

## Artículo

# Mapeando oportunidades: análisis geoespacial del balsa en Colombia para la planificación forestal

Valentina Cardona <sup>1</sup><sup>1</sup> Universidad Nacional de Colombia; vcardonav@unal.edu.co

**Resumen:** *Ochroma pyramidale* (balsa) es una especie de rápido crecimiento y alto valor comercial, con potencial para el desarrollo forestal en Colombia. Este estudio tiene como objetivo analizar los patrones espaciales de distribución del balsa y comprender las variables ambientales asociadas a su presencia, como base para identificar zonas aptas para nuevas plantaciones. Se utilizaron 8.380 registros de ocurrencia provenientes de GBIF, variables bioclimáticas de WorldClim y modelos espaciales implementados en R. Se aplicaron análisis de patrón de puntos (Kernel, Ripley, DBSCAN), regresión logística (GLM) y modelos bayesianos espaciales (CARBayes e INLA), considerando dependencia y heterogeneidad espacial. Los resultados muestran un patrón de distribución agrupado, con alta concentración de registros en los valles interandinos y el piedemonte llanero. Las variables más influyentes fueron: precipitación del mes más seco, isothermalidad, estacionalidad de la temperatura y estacionalidad de la precipitación. El modelo INLA (BYM + heterogeneidad) presentó el mejor ajuste (DIC y WAIC). Se concluye que la distribución del balsa responde a factores climáticos con marcada dependencia espacial. Los modelos geoespaciales son herramientas clave para entender esta dinámica y apoyar la planificación de plantaciones forestales estratégicas en Colombia.

**Palabras claves:** Geoespacial, Biogeografía, Distribución, *Ochroma pyramidale*, Balsa, Temperatura, Precipitación, Colombia

## 1. Introducción

El balsa (*Ochroma pyramidale* (Cav. ex Lam.) Urb.) pertenece a la familia Malvaceae. Se han reportado dos variedades para esta especie: *Ochroma pyramidale* var. *bicolor* (Rowlee) Brizicky y *Ochroma pyramidale* var. *concolor* (Rowlee) R.E. Schult (Missouri Botanical Garden, 2025). Está ampliamente distribuida desde el sur de México, Centroamérica, el norte de los Andes, las Antillas y Brasil (Figura 1) (Elbert L. Little & Frank H. Wadsworth, 1964; Francis et al., 2000; «Listado de árboles de El Salvador», 2009)

Es una especie típica de climas cálidos y húmedos, que crece principalmente en zonas bajas (<1.000 m s.n.m.). No tolera heladas y requiere suelos ricos en nutrientes y con buen drenaje. Se caracteriza por su rápido crecimiento y la producción de madera de baja densidad, la cual se emplea en la elaboración de juguetes, artesanías y materiales aislantes. Además, sus frutos contienen una fibra algodonosa utilizada en la fabricación de almohadas. Es comúnmente conocida como “algodón”, “balsa” o “corcho”, nombres relacionados con sus características y usos tradicionales.

El principal producto de comercialización es la madera; sin embargo, algunas comunidades de la Amazonía ecuatoriana también le atribuyen usos medicinales (Ramos Corrales, 2016). Debido a su alta tasa de crecimiento, *O. pyramidale* ha sido introducida con

Academic Editor: Firstname Last-name

Received: date

Revised: date

Accepted: date

Published: date

**Citation:** To be added by editorial staff during production.

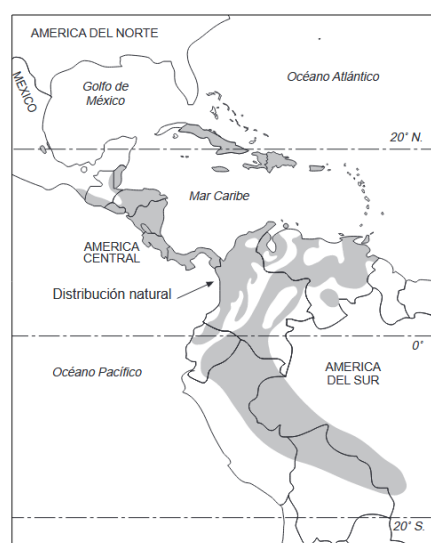
**Copyright:** © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

éxito en regiones fuera de su distribución natural, incluyendo países de Asia y Oceanía como India, Sri Lanka, Malasia, Vietnam, Borneo, Fiyi, Filipinas, Islas Salomón y Papúa Nueva Guinea (Francis et al., 2000)

En el contexto latinoamericano, Ecuador es actualmente el principal productor y exportador de balsa, siendo el único competidor regional en el mercado internacional. Por sus características (rápido crecimiento, turno de cosecha de 4–6 años, productos maderables y no maderables, y demanda comercial), el balsa representa una especie con alto potencial para el desarrollo forestal en Colombia, donde existen incentivos gubernamentales como el Certificado de Incentivo Forestal (CIF) para el establecimiento de nuevas plantaciones (Boletín estadístico forestal, 2024)

Ante el reto de identificar zonas óptimas para el establecimiento de plantaciones de balsa en Colombia, el objetivo de este estudio es: (1) Analizar los patrones espaciales de la distribución actual de *O. pyramidale* en el país, y (2) Comprender los procesos (factores ambientales) que explican dicha distribución, con el fin de establecer una línea base para proponer áreas con mayor aptitud para nuevas plantaciones.

Para ello, se integran teorías y herramientas de estadística geoespacial, partiendo de la premisa de que los datos ambientales presentan dependencia y heterogeneidad espacial. Se emplean registros de ocurrencia disponibles en plataformas abiertas (GBIF), datos climáticos satelitales y modelos espaciales avanzados. Este proyecto busca ir más allá de la simple identificación de presencia, proponiendo una aproximación orientada a comprender las causas que explican la presencia o ausencia del balsa en el territorio colombiano.



