REMOTELY-SENSED VEGETATION AND HABITAT STRUCTURE CAN SERVE AS SUITABLE SURROGATES TO PREDICT THE DISTRIBUTION OF A MICRO-ENDEMIC BIRD SPECIES

Edwin Zárate, Christine Wallis, Vinicio Santillán, Roland Brandl, Nina Farwig and Jörg Bendix

With 5 figures, 5 tables and appendix
Received: 18 September 2025 · Accepted: 07 May 2026

Summary: Biodiversity is increasingly threatened by human land use and climate change, making predictive modeling crucial for conservation planning and the identification of priority conservation areas. Large-scale species distribution projections can be improved by integrating remotely sensed vegetation indices, as they reflect important vegetation and habitat characteristics. The Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) indicate vegetation productivity, while texture metrics derived from these indices reveal habitat structure. When combined with climatic variables, these indicators can significantly improve the accuracy of species distribution models (SDMs). In this study, we evaluated the performance of SDMs using vegetation indices, texture metrics, and climatic variables to predict the distribution of the Violet-throated Metaltail (*Metallura baroni*), a microendemic hummingbird restricted to the environmentally complex high-altitude regions of the southern Ecuadorian Andes. Using a backward-selection and cross-validation approach for predictor selection and the Maximum Entropy (MaxEnt) algorithm, we compared model performance using AUC values. Our results demonstrate that incorporating habitat structure indicators (NDVI- and NDWI-derived texture metrics) together with climatic variables significantly improves SDM performance, allowing better discrimination of shrub ecosystems where *M. baroni* occurs. This approach highlights the importance of integrating key aspects of habitat structure as indicators of resource availability, which directly influence the distribution of bird species in complex landscapes.

Keywords: Spectral vegetation indices, highlands, Maxent, remote sensing, species distribution models, habitat, texture metrics

1 Introduction

Biodiversity is facing unprecedented threats from human land use and climate change, which are driving habitat loss, fragmentation, and shifts in species distributions (He et al. 2019). In this context, predictive modeling of biodiversity distributions has become a critical tool for developing effective conservation strategies and identifying priority areas for protection (Onoh et al. 2024). Among the most widely used tools for this purpose are Species Distribution Models (SDMs), which predict species' geographic ranges by combining occurrence data with environmental predictors such as climate, habitat structure, and topography (Elith et al. 2006, Franklin 2023). These models are grounded in ecological niche theory, which posits that species distributions are shaped by the interplay of abiotic and biotic factors within their fundamental and realized niches (Hutchinson 1957, 1959).

The Niche concept encompasses diverse aspects of ecology, evolution, and conservation biology. It refers to the diversity of resources or environments used by an individual, population, species, or clade (Hutchinson 1959, MacArthur & MacArthur 1961). However, biotic interactions and resource availability further constrain species distributions within their fundamental niche (Lembrechts & Lenoir 2019, Colwell & Rangel 2009). SDMs leverage this theoretical framework to predict species distributions by correlating occurrence data with environmental predictors, providing valuable insights into habitat suitability and species-environment relationships (Carscadden et al. 2020, Guisan & Zimmermann 2000).

Due to the importance of climatic variables for species distributions, especially in mountain systems (Santillán et al. 2018), species distribution models have traditionally been based on coarse-grained climate data (e.g., WorldClim) with resolutions above 0.5 km², which often fail to capture microclimatic variability and are inadequate for modelling species with restricted ranges smaller than 0.1 km² (Tomlinson et al. 2020). However, recent advances in remote sensing and habitat modelling have enabled the development of high-resolution predictors, such as those derived from

https://doi.org/10.3112/erdkunde.2026.02.04 ISSN 2702-5985 (Online)

LiDAR (Acebes et al. 2021, Farrell et al. 2013) or hyperspectral satellites imagery and vegetation indices, which significantly improve the ability of habitat models to predict species distributions in topographically complex regions (Barry 2008, Barbet-Massin & Jetz 2014). Vegetation indices, such as the normalized difference vegetation index (NDVI) and the normalized difference water index (NDWI), provide quantitative measures of vegetation density, productivity and moisture content, and serve as indicators of habitat quality and resource availability (GAO 1996, Pettorelli et al. 2011). Furthermore, texture metrics derived from gray-level co-occurrence matrices (GLCMs) capture fine-scale habitat heterogeneity, which is critical for species that rely on specific habitat structures (Haralick et al. 1973, Wallis et al. 2017, Farwell et al. 2020). Habitat structure and heterogeneity are aspects closely linked to the availability of resources that support bird populations (Mills et al. 2023). The integration of vegetation proxies with climatic variables has proven effective in modeling species distributions, particularly for critically endangered species such as the kakapo (*Strigops habroptilus*) (Fisher et al. 2019) and the California condor (*Gymnogyps californianus*), (Phillips 2005). Wood et al. (2013) used image texture to predict avian density and species richness. In the Andes of South America, studies have successfully modeled the ecological niches of endemic hummingbirds using climatic variables and NDVI (Rodríguez & Bonaccorso 2016, Tinoco et al. 2023). However, for microendemic species with highly restricted distributions, SDMs that incorporate both habitat structure and climatic variables remain scarce.

This study aims to evaluate the performance of SDMs by combining habitat predictors such as climate (temperature, precipitation), vegetation indices (NDVI, NDWI) and habitat structure metrics (textural metrics) to model the distribution of *Metallura baroni*, a microendemic hummingbird restricted to the high-altitude regions of the southern Andes of Ecuador. We hypothesize that the inclusion of habitat structure proxies will significantly improve model accuracy, enabling better discrimination of the species' habitat conditions (Rahbek & Graves 2000, Wallis et al. 2017, Marston et al. 2023). By integrating high-spatial-resolution remote sensing data with traditional climatic predictors, this study seeks to advance our understanding of species-environment relationships in complex ecosystems and inform conservation strategies for range-restricted species.

2 Methods

Workflow for constructing and evaluating species distribution models (SDMs) for *M. baroni* using environmental predictors. The process begins with data preparation, including the review and filtering of occurrence records, acquisition of bioclimatic and Landsat imagery, and the calculation of vegetation (NDVI) and water (NDWI) indices. Predictors are selected and refined to address collinearity, followed by the construction of four MaxEnt models combining bioclimatic variables with NDVI, NDWI, or both. Finally, the models are evaluated to assess their performance.

2.1 Study area

The area where *M. baroni* inhabits is a tropical high-mountain zone characterized by environmental complexity due to steep elevation gradients. These conditions produce a variety of montane ecosystems (Tinoco et al. 2009, Carrasco et al. 2022). Specifically, *M. baroni* is found in the western and eastern Andean ranges (Fig. 1), above 3000 m above sea level in the central-southern Ecuadorian provinces of Azuay, Cañar, and Morona Santiago (Tinoco et al. 2009, Carrasco et al. 2022). The area is mainly covered by montane forest and páramo, the latter including a wide variety of vegetation types (scrub, *Polylepis* forest, herbaceous páramo) (MAE 2013, Carrasco et al. 2022). The climate is typical of the high tropical mountains of the Andes, where temperature and precipitation are determined by altitude, solar exposure and synoptic forcing (Navarrete et al. 2022). The mean monthly air temperature ranges from 5 to 12 °C, and the daily air temperature can fluctuate considerably from 0 to 20 °C. Annual precipitation averages between 1200 and 1500 mm annually (Córdova et al. 2018).

2.2 Occurrence database

Records of *M. baroni* occurrence in Ecuador were obtained from the Global Biodiversity Information Facility (GBIF, 2021) using the gbif function of the *dismo* package in R (Hijmans et al. 2024). Initially, 356 records were collected and subjected to statistical filtering (Sullivan et al. 2014) to reduce potential autocorrelation among points (Aiello-Lammens et al. 2015, Feng et al. 2019). After filtering, a total of 63 occurrence records were retained for SDM devel-

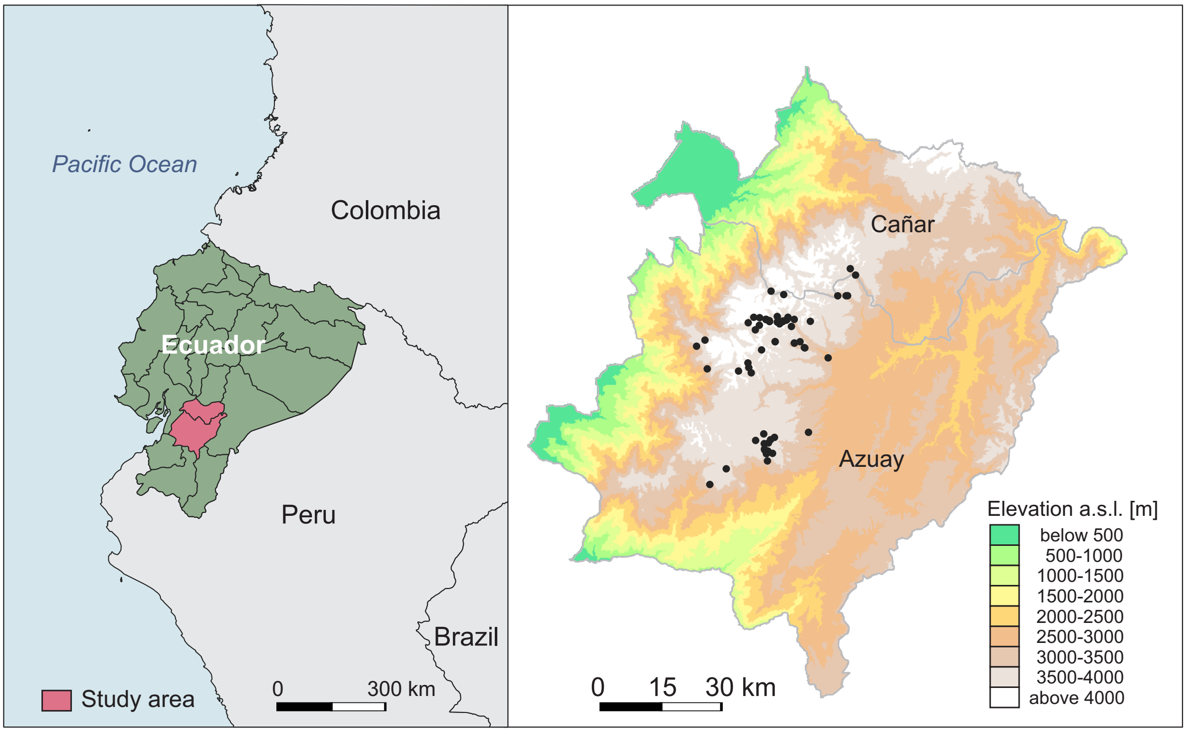

Fig. 1: Study area in the Azuay and Cañar provinces, southern Ecuador. The black dots indicate occurrence records of *M. Baroni* from Global Biodiversity Information Facility (GBIF) through *gbif* function of the *dismo* R package (Hijmans et al. 2024).

opment using the MaxEnt platform. This number is considered sufficient, as MaxEnt performs well even with relatively small sample sizes (van Proosdij et al. 2016, Cavalheri et al. 2024).

