

Distribution and conservation of the species of Marmosini (Didelphimorphia, Didelphidae) from Colombia

Baltazar González (✉ baltazargch@gmail.com)

Centro de Investigación Esquel de Montaña y Estepa Patagónica: Centro de Investigación Esquel de Montana y Estepa Patagonica <https://orcid.org/0000-0001-5805-8107>

Federico Brook

Centro de Investigación Esquel de Montana y Estepa Patagonica

Gabriel M. Martin

Centro de Investigación Esquel de Montana y Estepa Patagonica

Research Article

Keywords: ENM, Marmosa, marsupials, Monodelphis, Neotropics, Andes

Posted Date: May 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-557895/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Marmosini species were taxonomically revised recently, however, little is known about their distribution and conservation. The aim of this research was to delimit the distribution of all species of Marmosini that inhabit Colombia, and analyze their conservation using potential distributions, protected areas, and human pressure data of the country. We used the widely known ecological niche modeling algorithm maxent to model the distribution of each species using two approaches to estimate the modeling area: a buffer-derived and an ecoregion-derived. After selecting a final model, we used data on protected areas and human pressure specific to Colombia, and analyze their conservation and pressure scenarios. Finally, we generated a species richness map for Marmosini in Colombia. We found that most species of Marmosini from Colombia co-occur at mid-elevations of the Andes with an upper elevation limit of maximum richness at ~ 2000 m. Marmosini species' distribution covers 91% of the country's continental area, and the maximum area protected for any species of this group is between 29–5.4% of their modeled distribution. Most of the protected areas under strict and national conservation types presented small areas of high human pressure, while other categories (conservation units under managed resources and other conservation types) presented large areas of high human pressure. These species are poorly protected by the Natural Protected Areas of Colombia. Future reserves that cover Andean regions below 2000 m could help optimize their conservation.

Declarations

Funding

No funding was received for conducting this study. Acknowledgment to indirect financial support for the authors are given in the *Acknowledgments* section.

Conflicts of interest/Competing interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Availability of data and material

All data generated or analysed during this study are included in this published article and its supplementary information files. Other public sources of information used were cited and declared.

Code availability

All codes generated for this research are available at the first author GitHub repository (https://github.com/baltazargch/sdm_marmosini_colombia).

Authors' contributions

BG, FB, and GM conceived the study. BG and GM design the method. BG, FB, and GM clean and prepared the data. BG developed the code and analyzed. BG redacted the first draft. All authors reviewed and discussed the results and read and approved the submitted manuscript.

ACKNOWLEDGMENTS

The authors thank Adrian Monjeau and Nature Map Argentina for providing the computer power needed for this project. BG and FB made this research while a Doctoral grant awarded by the National Science Council of Argentina (CONICET). GMM thanks E. Watkins and M. Simeon for the economic support.

Introduction

There are 40 marsupial species recognized as occurring in Colombia, of which 19 belong to tribe Marmosini (Solari et al. 2013; Ramírez-Chaves et al. 2016; Voss and Giarla 2021). However, there are no detailed maps of the current distribution for most species (but see Gutiérrez et al. (2014), that as for any other organism represents fundamental information for species conservation planning and basic ecological research. Marmosini includes the genera *Marmosa*, *Monodelphis* and *Tlacuatzin* (Voss and Jansa 2009; Rossi et al. 2010; Voss et al. 2020; Voss and Giarla 2021), the latter exclusive to Central America, while *Monodelphis* is restricted to South America. Recently, both *Marmosa* and *Monodelphis* have been taxonomically reviewed, providing clarity about species identities and occurrence localities (Pavan et al. 2014; Voss et al. 2014, 2020; Voss and Giarla 2021)

Colombia supports the highest species richness of Marmosini in South America with 19 species, followed by Perú with 18, Brazil with 16, Bolivia with 12, Venezuela with 10, and the rest of countries with less than 10 species (Solari et al. 2013; Tirira et al. 2020; IUCN 2021; Voss and Giarla 2021). Regarding their national threatened status, only one Marmosini species (*Marmosa xerophila*) is listed as Data Deficient in the last national assessment of threatened mammals in Colombia (Alberico and Rojas-Díaz 2006). At a global scale, the International Union for the Conservation of Nature (IUCN) lists 12 species for Colombia, of which two are Vulnerable (*M. xerophila* and *M. phaea*), one is Data Deficient (*M. rubra*), and the remaining species are Least Concern (IUCN 2021). In the last 12 years, new species have been identified, mainly as separation of widely distributed taxa (Rossi et al. 2010; Voss et al. 2020), but no assessment exists for these species yet. Consequently, to study the conservation biogeography of this group globally and in Colombia is key for any future efforts of preserving these species.

There are several ways to assess species' distributions. Recently, a variety of approaches including what is known as Ecological Niche Models (ENM) have gained strength in the scientific community to address this and related subjects (Elith et al. 2006; Qiao et al. 2015; Urbina-Cardona et al. 2019). These methods vary in their input requirements and output interpretation, but are based in a sound conceptual framework about species biogeography and their realized niche (Peterson et al. 2011; Peterson and Soberón 2012; Norberg et al. 2019). Moreover, these methods allow generating species distribution models based on associated niche characteristics (Peterson and Soberón 2012; Soberón et al. 2017).

One of the most widely used and tested algorithm for ENM is maxent (Elith et al. 2011; Merow et al. 2013; Phillips et al. 2017), but its implementation and results depend on several variables. For example, the area where the model is estimated on (Barve et al. 2011), the complexity of the model (Merow et al. 2014) and how accuracy is tested (Elith et al. 2011; Qiao et al. 2015), among others.

In general, the ecology and chorology of New World marsupials has been partially studied (poorly studied compared to other, most conspicuous mammals), and conservation efforts are hindered due to this lack of knowledge (Cayuela et al. 2009). Although the tribe Marmosini has been recently revised regarding its taxonomy, there is little knowledge about their distribution, which could provide a baseline of information such as where they occur and co-occur, essential to any future conservation task.

Altogether, the availability of powerful algorithms for estimating species niches and distribution, the increased clarity of species identities and the updated list of marsupials that inhabit Colombia, provides the opportunity to develop detailed distribution maps of the Marmosini species, and analyze their conservation.

The main objective of this work was to model the distribution of Marmosini species that inhabit Colombia, analyze their spatial richness patterns and conservation throughout the country. To accomplish this we estimated how much of their modeled distribution is under different types of protection and/or under different levels of human pressure.

Methods

When modeling species distributions with maxent (Phillips et al. 2017) and other methods, there are many configurations, inputs, and decisions that can affect final results (Elith et al. 2006, 2011; Elith and Leathwick 2009; Barve et al. 2011; Merow et al. 2013; Norberg et al. 2019). In this study, we developed a workflow that seeks to include the majority of these considerations, and evaluate them in many scenarios (Fig. 1).

- **Occurrence and background data**

For the species of Marmosini that inhabit Colombia we follow Solari et al. (2013), Ramírez-Chávez et al. (2016), and Voss and Giarla (2021), although the genus-level classification follows Voss and Jansa (2009) (i.e., we considered *Micoureus* a subgenus of *Marmosa*). We gathered occurrence records from two sources: localities from the literature and from the Global Biodiversity Information Facility (GBIF.org 2021). Occurrence data from literature was gathered using recent taxonomic revisions of the genera *Marmosa* and *Mondelphis* (see Appendix 1 – Sheet 1 for full list of references). The GBIF data was obtained by querying the database for records based on preserved specimens and material samples for the category 'basis of record', and with less than or equal to 1000 m for the category 'coordinate uncertainty'.

For GBIF data, we excluded any records without geographic coordinates, and corrected only obvious georeferencing errors such as positive-to-negative longitudinal or latitudinal coordinates. Otherwise, we deleted the entry. Additionally, we used standard best-practices to clean and prepare GBIF data using the R package 'CoordinateCleaner' (Zizka et al. 2019). When available, we manually inspected each species locality coordinates based on their taxonomic revision for possible inconsistencies. A complete account of the localities used for this study is presented in Appendix 1 – Sheet 1.