**Figura 1. Distribución natural de *Ochroma pyramidale* (balsa).** El sombreado gris representa el rango geográfico conocido de la especie en el Neotrópico. (Francis et al., 2000)

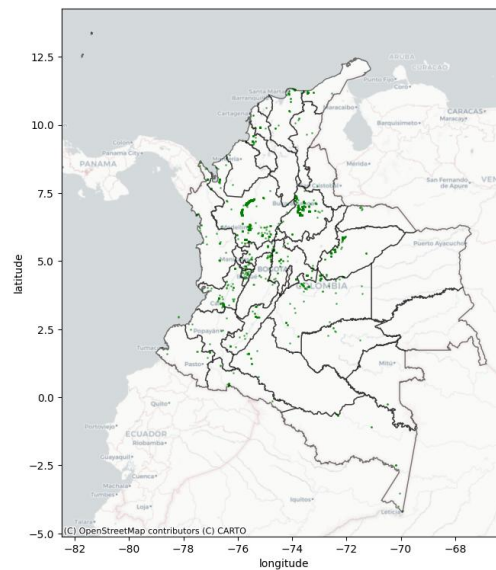
## 2. Materiales y Métodos

### 2.1. Datos

#### 2.1.1. Datos de ocurrencia de la especie.

Los datos de ocurrencia de *Ochroma pyramidale* se obtuvieron de los registros disponibles en el Global Biodiversity Information Facility (GBIF) (Figura 2). Cada registro se interpretó como un evento ocurrido en una ubicación específica, pero que podría haberse producido en un conjunto más amplio de sitios. Para la descarga de los datos se aplicaron los siguientes filtros: "Country is Colombia", "HasCoordinate is true",

"HasGeospatialIssue is false", y "TaxonKey is Ochroma pyramidale (Cav.) Urb.". Referencia de descarga: GBIF.org (25 mayo 2025). GBIF Occurrence Download <https://doi.org/10.15468/dl.3g5j8x>.



**Figura 2. Registros geográficos de *Ochroma pyramidale* (balsa) en Colombia.**

Se muestran 8.380 ocurrencias reportadas en la base de datos GBIF. Los puntos indican la ubicación de los registros, mientras que los límites departamentales de Colombia se encuentran resaltados.

### 2.1.2. Variables bioclimáticas

Se utilizaron datos climáticos de WorldClim versión 2.1 correspondientes al período 1970–2000, publicados en enero de 2020. Estos datos tienen una resolución espacial de 2.5 minutos ( $\sim 21.4 \text{ km}^2$ ).

A partir de los polígonos de los 1.121 municipios de Colombia, se calculó el valor promedio de cada variable climática con base en los valores de los píxeles contenidos en cada municipio. Se consideraron las siguientes diez variables:

- Elevación media
- Temperatura media anual
- Precipitación media anual
- Rango medio diurno (promedio mensual de  $(T_{\text{max}} - T_{\text{min}})$ )
- Precipitación del mes más lluvioso
- Precipitación del mes más seco
- Isotermalidad (rango medio diurno / rango anual)
- Estacionalidad de la temperatura (desviación estándar  $\times 100$ )
- Rango anual de temperatura ( $T_{\text{max}}$  del mes más cálido -  $T_{\text{min}}$  del mes más frío)
- Estacionalidad de la precipitación (coeficiente de variación)

### 2.1.3. Capas vectoriales

Se utilizó información vectorial en formato .shp de la división político-administrativa de Colombia, suministrada por el geoportal del DANE. El archivo, en formato poligonal, contiene los 1.121 municipios y sus respectivos departamentos.

## 2.2. Análisis de patrón de puntos

Los registros de balsa se analizaron como ocurrencias de un proceso estocástico, en el que cada punto representa la realización de un evento espacial. Desde esta perspectiva,

el dominio no es fijo, ya que la ubicación de cada punto es el resultado de un proceso aleatorio. Este enfoque permite identificar patrones subyacentes como agregación, dispersión o distribución aleatoria.

Se realizaron los siguientes análisis:

- Análisis de dispersión
- Estimación de densidad de Kernel
- Función de Ripley ( $G(d)$  y  $F(d)$ ): Para reducir la complejidad computacional, se realizó un submuestreo aleatorio del 50 % de los registros. El rango de distancias evaluado fue de 0 a 400 km.
- Agrupamiento DBSCAN: Se utilizó el algoritmo DBSCAN (Density-Based Spatial Clustering of Applications with Noise), que asigna un mismo identificador a puntos espacialmente agrupados. Se definieron los siguientes parámetros: población mínima viable de 50 individuos y distancia máxima de 50 km para pertenecer al mismo clúster.

### 2.3. Regresión sin espacialidad (GLM – Regresión logística)

Se construyó una variable de conteo a nivel municipal, correspondiente al número de reportes de balso en cada municipio. Dado que se observó sobre dispersión en los datos, posiblemente asociada al sobremuestreo desigual entre regiones; a partir de la variable conteo, se creó una nueva variable binaria: 1 si el municipio presenta al menos un registro (presencia), y 0 si no presenta registros (ausencia).

Se ajustó un modelo de regresión logística utilizando la función `glm()` en R, con el objetivo de identificar las variables más influyentes en la presencia o ausencia del balso a nivel municipal. Se evaluó el desempeño del modelo mediante la curva ROC y el valor del AUC.