2.3 Climate predictors

Climate information was obtained from WorldClim (Worldclim 2021), a spatially interpolated monthly climate dataset for global land areas at a spatial resolution of 30 seconds (~1 km²), based on records from 9,000–60,000 weather stations worldwide (Fick & Hijmans 2017). This spatial resolution corresponds to one of the finest scales currently available for global bioclimatic variables and has been widely used in species distribution modeling studies (Fick & Hijmans 2017). The predicted distribution based on these data is also broadly consistent with the known distribution range for *Metallura baroni* (Tinoco et al. 2009, Carrasco et al. 2022); however, higher-resolution climate data would be preferable for species with such restricted distributions. Regional initiatives, such as Enhancing Adaptive Capacity of Andean Communities through Climate Services (ENANDES) initiative, aim to generate higher-resolution climatic datasets for the Andean region; however, these products are still in the implementation phase and therefore are not yet available for download (CRC-OSA/CIIFEN 2024, CIIFEN n.d.). Hence, WorldClim is the most accessible dataset currently available.

The full set of 19 climatic variables available from WorldClim expresses spatial variation in annual means, seasonality, and interactions among climatic factors (Stoica 2018). Among these, temperature and precipitation are particularly relevant for the distribution and ecology of birds in high-mountain environments such as the Andes (Fjeldså et al. 2012, 2023). In our case, we analyzed the collinearity among the 19 WorldClim variables to select a subset of predictors. This analysis was based on occurrence records, which reflect the environmental space occupied by the species, although they likely do not fully represent all available environmental conditions. To address this limitation, additional variables related to habitat structure were incorporated in subsequent modeling steps. The climatic predictors showing the lowest collinearity were maximum temperature of the warmest month (bio5), precipitation seasonality (bio15), and precipitation of the wettest quarter (bio16) (Tab. 1).

Tab: 1: Climatic variables from Wordclim data (resolution ~1 km²) and texture metrics derived from gray-level co-occurrence matrices (GLCMs) calculated from vegetation indices (NDVI, NDWI)

Climatic Variable		Description
BIO 5		Maximum Temperature. Warmest Month
BIO 15		Seasonality in precipitation. Coefficient of variation
BIO 16		Precipitation of the wettest quarter
Remote Sensing data (Landsat TM)	**Formula**	**Description**
Normalized difference vegetation index (NDVI)	$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$	Sensitive to chlorophyll pigments
Normalized difference water index (NDWI)	$NDWI = \frac{(NIR - SWIR)}{(NIR + SWIR)}$	Sensitive to leaf water content
Habitat structure. Variance, texture metric	$VA = \sum_{i,j=0}^{N-1} Pi, j(i - ME)^2$	This is a measure of the dispersion of the values around the mean
Habitat structure. Entropy, texture metric	$EN = \sum_{i,j=0}^{N-1} Pi, j(-lnPi, j)$	It measures the disorder of the values

Pi= probability of each pixel value, *ME* = mean, *j* and *n* are the number of rows or columns.

These data were resampled to match the spatial resolution of Landsat ETM+ (30 m pixel size) using the 'disaggregate' function with bilinear interpolation from the 'raster' package in R (R Development Core Team 2019). This procedure was applied to ensure spatial alignment among predictors within the modeling framework and does not imply any increase in the resolution or information content of the climatic variables. The bilinear interpolation method is based on a distance-weighted average of values in both the x and y directions (Fletcher & Fortin 2018).

Similarly, the selected climatic predictors were considered ecologically relevant, as they capture key patterns of temperature and precipitation that strongly influence the distribution of *M. baroni* (Tinoco et al. 2009, Carrasco et al. 2022), particularly in high-Andean environments where climatic conditions constrain species physiology and resource availability.

2.4 Remote sensing predictors

To construct the vegetation indices, satellite images from the Landsat program were used because of the large amount of information available across multiple years, in addition to its wide application in ecological studies (Pettorelli et al. 2011). Landsat 7 ETM+ multispectral images with a spatial resolution of 30 m were selected from the High-Resolution Global Maps Project (Hansen et al. 2013). To generate a cloud-free mosaic, Landsat images acquired between 1999 and 2012 from the Global Forest Change dataset (Global Forest Change 2021) were used and subsequently subjected to orthorectification and atmospheric corrections. Subsequently, a topographic C-correction (Riaño et al. 2003) was applied to compensate for the effects of slope and aspect on the land-surface reflectance, which is known to strongly affect classification accuracy in mountainous areas (Tassi et al. 2021).

Subsequently, spectral bands 4, 5, 6, and 7 available from the already-corrected Landsat images, corresponding respectively to Red (RED), Near Infrared (NIR), Shortwave Infrared 1 (SWIR1) and Shortwave Infrared 2 (SWIR2), were used to calculate two normalized indices: NDVI and NDWI (Tab. 1). These indices are calculated as the normalized difference between specific spectral bands. NDVI is the most recognized vegetation index and indicates vegetation productivity and health (Wallis et al. 2016, Huang et al. 2021). NDVI is derived from the normalized difference between the RED channel (0.66 μm) and NIR channel (0.86 μm), where the contrast is obtained between the high reflectance of the vegetation in the infrared and its low reflectance in the red region of the spectrum. NDWI is derived from the difference between the NIR (0.86 μm) and the SWIR (1.24 μm) bands and indicates canopy water content (water stress) and ground-cover moisture (Gao 1996). NDWI is known to be independent of and complementary to NDVI because it extends the spectral information from the 0.86 µm band used NDVI to 1.24 µm (Qiao et al. 2012, Ghalehteimouri et al. 2024), and the contrast is obtained between the moderate reflectance of water in the SWIR and its strong absorption in the infrared. When used together with NDVI, these indices improve the ability to discriminate among different vegetation characteristics (Zeng et al. 2020), which is suitable for the study area as it presents a variety of vegetation types.

From these two vegetation indices, texture statistics derived from the gray-level co-occurrence matrix were calculated using the glcm package in R (Zvoleff 2015, Liang et al. 2020). The matrix can be calculated for each pixel and recalculated by statistics applied to the surrounding neighborhood of pixels through a moving or fixed window algorithm (Farwell et al. 2020, Wallis et al. 2023). These textural metrics have been assessed, particularly in bird studies, in multiple habitat types (Farrell et al. 2013, Culbert et al. 2012)), especially in tropical mountain forests, where they offer the possibility of selecting the area or number of pixels according to the home range of the species studied (Culbert et al. 2012, Wallis et al. 2016, 2017).

To adjust texture statistics to the habitat scales of *M. baroni*, a 9 × 9 pixel moving window was used, corresponding to 270 × 270 m. Among the eight first- and second-order texture metrics, two statistics were selected: variance and entropy of NDVI and NDWI, as these were identified as the least correlated (Tab. 1). Furthermore, these metrics provide information on the structural contrast of the vegetation (variance) and its complexity or disorder (entropy) (Haralick et al. 1973, Culbert et al. 2012), which may reveal preferences in vegetation structure for animal species (Wood et al. 2012, 2013), and in this case, when compared with occurrence points, may reveal the habitat preferences of *M. baroni.*

2.5 Selection of climatic and habitat predictor

To select the most suitable variables for the SDMs, variables were first grouped according to their nature and characteristics (Vaughan & Ormerod 2005): (1) bioclimatic variables, (2) the NDVI index and their textures, and (3) the NDWI index and their textures, as these groups represent distinct environmental processes. This grouping helps preserve the diversity of ecological processes during variable selection.

An initial removal of redundant predictors was performed using the 'findCorrelation' function from the caret package in R (Kuhn 2008). This technique is widely used as a preliminary step to reduce redundancy among predictors before implementing predictive models. The algorithm examines the correlation matrix and, when two variables exhibit an absolute correlation greater than a predefined threshold (cutoff = 0.7), commonly used threshold in ecological modeling to reduce redundancy among predictors, it removes the variable with the highest average absolute correlation with the remaining variables (Kuhn 2008). The purpose of this procedure is to reduce predictors with redundant information within each group, preserving the most representative variables in each group while retaining all those variables that represent different environmental processes.

Subsequently, partial least squares (PLS) regression was applied to the remaining predictors as an exploratory step to reduce dimensionality and identify the predictors most strongly associated with the environmental conditions of the presence locations. PLS is a multivariate regression technique particularly effective for handling multicollinearity, as it generates orthogonal latent vectors that maximize the covariance between the predictors and the response variable (Wallis et al. 2017, Hair & Alamer 2022). It is especially useful when working with a small number of samples (63 presence points) and a large number of predictors, a common condition in ecological studies (Carrascal et al. 2009). This technique is widely applied in remote sensing, where correlated spectral bands can complicate analyses, as it reduces a large set of predictors to a smaller set of uncorrelated components (Carrascal et al. 2009, Wallis et al. 2016, Hair & Alamer 2022).

This final selection of predictors was carried out using the automated PLS regression approach implemented in the autopls package in R (Pereira et al. 2018). This package incorporates a selection procedure based on the VIP (Variable Importance in Projection) statistic and significance tests using jackknifing (Schmidtlein et al. 2012). The autopls algorithm was trained with 63 *M. baroni* presence points. During this process, a substantial number of predictors were removed, retaining only seven predictors: three bioclimatic variables and four textural statistics related to habitat structure (Tab. 2). This approach improves SDM performance, as

Tab. 2: Variable selection through 'find correlation' analysis and partial-least squares regressions (AutoPLS) with implemented backward selection and leave-one-out (LOO) cross validation, for the three sets of predictors variable combinations: climatic (Bioclimatic), NDVI and textural metrics (vegetation productivity and habitat structural indicator), and NDWI and textural metrics (vegetation water content and habitat structure indicator).

Variable Type	Find Correlation Selection	AutoPLS Selection	AutoPLS Coefficients	LOO R^2 AutoPLS	Significance level
19 Bioclimatics	BC 4	-----	-----	0.13	
	BC 5	Bio 5	-0.12		***
	BC 15	Bio 15	0.01		**
	BC 16	Bio 16	0.05		***
	BC 19	-----	-----		
Vegetation and	NDVI_contrast	-----	-----	-----	-----
habitat structure	NDVI_correlation	-----	-----	-----	-----
(NDVI and 8	NDVI_entropy	NDVI_entropy	-0.003	0.02	***
textural)	NDVI_variance	NDVI_variance	-0.046		
Vegetation and	NDWI_correlation	-----	-----	-----	-----
habitat structure	NDWI_ entropy	NDWI_ entropy	0.008	0.04	***
(NDWI and 8	NDWI_variance	NDWI_variance	-0.058		
textural)					

model accuracy can be compromised when using high-dimensional predictors with significant inter-correlation, such as remotely sensed imagery and bioclimatic variables (Beyer et al. 1999, Wallis et al. 2023).

