



Pontificia Universidad
JAVERIANA
Cali

**VARIACIONES GEOESPACIALES Y TEMPORALES DE CONDICIONES
ATMOSFÉRICAS, PARA EL ESTABLECIMIENTO DE CULTIVOS EN COLOMBIA
UTILIZANDO MODELOS DE APRENDIZAJE ESTADÍSTICO**

Diego Felipe Feria Gómez; 901498
Oscar Eduardo Ramírez Zambrano; 8992256
Yuli Alexandra Deaquiz Oyola; 9014178

*Proyecto Aplicado para optar al título de
Magister en Ciencia de Datos*

Director
David Arango Londoño

FACULTAD DE INGENIERÍA Y CIENCIAS
MAestrÍA EN CIENCIA DE DATOS
SANTIAGO DE CALI, NOVIEMBRE 1 DE 2025

TABLA DE CONTENIDO

| | | |
|-----------|---|-----------|
| 1 | DEFINICIÓN DEL PROBLEMA | 2 |
| 1.1 | PLANTEAMIENTO DEL PROBLEMA | 2 |
| 1.2 | FORMULACIÓN DEL PROBLEMA | 2 |
| 1.2.1 | Sistematización..... | 2 |
| 2 | OBJETIVOS DEL PROYECTO | 4 |
| 2.1 | OBJETIVO GENERAL | 4 |
| 2.2 | OBJETIVOS ESPECÍFICOS | 4 |
| 3 | MARCO TEÓRICO Y ANTECEDENTES | 5 |
| 3.1 | MARCO TEÓRICO | 5 |
| 3.1.4 | Evaluación del desempeño de modelos de clasificación | 8 |
| 3.2 | ANTECEDENTES | 17 |
| 4. | METODOLOGÍA | 20 |
| 5. | CARACTERIZAR LA UBICACIÓN GEOGRÁFICA ACTUAL DE LOS CULTIVOS DE CAFÉ, BANANO Y AGUACATE EN COLOMBIA, ASÍ COMO LOS RANGOS CLIMÁTICOS ACTUALES EN LOS QUE SE DESARROLLAN. | 22 |
| 5.1 | Actividades | 22 |
| 5.2 | Resultados | 23 |
| 6. | DETERMINACIÓN DE LAS VARIACIONES ATMOSFÉRICAS DURANTE UN PERÍODO DE 30 A 50 AÑOS | 28 |
| 6.1 | Actividades | 28 |
| 6.2 | Resultados | 28 |

| | | |
|-----------|--|----|
| 7. | METODOLOGÍAS DE LOS MODELOS DE DISTRIBUCIÓN DE ESPECIES (SDM) PARA LAS ZONAS DE PRODUCCIÓN EN LOS PRÓXIMOS 30 Y 50 AÑOS | 34 |
| 7.1. | Actividades..... | 34 |
| 7.2. | Resultados..... | 36 |
| 8. | ANÁLISIS DE VIABILIDAD DE LAS ZONAS GENERADAS POR LOS MODELOS, CONSIDERANDO POSIBLES RESTRICCIONES NO CLIMÁTICAS | 43 |
| 8.1 | Actividades..... | 43 |
| 8.2 | Resultados..... | 43 |
| 9 | CONCLUSIONES Y TRABAJOS FUTUROS | 46 |
| 9.1 | CONCLUSIONES..... | 46 |
| 9.2 | TRABAJOS FUTUROS..... | 46 |
| 10 | REFERENCIAS BIBLIOGRÁFICAS | 48 |

LISTA DE FIGURAS

| | | |
|--------------------|--|----|
| Figura 3.1: | Modelación de distribución de especies(Fuente: [17])..... | 5 |
| Figura 5.1: | Mapas de aptitud de la tierra para los cultivos en estudio en Colombia, dispuestos por La Unidad de Planificación Agropecuaria (UPRA). | 24 |
| Figura 5.2: | Mapas de aptitud alta para los cultivos en estudio en Colombia. | 25 |
| Figura 5.3: | Variables climáticas de precipitación y temperatura para Colombia. | 26 |
| Figura 5.4: | Variables climáticas de precipitación y temperatura aptas para el cultivo de aguacate (A), banano (B) y café (C). | 27 |
| Figura 6.1: | Modelo EC-Earth3-Veg para las variables climáticas entres los años 2021 – 2040 (A) y 2041 – 2060 (B) (SPP2-4.5)..... | 33 |
| Figura 7.1: | Importancia media de las variables ambientales en modelos para aguacate. | 38 |
| Figura 7.2: | Importancia media de las variables ambientales en modelos para banano..... | 39 |
| Figura 7.3: | Importancia media de las variables ambientales en modelos para café. | 39 |
| Figura 7.4: | Proyecciones de distribución de aguacate bajo escenarios climáticos futuros. | 40 |
| Figura 7.5: | Proyecciones de distribución de banano bajo escenarios climáticos futuros. | 41 |
| Figura 7.6: | Proyecciones de distribución de café bajo escenarios climáticos futuros..... | 41 |
| Figura 8.1: | Mapa de vías de Colombia y sus áreas protegidas. | 44 |
| Figura 8.2: | Áreas limitadas por la viabilidad operativa y legal de los cultivos..... | 45 |

LISTA DE TABLAS

| | |
|--|----|
| Tabla 4.1. Actividades por objetivos específicos. | 20 |
| Tabla 5.1. Categorías de aptitud en evaluación de tierras..... | 23 |
| Tabla 6.1. Modelos de pronóstico climático para la zona de Colombia | 29 |
| Tabla 6.2. Contextos de escenarios climáticos. | 31 |
| Tabla 7.1. Métricas asociadas a los modelos en el cultivo de aguacate | 36 |
| Tabla 7.2. Métricas asociadas a los modelos en el cultivo de banano..... | 36 |
| Tabla 7.3. Métricas asociadas a los modelos en el cultivo de café..... | 37 |
| Tabla 7.4. Áreas por rango de idoneidad por cultivo..... | 42 |

INTRODUCCIÓN

El fenómeno del cambio climático se ha venido intensificando, en el año 2021 el Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC) realizó importantes anuncios sobre el estado actual del clima, posibles escenarios futuros y posibles impactos en la cual este fenómeno ha sido inducido por la actividad humana, indican que la influencia humana ha provocado un calentamiento en el clima a un ritmo sin precedentes en al menos 2000 años [1]. Por otra parte, la Organización Meteorológica Mundial (WMO) en su informe Estado del Clima Global [2] confirma que el año 2023 fue el más caluroso con $1.45 \pm 12^\circ$ C por encima del promedio, las concentraciones de los gases productores de efecto invernadero alcanzaron niveles récord respecto al año 2022, el contenido de calor del océano alcanzó su nivel más alto en el registro de observaciones de 65 años, el nivel medio del mar alcanzó un máximo histórico respecto a los 10 últimos años.

Ante los cambios drásticos del clima, Colombia es un país vulnerable, el aumento en los niveles del mar tanto en el Caribe como en el Pacífico causan erosión, salinización e inundaciones en ambas regiones, el aumento en la temperatura ocasiona incremento de precipitaciones en la región andina y sequías en otras regiones del país, lo que tendrá impactos significativos en la agricultura y en la economía nacional, las tasas de pobreza rural y la seguridad alimentaria, ante este escenario probable, se requiere implementar un Plan Nacional de Adaptación (PNA) el cual asigne responsabilidades a los entes gubernamentales (nacionales y locales) [3].

Ante estos escenarios probables previstos, se tienen posibles impactos y a su vez medidas de adaptación entre las cuales se encuentran; reubicación de actividades de acuerdo con nuevos planes de ordenamiento territorial por posible inundación de tierras agrícolas, debido al aumento del nivel del mar, manejo agronómico sostenible y mejorado para la resiliencia del suelo como respuesta a los procesos de desertificación de la tierra producto de las sequías.

En vista del alto grado de incertidumbre, Colombia debe invertir en modelación climática para identificar posibles escenarios con el fin de generar planes de adaptación ajustados a la geografía colombiana y los diferentes cultivos. Como respuesta a esta necesidad la presente investigación busca mediante el modelado de especies (SDM) identificar las variaciones geoespaciales de las condiciones atmosféricas para los cultivos de Café, Banano y Aguacate siendo estos de importancia relevante en el desarrollo económico del país.

La utilización de modelado de especies pretende mapear áreas que sean adecuadas para la producción de cultivos de Café, Banano y Aguacate, utilizando las proyecciones climáticas de los modelos de las bases de datos en WordClim para predecir cambios geográficos y de esta manera modelar el impacto climático facilitando la planificación y adaptación del sector agrícola. Además, con el modelado se pretende identificar la viabilidad de las áreas que sean adecuadas para la producción de los cultivos mencionados, contra variables no climáticas las cuales ayudan a identificar posibles puntos de mejora para establecer estrategias de implementación en el mediano plazo.

1 DEFINICIÓN DEL PROBLEMA

1.1 PLANTEAMIENTO DEL PROBLEMA

El cambio climático es un fenómeno global que afecta al medio ambiente, la economía y la sociedad, siendo los cambios cada vez más severos en el tiempo [4]. Para el caso de la agricultura que consume alrededor del 70% de toda el agua dulce mundial [5], ha generado efectos negativos significativos a nivel mundial y Colombia no es la excepción, dado a que el desarrollo de cultivos está limitado por la variabilidad extrema de factores climáticos como las precipitaciones intensas, los pocos eventos lluviosos, la inadecuada distribución espacial y temporal de las precipitaciones [6].

Además, a esto se le suma el aumento de temperaturas de acuerdo a la zona [7], donde se prevé que empeoren a medida que aumentan las temperaturas globales promedio y se espera que este aumento esté entre 2 y 4 °C en el próximo siglo [8]. Así mismo, en cultivos de exportación como el café se proyecta que las áreas aptas podrían reducirse en hasta un 30% para el año 2050, como resultado del aumento de las temperaturas y la alteración de las condiciones climáticas óptimas para su crecimiento [9] [10][11].

También, la frecuencia creciente de fenómenos climáticos extremos, como sequías prolongadas e inundaciones, aumento de temperatura, está afectando de manera directa la productividad agrícola, especialmente en cultivos clave como el banano, el café y el aguacate [12]. Así mismo, la disminución general de la huella hídrica con el aumento de la temperatura se puede explicar por el impacto de la temperatura en la pérdida de agua por evapotranspiración, lo que resulta en una mayor demanda de riego y menores rendimientos finales [13].

Estos impactos no solo ponen en riesgo la producción agrícola, sino también la estabilidad económica y el bienestar de las comunidades rurales que dependen de esta actividad para su subsistencia [14]. Por lo tanto, es fundamental desarrollar estrategias de adaptación basadas en ciencia de datos, que incluyan la implementación de tecnologías agrícolas más resilientes y una gestión sostenible de los recursos del suelo y el agua.

1.2 FORMULACIÓN DEL PROBLEMA

¿Cómo mediante la aplicación de metodologías de modelado espacial, se pueden identificar la variación de las zonas recomendadas para la producción de los cultivos, por efecto del cambio climático?

1.2.1 Sistematización

¿Cuál es la distribución geoespacial actual de los cultivos de importancia económica en Colombia, como el café, el banano y el aguacate; y bajo qué condiciones climáticas se desarrollan?

¿Cuál es el comportamiento pronosticado de las variables atmosféricas por efecto del cambio climático en Colombia?

¿Cuáles metodologías de modelado de distribución de especies (SDM), se pueden utilizar para pronosticar las futuras dinámicas geoespaciales de las condiciones recomendadas para la producción de café, banano y aguacate, considerando el cambio en las condiciones climáticas?

¿Cuál modelo tiene mejor desempeño al pronosticar las zonas con mejores condiciones climáticas para los cultivos?

¿Existen restricciones no climáticas en las zonas de producción pronosticadas, que puedan afectar el establecimiento de los cultivos?

2 OBJETIVOS DEL PROYECTO

2.1 OBJETIVO GENERAL

Evaluar las posibles dinámicas geoespaciales del cambio climático en el tiempo, en relación con las condiciones recomendadas para la producción de cultivos de importancia económica en Colombia (café, banano y aguacate), mediante el uso de metodologías de modelado de distribución de especies (SDM, Species Distribution Modeling).

2.2 OBJETIVOS ESPECÍFICOS

- Caracterizar la ubicación geográfica actual de los cultivos de café, banano y aguacate, y los rangos actuales climáticos donde se desarrollan en Colombia.
- Determinar las variaciones atmosféricas durante un periodo de 30 (2021 a 2040) y 50 (2041 a 2060) años a partir de modelos de pronóstico existentes (Wordclim).
- Aplicar las metodologías de modelos de distribución de especies (SDM) para las zonas de producción recomendadas y realizar una proyección para los cultivos en los próximos 30 y 50 años, estableciendo el modelo más adecuado a partir de métricas de desempeño y validación cruzada.
- Analizar la viabilidad de las zonas generadas por los modelos considerando posibles restricciones no climáticas.

3 MARCO TEÓRICO Y ANTECEDENTES

3.1 MARCO TEÓRICO

3.1.1 Modelado de distribución de especies (Species Distribution Modeling SDM)

Es una metodología construida sobre conceptos ecológicos y biogeográficos básicos sobre la relación entre las distribuciones de especies (u otras variables de respuesta biótica que describen aspectos de la biodiversidad) y el ambiente físico (abiótico). Son modelos cuantitativos y empíricos de las relaciones entre especies y ambiente que se desarrollan típicamente utilizando datos de ubicación de especies (ocurrencia, abundancia) y aquellas variables ambientales que se cree que influyen en las distribuciones de especies. Siempre que se disponga de datos adecuados, los SDM se pueden aplicar a cualquier Especie [15].

Los SDM se utilizan para la evaluación del impacto ambiental y la planificación del uso de la tierra para determinar ubicaciones adecuadas para la restauración del hábitat y la reintroducción de especies y para predecir el riesgo de establecimiento de especies en nuevas áreas, patógenos y vectores de enfermedades [15]. Uno de los usos de más rápido crecimiento de los SDM para el análisis de riesgos es predecir el impacto potencial de los patrones proyectados de cambio climático antropogénico en las distribuciones de especies [16].

Los modelos de distribución de especies predicen la ocurrencia o abundancia de organismos a partir de información ambiental [17] y, por lo tanto, están estrechamente vinculados a la heterogeneidad del hábitat. Estos son generalmente modelos empíricos que se construyen a partir de relaciones hábitat-organismo observadas (Figura 1.1). Se han utilizado varios algoritmos, incluidos modelos lineales generalizados (GLM), modelos aditivos generalizados (GAM), modelos de bosque aleatorio, modelos de máxima entropía y modelos de redes neuronales artificiales (ANN) para predecir las distribuciones de organismos marinos [15].

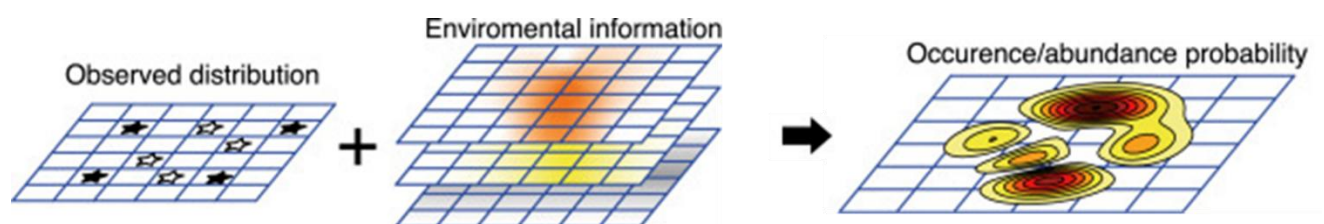


Figura 3.1: Modelación de distribución de especies (Fuente: [17]).

Estos modelos pueden estar impulsados por las distribuciones reales de los organismos y las variables ambientales, como las temperaturas de la superficie del mar derivadas de observaciones satelitales. El resultado del modelo (las distribuciones de los organismos) se obtiene en algunos casos mediante la regresión a una función dada (GLM y GAM), o mediante procesos de aprendizaje iterativo en otros casos [18].

La mejor práctica en los modelos de distribución de especies para climas futuros es modelar solo las

especies cuyo rango completo se encuentre dentro de la región de estudio. Esto se debe a que una especie que tiene un rango fuera de la región modelada tiene tolerancias climáticas que no se pueden capturar en las estadísticas de los modelos de distribución de especies [19].

3.1.2 Patrones puntuales

El análisis de patrones puntuales estudia la distribución espacial de los puntos, que contiene dos tipos principales de componentes: los puntos que representan los objetos de estudio y la zona geográfica en la que se encuentran [20]. Estos patrones se caracterizan por dos propiedades específicas: la de primer orden (la intensidad, o densidad) y la de segundo orden (suma de interacciones de atracción o repulsión entre los puntos).

El análisis se realiza a través de procesos estocásticos, que genera patrones de puntos aleatorios que comparten la misma estructura espacial (la ley del proceso) [21], los más conocidos incluyen los procesos de Markov de los cuales se derivan otros procesos como lo son los procesos de Poisson o los procesos de difusión, cada uno de ellos con sus propiedades específicas e idóneos para modelizar diferentes tipos de fenómenos [22]. Según [23], el objetivo general es comprobar si los puntos exhiben algún tipo de patrón como los siguientes:

- (1) Patrón aleatorio: los puntos se distribuyen aleatoriamente en el espacio.
- (2) Patrón regular: existe una distancia media entre los puntos que tiende a ser constante.
- (3) Patrón agregado: existen aglomeraciones de puntos en el espacio.

- **Métodos basados en el análisis de primer orden**

Intensidad En el análisis exploratorio el primer paso es revisar la intensidad del patrón, esta medida describe número esperado de eventos por unidad de área. Donde la densidad empírica de los puntos está dada por la siguiente ecuación:

$$\lambda = \frac{n(x)}{|W|} \quad (1)$$

Donde $n(x)$ corresponde al número de puntos en la región de análisis y $|W|$ su respectiva **área** [24].

- **Estimación de la función de intensidad por suavizamiento de Kernel**

La estimación no paramétrica de la función de intensidad se realiza a través del estimador de Kernel propuesto por [23] y dado por la siguiente expresión:

$$\lambda(u) = \sum_{i=1}^n \frac{1}{e(x_i)} k(u - x_i) \quad (2)$$

Donde u localización dentro de la ventana de observación; λ representa la Ventana de observación; $k(u)$ representa la función Kernel que debe ser una densidad de probabilidad $k(u) \geq 0$ para todas las localizaciones u [24].

- **Método de conteo por cuadrantes para Aleatoriedad espacial completa (CSR)**

Este método consiste en dividir la ventana de observación W en subregiones B_1, \dots, B_m llamadas cuadrantes que, en principio, se consideran de igual área. A partir de esto, se realiza el conteo de los puntos que caen en cada subregión $n_j = n(X \cap B_j)$ para $j = 1, \dots, m$. Al tratarse de intensidad homogénea, los conteos deberían ser iguales o cercanos en cada subregión, por el contrario, los cambios fuertes en los conteos pueden sugerir tendencias espaciales. Con el objetivo de contrastar la hipótesis de homogeneidad.

Se aplica la prueba χ^2 (Chi-cuadrado), dado el número total de puntos $n = \sum_j n_j$, el área total de la ventana de observación $a = \sum_j a_j$, la intensidad estimada $\lambda = n/a$ y el número esperado de conteos en el cuadrante B_j es $e_j = \lambda a_j = n a_j / a$ [24].