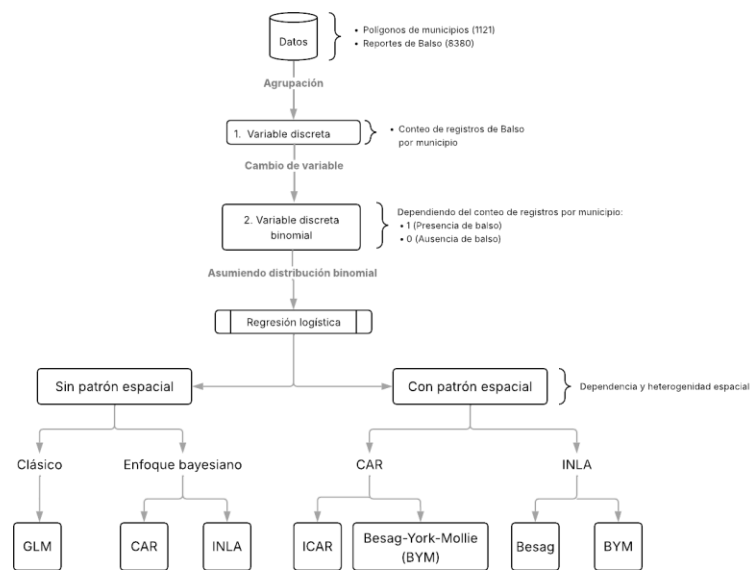
### 2.4. Análisis de patrones areales

Se crea matriz de pesos basada en contigüidad tipo Queen (considera dos observaciones como vecinas si comparten al menos un punto de su límite). Se analizaron patrones areales de la variable respuesta y las variables predictoras significativas según el modelo de regresión GLM (regresión logística). Principalmente, visualización de las variables usando mapas de cloropletas (Las clases están organizadas por cuantiles. Evita problema de clases dispersas), posteriormente usando la matriz de pesos se evalúa la autocorrelación espacial (Gráfico e índice de Moran, y gráfico de LISA)

### 2.5. Regresión espacial

Con base en la intuición obtenida sobre los patrones espaciales observados, se ajustaron diferentes modelos de regresión clásica y bayesiana, incorporando estructura espacial y heterogeneidad.

Se emplearon los paquetes **CARBayes** e **INLA** en R. Los modelos se evaluaron utilizando los criterios **DIC** (Deviance Information Criterion) y **WAIC** (Watanabe-Akaike Information Criterion), con el fin de seleccionar el modelo con mejor ajuste (Figura 3).

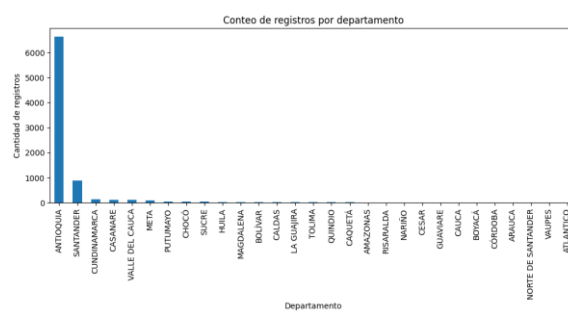


**Figura 3. Esquema del flujo de análisis de datos aplicado al estudio de *Ochroma pyramidale*.** Se representa el proceso completo, desde la recopilación y procesamiento de datos hasta la aplicación de modelos de regresión clásicos y espaciales. Elaboración propia.

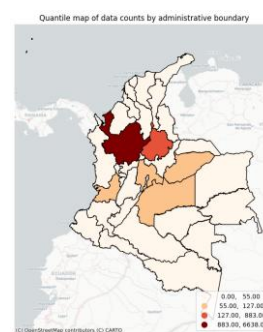
### 3. Resultados

#### 3.1. Análisis de patrón de puntos

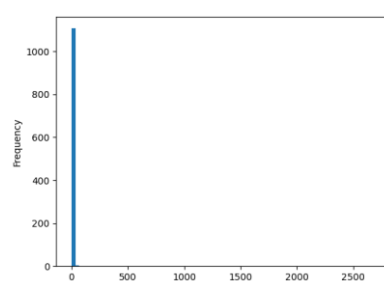
La mayoría de los municipios reportan un bajo número de observaciones de *Ochroma pyramidale* (balso), mientras que solo unos pocos presentan conteos elevados, principalmente en los departamentos de Antioquia y Santander. La distribución presenta un fuerte sesgo a la derecha: 897 municipios (equivalente al 80 % de las unidades) no reportan registros de balso (Figura 4).



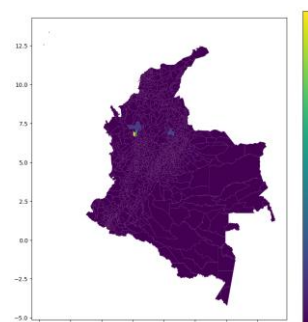
(a)



(b)



(c)



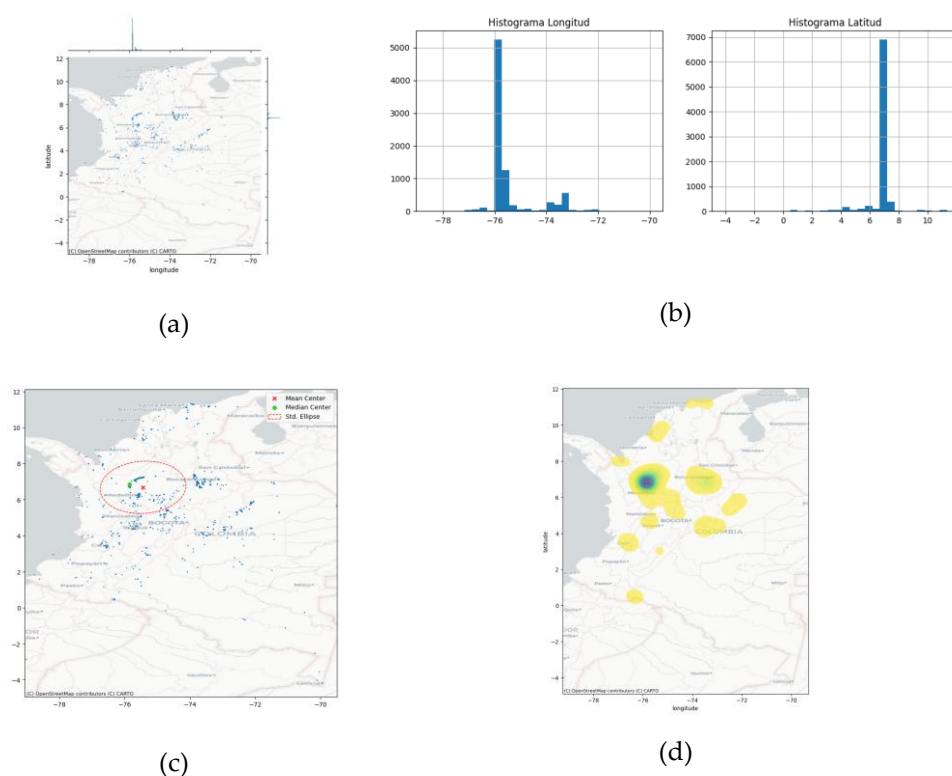
(d)

