

# Acta de Reunión

Fecha: Miércoles, 1 de octubre de 2025

Lugar: Sesión virtual (Teams)

Proyecto: Mapa de Vulnerabilidad Faunística - INVÍAS

Herramientas: R (paquete 'dismo' con MaxEnt), Python, Google Earth Engine, fuentes IGAC/Colombia en Mapas y Global Forest Watch

Duración estimada: 1 h 08 min (según transcripción)

## 1. Propósito de la reunión

Revisar avances de modelación y preprocesamiento de insumos, validar la arquitectura técnica (ingesta → procesamiento → publicación), y acordar la estrategia de ejecución/actualización del modelo MaxEnt (periodicidad, segmentación por extensión y requerimientos de cómputo/almacenamiento). Adicionalmente, definir ajustes al inventario de insumos (nuevas columnas de capacidad y periodicidad) y preparar la presentación técnica para el comité de resultados del viernes.

## 2. Asistentes

- Luis Esteban Gómez Cadavid
- Andrés Felipe Hernández Marulanda
- Jairo Iván Coy Coy
- Nelson Aníbal Miranda Ríos
- Jasmín Marín Perez

## 3. Desarrollo de la reunión (narrativa)

Se inició con el reporte técnico del frente de modelación en R usando el paquete **\*\*dismo\*\*** (interfaz de **\*\*MaxEnt\*\***). Se aclaró que **\*\*MaxEnt\*\*** históricamente corría sobre **\*\*Java\*\***, y hoy se invoca desde R a través de **\*dismo\***, permitiendo **\*\*réplicas\*\*** del modelo y reportes de **\*\*jackknife\*\*** (importancia de variables) y **\*\*AUC\*\*** (área bajo la curva ROC). La transcripción mencionaba "DISMO", "GEDAL", "RAST" y "Abuserock"; se precisa la terminología correcta: **\*\*dismo\*\*** (paquete de R), **\*\*raster/terra\*\*** (paquetes de R para ráster), **\*\*jackknife\*\*** y **\*\*AUC-ROC\*\*** respectivamente.

Con fines de robustez se definieron **\*\*5 réplicas\*\*** del modelo (configurables). La partición de datos se aclaró como **\*\*80% entrenamiento / 20% validación\*\*** (la transcripción mencionaba "70%" de forma inconsistente). El **\*\*jackknife\*\*** mostró alta contribución de bandas de Sentinel-2, en particular **\*\*B8A (NIR estrecha)\*\*** y **\*\*B11 (SWIR)\*\***. Se emplean ~**\*\*30 variables\*\*** entre: **\*\*índices espectrales\*\*** (NDVI, etc.), **\*\*derivados estáticos\*\*** (distancias a cuerpos de agua, red vial, áreas protegidas y núcleos de pérdida/ganancia de bosque) y **\*\*morfometría\*\*** (pendiente; **\*surface model\***). Las **\*\*distancias\*\*** se calculan como **\*\*euclidianas\*\*** sobre ráster previamente rasterizados y normalizados.

Para la salida, se promedia el raster de probabilidad de las réplicas y se reclasifica en **\*\*deciles 0.0-1.0\*\*** (0-0.1, 0.1-0.2, ...). Luego se **\*\*transfiere\*\*** al eje vial mediante un **\*\*buffer\*\*** ( $\pm 60$  m) para asignar una probabilidad por **\*\*segmento\*\*** (vector final), que será la capa visible en el visor de INVÍAS.

Respecto a desempeño, se documentó que una prueba regional (Orinoquía) en un **\*\*Core i7 3ra gen, GPU 2 GB\*\*** demora ~**\*\*2 días\*\*** para 5 réplicas; se estima ~**\*\*10 horas\*\*** para una sola, a resolución de **\*\*30 m\*\***. Se discutió explorar **\*\*100 m\*\*** como resolución operativa inicial para reducir tiempos y almacenamiento, y **\*\*recortar\*\*** insumos a un **\*\*buffer de la red\*\***.

vial de INVÍAS\*\*, ya que no interesan áreas sin vía.

