



**Uso de los modelos de distribución de especies como herramienta para la conservación en
Colombia**

Yuleimis Martínez Caballero

Monografía presentada para optar al título de Especialista en Medio Ambiente y Geoinformática

Asesor

Sebastián Ríos Cortes, Especialista (Esp) en Medio ambiente y Geoinformática

Universidad de Antioquia
Facultad de Ingeniería
Especialización en Medio Ambiente y Geoinformática
Medellín, Antioquia, Colombia
2022

Cita	(Martínez Caballero, 2022)
Referencia	Martínez Caballero, Y. (2022). <i>Uso de los modelos de distribución de especies como herramienta para la conservación en Colombia</i> [Trabajo de grado especialización]. Universidad de Antioquia, Medellín, Colombia.
Estilo APA 7 (2020)	



Especialización en Medio Ambiente y Geoinformática, Cohorte XVII.
 Grupo de Investigación Seleccione grupo de investigación UdeA (A-Z).
 Seleccione centro de investigación UdeA (A-Z).



Elija un elemento.

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano/Director: Jesús Francisco Vargas Bonilla.

Jefe departamento: Diana Catalina Rodríguez Loaiza.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Contenido

Resumen	6
Abstract	7
1. Introducción	8
2. Objetivos	10
2.1. Objetivo general	10
2.2. Objetivos específicos.....	10
3. Marco teórico	11
4. Metodología	14
4.1. Revisión bibliográfica	14
4.2. Modelación de distribución.....	14
5. Resultados	15
5.1. Revisión bibliográfica	15
5.1.1. Datos de ocurrencia.....	15
5.1.2. Variables ambientales	16
5.1.3. Definición de las áreas de modelado.....	17
5.1.4. Modelamiento y evaluación de modelos.....	17
5.2. Modelación de la especie <i>Hyalinobatrachium colymbiphyllum</i>	19
6. Análisis de resultados	26
6.1. Análisis de los modelos.....	26
6.2. Modelado de la distribución de la especie <i>H. colymbiphyllum</i>	27
6.3. Limitaciones de los SDM.....	27
7. Conclusiones	29
8. Referencias.....	30

Lista de tablas

Tabla 1. Resumen de artículos seleccionados	15
Tabla 2. Resumen de datos empleados en el modelado	16
Tabla 3. Variables bioclimáticas empleadas en el modelo.....	21
Tabla 4. Límite de la predicción.....	23
Tabla 5. Análisis de la contribución de las variables.	24

Lista de figuras

Figura 1. Mapas de distribución de las especies objeto de estudio producto de las modelaciones empleando MAXENT. a) Área potencial de distribución de <i>Panthera onca</i> en la Sierra Nevada de Santa Marta (Zárrate et al., 2018) b) Distribución potencial de 16 especies de los géneros Marmosa and Monodelphis en Colombia (González, et al., 2022) c) Distribución potencial de <i>Batrachochytrium dendrobatidis</i> en Colombia (Flechas et al., 2017) d) Distribución potencial del pez invasor <i>Pangasianodon hypophthalmus</i> en Colombia (Castellanos et al., 2021).....	19
Figura 2. Distribución de la especie <i>Hyalinobatrachium colymbiphyllum</i>	20
Figura 3. Resultado modelo de distribución de la especie <i>H.colymbiphyllum</i>	22
Figura 4. Especificidad del modelo para <i>H. colymbiphyllum</i>	23
Figura 5. Prueba de Jackknife para a) el entrenamiento del modelo, b) prueba del modelo.	25

Resumen

Los modelos de distribución de especies son herramientas que se han desarrollado para predecir la distribución potencial de especies y los factores que pueden afectarla, ya sean bióticos o abióticos. Información que se está empleando como base en la toma de decisiones de conservación. En Colombia se han desarrollado diversas investigaciones cuyo objeto principal ha sido aumentar el conocimiento de la distribución de especies de interés (en peligro, endémicas, etc) para realizar esfuerzos en torno a su conservación. En el presente trabajo se realiza el análisis de investigaciones relevantes desarrolladas en el tema en Colombia, teniendo en cuenta los modelos empleados, la adquisición y preparación tanto de las variables dependientes como predictoras, así como los comentarios y sugerencias. Adicionalmente, utilizando el algoritmo de máxima entropía Maxent - empleado en los artículos seleccionados., se realizó la modelación de la distribución de la especie de Anuro *Hyalinobatrachium colymbiphyllum* en Colombia, encontrando resultados acordes con las observaciones en campo y la biología de la especie de estudio.

Palabras clave: Modelos de distribución de especies, Maxent, Conservación.

Abstract

Species distribution models have been developed to predict the potential distribution of species and the factors that may affect this distribution, such as biotic or abiotic variables. This information is being used as a basis for conservation decision-making. In Colombia, several investigations have been developed whose main objective has been to increase knowledge of the distribution of species of interest (endangered, endemic, etc.) to carry out efforts around their conservation. We analyze relevant research developed in Colombia, taking into account the models used, the acquisition and preparation of both the dependent and predictor variables, as well as the comments and suggestions. Additionally, using the Maxent maximum entropy algorithm -used in the selected articles-, We model the distribution of the Anuran species *Hyalinobatrachium colymbiphyllum* in Colombia, finding results consistent with the observations in the field and the biology of the study specie.

Keywords: Species distribution models, Maxent, Conservation

1. Introducción

En las últimas décadas ha aumentado el uso de herramientas de información geográfica para tratar de solucionar problemas ecológicos que antes no podían ser resueltos por limitaciones en la obtención de la información, problemas como conocer de qué depende que un organismo exista en un determinado sitio (Naoki et al., 2006), o de qué factores depende su permanencia en el mismo; para dar respuesta a estos interrogantes surgen los modelos de distribución, de hábitat o de nicho ecológicos de especies.

Los modelos de distribución de especies (SDM) son herramientas que se han empleado para explicar y predecir la distribución en un hábitat determinado (Cayuela et al., 2009), ya sea natural o invadido; así como determinar las relaciones que limitan la ocurrencia de dichas especies con diferentes factores bióticos y abióticos (Krebs, 1985), basados en correlaciones entre registros de ocurrencia conocidos y las condiciones ambientales en las localidades de ocurrencia (Phillips et al., 2006; Elith & Leathwick, 2009; Gomes et al., 2018); establecidas bajo el principio de nicho ecológico. Como resultado de la aplicación de los SDM se obtienen mapas geográficos de la distribución de la o las especies objeto de estudio (Miller, 2010).

Los modelos de distribución de especies son una de las áreas de investigación con mayor desarrollo en el campo de la biogeografía de la conservación (Richardson & Whittaker, 2010); actualmente son empleados en diferentes áreas del conocimiento: macroecología, biogeografía, investigación de biodiversidad, conservación, restauración, entre otros (Gomes et al., 2018), para dar respuesta a un sinnúmero de interrogantes tales como la selección de áreas para protección y conservación, la predicción de los efectos del cambio climático en el rango distribución de una especie, la determinación del riesgo de invasiones (Araujo & Peterson, 2012; Pearson, 2010), la evaluación de amenazas a las que está sometida la especie de interés (Kremen et al., 2008), etc,

Colombia es un país megadiverso con aproximadamente 67.000 especies, ocupando el segundo lugar en el mundo con mayor número de especies registradas; sin embargo, actualmente su diversidad se está viendo afectada por diferentes presiones, por lo que aproximadamente el 2% de las especies se encuentran dentro de alguna categoría de amenaza (Aprox 1.302); adicionalmente se reporta la introducción de especies exóticas, alrededor de 508 especies de las cuales 22 son consideradas invasoras (Baptiste et al. 2020). Por este motivo las investigaciones que aporten a dar respuesta a problemas de conservación en el país cada día toman mayor relevancia.