To match the resolution of occurrences to that of predictors, we filtered records that were separated by less than 1 km (i.e., each record corresponded to a unique environmental raster pixel in this study). We chose this resolution since many of the species we analyzed are found in forested areas in the Andean cordilleras, and more than one single horizontal kilometer could represent important variation due to high slopes.

We used maxent v. 3.4.3 (Phillips et al. 2017) to generate the models, with 10,000 random unique background points for each species as pseudo-absences (herein referred as background points). We used two types of modeling areas (see 2.3 below and Online Resource 1 – Fig S1), a unique set of 10,000 background points for each species and modeling area, and extracted the environmental data for presence and background points for each modeling scenario (see 2.2 below).

- **Environmental data**

In this study we modeled the distribution of the species of Marmosini based on four of what we call here predictors scenarios (i.e., different sets of environmental predictors) (Fig. 1b): three user-defined scenarios based on previous findings and our informed criteria about which variables could have higher explanation power, and one statistically-defined scenario in which predictors were selected in order to reduce collinearity among variables.

Current climatic conditions were represented by subsets of the following databases: the bioclimatic variables of WorldClim v. 2 (Fick and Hijmans 2017), ENVIREM (Title and Bemmels 2018) and one vegetation index, the Modified Soil Adjusted Vegetation Index (MSAVI). MSAVI was derived from the red and near infra-red bands of MODIS Terra imagery (Daac 2017). MSAVI was calculated as the mean for the year 2000 following Qi et al. (1994), which is the closest year available from MODIS Terra imagery to the bioclimatic variables mentioned above.

From all above-mentioned variables we selected 8 WorldClim and 2 ENVIREM, based on those that have explanation power for marsupials (Martin 2010, 2011; Gutiérrez et al. 2014; Tocchio et al. 2015), and those that may vary widely in the study zone such as topography, due to the Andes mountains. For the last case, we opted to use ecologically-oriented variables (Title and Bemmels 2018), rather than other more common topographic variables such as digital elevation models or slope. The full set of predictors resulted in 11 variables that were used in the four predictors scenarios as follows: 1) 'onlywc': only with WorldClim (8 variables), 2) 'ud.noplants': user defined with WorldClim + ENVIREM (10 variables), 3) 'ud.all': user defined with WorldClim + ENVIREM + MSAVI (11 variables), and 4) 'uncorr': collinearly

reduced variables from the full set (species-specific number and type of variables). A complete account of the variables used in each scenario is given in Appendix 1 – Sheet 2. To avoid high collinearity between the predictors in the ‘uncorr’ case, we used a Pearson test to randomly choose pairs of variables below $p=0.75$, by sampling within each species modeling area 500,000 random values. For models that did not include space and/or time extrapolation, previous studies found that collinearity among predictors may not significantly affect maxent models (Feng et al. 2019a) Thus, for the first three scenarios, predictors were chosen according to our informed criteria of potential explanation power (Fourcade et al. 2018) and not tested for collinearity. Each of the four scenarios were used for each species and each type of modeling area (see 2.3 below). All environmental data were downloaded as rasters of $\sim 1 \text{ km}^2$ spatial resolution at the equator (~ 30 arc-seconds), and were processed using the coordinate system WGS84.

- **Definition of modeling areas**

Barve et al. (2011) discussed how the modeling area from which predictors are sampled for background points affects modeling results. We used two methods to estimate this area, herein referred as M area, after the BAM diagram from Peterson and Soberón (2011). The M area represents a geographic space with suitable abiotic conditions historically available for the species to disperse (including currently occupied and unoccupied areas) (Peterson and Soberón 2012). Ideally, information from fossils are used to estimate these areas but in the absence of them, buffered methods have been used in these species (Gutiérrez et al. 2014; Tocchio et al. 2015).

The methods used to estimate M areas for model fitting and prediction are called here M1 and M2 (see Online Resource 1 – Fig. S1). M1 was generated as a minimum convex polygon from all the occurrence points plus a buffer of $\sim 330 \text{ km}$ at the equator in each direction (a similar approach was used for a couple of Marmosini species by Gutierrez et al. 2014). M2 was generated by overlapping a minimum convex polygon buffered by $\sim 50 \text{ km}$ in each direction with an ecoregion map (Dinerstein et al. 2017) and selecting the ecoregions that overlapped. Minimum convex polygons were generated with the R package ‘adehabitatHR’ (Calenge 2006). These methods represent two ways of estimating the M area for each species, and its performance was evaluated according to four evaluation metrics as described below.

- **Modeling procedure and evaluation**

For model tuning and fitting we used the algorithm maxent through the function *ENMevaluate* from the R package ‘ENMeval’ (Muscarella et al. 2014) using the features: Linear (L), Quadratic (Q), Product (P), Hinge (H), and Threshold (T), depending on the number of occurrences. For species with less than or equal to 80 occurrence records, we used 'L', 'LQ', and 'LQP'; for species with more than 80 occurrences we used 'L', 'LQ', 'LQP', 'H', 'LQH', 'LQHP', and 'LQHPT' (Merow et al. 2013). These features were combined with different regularization multipliers (rm): from 0.5 to 5 in increasing steps of 0.5. Due to the difference in number of records between species, we also tested for different methods of cross-validation using a conditional approach (Fig. 1c). For all species, a block-type partitioning cross-validation method was

used, if the species had less than 25 records we used a jackknife approach, and if the species had 25 or more records a random k-fold approach with 5 partitions was used.

To select the best configuration for each species (features and rm), M area, and predictors scenarios, we evaluated the results through the next ordered-steps: i) choosing the 75% of higher values of the average test area under the receiver-operator curve (AUC), ii) choosing the models with minimum average difference AUC between test and training models, iii) models with minimum average test of omission rate at the minimum training presence (orMTP), iv) models with the minimum value for the corrected Akaike information criterion (AICc), and v) if more than one model remained after the previous steps, we chose the model that minimized the regularization multiplier (rm), maximized the train AUC, and minimized the number of parameters. In the rare cases where all previous filters ended in more than one model, a final model was chosen at random.

To transform continuous maps of prediction several thresholds have been proposed (Liu et al. 2005, 2016). In this study we used the value that maximized the sum of sensitivity and specificity (maxSSS) to generate a binary map of 0 (absence) and 1 (presence). We then downscaled the resolution from 1 km² to 4 km² to avoid zones with many isolated presence cells. Then, prediction rasters were converted to polygon data and their limits were smoothed using the *ksmooth* function from the R package 'smoothr' (Strimas-Mackey 2020), using a smoothing index of 2. Additionally, holes that were less than 100 km² and crumbs that were less than 50 km² were removed. For each species, four final models were inspected manually, the two best for each M area. Finally, when deemed necessary based on the knowledge of each species, final ranges were adjusted using known geographic barriers (e.g., Hazzi et al. (2018)(Fig. 1d, Online Resource 3).

- **Richness and conservation metrics**

To explore the spatial richness of Marmosini within the country, we generated a richness map by dividing continental Colombia in a grid of 25 km² pixel size, and counted how many species occurred per pixel based on final species ranges.

For conservation analyzes, based on the final maps we calculated the area of each species distribution within continental Colombia, the area within the Natural Protected Areas (NPA) of the country, and the area under high and low human pressure. NPA data set was downloaded from the World Database on Protected Areas (UNEP-WCMC and IUCN 2021), version February 2021), and was cleaned according to standard best-practices (Butchart et al. 2015; Runge et al. 2015) through the R package 'wdpar' (Hanson 2020). Additionally, marine NPA were excluded from the analysis. We then calculated the overlapped area for each species according to two criteria: (1) Conservation-only, where the NPA data set was divided firstly, according to their governance type (GOV_TYPE column from the WDPA) and secondly, according to the IUCN category (IUCN_CAT column from the WDPA). Governance type categories were pooled as national, sub-national, indigenous territories and local communities, shared and private governance (UNEP-WCMC and IUCN 2020). IUCN categories were pooled as strict-conservation (categories Ia, Ib, and

II) and managed resources (all other categories, excluding Non-Applicable and Non-Reported from WDPA; UNEP-WCMC and IUCN 2021). (2) Conservation-pressure, where the same IUCN category pooling criteria of the conservation-only approach was used, but differentiating between high and low pressure areas within each NPA.

