat association among Amazonian tree species: a landscape-scale  
814 approach. *Journal of Ecology* 91(5): 757–775. [https://doi.org/10.1046/j.1365-](https://doi.org/10.1046/j.1365-2745.2003.00815.x)  
815 [2745.2003.00815.x](https://doi.org/10.1046/j.1365-2745.2003.00815.x)
- 816 Poggio L, Sousa LM de, Batjes NH, Heuvelink GBM, Kempen B, Ribeiro E, Rossiter D  
817 (2021) SoilGrids 2.0: producing soil information for the globe with quantified spatial  
818 uncertainty. *SOIL* 7(1): 217–240. <https://doi.org/10.5194/soil-7-217-2021>
- 819 QGIS Development Team (2022) QGIS Geographic Information System. <https://qgis.org/>  
820 [accessed 12.01.2022]
- 821 Quesada CA, Phillips OL, Schwarz M, Czimczik CI, Baker TR, Patiño S, Fyllas NM, Hodnett  
822 MG, Herrera R, Almeida S, Alvarez Dávila E, Arneth A, Arroyo L, Chao KJ, Dezzeo  
823 N, Erwin T, Fiore A di, Higuchi N, Honorio Coronado E, Jimenez EM, Killeen T,  
824 Lezama AT, Lloyd G, López-González G, Luizão FJ, Malhi Y, Monteagudo A, Neill  
825 DA, Núñez Vargas P, Paiva R, Peacock J, Peñuela MC, Peña Cruz A, Pitman N, Priante  
826 Filho N, Prieto A, Ramírez H, Rudas A, Salomão R, Santos AJB, Schmerler J, Silva  
827 N, Silveira M, Vásquez R, Vieira I, Terborgh J, Lloyd J (2012) Basin-wide variations in  
828 Amazon forest structure and function are mediated by both soils and climate.  
829 *Biogeosciences* 9(6): 2203–2246. <https://doi.org/10.5194/bg-9-2203-2012>
- 830 R Core Team (2024) R: A Language and Environment for Statistical Computing. R  
831 Foundation for Statistical Computing. <https://www.r-project.org/> [accessed 15.01.2025]
- 832 Rodrigues WA (1974) Subsídios para o estudo das Sapotáceas da Amazônia. *Acta Amazonica*  
833 4(3): 9–18. <https://doi.org/10.1590/1809-43921974043009>
- 834 Rosas F, Quesada M, Lobo JA, Sork VL (2011) Effects of habitat fragmentation on pollen  
835 flow and genetic diversity of the endangered tropical tree *Swietenia humilis*  
836 (Meliaceae). *Biological Conservation* 144(12): 3082–3088.  
837 <https://doi.org/10.1016/j.biocon.2011.10.003>
- 838 Salm R, Salles NV de, Alonso WJ, Schuck-Paim C (2007) Cross-scale determinants of palm  
839 species distribution. *Acta Amazonica* 37(1): 17–26. [https://doi.org/10.1590/S0044-](https://doi.org/10.1590/S0044-59672007000100002)  
840 [59672007000100002](https://doi.org/10.1590/S0044-59672007000100002)

- 841 Stein A, Gerstner K, Kreft H (2014) Environmental heterogeneity as a universal driver of  
842 species richness across taxa, biomes and spatial scales. *Ecology Letters* 17(7): 866–880.  
843 <https://doi.org/10.1111/ele.12277>
- 844 Stropp J, Umbelino B, Correia RA, Campos-Silva JV, Ladle RJ, Malhado ACM (2020) The  
845 ghosts of forests past and future: deforestation and botanical sampling in the Brazilian  
846 Amazon. *Ecography* 43(7): 979–989. <https://doi.org/10.1111/ecog.05026>
- 847 Swenson U, Lepschi B, Lowry II PP, Terra-Araujo MH, Santos K, Nylinder S, Alves-Araújo  
848 A (2023) Reassessment of generic boundaries in Neotropical Chrysophylloideae  
849 (Sapotaceae): Eleven reinstated genera and narrowed circumscriptions of  
850 *Chrysophyllum* and *Pouteria*. *Taxon* 72(2): 307–359. <https://doi.org/10.1002/tax.12894>
- 851 ter Steege H, Pitman NCA, Phillips OL, Chave J, Sabatier D, Duque A, Molino JF, Prévost  
852 MF, Spichiger R, Castellanos H, Hildebrand P von, Vásquez R (2006) Continental-scale  
853 patterns of canopy tree composition and function across Amazonia. *Nature* 443(7110):  
854 444–447. <https://doi.org/10.1038/nature05134>
- 855 ter Steege H, Pitman NCA, Sabatier D, Baraloto C, Salomão RP, Guevara JE, Phillips OL,  
856 Castilho CV, Magnusson WE, Molino JF, Monteagudo A, Vargas PN, Montero JC,  