**Figura 4. Distribución espacial de los registros de *Ochroma pyramidale* por unidades administrativas.** (a–b) Conteo de registros por departamento: Antioquia presenta el mayor número de ocurrencias; (c–d) Conteo de registros por municipio: la mayoría de los municipios no reportan registros de balso o presentan cantidades muy bajas.

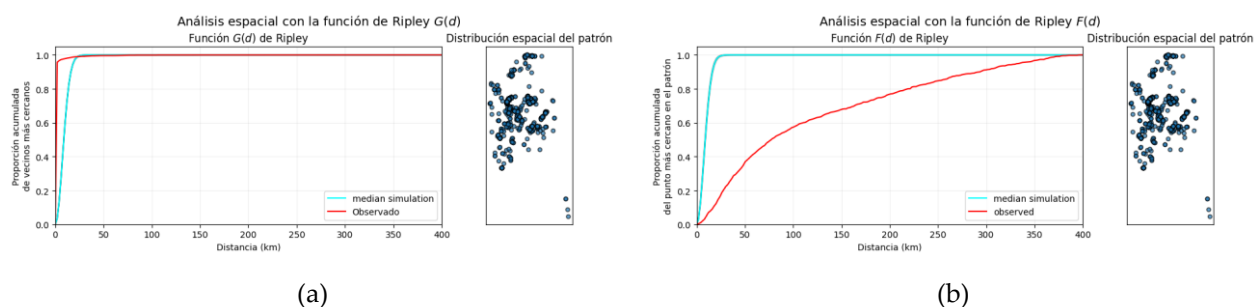
La especie se encuentra ampliamente distribuida en Colombia, con mayor concentración entre los 6° y 8° de latitud norte, y alrededor de los 76° de longitud oeste (Figura 5). Predomina en zonas de baja altitud (0–1.200 m s.n.m.), con temperaturas medias anuales de  $24 \pm 3$  °C, temperaturas máximas en el mes más cálido de  $30 \pm 3$  °C, y temperaturas mínimas en el mes más frío de  $18 \pm 3$  °C. La precipitación media anual es de aproximadamente  $2.400 \pm 800$  mm, con valores cercanos a 400 mm en los meses más lluviosos y hasta 100 mm en los más secos.

Una medida común en la centrografía es la distancia estándar, que representa la dispersión media desde el centroide de la nube de puntos. En este estudio, la distancia estándar fue de 158.71 km (Figura 5b).

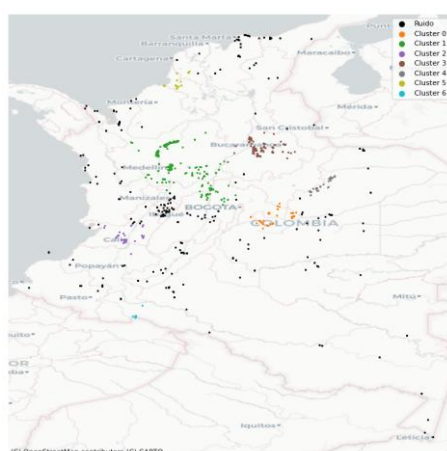
A distancias cortas, los puntos tienden a agruparse, mientras que a distancias mayores se dispersan, generando vacíos espaciales o "gaps" (Figura 6). El análisis de agrupamiento espacial mediante DBSCAN identificó siete clústeres, destacándose los asociados con Antioquia (clúster verde) y Santander (clúster café) como los más extensos (Figura 7).



**Figura 5. Análisis de patrón de puntos.** (a–b) Gráfico de dispersión de los registros (longitud vs. latitud) con histogramas marginales para cada variable; permite visualizar la distribución espacial general.; (c) Centrografía y distancia estándar desde el centroide de los puntos. (d) Estimación de densidad de Kernel; el mapa resalta las áreas con mayor concentración de registros.



**Figura 6. Análisis de la distribución espacial mediante funciones de Ripley.** (a) Función  $G(d)$  de Ripley: evalúa la distribución de las distancias al vecino más cercano, útil para identificar agrupamientos. (b) Función  $F(d)$  de Ripley: analiza la presencia de vacíos espaciales ("gaps") en el patrón de puntos.



**Figura 7. Agrupamiento espacial de los registros de *Ochroma pyramidale* mediante DBSCAN.** Se identificaron siete clústeres utilizando el algoritmo DBSCAN, con una población mínima viable de 50 individuos y una distancia máxima de 50 km para considerar puntos dentro de un mismo grupo.

### 3.2. Regresión sin espacialidad (GLM – Regresión logística)

Se ajustó un modelo de regresión logística clásica utilizando la función `glm` de R para evaluar la probabilidad de ocurrencia de *Ochroma pyramidale* en los 1.121 municipios de Colombia (Tabla 1). Los resultados indican que las variables climáticas significativamente asociadas con la presencia de la especie son: la precipitación del mes más seco, la isoterma, la estacionalidad de la temperatura y la estacionalidad de la precipitación.

El mapa de predicción generado a partir del modelo muestra una mayor probabilidad de ocurrencia en la región andina, en el piedemonte oriental y en la Región Caribe Baja. La curva ROC mostró un buen poder discriminativo entre presencia y ausencia, con un valor de AUC (Área Bajo la Curva) superior a 0.80, lo que indica una capacidad predictiva adecuada (Figura 8).