2.6 Species distribution modeling

We used maximum entropy modeling (MaxEnt) (Phillips et al. 2006, Bertram et al. 2019), implemented using the dismo (Hijmans et al. 2024) and rJava (Urbanek 2026) packages in R, to develop species distribution models (SDMs). Maxent has demonstrated excellent performance in identifying species distributions and habitat patterns using presence-only occurrence data (Elith et al. 2006, SU et al. 2021). Furthermore, MaxEnt performs well with few occurrence records in environmentally heterogeneous areas by using background points, defined as randomly sampled locations representing the environmental space available to the species within its accessible range (Vanderwal et al. 2009, Velazco et al. 2022).

In this study, 10,000 background points were randomly generated within the study area, following the default configuration commonly used in Maxent modeling. Several studies have shown that including background points representing the available environmental space improves the model's ability to characterize species-environment relationships and enhances predictive performance (Phillips & Dubík 2008, Velazco et al. 2022). These points represent the available environmental conditions against which the environmental characteristics of the presence locations are contrasted (Velazco et al. 2022). Maxent models were run using the algorithm's default configuration, including the default regularization multiplier and automatically selected feature classes.

The excellent computational performance of Maxent enables the use of large numbers of input variables and allows the modeling of complex responses to environmental variables (Cord & Rödder 2011, Ahmed et al. 2020). Under the premise that species distribution models (SDMs) improve when climatic predictors are combined with habitat structure indicators (Elith & Leathwick 2009, Valerio et al. 2020), we constructed four model combinations to assess the relative influence of climatic variables and remotely sensed texture metrics on the distribution of *M. baroni*. Each model and their associated predictors are presented in Table 3.

The baseline model (BC) included only bioclimatic variables (bio5, bio15, bio16), whereas the remaining models incrementally incorporated NDVI textures metrics (BC + NV; vegetation structure), NDWI textures (BC + NW; water-related features), and both (BC + NV + NW). This design tested whether texture metrics enhanced predictive accuracy beyond climate data alone while minimizing multicollinearity and capturing distinct habitat attributes. All Maxent runs used consistent settings: convergence threshold (10^{-5}), maximum iterations (1000), regularization ($\beta = 10^{-4}$), and linear, quadratic, product, and hinge features (Phillips et al. 2006, 2017).

Tab. 3: Combination of selected predictors to model distribution area of *Metallura baroni* using MaxEnt (four models)

Model	Predictor	Data set	Abbreviation
1	Bioclimatic (Wordlclim)	Max Temperature of Warmest Month (bio5), Precipitation Seasonality (Coefficient of Variation) (bio15), Precipitation of Wettest Quarter (bio16)	BC
2	BC + NDVI texture	Max Temperature of Warmest Month (bio5), Precipitation Seasonality (Coefficient of Variation) (bio15), Precipitation of Wettest Quarter (bio16), NDVI-entropy, NDVI-variance	BC + NV
3	BC + NDWI textures	Max Temperature of Warmest Month (bio5), Precipitation Seasonality (Coefficient of Variation) (bio15), Precipitation of Wettest Quarter (bio16), NDWI-entropy, NDWI-variance	BC + NW
4	BC + (NDVI + NDWI) textures	Max Temperature of Warmest Month (bio5), Precipitation Seasonality (Coefficient of Variation) (bio15), Precipitation of Wettest Quarter (bio16), NDVI-entropy, NDVI-variance, NDWI-entropy, NDWI-variance	BC + NV + NW

2.7 Model evaluation

We tested the SDMs performance (discrimination capacity and reliability), using the area under curve (AUC) of receiver operating characteristic (ROC) plots, obtained by plotting the sensitivity as a function of the commission error (1-specificity), for all possible thresholds of a probabilistic occurrence prediction (Mcpherson et al. 2004, Phillips et al. 2017). The ROC curve provides a single measure of overall model accuracy that is independent of a particular threshold (Mcpherson et al. 2004, Phillips et al. 2017). For all four models, we estimate the variance of the AUC by a bootstrapping approach (999 iterations), based on the fact that the AUC is equal to the probability that a true positive is ranked higher than a true negative (Cord & Rödder 2011, Muschelli 2020). The AUC value is a threshold-independent measure of accuracy commonly used to evaluate SDMs performance (ranging 0 to 1, where 0.5 indicates model accuracy not better than random and 1.0 indicates perfect model fit), (Phillips et al. 2017).

To evaluate the significance of the *M. baroni* SDMs, we compared their AUC values with those of null models generated from random background points (Raes & ter Steege et al. 2007, Bohl et al. 2019). For each of the four SDMs, null distributions were generated by randomly sampling 999 sets of background points (without replacement) and calculating their corresponding AUC values. The upper limit of the 95% confidence interval (C.I.) was determined by ranking these null AUC values. Observed AUC values exceeding this threshold indicated that the models performed significantly better than random, suggesting that *M. baroni* occurrence records reflect specific niche requirements rather than chance (Raes & Steege et al. 2007, Osborne et al. 2022).

Traditionally, model the accuracy has been interpreted on arbitrary AUC categories defined by researchers where values of 0.90-1.00 are considered excellent, 0.80–0.90 good, 0.70–0.80 fair, 0.60–0.70 poor and 0.50–0.60 = fail (Su et al. 2021). However, these classifications may lack ecological relevance (Liu et al. 2015, Osborne et al. 2022). To address this limitation, we select the maxSSS method, which determines model thresholds by maximizing the sum of sensitivity and specificity (Liu et al. 2015). The maxSSS method is considered a threshold-selection approach for presence-only data because it accounts for the relative importance of omission and commission errors (Liu et al. 2015, Hintze et al. 2021). To compare the differences between thresholds derived from AUC and the maxSSS statistic across the four models with the different combination of predictors, we performed a permutation test. We evaluated how the species presence points were associated with raster values exceeding the select threshold that would represent areas of high suitability for the species, according to MaxEnt and the maxSSS statistic. For this analysis we used the 'raster' (Hijmans 2024) and 'sp' packages in R (Pebesma & Bivand 2005). The test generated 999 random points for comparison with the real presence points of the species (Meirmans 2021, Astudillo et al. 2024)

Finally, we evaluated the ecosystem areas falling within the maxSSS threshold by cross-referencing this information with the official ecosystems map of Ecuador (MAE 2013). We then performed a coefficient of variation analysis to determine whether differences in total area and in each ecosystem type among the different models were significant (Fig. 2).

3 Results

3.1 Predictor selection

To reduce the number of predictors and ensure the optimal performance of the Maxent models, two variable-selection analyses combining 'find-correlation' and 'PLS' regression were performed. These analyses aimed to identify predictors with low intercorrelation and strong associations with the environmental conditions of *M. baroni* presence locations. These methods were applied to three different sets of predictors. In the first set, we used 19 climatic variables from WorldClim. In the second, we generated the NDVI index together with its eighth textural structure variables. In the third, we generated nine variables from NDWI, including its eight texture variables. Different suitability patterns were identified among three types of predictors in relation to the focal species occurrence points, ranging from the highest LOO R^2 values for climatic predictors to the lowest values for vegetation predictors (Tab. 2). The climatic variables maximum temperature of the warmest month (bio5), precipitation seasonality (bio15), and precipitation of the wettest quarter (bio16) showed the lowest correlation, variables known to be important determinants of bird species distribution (Fjeldså et al. 2012) (Tab. 2).

In the second and third analyses, which included the vegetation indices NDVI, NDWI, (indicators of vegetation productivity) together with their textural metrics (indicator habitat structure), the entropy and variance metrics of both vegetation indices exhibited the lowest correlation. Thus, habitat structure indicators were selected as the most relevant predictors for modeling the distribution of *M. baroni* over vegetation productivity indicators.

The weighted regression coefficients of the PLS analysis identified the magnitude and direction of the relationships between the predictors and the occurrence of *M. baroni* (Fig. 3). The temperature vari-

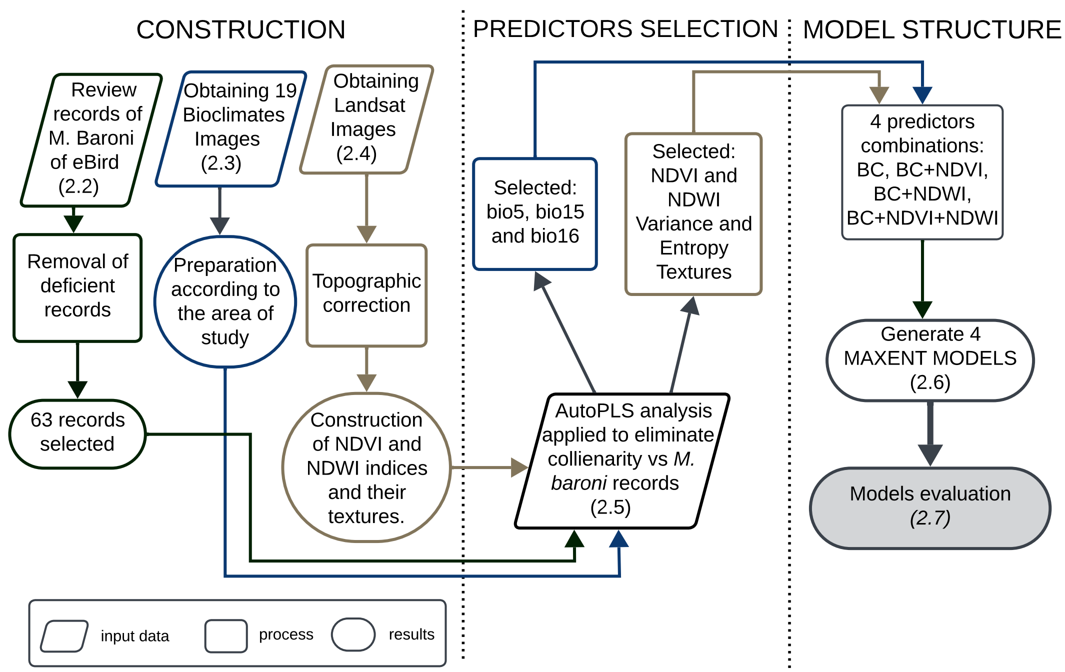

Fig. 2: Flowchart of the modeling process. The methodology is divided into model construction and variable selection stages, including the acquisition and filtering of species occurrence records, the acquisition and selection of bioclimatic variables and vegetation indices using PLS analysis, the development of four models, the statistical evaluation of the models, and the analysis of the results. Numbers in parentheses refer to the corresponding sections in the text.