El presente trabajo tiene como finalidad analizar el uso de los SDM como herramienta para obtener información que aporte la conservación de fauna silvestre en Colombia, a través del análisis de los métodos empleados en cuatro investigaciones realizadas en los diferentes grupos faunísticos en los que se analizó el potencial invasor de una especie de pez (Castellanos-et al.,2021), las posibles áreas de conservación para proteger a una especie de felino (Zárrate et al., 2018), la distribución potencial de especies de la tribu Marmosini (González, et al, 2022) y la relación de la distribución de un hongo que afecta a los anfibios del país con algunas variables ambientales (Flechas et al., 2017). Y con base en las metodologías empleadas y los hallazgos de dichas investigaciones, realizar el modelado de la distribución potencial de una especie de anuro de la familia Centrolenidae, *Hyalinobatrachium colymbiphllum* en Colombia.

2. Objetivos

2.1. Objetivo general

Realizar una revisión bibliográfica del uso de los modelos de distribución de especies como herramienta para la conservación de fauna silvestre en Colombia.

2.2. Objetivos específicos

- Seleccionar los artículos más relevantes de investigaciones realizadas en Colombia que empleen modelos de distribución de especies como aporte a la conservación de especies.
- Analizar la información y metodologías empleadas en la elaboración de los modelos de distribución de especies en cada uno de los artículos seleccionados.
- Realizar la modelación de la distribución de una especie de fauna empleando la metodología más aceptada en las investigaciones encontradas.
- Se describen algunos ejemplos de verbos comunes que se utilizan en el planteamiento de objetivos, los cuales cambiarán dependiendo de su investigación.

3. Marco teórico

La biogeografía es la disciplina de la biología comparada que se encarga del estudio de la distribución geográfica de los seres vivos en el tiempo y el espacio, tomando en cuenta los diversos factores que dieron lugar a tal distribución (Contreras, 2006); identificar y caracterizar las áreas de distribución de las especies ha sido el principio fundamental en el desarrollo del conocimiento biogeográfico (Lomolino et al., 2010).

Por su parte, el nicho ecológico se define como todas aquellas condiciones óptimas en el hipervolumen n-dimensional (condiciones bióticas y abióticas) en las cuales la especie puede y podría lograr su desarrollo y subsistencia (Hutchinson 1957). El nicho ecológico y la distribución de las especies están estrechamente relacionados; el área de distribución es generalmente representado mediante mapas, los cuales son abstracciones de diferentes observaciones de patrones temporales y espaciales de los sitios que ocupan las especies.

Para estimar la distribución potencial de las especies, se desarrollaron los denominados modelos de distribución de especies (SDM, species distribution models) o modelos de nicho ecológico (ENM, ecological niche model) (Guisan y Zimmermann, 2000; Soberón y Peterson 2005).

Los **modelos de distribución de especies** son representaciones cartográficas de la idoneidad de un espacio para la presencia de una especie, en función de las variables empleadas para generar dicha representación. Entendiéndose como idoneidad la relación matemática o estadística entre la distribución real conocida y un conjunto de variables independientes que se usan como indicadores. Dichas variables suelen ser geológicas, topográficas o climáticas, y se espera que, con algunas de ellas, individualmente o en combinación, se puedan definir los factores ambientales que delimiten las condiciones favorables para la presencia de la especie (Guisan & Zimmermann 2000).

La construcción de modelos de distribución de especies es un proceso de clasificación. La variable dependiente es dicotómica (presencia/ ausencia) y las independientes o predictoras pueden ser cuantitativas (e.g., temperatura o elevación) o nominales (e.g., litología o uso del suelo). Los métodos implicados son clasificadores que deben generar un valor numérico para cada punto del terreno. Dicho valor refleja, directa o indirectamente, la idoneidad de presencia de la especie en función de los valores locales de las variables independientes (Guisan & Zimmermann 2000).

Existen diferentes tipos de modelos en función de sus planteamientos teóricos y el tipo de datos que utilizan, presencias y pseudoausencias (Mota et al., 2020) (tabla 1).

Tabla 1. Algunos algoritmos usados para el ENM y los SDM. Adaptado de Mota et al., 2020.

Nombre	Algoritmo	Fuente	Datos biológicos
GLM	Modelos lineales generalizados	s McCullag y Nelder (1989); Guisan et al. (2002)	Presencias/Pseudoausencias
GAM	Modelos Aditivos Generalizados	Hastie y Tibshirani (1990); Guisan et al. (2002)	Presencias/Pseudoausencias
BRT	Árboles de regresión impulsado	Elith et al. (2008)	Presencias/Pseudoausencias
BIOCLIM	Envolturas bioclimáticas	Busby (1991)	Presencias
ENFA	Análisis factorial del nicho ecológico	Hirzel et al. (2002)	Presencias
MAXENT	Máxima Entropía	Phillips et al. (2006)	Presencias/Entornos
GARP	Algoritmos Genéticos	Stockwell y Peters (1999)	Presencias/Entornos
CONVEXHULL	Elipsoides de volumen	Guisan y Zimmermann, (2000)	Presencias
MVE	Elipsoides de volumen mínimo	Van Aelst y Rousseeuw (2009); Qiao et al. (2016)	Presencias
KDE	Estimación de la densidad de Kernell	Blonder et al. (2014)	Presencias
MA	Marble	Qiao et al. (2015b)	Presencias
DOMAIN	Distancia de Gower (disimilaridad entre ambientes)	Carpenter et al. (1993)	Presencias

Los SDM también se pueden discriminar cuatro grupos: las relacionadas a **modelos estadísticos de regresiones** (Modelos Lineales Generalizados (GLM), Modelos Aditivos Generalizados (GAM)), los **métodos de clasificación** (Boosted regression trees (BRT)), los **métodos de "sobre"** (BIOCLIM, ENFA) y aquellos basados en algoritmos específicos (GARP, MAXENT) (Pliscoff & Fuentes, 2011).

Dentro de los SDM basados en algoritmos específicos, se encuentra **MaxEnt**, el cual es un método de inteligencia artificial que aplica el principio de máxima entropía para calcular la distribución geográfica más probable para una especie. Estima la probabilidad de ocurrencia de la especie buscando la distribución más uniforme posible, sujeta a la condición de que el valor esperado de cada variable ambiental según esta distribución coincida con su media empírica. El resultado del

modelo expresa el valor de idoneidad del hábitat para la especie como una función de las variables ambientales. Un valor alto de la función de distribución en una celda determinada indica que ésta presenta condiciones muy favorables para la presencia de la especie. MaxEnt puede utilizar variables cualitativas, otorgando a cada valor de la variable un peso relativo al número total de puntos de presencia que contiene. El programa proporciona las curvas de respuesta de la especie ante las distintas variables ambientales y estima la importancia de cada variable en la distribución de la especie (Phillips et al., 2006, Benito & Peñas, 2007).

Los SDM se generan utilizando dos tipos de información, los datos de presencia/ausencia de la especie que se quiere modelar y las variables descriptoras o predictoras que definirán el espacio ambiental donde se distribuirá en el espacio geográfico (Pliscoff & Fuentes, 2011).