- **Métodos basados en el análisis de segundo orden**

K de cualquier punto típico del patrón. Para una determinada escala (es decir, para un determinado radio r) $N(r)$ estima la probabilidad de encontrar un punto en las inmediaciones de otro el cual se expresa de la siguiente manera [21]:

$$K(r) = \frac{N(r)}{\lambda} \quad (3)$$

Siendo λ la intensidad del patrón (en la mayoría de los casos equivalente a la densidad de puntos).

- **La función de correlación de par**

Es una función que caracteriza de forma no acumulativa el patrón, es decir que tenga en cuenta tan sólo los pares de puntos que se encuentran separados por una distancia exactamente igual o similar a la distancia r . La función de correlación de par $g(r)$ es la herramienta apropiada en este caso, la cual tiene la siguiente expresión:

$$g(r) = \frac{K'(r)}{2\pi r} \quad (4)$$

Siendo $K'(r)$ la derivada de $K(r)$. En un patrón de Poisson, $g(r) = 1$ para todas las distancias r ; en cualquier patrón empírico, valores de $g(r) > 1$ indican que las distancias alrededor de r son relativamente más frecuentes de lo que serían en un patrón aleatorio (por ejemplo, si se tratase de valores pequeños de r indicaría agrupamiento) mientras que valores de $g(r) < 1$ indican que las correspondientes distancias son menos frecuentes de lo esperado, lo que podría indicar algún tipo de inhibición [25].

- **Métodos basados en la distancia al vecino más cercano y otras relacionadas**

Busca determinar cuál sería la distribución acumulada de distancias que presentaría un patrón de Poisson homogéneo, se han desarrollado varios test que se basan en comparar dicha distribución con la del patrón objeto de estudio. Una aproximación de la función de distribución acumulada de distancias al vecino más cercano es [21] [23].

$$G(y) = 1 - e^{-\lambda \pi y^2} \quad (5)$$

Siendo y la distancia desde un punto cualquiera del patrón hasta el punto más cercano y λ la intensidad (número esperado de puntos por unidad de área).

3.1.3 Método de nichos ecológicos

El nicho se puede considerar como un hipervolumen multidimensional, en donde cada eje que lo conforma corresponde a un estado del ambiente que permite a una especie existir indefinidamente. Esta multidimensionalidad está dada por dos tipos de variables, las que son independientes y están ligadas a la presencia de las especies, que actúan a diferentes escalas y que pueden estar ligadas de manera dinámica o no a la presencia de la especie [26].

Las variables independientes (factores abióticos) como la altitud, la precipitación y la temperatura han sido llamadas escenopoéticas y estas también pueden incluir factores bióticos. Por otra parte, las variables ligadas a la presencia de las especies son llamadas bionómicas y están relacionadas con las interacciones bióticas que pueden existir, estos dos tipos de variables forman parte de la multidimensionalidad del nicho [26].

Para poder ilustrar de manera sencilla los aspectos relacionados con la distribución de las especies y su nicho ecológico se ha propuesto el diagrama BAM (elementos bióticos, abióticos y movilidad), el primer componente B, representa la región en donde ocurren para que la especie pueda subsistir, el segundo componente A hace referencia a las condiciones bióticas y el componente M incluye el área que es accesible para la especie, dada la capacidad de movilidad así como los factores antropogénicos que pueden intervenir. Por esta razón, la mayoría de los algoritmos de modelado de nicho relacionan los datos biológicos con características geográficas representadas por el factor abiótico [27].

3.1.4 Evaluación del desempeño de modelos de clasificación.

La evaluación del desempeño es una etapa importante en el desarrollo y validación de modelos de clasificación, ya que permite cuantificar la capacidad predictiva de un modelo y comparar alternativas metodológicas. En contextos como la epidemiología, la ecología, la medicina y las ciencias agropecuarias, donde las decisiones se basan en la correcta identificación de eventos (presencia/ausencia, enfermo/sano, apto/no apto), se emplean métricas derivadas de la matriz de confusión que resumen diferentes aspectos del rendimiento del modelo [28].

- CURVA ROC (Receiver Operating Characteristic)

Es un método estadístico que permite evaluar el desempeño de métodos clasificatorios, además de determinar la exactitud diagnóstica de test que utilizan escalas continuas, siendo utilizadas con 3 propósitos específicos: determinar el punto de corte en el que se alcanza la sensibilidad y especificidad más altas que garantizan el mejor desempeño de la prueba [29], evaluar la capacidad discriminativa del test diagnóstico y comparar la capacidad discriminativa de 2 o más test diagnósticos que expresan sus resultados como escalas continuas. Así mismo, estos métodos tratan de identificar a que tipo de eventos pertenecen las observaciones, el ROC resume esta relación en un único valor que varía entre 0.5

(desempeño aleatorio) y 1 (clasificación perfecta) [30] [31]. No obstante, la curva ROC, puede ocultar deficiencias en rangos específicos de probabilidad que sean críticos para la toma de decisiones. Además, en conjuntos de datos altamente desbalanceados, el ROC puede ofrecer una visión optimista del desempeño real del modelo en la clase minoritaria [28].

- **Sensibilidad y especificidad**

La sensibilidad también denominada tasa de verdaderos positivos o *recall* representa la capacidad del modelo de clasificación de detectar correctamente los valores verdaderos, mientras que la especificidad corresponde a la probabilidad de casos negativos correctamente clasificados [32]. Ambas métricas se definen a partir de la matriz de confusión y reflejan el comportamiento del modelo frente a cada clase.

Una de las principales ventajas de la sensibilidad y la especificidad es su interpretación directa y su utilidad en contextos donde el costo de los errores es asimétrico, por ejemplo, cuando es más crítico minimizar los falsos negativos que los falsos positivos. Sin embargo, estas métricas dependen del umbral de decisión utilizado para convertir probabilidades en clasificaciones binarias, lo que puede dificultar la comparación entre modelos si no se fija un criterio común [33].

- **True Skill Statistic (TSS)**

El TSS, también conocido como índice de Youden, se define como la suma de la sensibilidad y la especificidad menos uno. Este estadístico toma valores entre -1 y 1, donde 0 indica un desempeño no mejor que el azar y 1 una clasificación perfecta. Es especialmente útil en problemas con clases desbalanceadas, como los modelos de distribución de especies o la detección de enfermedades raras. Además, combina en una sola métrica la información de sensibilidad y especificidad. Sin embargo, el TSS depende de la selección de un umbral de clasificación y no penaliza explícitamente los errores de comisión y omisión de manera diferenciada [34].

3.1.5 Supuestos estadísticos de los modelos GLM y GAM en Modelos de Distribución de Especies (SDM)

- **Modelo Lineal Generalizado (GLM) [36]**

El Modelo Lineal Generalizado (Generalized Linear Model, GLM) constituye una extensión del modelo lineal clásico que permite modelar variables respuesta cuya distribución no necesariamente es normal, lo cual resulta especialmente apropiado para datos ecológicos de presencia-ausencia o probabilidades de ocurrencia.

En SDM binarios, el GLM más utilizado es la regresión logística, donde la probabilidad de presencia de una especie se modela en función de variables ambientales.

Matemáticamente:

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}$$

donde $g()$ es la función de enlace, $\mu_i = E(Y_i)$ y η_i el predictor lineal.

Componentes estructurales del GLM [37]

Un GLM está compuesto por tres elementos:

- **Componente aleatoria:** variable respuesta con distribución de la familia exponencial (Bernoulli para presencia/ausencia).
- **Componente sistemática:** combinación lineal de predictores ambientales.
- **Función de enlace:** relaciona la media de la respuesta con el predictor lineal.

- **Supuestos estadísticos del GLM en SDM**

1. Independencia de observaciones

Las observaciones Y_i deben ser independientes entre sí.

En SDM esto implica ausencia de autocorrelación espacial fuerte (un supuesto frecuentemente violado en ecología) [36].

2. Distribución de la variable respuesta

La variable respuesta pertenece a la familia exponencial (en SDM presencia-ausencia: Bernoulli/binomial) [38].

3. Relación lineal en el predictor

El modelo asume que la relación entre variables ambientales y la respuesta es lineal en el espacio del enlace:

$$\text{logit}(\pi_i) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

Esto significa que el efecto ecológico es monótonico (no óptimo ambiental unimodal).

1. Correcta especificación de la función de enlace

Debe elegirse un link adecuado al tipo de variable; para presencia-ausencia se utiliza:

$$g(\mu) = \log\left(\frac{\pi}{1-\pi}\right)$$

2. Homogeneidad de la varianza (condicionada)

La varianza depende de la media según la distribución asumida:

$$\text{Var}(Y) = \pi(1-\pi)$$

Por tanto, no se asume varianza constante como en regresión lineal clásica.

3. Ausencia de multicolinealidad severa

Los predictores no deben ser combinaciones lineales entre sí (condición de identificabilidad del modelo).

- **Interpretación ecológica del GLM**

El GLM modela respuestas ecológicas monótonicas (gradientes ambientales simples).

Por tanto, describe adecuadamente:

- límites climáticos
- gradientes de tolerancia
- respuestas lineales o logísticas

Pero falla ante nichos ecológicos con óptimos complejos (respuestas unimodales o multimodales).

- **Modelo Aditivo Generalizado (GAM)**

El Modelo Aditivo Generalizado (Generalized Additive Model, GAM) es una extensión semiparamétrica del GLM que reemplaza la relación lineal por funciones suavizadas no paramétricas.

$$g(\mu) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_p(x_p)$$

donde f_k son funciones spline suavizadas.

Los GAM fueron introducidos para capturar relaciones ecológicas no lineales comunes en SDM [39].

Supuestos estadísticos del GAM

1. Aditividad funcional

La respuesta es la suma de funciones suaves independientes:

$$\eta = \sum f_k(x_k)$$

Es el supuesto fundamental del modelo [39].

2. Suavidad de las funciones ecológicas

Las respuestas ambientales deben ser continuas y suaves, no discontinuas (ej. tolerancias ecológicas). Esto representa la hipótesis biológica de nicho ecológico continuo.

3. Independencia de observaciones

Al igual que el GLM, requiere independencia (sin autocorrelación espacial fuerte).

4. Distribución de la respuesta

Mantiene la familia exponencial (Bernoulli para SDM).

5. Baja colinealidad entre predictores

Los predictores no deben estar altamente correlacionados, pues afecta la estimación de splines.

6. Selección del grado de suavizado

El modelo asume que existe un nivel óptimo de suavizado (penalización), el cual controla el sobreajuste

- **Interpretación ecológica del GAM**

El GAM representa mejor la teoría del nicho ecológico porque permite:

- óptimos ambientales
- tolerancias
- umbrales
- respuestas unimodales

Por ello suele superar al GLM en SDM climáticos.

Supuestos y fundamentos metodológicos de MAXENT/MAXNET en SDM

Enfoque general: presencia-fondo y máxima entropía

MaxEnt (Maximum Entropy) es un método ampliamente usado para SDM cuando se dispone de presencias y se contrasta contra un conjunto de puntos de fondo (background) que representan el ambiente disponible. El objetivo se formula como un problema de estimación de una distribución sobre el espacio geográfico/ambiental que sea lo más “uniforme” posible (máxima entropía) sujeta a restricciones

impuestas por los valores promedio de variables ambientales en los sitios de presencia (y sus transformaciones). Esta formulación conduce a un modelo de la familia exponencial con fuerte relación con enfoques estadísticos clásicos

Un resultado teórico importante es que MaxEnt es estrechamente equivalente (bajo condiciones) a un modelo de Proceso de Poisson inhomogéneo (IPP) o a una regresión de Poisson en el marco presencia-fondo; esto aclara que lo estimable con presencia-fondo es principalmente intensidad relativa / idoneidad relativa, y no necesariamente “probabilidad absoluta de presencia” sin supuestos adicionales [40].

Supuestos principales de MaxEnt en SDM (presencia-fondo)

1. Representatividad del muestreo de presencias

Se asume que las presencias disponibles constituyen una muestra informativa del nicho realizado en el área accesible (M) y que, dadas las covariables, el proceso de muestreo no introduce distorsiones severas. En la práctica, sesgos (carreteras, cercanía a centros poblados, accesibilidad) pueden inducir sobreajuste espacial si no se controlan. La literatura enfatiza explícitamente que las decisiones de insumos y configuración afectan la interpretación y el rendimiento [41].

2. Definición adecuada del “ambiente disponible” (background)

El conjunto de fondo debe aproximar el ambiente disponible para la especie en el área de calibración. Un fondo demasiado amplio o restringido cambia el contraste presencia-fondo y puede alterar la inferencia. Esto es central en guías prácticas de MaxEnt, donde se discute el papel del background y de la región de calibración [41].

3. Supuesto de relación funcional capturable por “features”

MaxEnt no asume linealidad simple: el método construye transformaciones (feature classes) de las covariables (por ejemplo: lineales, cuadráticas, “hinge”, producto, umbral, etc.) y ajusta una combinación ponderada. La hipótesis práctica es que la respuesta especie-ambiente puede aproximarse razonablemente mediante esas bases [42].

4. Estabilidad de la relación especie-ambiente (transferencia)

Para proyecciones (p. ej., escenarios climáticos futuros), se asume que la relación estimada (en el espacio de covariables) es transportable. La extrapolación a combinaciones climáticas nuevas es una fuente conocida de incertidumbre, especialmente cuando el modelo es complejo (muchas features, baja penalización). Esta preocupación está estrechamente ligada al control de complejidad vía regularización [42].

- Regularización y control de complejidad en MaxEnt

MaxEnt incorpora regularización para evitar que el modelo se ajuste excesivamente a las presencias (sobreajuste), penalizando el uso de demasiadas features o pesos muy grandes. La práctica estándar es ajustar dos elementos: (i) clases de features y (ii) multiplicador de regularización (RM), donde valores mayores de RM implican modelos más suaves (menos complejos).

En términos metodológicos, la regularización opera como un compromiso sesgo-varianza: mayor complejidad mejora el ajuste aparente en calibración, pero puede reducir generalización espacial/temporal. Este punto es explícito en la evaluación y extensiones de MaxEnt [42].

- **MAXNET: implementación en R y vínculo con GLM regularizados**

Maxnet es una implementación en R del enfoque MaxEnt que ajusta el modelo usando glmnet, es decir, un marco de modelos lineales generalizados regularizados (L1/L2/elastic-net) aplicado a las features equivalentes de MaxEnt. Por defecto busca replicar los conjuntos de features y regularización del software Java de MaxEnt.

Desde el punto de vista computacional y estadístico:

- Maxnet hereda la lógica de penalización de glmnet (regularización convexa para controlar complejidad) [37]
- se integra naturalmente a flujos reproducibles en R, manteniendo coherencia con el marco de MaxEnt abierto y documentado [37].

Supuestos, funcionamiento y riesgos de sobreajuste en Random Forest para SDM

- **Fundamento del método Random Forest para SDM**

Random Forest (RF) es un método de ensamble que combina muchos árboles de decisión construidos sobre muestras Bootstrap, incorporando aleatoriedad en la selección de predictores en cada partición del árbol. Esta combinación reduce varianza y suele mejorar desempeño predictivo respecto a un solo árbol. La formulación clásica y sus propiedades de error generalizado están establecidas por Breiman [43]. En ecología, RF se ha posicionado como un clasificador potente para problemas de clasificación y regresión, incluyendo distribución de especies, destacándose por su capacidad para modelar relaciones no lineales e interacciones [44].

Supuestos prácticos de RF en SDM (más que supuestos “paramétricos”)

A diferencia de GLM/GAM, RF no impone supuestos paramétricos fuertes (normalidad, homocedasticidad, linealidad). No obstante, en SDM sí existen supuestos prácticos decisivos:

1. Independencia efectiva de las observaciones para evaluar desempeño

Aunque RF puede entrenarse con datos espacialmente auto correlacionados, si la validación no controla esa estructura, la precisión puede resultar artificialmente alta. Estudios sobre validación espacial muestran que el desempeño suele caer cuando se usa validación espacial en lugar de Cross-Validation aleatoria, evidenciando el sesgo optimista de la validación estándar en datos espaciales [45].

2. Representatividad del muestreo y balance presencia/ausencia

En presencia-ausencia (o pseudoausencia), el modelo asume que las ausencias/pseudoausencias representan adecuadamente el contraste ambiental con las presencias. Desbalances severos tienden a sesgar la clasificación hacia la clase mayoritaria, afectando sensibilidad/especificidad y umbrales.

3. Estabilidad de relaciones para transferencia y proyección

RF aprende reglas empíricas del dominio observado. Fuera del rango de entrenamiento (extrapolación climática o geográfica), puede degradarse de manera difícil de diagnosticar. En SDM, este problema se agrava cuando la complejidad del bosque captura patrones locales o sesgos de muestreo.

Riesgos de sobreajuste en Random Forest (en SDM geoespacial):

1. Sobreajuste aparente por autocorrelación espacial (leakage espacial)

El riesgo más frecuente no es “demasiados árboles” (RF suele estabilizar con más árboles), sino que el modelo explota similitudes espaciales (vecindad) y el esquema de validación aleatorio permite que puntos cercanos queden en entrenamiento y prueba, inflando métricas. La evidencia empírica en contextos espaciales muestra que la validación espacial suele reportar errores mayores (más realistas) [45].

RF puede ajustar fronteras de decisión muy irregulares si los árboles son profundos, esto puede capturar ruido o sesgos locales del muestreo; en SDM esto se manifiesta como “parches” de idoneidad altamente fragmentados.

2. Alta dimensionalidad, predictores correlacionados y “importancia” sesgada

RF tolera colinealidad mejor que métodos lineales, pero la presencia de predictores correlacionados puede:

- redistribuir importancia entre variables correlacionadas,
- favorecer splits sobre variables con más puntos de corte,
- Dificultar interpretación ecológica, en ecología se discuten ventajas y advertencias de RF para interpretación y selección de variables [44].

3. Sensibilidad a la definición de pseudoausencias

Cuando no hay ausencias verdaderas, la generación de pseudoausencias puede inducir sobreajuste si se introducen contrastes demasiado fáciles (p. ej., pseudoausencias muy lejos del dominio ambiental de presencia), lo que eleva AUC/TSS sin reflejar capacidad real de discriminación.

- Estrategias metodológicas recomendadas

Para reducir sobreajuste y asegurar validez espacial:

1. Validación cruzada espacial (bloques/particiones espaciales) para estimar desempeño realista [45].
2. Ajuste de hiperparámetros para controlar complejidad; RF es robusto, pero no “inmune” a configuraciones que inducen fronteras demasiado locales [44].
3. Manejo de desbalance (submuestreo, ponderación de clases) y evaluación con métricas sensibles a la clase minoritaria (sensibilidad/TSS).
4. Control de sesgo de muestreo (p. ej., background/pseudoausencias con sesgo similar a presencias, o filtros espaciales) para evitar que el modelo aprenda accesibilidad en lugar de nicho

- Validación cruzada espacial

En validación tradicional, los datos de entrenamiento y prueba pueden estar separados solo por metros, por lo que la autocorrelación permanece presente.

La validación espacial divide el área de estudio en bloques geográficos independientes:

$$\text{Entrenamiento} \cap \text{Prueba} = \emptyset \text{ espacialmente}$$

De esta forma el modelo es evaluado en regiones donde nunca observó datos durante el entrenamiento. Este procedimiento simula mejor la transferencia hacia condiciones climáticas futuras.

la validación aleatoria puede inflar el AUC entre 0.1 y 0.3 unidades en modelos ecológicos, generando conclusiones erróneas sobre la calidad predictiva [46].