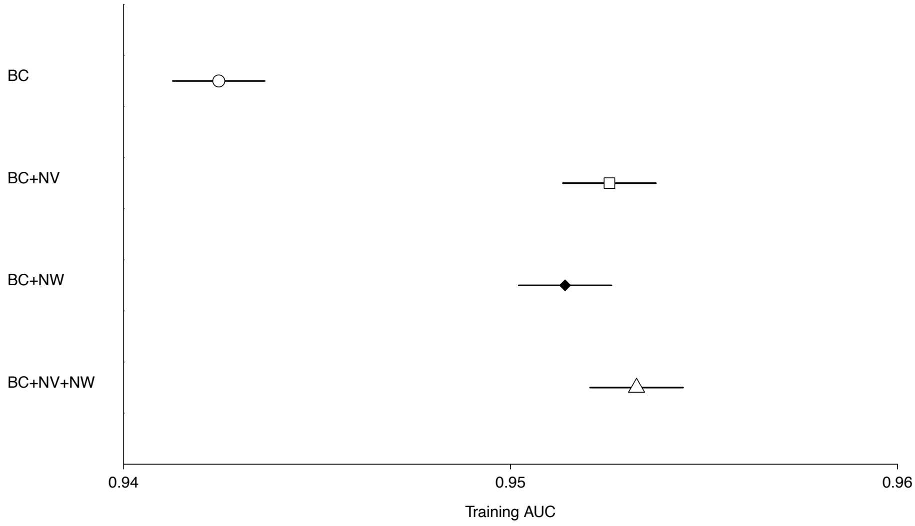

Fig. 3: Magnitude and direction of the weighted regression coefficients from the PLS analysis on the relationship between the predictors (temperature, precipitation, entropy, and variance texture metrics of NDVI and NDWI) and the *M. baroni* occurrence points. Significance corresponds to: ** p < 0.05, *** p < 0.001.

able maximum temperature of the warmest month (bio5) had a significant negative effect, while the precipitation variables precipitation seasonality (bio15) and precipitation of the wettest quarter (bio16) had a significant positive effect on the occurrence of *M. baroni*. The entropy metrics of NDVI and NDWI had no significant effect, whereas the variance metrics of NDVI and NDWI showed significant negative effect on the occurrence points of *M. baroni* (Fig. 3, Tab. 2). Regarding statistical performance, the percentage of contribution and the permutation importance of variables in species distribution models (SDMs), indicated that climatic variables showed a slightly greater contribution, while habitat structure variables showed a lower percentage of contribution to the occurrence of *M. baroni* (Tab. A1).

3.2 SDM model performance

The distribution of *M. baroni* was modeled using four combinations of climatic and habitat structure predictors (Tab. 3). All models showed high predictive performance (AUC > 0.94), the models including habitat structure predictors achieving slightly higher AUC scores (AUC 0.98) than the climatic model (AUC = 0.94) (Fig. 4, Tab. 4).

The AUC values significantly differed from null models, confirming the niche specificity of *M. baroni* (Fig. 5).

Permutation analysis applied to Maxent, using predictors related to habitat structure and considering the thresholds defined by Maxent and the maxSSS statistic, revealed that the actual presence points of the species were significantly associated ($p < 0.05$) with pixels exceeding the threshold considered suitable for the species according to Maxent and the maxSSS statistic. This association was significantly stronger than that observed for the random points generated during the permutation test. In contrast, the model based solely on climatic predictors, using the threshold determined by Maxent (0.94), did not show a significant association ($p = 1$). This suggests that bioclimatic variables alone may be insufficient to accurately predict the species‘ distribution (Tab. A2).

Thresholds based on the maxSSS method showed minor variations across models (Tab. 4, Fig. A1), but notable differences were observed in the area and vegetation types delineated from the official vegetation cover map of Ecuador. For example, the model using only climatic predictors (BC) covered a larger area (354.7 Km^2) compared to models incorporating habitat structure variables (238.4-148.5 km^2) (Tab. 4). In particular, the model developed using the BC+NV+NW combination delineated the smaller areas. The most notable change was the reduction in páramo areas, allowing a more precise delineation of the forest and scrub habitats where *M. baroni* is found (Tinoco et al. 2009, Astudillo et al. 2024) (Tab. 4).

On the other hand, coefficient of variation analysis indicated that the total areas captured within the thresholds defined by the maxSSS method dif-

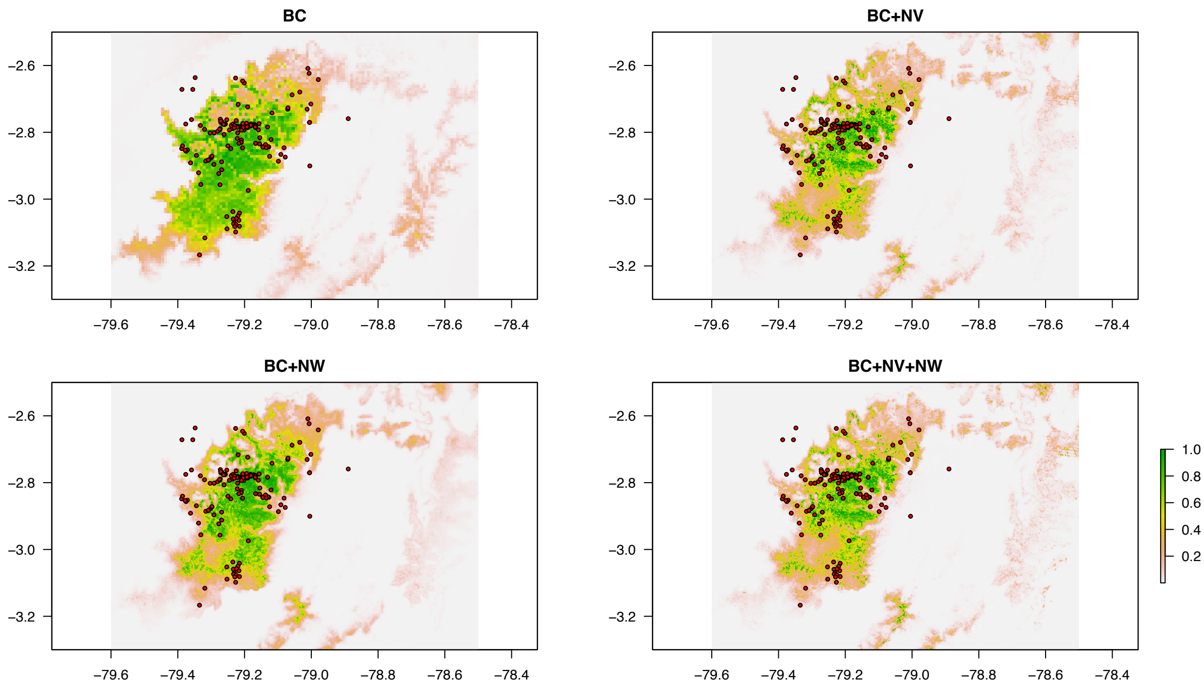

Fig. 4: Prediction maps of *M. baroni* from Maxent models evaluating (a) BC variables, (b) BC+NV variables, (c) BC+NW variables, and (d) BC+NV+NW variables. Points indicate occurrence records of *M. baroni.* Color legends indicate AUC values, values of 0.5 indicate that the model does not perform better than expected by chance and 1.0 indicate near-perfect predictive performance.

Tab. 4: Threshold values determined by the maxSSS index of the four SDMs elaborated for *M. baroni* and their corresponding areas extracted from the official ecosystem map of Ecuador (MAE 2013) through cartographic overlay analysis

Variables	Maxent AUC	Threshold	Total area	Shrubland		Forest		Paramo		Other	
		maxSSS	km²	km²	%	km²	%	km²	%	km²	%
BC	94.9	0.78	354.79	26.59	7.49	0.78	0.20	311.36	87.76	16.05	4.52
BC+NV	96.1	0.79	179.34	22.37	12.47	4.32	2.41	138.53	77.24	14.10	7.86
BC+NW	95.3	0.79	238.4	18.72	7.85	5.07	2.12	197.63	82.90	16.96	7.11
BC+NV+NW	95.9	0.81	148.53	18.98	12.78	6.51	4.38	109.22	73.53	13.80	9.29

fered significantly among the models developed with different combinations of predictors. The ecosystem types showing the greatest differences were the páramo and forests (CV > 30%). In contrast, shrub areas did not show significant differences (CV < 30%) (Tab. 5).

4 Discussion

4.1 Predictor selection

Our results revealed that climatic variables and habitat structure predictors best explained the current distribution of *M. baroni*. Model performance improved following a structured predictor-selection process which reduced redundancy and identified the variables most closely associated with the species' occurrence. This process involved the separate evaluation of climatic variables, vegetation indices, and habitat structure metrics prior to integrating the final set of predictors. Unlike previous studies that relied primarily on NDVI as an indicator of vegetation productivity (Tinoco et al. 2023), our approach incorporated texture metrics that provided additional information on habitat structure and environmental heterogeneity. This methodological advancement allowed us to better capture the ecological requirements of *M. baroni*, particularly in complex ecosystems such as the tropical Andes.

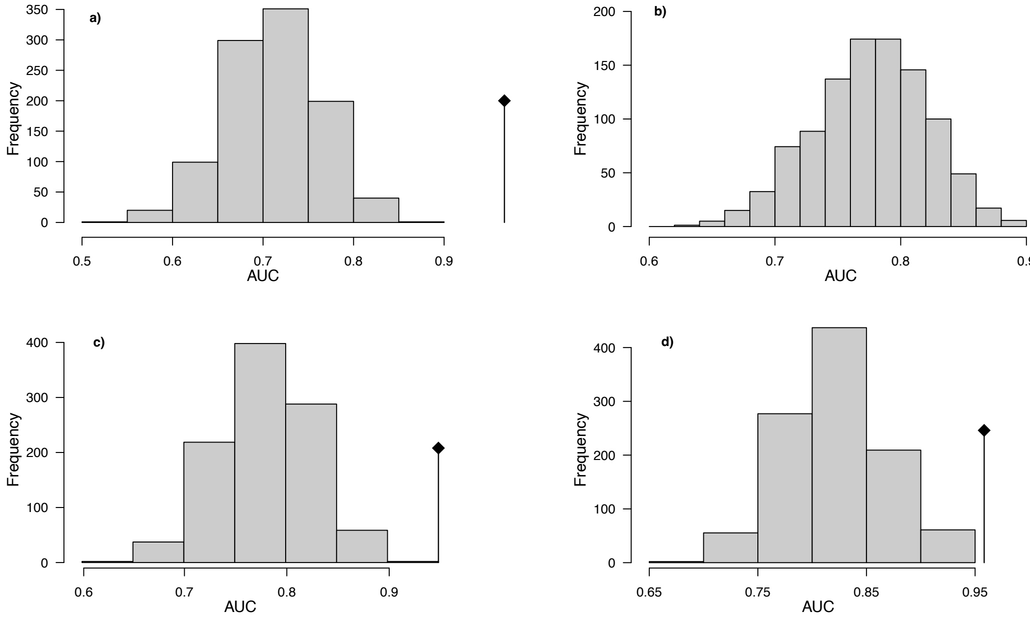

Fig. 5: Frequency histograms of expected 95% confidence interval (C.I.) of upper limit AUC values under a null model (Meirmans 2021), compared with the observed AUC values of Maxent models. The null AUC was computed 999 times with an equal number of points (63) randomly drawing without replacement from all available cells (histograms). The black pin indicates the mean observed AUC scores. The observed AUC values differed significantly from the distribution of AUC values under the null model (Monte Carlo test, simulated p-value = 0.001), for the following models: a) BC model, b) BC+NV model, c) BC+NW model, and d) BC+NV+NW model. The p-value was estimated as: (number of random values equal to or greater than the observed one + 1)/(number of permutations + 1).