To estimate areas of high and low human pressure, we generated a binary map from the Colombian Human Impact Index of 2015 (Correa Ayram et al. 2020) by selecting the lowest 40% of the values as low pressure, and the remaining as high pressure. Then, areas were calculated using a proportional approach: for each pixel of the human pressure raster, we calculated the proportion of it inside each species range and/or each protected area, and then multiplied the area of each pixel to its fraction inside a given polygon. Proportions were calculated by dividing each pixel in 100 equal-sized sub-units, and counting how many of them fall inside each polygon.

All areas were calculated in square kilometers using a geodesic approach, and based on a WGS84 projection. All analyzes were run in R v. 4.0.2 (R Core Team 2020) and rstudio 1.4 (RStudio Team 2020). We used qgis v. 3.18.1 to manually modify final range maps (QGIS.org 2021). The entire code used for this study will be made available at the first author's GitHub repository (https://github.com/baltazargch/sdm_marmosini_colombia) upon publishing.

Results

- **Occurrence data**

From the original list of 19 species of Marmosini from Colombia (Solari et al. 2013; Voss and Giarla 2021), we gathered information on 16 species, 13 from the genus *Marmosa* and 3 from *Monodelphis*. A total of 648 records were gathered and visually verified. The number of species records varied from 9 (*Marmosa phaea* and *M. jansae*) to 199 (*M. robinsoni*) after cleaning the data. From the 16 species included, 13 species had records within Colombia, while for the species *M. regina*, *Monodelphis brevicaudata*, and *Monodelphis palliolata* no records within the country were gathered, but were included because records fall within ecosystems that occur in the country.

The records of Marmosini, gathered from 19 different sources (Appendix 1 – Sheet 1), extended from 17.50° S and 86.1° W to 13.45° N and 51.763° W, and occurred within fourteen countries: Bolivia (n=17), Brazil (n=28), Colombia (n=101), Costa Rica (n=30), Ecuador (n=49), French Guyana (n=1), Grenada (n=2), Guyana (n=8), Nicaragua (n=8), Panama (n=71), Peru (n=88), Suriname (n=6), Trinidad and Tobago (n=16) and Venezuela (n=208).

Following are the analyses of the different aspects of models taken into account in this work, they represent the filtered results by selecting the upper quantile of the average test AUC metric, except otherwise stated. A complete account of the unfiltered results are available in Appendix 1 – Sheet 3.

- **Predictors scenarios**

Predictors scenarios varied when model performance metrics were analyzed. The two best scenarios based on train AUC and average test AUC were user-defined with and without vegetation ('ud.all' and 'ud.noplants'), which performed similarly. While 'onlywc' and 'uncorr' scenarios had the worst performance, irrespective to the rm analyzed (Fig. 2a-b). In contrast, for the orMTP and at low rm values, all models performed similarly. As the rm value increased they gradually differentiated, with all models being similarly good except for 'onlywc' scenario which was consistently worst from a rm value of 2.5, thus with a higher rate of omitting presences (Fig. 2c). For the AICc metric and at the lowest rm value (0.5), models performed distinctively, but as the rm increased models seemed to become similar in their performance, but with both user-defined scenarios being slightly better based on this metric (Fig. 2d).

- **Modeling areas**

When the different modeling areas were compared regarding their overall performance with the metrics used here, M2 outperformed M1 in most of the cases with different predictor scenarios, cross-validation type and rm (Online Resource 2 – Fig. S2). For train and test AUC, M2 was consistently better than M1, except for jackknife cross-validation with which models from M1 and M2 seemed to performed comparably (Online Resource 2 – Fig. S2a-b). When evaluated with orMTP, both areas performed similarly with only small differences in the high values obtained for M2, jackknife and 'onlywc' case from rm 1 to 2.5, and in the block type cross-validation, where M2 was slightly better than M1 (Online Resource 2 – Fig. S2c). However, for AICc M1 was better than M2 in most of the comparisons (Online Resource 2 – Fig. S2d).

- **Model results**

A total of 8,320 models were generated, 2,080 for each predictor scenarios, and between 120 and 280 models for each species, depending on numbers of records (Fig. 1c). A total of 64 models were visually and critically inspected to decide based on biological data, which best represented each species' distribution. Most of the chosen models were from the predictors scenarios based on 'ud.all' (n = 7), followed by 'onlywc' (n = 4), 'ud.noplants' (n = 3), and finally 'uncorr' (n = 2). Regarding cross-validation, 9 out of 16 final models were from those using random k-fold method, and 7 out of 16 used block method. Most of the models were from regularization multipliers smaller than two and had a train AUC > 0.7. Other metrics and configurations of the final models are presented in Table 1 and Appendix 1 – Sheet 3.

Variable contribution and permutation importance varied greatly between species (Fig. 3). Precipitation related variables were the most important (with values above 50%) in the models generated for 11 species, and temperature related variables were the most important in 4 (*M. isthmica*, *M. phaea*, *M. rutteri*, and *M. xerophila*) (Fig. 3 and Appendix 1 – Sheet 4). In *M. regina*, temperature related variables, terrain roughness (tri) and MSAVI were the most important variables (Fig. 3 and Appendix 1 – Sheet 4). In seven species precipitation related variables represented the most important contribution and permutation values, while only three species had contribution and permutation values represented by temperature variables exclusively (Appendix 1 – Sheet 4). Six species had a combination of precipitation and temperature (*M. alstoni*, *M. lepida* and *M. waterhousei*), or temperature and precipitation (*M. xerophila*, *M.*

zeledoni, and *M. brevicaudata*), as the highest contribution and permutation variables, respectively. The first three species had Mean Temperature of the Coldest Quarter (bio11) as the variable with the highest information not present in the others (permutation importance) but different variables with the highest contribution to their models (i.e., bio15 for *M. waterhousei*, bio16 for *M. alstoni*, and bio17 for *M. lepida*) (Appendix 1 – Sheet 4). The variable “topographic wetness” was markedly high for *Monodelphis* species, especially for *M. adusta* and *M. palliolata*. The best model configuration for *M. phaea* resulted in a single variable contributing to the model, Mean Temperature of the Warmest Quarter, and the model for *M. germana* with only two variables (Precipitation Seasonality (Coefficient of Variation) and Precipitation of the Driest Quarter) (Appendix 1 – Sheet 4).

- **Species’ distribution**

After visual inspection of final models, geographical barriers were taken into account to modify and generate final range maps. A complete account of the geographical barriers proposed for delimiting each species and final range maps are in Online Resource 3. Different spatial patterns were found regarding their distribution in the country. To describe them, we followed the national categorization of continental biogeographic regions for local relevance and clarity, published by the Instituto de Investigación de Recursos Biológicos Alexander von Humboldt of Colombia (maps available at <https://www.redalyc.org/articulo.oa?id=49150103> and reproduced in inset map of Fig 4a).

Marmosa germana, *M. jansae*, *M. regina*, *M. rubra*, *M. rutteri*, and *M. brevicaudata* showed a distribution restricted to the Amazonian region, with several species showing small distribution areas in the Amazonian-Andean transition, east of Nudo de Los Pastos. *Marmosa lepida* and *M. phaea* were distributed in the Andean, Orinoquia and Amazonian regions, with the former being mainly distributed in the Amazonian and Orinoquia regions, and a distribution area at mid-elevations of the Eastern Andean region, while *M. phaea* was mainly distributed in the Andean region and few areas in the Amazonian and Orinoquia region, especially east of Nudo de los Pastos and Sierra de La Macarena. *Marmosa waterhousei* and *M. adusta* were distributed mainly in the Andean, Pacific, and Amazonian regions, with the former being widely distributed in the Andean and Amazonian regions and partly in the Pacific and Orinoquia regions, with few areas in the Caribbean region, while the latter is mainly distributed in the Pacific and Andean region, including the Sierra Nevada de Santa Marta at the Caribbean region, and a few spots predicted in the Amazonian region. *Marmosa alstoni* and *M. zeledoni* showed an Andean-Pacific distribution, with the former being more widely distributed in the Andean region compared to the latter, and with some areas in the Serranía de San Lucas (Caribbean). *Marmosa robinsoni* and *M. isthmica* were distributed mainly in the Caribbean region and the inter-Andean valleys, especially the Magdalena river valley, with the former having more predicted areas in the Andean region, and the latter with more predicted areas in the Pacific region. The distribution of *Monodelphis palliolata* was mainly concentrated in the Andean and Caribbean regions, limited to mid-low elevations at the northern end of the Eastern Andean region, Catatumbo, and low elevations of the Sierra Nevada de Santa Marta. Finally, *M. xerophila* with a restricted distribution, concentrated in the Guajira, north of the Caribbean region.