**Tabla 1.** Resultados de la regresión logística (Distribución binomial)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-1.66451	0.09648	-17.252	< 2e-16 ***
elev_mean	0.99742	0.84167	1.185	0.235998
temperatura_media_anual_mean	1.62849	0.87268	1.866	0.062032 .
precipitacion_anual_mean	0.13637	0.53561	0.255	0.799030



---

rango_medio_diurno_mean	-1.98191	1.14619	-1.729	0.083785	.
precipitacion_mes_mas_lluvioso_mean	0.24056	0.42101	0.571	0.567739	
precipitacion_mes_mas_seco_mean	-0.77394	0.24692	-3.134	0.001723	**
isotermalidad_mean	1.86700	0.74734	2.498	0.012483	*
estacionalidad_de_la_temperatura_mean	0.51218	0.15975	3.206	0.001345	**
rango_anual_de_temperatura_mean	2.24701	1.48055	1.518	0.129094	
estacionalidad_de_la_precipitacion_mean	-0.71677	0.21283	-3.368	0.000758	***

---

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

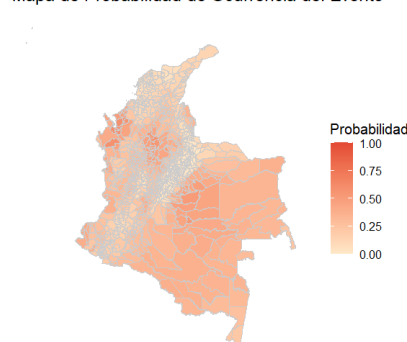
Null deviance: 1121.3 on 1120 degrees of freedom

Residual deviance: 1003.7 on 1110 degrees of freedom

AIC: 1025.7

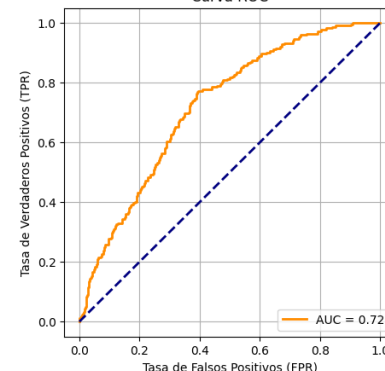
---

Mapa de Probabilidad de Ocurrencia del Evento



(a)

Curva ROC



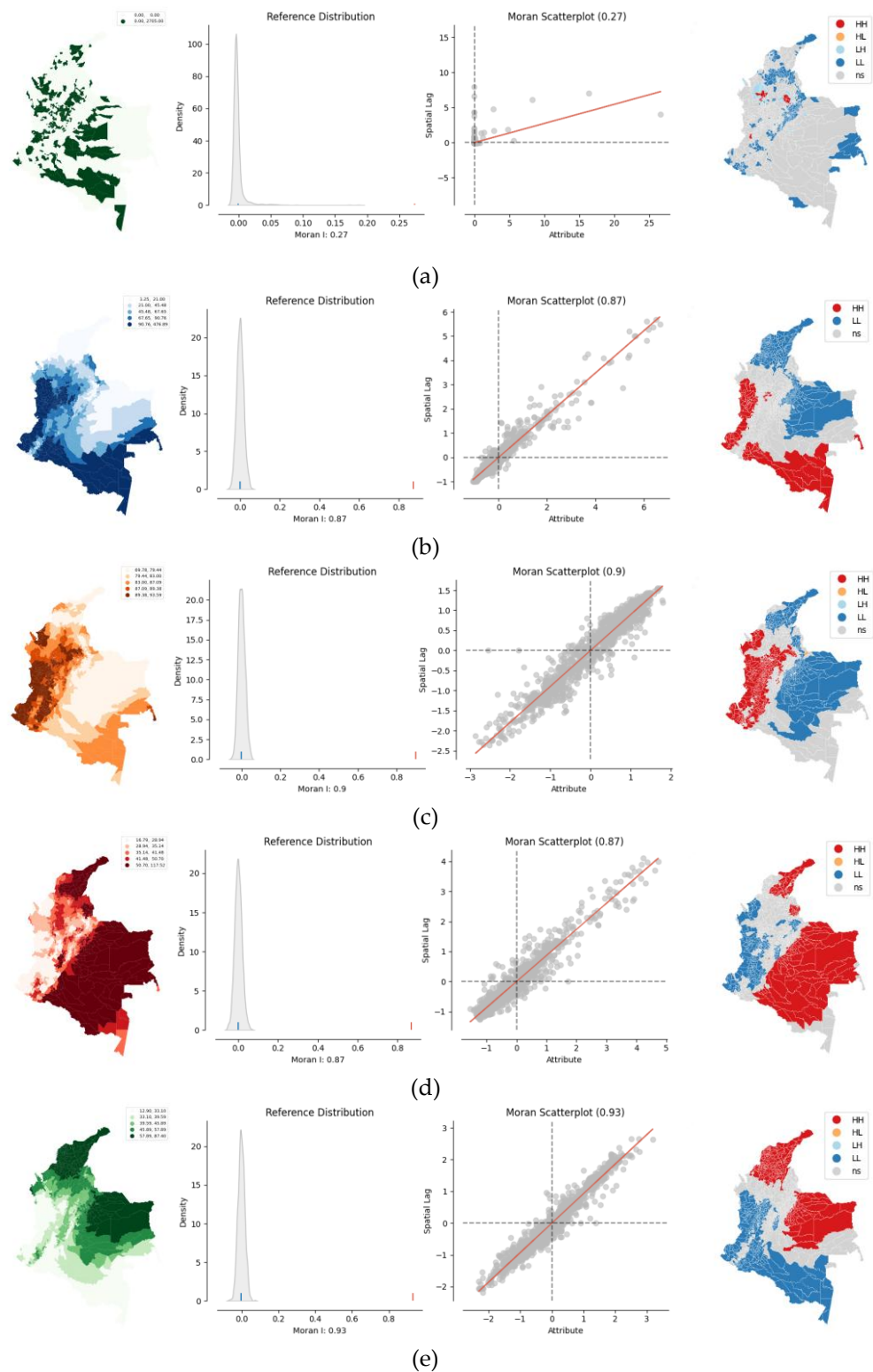
(b)

**Figura 8. Resultados del modelo de regresión logística para la presencia de *Ochroma pyramidale*.** Mapa de probabilidades de ocurrencia del evento según el modelo GLM (Tabla 1) y curva ROC utilizada para evaluar el desempeño predictivo del modelo.