In this study, by combining climatic variables and vegetation indices and their associated textures metrics derived from remote sensing data, we obtained a more comprehensive representation of the environmental conditions influencing the distribution of *M. baroni*. This approach enabled a better characterization of the environmental context associated with the species' occurrences and habitat requirements. This type of methodology, which integrates multiple sources of environmental data, provides a more comprehensive understanding of species' ecological niches and should therefore be considered in future species distribution modeling studies (Pettorelli et al. 2014, Franklin 2023).

Tab. 5: Coefficient of variation analysis comparing the four types among models, to determine whether significant differences existed between them

Comparative values of the 4 models	CV value [%]
Total Area	39.52
Shrubland	19.98
Forest	58.42
Paramo	47.24
Other	10.00

CV > 30% indicates high variability among values and CV < 30% indicates low variability among values.

The predictor-selection process highlighted temperature, precipitation, and texture metrics derived from NDVI and NDWI emerged as the most relevant variables for explaining the distribution of *M. baroni*. These findings emphasize the combined role of climatic conditions and habitat structure in shaping species distribution in tropical mountain ecosystems. Similar patterns have been reported in previous studies, which have highlighted the importance of vegetation-related variables in determining species distribution in mountainous regions (Jetz et al. 2007, Valerio et al. 2020).

Although the R^2 values obtained from the PLS analysis were relatively low, this was expected given the primary objective of identifying the least correlated predictors for use in MaxEnt models (Shmueli et al. 2019). Low R^2 values are not uncommon in PLS, as the method prioritizes reducing multicollinearity and identifying relevant predictors rather than maximizing explained variance (Abdi 2010, Goktas & Dirsehan 2024). This approach helps ensure that the predictors included in the final SDM are both ecologically meaningful and statistically robust, reducing the risk of overfitting and improving model generalizability (Merow et al. 2013, Omweno et al. 2021).

Our predictor selection process, guided by PLS analysis, highlights the value of integrating multiple data sources and applying dimensionality-reduction techniques to improve the ecological relevance and predictive performance of SDMs. Both climatic variables and texture metrics revealed important environmental patterns associated with the distribution and habitat requirements of *M. baroni*. Among the climatic predictors, whereas bio15 (precipitation seasonality) and bio16 (precipitation of the wettest quarter) showed a positive association. These relationships suggest that the distribution of *M. baroni* is closely linked to specific temperature and precipitation regimes characteristic of High-Andean ecosystems. Given that this hummingbird is restricted to elevation between 3000 and 4100 m a.s.l., these results are consistent with its occurrence in cool, humid mountain environments such as Andean páramo (Tinoco et al. 2009, Carrasco et al. 2022). Similar patterns have been reported for the other mountain hummingbirds, whose distributions are strongly influenced by temperatures and precipitation gradients (Santillán et al. 2018). Furthermore, structurally complex habitats may contribute to maintenance of favorable microclimatic conditions, which are important for thermoregulation and the health of these species (Cabello et al. 2012, Jirinec et al. 2022).

For its part, the inclusion of NDVI and NDWI derived texture metrics highlights the importance of habitat structure and vegetation dynamics in shaping species distributions in complex landscapes such as the tropical Andes, particularly for *M. baroni*. This is especially relevant because habitat structure influences the availability of suitable space, shelter, and resources within the species' distribution range, particularly in areas dominated by shrubby vegetation (Tinoco et al. 2009, Carrasco et al. 2022). Previous studies have demonstrated the utility of remote sensing-derived metrics in capturing fine-scale environmental variability (Pettorelli et al. 2011, Wood et al. 2012, Wood et al. 2013, Franklin 2023, Buthelezi et al. 2024). By incorporating these metrics, our models provide a more nuanced understanding of the ecological drivers of *M. baroni* distribution.

The inclusion of texture metrics, particularly variance and entropy, proved highly informative. These metrics capture vegetation heterogeneity, a critical factor for species that depend on specific habitat structures (MacArthur & MacArthur 1961, Kissling et al. 2008, Ciarle & Burns 2024). The variance metrics derived from NDVI and NDWI showed significant negative relationships in the PLS analysis, suggesting that *M. baroni* is associated with relatively homogeneous vegetation structures rather than highly heterogeneous landscapes. The importance of habitat structure for montane bird species is supported by numerous studies highlighting its role in key ecological functions such as nesting, foraging, and the maintenance of suitable microclimatic conditions (Daily et al. 2009, Leurs et al. 2023). These findings suggest that the occurrence of *M. baroni* is influenced not only by climatic conditions but also by the structural characteristics of the vegetation, which may provide suitable shelter, foraging opportunities, and favorable microhabitats within high-Andean ecosystems (Braunisch & Suchant 2010, Moudrý et al. 2021).

4.2 SDM performance

One key reason for the improved performance of our models was the inclusion of texture metrics derived from NDVI and NDWI, which provided a more detailed representation of habitat structure. Although climatic variables accounted for the largest relative contribution to model predictions (94.9%), the addition of texture metrics - despite their modest contribution (0.4–1.2%) resulted in ecologically meaningful refinements of predicted habitat suitability. These improvements enhanced the delineation of fine-scale habitat conditions associated with *M. baroni*, particularly in shrub and forest habitats where vegetation structure may influence resource availability, shelter, and microclimatic conditions. This finding is consistent with Wallis et al. (2017), who demonstrated that texture metrics are strongly associated with habitat complexity and species diversity in montane ecosystems. For microendemic species such as *M. baroni*, even modest improvements in model precision can be important for conservation planning, as they may help identify suitable habitats that would otherwise remain undetected.

In comparison with studies conducted in less complex ecosystems, such as lowland forests or temperate regions, our models faced additional challenges due to the high environmental heterogeneity of the tropical Andes. Models based solely on climatic predictors predicted more extensive areas of suitable habitat, whereas the incorporation of habitat structure variables resulted in a more restricted and ecologically realistic representation of the species' potential distribution. This refinement increased model specificity and improved the discrimination of habitats associated with *M. baroni*. These findings underscore the importance of incorporating habitat structure indicators into species distribution models (SDMs), particularly in topographically and ecologically complex regions (Rödder & Lötters 2009, Valerio et al. 2020, Su et al. 2021).

Our results also align with broader ecological studies highlighting the role of vegetation structure in shaping species distributions. For instance, Basile et al. (2021), MacArthur & MacArthur (1961), and Rotenberry (1985) demonstrated that bird species richness and abundance are closely associated with vegetation structure and complexity. Similarly, Buermann et al. (2008) found that remote sensing data, including texture metrics, significantly improved the performance of SDMs for Andean bird species. By incorporating texture metrics into our modeling framework, we achieved a more detailed representation of habitat structure and environmental heterogeneity, improving the ecological realism of model predictions.

In contrast, the lack of a significant association for the model based solely on climatic predictors when evaluated using the MaxEnt threshold ($p = 1$, Tab. A1) suggests that bioclimatic variables alone may be insufficient to accurately predict the distribution of *M. baroni*. This result is consistent with previous studies showing that climatic variables, although highly informative, do not always capture the full complexity of species' ecological requirements, particularly when distributions are strongly influenced by habitat structure and other fine-scale environmental characteristics (Elith & Leathwick 2009, Valerio et al. 2020).

While our model performed well, there are opportunities for further refinement. For example, incorporating additional remote sensing data, such as LiDAR-derived canopy height or hyperspectral imagery, could provide even more detailed information on habitat structure (Farell et al. 2013, Wallis et al. 2017). Future studies could also explore the temporal dynamics of vegetation and climate, as seasonal changes may influence the distribution of *M. baroni* and other montane species (Meirmans 2021, Buthelezi et al. 2024). Furthermore, the development of high-resolution climatic datasets for the Andes may help capture microclimatic variation that is not represented in currently available global products, thereby improving predictions for microendemic species.

5 Conclusions

In conclusion, our study highlights the importance of integrating complementary environmental predictors to model species distributions in complex ecosystems. Given the limited availability of high-resolution climatic data, we combined bioclimatic variables with remote sensing-derived texture metrics from NDVI and NDWI to better represent both climatic conditions and habitat structure. This approach improved the ecological realism and predictive performance of species distribution models (SDMs) for *M. baroni*, a microendemic hummingbird of the tropical Andes, providing a more refined representation of its ecological niche and valuable information to support conservation planning in a region characterized by high levels of endemism.

In ecosystems such as the Andean páramo, altitudinal gradients often produce abrupt changes not only in climatic conditions but also in habitat structure, including the availability of shelter and food resources. Consequently, incorporating habitat structure variables can improve the representation of fine-scale environmental conditions influencing species distributions. While many SDM studies have focused on species with broad geographic ranges, microendemic species with highly restricted distributions present additional challenges. In these cases, the spatial resolution of currently available climatic datasets may be insufficient to capture fine-scale environmental variability (Hemp & Hemp 2024), potentially leading to an overestimation of suitable habitat and a less accurate representation of species-environment relationships.

Including habitat structure variables in SDMs reduced the extent of predicted suitable areas compared with models based solely on climatic predictors, suggesting that the latter may overestimate species distributions, partly due to the limited spatial resolution of currently available climatic datasets. The incorporation of habitat structure information produced more spatially refined predictions, which could help optimize field surveys by reducing the area requiring in situ verification and enabling a more precise delineation of priority conservation areas. Overall, our results highlight the value of integrating multiple sources of environmental information to improve species distribution modeling and support conservation planning in regions characterized by high biodiversity and environmental complexity, such as the tropical Andes. Species distributions are influenced by climatic conditions, habitat characteristics, and land-cover dynamics. The interaction among these factors can strongly affect biodiversity patterns (He et al. 2019), highlighting the importance of integrating complementary environmental predictors in species distribution models.

Acknowledgments

We gratefully acknowledge the support of the University of Azuay, particularly Rector Francisco Salgado and Vice-Rector for Research Rafaella Ansaloni, for providing the resources that made this study possible. We also extend our sincere thanks to Monserrath Santillán for her valuable assistance with manuscript editing.

