

Supplement S2

Gonzalez et al. 2022. Updated distribution and conservation perspectives of marmosine opossums from Colombia.

SUPPLEMENT S2: DETAILED METHODS AND PERFORMANCE METRICS

Methods. ---Since each model should use appropriate background data for the modeling area, for each species and each area M (see below), a unique set of 10,000 background points was generated to represent the environmental data in each predictor's scenario. That is, background points varied between types of area M but not between predictor's scenarios, to allow a more straightforward comparison (specially when using AICc), when choosing the best models across scenarios.

Predictors scenarios come from a subselection of current climatic conditions represented by subsets of WorldClim v. 2 (Fick and Hijmans, 2017), ENVIREM (Title and Bemmels, 2018), and one vegetation index, the Modified Soil Adjusted Vegetation Index (MSAVI). From all the above-mentioned variables we selected 8 WorldClim and 2 ENVIREM, based on those that have explanation power for marsupials (Prieto-Torres and Pinilla-Buitrago, 2017; Tocchio et al., 2015; Gutiérrez et al., 2014; Martin, 2011, 2010), 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.

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 ‘onlywc’, ‘ud.noplants’, and ‘ud.all’ scenarios, predictors were chosen according to our informed criteria of potential explanation power (Fourcade et al., 2018) and not tested for collinearity.

Barve et al. (2011) discussed how the modeling area from which predictors are sampled for background points affects modeling results. Here we used two methods to estimate this area, herein referred as area M, after the BAM diagram from Peterson and Soberón (2012). The M area represents a geographic space with suitable abiotic conditions historically available for the species to disperse (including currently occupied and unoccupied areas) (Barve et al., 2011). Ideally, information from fossils are used to estimate these areas but in the absence of them, buffered (Tocchio et al. 2015; Gutiérrez et al. 2014) or ecoregion (Prieto-Torres and Pinilla-Buitrago, 2017) methods have been used in these species. In the absence of information of their comparative performance, we used an adaptation of both approaches and compared them according to four evaluation metrics.

Below is the example of two methods for estimating modeling areas (area M): a simple approach where point localities are buffered by $\sim 330 \text{ km}^2$ generating a minimum convex polygon of the buffers (Fig. S1, left), and ecoregion-derived area where a minimum convex polygon from the localities is buffered by $\sim 55 \text{ km}^2$ and overlapped to ecoregions (Fig. S1, right). Note how the ecoregion-derived method represents natural continuous areas, while the simple approach delimits the model area arbitrarily.