Los datos de presencias o ausencias se pueden generar mediante la obtención de información georreferenciada en campo o recurriendo a las fuentes directas de información de especies, como son los museos o herbarios o de forma indirecta mediante el uso de bases de datos de libre acceso, como lo es la base de datos de especies GBIF (Global Biodiversity Information Facility) que recopila la información de colectas para todo tipo de organismos, desde diferentes fuentes (museos, herbarios, estudios, entre otros) (Pliscoff & Fuentes, 2011).

Por su parte la elección del tipo de variable dependerá de los objetivos que busca la identificación y caracterización del espacio ambiental donde se distribuye la especie. Existen principalmente dos tipos de variables que se utilizan comúnmente, las variables climáticas (bioclimáticas) y las variables topo-edáficas. Las primeras son las más utilizadas, principalmente por la facilidad de su obtención, ya que existen bases de datos globales a alta resolución (~1 km) de libre acceso y uso. Una de las estrategias de selección más utilizada es trabajar con un set de variables bioclimáticas predefinidas, como es el caso de las 19 variables bioclimáticas disponibles en la base de datos global de superficies climáticas Worldclim (Fick & Hijmans, 2017).

4. Metodología

4.1.Revisión bibliográfica

Se realiza la búsqueda de los artículos científicos de investigaciones que emplearan el uso de modelos de distribución como herramienta para aportar a la conservación de fauna silvestre en los últimos 10 años.

La búsqueda fue ejecutada en las plataformas de información bibliográfica SpringerLink y Google Scholar, usando los campos: título, palabras clave y resumen, estableciendo los años de búsqueda 2017-2022 y las palabras clave en inglés: distribution models y Colombia.

Se tuvieron en cuenta las investigaciones cuyo cumplimiento de sus objetivos aportara a la conservación de las especies de estudio y que estas fueran de interés, ya sea por ser sensibles, presentar muy poca información o encontrarse bajo alguna categoría de amenaza.

Se realiza un análisis conceptual y metodológico de cada uno de los artículos seleccionados, describiendo los aspectos tenidos en cuenta por los autores para realizar los modelos: selección de especies, información de referencia, variables predictoras, métodos de selección de variables, entre otros; así como las sugerencias y comentarios de dichos modelos.

4.2.Modelación de distribución

Teniendo en cuenta los hallazgos de los autores se realiza la selección de una metodología para modelar la distribución de una especie.

Para el proceso de modelación, se definió como área de estudio el territorio continental de Colombia. La información de la distribución de la especie se extrajo de la página web Global Biodiversity Information Facility – GBIF (2022), mediante el complemento para QGIS, GBIF Occurrences desarrollado por el Belgian Biodiversity Platform.

Para definir las variables que alimentaran el modelo de distribución de la especie, se emplearon las diecinueve (19) variables bioclimáticas disponibles en la plataforma WorldClim (Fick y Hijmans, 2017). El área de modelado fue definida mediante ARCGIS.

Para la modelación de la especie se empleó el software libre Maxent (Phillips et al., s.f), empleando como parámetros los datos de ocurrencia de la especie y las variables bioclimáticas. Para conocer las variables que mejor explicaban la distribución de la especie se realizó en Maxent, una prueba de Jackknife.

5. Resultados

5.1.Revisión bibliográfica

Se seleccionaron cuatro artículos de investigaciones realizadas en Colombia, en los que se emplearon los modelos de distribución de especies para aportar información sobre la conservación, tales como, el potencial invasor de una especie de pez (Castellanos-et al.,2021), la distribución de un patógeno sobre los anfibios (Flechas et al., 2017), identificación de áreas de conservación para una especie de felino (Zárrate et al., 2018) y aumento en el conocimiento de la distribución de un grupo de especies de mamíferos (González, et al, 2022)

En la Tabla 1 se presentan los objetivos de cada uno de los artículos seleccionados, así como la principal especie o grupo de especies de estudio.

Tabla 1. Resumen de artículos seleccionados

Artículo	Especie principal de estudio	Objetivos
Castellanos-et al.,2021	<i>Pangasianodon hypophthalmus</i>	Evaluar el potencial invasor de <i>P. hypophthalmus</i> en algunas de las cuencas de la cuenca del Mar Caribe. (1) hacer una comparación entre el nicho nativo y el nicho de introducción en la cuenca del Caribe, (2) Calcular la posible área de invasión en Colombia (3) comparar la distribución geográfica de <i>P. hypophthalmus</i> con la distribución de tres especies nativas presentes en ríos de la cuenca del Caribe.
Flechas et al., 2017	<i>Batrachochytrium dendrobatidis</i>	Evaluar la relación de las variables climáticas sobre la probabilidad de ocurrencia de Bd
Zárrate et al., 2018	<i>Panthera onca</i>	Identificar áreas donde la conservación y el manejo serían más eficientes con base en la identificación y relación de tres criterios espaciales: (1) parches de hábitat adecuados, (2) áreas de conectividad y (3) áreas con una mayor probabilidad de ocurrencia de conflictos entre humanos y jaguares en la ecorregión de Sierra Nevada.
González, et al, 2022	Tribu Marmosini (Géneros <i>Marmosa</i> and <i>Monodelphis</i>)	Modelar y analizar la distribución potencial de las especies de Marmosini en Colombia. Analizar el patrón espacial de riqueza y conservación de especies en todo el país.

5.1.1. Datos de ocurrencia

Los datos de ocurrencia de los artículos fueron tomados fuentes tales como el Global Biodiversity Information Facility (GBIF.org, 2021), información secundaria de investigaciones realizadas, reportes de autoridades y muestreos en campo.

5.1.2. Variables ambientales

La información de variables predictoras fue tomada de EarthEnv, WorldClim, ENVIREM, Modified Soil Adjusted Vegetation Index (MSAVI).

En cada uno de los artículos, los investigadores emplearon diferentes criterios para seleccionar las variables predictoras que se emplearían en los modelos. Flechas et al. (2017) aplicaron el coeficiente de correlación de Spearman bilateral, para conocer las variables que se encontraban relacionadas con la distribución de la especie de estudio; Castellanos-et al.,2021 y González et al (2022) seleccionaron variables que tuvieran relación con la biología de las especies analizadas, de acuerdo con la bibliografía consultada y Zárrate et al, (2018) seleccionaron las variables de acuerdo con un análisis Jackknife y de contribución variable (eligiendo las que presentaron la mayor contribución) una vez corrieron el modelo.

En la Tabla 2 se presenta el resumen del número de especies, registros y variables con los que se construyeron los modelos, así como el área de estudio, grupo taxonómico y resolución de los datos.

Tabla 2. Resumen de datos empleados en el modelado

Artículo	Región	Registro de especies			Variables predictoras	
		Grupo faunístico	Número de especies	Número de registros	Clase amplia	Tamaño de celda
González, et al., 2022	Colombia	Mamíferos	16	618	8 climáticas 2 topográficas 1 vegetación	1km ²
Castellanos et al.,2021	Cuenca Caribe	Peces	4	365	31 climáticas 8 topográficas 20 suelo 9 cobertura 2 caudal 1 corriente	1km ²
Zárrate et al., 2018	Sierra Nevada Santa Marta	Mamíferos	1	69	Climáticas 1 topográfica 7 distancia	1km ²
Flechas et al., 2017	Colombia (Excepto Llanos Orientales y Amazonas)	Fungi	1	42	10 climáticas	5km ²

5.1.3. Definición de las áreas de modelado

Teniendo en cuenta que el área de estudio puede condicionar los métodos de evaluación de los modelos y su precisión (Barve et al., 2011), previo a la implementación, los investigadores realizaron la definición de las áreas en las que trabajarían los modelos.