References

Abdi H (2010) Partial least squares regression and projection on latent structure regression (PLS Regression) *Wiley Interdisciplinary Reviews: Computational Statistics* 2: 97-106. https://doi.org/10.1002/wics.51

Acebes P, Lillo P, Jaime-González C (2021) Disentangling lidar contribution in modeling species-habitat structure relationships in terrestrial ecosystems worldwide. A systematic review and future directions. *Remote Sensing* 13: 3447. https://doi.org/10.3390/rs13173447

Ahmed N, Atzberger C, Zewdie W (2020) Integration of remote sensing and bioclimatic data for prediction of invasive species distribution in data-poor regions: A review on challenges and opportunities. *Environmental Systems Research* 9: 32. https://doi.org/10.1186/s40068-020-00195-0

Aiello-Lammens ME, Borja RA, Radosavljevic A, Vilela B, Anderson RP (2015) spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 38: 541-545. https://doi.org/10.1111/ecog.01132

Astudillo PX, Barros S, Mejía D, Villegas FR, Siddons DC, Latta SC (2024) Using surrogate species and MaxEnt modeling to prioritize areas for conservation of a páramo bird community in a tropical high Andean biosphere reserve. *Arctic, Antarctic, and Alpine Research* 56: 2299362. https://doi.org/10.1080/15230430.2023.2299362

Barbet-Massin M, Jetz W (2014) A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modeling. *Diversity and Distributions* 20: 1285-1295. https://doi.org/10.1111/ddi.12229

Barry RG (2008) Mountain weather and climate. Cambridge. https://doi.org/10.1017/CBO9780511754753

Basile M, Storch I, Mikusiński G (2021) Abundance, species richness and diversity of forest bird assemblages - the relative importance of habitat structures and landscape context. *Ecological Indicators* 133: 108402. https://doi.org/10.1016/j.ecolind.2021.108402

Bertram J, Newman EA, Dewar RC (2019) Comparison of two maximum entropy models highlights the metabolic structure of metacommunities as a key determinant of local community assembly. *Ecological Modeling* 407: 108720. https://doi.org/10.1016/j.ecolmodel.2019.108720

Beyer K, Goldstein J, Ramamakrishnan R, Shaft U (1999) When is "Nearest Neighbor" meaningful? Beeri C, Buneman P (eds) Database theory — ICDT'99. ICDT 1999. Lecture notes in computer science 1540: 217-235. Berlin, Heidelberg. https://doi.org/10.1007/3-540-49257-7_15

Bohl CL, Kass JM, Anderson RP (2019) A new null model approach to quantify performance and significance for ecological niche models of species distributions. *Journal of Biogeography* 46: 1101-1111. https://doi.org/10.1111/jbi.13573

Braunisch V, Suchant R (2010) Predicting species distributions based on incomplete survey data: The trade-off between precision and scale. *Ecography* 33: 826-840. https://doi.org/10.1111/j.1600-0587.2009.05891.x

Buermann W, Saatchi S, Smith TB, Brian R, Graham CH, Chaves JA, Mila B (2008) Predicting species distributions across the Amazonian and Andean regions using remote sensing data. *Journal of Biogeography* 35: 1160-1176. https://doi.org/10.1111/j.1365-2699.2007.01858.x

Buthelezi MNM, Lottering RT, Peerbhay KY, Mutanga O (2024) A machine learning approach to mapping suitable areas for forest vegetation in the eThekwini municipality. *Remote Sensing Applications: Society and Environment* 35: 100898. https://doi.org/10.1016/j.rsase.2024.101208

Cabello J, Fernández N, Alcaraz-Segura D, Oyonarte C, Piñero G, Altesor A, Delibes M, Paruelo JM (2012) The ecosystem functioning dimension in conservation: Insights from remote sensing. *Biodiversity and Conservation* 21: 3287–3305 https://doi.org/10.1007/s10531-012-0370-7

Carrascal LM, Galván I, Gordo O (2009) Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos* 118: 681–690. https://doi.org/10.1111/j.1600-0706.2008.16881.x

Carrasco-Ugalde A, Molina-Abril P, Pacheco D, Tinoco BA (2022) Nesting biology of an Ecuadorian endemic hummingbird, the endangered Violet-throated Metaltail *Metallura baroni*. *Revista Ecuatoriana de Ornitología* 1: 31–40. https://doi.org/10.18272/reo.v8i1.2320

Carscadden KA, Emery NC, Arnillas CA, Cadotte MW, Afkhami ME, Gravel D, Livingstone SW, Wiens JJ (2020) Niche breadth: Causes and consequences for ecology, evolution, and conservation. *The Quarterly Review of Biology* 95: 179–214. https://doi.org/10.1086/710388

Cavalheri DG, Ceron K, Carrillo JFC, Neves MO, Santana DJ (2024) Ecological niche modeling of *Pseudopaludicola Motorzinho* (Anura, Leptodactylidae), with two new distribution records and comments on its advertisement call variation. *Southwestern Naturalist* 67: 283–290. https://doi.org/10.1894/0038-4909-67.4.283

Ciarle R, Burns KC (2024) Island biogeography of birds in the South West Pacific: Direct and indirect effects of physical geography and co-occurring vegetation. *Journal of Biogeography* 51: 1623–1631. https://doi.org/10.1111/jbi.14731

CIIFEN (Centro Internacional para la Investigación del Fenómeno de El Niño) (s.f.) Generador de mapas. https://ciifen.org/generador-de-mapas/

Colwell RK, Rangel TF (2009) Hutchinson's duality: The once and future niche. *PNAS* 106: 19651–19658. https://doi.org/10.1073/pnas.0901650106

Cord AF, Rödder D (2011) Inclusion of habitat availability in species distribution models through multi-temporal remote-sensing data? *Ecological Applications* 21: 3285–3298. https://doi.org/10.1890/11-0114.1

Córdova M, Célleri R, Shellito CJ, Orellana-Alvear J, Abril A, Carrillo-Rojas G (2018) Near-surface air temperature lapse rate over complex terrain in the southern Ecuadorian Andes: Implications for temperature mapping. *Arctic, Antarctic, and Alpine Research* 48: 673–684. https://doi.org/10.1657/AAAR0015-077

CRC-OSA/CIIFEN (Centro Regional del Clima para el Oeste de Sudamérica/ Centro Internacional para la Investigación del Fenómeno de El Niño) (2024) Seven gridded databases for western South America are delivered to meteorological services, key information for the generation of climate services! https://crc-osa.ciifen.org/en/2024/01/26/se-entregan-a-los-servicios-meteorologicos-siete-bases-de-datos-grillados-para-el-oeste-de-sudamerica-informacion-clave-para-la-generacion-de-servicios-climaticos/

Culbert PD, Radeloff VC, St-Louis V, Flather CH, Rittenhouse CD, Albright TP, Pidgeon AM (2012) Modeling broad-scale patterns of avian species richness across the Midwestern United States with measures of satellite image texture. *Remote Sensing of Environment* 118: 140–150. https://doi.org/10.1016/j.rse.2011.11.004

Daily GC, Polasky s, Goldstein j, Kareiva PM, Mooney ha, Pejchar L, Ricketts TH, Salzman J, Shallenberger R (2009) Ecosystem services in decision making: Time to deliver. *Frontiers in Ecology and the Environment* 7: 21–28 https://doi.org/10.1890/080025

Elith J, Leathwick JR (2009) Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129–151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

Farell SL, Collier BA, Skow K.L, Long AM, Campomizzi AJ, Morrison ML, Hays KB, Wilkins RN (2013) Using LiDAR-derived vegetation metrics for high-resolution, species distribution models for conservation planning. *Ecosphere* 4: 1–18 https://doi.org/10.1890/ES12-000352.1

Farwell LS, Elsen PR, Razenkova E, Pidgeon AM, Radeloff VC (2020) Habitat heterogeneity captured by 30-m resolution satellite image texture predicts bird richness across the United States. *Ecological Applications* 30: e02157. https://doi.org/10.1002/eap.2157

Feng X, Park DS, Walker C, Peterson AT, Merow C, Papes M (2019) A checklist for maximizing reproducibility of ecological niche models. *Nature Ecology and Evolution* 3: 1382–1395. https://doi.org/10.1038/s41559-019-0972-5

Fick SE, Hijmans RJ (2017) WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302–4315. https://doi.org/10.1002/joc.5086

Fisher A, Rudin C, Dominici F (2019) All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research* 20: 1–81.

Fjeldså J, Bowie RC, Rahbek C (2012) The role of mountain ranges in the diversification of birds. *Annual Review of Ecology, Evolution, and Systematics* 43: 249–265. https://doi.org/10.1146/annurev-ecolsys-102710-145113

Fjeldså J, Sonne J, Rahbek C (2023) The alpine avifauna of tropical mountains. Chamberlain D, Lehikoinen A, Martin K (eds) *Ecology and conservation of mountain birds:* 336–371. Cambridge. https://doi.org/10.1017/9781108938570.010

Fletcher R, Fortin MJ (2018) Spatial ecology and conservation modeling. Cham. https://doi.org/10.1007/978-3-030-01989-1

Franklin J (2023) Mapping species distributions: Spatial inference and prediction. Cambridge. https://doi.org/10.1017/CBO9780511810602

Gao BC (1996) NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* 58: 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Ghalehteimouri KJ, Ros FC, Rambat S, Nasr T (2024) Spatial and temporal water pattern change detection through the normalized difference water index (NDWI) for initial flood assessment: A case study of Kuala Lumpur 1990 and 2021. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences* 114: 178–187. https://doi.org/10.37934/arfmts.114.1.178187

GBIF (Global Biodiversity Information Facility) (2021) What is GBIF? Copenhagen. https://www.gbif.org/what-is-gbif

GLOBAL FOREST CHANGE (2013) Global Forest Change. University of Maryland. http://earthenginepartners.appspot.com/science-2013-global-forest

Goktas P, Dirsehan T (2024) Using PLS-SEM and XAI for causal-predictive services marketing research. *Journal of Services Marketing* 39*:* 53–60. https://doi.org/10.1108/JSM-10-2023-0377

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modeling* 135: 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hair J, Alamer A (2022) Partial least squares structural equation modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics* 1. https://doi.org/10.1016/j.rmal.2022.100027

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A (2013) High-resolution global maps of 21st-century forest cover change. *Science* 342: 850–853. https://doi.org/10.1126/science.1244693

Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics 3(6): 610–621. https://doi.org/10.1109/TSMC.1973.4309314

He X, Liang J, Zeng G, Yuan Y, Li X (2019) The effects of interaction between climate change and land-use/cover change on biodiversity-related ecosystem services. *Global Challenges* 3: 1800095. https://doi.org/10.1002/gch2.201800095

Hemp A, Hemp J (2024) Weather or not - Global climate databases: Reliable on tropical mountains? *PLoS ONE* 19*:* e0299363. https://doi.org/10.1371/journal.pone.0299363

Hijmans R (2024) _raster: Geographic data analysis and modeling. R package version 3.6-30. https://CRAN.R-project.org/package=raster