Por ello, la validación cruzada espacial es actualmente considerada obligatoria en SDM orientados a cambio climático.

- **Generación de pseudoausencias**

En la mayoría de los casos, los datos biológicos disponibles corresponden únicamente a registros de presencia. Sin embargo, los algoritmos de clasificación requieren información de ausencia. Para resolver este problema se generan pseudoausencias.

Las pseudoausencias representan ubicaciones donde la especie no ha sido registrada, pero cuya ausencia no es necesariamente confirmada. Por tanto, no equivalen a ausencias verdaderas sino a muestras del fondo ambiental.

La selección de pseudoausencias influye directamente en la estimación del nicho ecológico:

- Pseudoausencias cercanas → modelo restringido
- Pseudoausencias lejanas → modelo generalista

El número y distribución espacial de pseudoausencias modifica significativamente el desempeño del modelo y la extensión proyectada del rango futuro [47].

En SDM climáticos se recomienda seleccionar pseudoausencias dentro del área accesible de la especie (área M), evitando regiones biogeográficamente inalcanzables. Este enfoque reduce sesgos de muestreo y mejora la interpretación ecológica del nicho.

- **Problemas de extrapolación en escenarios climáticos futuros**

Cuando un modelo se proyecta hacia condiciones climáticas futuras, frecuentemente debe predecir en combinaciones ambientales que no existen en el presente:

$$X_{futuro} \notin \text{dominio}(X_{entrenamiento})$$

Este proceso se denomina extrapolación ambiental.

El principal riesgo es que el modelo no está estimando el nicho real, sino prolongando relaciones estadísticas fuera de su rango de validez. Como resultado, pueden generarse:

- áreas de idoneidad inexistentes
- desplazamientos irreales de rango
- expansión hacia climas incompatibles con la fisiología de la especie

el error por extrapolación es la principal fuente de incertidumbre en proyecciones biogeográficas futuras. [20]

La confiabilidad de un SDM bajo cambio climático depende de la combinación de tres componentes:

1. Evaluación estadística (AUC, TSS, sensibilidad y especificidad)
2. Control espacial (bloqueo espacial y reducción de autocorrelación)
3. Consistencia ecológica (pseudoausencias realistas y control de extrapolación)

Un modelo solo puede considerarse robusto si cumple simultáneamente estos tres criterios. La ausencia de cualquiera de ellos puede producir mapas futuros aparentemente precisos, pero ecológicamente inválidos.

- **Incertidumbre y limitaciones en la modelación de distribución de especies bajo cambio climático**

Los modelos de distribución de especies constituyen aproximaciones probabilísticas al nicho climático y, por tanto, sus proyecciones futuras deben interpretarse como hipótesis espaciales condicionadas a supuestos ecológicos y estadísticos. La incertidumbre en SDM no proviene de una única fuente, sino de la acumulación de errores derivados de datos biológicos, variables ambientales, algoritmos de modelación y escenarios climáticos [17]

- **Incertidumbre asociada a los datos biológicos**

Los registros de presencia representan observaciones oportunistas y no un muestreo sistemático del espacio ambiental. En consecuencia, tienden a concentrarse en zonas accesibles o cercanas a centros poblados, generando sesgos espaciales de muestreo. Aunque la rarefacción espacial reduce este efecto, no elimina completamente la influencia de accesibilidad humana sobre la estimación del nicho.

Adicionalmente, la ausencia verdadera raramente es conocida. La utilización de pseudoausencias implica asumir que las localidades sin registros representan condiciones no ocupadas, lo cual introduce incertidumbre estructural en la estimación de la función de idoneidad ambiental [47].

- **Incertidumbre ambiental**

Las variables climáticas utilizadas representan promedios temporales interpolados a partir de estaciones meteorológicas. Estas superficies poseen errores derivados de densidad de estaciones, interpolación y resolución espaciales. Como resultado, el modelo aprende relaciones con aproximaciones del ambiente real y no con mediciones exactas.

Asimismo, la colinealidad residual entre variables puede inducir relaciones espurias, particularmente en algoritmos flexibles, afectando la interpretación ecológica del nicho estimado.

- **Incertidumbre algorítmica**

Cada algoritmo de modelación representa el nicho mediante estructuras matemáticas diferentes. Por ello, distintos métodos pueden producir mapas divergentes aun utilizando los mismos datos. Este fenómeno, conocido como incertidumbre estructural, implica que no existe una única solución al problema de modelación ecológica.

Adicionalmente, los modelos pueden capturar estructuras espaciales en lugar de relaciones ecológicas

cuando existe autocorrelación espacial residual. Aunque la validación cruzada espacial reduce este efecto, no lo elimina completamente, especialmente en especies con rangos geográficos restringidos.

- **Incertidumbre por extrapolación climática**

La principal fuente de incertidumbre en SDM futuros es la transferencia del modelo hacia condiciones ambientales no observadas en el presente. Los modelos aprenden relaciones dentro del dominio ambiental disponible, pero al proyectarse bajo cambio climático pueden enfrentarse a combinaciones novedosas de temperatura y precipitación.

En estas situaciones el modelo extrapola matemáticamente el nicho sin garantía de validez ecológica, pudiendo generar áreas climáticamente adecuadas donde la especie no podría sobrevivir debido a limitaciones fisiológicas, interacciones bióticas o barreras de dispersión [17].

- **Incertidumbre de escenarios climáticos**

Los escenarios climáticos provienen de modelos de circulación global que difieren en estructura física, parametrización y supuestos de emisiones. Por tanto, la variabilidad entre escenarios no representa error del modelo ecológico sino incertidumbre inherente al sistema climático.

En consecuencia, las diferencias entre proyecciones deben interpretarse como un rango de posibles condiciones futuras y no como predicciones excluyentes.

- **Limitaciones ecológicas del enfoque correlativo**

Los SDM correlativos asumen equilibrio entre la especie y el clima actual. Sin embargo, las distribuciones reales pueden estar restringidas por dispersión limitada, historia evolutiva o interacciones bióticas. Bajo cambio climático, estos factores pueden impedir que la especie ocupe todas las áreas climáticamente adecuadas proyectadas.

Por ello, los mapas generados representan idoneidad climática potencial y no distribución futura garantizada.

3.2 ANTECEDENTES

La práctica del SDM ha florecido en las últimas cuatro décadas en parte porque la localidad de especies geoespaciales y los datos ambientales se han vuelto tan ampliamente disponibles. Las predicciones espaciales, así como los propios modelos, pueden servir de base para la investigación básica y aplicada sobre biodiversidad [15].

Se han utilizado múltiples algoritmos estadísticos para la modelización de la distribución de especies. Debido a las deficiencias de los conjuntos de datos sobre la presencia de especies, cada vez se utilizan más los métodos basados únicamente en la presencia (como MaxEnt). Sin embargo, el sesgo de muestreo sigue siendo un problema, sobre todo en los métodos basados en la densidad. El algoritmo Isolation Forest

(iForest) es un método basado exclusivamente en la presencia menos sensible a los patrones de muestreo y al sobreajuste, ya que ajusta el modelo describiendo las condiciones inadecuadas en lugar de las adecuadas que proporciona herramientas convenientes para el diagnóstico del modelo y el análisis posterior a la modelización. También ayuda a los usuarios a comprender las relaciones entre las especies y el entorno vital utilizando los valores de Shapley, una técnica sugerida en la inteligencia artificial explicable (xAI). Además, puede elaborar mapas de respuesta espacial que indican cómo responden las especies a las variables ambientales en el espacio y detectan áreas potencialmente afectadas por un entorno cambiante [28].

Por otra parte [29], realizo un estudio sobre modelado de nicho ecológico para caracterizar la expansión del género *Lonchura* en México, donde la base de datos fue tratada para reducir la autocorrelación espacial de los datos, para el sobreajuste se utilizó distancias geográficas (10 km) que generó la capacidad predictiva del modelo, dentro del análisis de utilizaron capas ambientales del WorldClim y una resolución espacial de 5 minutos de arco. El modelo se realizó utilizando el algoritmo MaxEnt, esto para cada especie y se proyectó el modelo seleccionado, además, se realizó el análisis de similitud de nicho, el MOP, las curvas ROC Y ROC parcial y las curvas de tasa de omisión; el uso de estas técnicas de correlación permite conocer el riesgo de la presencia de especies no nativas en áreas geográficas nuevas.

Los modelos de nicho ecológico permiten hacer inferencias sobre la distribución de especies y comunidades. Es por esto que el objetivo de esta investigación fue generar modelos de nicho ecológico como herramienta para estimar la distribución de comunidades vegetales en Michoacán. Para delimitar su distribución, se realizaron 76 modelos de nicho ecológico. “Los modelos individuales fueron ensamblados para obtener el modelo de distribución del bioma y el mapa resultante fue comparado con las propuestas del SDTF. El uso de especies de amplia distribución en la definición de comunidades resulta en modelos con mayor sobreestimación” [30].

Por otra parte, el modelado de nichos ecológicos implica la transferencia de modelos calibrados con datos ambientales de una región a otras regiones que pueden incluir nuevas condiciones ambientales. Cuando existen nuevas condiciones, la transferibilidad implica extrapolación, mientras que, en ausencia de tales condiciones, la transferibilidad es solo un paso de interpolación. En este estudio se evaluó la transferibilidad de modelos producidos utilizando 11 algoritmos de nichos ecológicos desde la perspectiva de la interpolación y la extrapolación en un marco de especies virtuales distribuidas en Eurasia, estos modelos se evaluaron en el espacio ambiental, para complementar la evaluación geográfica tradicional de los modelos, concluyendo que estos modelos pueden anticipar invasiones, enfermedades, distribución de especies en condiciones climáticas futuras [31].

Así mismo, en el estudio *Combining physiological threshold knowledge to species distribution models is key to improving forecasts of the future niche for macroalgae*, se combinaron SDM correlativos y conocimiento sobre límites fisiológicos para proporcionar predicciones más robustas. Los umbrales térmicos obtenidos en experimentos de crecimiento y supervivencia se utilizaron como indicadores de los nichos fundamentales de dos macrófitos marinos fundacionales. Las proyecciones geográficas de las distribuciones de estas especies obtenidas utilizando estos umbrales y SDM existentes fueron similares en áreas donde las especies están ausentes-raras o frecuentes y donde sus nichos potenciales y realizados coinciden, alcanzando predicciones de consenso. Se predijo que el alga fundacional templada-fría *Himanthalia elongata* se extinguiría en su límite sur en el norte de España en respuesta al calentamiento global, mientras que se esperaba que aumentara la ocupación de *Bifurcaria bifurcata*, una

especie lusitana del sur [32].

4. METODOLOGÍA

El proyecto de investigación tiene un enfoque teórico-práctico y busca alcanzar un objetivo mediante la aplicación de conocimientos, metodologías y herramientas existentes para el manejo de bases de datos. Principalmente, el proyecto se centra en determinar las variaciones espaciales y temporales de las condiciones climáticas ideales para cultivos como el banano, el aguacate y el café, ya que, debido al cambio climático, estas condiciones podrían cambiar en un plazo de 30 (2021 a 2040) y 50 (2041 - 2060) años, lo que implicaría la reubicación de las zonas actuales de producción.

Para lograr el desarrollo del proyecto, se plantearon las actividades que se detallan en la Tabla 4.1, correspondientes a cada uno de los objetivos específicos presentados anteriormente.

Tabla 4.1. Actividades por objetivos específicos.

| | |
|-------------------------------------|--|
| <p>Objetivo específico 1</p> | <p>Caracterizar la ubicación geográfica actual de los cultivos de café, banano y aguacate en Colombia, así como los rangos climáticos actuales en los que se desarrollan.</p> <p>Actividades</p> <ol style="list-style-type: none"> 1. Acceder a los mapas de aptitud de cultivos: La Unidad de Planificación Agropecuaria (UPRA), a través de la plataforma de Datos Abiertos del Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia, pone a disposición los mapas de aptitud para los cultivos de banano, aguacate y café, los cuales se descargaron para su posterior análisis. 2. Acceder a datos climáticos: Para conocer la información climática de todo el territorio, se debió descargar las interpolaciones históricas (1970-2000) de precipitación, temperatura máxima y temperatura mínima, proporcionadas en la plataforma WorldClim. 3. Realizar zonificación: Los mapas de aptitud debieron cruzarse con los mapas climáticos para identificar los rangos climáticos óptimos actuales para los cultivos. |
| <p>Objetivo específico 2</p> | <p>Determinar las variaciones atmosféricas durante un período de 30 a 50 años, utilizando modelos de pronóstico existentes.</p> <p>Actividades</p> <ol style="list-style-type: none"> 1. Acceso a información climática futura: Se utilizaron los datos climáticos futuros de los modelos proporcionados por WorldClim, correspondientes a los períodos 2021-2040 (30 años) y 2041-2060 (50 años). Entre estos modelos existen ACCESS-CM2, BCC-CSM2-MR, CMCC-ESM2, EC-Earth3-Veg, FIO-ESM-2-0, GFDL-ESM4, GISS-E2-1-G, HadGEM3-GC31-LL, INM-CM5-0, IPSL-CM6A-LR, MIROC6, MPI-ESM1-2-HR, MRI-ESM2-0 y UKESM1-0-LL. 2. Seleccionar un modelo de pronóstico climático: Se realiza un estudio teórico para seleccionar el modelo de predicción climática existente, más adecuado para las condiciones climáticas de Colombia. 3. Realizar la limpieza de datos: Se extrajeron los datos únicamente para la zona geográfica de Colombia. |

| | |
|-------------------------------------|---|
| <p>Objetivo específico 3</p> | <p>Aplicar las metodologías de los modelos de distribución de especies (SDM) para las zonas de producción recomendadas para los cultivos, para ser usados en las proyecciones actuales y futuras (30 y 50 años), estableciendo el modelo más adecuado a partir de métricas de desempeño y validación cruzada.</p> <hr/> <p>Actividades</p> <ol style="list-style-type: none"> 1. Seleccionar el modelo: Elección de los modelos de distribución de especies (SDM). 2. Realizar el entrenamiento: se generaron modelos de entrenamiento para probar los modelos disponibles, sobre los datos climáticos futuros, para identificar posibles zonas donde se presentarán condiciones óptimas para el banano, el aguacate y el café. 3. Calcular métricas y validaciones: Se calcularon estos parámetros estadísticos en cada modelo para poder compararlos y seleccionar el más apropiado en función del objetivo final. 4. Realizar la predicción: se realizó la predicción a partir del modelo que presentó un mejor desempeño y métricas. |
| <p>Objetivo específico 4</p> | <p>Analizar la viabilidad de las zonas generadas por los modelos, considerando posibles restricciones no climáticas.</p> <hr/> <p>Actividades</p> <ol style="list-style-type: none"> 1. Acceder a los mapas las variables no climáticas: Se descargaron los mapas existentes de puertos marítimos, la presencia de vías, zonas protegidas, tipo de suelo. 2. Evaluar los mapas: Se evaluó la viabilidad de las ubicaciones generadas por los modelos, teniendo en cuenta las variables no climáticas. |

5. CARACTERIZAR LA UBICACIÓN GEOGRÁFICA ACTUAL DE LOS CULTIVOS DE CAFÉ, BANANO Y AGUACATE EN COLOMBIA, ASÍ COMO LOS RANGOS CLIMÁTICOS ACTUALES EN LOS QUE SE DESARROLLAN.

5.1 Actividades

Acceso y recopilación de mapas de aptitud agroclimática.

Como parte del análisis espacial de la idoneidad agrícola en Colombia, se accedió a los mapas oficiales de aptitud agroclimática generados por la Unidad de Planificación Rural Agropecuaria (UPRA). Dichos insumos fueron descargados a través de la plataforma de Datos Abiertos del Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia.

Se seleccionaron los mapas correspondientes a los cultivos de interés: banano (*Musa spp.*), aguacate (*Persea americana*) y café (*Coffea spp.*). Estos productos cartográficos contienen la evaluación técnica de las condiciones edafoclimáticas y topográficas del territorio nacional, clasificadas en categorías de aptitud (alta, media, baja y no apta).

Los mapas de aptitud constituyen un insumo fundamental para la identificación de zonas con potencial productivo bajo condiciones actuales, dado que integran criterios biofísicos relevantes para el establecimiento y desarrollo de los cultivos analizados. Previo a su utilización, se verificó la integridad de los archivos, el sistema de referencia espacial y la coherencia de la clasificación temática.

Obtención de datos climáticos interpolados.

Con el propósito de caracterizar las condiciones climáticas actuales del territorio colombiano, se descargaron capas climáticas interpoladas de precipitación anual, temperatura mínima y temperatura máxima desde la base de datos de WorldClim. Estas variables bioclimáticas fueron seleccionadas por su reconocida influencia en la determinación de la aptitud agroclimática y en la distribución espacial de cultivos tropicales. Los datos presentan una resolución espacial adecuada para estudios regionales y han sido ampliamente utilizados y validados en aplicaciones de modelación ecológica y agrícola. Posteriormente, las capas climáticas fueron recortadas al límite político-administrativo de Colombia con el fin de garantizar la correspondencia espacial con los mapas de aptitud agroclimática y optimizar la eficiencia computacional en los análisis posteriores.

Zonificación agroclimática mediante integración espacial.

Una vez obtenidos los mapas de aptitud y las variables climáticas interpoladas, se procedió a realizar la zonificación agroclimática para los cultivos de banano, aguacate y café en Colombia. Este proceso consistió en el cruce espacial de las capas de aptitud con las variables climáticas, mediante técnicas de análisis geoespacial en sistemas de información geográfica (SIG). El objetivo fue identificar los rangos climáticos óptimos actuales asociados a las zonas con mayor aptitud para cada cultivo, permitiendo establecer umbrales de temperatura y precipitación que caracterizan las condiciones ideales de producción.

5.2 Resultados

Mapas de aptitud de cultivos

La Unidad de Planificación Rural Agropecuaria (UPRA) define la aptitud de la tierra como el “potencial para el establecimiento de un tipo específico de uso del suelo, determinado a partir del análisis de combinaciones de criterios físicos, socioecosistémicos y socioeconómicos, enmarcados en una frontera agrícola”. Esta metodología tiene como propósito categorizar espacialmente el territorio en cinco niveles de aptitud (ver Tabla 5.1), que reflejan distintos grados de idoneidad del terreno para el cultivo objeto de estudio. La categorización se realiza mediante un análisis de superposición de múltiples variables asociadas a cada uno de los criterios mencionados, lo que permite identificar las zonas más adecuadas para dicho uso agrícola.

Tabla 5.1. Categorías de aptitud en evaluación de tierras.

| Categoría | Concepto |
|------------------------|---|
| Aptitud alta | Presenta limitaciones menores para implementar el Tipo de Utilización de la Tierra (TUT). Productividad y sostenibilidad recomendada con aptitudes sociales y económicas favorables en el área de estudio. |
| Aptitud media | Presenta limitaciones para implementar el TUT. Procesos de inversión y productividad moderados. Aptitudes sociales y económicas reducidas en el área de estudio. |
| Aptitud baja | Presenta altas limitaciones para implementar el TUT con procesos de inversión variables. Puede disminuir considerablemente la productividad. Aptitudes sociales y económicas muy limitadas en el área de estudio. |
| No apto | Presenta condiciones marginales para la implementación del TUT, que afectan considerablemente la sostenibilidad y productividad. |
| Exclusión legal | Restringe totalmente la aptitud del TUT por presentarse en zonas donde normativamente no puede realizarse usos agropecuarios. |

Los mapas resultantes han sido puestos a disposición del público a través del portal del Sistema de Información para la Planificación Rural Agropecuaria (SIPRA). A partir de esta plataforma, se obtuvieron los correspondientes a los cultivos de aguacate (*Persea americana* var. Hass) y banano (*Musa* sp. AAA) del año 2019, así como el del cultivo de café (*Coffea arabica* L.) del año 2020. Estos muestran que Colombia tiene territorios con aptitud alta y media, para cultivos de aguacate de 2,410,321 ha, para cultivos de banano de 2,031,421 ha y para café de 4,784,305 ha (Figura 5.1).