857 Feldpausch TR, Coronado EN, Killeen TJ, Mostacedo B, Vasquez R, Assis RL,  
858 Terborgh J, Wittmann F, Andrade AC, Laurance WF, Laurance SGW, Marimon  
859 BSBHS, Marimon BSBHS, Vieira ICG, Amaral IL, Brien R, Castellanos H, López  
860 DC, Duivenvoorden JF, Mogollón HF, Matos FDDA, Dávila N, García-Villacorta R,  
861 Diaz PRS, Costa F, Emilio T, Levis C, Schiatti J, Souza P, Alonso A, Dallmeier F,  
862 Montoya AJD, Piedade MTF, Araujo-Murakami A, Arroyo L, Gribel R, Fine PVA,  
863 Peres CA, Toledo M, Aymard C. GA, Baker TR, Cerón C, Engel J, Henkel TW, Maas  
864 P, Petronelli P, Stropp J, Zartman CE, Daly DC, Neill D, Silveira M, Paredes MR,  
865 Chave J, Lima Filho DDA, Jørgensen PM, Fuentes A, Schöngart J, Valverde FC, Fiore  
866 A Di, Jimenez EM, Mora MCP, Phillips JF, Rivas G, Andel TR van, Hildebrand P von,  
867 Hoffman B, Zent EL, Malhi Y, Prieto A, Rudas A, Ruschell AR, Silva N, Vos V, Zent  
868 S, Oliveira AA, Schutz AC, Gonzales T, Nascimento MT, Ramirez-Angulo H, Sierra R,  
869 Tirado M, Medina MNU, Heijden G Van Der, Vela CIA, Torre EV, Vriesendorp C,  
870 Wang O, Young KR, Baider C, Balslev H, Ferreira C, Mesones I, Torres-Lezama A,  
871 Giraldo LEU, Zagt R, Alexiades MN, Hernandez L, Huamantupa-Chuquimaco I,  
872 Milliken W, Cuenca WP, Pauletto D, Sandoval EV, Gamarra LV, Dexter KG, Feeley K,  
873 Lopez-Gonzalez G, Silman MR, Núñez-Vargas P, Montero JC, Feldpausch TR,  
874 Coronado ENH, Killeen TJ, Mostacedo B, Vasquez R, Assis RL, Terborgh J, Wittmann  
875 FK, Andrade ACS, Laurance WF, Laurance SGW, Marimon BSBHS, Marimon  
876 BSBHS, Guimarães-Vieira IC, Amaral IL, Brien R, Castellanos H, Cárdenas-López  
877 D, Duivenvoorden JF, Mogollón HF, Matos FDDA, Dávila N, García-Villacorta R,  
878 Stevenson-Diaz PR, Costa F, Emilio T, Levis C, Schiatti J, Souza P, Alonso A,  
879 Dallmeier F, Montoya AJD, Fernandez-Piedade MT, Araujo-Murakami A, Arroyo L,  
880 Gribel R, Fine PVA, Peres CA, Toledo M, Aymard CGA, Baker TR, Cerón C, Engel J,  
881 Henkel TW, Maas P, Petronelli P, Stropp J, Zartman CE, Daly DC, Neill D, Silveira M,  
882 Paredes MR, Chave J, Lima-Filho DA, Jørgensen PM, Fuentes A, Schöngart J, Cornejo-  
883 Valverde F, Fiore A Di, Jimenez EM, Peñuela-Mora MC, Phillips JF, Rivas G, Andel  
884 TR van, Hildebrand P von, Hoffman B, Zent EL, Malhi Y, Prieto A, Rudas A, Ruschell  
885 AR, Silva N, Vos V, Zent S, Oliveira AA, Schutz AC, Gonzales T, Trindade-  
886 Nascimento, M. Ramirez-Angulo H, Sierra R, Tirado M, Umaña-Medina MN, Heijden  
887 MGA van der, Vela CIA, Vilanova-Torre E, Vriesendorp C, Wang O, Young KR,  
888 Baider C, Balslev H, Ferreira C, Mesones I, Torres-Lezama A, Urrego-Giraldo LE, Zagt  
889 R, Alexiades MN, Hernandez L, Huamantupa-Chuquimaco I, Milliken W, Palacios-  
890 Cuenca W, Pauletto D, Valderrama-Sandoval E, Valenzuela-Gamarra L, Dexter KG,

- 891 Feeley K, Lopez-Gonzalez G, Silman MR (2013) Hyperdominance in the Amazonian  
892 tree flora. *Science* 342(6156): 1243092. <https://doi.org/10.1126/science.1243092>
- 893 Therneau T, Atkinson B (2023) Rpart: Recursive Partitioning and Regression Trees. Version  
894 4.1.23. <https://cran.r-project.org/package=rpart> [accessed 11.04.2025]
- 895 Toledo JJ, Castilho C V, Magnusson WE, Nascimento HEM (2017) Soil controls biomass and  
896 dynamics of an Amazonian forest through the shifting of species and traits. *Brazilian*  
897 *Journal of Botany* 40(2): 451–461. <https://doi.org/10.1007/s40415-016-0351-2>
- 898 Toledo M, Peña-Claros M, Bongers F, Alarcón A, Balcázar J, Chuvina J, Leão C, Licona JC,  
899 Poorter L (2012) Distribution patterns of tropical woody species in response to climatic  
900 and edaphic gradients. *Journal of Ecology* 100(1): 253–263.  