Hijmans R.J, Phillips S, Leathwick J, Elith J (2024) dismo: Species Distribution Modeling. R package version 1.3-16. https://doi.org/10.32614/CRAN.package.dismo

Hintze F, Machado RB, BERNARD E (2021) Bioacoustics for in situ validation of species distribution modeling: An example with bats in Brazil. *PLoS ONE* 16. https://doi.org/10.1371/journal.pone.0248797

Huang S, Tang L, Hupy JP, Wang Y, Shao G (2021) A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research* 32. https://doi.org/10.1007/s11676-020-01155-1

Hutchinson GE (1957) Concluding remarks. Population studies: animal ecology and demography. *Cold Spring Harbor Symposium on Quantitative Biology* 22: 415. https://doi.org/10.1101/SQB.1957.022.01.039

Hutchinson GE. (1959) Homage to Santa Rosalia or why are there so many kinds of animals? *The American Naturalist* 93: 145–149. https://doi.org/10.1086/282070

Jetz W, Wilcove DS, Dobson AP (2007) Projected impacts of climate and land-use change on the global diversity of birds. *PLoS Biology* 5: e157. https://doi.org/10.1371/journal.pbio.0050157

Jirinec V, Elizondo EC, Rodrigues PF, Stouffer PC (2022) Climate trends and behavior of a model Amazonian terrestrial insectivore, Black-faced Antthrush, indicate adjustment to hot and dry conditions. *Journal of Avian Biology* 2022: e02946. https://doi.org/10.1111/jav.02946

Kissling WD, Field R, Böhning-Gaese K (2008) Spatial patterns of woody plant and bird diversity: functional relationships or environmental effects? *Global Ecology and Biogeography* 17: 327–339. https://doi.org/10.1111/j.1466-8238.2007.00379.x

Kuhn M (2008) Building predictive models in R Using the caret package. *Journal of Statistical Software* 28: 1–-26. https://doi.org/10.18637/jss.v028.i05

Lembrechts JJ, Nijs I, Lenoir J (2019) Incorporating microclimate into species distribution models. *Ecography* 42: 1267–1279. https://doi.org/10.1111/ecog.03947

Leurs G, Nieuwenhuis BO, Zuidewind TJ, Hijner N, Olff H, Govers LL (2023) Where land meets sea: Intertidal areas as key-habitats for sharks and rays. *Fish and Fisheries* 24: 407–426. https://doi.org/10.1111/faf.12735

Liang M, Pause M, Prechtel N, Schramm M (2020) Regionalization of coarse scale soil moisture products using fine-scale vegetation indices-prospects and case study. *Remote Sensing* 12. https://doi.org/10.3390/rs12030551

Liu C, Newell G, White M (2015) On the selection of thresholds for predicting species occurrence with presence-only data. *Ecology and Evolution* 5: 337–348. https://doi.org/10.1002/ece3.1878

Macarthur RH, Macarthur JW (1961) On bird species diversity. *Ecology* 42: 594–598. https://doi.org/10.2307/1932254

MAE (Ministerio Del Ambiente Ecuador) (2013) Sistema de Clasificación de Ecosistemas para el Ecuador Continental. Quito.

Marston C, Raoul F, Rowland C, Quéré JP, Feng X, Lin R, Giraudoux P (2023) Mapping small mammal optimal habitats using satellite-derived proxy variables and species distribution models. *PLoS ONE* 18: e0289209. https://doi.org/10.1371/journal.pone.0289209

Mcpherson JM, Jetz W, Rogers DJ (2004) The effects of species' range sizes on the accuracy of distribution models: Ecological phenomenon or statistical artefact? *Journal of Applied Ecology* 41: 811–823. https://doi.org/10.1111/j.0021-8901.2004.00943.x

Meirmans PG (2021) Niche divergence contributes to geographical parthenogenesis in two dandelion taxa. *Journal of Evolutionary Biology* 34: 1071–1086. https://doi.org/10.1111/jeb.13794

Merow C, Smith MJ, Silander JA Jr (2013) A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* 36: 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Mills SC, Socolar JB, Edwards FA, Parra E, Martínez-Revelo DE, Ochoa-Quintero JM, Edwards DP (2023). High sensitivity of tropical forest birds to deforestation at lower altitudes. *Ecology* 104: e3867. https://doi.org/10.1002/ecy.3867

Moudrý V, Moudrá L, Barták V, Bejček V, Gdulová K, Hendrychová M, Moravec D, Musil P, Rocchini D, Šťástný K, Volf O, Šálek M (2021) The role of the vegetation structure, primary productivity and senescence derived from airborne LiDAR and hyperspectral data for bird's diversity and rarity on a restored site. *Landscape and Urban Planning* 210. https://doi.org/10.1016/j.landurbplan.2021.104064

Muschelli J (2020) ROC and AUC with a binary predictor: A potentially misleading metric. *Journal of Classification* 37: 696–708. https://doi.org/10.1007/s00357-019-09345-1

Navarrete E, Morante-Carballo F, Dueñas-Tovar J, Carrión-Mero P, Jaya-Montalvo M, Berrezueta E (2022) Assessment of geosites within a natural protected area: A case study of Cajas National Park. *Sustainability* 14: 3120. https://doi.org/10.3390/su14053120

Omweno JO, Getabu A, Orina PS, Omasaki SK, Zablon WO (2021) PLS regression to compare the relative importance of water quality variables in fish growth. *Journal of Applied Structural Equation Modeling* 5: 1-8. https://doi.org/10.47263/JASEM.5(1)01

Onoh UC, Ogunade J, Owoeye E, Awakessien S, Asomah JK (2024) Impact of climate change on biodiversity and ecosystems services. *IIARD International Journal of Geography and Environmental Management* 10: 77–93.

Osborne OG, Fell HG, Atkims H, Van tol J, Phillips D, Herrera-Alsina L et al. (2022) Fauxcurrence: simulating multi-species occurrences for null models in species distribution modeling and biogeography. *Ecography* 2022: e05880. https://doi.org/10.1111/ecog.05880

Pebesma E, Bivand RS (2005). Clases y métodos S para datos espaciales: el paquete sp. *R news*, *5* (2), 9-13.

Pereira FR De S, Kampel M, Soares MLG, Estrada GCD, Bentz C, Vincent G (2018) Reducing uncertainty in mapping of mangrove aboveground biomass using airborne discrete return lidar data. *Remote Sensing* 10: 637. https://doi.org/10.3390/rs10040637

Pettorelli N, Laurance WF, O'brien TG, Wegmann M, Nagendra H, Turner W (2014) Satellite remote sensing for applied ecologists: Opportunities and challenges. *Journal of Applied* Ecology 51: 839–848. https://doi.org/10.1111/1365-2664.12261

Pettorelli N, Ryan S, Mueller T, Bunnefeld N, Jedrzejewska B, Lima M, Kausrud K (2011) The normalized difference vegetation index (NDVI): Unforeseen successes in animal ecology. *Climate Research* 46: 15–27. https://doi.org/10.3354/cr00936

Phillips SJ (2005) A brief tutorial on Maxent. New Jersey.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modeling* 190: 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography* 31: 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME (2017) Opening the black box: An open-source release of Maxent. *Ecography* 40: 887–893. https://doi.org/10.1111/ecog.03049

Qiao C, Luo J, Sheng Y (2012) An adaptive water extraction method from remote sensing image based on NDWI. *ournal of the Indian Society of Remote Sensing* 40: 421–433. https://doi.org/10.1007/s12524-011-0162-7

R Development Core Team (2019) A language and environment for statistical computing (3.6.1.) R Foundation for Statistical Computing. https://cran.r-project.org/

Raes N, Steege H (2007) A null-model for significance testing of presence-only species distribution models. *Ecography* 30: 727–736. https://doi.org/10.1111/j.2007.0906-7590.05041.x

Rahbek C, Graves GR (2000) Detection of macro-ecological patterns in South American hummingbirds is affected by spatial scale. *Proceedings of the Royal Society B: Biological Sciences* 267: 2259–2265. https://doi.org/10.1098/rspb.2000.1277

Riaño D, Chuvieco E, Salas J, Aguado I (2003) Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. *IEEE Transactions on Geoscience and Remote Sensing* 41: 1056–1061. https://doi.org/10.1109/TGRS.2003.811693

Rödder d, Lötters S (2009) Niche shift versus niche conservatism? climatic characteristics of the native and invasive ranges of the mediterranean house gecko *(Hemidactylus turcicus). Global Ecology and Biogeography* 18*:* 674–687. https://doi.org/10.1111/j.1466-8238.2009.00477.x

Rodríguez Saltos CA, Bonaccorso E (2016) Understanding the evolutionary history of a high Andean endemic: The Ecuadorian hillstar *Oreotrochilus chimborazo. Neotropical Biodiversity* 2: 37–50. https://doi.org/10.1080/23766808.2016.1155280

Rotenberry JT (1985) The role of habitat in avian community composition: Physiognomy or floristics? *Oecologia* 67: 213–217.

Santillán V, Quitián M, Tinoco Ba, Zárate E, Schleuning M, Böhning-Gaese K et al. (2018) Spatio-temporal variation in bird assemblages is associated with fluctuations in temperature and precipitation along a tropical elevational gradient. *PLoS ONE* 13: e0196179. https://doi.org/10.1371/journal.pone.0196179

Shmueli G, Sarstedt M, Hair JF, Cheah JH, Ting H, Vaithilingam S, Ringle CM (2019) Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing* 53: 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

Stoica IA (2018) An interpretation of multi-model future climate predictions for bioclim variables in Romania. *Contributii Botanice* 53. https://doi.org/10.24193/Contrib.Bot.53.8

Su H, Bista M, Li M (2021) Mapping habitat suitability for Asiatic black bear and red panda in Makalu Barun National Park of Nepal from Maxent and GARP models. *Scientific Reports* 11: 14135. https://doi.org/10.1038/s41598-021-93540-x

Sullivan BL, Aycrigg JL, Barry JH, Bonney RE, Bruns N, Cooper CB, Damoulas et al. (2014) The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation* 169: 31-40. https://doi.org/10.1016/j.biocon.2013.11.003

Tassi A, Gigante D, Modica G, Di Martino L, Vizzari M (2021) Pixel-vs. Object-based Landsat 8 data classification in google earth engine using Random Forest: The case study of Maiella national park. *Remote Sensing* 13. https://doi.org/10.3390/rs13122299

Tinoco Ba, Freile Jf, Molina P, Carrasco A, Ordoñez N, Bonaccorso E (2023) Ecology and distribution of the "Critically Endangered" Blue-throated Hillstar *Oreotrochilus cyanolaemus*. *Bird Conservation International* 33: e60. https://doi.org/10.1017/S0959270923000114

Tinoco BA, Astudillo PX, Latta SC, Graham CH (2009) Distribution, ecology and conservation of an endangered Andean hummingbird: The Violet-throated Metaltail (*Metallura baroni*). *Bird Conservation International* 19: 63–76. https://doi.org/10.1017/S0959270908007703

Tomlinson S, Lewandrowski W, Miller BP, Turner SR, Elliott CP (2020) High-resolution distribution modeling of a threatened short- range endemic plant informed by edaphic factors. *Ecology and Evolution* 10: 763–777. https://doi.org/10.1002/ece3.5933

Urbanek S (2026) rJava: Low-Level R to Java Interface. https://doi.org/10.32614/CRAN.package.rJava, R package version 1.0-18.