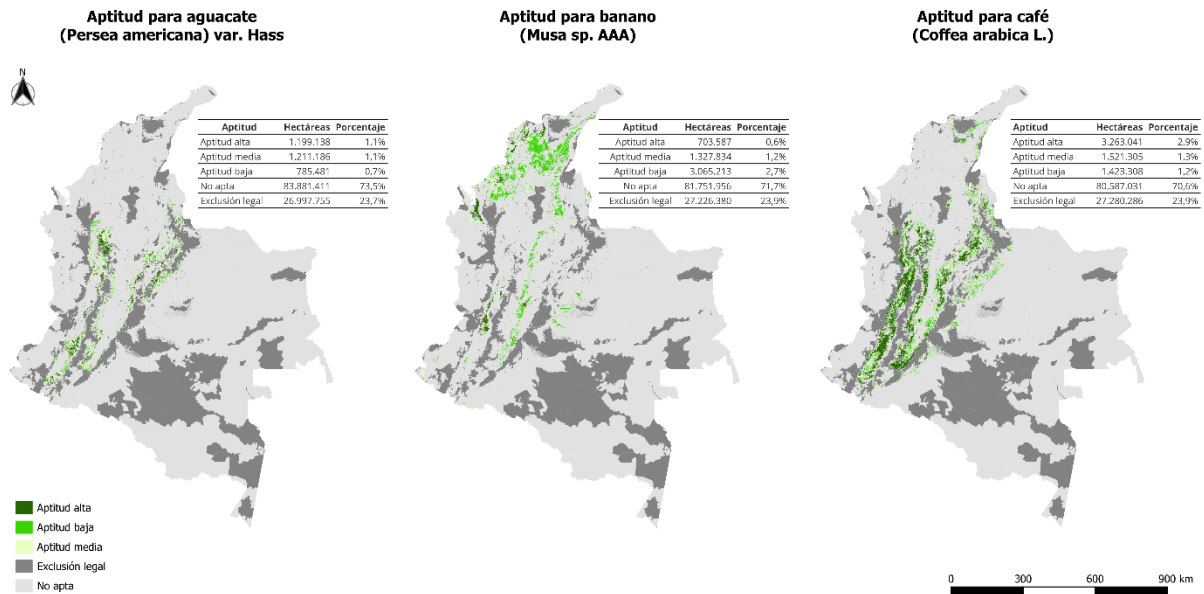


Figura 5.1: Mapas de aptitud de la tierra para los cultivos en estudio en Colombia, dispuestos por La Unidad de Planificación Agropecuaria (UPRA).

Los archivos obtenidos corresponden al tipo Shapefile, los cuales contienen información vectorial georreferenciada que incluye límites, polígonos y atributos asociados a las diferentes clases de aptitud. Estos archivos constituyen la base principal para el análisis espacial y geográfico dentro del proceso de zonificación agroclimática, permitiendo su integración con datos como temperatura y precipitación. Con el propósito de generar un análisis multicriterio.

La extracción de las zonas de alta aptitud (Figura 5.2) permite identificar y delimitar las áreas con condiciones óptimas para el desarrollo de cada cultivo. Esta información será fundamental para realizar recortes espaciales precisos respecto a cada especie agrícola (aguacate, banano y café), contribuyendo al análisis más detallado de potencial agroclimático, potencial productivo y planificación territorial. Además, estos recortes facilitan la aplicación de modelos predictivos de producción, así como la priorización de zonas para proyectos de establecimiento de cultivos en función de su aptitud agroecológica.

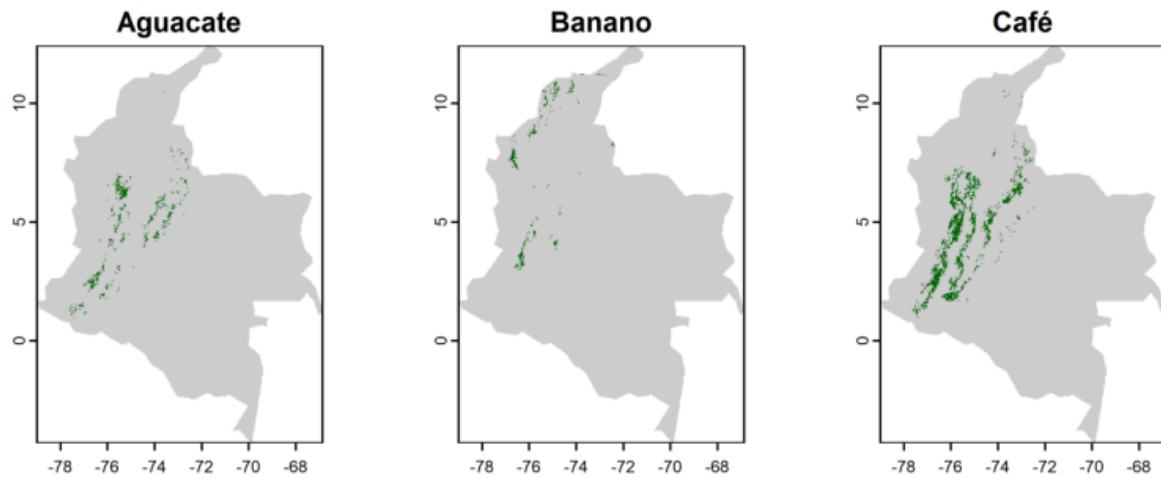


Figura 5.2: Mapas de aptitud alta para los cultivos en estudio en Colombia.

Datos climáticos

Se tomaron las variables climáticas de precipitación anual, temperatura mínima anual y temperatura máxima anual para todo el territorio colombiano, tal como se observa en la Figura 5.3. Estas variables se emplearon para caracterizar las condiciones agroclimáticas que influyen directamente en la aptitud de los cultivos seleccionados. Los resultados muestran que, en las zonas de alta aptitud identificadas para los cultivos de aguacate, banano y café, las precipitaciones se encuentran en un rango aproximado de 1000 a 2000 mm anuales, valores que garantizan una disponibilidad hídrica adecuada para el desarrollo de dichos cultivos.

Asimismo, se observó que la temperatura mínima promedio varía entre 10 y 20 °C, mientras que la temperatura máxima promedio oscila entre 20 y 35 °C. Estas condiciones térmicas son favorables para el crecimiento y productividad de las especies evaluadas. En conjunto, estos resultados confirman que las áreas clasificadas como de alta aptitud agroclimática presentan una distribución adecuada entre temperatura y precipitación, lo que sustenta su potencial para el establecimiento de los cultivos de estudio.

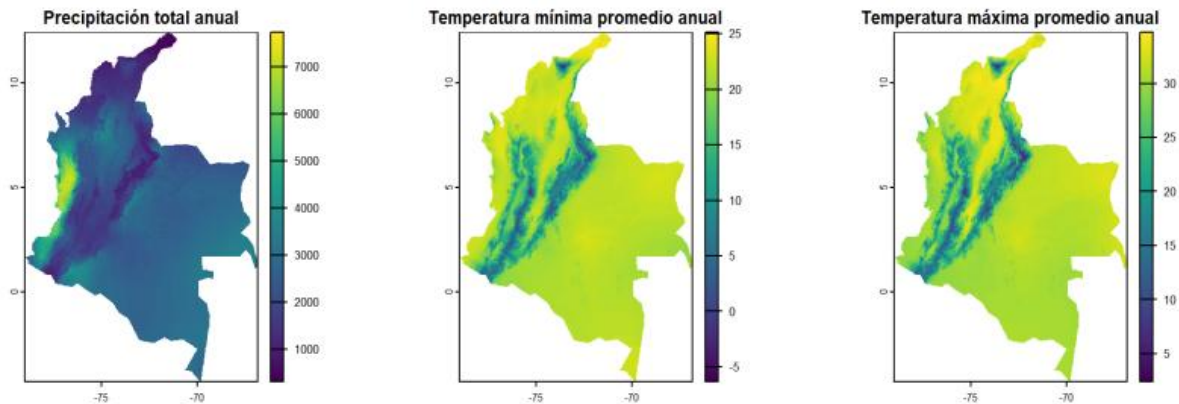


Figura 5.3: Variables climáticas de precipitación y temperatura para Colombia.

Zonificación climática de los cultivos

Al obtener los datos climáticos para Colombia se procedió a determinar para cada cultivo las zonas óptimas y sus rangos climáticos para el cultivo de aguacate (A), banano (B) y café (C) como se evidencia en la figura 5.4. Las zonas aptas para el aguacate se concentran principalmente en regiones de clima templado de la cordillera central y occidental. Las precipitaciones en estas áreas varían entre 1500 y 3000 mm anuales, mientras que la temperatura mínima oscila entre 8 y 16 °C y la temperatura máxima entre 22 y 26 °C. Estas condiciones son coherentes con los requerimientos agroclimáticos del cultivo, que prefiere climas moderadamente húmedos y temperaturas templadas, coincidiendo con los rangos óptimos reportados para aguacate de 20 a 25 °C y de 1200 a 1500 mm/año [33].

En el caso del banano, las zonas de mayor aptitud se ubican en áreas de clima cálido y húmedo, principalmente en el Urabá antioqueño y el litoral Pacífico. Las precipitaciones superan los 3000 mm anuales, alcanzando valores cercanos a 5000 mm, con temperaturas mínimas entre 18 y 22 °C y máximas entre 30 y 34 °C. Estas condiciones reflejan un ambiente tropical, con alta disponibilidad de humedad y temperaturas elevadas durante todo el año, factores esenciales para el crecimiento continuo del cultivo, lo cual coincide con sus condiciones óptimas de establecimiento para banano de 25 a 28 °C y 1200 a 1900 mm/año [34].

Las zonas aptas para café se localizan en las regiones de ladera de la Cordillera Central y Oriental, donde la precipitación varía entre 1500 y 3500 mm anuales, la temperatura mínima se sitúa entre 10 y 18 °C, y la máxima entre 20 y 28 °C. Estos rangos térmicos y pluviométricos son característicos de las zonas cafetaleras tradicionales del país, donde se logra un equilibrio entre humedad, temperatura y altitud, garantizando una buena calidad del grano y estabilidad en la producción y según lo reportado las condiciones ideales de establecimiento y desarrollo de las plantas están entre de 18 a 22 °C y 1500 a 1800 mm/año [35].

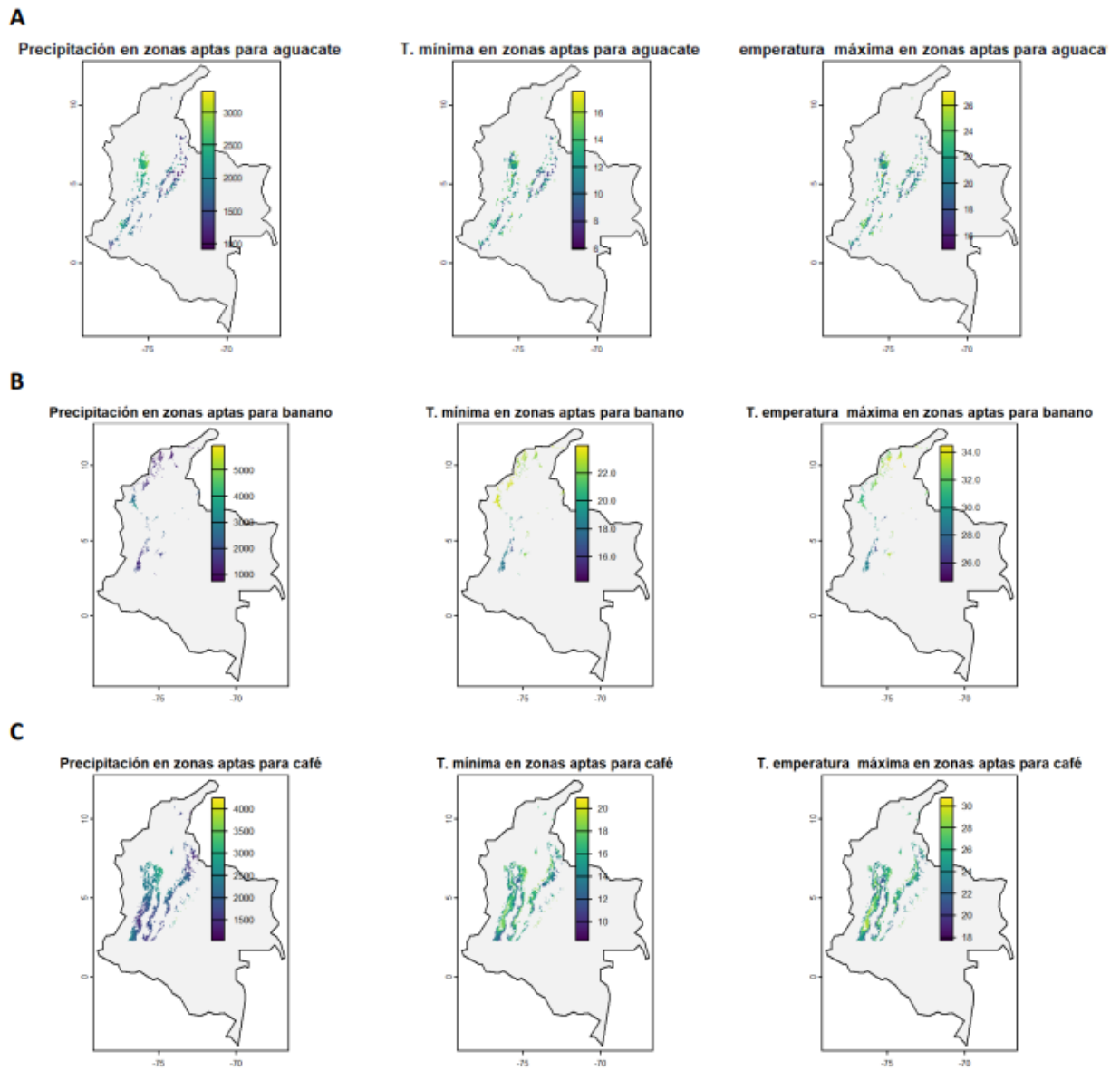


Figura 5.4: Variables climáticas de precipitación y temperatura aptas para el cultivo de aguacate (A), banano (B) y café (C).

6. DETERMINACIÓN DE LAS VARIACIONES ATMOSFÉRICAS DURANTE UN PERÍODO DE 30 A 50 AÑOS

6.1 Actividades

Acceso y recopilación de datos de modelos de pronóstico climático.

Con el fin de incorporar escenarios futuros de cambio climático en el análisis de distribución potencial de especies, se accedió a los datos de proyección climática disponibles en la plataforma WorldClim. Esta base de datos proporciona información climática derivada de múltiples Modelos de Circulación General (GCMs) utilizados en el marco del Coupled Model Intercomparison Project Phase 6 (CMIP6). Para este estudio, se seleccionaron dos horizontes temporales representativos: el período 2021–2040 (denominado horizonte de 30 años) y el período 2041–2060 (horizonte de 50 años), con el objetivo de evaluar posibles variaciones en la distribución espacial bajo distintos escenarios de cambio climático. Los modelos incluidos en el análisis fueron: ACCESS-CM2, BCC-CSM2-MR, CMCC-ESM2, EC-Earth3-Veg, FIO-ESM-2-0, GFDL-ESM4, GISS-E2-1-G, HadGEM3-GC31-LL, INM-CM5-0, IPSL-CM6A-LR, MIROC6, MPI-ESM1-2-HR, MRI-ESM2-0 y UKESM1-0-LL.

Selección teórica del modelo climático más adecuado.

Con el propósito de reducir la incertidumbre asociada a la variabilidad entre modelos climáticos, se llevó a cabo una revisión teórica y comparativa de los modelos disponibles, considerando su desempeño histórico en la región andina y su capacidad para representar adecuadamente las condiciones climáticas de Colombia. Esta evaluación incluyó criterios como la resolución espacial, la representación de variables clave (precipitación, temperatura), y la validación empírica en estudios previos. A partir de este análisis, se seleccionó el modelo que presenta mayor coherencia con los patrones climáticos observados en el país, el cual será utilizado como base para la proyección de escenarios futuros en los modelos de distribución de especies.

Procesamiento y depuración de datos climáticos.

Una vez obtenidos los datos climáticos proyectados, se procedió a realizar un proceso de limpieza y depuración de la información, con el fin de garantizar su pertinencia geográfica y su adecuación al área de estudio. Para ello, se extrajeron exclusivamente los datos correspondientes al territorio colombiano, delimitado mediante un shapefile oficial del país. Esta etapa incluyó la verificación de la integridad espacial de los datos, la eliminación de valores faltantes o inconsistentes, y la estandarización de las capas climáticas para su posterior integración en los modelos de distribución de especies. Este procedimiento asegura que los insumos utilizados en el modelado sean representativos y comparables en términos espaciales y temporales.

6.2 Resultados

Colombia, tiene una geografía que se distribuye en cordilleras, selvas, costas y Amazonía, por lo cual es ideal elegir modelos que representen bien estas regiones tropicales y montañosas, y que tengan resolución

relativamente alta. Por lo anterior, se identificaron 4 modelos que se pueden adaptar a las condiciones agroecológicas del país, estos son utilizados en el marco del CMIP6 (Proyecto de Intercomparación de Modelos Acoplados, Fase 6) para simular y proyectar el clima pasado, presente y futuro. Cada uno tiene características y fortalezas particulares en la representación de procesos físicos y biogeoquímicos. Los modelos EC-Earth3-Veg y MPI-ESM1-2-HR son consistentes para simular precipitación y temperatura en la Amazonía y el norte de Sudamérica, regiones que incluyen Colombia, ambos presentan bajos sesgos y alta correlación con datos observados, no obstante, este último no tiene en cuenta la dinámica de vegetación [37], siendo esto un factor importante para la simulación del modelo.

En el caso del modelo IPSL-CM6A-LR ha mejorado respecto a versiones anteriores, pero aún muestra sesgos en la representación de la Zona de Convergencia Intertropical y en extremos de precipitación. Y el modelo UKESM1-0-LL se destaca en la simulación de extremos de precipitación y temperatura, pero puede presentar sesgos en la distribución espacial de la precipitación, y su desempeño general es ligeramente inferior a EC-Earth3-Veg y MPI-ESM1-2-HR en la región [38].

Teniendo en cuenta cada una de las características, ventajas y desventajas de los modelos, se eligió el modelo EC-Earth3-Veg, dado a que este se destaca por su integración interactiva de vegetación dinámica, lo que permite simular de manera más realista los intercambios entre la atmósfera, la superficie terrestre y la biosfera, además cumple con las características de resolución espacial, simulación de fenómenos tropicales, capacidad para simular zonas montañosas, vegetación dinámica y cobertura global y proyecciones futuras, que permiten realizar un buen modelo de pronóstico (Tabla 6.1).