Zárrate et al, (2018) identificaron todas las unidades biogeográficas donde la especie (*Panthera onca*) presentaba al menos un registro de ocurrencia dentro del área de estudio, de esta manera seleccionaron como área de modelado de los límites políticos de los tres departamentos que constituyen la Sierra Nevada. Por otro lado, Castellanos et al (2021) calcularon el área de movilidad (M) de la especie invasora (*P. hypophthalmus*), realizando una intersección entre los registros de presencia de las especies nativas también evaluadas (*Ageneiosus pardalis*, *Sorubim cuspicaudus* y *Pseudoplatystoma magdaleniatum*) con un mapa de cuencas hidrográficas y seleccionaron como área del modelo la M más grande dentro de las especies nativas (*A. pardalis*). González, et al., 2022 emplearon dos métodos para estimar el espacio geográfico con condiciones abióticas adecuadas históricamente disponibles para que las especies se dispersen (M): M1 y M2; M1 se generó realizando una zona de amortiguamiento de 300km² a las localidades puntuales donde se registraron las especies y M2 agregando una zona de influencia de 50 km² y seleccionando las ecorregiones que se superponían con ella. M1 fue diseñado para incluir áreas contiguas con condiciones climáticas similares basadas en la exploración visual para todas las especies y M2 siguió una combinación de dos áreas de modelado utilizadas previamente para modelados en marsupiales.

Por su parte, Flechas et al (2017) limitaron el área de estudio a Colombia, y seleccionaron las localidades basadas en los sitios donde obtuvieron muestras positivas para el patógeno (42 localidades) y donde se tuviera registros en la bibliografía de presencia del hongo (10 localidades).

5.1.4. Modelamiento y evaluación de modelos

Una vez seleccionadas las variables y registros y delimitada el área de estudio, se construyeron los modelos empleando el algoritmo de máxima entropía Maxent (Phillips et al. 2006).

Zárrate et al, (2018) emplearon diferentes multiplicadores de regularización y combinaciones de clases para estimar la complejidad óptima de los modelos, seleccionaron el modelo óptimo utilizando los Criterios de información de Akaike (AIC, donde se elige el que presenta el menor valor), y probaron su rendimiento utilizando un subconjunto de ocurrencias previamente

seleccionado al azar independiente del 15 % de acuerdo con lo propuesto por Guisan et al (2017). Verificaron la autocorrelación espacial (SA) usando el Índice Moran Global y evitaron el uso del modelo con presencia de esta y que se encontraran por fuera de los límites del área y prueba de calibración. Posteriormente transformaron el modelo final en una salida binaria, la cual fue definida en función de la altitud máxima y mínima notificada para la especie (0–3800 msnm) y finalmente la distribución potencial resultante fue una intersección entre nuestro modelo y los rangos de distribución existentes combinados.

Flechas et al. (2017) ajustaron los modelos empleando 80% de los datos para la construcción y el 20% para probarlos. Emplearon el Área Bajo la Curva (AUC) para medir la precisión predictiva del modelo en términos de sensibilidad y especificidad, cuyos valores varían de 0.5 (predicción no mejor que aleatoria) a 1 (discriminación perfecta), de acuerdo con Fielding & Bell (1997). Transformaron la distribución predicha de *B. dendrobatidis* como un mapa binario de presencia/ausencia, seleccionando como umbral la predicción más baja para datos de entrenamiento por encima del percentil 10. Finalmente, las preferencias climáticas del hongo se infirieron extrayendo valores de diez variables bioclimáticas de 10 000 puntos aleatorios del mapa binario y comparándolos valores de los tomados de los sitios predichos como adecuados con los predichos como inadecuados.

Castellanos et al (2021) seleccionaron el mejor modelo para cada especie, teniendo en cuenta el valor de AIC más bajo y construyeron un mapa binario utilizando el valor umbral de corte 10P (décimo percentil) para *P. hypophthalmus* y el valor umbral MTP (presencia mínima de entrenamiento) para las especies nativas.

González, et al (2022) realizaron los modelos para cada especie y los validaron empleando métodos de validación cruzada de bloques de acuerdo con el número de registros. Para seleccionar la mejor combinación de escenarios evaluaron los resultados a través de pasos: i) Con base en el valor de prueba promedio AUC, ii) eligiendo los modelos con mínima diferencia promedio AUC entre los modelos de prueba y entrenamiento, iii) modelos con prueba promedio mínima de tasa de omisión en la presencia mínima de entrenamiento (o MTP), iv) modelos con el valor mínimo para AIC corregido, y v) si quedaba más de un modelo después de los pasos anteriores, eligieron el modelo que maximizaba el multiplicador de regularización. Para transformar los mapas continuos en predicciones binarias, emplearon el valor que maximiza la suma de la sensibilidad y especificidad

y finalmente los rangos finales se ajustaron utilizando barreras geográficas conocidas para cada especie.

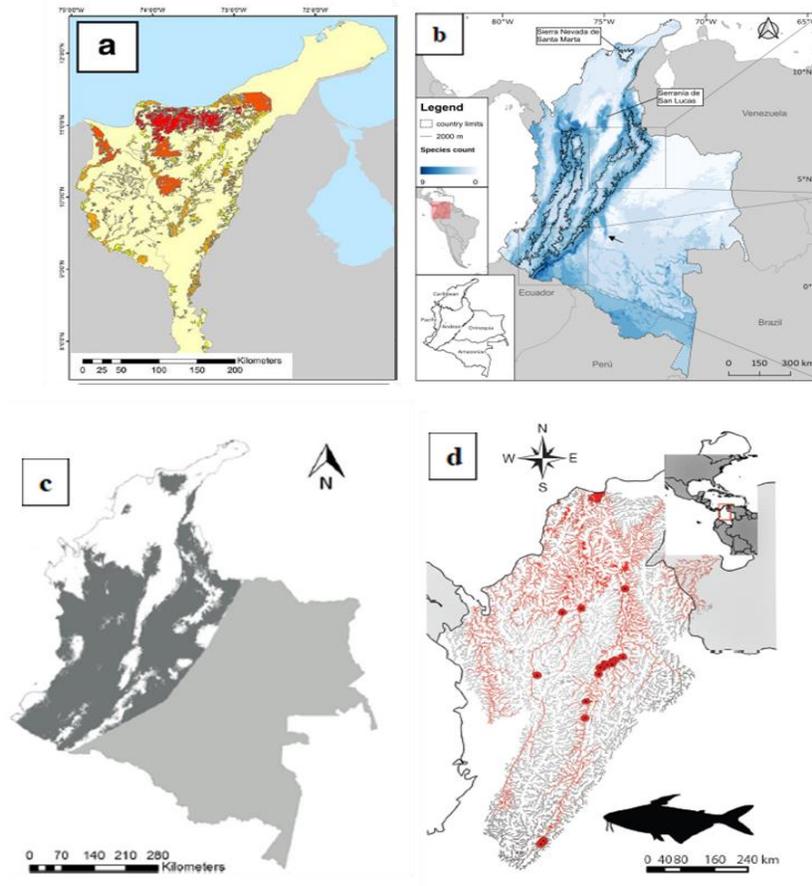


Figura 1. Mapas de distribución de las especies objeto de estudio producto de las modelaciones empleando MAXENT. a) Área potencial de distribución de *Panthera onca* en la Sierra Nevada de Santa Marta (Zárrate et al., 2018) b) Distribución potencial de 16 especies de los géneros *Marmosa* and *Monodelphis* en Colombia (González, et al., 2022) c) Distribución potencial de *Batrachochytrium dendrobatidis* en Colombia (Flechas et al., 2017) d) Distribución potencial del pez invasor *Pangasianodon hypophthalmus* en Colombia (Castellanos et al., 2021).