Valerio F, Ferreira E, Godinho S, Pita R, Mira A, Fernandes N, Santos SM (2020) Predicting microhabitat suitability for an endangered small mammal using Sentinel-2 data. *Remote Sensing* 12: 562. https://doi.org/10.3390/rs12030562

Vanderwal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modeling* 220: 589–594. https://doi.org/10.1016/j.ecolmodel.2008.11.010

van Proosdij ASJ, Sosef MSM, Wieringa JJ, Raes N (2016) Minimum required number of specimen records to develop accurate species distribution models. *Ecography* 39: 542–552. https://doi.org/10.1111/ecog.01509

Vaughan IP, Ormerod SJ (2005) Increasing the value of principal components analysis for simplifying ecological data: A case study with rivers and river birds. *Journal of Applied Ecology* 42: 487–497. https://doi.org/10.1111/j.1365-2664.2005.01038.x

Velazco SJE, Rose MB, De Andrade AFA, Minoli I, Franklin J (2022) Flexsdm: An r package for supporting a comprehensive and flexible species distribution modeling workflow. *Methods in Ecology and Evolution* 13: 1661–1669. https://doi.org/10.1111/2041-210X.13874

Wallis CIB, Paulsch D, Zeilinger J, Silva B, Curatola Fernandez GF, Brandl R, Farwig N, Bendix J (2016) Contrasting performance of Lidar and optical texture models in predicting avian diversity in a tropical mountain forest. *Remote Sensing of Environment* 174: 223–232. https://doi.org/10.1016/j.rse.2015.12.019

Wallis CIB, Brehm G, Donoso DA, Fiedler K, Homeier J, Paulsch D, Sÿssenbach D, Tiede Y, Brandl R, Farwig N, Bendix J (2017) Remote sensing improves prediction of tropical montane species diversity but performance differs among taxa. *Ecological Indicators* 83: 538–549. https://doi.org/10.1016/j.ecolind.2017.01.022

Wallis CIB, Crofts AL, Inamdar D, Arroyo-Mora JP, Kalacska M, Laliberté É, Vellend M (2023) Remotely sensed carbon content: The role of tree composition and tree diversity. *Remote Sensing of Environment* 284: 113333. https://doi.org/10.1016/j.rse.2022.113333

Wood EM, Pidgeon AM, Radeloff VC, Keuler NS (2013) Image texture predicts avian density and species richness. *PLoS ONE* 8: e63211. https://doi.org/10.1371/journal.pone.0063211

Wood EM, Pidgeon A M, Radeloff VC, Keuler NS (2012) Image texture as a remotely sensed measure of vegetation structure. *Remote Sensing of Environment* 121: 516–526. https://doi.org/10.1016/j.rse.2012.01.003

Worldclim (2021) Bioclimatic variables. Global climate and weather data. https://www.worldclim.org/data/bioclim.html

Zeng L, Wardlow BD, Xiang D, Hu S, Li D (2020) A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment* 237: 11511. https://doi.org/10.1016/j.rse.2019.111511

Zvoleff A (2015) GLCM: Calculate textures from grey-level co-occurrence matrices (GLCMs) in R (Version 1.2) https://doi.org/10.32614/CRAN.package.glcm

Authors

Prof. Edwin Zárate, M.Sc.
https://orcid.org/0000-0001-5124-5436
ezarate@uazuay.edu.ec
Universidad del Azuay
Escuela de Biología
Av. 24 de mayo 7-77 y Hernán Malo
Cuenca
Ecuador

Dr. Christine Wallis
https://orcid.org/0000-0002-0794-3146
christine.wallis@tu-berlin.de
Technische Universität Berlin
Geoinformation in Environmental Planning
Berlin
Germany

Dr. Vinicio Santillán
https://orcid.org/0000-0002-4296-580X
vinicio.santillanr@ucacue.edu.ec
Universidad Católica de Cuenca
Av. de las Américas y Humboldt
Cuenca
Ecuador

Prof. Dr. Roland Brandl
brandl@staff.uni-marburg.de
Philipps-Universität Marburg
Faculty of Biology
Department of EcologyAnimal Ecology
Karl-von-Frisch-Straße 8
D-35043 Marburg
Germany

Prof. Dr. Nina Farwig
https://orcid.org/0000-0002-0554-5128
farwig@biologie.uni-marburg.de
Philipps-Universität Marburg
Faculty of Biology
Conservation Ecology
Karl-von-Frisch-Straße 8
D-35043 Marburg
Germany

Prof. Dr. Jörg Bendix
https://orcid.org/0000-0001-6559-2033
bendix@staff.uni-marburg.de
Philipps-Universität Marburg
Faculty of Geography
Laboratory for Climatology and Remote Sensing
Deutschhausstraße 12
D-35037 Marburg
Germany

Appendix

Tab. A1: Variable contribution in Maxent models using four different predictor combinations. Estimates of the relative contributions (% contrib) of the environmental variables are presented. The model was evaluated on the permuted data, and the resulting decrease in training AUC is shown in the table, normalized to percentages (permutation importance), (Predictor selection 3.1).

Environmental variables	BC		BC+NV		BC+NW		BC+NV+NW	
	% contrib	Perm important	% contrib	Perm important	% contrib	Perm important	% contrib	Perm important
Max temp warmest month (bio5)	78.9	48.2	73.6	42.3	75.8	54.4	71.4	39.7
Precipitation seasonality (bio15)	8.4	14.1	7.9	23.3	9.3	18.6	8.3	16.1
Precipitation wettest quarter (bio16)	12.7	37.7	9.6	29.3	11.7	24.9	10.2	34
NDVI entropy	n.a.	n.a.	8.2	4.3	n.a.	n.a.	8.5	4.8
NDVI variance	n.a.	n.a.	0.7	0.9	n.a.	n.a.	1	2
NDWI entropy	n.a.	n.a.	n.a.	n.a.	2.3	1.1	0.2	0.4
NDWI variance	n.a.	n.a.	n.a.	n.a.	1	1	0.4	2.9

n.a = no aport

Tab. A2: Permutation analysis used to determine the significance of AUC values between the Maxent and maxSSS models (Model evaluation 3.2, 4.2)

Variable	Maxent Threshold	P_value	maxSSS Threshold	*p_value*
BC	0.94	1	0.78	0
BC + NDVI	0.97	0	0.79	0
BC + NDWI	0.95	0	0.79	0
BC + NDVI + NDWI	0.96	0	0.81	0

If p_value < 0.05, the real presence points are significantly more associated with pixels above the threshold than random points. If p-value is high, the association may be attributed to chance.

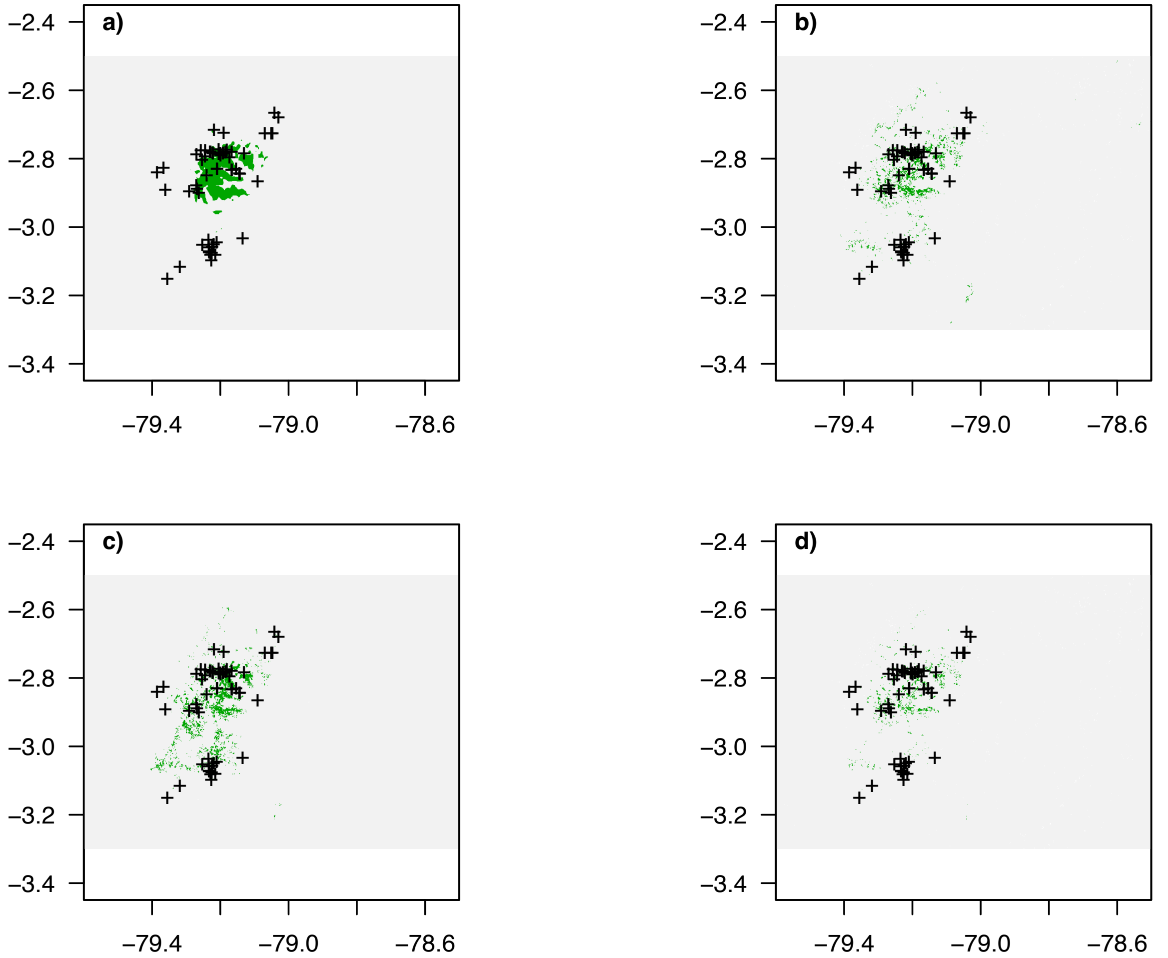

Fig. A1: Binary map (green) represents areas selected by the maxSSS threshold from MaxEnt models for the predicted presence of *M. baroni*, using (a) BC variables, (b) BC+NV variables, c) BC+NW variables, and (d) BC+NV+NW variables. Crosses indicate records of occurrence of *M. baroni*.