Tabla 6.1. Modelos de pronóstico climático para la zona de Colombia

| Modelo | Resolución Espacial | Simulación de Fenómenos Tropicales | Rendimiento en Precipitaciones | Capacidad para Simular Zonas Montañosas | Vegetación Dinámica | Cobertura Global y Proyecciones Futuras |
|----------------------|---------------------|------------------------------------|--------------------------------|---|---------------------|---|
| MPI-ESM1-2-HR | ~0.5° | Alta | Alta | Alta | No | Alta |
| IPSL-CM6A-LR | ~1° | Alta | Alta | Moderada | No | Alta |
| EC-Earth3-Veg | ~0.7° | Alta | Alta | Alta | Si | Alta |
| UKESM1-0-LL | ~1.25° | Moderada | Moderada | Moderada | Si | Alta |

Este modelo EC-Earth3-Veg acopla el modelo atmosférico IFS (ECMWF), el modelo oceánico NEMO3.6, el modelo de hielo marino LIM3 y el modelo de vegetación dinámica LPJ-GUESS. Esta integración permite simular procesos físicos, biogeoquímicos y de uso del suelo de forma interactiva [39]. A diferencia de otras configuraciones, EC-Earth3-Veg incluye la retroalimentación de la vegetación, lo que mejora la representación de los ciclos de carbono y agua, y la respuesta de los ecosistemas al cambio climático [40]. Además, ofrece alta resolución espacial (atmósfera ~80 km, océano ~1°) y puede configurarse para diferentes experimentos, desde simulaciones climáticas hasta proyecciones futuras [41]. EC-Earth3-Veg ha sido utilizado para simular el clima del Holoceno medio, el Plioceno y escenarios futuros, mostrando mejoras en la representación de lluvias monzónicas y retroalimentaciones vegetación-clima [39].

Escenarios climáticos

Los escenarios climáticos son herramientas fundamentales para proyectar posibles futuros del clima global, basados en diferentes trayectorias de emisiones de gases de efecto invernadero y cambios socioeconómicos. Estos escenarios guían la investigación, la política y la planificación de la adaptación y mitigación. Los escenarios más usados son los RCP (Representative Concentration Pathways) y los SSP (Shared Socioeconomic Pathways), que combinan trayectorias de emisiones y contextos socioeconómicos. Los SSPs, son adoptados en CMIP6 y el IPCC AR6, permiten explorar futuros desde el desarrollo sostenible (SSP1) hasta desarrollo basado en combustibles fósiles (SSP5), estos escenarios cubren desde trayectorias de bajas emisiones (SSP1-1.9, compatible con 1.5 °C) hasta escenarios de altas emisiones (SSP5-8.5, >4 °C de calentamiento) [42] [43].

El escenario SSP1-2.6 (Bajo) ve un mundo sostenible, con cooperación internacional, energías limpias y políticas climáticas ambiciosas. Las emisiones de gases de efecto invernadero disminuyen rápidamente, limitando el calentamiento global a menos de 2 °C, los cambios en precipitación son menores comparados con escenarios de altas emisiones [44]. Se proyectan ligeras disminuciones o estabilidad en la precipitación anual, con variaciones estacionales: posibles aumentos en primavera e invierno y leves reducciones en verano y otoño. La menor variabilidad climática y la reducción de eventos extremos favorecen la productividad agrícola y la seguridad alimentaria. Se promueve la transición a sistemas agrícolas sostenibles y eficientes en el uso del agua, además, menor frecuencia de olas de calor, inundaciones y sequías reduce riesgos sanitarios, enfermedades transmitidas por vectores y mortalidad asociada al clima y en el caso de infraestructura menos eventos extremos disminuyen daños a infraestructuras críticas, reduciendo costos de reparación y aumentando la resiliencia urbana y rural [45].

El escenario SSP2-4.5 proyecta un calentamiento global entre 2 °C y 3 °C para 2100, con políticas y desarrollo siguiendo tendencias actuales. Los impactos sectoriales son intermedios: más severos que en un mundo sostenible, pero menos que en escenarios de altas emisiones, además, se proyectan descensos o cambios leves en la precipitación anual en muchas regiones, con mayor variabilidad estacional y espacial. En algunos lugares, la precipitación disminuye hasta un 21% y los veranos tienden a ser más secos, aumentando el riesgo de sequías y eventos extremos [44].

En cuanto a la productividad agrícola se torna más vulnerable a sequías y olas de calor, con mayor variabilidad en los rendimientos y necesidad de adaptación (cultivos resistentes, tecnologías de riego eficiente) y el riesgo de inseguridad alimentaria aumenta en regiones sensibles, así mismo, se intensifican las amenazas a la biodiversidad y los servicios ecosistémicos, con mayor frecuencia de sequías, degradación de suelos y fragmentación de hábitats. Los humedales y bosques sufren por menor disponibilidad de agua y aumento de incendios, aumentan los riesgos de enfermedades transmitidas por vectores, golpes de calor y problemas respiratorios debido a olas de calor y contaminación y la mayor exposición a eventos extremos (inundaciones, sequías, erosión) incrementa los daños a infraestructuras críticas y eleva los costos de mantenimiento y adaptación [46].

El escenario SSP3-7.0 (Desigualdad alta) proyecta un mundo fragmentado, con alta desigualdad, baja cooperación internacional y un calentamiento global de 3.5–4.5 °C, lo que intensifica los riesgos climáticos y sociales en todos los sectores, por lo que se prevén aumentos significativos en la frecuencia e intensidad de eventos extremos, como lluvias intensas seguidas de olas de calor, así como una mayor variabilidad

estacional [44]. Además, el estrés hídrico se agrava por la mayor demanda y la irregularidad de las lluvias, lo que incrementa las sequías agrícolas e hidrológicas, afectando la disponibilidad de agua para consumo humano, riego y ecosistemas, por lo que la productividad agrícola disminuye por el aumento de olas de calor, sequías y eventos extremos. El riesgo de pérdidas de cosechas y de inseguridad alimentaria es alto, especialmente en regiones vulnerables y con baja capacidad de adaptación [46].

Así mismo, en este escenario se proyecta que se intensifique la degradación de hábitats, pérdida de biodiversidad y fragmentación de ecosistemas, donde los servicios ecosistémicos, como la regulación hídrica y la calidad del suelo, se ven gravemente afectados, además, aumentan los riesgos de enfermedades transmitidas por vectores, golpes de calor, desnutrición y mortalidad asociada a eventos extremos, y la mayor frecuencia de inundaciones, sequías y olas de calor provoca daños recurrentes a infraestructuras críticas (agua, energía, transporte), elevando los costos de reparación y adaptación [47].

El escenario SSP5-8.5 (Muy alto) proyecta un desarrollo económico acelerado basado en combustibles fósiles, sin políticas de mitigación, con un calentamiento potencial de hasta 5.7 °C para 2100, esto implica impactos severos y generalizados en todos los sectores clave, donde se prevé un aumento en la frecuencia e intensidad de eventos extremos, como lluvias torrenciales y sequías prolongadas, en algunas regiones, la precipitación anual disminuye hasta un 22%, mientras que en otras puede aumentar hasta un 27%, intensificando la variabilidad y los riesgos de inundaciones y sequías [44].

En este escenario se proyecta que el estrés hídrico se agrave por la mayor demanda de agua para riego (hasta 12% más en 2100), disminución de caudales y recarga de acuíferos, y mayor duración e intensidad de las sequías, especialmente en zonas áridas y semiáridas. Aumentan las pérdidas de rendimiento por sequía, olas de calor y eventos extremos, donde se incrementa la demanda de agua para riego y la inseguridad alimentaria. Además, se intensifica la degradación de hábitats, pérdida de biodiversidad y fragmentación de ecosistemas. Los servicios ecosistémicos, como la regulación hídrica y la retención de suelos, se ven gravemente afectados, con mayor erosión y sedimentación. Así mismo, se aumentan los riesgos de enfermedades transmitidas por vectores, golpes de calor, desnutrición y mortalidad asociada a eventos extremos y mayor frecuencia de inundaciones, sequías y olas de calor provoca daños recurrentes y costosos a infraestructuras críticas (agua, energía, transporte), especialmente en áreas urbanas densamente pobladas [46].

Luego de revisar los diferentes contextos de escenarios climáticos, se eligió el escenario SSP2-4.5, que puede tener mejores rangos climáticos para generar el mejor pronóstico para el estudio, teniendo en cuenta cada una de las características descritas en la Tabla 6.2.

Tabla 6.2. Contextos de escenarios climáticos.

| Variable / Sector | SSP1-2.6 (Bajo) | SSP2-4.5 (Intermedio) | SSP3-7.0 (Desigualdad alta) | SSP5-8.5 (Muy alto) |
|--------------------------------|--|---|--|--|
| Temperatura media anual | +1.0 – 1.5 °C | +2.0 – 2.5 °C | +3.0 – 4.0 °C | +3.5 – 5.0 °C |
| Precipitación | Ligeros cambios; más lluvias en selvas | Aumentos regionales; disminución en Caribe seco | Aumentos desorganizados; más extremos regionales | Cambios extremos; lluvias intensas y sequías recurrentes |

| | | | | |
|------------------------------------|------------------------------------|---|--|--|
| Agricultura | Adaptación viable con políticas | Pérdidas moderadas sin adaptación | Alta vulnerabilidad rural; mayor inseguridad alimentaria | Colapso en rendimiento sin medidas; pérdidas masivas |
| Recursos hídricos | Manejables con gestión sostenible | Mayor estrés en zonas secas | Escasez hídrica crítica en regiones vulnerables | Crisis hídrica generalizada, especialmente en Andes y Caribe |
| Ecosistemas y biodiversidad | Conservación posible | Estrés en ecosistemas montañosos y costeros | Pérdidas aceleradas por uso de suelo y clima | Extinción de especies clave; colapso de hábitats |
| Fenómenos extremos | Eventos controlables | Más lluvias intensas y calor | Sequías prolongadas, incendios forestales, inundaciones | Aumento exponencial de eventos extremos e impactos |
| Salud pública | Riesgo bajo; buen manejo sanitario | Aumento controlable de enfermedades vectoriales | Alta exposición a enfermedades tropicales y estrés térmico | Brotos frecuentes de dengue, malaria y olas de calor graves |
| Infraestructura | Adaptación sencilla | Riesgos en zonas costeras y laderas | Destrucción recurrente de infraestructuras vulnerables | Daños graves a ciudades costeras, vías y servicios básicos |

Basados en el modelo EC-Earth3-Veg y el escenario SSP2-4.5, se obtuvo la distribución espacial proyectada de las variables climáticas futuras (Figura 6.1). Este modelo permite simular con alta resolución los cambios en temperatura y precipitación bajo condiciones intermedias de emisiones.

Los resultados muestran tendencias de incremento en la temperatura media anual y variaciones espaciales en la precipitación, con patrones heterogéneos según la región. En zonas andinas se proyecta un aumento térmico más pronunciado, mientras que las áreas de baja altitud podrían experimentar reducciones o redistribuciones de las lluvias.

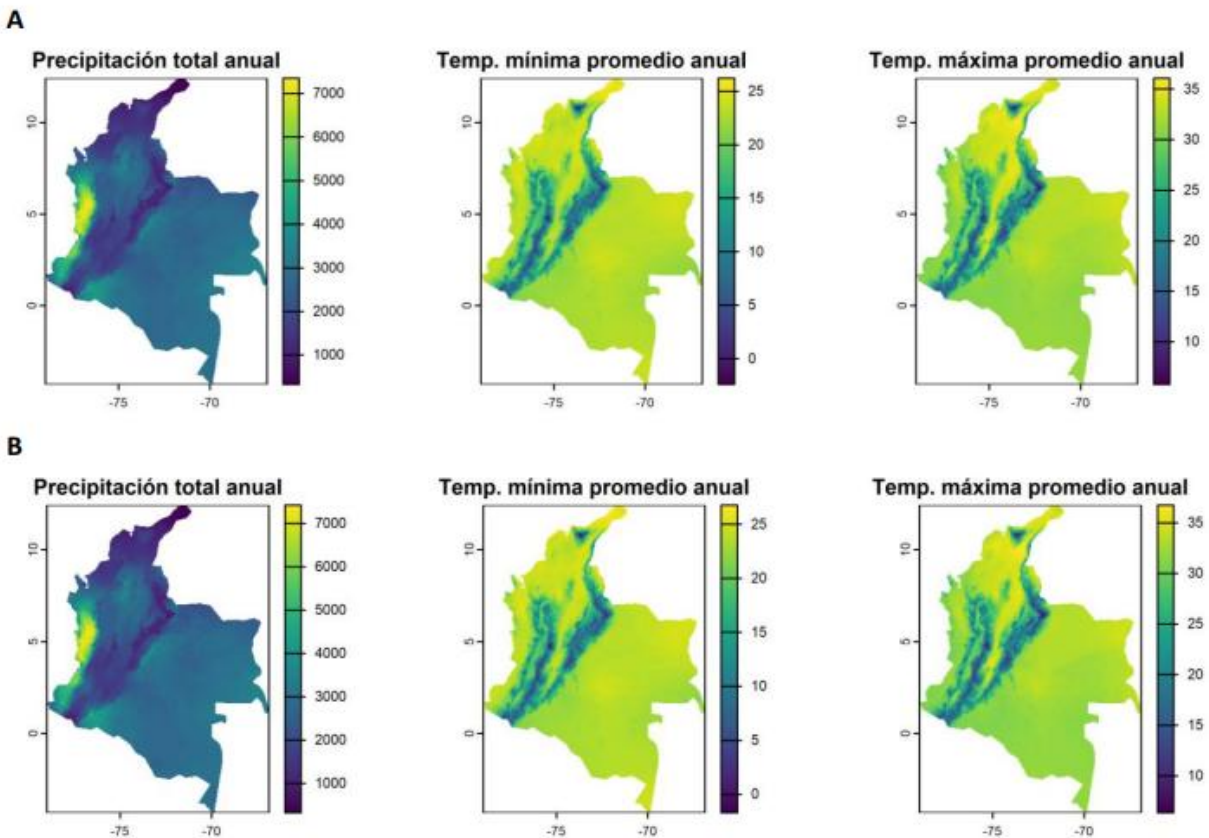


Figura 6.1: Modelo EC-Earth3-Veg para las variables climáticas entre los años 2021 – 2040 (A) y 2041 – 2060 (B) (SPP2-4.5).

De acuerdo a la figura 6.1A durante el periodo 2021–2040, se observa que las mayores precipitaciones continúan concentrándose en las zonas del Pacífico y el piedemonte amazónico, superando los 5000–7000 mm anuales, mientras que las regiones del Caribe, centro del país y oriente registran valores menores, entre 1000 y 3000 mm. Las temperaturas mínimas promedio oscilan entre 10 °C y 20 °C en la región andina y aumentan hacia las tierras bajas, alcanzando valores superiores a 24 °C. Las temperaturas máximas se distribuyen principalmente entre 25 °C y 35 °C, con los mayores valores en la región Caribe y Orinoquía.

En el segundo periodo, 2041–2060 (figura 6.1B), el modelo proyecta un incremento generalizado de la temperatura, con aumentos de aproximadamente 0.5 °C a 1.5 °C en la media nacional. Este calentamiento es más evidente en las zonas interandinas y del Caribe, donde la temperatura máxima podría superar los 34 °C. La precipitación muestra un patrón espacial similar al del periodo anterior, y aumentos marginales en regiones de montaña y la vertiente pacífica.

7. METODOLOGÍAS DE LOS MODELOS DE DISTRIBUCIÓN DE ESPECIES (SDM) PARA LAS ZONAS DE PRODUCCIÓN EN LOS PRÓXIMOS 30 Y 50 AÑOS

7.1. Actividades

Desarrollo de los modelos de nicho ecológico

La modelación del nicho climático se realizó siguiendo un flujo estructurado que incluyó: (i) compilación y depuración de registros de presencia, (ii) selección y preprocesamiento de variables ambientales, (iii) partición de datos en conjuntos de entrenamiento y validación, (iv) calibración de los modelos mediante el paquete biomod2, (v) evaluación del desempeño usando métricas ROC, TSS, Kappa, sensibilidad y especificidad, y (vi) proyección espacial de la idoneidad climática.

Selección de modelos.

En este estudio se emplearon cuatro algoritmos de modelado de distribución de especies con enfoques complementarios:

- **GAM (Modelos Aditivos Generalizados)**, que permiten capturar relaciones no lineales entre las variables ambientales mediante funciones suavizadas.
- **GLM (Modelos Lineales Generalizados)**, ampliamente utilizados por su interpretabilidad y flexibilidad. Para este modelo se utilizó una distribución binomial con función de enlace logit, apropiada para modelar la probabilidad de presencia de los cultivos en función de las variables ambientales.

La suposición de linealidad en el logit para las variables continuas fue evaluada mediante gráficos de residuos parciales y análisis exploratorios. Cuando se detectaron desviaciones de la linealidad, se consideró la inclusión de términos cuadráticos y transformaciones funcionales para mejorar el ajuste del modelo.

- **MAXNET**, implementación moderna del algoritmo MaxEnt basada en regularización tipo LASSO, adecuada para datos de presencia únicamente.
- **RF (Random Forest)**, método de aprendizaje automático basado en árboles de decisión, reconocido por su capacidad para capturar interacciones complejas y manejar relaciones no lineales.

La combinación de estos modelos permitió evaluar el desempeño predictivo desde distintas perspectivas estadísticas y computacionales, fortaleciendo la robustez de los resultados obtenidos.

Diagnóstico del modelo GLM

Se evaluó la posible sobredispersión del modelo mediante la razón entre la desviación residual y los grados de libertad. Los valores obtenidos indicaron que la dispersión se encontraba dentro de rangos aceptables, por lo que no fue necesario ajustar modelos cuasi-binomiales.

Adicionalmente, se realizó un análisis de diagnóstico mediante la inspección de residuos de Pearson y deviance, así como la evaluación de valores de leverage y distancia de Cook, con el fin de identificar observaciones influyentes. No se evidenciaron puntos con influencia desproporcionada sobre el ajuste del modelo.

Preparación de los datos de presencia–ausencia

La capa de presencia (aptitud alta del cultivo) fue transformada en un conjunto de puntos con coordenadas geográficas (XY) y valores binarios que indican presencia (1) o ausencia (0). Esta conversión permitió estructurar los datos en formato tabular requerido por los algoritmos de modelación.

La extracción de coordenadas facilitó la vinculación de cada observación con sus condiciones ambientales locales, constituyendo un paso fundamental para la inferencia ecológica y la calibración de los modelos predictivos.

Estrategia de partición y validación espacial

Dado que los registros de presencia pueden presentar autocorrelación espacial, la validación aleatoria simple puede generar estimaciones optimistas del desempeño del modelo. Para reducir este sesgo, se implementó un esquema de validación espacial mediante partición por bloques geográficos, garantizando la independencia espacial entre los conjuntos de entrenamiento y validación.

Este enfoque permitió obtener métricas de evaluación más realistas y robustas frente a la dependencia espacial de los datos.

Realización del entrenamiento de modelos.

El flujo de trabajo fue estructurado utilizando el paquete biomod2, que permite integrar múltiples algoritmos de modelado de distribución de especies dentro de un marco reproducible y estandarizado. A través de esta herramienta se realizó:

- el formateo de datos de presencia-ausencia,
- la calibración de modelos individuales (GAM, GLM, MAXNET y RF),
- la validación cruzada espacial,
- y la evaluación del desempeño.

El uso de este marco integrador facilitó la comparación entre algoritmos y la selección de aquellos con mayor capacidad predictiva.

Cálculo de métricas y validaciones.