5.2. Modelación de la especie *Hyalinobatrachium colymbiphyllum*

Se seleccionó la especie de anuro *Hyalinobatrachium colymbiphyllum* para realizar la modelación de su distribución. Esta especie, perteneciente a la familia Centrolenidae, se distribuye en Colombia, Costa Rica, Honduras y Panamá. En Colombia se distribuye en las selvas húmedas tropicales, bosques andinos y subandinos de las vertientes occidental de la Cordillera Oriental y vertiente Oriental de la cordillera Central en el Valle del Magdalena, norte y centro de la región

del Pacífico en los Departamentos de Antioquia, Boyacá, Caldas, Chocó, Córdoba, Quindío, Nariño, Risaralda y Valle del Cauca, entre los 0 y 1800 msnm (Bernal & Lynch, 2008).

A través de la página web GBIF se obtuvieron 126 registros de ocurrencia de la especie para Colombia. Se realizó la limpieza de la base de datos, eliminando los registros duplicados, de esta forma se obtuvo una base de datos de 27 registros (Figura 2).

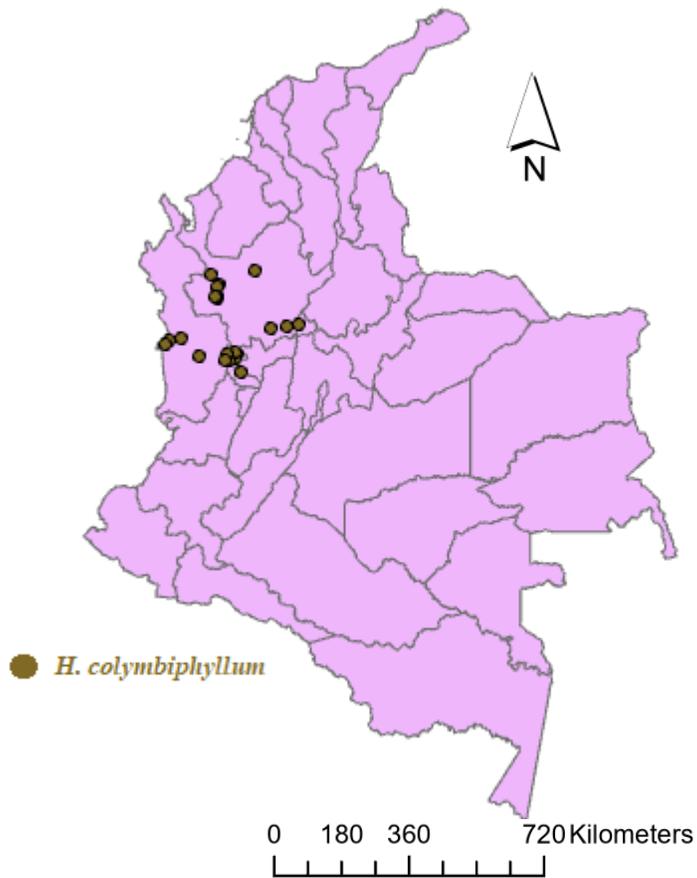


Figura 2. Distribución de la especie *Hyalinobatrachium colymbiophyllum*.

Se emplearon las diecinueve (19) variables bioclimáticas disponibles en la plataforma WorldClim (Fick y Hijmans, 2017) para ejecutar el modelo (Tabla 3).

Tabla 3. Variables bioclimáticas empleadas en el modelo.

Variable	Descripción
bio1	Temperatura media anual
bio2	Rango promedio de temperaturas diarias (Tmax-Tmin)
bio3	Isotermalidad $(\text{bio2}/\text{bio7}) \times 100$
bio4	Estacionalidad de la temperatura (Desviación estándar $\times 100$)
bio5	Temperatura máxima del mes más cálido
bio6	Temperatura mínima del mes más frío
bio7	Rango anual de la temperatura (bio5-bio6)
bio8	Temperatura media del trimestre más húmedo
bio9	Temperatura media del trimestre más seco
bio10	Temperatura media del trimestre más cálido
bio11	Temperatura media del trimestre más frío
bio12	Precipitación anual
bio13	Precipitación del mes más húmedo
bio14	Precipitación del mes más seco
bio15	Estacionalidad de la precipitación
bio16	Precipitación del trimestre más húmedo
bio17	Precipitación del trimestre más seco
bio18	Precipitación del trimestre más cálido
bio19	Precipitación del trimestre más frío

A través de ARCGIS se realizó la delimitación del área de modelado, se cortó de cada una de las capas de las variables bioclimáticas al área de estudio (Colombia); se unificó el tamaño de píxel a 0,0083 metros.

En MAXENT se tomó el 80% de los registros para realizar los modelos y el 20% para probarlos. Se construyeron curvas de respuesta y se realizó la prueba de Jacknife para medir la importancia de cada variable.

Con la definición de la distribución de la especie, con los datos de ocurrencia y variables bioclimáticas en el software Maxent, se obtuvo un mapa de probabilidad de ocurrencia o distribución para *H.colymbiphyllum* en el área de estudio. Como se observa en la Figura 3, el modelo de distribución es congruente con lo reportado en la literatura para la distribución de la especie, vertiente occidental de la Cordillera Oriental y vertiente Oriental de la cordillera Central (Acosta, 2000).