The areas of the final ranges varied from 681,717 km² in *M. waterhousei* to 8,526 km² in *M. rubra*, with a median area of 135,517 km². Taken together, the combined distribution of all Marmosini species' distribution covers 1,038,318 km² of continental Colombia, about 91.4% of the continental area of the country, being absent only in parts of the Orinoquia region, North to the Meta river and East of the Eastern Andes near the border with Venezuela, and two small portions one in the northwestern end of the Western Andes (Chocó department), and the other in the southwestern coast near Tumaco (Nariño department).

- **Richness and conservation metrics**

Richness of Marmosini in Colombia varied from a maximum of 10 species to a minimum of 1 species per 25 km² (Fig. 4). We found a clear pattern of maximum and sub-maximum richness concentrated at mid-elevation slopes of the Andes, with an approximate upper elevation limit of richness at 2000 m, especially for the Central and Eastern Andes (Fig. 4b). The highest richness (10-8 species) was found in the Amazon-Andes transition, east of Nudo de los Pastos, and southeast of the Colombian massif (Fig. 4c). The mid-high richness (7-5 species) was found mainly at the eastern slope of the Eastern Andes, Sierra de La Macarena, Catatumbo (northwest of the Táchira depression at the limit between Colombia and Venezuela in the Eastern Andes), Serranía de San Lucas northwest of the Central Andes, in the northern end of the Western Andes, the northwestern coast of the Pacific region and southern Amazonas in the region contained between the Putumayo and Caquetá rivers (Fig. 4a-b). The mid-low richness (4-2 species) was found mainly at the Amazon (north of Caquetá river), the transition zone between Amazonas and Orinoquia regions, inter-Andean valleys, Central Andes, high elevations of the Andes including Eastern Andes and most of its western slope, Pacific region (excluding areas in the northwest with mid-high richness), Sierra Nevada de Santa Marta, and much of the Caribbean region (Fig. 4a). Interestingly but with a lower richness compared to other areas, a local upper limit of high richness occurs at ~2000 m in Sierra Nevada de Santa Marta (Fig. 4a). The lowest richness (1 species) was found mainly in lowlands of the Orinoquia region, mid/low Magdalena river valley, Guajira, some areas of the Pacific region, and in the highest elevations of the Andes and Sierra Nevada de Santa Marta (Fig. 4b).

Of the total modeled areas, conservation within each species was highly variable, ranging from 29.9% in *M. rutteri* to 5.3% in *M. xerophila*, and a median of 15% (Table 2). This shows that roughly 85% of all distribution areas for Marmosini lack effective protection. Of the total preserved areas for each species, areas with strict conservation preserved between 28.9% in *M. rutteri*, and 2.3% in *M. xerophila*, with a median of 10.4% (Table 2), with large unprotected areas for all species. The areas preserved under managed resources ranged from 6.9% in *M. zeledoni* to 0% in *M. brevicaudata*, with a median value of 2.7% (Table 2). Human pressure within each species area ranged from 76.2% to 0% (*M. robinsoni* and *M. brevicaudata*, respectively) and a median of 27.9% for high pressure, and from 94.3% to 23% (*M. germana* and *M. robinsoni*) and a median of 71.1% for low pressure (Table 2). No data values regarding pressure varied from a maximum of 7.6% in *M. brevicaudata* to 0.3% in *M. phaea* (Table 2).

Our conservation-only analysis based on IUCN criteria showed that most of the species had more preserved area under strict conservation (median of 11.98%), than area under managed-resources

(median of 2.24%). Conservation based on governance showed that most protection comes from national governed areas (median of 11.66%), followed by sub-national governed areas (median of 3.83%), with private areas representing very little of the protected areas (median <0.1%) (Table 3). Among species, *M. rutteri* had the highest percentage of its area strictly protected while *M. xerophila* had the lowest, which were mostly under national governed areas. The percentage of sub-national governed areas was higher than national governed areas in *M. isthmica*, *M. robinsoni* and *M. xerophila* (the last two with values slightly above the national governed areas; Table 3). No protected areas of Indigenous territories and local communities were found throughout the distribution of Marmosini in Colombia.

Conservation-pressure analysis showed that regions within species ranges that are under strict-conservation have a lower area under high pressure (median of 0.54%) compared to areas under managed-resources (median of 38.9%) (Table 4). Within governance-types, the lower median values were for areas under national governance (1.06%), followed by sub-national areas (33.72%) and private areas (62.09%). This pattern is consistently found throughout all species of Marmosini analyzed (Table 4). These results show that Marmosini in Colombia are more exposed to higher human pressure in managed, sub-national and private areas, while strict reserves and national governance areas have smaller percentages of areas under high pressure throughout the species' ranges.

Discussion

To our knowledge, this is the first time that the distribution and conservation of species of the tribe Marmosini were assessed specifically and at a national scale (but see Gutiérrez et al. (2014)). Also, no other assessment of this kind was done within any other marsupial group in the country. In our study, we found that the species of Marmosini from Colombia have relatively few areas covered by the Colombian NPA network (Table 2). Moreover, only strict and national reserves represent desirable scenarios for their conservation, with most of their overlapped area having low human pressure (Table 3). Regions identified in this work with the greatest richness of Marmosini are also known as important zones for threatened and endemic vertebrates in the country (Kattan et al. 2004; Forero-Medina and Joppa 2010) giving support to the idea of Andean cordilleras as highly important spots for biodiversity and conservation in general (Myers et al. 2000), and for Marmosini species in particular.

- **Spatial patterns of biodiversity**

Our finding of the highest richness concentrated at mid-elevations of the Andean region is a common pattern found in other groups of plants and animals (Kattan et al. 2004; Forero-Medina and Joppa 2010; Rangel - Ch. 2015) but to our knowledge, this is the first time it is described for Marmosini. Yet, the prevalence of such richness in the eastern slope of the Eastern Andes needs to be explored and explained further on. This pattern is likely related to the many Cis-Andean taxa in this group (Gutiérrez et al. 2010; Voss and Giarla 2021), but also to other non Cis-Andean species contributing to this richness, which can be related to overlapping environments in western Amazonia (Online Resource 3). Recently, Clerici et al. (2019) discussed the importance of the Páramo and Imerí provinces (sensu (Morrone 2014)), with which

our region of highest and mid-high Marmosini richness overlaps. These authors mentioned the area as a priority corridor to be preserved for Amazonian-Andean biodiversity in Colombia, especially within the current social context of the country (Clerici et al. 2016, 2019). This area could have played an important role in the diversification and distribution of this group of marsupials, as in other mammals (Patterson and Velazco 2008).

Bax and Francesconi (2019) studied tropical Andes faunas, and found that areas in the Colombian Andes have high levels of irreplaceability and vulnerability. Forero-Medina and Joppa (2010) found that for mid-elevation biomes of the Andes ('Orobiomas medios de los Andes'), only 13% are covered by the NPA network. Moreover and specifically for marsupials, spatial data of conservation values (Martin et al. 2021) and phylogenetic diversity and taxonomic richness (Fergnani and Ruggiero 2015) reinforce the idea of the Andes as a critical region for the conservation of New World marsupials (including Marmosini) in the country. Our results and those above strengthen the argument of the Andes, and specifically its mid-elevations habitats, as of crucial importance for the conservation of Marmosini (Fig. 4). However, it is also important to acknowledge other regions such as the Pacific, Serranía de San Lucas, Sierra de La Macarena, and Sierra Nevada de Santa Marta (Fig. 4), all of these with different richness values.