La evaluación del desempeño se realizó mediante validación cruzada espacial por bloques. Para cada algoritmo se calculó la media y desviación estándar de las métricas (AUC, TSS y Kappa) a través de los pliegues, con el fin de cuantificar la estabilidad y robustez de los modelos. Estas métricas fueron calculadas sobre los conjuntos de validación espacialmente independientes, con el fin de evitar la sobreestimación del desempeño asociada a la autocorrelación espacial.

Realización de la predicción.

Los modelos calibrados se proyectaron sobre capas climáticas actuales y escenarios futuros (horizontes de 30 y 50 años), con el fin de estimar la distribución potencial de los cultivos bajo condiciones de cambio climático.

Las proyecciones se realizaron utilizando capas ráster previamente armonizadas en resolución, extensión y sistema de referencia espacial, garantizando la coherencia geográfica de los resultados. Esta etapa

permitió identificar posibles desplazamientos en la idoneidad climática y evaluar implicaciones agroecológicas para la planificación territorial.

7.2. Resultados

Desempeño y selección del mejor modelo

Se evaluaron cuatro algoritmos de modelado (GAM, GLM, MAXNET y RF), los valores reportados corresponden al promedio y la desviación estándar obtenidos a partir de cinco particiones de validación cruzada espacial tipo block. La variabilidad observada entre repeticiones fue moderada, lo que indica estabilidad en el desempeño de los modelos, particularmente para GLM y MAXNET. Random Forest mostró alta variabilidad relativa y bajo poder predictivo, sugiriendo sobreajuste y limitada capacidad de generalización. Los resultados (Tabla 7.1) permiten establecer una comparación integral del desempeño predictivo de cada modelo para el cultivo de aguacate.

Tabla 7.1. Desempeño de los modelos bajo validación cruzada espacial en el cultivo de aguacate

| Algoritmo | Kappa (\pm DE) | ROC (\pm DE) | TSS (\pm DE) | Especificidad (\pm DE) | Sensibilidad (\pm DE) |
|-----------|-------------------|------------------|-------------------|---------------------------|--------------------------|
| GAM | 0.178 \pm 0.080 | 0.95 \pm 0.029 | 0.764 \pm 0.066 | 91.92 \pm 2.90 | 91.10 \pm 15.25 |
| GLM | 0.184 \pm 0.056 | 0.94 \pm 0.033 | 0.812 \pm 0.094 | 91.47 \pm 3.28 | 90.43 \pm 15.68 |
| MAXNET | 0.184 \pm 0.055 | 0.94 \pm 0.032 | 0.796 \pm 0.084 | 91.52 \pm 3.23 | 89.71 \pm 17.46 |
| RF | 0.030 \pm 0.060 | 0.90 \pm 0.042 | 0.031 \pm 0.061 | 83.33 \pm 38.92 | 100 \pm 0.0 |

La validación cruzada espacial evidenció que los modelos GLM, MAXNET y GAM presentaron una capacidad predictiva adecuada (TSS > 0.75), siendo GLM el algoritmo con mejor desempeño global (TSS = 0.812 \pm 0.094). Por el contrario, Random Forest mostró un rendimiento muy bajo (TSS = 0.031), lo que sugiere una pobre capacidad de generalización bajo partición espacial.

Los valores de Kappa fueron moderados (< 0.20), lo cual es esperable en modelos de distribución de especies debido a la sensibilidad de este índice a la prevalencia y al uso de pseudo-ausencias. En este contexto, el TSS se considera una métrica más robusta para la evaluación del desempeño.

En función de la consistencia, capacidad de generalización y equilibrio entre sensibilidad y especificidad, el modelo de GLM se posiciona como el modelo más robusto y confiable para aplicaciones predictivas y proyecciones bajo escenarios de cambio climático. Su desempeño sostenido en todas las etapas lo convierte en el candidato ideal para ser usado como base principal en el análisis de distribución potencial.

Cultivo de banano

Bajo validación cruzada espacial, los modelos mostraron un desempeño moderado para el cultivo de banano, con valores de ROC entre 0.65 y 0.74. Sin embargo, los bajos valores de TSS (0.03–0.29) y Kappa (0.01–0.14) indican una capacidad limitada de clasificación. MAXNET presentó el comportamiento más

equilibrado entre discriminación y estabilidad, al combinar valores relativamente superiores de Kappa (0.137) y ROC (0.71) con una variabilidad moderada entre bloques espaciales. En contraste, Random Forest evidenció una sensibilidad perfecta pero un TSS muy bajo, lo que sugiere sobrepredicción de la presencia y baja capacidad discriminativa. Estos resultados indican que el nicho climático del banano podría ser relativamente amplio o estar influenciado por factores no incluidos en el modelo (Tabla 7.2).

En conjunto, MAXNET se considera el modelo más robusto para la predicción de la presencia del banano en las áreas evaluadas, debido a su mejor balance entre capacidad discriminativa y estabilidad bajo validación cruzada espacial, aunque las diferencias entre algoritmos no son marcadamente amplias.

Tabla 7.2. Desempeño de los modelos bajo validación cruzada espacial en el cultivo de banano.

| Algoritmo | Kappa (\pm DE) | ROC (\pm DE) | TSS (\pm DE) | Especificidad (\pm DE) | Sensibilidad (\pm DE) |
|-----------|-------------------|-----------------|-----------------|---------------------------|--------------------------|
| GAM | 0.112 \pm 0.11 | 0.65 \pm 0.10 | 0.29 \pm 0.12 | 88.99 \pm 4.44 | 91.27 \pm 8.88 |
| GLM | 0.063 \pm 0.09 | 0.66 \pm 0.03 | 0.24 \pm 0.10 | 83.19 \pm 11.42 | 77.57 \pm 21.47 |
| MAXNET | 0.137 \pm 0.14 | 0.71 \pm 0.17 | 0.28 \pm 0.19 | 87.82 \pm 4.32 | 88.34 \pm 10.16 |
| RF | 0.010 \pm 0.02 | 0.74 \pm 0.14 | 0.03 \pm 0.04 | 75.00 \pm 45.22 | 100 \pm 0.0 |

Cultivo de café

Como se observa en la Tabla 7.3, el modelo GAM presentó un rendimiento aceptable en todas las métricas, con un TSS de 0.817 \pm 0.09. Estos resultados indican buena capacidad de discriminación (ROC = 0.94 \pm 0.036), aunque la concordancia más allá del azar es moderada (Kappa = 0.365 \pm 0.09), sugiriendo cierta variabilidad en la predicción de presencia y ausencia del café.

El modelo GLM mostró un desempeño consistente y robusto, con un TSS de 0.804 \pm 0.11. Estos valores reflejan un modelo equilibrado, con alta capacidad de discriminación (ROC = 0.94 \pm 0.038), buena concordancia (Kappa = 0.429 \pm 0.14) y sensibilidad sólida (88.46 \pm 15.10%), sin evidencias claras de sobreajuste. El modelo MAXNET presentó un comportamiento similar al de GLM, con un TSS de 0.797 \pm 0.11 y un Kappa de 0.336 \pm 0.04, ligeramente inferior al de GLM. Esto lo posiciona como una alternativa sólida y confiable para predecir la distribución del café.

En el caso del modelo Random Forest (RF), su TSS cayó a 0.180 \pm 0.14, evidenciando limitada capacidad de generalización a nuevas áreas, a pesar de mostrar sensibilidad perfecta (100 \pm 0.0%) y especificidad muy alta (99.34 \pm 1.94%), lo que sugiere un comportamiento sesgado hacia la predicción de presencia. Después de la evaluación de los modelos de distribución, el GLM se seleccionó como el modelo más adecuado debido a su desempeño consistente y equilibrado en todas las métricas, reflejando buena capacidad de discriminación, concordancia y generalización a nuevas áreas.

Tabla 7.3. Desempeño de los modelos bajo validación cruzada espacial en el cultivo de café.

| Algoritmo | Kappa (\pm DE) | ROC (\pm DE) | TSS (\pm DE) | Especificidad (\pm DE) | Sensibilidad (\pm DE) |
|-----------|-------------------|------------------|------------------|---------------------------|--------------------------|
| GAM | 0.365 \pm 0.09 | 0.94 \pm 0.036 | 0.817 \pm 0.09 | 91.77 \pm 3.79 | 89.64 \pm 13.39 |
| GLM | 0.429 \pm 0.14 | 0.94 \pm 0.038 | 0.804 \pm 0.11 | 91.43 \pm 4.14 | 88.46 \pm 15.10 |
| MAXNET | 0.336 \pm 0.04 | 0.94 \pm 0.038 | 0.797 \pm 0.11 | 91.80 \pm 4.15 | 87.54 \pm 17.05 |
| RF | 0.269 \pm 0.12 | 0.93 \pm 0.040 | 0.180 \pm 0.14 | 99.34 \pm 1.94 | 100 \pm 0.0 |

Importancia media de las variables ambientales

Cultivo de aguacate

El análisis de importancia de las variables ambientales indicó que la temperatura mínima es el factor más determinante para la presencia del cultivo de aguacate, con un valor de importancia media de 0.95. Esto deja ver la alta sensibilidad del cultivo a heladas o bajas temperaturas, especialmente durante las etapas de desarrollo de la planta. En el caso de la temperatura máxima mostró una importancia media de 0.31, indicando que los valores extremos de calor influyen de manera moderada en la presencia del cultivo, mientras que la precipitación, con un valor de 0.13, tiene una contribución relativamente baja, sugiriendo que la disponibilidad de agua no es un factor crítico para determinar la distribución del aguacate en las áreas evaluadas (Figura 7.1).

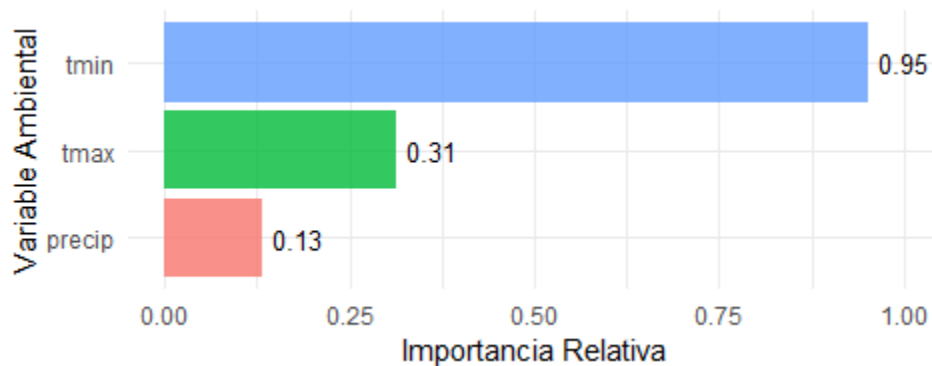


Figura 7.1: Importancia media de las variables ambientales en modelos para aguacate.

Cultivo de banano

El análisis de importancia de las variables ambientales indicó que la temperatura mínima es un factor determinante para la presencia del cultivo de banano, con un valor de importancia media de 0.63. En el caso de la precipitación y la temperatura máxima su importancia media presenta el mismo valor 0.61 (Figura 7.2). Estos resultados indican que las condiciones de temperatura mínima, pueden estar asociadas a la sensibilidad del cultivo a bajas temperaturas o heladas. Por su parte, la precipitación y la temperatura máxima también influyen, aunque en menor medida, indicando que la disponibilidad de agua y los aumentos de calor contribuyen al establecimiento y desarrollo del banano, pero no son factores tan limitantes como las temperaturas mínimas.

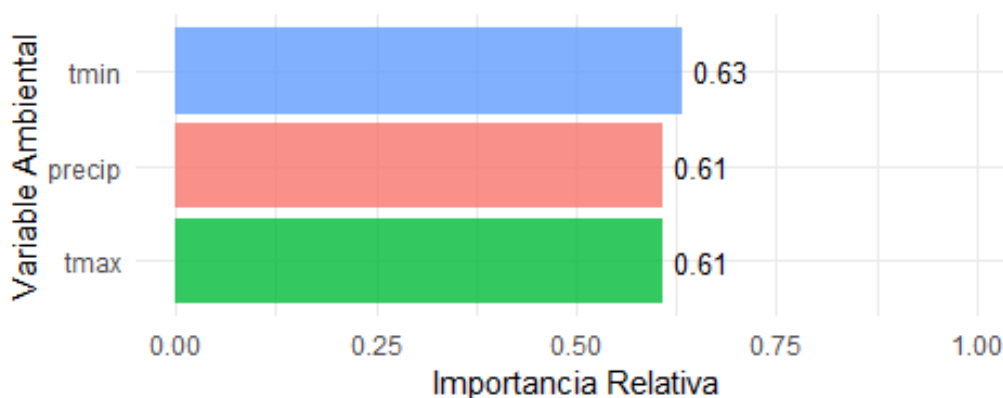


Figura 7.2: Importancia media de las variables ambientales en modelos para banano.

Cultivo de café

De acuerdo a la importancia media la variable ambiental más relevante es la temperatura mínima con un valor de 0.94, lo que refleja la alta sensibilidad del cultivo a bajas temperaturas, mientras que la temperatura máxima tiene un valor de importancia menor (0.25) y por su parte la precipitación esta tiene una contribución relativamente baja con un valor de 0.15, lo que sugiere que la disponibilidad de agua no es un factor crítico para determinar la distribución del cultivo en las áreas evaluadas (Figura 7.3). Esto indica que la variable más crítica para el establecimiento y manejo del cultivo es la temperatura mínima.

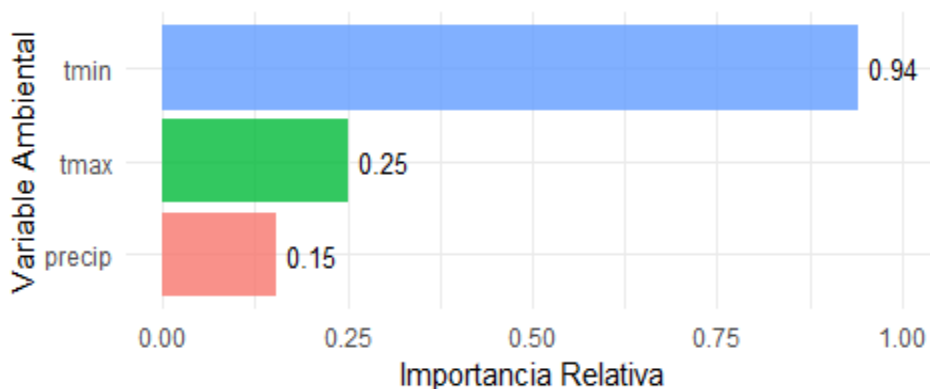


Figura 7.3: Importancia media de las variables ambientales en modelos para café.

Predicción de áreas de idoneidad en condiciones climáticas futuras

El análisis de aptitud climática para el cultivo de aguacate, de acuerdo con el modelo lineal generalizado (GLM), muestra una tendencia general a la reducción progresiva del área apta, especialmente en las categorías de mayor idoneidad (Figura 7.4). Aunque la categoría de 0–200 se mantiene relativamente estable, con incrementos marginales del 0.4% y 1.3% a 30 y 50 años respectivamente, las zonas con condiciones óptimas de altitud intermedia (600–1000) evidencian una disminución significativa, con pérdidas que alcanzan hasta el -21.1% en la categoría de 800–1000 en el horizonte de 50 años. Estos

resultados sugieren que el aumento de la temperatura y los posibles cambios en el régimen hídrico podrían desplazar la distribución del cultivo hacia zonas más bajas o restringir su rango altitudinal óptimo, afectando su sostenibilidad productiva a largo plazo (Tabla 7.4).

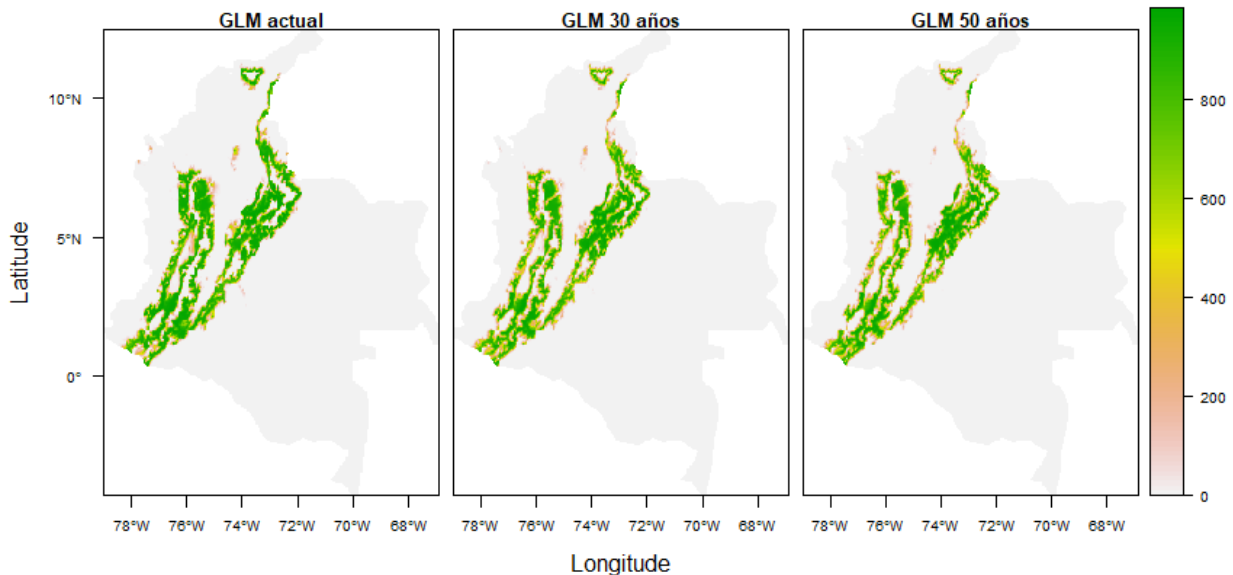


Figura 7.4: Proyecciones de distribución de aguacate bajo escenarios climáticos futuros.

Las proyecciones de distribución del cultivo de banano bajo escenarios climáticos futuros, obtenidas mediante el modelo MAXNET, evidencian una reducción progresiva en la idoneidad climática de las áreas actualmente productoras en Colombia. En el escenario actual, las zonas con mayor aptitud se concentran principalmente en la región del Caribe y el noroccidente del país, mientras que a 30 y 50 años se observa una notable contracción y desplazamiento de estas áreas hacia zonas más elevadas (Figura 7.5).

Así mismo, las zonas ubicadas en la categoría de 0–200, actualmente dominantes, presentan reducciones del -26.5% y -30.0%, mientras que las categorías intermedias (200–400) muestran un notable incremento de 272.9% y 358.2%, reflejando un posible desplazamiento del cultivo hacia regiones con condiciones menos restrictivas. Sin embargo, las áreas de mayor altitud (600–1000) experimentan una contracción significativa, con reducciones de hasta -85.2%, lo que evidencia la vulnerabilidad del banano frente al incremento térmico y la variabilidad climática. Estos resultados permiten inferir que en el futuro el cultivo podría concentrarse en zonas más bajas y húmedas, mientras las regiones de alta montaña perderían su potencial productivo (Tabla 7.4).