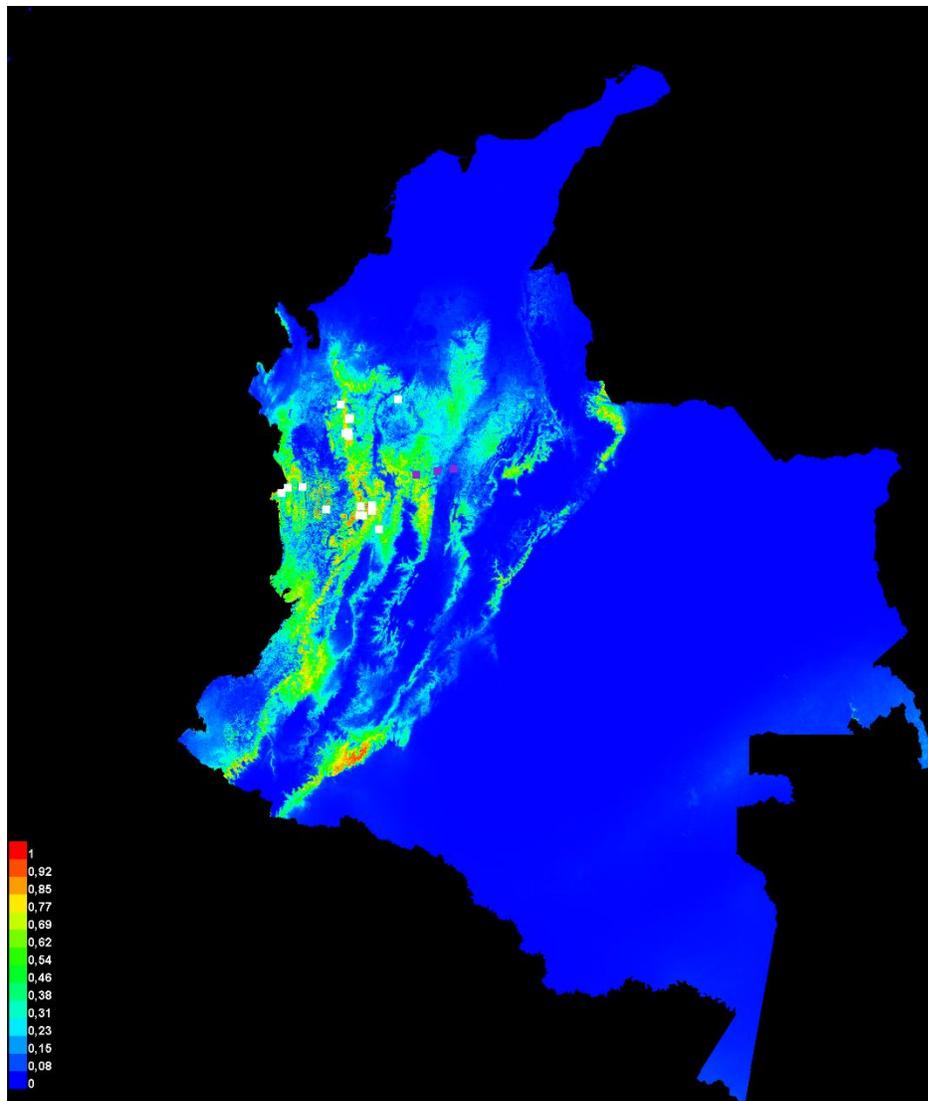


Figura 3. Resultado modelo de distribución de la especie *H.colymbiphyllum*.

De acuerdo con el análisis del límite de la predicción se encontró que los valores por encima de 0,129 son idóneos para la distribución de la especie (Tabla 4), tal como se observa en la Figura 3.

Tabla 4. Límite de la predicción.

Cumulative threshold	Logistic threshold	Description	Fractional predicted area
1.000	0.013	Fixed cumulative value 1	0.384
5.000	0.061	Fixed cumulative value 5	0.214
10.000	0.129	Fixed cumulative value 10	0.149
20.996	0.253	Minimum training presence	0.085
24.458	0.293	10 percentile training presence	0.073
22.209	0.268	Equal training sensitivity and specificity	0.081
20.996	0.253	Maximum training sensitivity plus specificity	0.085
6.889	0.089	Equal test sensitivity and specificity	0.183
6.885	0.088	Maximum test sensitivity plus specificity	0.183
4.185	0.050	Balance training omission, predicted area and threshold value	0.233
11.933	0.152	Equate entropy of thresholded and original distributions	0.133

En cuanto al ajuste del modelo, como lo ilustra la Figura 4, se puede decir que el modelo resultante es confiable para la predicción de distribución de *H.colymbiphyllum* en Colombia, pues tanto los datos de prueba como los de entrenamiento del modelo presentan curvas ajustadas y por encima de la máxima área bajo la curva (AUC) esperada.

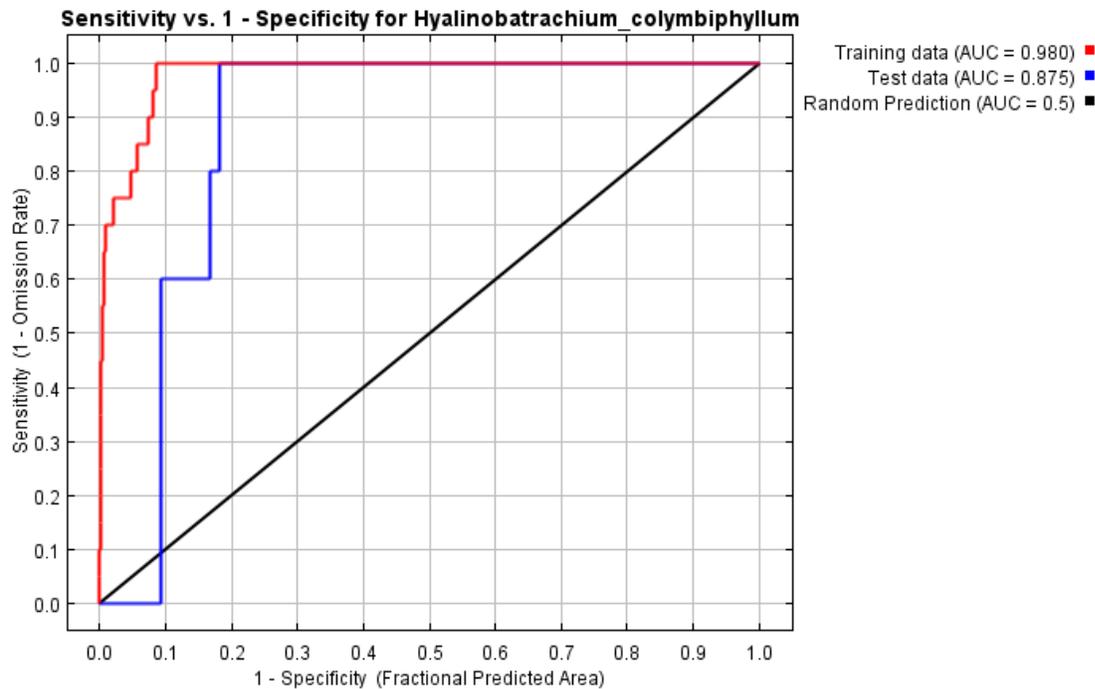


Figura 4. Especificidad del modelo para *H. colymbiphyllum*.

De acuerdo con el modelo, se obtuvo que 11 de las 19 variables presentan relación con la distribución de la especie (Tabla 5).

Tabla 5. Análisis de la contribución de las variables.

Variable	Percent contribution	Permutation importance
bio4	34	63.9
bio14	20	1.6
bio9	12.1	0
bio6	9.5	12.4
bio15	6.9	1.1
bio13	6.3	2.7
bio19	4.4	0
bio12	3.6	13.8
bio18	1.1	2.4
bio3	1.1	0.2
bio17	1	1.9

La variable que más se relacionó con la distribución de la especie fue bio4, que corresponde a la estacionalidad de la temperatura, contribuyendo en un 34% seguida bio14, la precipitación en el mes más seco con un 20%; bio 9, la temperatura media del trimestre más seco con un 12.1%; bio6, la temperatura mínima del mes más frío (9.5%) y bio15, la estacionalidad de la precipitación (6.9%).

Cabe resaltar que la variable que mejor puede explicar, por sí sola, la distribución de la especie en el entrenamiento del modelo es igualmente bio4 (Figura 5a). mientras que bio3, es la variable que por sí sola predice de manera más efectiva la distribución de los datos usados para la prueba del modelo (Figura 5b).