Our data shows that Marmosini is clearly restricted to low and middle elevations below 2000 m. Besides the upper elevation limit at 2000 m, other natural limits identified in this report are worthy of mention, like mid-high richness areas at the Amazon which are bounded North by the Caquetá river (Fig. 4a). Similarly, two species in the Amazon presented almost perfectly complementary distributions (see *M. rutteri* and *M. brevicaudata* in Online Resource 3), separated by the Vaupés river. These results may suggest an important role of rivers as natural barriers for these species, a hypothesis to be tested for this group of marsupials in Colombia. Different studies have shown the importance of rivers as barriers to the dispersion of New World marsupials (and other mammals), especially in forested habitats (Myers 1982; Patton et al. 2000), in what is known as the "Riverine Barrier Hypothesis" (Wallace 1854). Our results add support to this geographic hypothesis, especially for the Amazonian region in which large and wide rivers occur.

- **Marsupials and environmental variables**

The models we generated showed a higher relationship between precipitation variables and the distribution of Marmosini, with temperature variables only important in the models for a few species. This pattern of precipitation related variables as the most important in ENMs was also found in *Thylamys pallidior*, *Dromiciops gliroides*, and *Rhyncholestes raphanurus* (Martin 2008, 2010, 2011), species from clearly different environments than those of Colombian Marmosini, and not close phylogenetically. Also, precipitation variables were found to be important in ENMs of the semi-aquatic *Chironectes minimus* (Prieto-Torres and Pinilla-Buitrago 2017), a tropical and subtropical species with an upper altitude limit of 2000 m. Although Birney and Monjeau (2003) described the mean minimum extreme temperature as a possible limiting factor for marsupial richness (especially outside the tropics), they acknowledged that

precipitation differences could be a surrogate for habitat heterogeneity, thus supporting a higher richness in some latitudinal bands or areas. Our findings that ENMs of Marmosini were mostly influenced by precipitation variables can be related to Colombia's high environmental heterogeneity, especially in areas bounded by the different Cordilleras (Londoño-Murcia et al. 2010). These precipitation and geographic variations, and their consequence in the habitats/environments of Colombia, might help explain the high biotic species richness, of which our findings of Marmosini richness is only an example.

- **Colombian Marmosini richness, conservation, and pressure**

We compiled information from 16 species of the 19 listed for the country (Solari et al. 2013; Ramírez-Chaves et al. 2016), which means that 15.78% of the species could not be evaluated due to lack of information. Importantly, species with no records in Colombia but that are cited in the current species' list of the country (e.g., *M. rubra* and *M. palliolata*, Ramírez-Chaves et al. 2016) were predicted to occur within the country's limits (Online Resource 3), adding support to the country's currently recognized species richness. In the last national and global conservation assessment which included these species (Rodríguez-Mahecha, et al. 2006; IUCN 2021), only two were listed as globally threatened (*M. phaea* and *M. xerophila*) and only one nationally threatened (*M. xerophila*). Given that the last national assessment took place more than 15 years ago, we expect that the information provided here can be used in the upcoming and much-needed conservation assessments.

Human pressure scenarios for the country are challenging. Correa Ayram et al. (2020), based on the Legacy-adjusted Human Footprint Index (LHFI), found that this index increased 50% from 1970 to 2015. Areas at the foothills of the Andes, especially in the eastern slope, are among the areas with the highest proportion of LHFI preserving less natural habitats. This is mostly related with deforestation fronts (Clerici et al. 2019; Correa Ayram et al. 2020) and their effects on habitat continuity. This has a direct impact in the majority of species considered here due to their largely arboreal habits, except for the three species of *Monodelphis* (Astúa 2015). Other regions as the Pacific and Caribbean, that also present high levels of LHFI, are areas of high and mid-richness of Marmosini species (Fig. 4). What is more challenging in these areas, is that a variety of human activities affect marsupials species in different ways. For example, deforestation is the main cause of habitat loss in the Andean-Amazonian foothills (Dávalos et al. 2011, 2014) while illegal mining for gold and other metals is an important cause for habitat degradation and loss in the Pacific region (Servicio Geológico Colombiano 2012; Correa Ayram et al. 2020). Illegal coca and the expansion of African oil palm plantations are among other important causes for habitat loss or modification (Dávalos et al. 2011).

In general, conservation-pressure scenarios for Marmosini species are complex, as is the case for many species in the country. Especially with the social and cultural context in which these problems are immersed (Clerici et al. 2016). In particular, we suggest that species as *M. xerophila* and *M. isthmica* (among others), with low percentages of their ranges cover by the NPA network and large areas of high pressure, should receive special attention in future conservation efforts.

There are many ways to study the distribution of species (Peterson et al. 2011) as well as many algorithms among ENM methods (Qiao et al. 2015; Phillips et al. 2017). In this study we chose a widely used algorithm that we expect can be used to make direct comparisons, as more and better data become available and/or other groups are assessed. In this sense, we made a methodological framework that takes into account most of the discussed caveats of maxent and ENM research (Anderson and Gonzalez 2011; Merow et al. 2013; Feng et al. 2019b; Zurell et al. 2020). Furthermore, other results from different models can be complementary and help clarify cases where our maxent models seemed to have not performed appropriately (i.e., *Marmosa phaea*). It is important to note that distribution maps from models are limited in their skill of predicting “true” species’ distribution, especially for recent events such as deforestation, fires, or other natural and human-related pressures affecting their distribution. It is then desirable that future field work corroborate or reject our final species range hypotheses.

Throughout this work, we highlighted the lack of studies on Colombian Marmosini in particular and marsupials in general, both at local and national scales. We expect to have shown one of the possibilities that arises from the current increase in taxonomic clarity within the group (Voss and Jansa 2009, 2019; Gutiérrez et al. 2010; Voss et al. 2020; Voss and Giarla 2021). A subject that we will keep working on and expanding for other related groups in the near future.

- **Conclusions**

Although Colombia hosts a very high species richness within the tribe Marmosini, the distribution (and biology) of most species is poorly known. Also, the NPA network of Colombia preserve little of the overall species’ distribution. In this work, areas with the highest species richness for the tribe were identified, also describing an upper limit of 2000 m to maximize conservation efforts for these species. We hope our work can be included with other taxa to help prioritize the creation of new conservation areas in Colombia.

References

Please see the supplementary files section to view the references.

Tables

Table 1. Performance metrics of maxent models for Marmosini species of Colombia. Metrics presented correspond to the chosen models after filtering results and visual inspection. See text for abbreviations.

Species	Area M	Cross-validation	Features	rm	Train AUC	Average test AUC	Average test or MTP	AICc
<i>Marmosa alstoni</i>	M2	random	LQ	0.5	0.88	0.81	0.11	798.3
<i>Marmosa ermana</i>	M2	block	LQ	4.5	0.81	0.78	0.13	279.32
<i>Marmosa sthmica</i>	M1	random	LQP	3	0.86	0.83	0.03	1553.8
<i>Marmosa jansae</i>	M2	block	LQ	0.5	0.82	0.77	0.13	278.55
<i>Marmosa lepida</i>	M2	random	LQ	1	0.83	0.81	0.03	885.4
<i>Marmosa phaea</i>	M2	block	L	2.5	0.89	0.89	0	262.13
<i>Marmosa regina</i>	M2	random	LQP	0.5	0.89	0.85	0.07	1642.89
<i>Marmosa robinsoni</i>	M2	random	LQHPT	1	0.93	0.9	0.01	5189.94
<i>Marmosa rubra</i>	M2	block	LQP	0.5	0.94	0.9	0	438.19
<i>Marmosa rutteri</i>	M2	random	LQP	5	0.71	0.7	0.07	916.72
<i>Marmosa waterhousei</i>	M2	block	LQ	0.5	0.86	0.77	0.17	724.91
<i>Marmosa xerophila</i>	M2	block	LQP	3	0.94	0.92	0.23	313.52
<i>Marmosa zeledoni</i>	M2	block	LQP	1.5	0.81	0.76	0	892.64
<i>Monodelphis dusta</i>	M1	random	LQ	2	0.82	0.81	0.08	688.17
<i>Monodelphis revicaudata</i>	M2	random	LQP	0.5	0.75	0.72	0.05	1113.98
<i>Monodelphis alliolata</i>	M2	random	LQ	0.5	0.94	0.92	0.02	994.73

Table 2. Range, conservation, and pressure area for Marmosini species of Colombia based on maxent models. All values are presented in km² and were calculated based on a geodesic approach in a WGS 84 projection. Area under protection categories are based on World Database on Protected Areas IUCN_CAT, while pressure categories are a discretization of pressure map for Colombia for the year 2015.