### 3.3. Análisis de patrones areales

La variable respuesta binaria es la que menor patrón areal presenta, posiblemente este gráfico se confunda con la naturaleza del dato. Por el contrario, las variables climáticas (variables continuas) se visualizan patrones de dependencia espacial, con altos índices de Moran y evidente agrupación en el gráfico de LISA. En la región de la macrocuenca del Caribe y la Orinoquia se congregan las zonas con menores precipitaciones en los meses más secos e isotermalidad, y mayor estacionalidad de la temperatura y la precipitación. Por el contrario, la región de la macrocuenca del Pacífico y la Amazonía corresponde a las zonas como mayores precipitaciones en los meses más secos, mayor isotermalidad, y menor estacionalidad de la temperatura y precipitación. La región asociada a la macrocuenca de la región Andina no agrupa patrones tan claramente diferenciados como en las anteriores macrocuencas.





**Figura 9.** Patrones areales. (a) Variable binaria 1: Presencia de balso, 0: Ausencia de balso; (b) Precipitación del mes más seco; (c) Isotermalidad; (d) Estacionalidad de la temperatura; (e) Estacionalidad de la precipitación

### 3.4. Regresiones espaciales

**Tabla 2. Resultados de los modelos de regresión logística con y sin efectos espaciales para la ocurrencia de *Ochroma pyramidale* en Colombia.**

Se comparan modelos clásicos (GLM) y espaciales (CARBayes e INLA), considerando diferentes estructuras de dependencia espacial y heterogeneidad. Se incluyen los criterios de ajuste (DIC, WAIC), el número efectivo de parámetros (p.d), y las variables climáticas significativas identificadas en cada modelo.

						Marginal										
Modelo	Dependencia	Heterogeneidad	DIC	WAIC	p.d	log-lik.	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
GLM (Modelo logístico clásico)	No	No	1025.3 (AIC)									*	*	**		***
CARBayes (sin efectos aleatorios)	No	No	1024.1		10.34	-512.08				*		*	*	*		*
CARBayes ICAR	Sí Leroux (rho=1)	No	912.36		239.51	-478.48						*				*
CARBayes BYM	Sí (BYM)	Sí	685.37		342.68	-552.08						*	*	*		*
INLA (sin efectos aleatorios)	No	No	1025.01	1024.99	10.83	-553.48						*	*	*		*
INLA (Heterogeneidad)	No	Sí	1003.2	1003.5	26.13	-554.75						*		*		*
INLA (Besag)	Sí (Besag)	No	1024.96	1024.96	10.83	-1412.23						*	*	*		*
INLA (Besag + Heterogenidad)	Sí (Besag)	Sí	1003.22	1003.2	26.34	-1413.53						*		*		*
INLA																
(BYM)	Sí (BYM)	No	1024.36	1024.56	11.25	-384.45						*	*	*		*
INLA																
(BYM + Heterogeneidad)	Sí (BYM)	Sí	1002.03	1002.19	26.61	-385.78						*		*		*

<sup>1</sup> X1: Elevación media; X2: Temperatura media anual; X3: Precipitación media anual; X4: Rango medio diurno; X5: Precipitación media del mes más lluviosa; X6: Precipitación media del mes más seco; X7: Isotermalidad; X8: Estacionalidad de la temperatura; X9: Rango anual de temperatura; X10: Estacionalidad de la precipitación.



**Tabla 3. Modelo seleccionado con mejor ajuste: INLA (BYM) con heterogeneidad.** Resumen de los parámetros estimados por el modelo INLA bajo la especificación BYM (Besag–York–Mollié), que incorpora efectos estructurados y no estructurados, y regímenes espaciales a nivel de departamentos. Se presentan las estimaciones posteriores para los efectos fijos, los hiperparámetros y los criterios de ajuste (DIC, WAIC).

	mean	sd	0.025quant	0.5quant	0.975quant	mode	k1d
(Intercept)	-1.759	0.168	-2.110	-1.752	-1.444	-1.752	0
elev_mean	0.041	1.131	-2.223	0.057	2.216	0.057	0
temperatura_media_anual_mean	0.700	1.147	-1.591	0.714	2.913	0.715	0
precipitacion_anual_mean	-0.191	0.661	-1.501	-0.187	1.094	-0.187	0
rango_medio_diurno_mean	-0.837	1.353	-3.466	-0.846	1.841	-0.846	0
precipitacion_mes_mas_lluvioso_mean	0.434	0.497	-0.537	0.432	1.415	0.432	0
precipitacion_mes_mas_seco_mean	-0.627	0.290	-1.190	-0.629	-0.052	-0.629	0
isotermalidad_mean	1.051	0.929	-0.788	1.057	2.854	1.057	0
estacionalidad_de_la_temperatura_mean	0.447	0.187	0.078	0.448	0.811	0.448	0
rango_anual_de_temperatura_mean	0.739	1.794	-2.811	0.751	4.225	0.751	0
estacionalidad_de_la_precipitacion_mean	-0.575	0.259	-1.074	-0.578	-0.057	-0.578	0