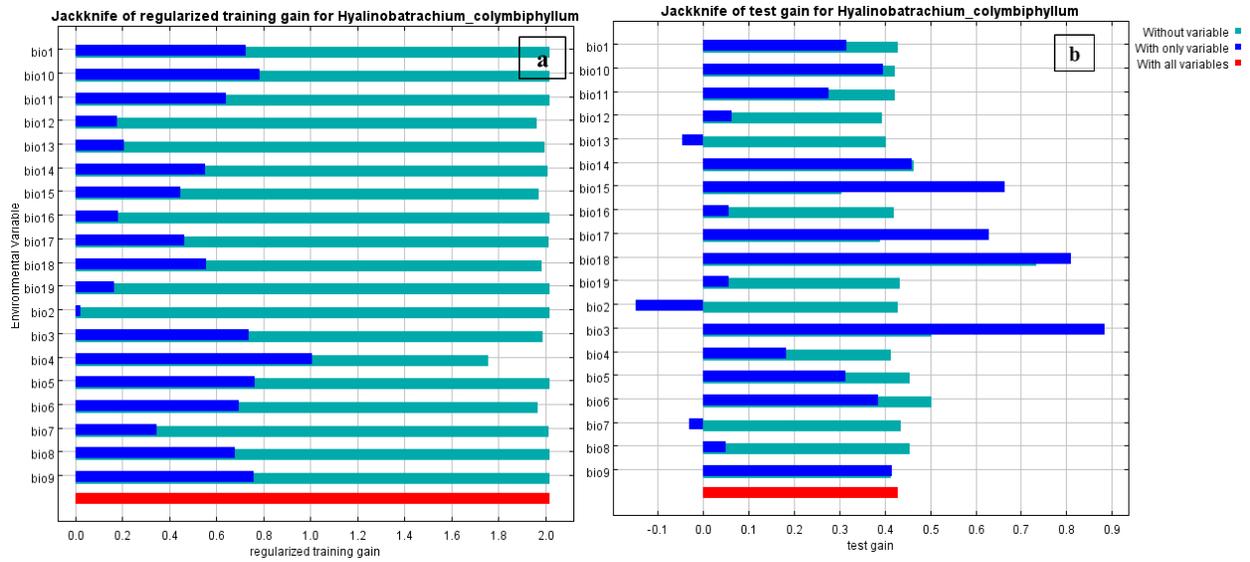


Figura 5. Prueba de Jackknife para a) el entrenamiento del modelo, b) prueba del modelo.

Teniendo en cuenta el análisis del modelo, se puede concluir que los datos de distribución y las variables bioclimáticas empleadas permiten modelar con un nivel de confianza alto la distribución de la especie en el área definida (Figura 4). Las variables asociadas a la temperatura media y precipitación pueden explicar por sí solas la distribución de la especie (Figura 5b).

6. Análisis de resultados

6.1. Análisis de los modelos

A pesar de aplicar el mismo algoritmo para cumplir con los objetivos, se encontraron ligeras diferencias en las metodologías empleadas por los autores.

En la selección de las variables predictoras se encontró que algunos autores realizaron revisiones bibliográficas para llevar al modelo aquellas variables de las que se tiene conocimiento previo que pueden explicar el comportamiento de las especies, tal como Castellanos et al (2021) y González, et al (2022); otros realizaron análisis de correlación (Flechas et al., 2017) para eliminar variables redundantes; otros incluyen todas las variables en los modelos y las seleccionaron de acuerdo con el Análisis de Contribución de Variables realizado por el algoritmo (Zárrate et al., 2018).

Por otro lado, Castellanos et al (2021) afirman que Maxent presenta una gran capacidad para mitigar las contribuciones de variables redundantes, soportado por diversos autores (Elith et al., 2011; Feng, et al., 2019) y emplearon en sus modelos un mayor número de variables que los demás autores, por su parte Zárrate et al (2018) sostienen que los modelos se construyen con demasiadas variables para el tamaño de muestra dado, tienden a sobreajustarse y no predicen bien los datos independientes, basados en Shirley et al (2013).

Para seleccionar los modelos más adecuados Flechas et al. (2017) tuvieron en cuenta los valores de AUC (Área bajo la curva), mientras que los demás autores emplearon los valores de AIC (Criterio de información de Akaike) y González, et al (2022) emplearon una secuencia de diferentes medidas (AIC, AUC, tasa de omisión).

Finalmente, González, et al., 2022 encontraron que para una de las especies estudiadas el MAXENT no presentó un modelo cercano a la distribución real de la especie y sugieren explorar otros tipos de modelaciones en estos casos.

De acuerdo con los autores el empleo del algoritmo de Maxent para estimar la distribución potencial de las especies objeto de estudio fue exitoso.

Flechas et al. (2017) encontraron que el hongo *B. dendrobatidis* podía distribuirse en un mayor rango altitudinal al reportado en la bibliografía. Por su parte Castellanos et al (2021) concluyeron que *P. hypophthalmus* presenta un amplio potencial invasor, que puede ocupar el hábitat y desplazar otras especies de peces endémicas y en peligro nativas en el país.

A partir de la modelación realizada, Zárrate et al (2018) adquirieron los insumos para realizar la estimación de las áreas de posible conflicto del Jaguar en la Sierra Nevada de Santa Marta, así como las áreas de conectividad, susceptibles a ser protegidas.

Finalmente, González, et al (2022) realizaron la distribución potencial de 16 especies de la tribu Marmosini en Colombia, discriminando las áreas con mayor diversidad y que se encuentran dentro de áreas protegidas.

Todos los resultados obtenidos en cada una de las investigaciones contribuyen al aporte del conocimiento para la conservación de la diversidad en Colombia, lo que destaca la importancia de la implementación de estos métodos en las investigaciones.

6.2. Modelado de la distribución de la especie *H. colymbiphyllum*

El modelado de la distribución de *H. colymbiphyllum* fue consistente con los datos reportados para la especie en Colombia, la cual es restringida dentro del país; adicionalmente el modelo arrojó que las variables bioclimáticas que mayor influyen en su distribución son aquellas asociadas a las temperaturas medias y temperaturas y precipitaciones en los meses más secos, variables que de acuerdo a la bibliografía se asocian estrechamente con los anfibios, ya que este grupo es altamente sensible a los cambios en el ambiente (Anguilleta, 2009), debido principalmente a sus particularidades biológicas (ciclo de vida complejo con fase acuática y terrestre, baja capacidad de dispersión, variados mecanismos de respiración y una alta permeabilidad de la piel) (Miaud & Merilä 2001); lo cual es otro indicio del ajuste adecuado de los resultados obtenidos mediante el software Maxent

6.3. Limitaciones de los SDM

Los modelos de distribución de especies presentan limitaciones y sesgos dependiendo de diferentes factores, que deben ser tenidos en cuenta a la hora de analizar la información obtenida, entre los que se encuentran:

-La calidad de los datos de ocurrencia, ya que mucha de la información disponible no es verificada y algunos registros corresponden a especies mal identificadas o mal georreferenciados.

-Sesgos con los datos de ausencia. Gran parte de los SDM no contempla la ausencia de las especies; en el caso de Maxent, este asume como ausencia todos los sitios donde no haya reporte de las especies, lo cual no es completamente correcto, ya que la ausencia de esta información también

puede significar que los individuos no han sido muestreados en dichos lugares, pero que si pueden ocurrir.

-Al realizar las modelaciones no se toman todas las variables que pueden afectar la distribución de una especie, por lo que pueden existir diferencias importantes entre lo que realmente ocurre en la naturaleza y lo que arrojan los modelos.

Adicionalmente, se debe tener en cuenta que los mapas de distribución de los modelos tienen una habilidad limitada para predecir la distribución "verdadera" de las especies, especialmente para eventos recientes como la deforestación, los incendios u otras presiones naturales y relacionadas con el hombre y que afectan su distribución (González, et al, 2022).