Species	Range area	Area under protection		Area under pressure ^a	
		Strict	Managed	High	Low
<i>Marmosa alstoni</i>	179,594	16,365	9,385	75,253	103,285
<i>Marmosa germana</i>	85,206	18,963	82	2,973	80,324
<i>Marmosa isthmica</i>	201,804	5,727	12,219	132,089	67,709
<i>Marmosa jansae</i>	120,495	23,942	42	7,077	110,265
<i>Marmosa lepida</i>	553,394	115,092	1,092	44,513	502,743
<i>Marmosa phaea</i>	322,640	39,343	18,246	148,150	173,610
<i>Marmosa regina</i>	46,498	8,386	5	5,071	40,905
<i>Marmosa robinsoni</i>	150,540	6,871	7,389	114,656	34,591
<i>Marmosa rubra</i>	8,526	480	58	2,924	5,507
<i>Marmosa rutteri</i>	314,613	90,961	292	16,309	295,067
<i>Marmosa waterhousei</i>	681,718	122,611	15,271	147,370	529,166
<i>Marmosa xerophila</i>	12,419	280	383	4,386	7,556
<i>Marmosa zeledoni</i>	45,270	3,976	3,102	8,819	35,511
<i>Monodelphis adusta</i>	380,587	46,593	21,952	159,858	218,789
<i>Monodelphis brevicaudata</i>	28,077	3,303	0	8	25,943
<i>Monodelphis palliolata</i>	23,799	1,667	477	15,249	8,402

^a Based on Correa Ayram et al., (2020) human pressure index.

Table 3. Analysis of protection type areas for Marmosini species of Colombia based on maxent models. Strict conservation category corresponds to Ia, Ib, and II, while Managed-resources corresponds to the remaining categories of the World Database on Protected Areas. All values are presented as percentages (%) of the total range area of each species.

Species	IUCN		Governance		
	Strict	Managed	National	Sub-national	Private
<i>Marmosa alstoni</i>	9.11	5.23	8.46	5.82	0.05
<i>Marmosa germana</i>	22.26	0.10	22.35	0.00	0.00
<i>Marmosa isthmica</i>	2.84	6.05	3.75	5.08	0.11
<i>Marmosa jansae</i>	19.87	0.04	20.28	0.00	0.01
<i>Marmosa lepida</i>	20.80	0.20	25.78	0.14	0.02
<i>Marmosa phaea</i>	12.19	5.66	11.57	6.50	0.09
<i>Marmosa regina</i>	18.03	0.01	26.68	0.00	0.01
<i>Marmosa robinsoni</i>	4.56	4.91	4.62	4.73	0.12
<i>Marmosa rubra</i>	5.62	0.68	6.95	0.03	0.04
<i>Marmosa rutteri</i>	28.91	0.09	37.70	0.00	0.01
<i>Marmosa waterhousei</i>	17.99	2.24	21.88	2.31	0.05
<i>Marmosa xerophila</i>	2.25	3.08	2.56	2.77	0.00
<i>Marmosa zeledoni</i>	8.78	6.85	10.41	5.15	0.07
<i>Monodelphis adusta</i>	12.24	5.77	13.44	6.05	0.10
<i>Monodelphis brevicaudata</i>	11.76	0.00	11.76	0.00	0.00
<i>Monodelphis palliolata</i>	7.00	2.01	6.04	2.93	0.04

Table 4. Analysis of conservation-pressure areas for the Marmosini species of Colombia based on maxent models. For each species, values are presented as the percentage of the protected area of each protection category that fall under high (clean value) and low pressure (square brackets). Values that do not add up to 100% mean there are no data values for that category.

Species	IUCN			Governance	
	Strict	Managed	National	Sub-national	Private
<i>Marmosa alstoni</i>	4.44 [94.14]	38.94 [59.53]	5.09 [93.28]	34.06 [64.78]	50.39 [43.01]
<i>Marmosa germana</i>	0 [99.7]	1.24 [98.08]	0 [99.7]	0	34.87 [55.13]
<i>Marmosa isthmica</i>	7.24 [88.61]	51.09 [46.3]	10.56 [83.39]	55.52 [43.29]	76.53 [18.32]
<i>Marmosa jansae</i>	0.05 [99.11]	23.1 [74.19]	0.03 [99.43]	0 [98.49]	82.28 [14.25]
<i>Marmosa lepida</i>	0.13 [99.65]	19.65 [79.49]	0.14 [99.74]	15.88 [83.05]	26.22 [70.83]
<i>Marmosa phaea</i>	5.43 [94.32]	41.88 [57.89]	6.12 [93.58]	35.03 [64.84]	51.24 [43.24]
<i>Marmosa regina</i>	0.08 [98.9]	95.9 [0]	0.06 [98.83]	0	95.9 [0]
<i>Marmosa robinsoni</i>	9.29 [87.77]	57.19 [39.78]	8.16 [87.61]	58.19 [40.09]	83.75 [10.99]
<i>Marmosa rubra</i>	0.06 [97.34]	3.58 [95.3]	0.05 [97.49]	0 [99.41]	56.04 [35.05]
<i>Marmosa rutteri</i>	0.07 [99.73]	17.76 [81.92]	0.08 [99.83]	0	89.02 [8.55]
<i>Marmosa waterhousei</i>	0.4 [99.3]	42.63 [56.73]	0.69 [99.1]	36.62 [63.06]	61.78 [34.89]
<i>Marmosa xerophila</i>	15.86 [72.77]	74.84 [16.22]	25.3 [63.95]	72.72 [18.04]	0
<i>Marmosa zeledoni</i>	0.69 [96.73]	22.31 [73.46]	1.44 [94.67]	27.45 [70.46]	33.46 [61.54]
<i>Monodelphis adusta</i>	3.36 [95.85]	38.2 [60.73]	3.94 [95.06]	33.39 [66.16]	62.41 [33.22]
<i>Monodelphis brevicaudata</i>	0 [98.78]	0	0 [98.78]	0	0
<i>Monodelphis palliolata</i>	8.53 [88.28]	42.35 [56.65]	8.62 [87.94]	30.41 [68.45]	92.81 [0.78]