Random effects:

```
Name      Model
dpto_cnabr IID model
id BYM model
```

Model hyperparameters:

	mean	sd	0.025quant	0.5quant	0.975quant	mode
Precision for dpto_cnabr	3.68	2.48	1.04	3.04	10.25	2.13
Precision for id (iid component)	2774.72	3271.73	231.24	1779.43	11329.60	630.30
Precision for id (spatial component)	2779.40	3286.46	233.36	1780.47	11366.71	635.27

Deviance Information Criterion (DIC) .....: 1002.03

Deviance Information Criterion (DIC, saturated) ....: 1000.41

Effective number of parameters .....: 26.61

Watanabe-Akaike information criterion (WAIC) ....: 1002.19

Effective number of parameters .....: 25.88

Marginal log-Likelihood: -385.78

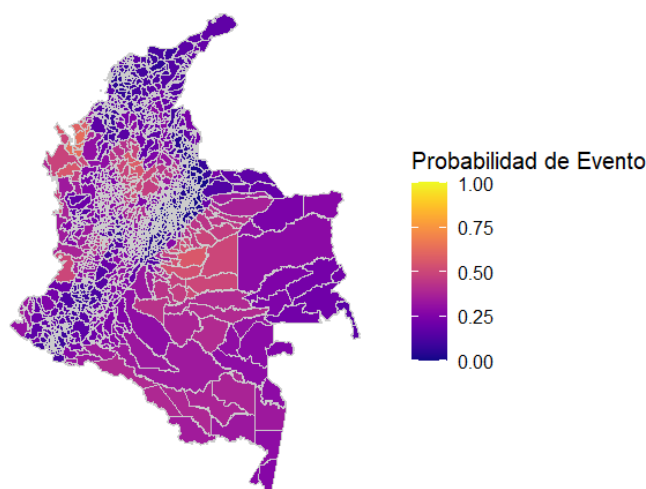
is computed

Posterior summaries for the linear predictor and the fitted values are computed

(Posterior marginals needs also 'control.compute=list(return.marginals.predictor=TRUE)')

### Mapa de Probabilidad de Ocurrencia del Evento

Predicciones generadas a partir del modelo INLA BYM + Heterogeneidad



**Figura 10. Mapa de predicción final del modelo INLA (BYM) con heterogeneidad.**

El mapa muestra la probabilidad estimada de ocurrencia de *Ochroma pyramidale* en los municipios de Colombia

## 4. Discusión

Los datos utilizados provienen de GBIF, una plataforma que integra información de ciencia ciudadana (observaciones en iNaturalist), especímenes de herbarios y, en el caso de Colombia, principalmente registros de inventarios forestales requeridos como parte de la reglamentación ambiental para la solicitud de licencias. Por lo tanto, la ausencia de registros de balso no implica necesariamente ausencia de la especie, sino que puede deberse a una falta de muestreo en determinadas zonas. A su vez, los datos muestran un claro sobremuestreo en algunas regiones del país, con muchos municipios sin registros y unos pocos con conteos altos, como ocurre en Antioquia, especialmente en los municipios de Buriticá, Sabanalarga y Liborina. Este patrón se asocia al Proyecto de Aprovechamiento Forestal en el área de influencia del Proyecto Hidroeléctrico Ituango, que impulsó un censo intensivo en la región.

Aunque la especie presenta una distribución amplia en Colombia, se concentra especialmente en los valles interandinos del Magdalena Medio y el río Cauca, en municipios de Antioquia y Santander. Un grupo adicional se ubica en el piedemonte llanero, donde el balso prospera en climas cálidos-húmedos, asociados a bosque seco tropical, bosque húmedo tropical y bosque montano bajo.

Los gráficos de las funciones de Ripley evidencian que existe un patrón espacial no aleatorio en los registros de *Ochroma pyramidale*. Se identifican zonas con alta densidad de registros y otras con grandes vacíos (*gaps*), lo que indica que el patrón observado ocurre con mayor frecuencia que lo esperado por azar. Este resultado coincide con lo reportado por otros autores: las especies con dispersión por viento tienden a exhibir patrones agregados, mientras que aquellas con dispersión por animales muestran distribuciones más dispersas. En este sentido, el balso —al presentar dispersión anemócora— manifiesta una agrupación espacial evidente a corta distancia (Arango et al., 2011). Este patrón fue reforzado por el análisis de agrupamiento DBSCAN (con umbrales de 50 individuos y 50 km de distancia), que identificó siete clústeres principales: Medio-Bajo Cauca (occidente

antioqueño), Magdalena Medio (Santander), piedemonte oriental (Cundinamarca y Casanare), Putumayo, Valle del Cauca y Caribe (Sucre).

El modelo de regresión logística ajustado permitió controlar parcialmente la sobredispersión en los datos. Aunque el valor de AUC (0.72) indica un desempeño aceptable, este modelo se considera una línea base. Los modelos con enfoque bayesiano (CAR e INLA) mejoraron significativamente el ajuste, particularmente aquellos que incorporan efectos de dependencia espacial y heterogeneidad. Los modelos CARBayes tipo Leroux ( $q = 1$ ) y BYM, así como los modelos INLA con efectos espaciales (Besag, BYM), redujeron los valores del DIC comparados con la regresión logística clásica. El mejor modelo correspondió al INLA (BYM + heterogeneidad), que presentó los valores más bajos de DIC y WAIC. Este modelo combina efectos aleatorios estructurados y no estructurados, y consideró como régimen espacial la división político-administrativa por departamentos.