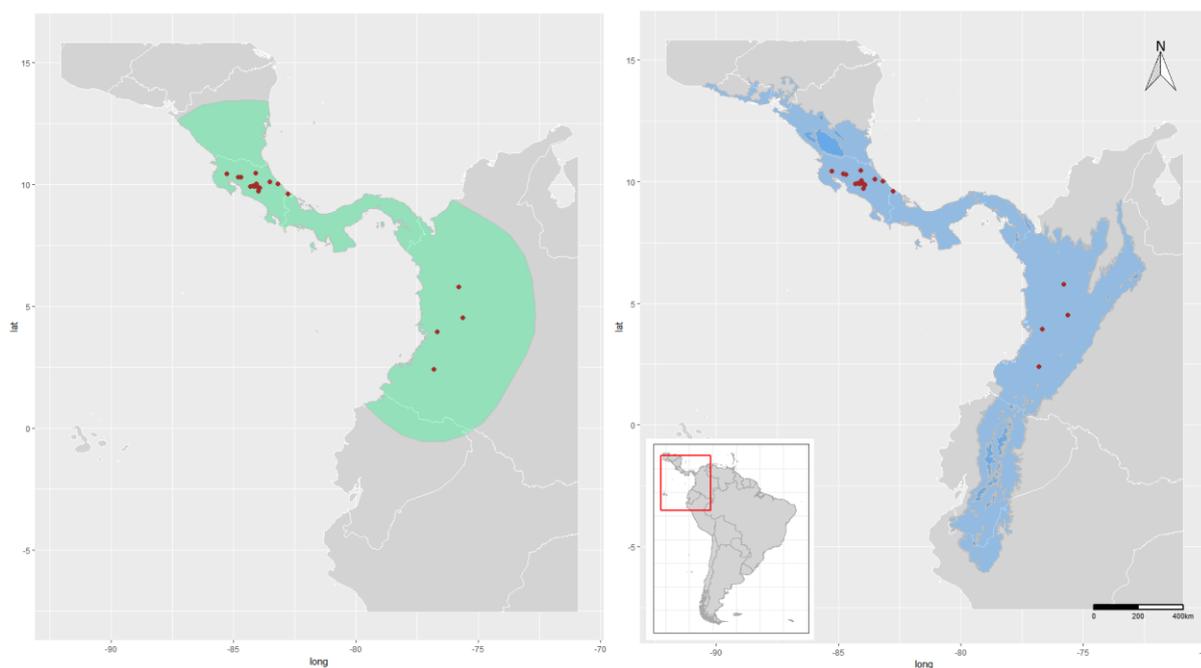


Fig. S2.1. Two modeling areas for maxent models of *Marmosa alstoni*. Simple buffer-derived area (left, green) represents a common practice in maxent models, while ecoregion-derived area (right, blue) represents an alternative to the latter method.

After applying thresholds to the models, we used the following protocol to convert raster predictions to polygons that represent the species ranges. First, we downscaled the resolution from 1 km² to 4 km² to avoid zones with many isolated or scattered presence cells. Then, prediction rasters were converted to polygon data using GDAL and their limits were smoothed using the `ksmooth` function from the R package ‘`smoothr`’ (Strimas-Mackey, 2020), by a smoothing index of 2. Additionally, holes that were less than 100 km² and crumbs that were less than 50 km² were removed.

Performance metrics. ---Summary of the performance metrics for the models used for final range construction of each species in this study (Table S1.1) and overall performance metrics across all models (Fig. S1.2)

Table S2.1. Performance metrics of maxent models for Marmosini species of Colombia.

Metrics presented correspond to the chosen models after filtering results and visual inspection. Abbreviations: CV, cross-validation; F, features; rm, regularization multiplier; t-AUC, training AUC; testAUC, average test AUC; orMTP, average omission rate at the minimum training presence; AICc, corrected Akaike Information Criterion.

Species	Area M	CV	F	rm	t-AUC	testAUC	orMTP	AICc
<i>Marmosa alstoni</i>	M2	random	LQ	0.5	0.88	0.81	0.11	798.3
<i>Marmosa germana</i>	M2	block	LQ	4.5	0.81	0.78	0.13	279.32
<i>Marmosa isthmica</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 adusta</i>	M1	random	LQ	2	0.82	0.81	0.08	688.17
<i>Monodelphis brevicaudata</i>	M2	random	LQP	0.5	0.75	0.72	0.05	1113.98
<i>Monodelphis palliolata</i>	M2	random	LQ	0.5	0.94	0.92	0.02	994.73