Figures

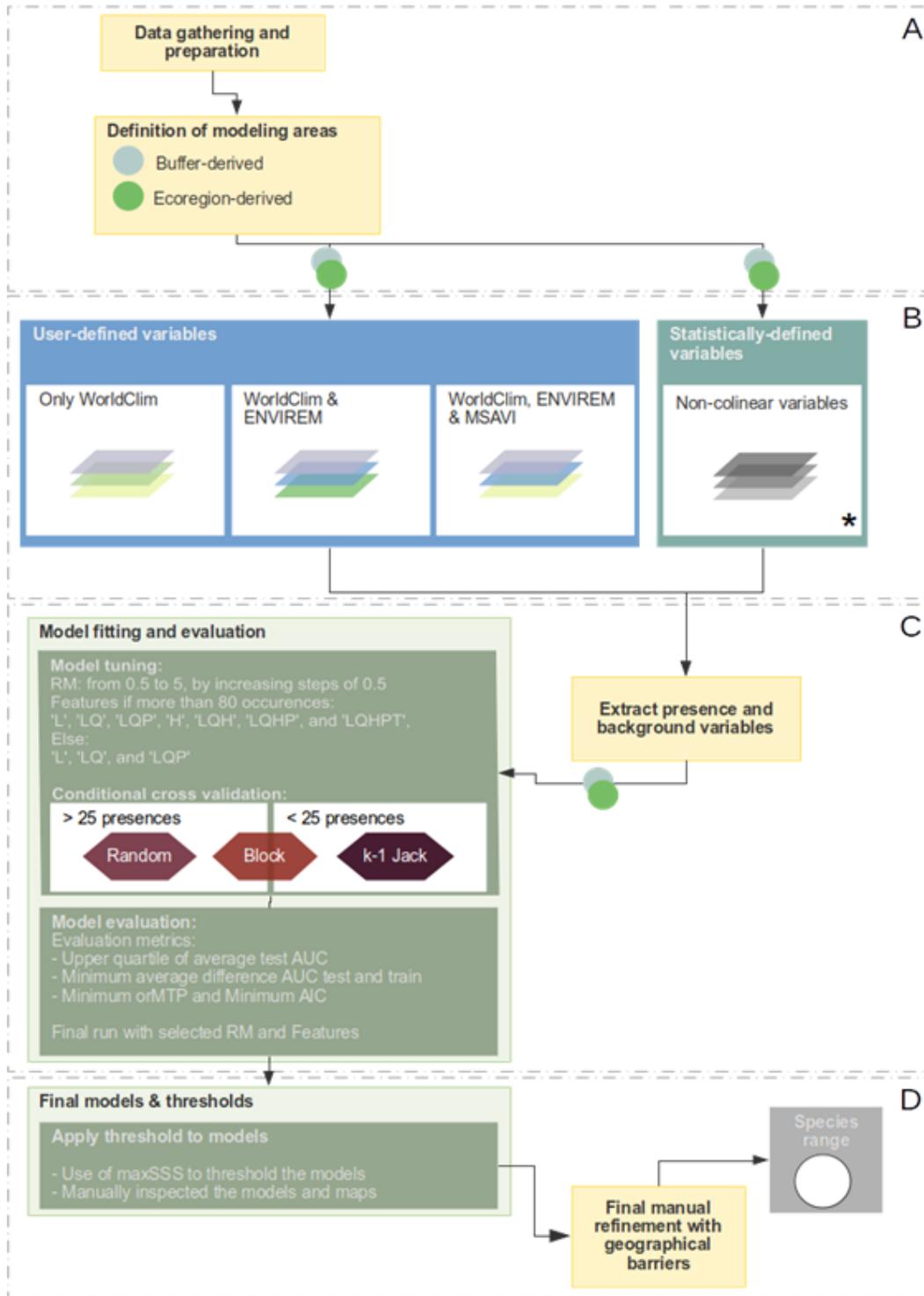


Figure 1

Workflow scheme for modeling distribution maps of the Marmosini species of Colombia. First step (a) is standard for most modeling exercises but note that the ‘Definition of modeling areas’ box presents two circles depicting two ways of estimating modeling areas (area M). For clarity, we depict the two same circles along the scheme to denote where models are being estimated for both areas. Second step (b) included estimation of four predictors scenarios. Although modeling is summarized by one arrow from b

to c, models are run for each scenario and for each area M until 'Model evaluation'. The third step (c) was model tuning, fitting, and evaluation. For each species, models were evaluated according to four metrics; models chosen passed to the next step. Last step (d) included the application of a threshold, visual evaluation and establishment of geographical barriers for final modifications. Note that the final arrow do not represent area M circles (M1 or M2), since each final species range is based on only one of them. Asterisk (*) in 'Statistically defined variables' is to clarify that correlation was evaluated for each species and for each area M (i.e., 32 sets of uncorrelated variables).

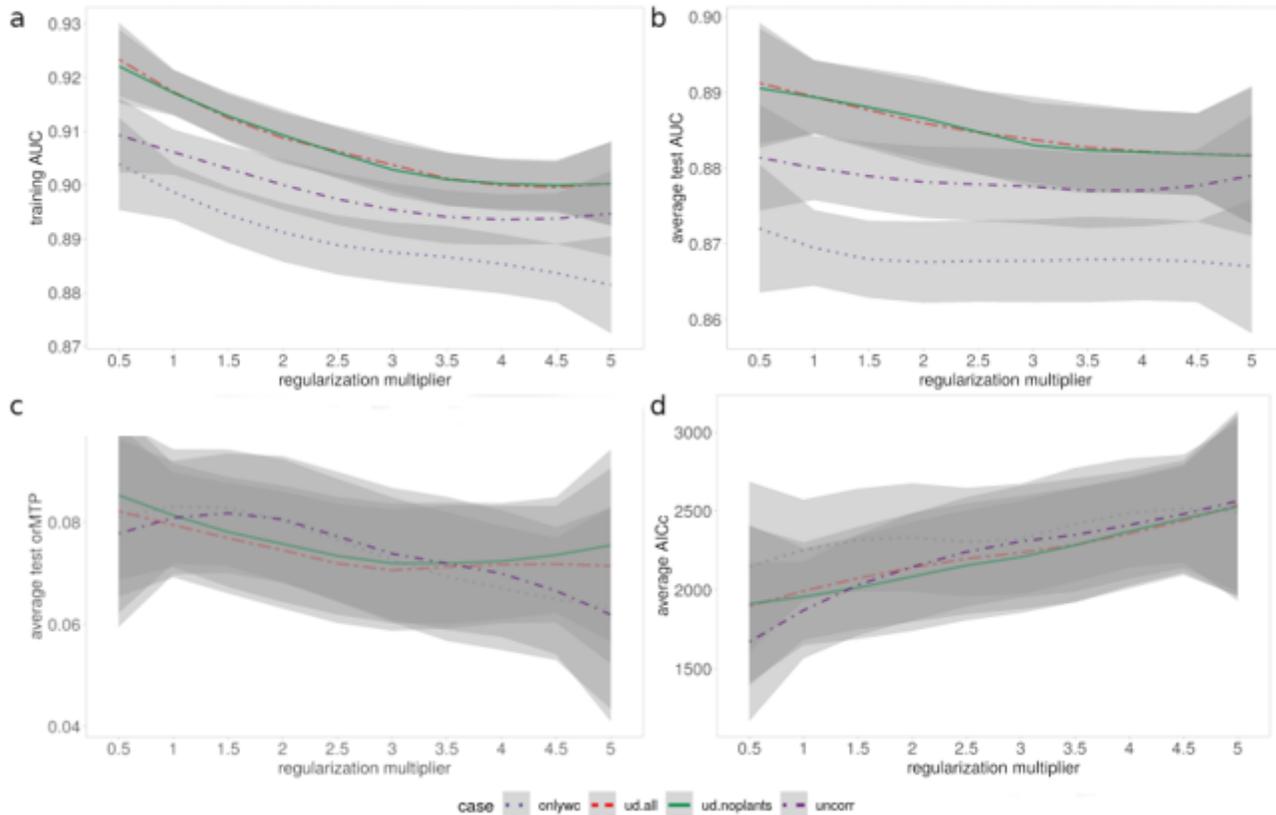


Figure 2

Performance of different predictors scenarios for maxent models of Marmosini species of Colombia, based on regularization multiplier and training AUC (a), average test AUC (b), average test orMTP (c), and corrected Akaike Information Criterion (d). Line represents a local polynomial regression among each predictors' scenario (case), with standard error represented by the surrounding shaded area of each line. Case abbreviations: onlywc, only WorldClim data; ud.all, user-defined variables including MSAVI; ud.noplants, user-defined variables excluding MSAVI; and uncorr, uncorrelated variables.