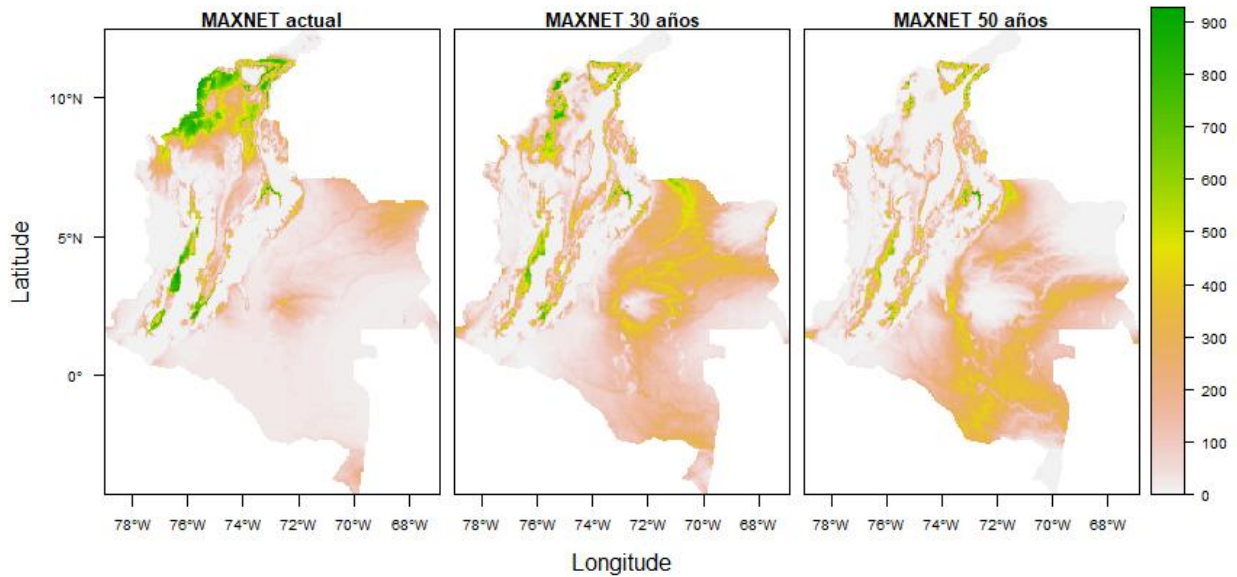


Figura 7.5: Proyecciones de distribución de banano bajo escenarios climáticos futuros.

La Figura 7.6 muestra las proyecciones de distribución del cultivo de café bajo escenarios climáticos futuros, en donde en el escenario actual, las zonas con mayor idoneidad climática se concentran principalmente en la región andina, abarcando el Eje Cafetero, partes de Antioquia, Tolima, Huila y Nariño. A 30 y 50 años, se observa que la distribución potencial del café se mantiene relativamente estable, con una leve expansión hacia áreas más elevadas.

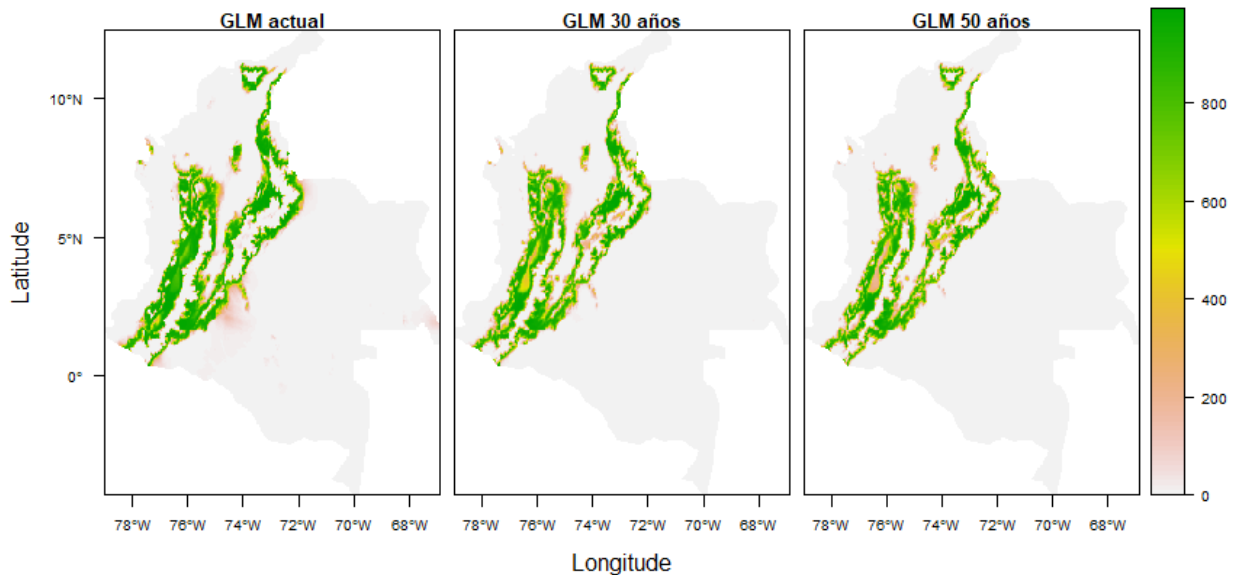


Figura 7.6: Proyecciones de distribución de café bajo escenarios climáticos futuros.

El modelo GLM indica que el cultivo de café mantendrá una alta estabilidad climática en las próximas décadas, con variaciones moderadas en las áreas aptas. La categoría de 0–200 presenta ligeros incrementos del 0.3% y 1.1% a 30 y 50 años, lo que indica que las zonas actualmente óptimas continuarán siendo adecuadas. No obstante, las áreas con mayor altitud (800–1000) muestran reducciones progresivas del -12.8% y -19.0%, lo cual podría asociarse al aumento de la temperatura mínima. En contraste, las zonas de altitud media (400–600) presentan leves fluctuaciones, evidenciando una posible relocalización del cultivo hacia pisos altitudinales más bajos (Tabla 7.4).

Tabla 7.4. Áreas por rango de idoneidad por cultivo.

| Cultivo | Categoría | Área (km ²) | | | Diferencia (%) | |
|--------------------------|-----------|-------------------------|---------|---------|----------------|---------|
| | | Actual | 30 años | 50 años | 30 años | 50 años |
| Aguacate GLM | 0–200 | 47,682 | 47,887 | 48,286 | 0.4% | 1.3% |
| | 200–400 | 909 | 850 | 848 | -6.5% | -6.7% |
| | 400–600 | 858 | 851 | 786 | -0.8% | -8.4% |
| | 600–800 | 1,369 | 1,228 | 1,221 | -10.3% | -10.8% |
| | 800–1000 | 4,445 | 3,830 | 3,505 | -13.8% | -21.1% |
| Banano MAXNET | 0–200 | 47,468 | 34,889 | 33,206 | -26.5% | -30.0% |
| | 200–400 | 4,239 | 15,808 | 19,422 | 272.9% | 358.2% |
| | 400–600 | 1,883 | 2,912 | 1,569 | 54.6% | -16.7% |
| | 600–800 | 1,262 | 876 | 388 | -30.6% | -69.3% |
| | 800–1000 | 411 | 161 | 61 | -60.8% | -85.2% |
| Café GLM | 0–200 | 46,592 | 46,754 | 47,113 | 0.3% | 1.1% |
| | 200–400 | 1,082 | 978 | 1,038 | -9.6% | -4.1% |
| | 400–600 | 984 | 1,048 | 912 | 6.5% | -7.3% |
| | 600–800 | 1,387 | 1,314 | 1,358 | -5.3% | -2.1% |
| | 800–1000 | 5,218 | 4,552 | 4,225 | -12.8% | -19.0% |

8. ANÁLISIS DE VIABILIDAD DE LAS ZONAS GENERADAS POR LOS MODELOS, CONSIDERANDO POSIBLES RESTRICCIONES NO CLIMÁTICAS.

8.1 Actividades

Acceso y recopilación de mapas de variables no climáticas.

Con el propósito de incorporar restricciones territoriales y criterios de viabilidad espacial en el análisis de distribución potencial de cultivos, se accedió a los mapas geoespaciales de variables no climáticas, específicamente las capas de vías de comunicación y áreas protegidas. Estos datos fueron descargados desde el portal de Datos Abiertos del Instituto Geográfico Agustín Codazzi (IGAC), entidad oficial encargada de la generación y custodia de información geográfica en Colombia. Las capas de vías permiten identificar zonas con accesibilidad logística para la producción agrícola, mientras que las áreas protegidas delimitan territorios con restricciones legales para el uso agropecuario. La integración de estas variables es fundamental para garantizar que las zonas identificadas como aptas por los modelos no solo sean climáticamente favorables, sino también viables desde el punto de vista territorial y normativo.

Evaluación espacial de la viabilidad territorial.

Una vez obtenidas las capas de variables no climáticas, se procedió a realizar una evaluación espacial de las ubicaciones generadas por los modelos de distribución. Este análisis consistió en el cruzamiento geoespacial entre los mapas de aptitud climática (*raster*) y las capas vectoriales de vías y áreas protegidas, utilizando herramientas de análisis en sistemas de información geográfica (SIG). El objetivo fue identificar aquellas zonas que, si bien presentan condiciones climáticas óptimas para el cultivo, podrían estar limitadas por factores de accesibilidad o restricciones legales. Esta etapa permite refinar los resultados del modelado, excluyendo áreas no viables y priorizando aquellas que cumplen simultáneamente con criterios climáticos, logísticos y normativos, fortaleciendo así la aplicabilidad de los resultados en contextos de planificación agropecuaria.

8.2 Resultados

Mapas de variables no climáticas

Se obtuvieron los mapas geoespaciales correspondientes a las variables no climáticas consideradas relevantes para la evaluación de la viabilidad territorial de los cultivos analizados. En particular, se descargaron las capas de infraestructura vial y áreas protegidas desde el portal de Datos Abiertos del Instituto Geográfico Agustín Codazzi (IGAC), en formatos compatibles con sistemas de información geográfica (SIG).

La capa de vías incluye la red nacional de carreteras, caminos secundarios y terciarios, lo cual permite evaluar la accesibilidad logística de las zonas agrícolas. Por su parte, la capa de áreas protegidas delimita los territorios con restricciones legales para el uso agropecuario, incluyendo parques nacionales naturales, reservas forestales y otras figuras de conservación ambiental. Ambas capas fueron re proyectadas al sistema de coordenadas estándar utilizado en el estudio y recortadas al límite político de Colombia,

garantizando su compatibilidad espacial con las demás variables del análisis (Figura 8.1). Estos insumos fueron integrados posteriormente en los modelos de distribución para refinar la identificación de zonas potencialmente aptas, considerando no solo criterios climáticos, sino también restricciones territoriales y de manejo.

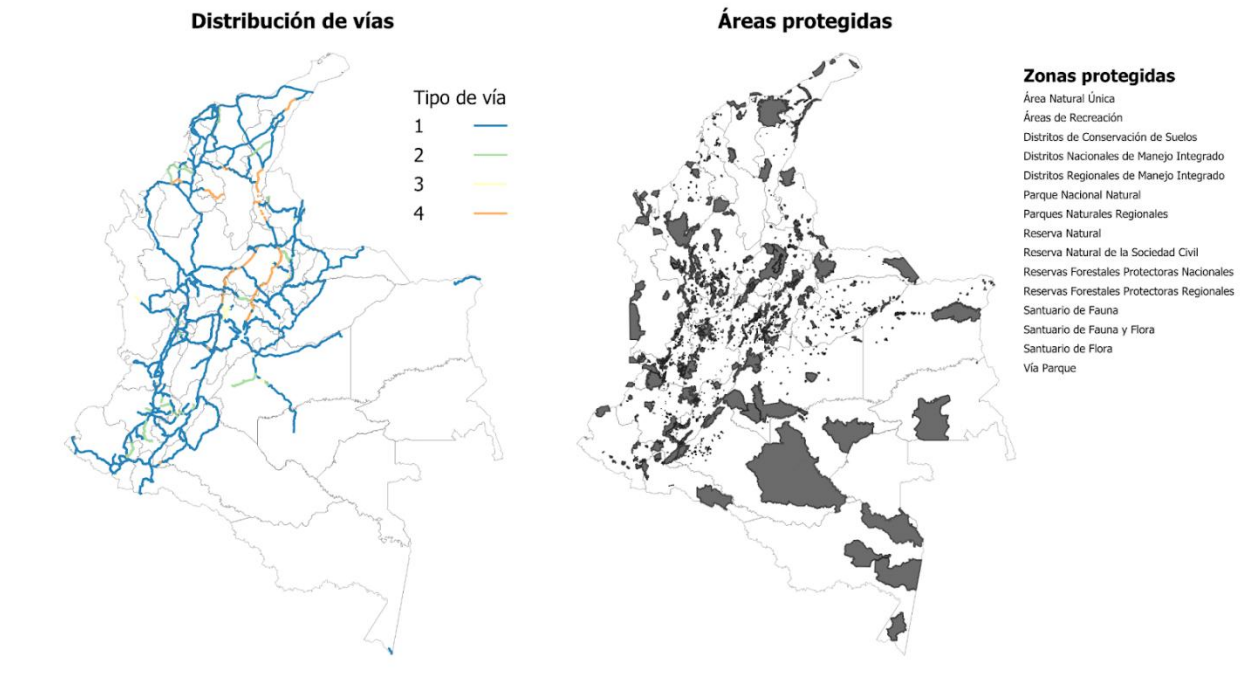


Figura 8.1: Mapa de vías de Colombia y sus áreas protegidas.

Evaluación de mapas

La incorporación de variables no climáticas en el análisis espacial permitió refinar significativamente las áreas inicialmente propuestas por los modelos de distribución climática para los cultivos de banano, aguacate y café (Figura 8.2). En particular, la superposición con las capas de áreas protegidas generó una reducción sustancial en la superficie potencialmente cultivable, al excluir territorios con restricciones legales para el uso agropecuario. Esta depuración espacial garantiza que las zonas identificadas como aptas no sólo respondan a condiciones climáticas favorables, sino que también cumplan con criterios de viabilidad normativa.

En el caso del banano, se observó que la región suroccidental de Colombia, especialmente en departamentos como Meta, Casanare y Vichada, presenta condiciones climáticas óptimas para el cultivo en los horizontes 2021–2040 y 2041–2060. Sin embargo, esta zona carece de una infraestructura vial adecuada para soportar operaciones logísticas de exportación, lo que limita su viabilidad operativa. Por el contrario, las áreas identificadas en el norte del país, particularmente en la región Caribe, además de mantener aptitud climática, cuentan con redes viales consolidadas y cercanía a puertos marítimos, lo que las convierte en zonas estratégicas para el desarrollo agroexportador del banano.

Para los cultivos de aguacate y café, las variaciones espaciales entre los escenarios climáticos evaluados fueron menos pronunciadas. Las zonas tradicionalmente aptas mantienen su idoneidad en los escenarios futuros, y la infraestructura vial existente en estas regiones, principalmente en el eje cafetero y zonas del altiplano cundiboyacense, puede ser aprovechada en el futuro para facilitar las operaciones productivas y comerciales. Esto sugiere una mayor estabilidad territorial para estos cultivos frente a las proyecciones climáticas, así como una mejor integración con la infraestructura actual.

En conjunto, esta evaluación espacial permitió priorizar zonas que cumplen simultáneamente con criterios climáticos, normativos y logísticos, fortaleciendo la aplicabilidad de los modelos en contextos reales de planificación agropecuaria y toma de decisiones.

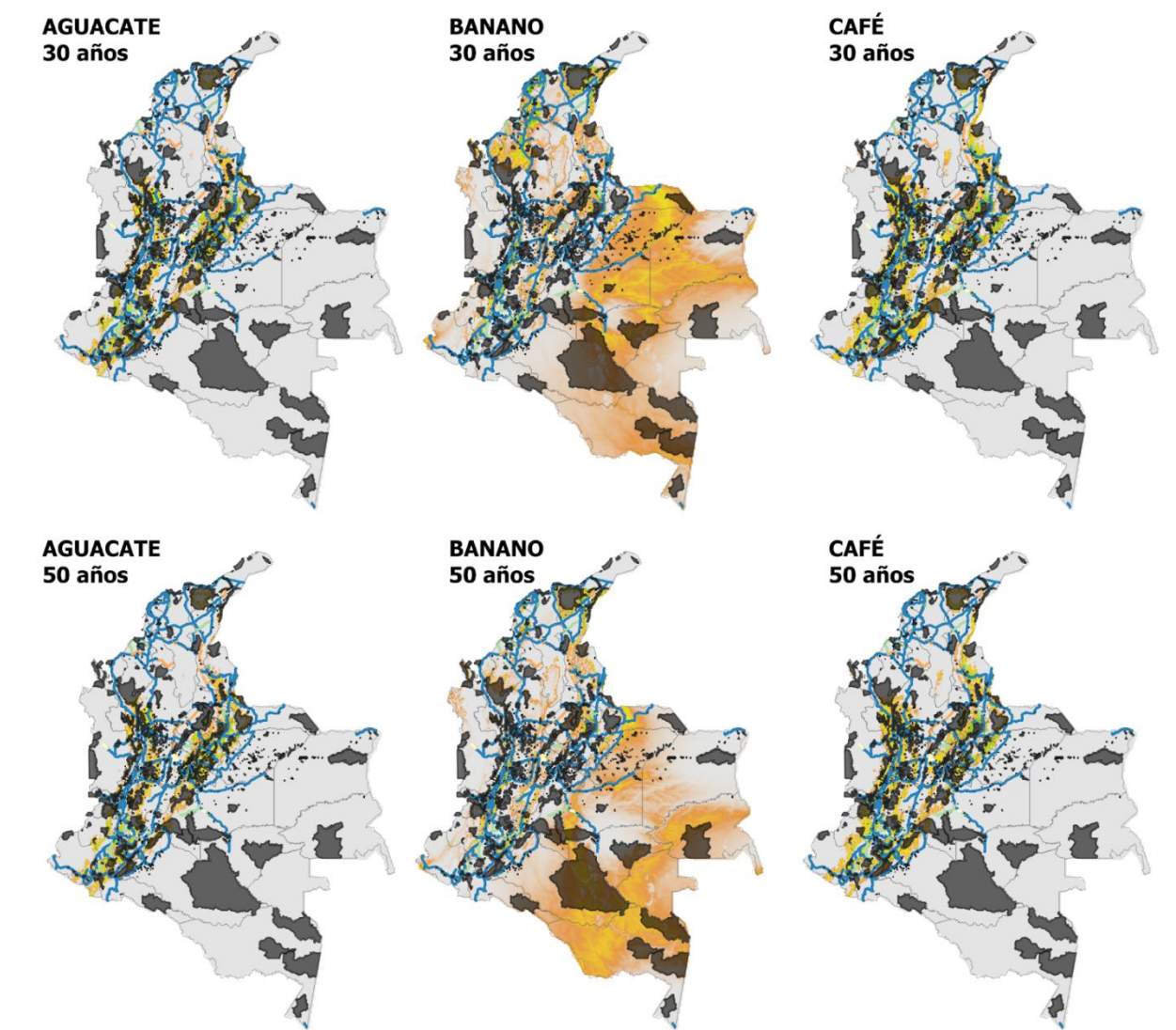


Figura 8.2: Áreas limitadas por la viabilidad operativa y legal de los cultivos.

9 CONCLUSIONES Y TRABAJOS FUTUROS

9.1 CONCLUSIONES

Las zonas de alta aptitud agrícola en Colombia presentan condiciones favorables para el cultivo de aguacate, café y banano, determinadas por sus rangos óptimos de temperatura y precipitación. El aguacate se adapta mejor en zonas interandinas con 10–30 °C y 1000–2000 mm de lluvia anual; el café en la región andina central con 18–22 °C y 1500–3000 mm; y el banano en regiones cálidas y húmedas del Caribe y noroccidente, con 26–34 °C y más de 3000 mm. Estas condiciones reflejan la diversidad agroclimática del país y su potencial para la planificación agrícola sostenible.