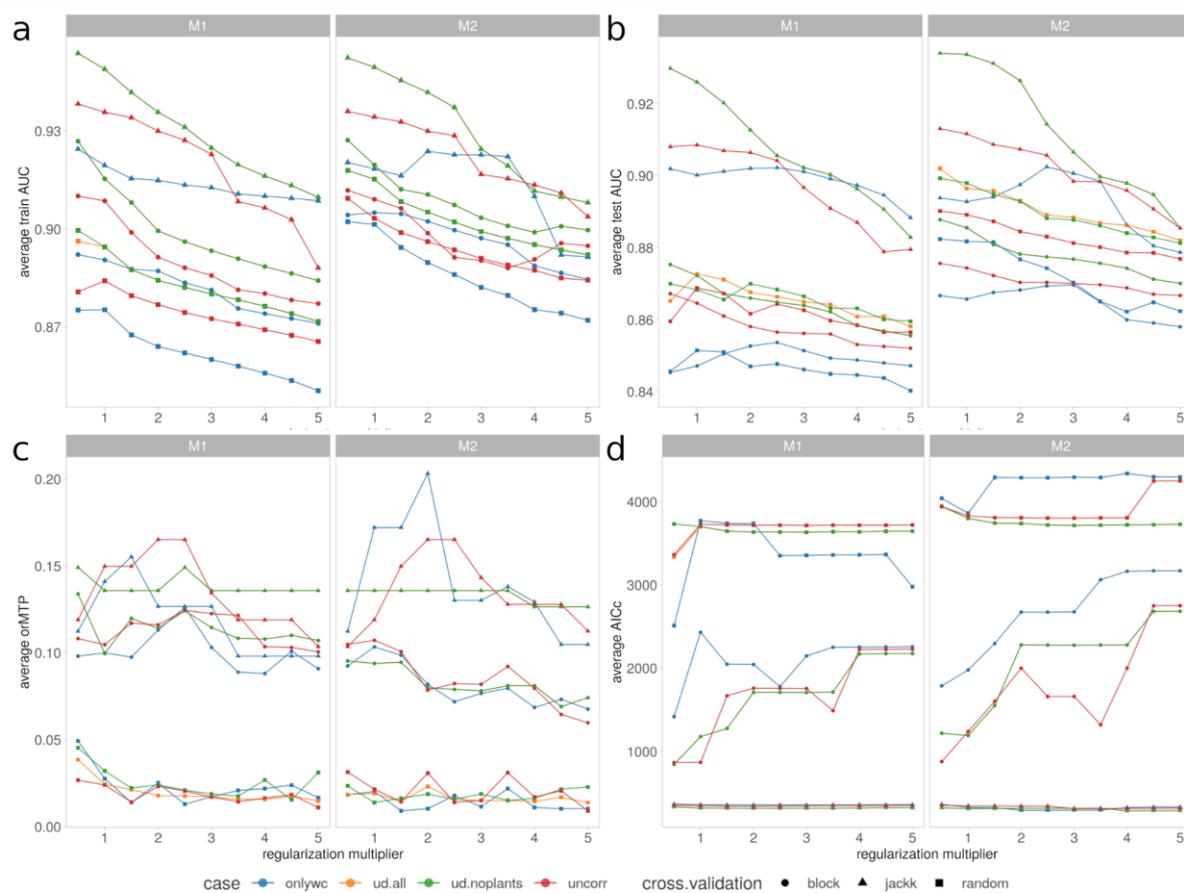


Fig. S2.2. Performance metrics of Maxent models for the Marmosini species of Colombia, contrasting between two modeling areas. M1 refers to a buffer-derived area and M2 to an ecoregion-derived area. Graphics are based on the models that are among the upper quartile of the test AUC metric, and show the mean value among grouped results based on predictors scenarios (case) and cross-validation regarding train AUC (a), test AUC (b), orMTP (c), and AICc (d).

LITERATURE CITED

- Barve N., Barve V., Jiménez-Valverde A., Lira-Noriega A., Maher S.P., Peterson A.T., Soberón J., Villalobos F., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol. Model.* 222(11): 1810–1819. doi:10.1016/j.ecolmodel.2011.02.011.

- Fick S.E., Hijmans R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37(12): 4302–4315. doi:10.1002/joc.5086.
- Fourcade Y., Besnard A.G., Secondi J., 2018. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Glob. Eco. Biogeography* 27:245–256.
- Gutiérrez E.E., Boria R.A., Anderson R.P., 2014. Can biotic interactions cause allopatry? Niche models, competition, and distributions of South American mouse opossums. *Ecography* 37(8): 741–753. doi:10.1111/ecog.00620.
- Martin G.M., 2010. Geographic distribution and historical occurrence of *Dromiciops gliroides* Thomas (Metatheria: Microbiotheria). *J. Mammal.* 91(4): 1025–1035. doi:10.1644/09-MAMM-A-347.1.
- Martin G.M., 2011. Geographic distribution of *Rhyncholestes raphanurus* Osgood, 1924 (Paucituberculata : Caenolestidae), an endemic marsupial of the Valdivian Temperate Rainforest. *Aust. J. Zool.* 59(2): 118–126. doi:10.1071/ZO11038.
- Peterson A.T., Soberón J., 2012. Species Distribution Modeling and Ecological Niche Modeling: Getting the Concepts Right. *Nat. Conserv.* 10(2): 102–107. doi:10.4322/natcon.2012.019.
- Prieto-Torres D.A., Pinilla-Buitrago G., 2017. Estimating the potential distribution and conservation priorities of *Chironectes minimus* (Zimmermann, 1780) (Didelphimorphia: Didelphidae). *Therya* 8(2): 131–144. doi:10.12933/therya-17-478.
- Strimas-Mackey M., 2020. smoothr: Smooth and Tidy Spatial Features. R package version 0.2.1. <https://CRAN.R-project.org/package=smoothr>.
- Title P.O., Bemmels J.B., 2018. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* 41(2): 291–307. doi:10.1111/ecog.02880.

Tocchio L.J., Gurgel-Gonçalves R., Escobar L.E., Peterson A.T., 2015. Niche similarities among white-eared opossums (Mammalia, Didelphidae): Is ecological niche modelling relevant to setting species limits? *Zoologica Scripta* 44:1–10.