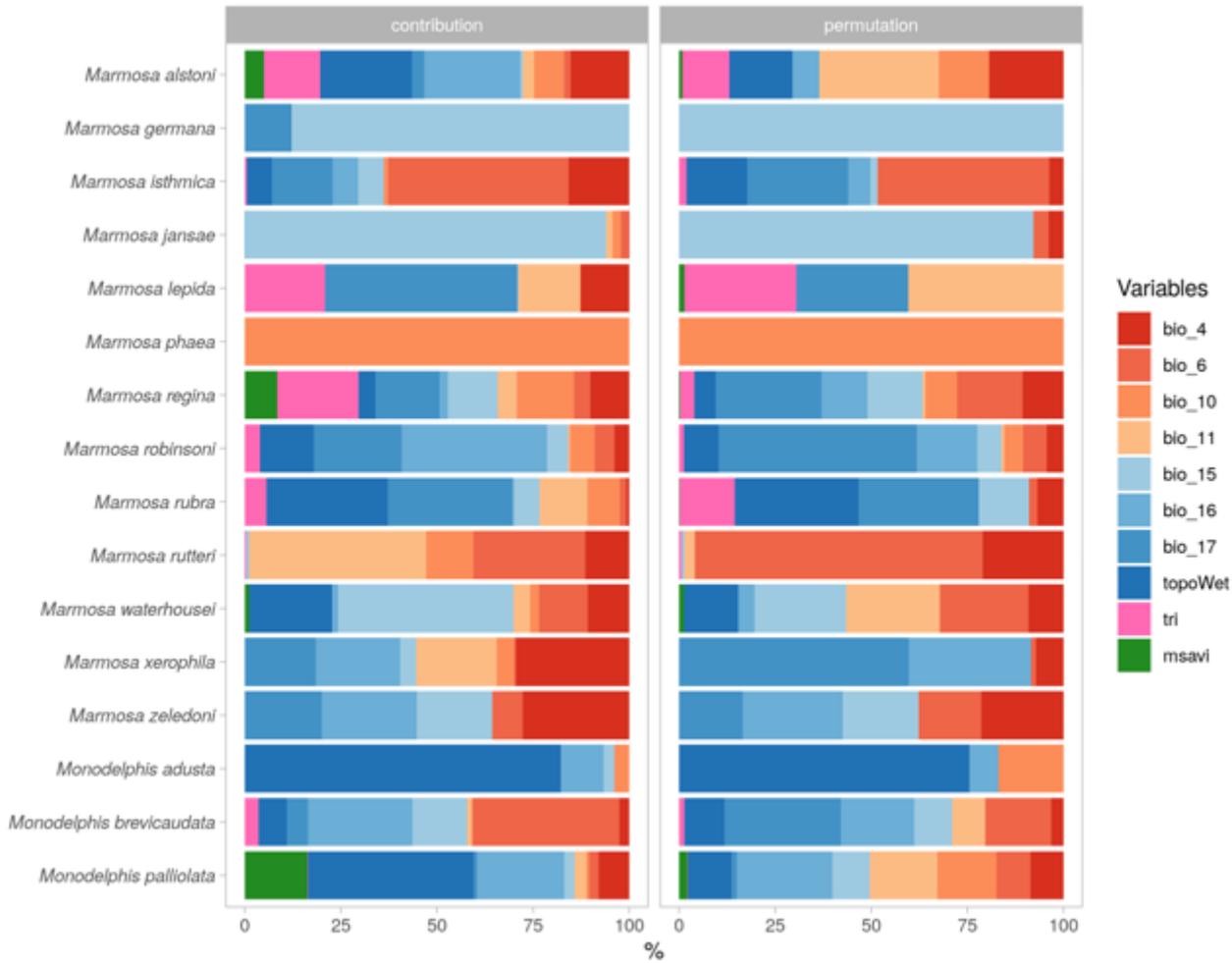


Figure 3

Percentage of contribution and permutation importance for maxent models of Marmosini species of Colombia. Species in the y-axis are ordered alphabetically. Variable definitions: bio_2, mean diurnal range; bio_4, temperature seasonality; bio_6, min temperature of the coldest month; bio_10, mean temperature of the warmest quarter; bio_11, mean temperature of the coldest quarter; bio_15, precipitation seasonality; bio_16, precipitation of the wettest quarter; bio_17, precipitation of the driest quarter; topowet, SAGA-GIS topographic wetness index; tri, terrain roughness index; and MSAVI, modified soil-adjusted vegetation index.

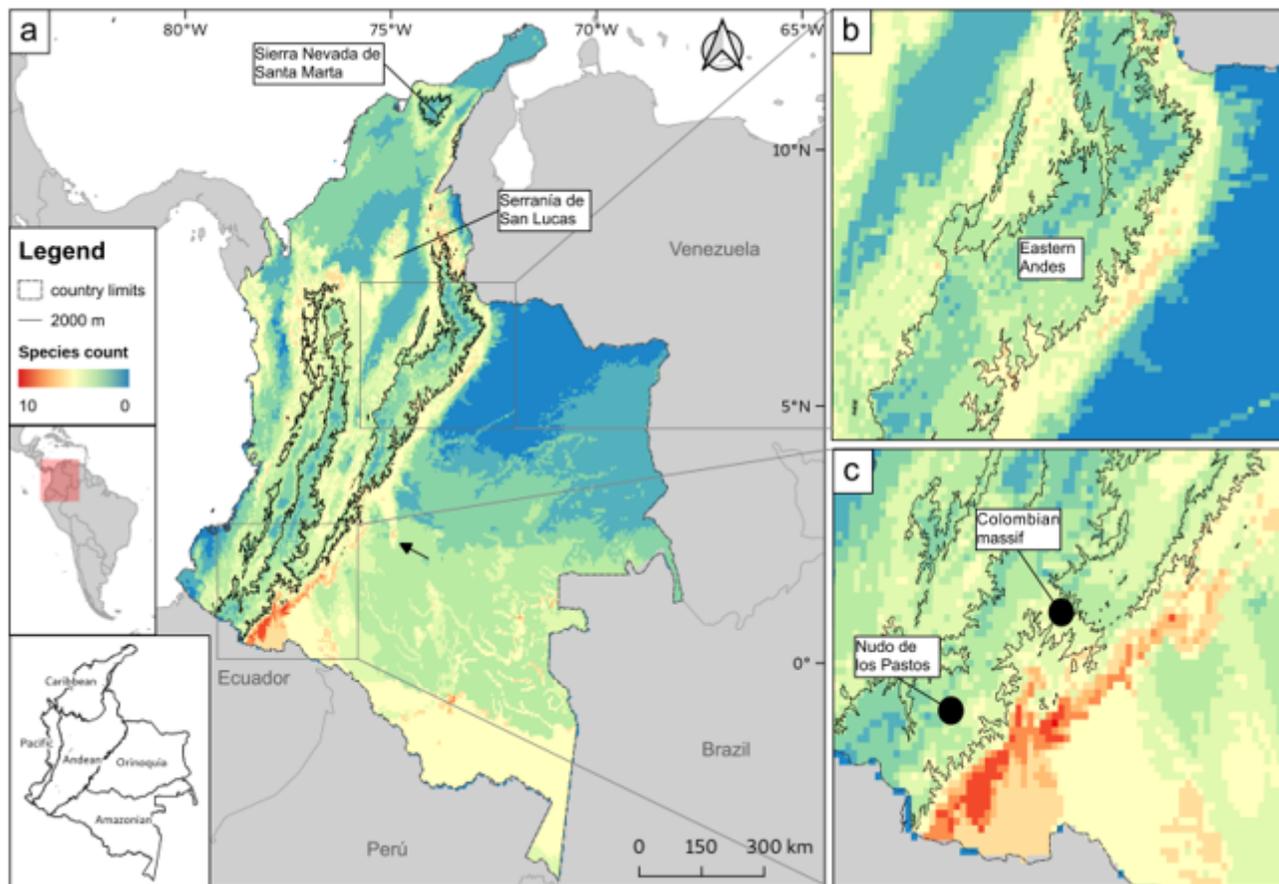


Figure 4

Spatial richness of Marmosini species from Colombia at a 25 km² pixel-size (a), note that for the Andean region, especially for Eastern Andes, Marmosini richness has a sub-maximum value that follows an elevation belt of 2000 m (b), while maximum value is reached south of the country, east to the Nudo de Los Pastos formation (c). Black arrow signals Sierra de La Macarena. Map is depicted in a Mollewide projection.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM13.pdf](#)
- [appendix1.xlsx](#)
- [References.docx](#)