Las variables climáticas más relevantes para explicar la probabilidad de presencia de balso en los municipios fueron:

- Precipitación del mes más seco: a menor precipitación, mayor probabilidad de presencia de balso. Esto es consistente con estudios previos que destacan la tolerancia del balso al déficit hídrico, lo cual le confiere una ventaja competitiva en condiciones de sequía.
- Isotermalidad: valores altos (mayor variación diaria frente a la variación anual de la temperatura) se asocian a una mayor probabilidad de ocurrencia. Esto sugiere que la especie podría verse afectada en zonas con alta estacionalidad térmica, ya que es susceptible a heladas y cambios térmicos abruptos (Francis et al., 2000)
- Estacionalidad de la temperatura: valores altos indican mayor dispersión anual de temperatura. En este caso, también se observó una relación positiva con la probabilidad de presencia.
- Estacionalidad de la precipitación: a mayor estacionalidad (mayor coeficiente de variación), disminuye la probabilidad de ocurrencia. Esto puede deberse a que la especie es susceptible a inundaciones (Francis et al., 2000), por lo que no prospera en sitios con lluvias intensas concentradas en periodos cortos.

Los resultados de los mapas coropléticos, el índice de Moran y los gráficos LISA confirman una fuerte autocorrelación espacial en estas variables predictoras (valores superiores a 0.85), lo que indica que altos valores tienden a agruparse espacialmente con otros altos, y lo mismo ocurre con los valores bajos. Esto refuerza la idea de que las variables ambientales presentan una naturaleza espacialmente dependiente y heterogénea, lo cual viola los supuestos de independencia y homogeneidad de la estadística clásica. Por lo tanto, los métodos de análisis geoespacial son herramientas esenciales para abordar este tipo de problemáticas en datos ambientales.

Entre los limitantes, retos y oportunidades, encontré que uno de los principales desafíos es la alta sobredispersión de los datos, que limita la capacidad de interpretar correctamente los patrones de distribución. Dado que aproximadamente el 80 % de los municipios (897 de 1.121) no presentan registros de balso, una alternativa sería restringir el análisis únicamente a los sitios donde sí se ha registrado la especie, con el fin de entender mejor las condiciones ambientales asociadas a su presencia. Esto permitiría construir una línea base más sólida para identificar otras zonas con condiciones similares y potencial para el establecimiento de plantaciones forestales de balso.

Además, sería recomendable explorar regímenes espaciales más adecuados desde el punto de vista ecológico, como macrocuencas, cuencas o subcuencas hidrográficas, en lugar de divisiones político-administrativas como los departamentos. Estas unidades reflejan de manera más precisa los procesos ecológicos que influyen en la distribución natural de la especie y podrían mejorar significativamente la calidad y la interpretación de los modelos espaciales.

## 5. Conclusiones y perspectivas

Este estudio evidenció un sobremuestreo por municipio, lo cual genera sobredispersión en los datos de conteo. Sin embargo, al tratarse de datos reales y disponibles, su análisis es fundamental para distinguir el patrón actual de distribución espacial del balso y comprender los procesos que lo generan, especialmente como insumo para la toma de decisiones en torno al establecimiento de plantaciones forestales futuras en Colombia.

La regresión logística se presenta como una herramienta robusta para modelar la presencia/ausencia de la especie, mitigando en parte los efectos del sobremuestreo. No obstante, los modelos que incorporan estructura espacial —como los enfoques bayesianos CARBayes e INLA con dependencia y heterogeneidad— mostraron mejores desempeños estadísticos (menores valores de DIC y WAIC), por lo que se recomienda su uso en estudios que trabajen con datos espaciales ambientales.

Entre las variables climáticas evaluadas, las que mejor explican la probabilidad de presencia de *Ochroma pyramidale* fueron: Precipitación en el mes más seco, Isotermalidad, Estacionalidad de la temperatura y Estacionalidad de la precipitación. Estos resultados coinciden con los patrones ecológicos conocidos para la especie: el balso prospera en zonas bajas, con climas cálido-húmedos, es tolerante al déficit hídrico pero susceptible a excesos de humedad prolongados. En este sentido, se favorece en entornos con baja estacionalidad climática.

Como proyección futura, resulta relevante analizar si los patrones identificados en este estudio se mantienen bajo distintos escenarios de cambio climático, considerando alteraciones en temperatura, precipitación y estacionalidad. Esta constituye una línea de investigación pertinente para evaluar la resiliencia y adaptabilidad del balso a largo plazo.

Finalmente, dada la alta concentración de registros en la región andina, una zona con intensa actividad antrópica (deforestación, agricultura, ganadería extensiva y urbanización), surge una pregunta clave: ¿Cómo se ven afectadas las poblaciones naturales de balso en términos de diversidad genética, ecología y dispersión, frente a estos procesos de transformación del paisaje? Responder esta pregunta podría aportar evidencia valiosa para la conservación de la especie y el diseño de estrategias de manejo sostenible en paisajes altamente intervenidos.

## References

1. Arango, H., Duque, Á., Cárdenas, D., & Barreto, J. S. (2011). *RELACIÓN ENTRE EL MECANISMO DE DISPERSIÓN DE SEMILLAS Y LA DISTRIBUCIÓN ESPACIAL DE ALGUNAS ESPECIES ARBÓREAS EN UN BOSQUE DE TIERRA FIRME DE LA AMAZONIA COLOMBIANA*.
2. *Boletín estadístico forestal*. (2024). [9 Edición]. Ministerio de Agricultura.
3. Elbert L. Little & Frank H. Wadsworth. (1964). Common Trees of Puerto Rico and the Virgin Islands. *Agriculture Handbook*, 249.
4. Francis, J. K., Lowe, C. A., & Trabanino, S. (2000). *Bioecología de Árboles Nativos y Exóticos de Puerto Rico y las Indias Occidentales*. 582.
5. Listado de árboles de El Salvador. (2009). *Englera*, 29, 17-225.
6. Missouri Botanical Garden. (2025, julio 25). *!Ochroma pyramidale* (Cav. Ex Lam.) Urb. [Tropicos.org]. *Tropicos v4.0.0*. <https://www.tropicos.org/name/3900204>
7. Ramos Corrales, P. C. (2016). «Balsa» *Ochroma pyramidale* (Cav. Ex Lam.) Urb. (*Bombacaceae*) [Tesis, Universidad Nacional de La Plata]. <https://doi.org/10.35537/10915/56011>