El modelo EC-Earth3-Veg bajo el escenario SSP2-4.5 evidencia una tendencia de incremento progresivo en la temperatura media anual y variaciones espaciales moderadas en la precipitación en Colombia para los periodos 2021–2040 y 2041–2060. Este comportamiento refleja un proceso de calentamiento sostenido, especialmente marcado en las zonas andinas, caribeñas y de piedemonte, donde las temperaturas máximas podrían superar los 34 °C hacia mediados de siglo.

El análisis de modelamiento de distribución potencial y aptitud climática para los cultivos de aguacate, banano y café permitió identificar patrones diferenciados de respuesta frente a las condiciones ambientales actuales y proyectadas a futuro. Los modelos seleccionados (GLM para aguacate y café, y MAXNET para banano) mostraron un buen desempeño en las métricas de ajuste y validación, destacándose por su equilibrio entre sensibilidad, especificidad y capacidad de generalización.

En cuanto a la importancia de las variables ambientales, la temperatura mínima se presentó como el principal factor determinante en la presencia de los tres cultivos, con valores de contribución altos en café (0.94) y aguacate (0.95), mientras que la temperatura máxima y la precipitación mostraron una influencia secundaria.

Los resultados en relación a la proyección y áreas de idoneidad por cultivo indican que el cambio climático modificará la distribución espacial y altitudinal de los cultivos a nivel de trópico en las próximas décadas, mostrando cambio en las zonas que actualmente se consideran óptimas para el establecimiento de estos cultivos.

La evaluación espacial de la viabilidad territorial permitió depurar las zonas propuestas por los modelos climáticos, excluyendo áreas protegidas y revelando limitaciones logísticas clave. Aunque el suroccidente del país muestra alta aptitud climática para el cultivo de banano, su infraestructura vial restringe su potencial exportador, a diferencia del norte, donde las vías existentes favorecen la operatividad. Para aguacate y café, las zonas aptas se mantienen estables y la infraestructura actual es funcional para futuros escenarios productivos.

9.2 TRABAJOS FUTUROS

Ampliar el conjunto de algoritmos de aprendizaje automático (*XGBoost*, *CatBoost*, *Support Vector Machines*, *Deep Neural Networks*) para comparar su desempeño frente a los modelos tradicionales (GLM,

GAM, MAXNET, RF), explorando técnicas de *ensemble learning* que integren los mejores predictores climáticos.

Crear aplicaciones web o paneles interactivos usando *google data studio* o *Power BI* que permitan a investigadores, productores y tomadores de decisiones explorar los resultados de los modelos, visualizar mapas de aptitud y evaluar escenarios de adaptación.

10 REFERENCIAS BIBLIOGRÁFICAS

- [1] IPCC, *Cambio Climático 2021: Bases físicas*. 2021.
- [2] World Meteorological Organization, *State of the Global Climate 2023*, no. 1347. 2024.
- [3] C. Lau, A. Jarvis, and J. Ramírez, “Agricultura colombiana: Adaptación al cambio climático. CIAT Políticas en Síntesis,” *Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.*, 2011. .
- [4] N. Saddique *et al.*, “A systematic review on farmers’ adaptation strategies in Pakistan toward climate change,” 2022, doi: 10.3390/atmos13081280.
- [5] J. Livsey *et al.*, “Do alternative irrigation strategies for rice cultivation decrease water footprints at the cost of long-term soil health?,” *Environ. Res. Lett*, 2019, doi: 10.1088/1748-9326/ab2108.
- [6] P. J. Kamdi, D. K. Swain, and S. P. Wani, “Efficient land water management practice and cropping system for increasing water and crop productivity in semi-arid tropics,” *Agron. J.*, vol. 112, no. 4, pp. 2552–2568, Jul. 2020, doi: 10.1002/AGJ2.20207.
- [7] C. Lesk, P. Rowhani, and N. Ramankutty, “Influence of extreme weather disasters on global crop production,” *Nat. 2016 5297584*, vol. 529, no. 7584, pp. 84–87, Jan. 2016, doi: 10.1038/nature16467.
- [8] H. Nguyen, M. Randall, and A. Lewis, “Factors Affecting Crop Prices in the Context of Climate Change—A Review,” *Agric.*, vol. 14, no. 1, pp. 1–17, 2024, doi: 10.3390/agriculture14010135.
- [9] L. C. Bernal Sánchez, “Artículo Original Estrategias de adaptación al cambio climático en el sector agrícola colombiano Climate change adaptation strategies in the colombian agricultural sector,” *Rev. Cuba. Meteorol.*, vol. 29, no. 3, pp. 2664–0880, 2023, Accessed: Nov. 03, 2024. [Online]. Available: <https://cu-id.com/2377/v29n3e03>.
- [10] G. Ubilla- Bravo, E. O. de la B. G. O. A. S. R. Méndez, and P. A. Maritano, “Desarrollo sustentable/sostenible – DS2 : diferencias y similitudes conceptuales e implicancias en el ordenamiento territorial,” *2do Semin. Exp. sobre la Planif. y el ordenamiento Territ. en Chile*, no. August, pp. 0–24, 2021, doi: 10.5281/zenodo.4937138.
- [11] O. Ovalle-Rivera, P. Läderach, C. Bunn, M. Obersteiner, and G. Schroth, “Projected shifts in *Coffea arabica* suitability among major global producing regions due to climate change,” *PLoS One*, vol. 10, no. 4, pp. 1–13, 2015, doi: 10.1371/journal.pone.0124155.
- [12] L. Perez *et al.*, “Determinants of vulnerability of bean growing households to climate variability in Colombia - CIAT - International Center for Tropical Agriculture Dataverse,” 2018. <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/E5UYJE> (accessed Nov. 03, 2024).
- [13] F. Ahmadi, S. Mehdizadeh, B. Mohammadi, Q. B. Pham, T. N. C. DOAN, and N. D. Vo, “Application of an artificial intelligence technique enhanced with intelligent water drops for monthly reference evapotranspiration estimation,” *Agric. Water Manag.*, vol. 244, p. 106622, Feb. 2021, doi: 10.1016/J.AGWAT.2020.106622.

- [14] Food and Agricultural Organisation of the United Nations FAO., "El estado mundial de la agricultura y la alimentación 2023," *El estado Mund. la Agric. y la Aliment. 2023*, Dec. 2023, doi: 10.4060/CC7724ES.
- [15] J. Elith and J. Franklin, "Species Distribution Modeling," *Encycl. Biodiversity, Third Ed. Vol. 1-7*, pp. V5-558-V5-572, Jan. 2024, doi: 10.1016/B978-0-12-822562-2.00379-0.
- [16] M. Pacifici *et al.*, "Assessing species vulnerability to climate change," *Nature Climate Change*, vol. 5, no. 3. 2015, doi: 10.1038/nclimate2448.
- [17] J. Elith and J. R. Leathwick, "Species distribution models: Ecological explanation and prediction across space and time," *Annu. Rev. Ecol. Evol. Syst.*, vol. 40, 2009, doi: 10.1146/annurev.ecolsys.110308.120159.
- [18] S. Itoh, A. Takeshige, A. Kasai, S. Kimura, J. Hayakawa, and N. Ohtsuchi, "Modeling Coastal Ecosystem Complexes," in *Treatise on Estuarine and Coastal Science (Second Edition)*, 2024.
- [19] L. Hannah, "Estimating Extinction Risk from Climate Change," in *Climate Change Biology*, 2015.
- [20] B. Boots and A. Getts, *Point Pattern Analysis*. 1988.
- [21] M. De la Cruz, "Métodos para analizar datos puntuales," no. May, 2015.
- [22] J. V. Zorraquino, "Una Breve Introducción a la Simulación de Patrones Puntuales Espaciales," 2021.
- [23] P. J. Diggle, "Statistical Analysis of Spatial and Spatio-Temporal Point Patterns -," 2015. https://books.google.com.ec/books?id=5FHSBQAAQBAJ&printsec=frontcover&hl=es&source=gbg_summary_r&cad=0#v=onepage&q&f=false (accessed Nov. 03, 2024).
- [24] L. Ramírez and W. Pineda, "Modelos de procesos puntuales para la identificación espacial de los sismos ocurridos en Colombia," pp. 1–16, 2019, [Online]. Available: <https://repository.usta.edu.co/bitstream/handle/11634/17617/2019lauraramirez.pdf?sequence=1&isAllowed=y>.
- [25] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, "Stochastic Geometry and Its Applications: Third Edition," *Stoch. Geom. Its Appl. Third Ed.*, pp. 1–555, Sep. 2013, doi: 10.1002/9781118658222.
- [26] C. A. Ríos-Muñoz, M. Vega-Flores, K. M. Vega-Flores, S. Hernández-Rubio, and D. V. Espinosa Martínez, "DEL CONCEPTO A SU APLICACIÓN: EL PROCESO DEL MODELADO DE NICHOS ECOLÓGICOS Y SUS ALGORITMOS," *Rev. Latinoam. Herpetol.*, vol. 4, no. 1, 2021, doi: 10.22201/fc.25942158e.2021.1.253.
- [27] R. P. Anderson *et al.*, "Ecological Niches and Geographic Distributions (MPB-49)," *Ecol. Niches Geogr. Distrib.*, Dec. 2011, doi: 10.1515/9781400840670.
- [28] A. Jiménez-Valverde and J. M. Lobo, "Threshold criteria for conversion of probability of species presence to either–or presence–absence," *Acta Oecologica*, vol. 31, no. 3, pp. 361–369, May 2007, doi: 10.1016/j.actao.2007.02.001.
- [29] I. A. Roy-García, C. Paredes-Manjarrez, J. Moreno-Palacios, R. Rivas-Ruiz, and A. A. Flores-Pulido, "Curvas ROC: características generales y su uso en la práctica clínica," *Rev. Med. Inst. Mex. Seguro Soc.*, vol. 61, no. Suppl 3, 2023.
- [30] D. Brzezinski, J. Stefanowski, D. Brzezinski, and J. Stefanowski, "Prequential AUC:

- properties of the area under the ROC curve for data streams with concept drift,” *Knowl Inf Syst*, vol. 52, pp. 531–562, 2017, doi: 10.1007/s10115-017-1022-8.
- [31] L. Curva Roc, J. A. Martínez Pérez, and P. S. Pérez Martin, “FORMACIÓN CONTINUADA-METODOLOGÍA Y TÉCNICAS,” *Med. Fam. Semer.*, vol. 49, p. 101821, 2023, doi: 10.1016/j.semerng.2022.101821.
- [32] D. Rizo, Miriela; Vuelta and A. Lorenzo, “Agricultura, Desarrollo Sostenible, Medioambiente, Saber Campesino Y Universidad,” *Cienc. en su PC*, vol. 1, no. 2, pp. 106–120, 2017, [Online]. Available: <http://www.redalyc.org/articulo.oa?id=181351615008>.
- [33] D. M. Vivanco Toala and R. G. Camero Berrones, “Determinación de un Modelo mediante Aprendizaje Automático para obtener la Idoneidad del Personal Naval Profesional,” *Cienc. Lat. Rev. Científica Multidiscip.*, vol. 9, no. 4, 2025, doi: 10.37811/cl_rcm.v9i4.19442.
- [34] S. De and E. De Posgrado, “UNIVERSIDAD DE COSTA RICA.”
- [35] G. Rau and Y. S. Shih, “Evaluation of Cohen’s kappa and other measures of inter-rater agreement for genre analysis and other nominal data,” *J. English Acad. Purp.*, vol. 53, no. 1, p. 101026, Sep. 2021, doi: 10.1016/j.jeap.2021.101026.
- [36] Morales, J y M. A. Martinez, “Modelos Estadísticos,” *Elche*, 2021.
- [37] S. y . A. R. McCue, T. Carruthers, E. Dawe., J., Liu, “Evaluation of generalized linear model assumptions using randomization,” 2008.
- [38] M. E. Toriz-Robles, N. y Ramírez-Guzmán, “Comparison of linear and nonlinear models to estimate the risk of soil contamination,” *Agrociencia*, 2019.
- [39] A. y T. C. E. Guisan, “Generalized linear and generalized additive models in studies of species distributions: setting the scene,” *Ecol. Modell.*, 2002.
- [40] T. FithianW., Hastie, “Finite-sample equivalence in statistical models for presence-only data,” *Ann. Appl. Stat.*, 1917.
- [41] M. J. Merow, C. y Smith, “A practical guide to MaxEnt for modeling species’ distributions: what it does, and why inputs and settings matter,” *Ecography (Cop.)*, 2013.
- [42] M. Phillips, S. J. y Dudík, “Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation,” *Ecography (Cop.)*, 2007.
- [43] L. Breiman, *Statistics, Random Forest*. .
- [44] T. Cuttler, D. R. y Edwards, “Random forests for classification in ecology,” *Ecology*, 2007.
- [45] A. B. Mushagalusa, C. A. Fandohan, “Random forest and spatial cross validation performance in predicting species abundance distributions,” *Environ. Syst. Res.*, 2024.
- [46] D. R. Roberts and V. Vahn, “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure,” *Ecography (Cop.)*, 2016.
- [47] M. Barbet-Massin and F. Jiguet, “Selecting pseudo-absences for species distribution models: how, where and how many?,” *Br. Ecol. Soc.*, 2012.
- [48] L. Song and L. Estes, “itsdm: Isolation forest-based presence-only species distribution modelling and explanation in r,” *Methods Ecol. Evol.*, vol. 14, no. 3, pp. 831–840, Mar. 2023, doi: 10.1111/2041-210X.14067.
- [49] S. Gay, “Modelado de nicho ecológico para caracterizar la expansión del género Lonchura en México,” *Cent. Investig. Científica Yucatán, A.C*, pp. 1–130, 2021, [Online]. Available: https://cicy.repositorioinstitucional.mx/jspui/bitstream/1003/1985/1/PCB_M_Tesis_202

- 1_Sebastian_Gay_Escalante.pdf.
- [50] M. Flores-Tolentino, E. Ortiz, and J. L. Villaseñor, “Ecological niche models as a tool for estimating the distribution of plant communities,” *Revista Mexicana de Biodiversidad*, vol. 90, no. 3. 2019, doi: 10.22201/ib.20078706e.2019.90.2829.
- [51] H. Qiao *et al.*, “An evaluation of transferability of ecological niche models,” *Ecography (Cop.)*, vol. 42, no. 3, 2019, doi: 10.1111/ecog.03986.
- [52] B. Martínez, F. Arenas, A. Trilla, R. M. Viejo, and F. Carreño, “Combining physiological threshold knowledge to species distribution models is key to improving forecasts of the future niche for macroalgae,” *Glob. Chang. Biol.*, vol. 21, no. 4, 2015, doi: 10.1111/gcb.12655.
- [53] S. Selim, D. Koc-San, C. Selim, and B. T. San, “Site selection for avocado cultivation using GIS and multi-criteria decision analyses: Case study of Antalya, Turkey,” *Comput. Electron. Agric.*, vol. 154, pp. 450–459, Nov. 2018, doi: 10.1016/J.COMPAG.2018.09.038.
- [54] G. H. de Olanda Souza, L. E. de Oliveira Aparecido, R. F. de Lima, G. B. Torsoni, A. G. Chiquitto, and J. R. C. de Moraes, “Agroclimatic zoning for bananas under climate change in Brazil,” *J. Sci. Food Agric.*, vol. 102, no. 14, pp. 6511–6529, Nov. 2022, doi: 10.1002/JSFA.12018;WEBSITE:WEBSITE:SCIJOURNALS;ISSUE:ISSUE:DOI.
- [55] J. Kath, V. Mittahalli Byrareddy, S. Mushtaq, A. Craparo, and M. Porcel, “Temperature and rainfall impacts on robusta coffee bean characteristics,” *Clim. Risk Manag.*, vol. 32, p. 100281, Jan. 2021, doi: 10.1016/J.CRM.2021.100281.
- [56] “Future climate data — WorldClim 1 documentation.”
https://www.worldclim.org/data/cmip6/cmip6climate.html#google_vignette (accessed Nov. 02, 2025).
- [57] M. S. Reboita, G. Willian de Souza Ferreira, J. Gabriel Martins Ribeiro, and S. Ali, “Assessment of precipitation and near-surface temperature simulation by CMIP6 models in South America,” *Environ. Res. Clim.*, vol. 3, no. 2, p. 025011, Apr. 2024, doi: 10.1088/2752-5295/AD3FDB.
- [58] O. Boucher *et al.*, “Presentation and Evaluation of the IPSL-CM6A-LR Climate Model,” *J. Adv. Model. Earth Syst.*, vol. 12, no. 7, p. e2019MS002010, Jul. 2020, doi: 10.1029/2019MS002010;WGROU:STRING:PUBLICATION.
- [59] R. Döscher *et al.*, “The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6,” *Geosci. Model Dev.*, vol. 15, no. 7, pp. 2973–3020, Apr. 2022, doi: 10.5194/GMD-15-2973-2022.
- [60] E. Berntell and Q. Zhang, “Mid-Holocene West African monsoon rainfall enhanced in EC-Earth simulation with dynamic vegetation feedback,” *Clim. Dyn.* 2024 627, vol. 62, no. 7, pp. 7001–7017, May 2024, doi: 10.1007/S00382-024-07262-7.
- [61] L. H. Caicedo-López, A. L. V. Aranda, D. Sáenz de la O, C. E. Z. Gómez, E. E. Márquez, and H. R. Zepeda, “Elicidores: implicaciones bioéticas para la agricultura y la salud humana,” *Rev. Bioética*, vol. 29, no. 1, 2021, doi: 10.1590/1983-80422021291448.
- [62] R. Pielke, M. G. Burgess, and J. Ritchie, “Plausible 2005–2050 emissions scenarios project between 2 °C and 3 °C of warming by 2100,” *Environ. Res. Lett.*, vol. 17, no. 2, p. 024027, Feb. 2022, doi: 10.1088/1748-9326/AC4EBF.

- [63] M. Meinshausen *et al.*, “A perspective on the next generation of Earth system model scenarios: Towards representative emission pathways (REPs),” *Geosci. Model Dev.*, vol. 17, no. 11, pp. 4533–4559, Jun. 2024, doi: 10.5194/GMD-17-4533-2024.
- [64] M. Meinshausen *et al.*, “The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500,” *Geosci. Model Dev.*, vol. 13, no. 8, pp. 3571–3605, Aug. 2020, doi: 10.5194/GMD-13-3571-2020.
- [65] M. R. Alizadeh, J. Adamowski, and A. Inam, “Integrated assessment of localized SSP–RCP narratives for climate change adaptation in coupled human-water systems,” *Sci. Total Environ.*, vol. 823, p. 153660, Jun. 2022, doi: 10.1016/J.SCITOTENV.2022.153660.
- [66] M. Farooq *et al.*, “Assessing Future Agricultural Vulnerability in Kashmir Valley: Mid- and Late-Century Projections Using SSP Scenarios,” *Sustain. 2024, Vol. 16, Page 7691*, vol. 16, no. 17, p. 7691, Sep. 2024, doi: 10.3390/SU16177691.
- [67] S. E. El-Mahroug *et al.*, “Predictive Modeling of Climate-Driven Crop Yield Variability Using DSSAT Towards Sustainable Agriculture,” *AgriEngineering 2025, Vol. 7, Page 156*, vol. 7, no. 5, p. 156, May 2025, doi: 10.3390/AGRIENGINEERING7050156.