Por lo tanto, a pesar de presentar una aproximación cercana a la realidad, no se debe dejar de lado la corroboración en campo de algunos de los hallazgos obtenidos al correr lo modelos. Sin embargo, no dejan de ser herramientas de gran importancia que han aportado al conocimiento de numerosos aspectos en la ecología, y que han servido de base en la toma de decisiones relacionadas con el manejo y conservación de algunas especies, incluso ecosistemas.

7. Conclusiones

Los modelos de distribución de especies son una herramienta de gran importancia y altamente utilizada por investigadores en el mundo, cada día son empleados con mayor frecuencia para dar respuesta a problemáticas en la biología. En Colombia se han utilizado para ampliar el conocimiento de la ecología de especies y para aportar a la toma de decisiones en materia de conservación.

De acuerdo con lo reportado por los autores y encontrado mediante la implementación del algoritmo, Maxent es uno de los más acertados al momento de modelar la posible distribución de una o varias especies, ya que se ha demostrado que sus resultados son cercanos a la realidad. Sin embargo, no se puede dejar de lado que presenta limitaciones, principalmente cuando se tiene pocos registros para algunas especies, por lo tanto, es importante aplicar diferentes modelos teniendo en cuenta la ecología de las especies de estudio.

8. Referencias

- Acosta-Galvis, A. R. (2000). Ranas, salamandras y caecilias (Tetrapoda: Amphibia) de Colombia. *Biota colombiana*, 1(3).
- Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93(7), 1527-1539.
- Angilletta, M. (2009). "Thermal adaptation a theoretical and empirical synthesis". Oxford University Press, USA. 289 pp.
- Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93(7), 1527-1539.
- Baptiste, M. P., García, L. M., Acevedo-Charry, O., Acosta, A., Alarcón, J., Arévalo, E., ... & Pagad, S. (2020). Global Register of Introduced and Invasive Species—Colombia. Version 1.5. *Invasive Species Specialist Group ISSG*. Available: <http://doi.org/10.15468/yznr8v>. Access, 6(07), 2021.
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., ..& Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological modelling*, 222(11), 1810-1819.
- Benito, B., & Peñas, J. (2007). Aplicación de modelos de distribución de especies a la conservación de la biodiversidad en el sureste de la Península Ibérica. *GeoFocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, (7), 100-119.
- Bernal, M. H., & Lynch, J. D. (2008). Review and analysis of altitudinal distribution of the Andean anurans in Colombia. *Zootaxa*, 1826(1), 1-25.
- Castellanos-Mejía, M. C., Herrera, J., Noguera-Urbano, E. A., Parra, E., & Jiménez-Segura, L. F. (2021). Potential distribution in Colombia of the introduced fish *Pangasianodon hypophthalmus* (Siluriformes: Pangasiidae) and implications for endangered native fish. *Revista de Biología Tropical*, 69(2), 573-587.
- Cayuela, L., Golicher, D. J., Newton, A. C., Kolb, M., De Albuquerque, F. S., Arets, E. J. M. M., & Pérez, A. M. (2009). Species distribution modeling in the tropics: problems, potentialities, and the role of biological data for effective species conservation. *Tropical Conservation Science*, 2(3), 319-352.

Contreras-Medina, R. (2006). Los métodos de análisis biogeográfico y su aplicación a la distribución de las gimnospermas en México. *Interciencia*, 31(3), 176-182.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, 40(1), 677-697.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and distributions*, 17(1), 43-57.

Feng, X., Park, D. S., Liang, Y., Pandey, R., & Papeş, M. (2019). Collinearity in ecological niche modeling: Confusions and challenges. *Ecology and evolution*, 9(18), 10365-10376.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, 24(1), 38-49.

Flechas, S. V., Paz, A., Crawford, A. J., Sarmiento, C., Acevedo, A. A., Arboleda, A., & Amézquita, A. (2017). Current and predicted distribution of the pathogenic fungus *Batrachochytrium dendrobatidis* in Colombia, a hotspot of amphibian biodiversity. *Biotropica*, 49(5), 685-694.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, 37(12), 4302-4315.

Gomes, V. H., IJff, S. D., Raes, N., Amaral, I. L., Salomão, R. P., de Souza Coelho, L. & Ramos, J. F. (2018). Species Distribution Modelling: Contrasting presence-only models with plot abundance data. *Scientific reports*, 8 (1), 1-12.).

González, B., Brook, F., & Martin, G. M. (2022). Updated distribution and conservation perspectives of marmosine opossums from Colombia. *Hystrix, the Italian Journal of Mammalogy*, 33(0).

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, 135(2-3), 147-186.

Hutchinson, G. E. (1957). Concluding remarks. population studies: animal ecology and demography. In *Cold Spring Harbor Symposia on Quantative Biology* (Vol. 22, pp. 415-427).

- Krebs, C. J. (1985). Ecology; the experimental analysis of distribution and abundance.
- Kremen, C., Cameron, A., Moilanen, A., Phillips, S. J., Thomas, C. D., Beentje, H., & Zjhra, M. L. (2008). Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science*, 320(5873), 222-226.
- Lomolino, M. V., Brown, J. H., & Sax, D. F. (2010). Island biogeography theory. The theory of island biogeography revisited, 13. Mateo, R. G., Felicísimo, Á. M., & Muñoz, J. (2011). Modelos de distribución de especies: Una revisión sintética. *Revista chilena de historia natural*, 84(2), 217-240.
- Miaud, C., & Merilä, J. (2001). Local adaptation or environmental induction? Causes of population differentiation in alpine amphibians. *Biota*, 2(1), 31-50.
- Miller, J. (2010). Species distribution modeling. *Geography Compass*, 4(6), 490-509.
- Mota Vargas, C., Encarnación Luévano, A., Ortega Andrade, H. M., Prieto Torres, D. A., Peña Peniche, A., & Rojas Soto, O. R. (2020). Una breve introducción a los modelos de nicho ecológico. Universidad Autónoma del Estado de Hidalgo.
- Naoki, K., Gómez, M. I., López, R. P., Meneses, R. I., & Vargas, J. (2006). Comparación de modelos de distribución de especies para predecir la distribución potencial de vida silvestre en Bolivia. *Ecología en Bolivia*, 41(1), 65-78.
- Pearson, R. G. (2010). Species distribution modeling for conservation educators and practitioners. *Lessons Conserv*, 3, 54-89.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190(3-4), 231-259.
- Pliscoff, P., & Fuentes-Castillo, T. (2011). Modelación de la distribución de especies y ecosistemas en el tiempo y en el espacio: una revisión de las nuevas herramientas y enfoques disponibles. *Revista de Geografía Norte Grande*, (48), 61-79.
- Richardson, D. M., & Whittaker, R. J. (2010). Conservation biogeography—foundations, concepts and challenges. *Diversity and Distributions*, 16(3), 313-320.

Shirley, S. M., Yang, Z., Hutchinson, R. A., Alexander, J. D., McGarigal, K., & Betts, M. G. (2013). Species distribution modelling for the people: unclassified landsat TM imagery predicts bird occurrence at fine resolutions. *Diversity and Distributions*, 19(7), 855-866.

Soberon, J., & Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas.

Zárrate-Charry, D. A., Massey, A. L., González-Maya, J. F., & Betts, M. G. (2018). Multi-criteria spatial identification of carnivore conservation areas under data scarcity and conflict: a jaguar case study in Sierra Nevada de Santa Marta, Colombia. *Biodiversity and Conservation*, 27(13), 3373-3392.