901 <https://doi.org/10.1111/j.1365-2745.2011.01890.x>
- 902 Tuanmu MN, Jetz W (2014) A global 1-km consensus land-cover product for biodiversity and  
903 ecosystem modelling. *Global Ecology and Biogeography* 23(9): 1031–1045.  
904 <https://doi.org/10.1111/geb.12182>
- 905 Tuanmu MN, Jetz W (2015) A global, remote sensing-based characterization of terrestrial  
906 habitat heterogeneity for biodiversity and ecosystem modelling. *Global Ecology and*  
907 *Biogeography* 24(11): 1329–1339. <https://doi.org/10.1111/geb.12365>
- 908 Tuomisto H, Ruokolainen K, Yli-Halla M (2003) Dispersal, environment, and floristic  
909 variation of western Amazonian forests. *Science* 299(5604): 241–244.  
910 <https://doi.org/10.1126/science.1078037>
- 911 van Roosmalen MGM, Garcia OM da CG (2000) Fruits of the Amazonian Forest. Part II:  
912 Sapotaceae. *Acta Amazonica* 30(2): 187–290. <https://doi.org/10.1590/1809-43922000302290>
- 914 Vasconcelos CC, Amôêdo SC, Ferraz IDK, Camargo JLC, Terra-Araujo MH (2020)  
915 Modelagem preditiva da distribuição atual de *Ecclinusa guianensis* Eyma (Sapotaceae,  
916 Chrysophylloideae) na bacia amazônica. In: Alves-Araujo A, Abreu KMP, Calazans  
917 LSB, Thomaz LD, Pereira OJ, Dutra VF (Org.) Anais do 40º Encontro Regional de  
918 Botânicos - Regional MG, BA, ES. Universidade Federal do Espírito Santo, Vitória, 1–  
919 6. <https://doi.org/10.29327/40ERBOT.302847>
- 920 Vasconcelos CC, Diele-Viegas ML, Camargo JLC, Ferraz IDK, Terra-Araujo MH (2024)  
921 Filling a distribution gap: a new record of *Pradosia ptychandra* (Sapotaceae) from  
922 eastern Brazilian Amazonia. *Acta Amazonica* 54(2): e54bc23421.  
923 <https://doi.org/10.1590/1809-4392202304211>
- 924 Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood  
925 estimation of semiparametric generalized linear models. *Journal of the Royal Statistical*  
926 *Society: Series B (Statistical Methodology)* 73(1): 3–36. <https://doi.org/10.1111/j.1467-9868.2010.00749.x>
- 928 Yee TW, Mitchell ND (1991) Generalized Additive Models in plant ecology. *Journal of*  
929 *Vegetation Science* 2(5): 587–602. <https://doi.org/10.2307/3236170>
- 930 Zizka A, Silvestro D, Andermann T, Azevedo J, Ritter CD, Edler D, Farooq H, Herdean A,  
931 Ariza M, Scharn R, Svantesson S, Wengström N, Zizka V, Antonelli A (2019)  
932 CoordinateCleaner: Standardized cleaning of occurrence records from biological  
933 collection databases. *Methods in Ecology and Evolution* 10(5): 744–751.  
934 <https://doi.org/10.1111/2041-210X.13152>
- 935 Zurell D Mecofun: Useful Functions for Macroecology and Species Distribution Modelling  
936 (2024). Version 0.7.1. <https://gitup.uni-potsdam.de/macroecology/mecofun> [accessed  
937 11.04.2025]
- 938 Zurell D, Franklin J, König C, Bouchet PJ, Dormann CF, Elith J, Fandos G, Feng X, Guillera-  
939 Arroita G, Guisan A, Lahoz-Monfort JJ, Leitão PJ, Park DS, Peterson AT, Rapacciuolo  
940 G, Schmatz DR, Schröder B, Serra-Diaz JM, Thuiller W, Yates KL, Zimmermann NE,

941 Merow C (2020) A standard protocol for reporting species distribution models.  
942 *Ecography* 43(9): 1261–1277. <https://doi.org/10.1111/ecog.04960>