Se socializó el inventario de fuentes (\*\*IGAC/Colombia en Mapas\*\* vía servicios \*\*REST\*\*; \*\*Global Forest Watch\*\* por \*\*teselas\*\* descargables; \*\*GEE\*\* para colecciones dinámicas). Varias descargas son \*\*manuales\*\* (p. ej. GFW por cuadrantes), por lo que se propuso clasificar capas por \*\*estáticas / semidinámicas / dinámicas\*\* y automatizar donde sea posible (APIs/robots). Se acordó agregar dos columnas al inventario: (i) \*\*Capacidad/volumen estimado\*\* (p. ej., MB-GB por día/mes) y (ii) \*\*Periodicidad de actualización\*\* esperada.

A nivel de arquitectura, se debatió si el \*\*backend\*\* procesará internamente \*\*MaxEnt\*\* o si consumirá resultados de un \*\*pipeline paralelo\*\* (para no sobrecargar el servidor de producción). Se enfatizó la necesidad de \*\*segmentar\*\* (por regiones, por extensión solicitada o por \*tiles\*) y de establecer ventanas de \*\*proceso nocturno\*\* cuando se ejecuten corridas completas (mantenimiento programado con limpieza de \*overheads\* y borrado de temporales).

## 4. Agenda (simplificada)

- Avances de modelación en R (dismo/MaxEnt), jackknife y AUC-ROC.
- Estandarización de insumos: rasterización, distancias euclidianas, resolución (30-100 m), buffer  $\pm 60$  m sobre vías.
- Inventario de datos: fuentes REST/teselas/APIs; clasificación estática-semidinámica-dinámica; nuevas columnas.
- Arquitectura y operación: ¿procesar en backend o consumir resultados externos? Segmentación/tiles y ventanas nocturnas.
- Preparación de presentación técnica (8-10 diapositivas) para el comité del viernes.
- Definición de reunión técnica de integración para el jueves 2 de octubre (2:30 p. m.).

## 5. Revisión de compromisos pasados

- Se presentó inventario de capas con origen, formato y resolución; se añaden columnas de 'capacidad' y 'periodicidad'.
- Se realizaron pruebas de MaxEnt con 5 réplicas y reporte de jackknife/AUC; pendiente consolidar promedio y deciles.
- Se avanzó en rasterización y cálculo de distancias (agua, bosque/no bosque), corrigiendo geometrías y valores NoData.

## 6. Próximos pasos y posibles fechas

- Agregar al inventario columnas: \*\*Capacidad/volumen estimado\*\* y \*\*Periodicidad de actualización\*\* por capa.
- Definir \*\*resolución operativa\*\* (propuesta inicial 100 m) y \*\*recorte\*\* por buffer de red vial INVÍAS ( $\pm 60$  m).
- Consolidar \*\*salida vectorial\*\* por segmento (promedio de réplicas → deciles → transferencia a ejes viales).
- Arquitectura: decisión entre \*\*procesamiento interno\*\* vs. \*\*pipeline paralelo\*\*; diseñar \*\*tiling/extentos\*\* y ventanas nocturnas.
- Infraestructura: solicitar a INVÍAS un \*\*nodo de pruebas\*\* (coordinación con Darwin) para medir CPU/RAM/IO y almacenamiento.
- Entorno de pruebas MaxEnt: Jairo enviará \*\*ambiente reproducible\*\* (scripts/paquetes) antes del \*\*domingo 5/10/2025\*\*.
- Reunión de integración técnica: \*\*jueves 2/10/2025 a las 2:30 p. m.\*\* (inicia Jasmín; Nelson se suma 6:00 p. m.).

- Preparar \*\*presentación 8-10 diapositivas\*\*: actividad, script, resultados, problemas y limitaciones para el \*\*viernes 3/10/2025\*\*.

## 7. Hitos / Conclusiones

- Criterios metodológicos fijados: réplicas en MaxEnt, promedio y decilización; vectorización por segmento vial.
- Clarificado el uso de \*\*dismo/raster(terra)\*\*, \*\*jackknife\*\* y \*\*AUC-ROC\*\* (corrección de términos transcritos).
- Definida la \*\*segmentación/tiles\*\* y \*\*ventanas nocturnas\*\* como estrategia de operación.
- Acuerdo de \*\*clasificación de fuentes\*\* y \*\*automatización\*\* progresiva (APIs/robots) donde aplique.

## 8. Análisis y recomendaciones

El enfoque de promediar réplicas y trasladar el ráster a eje vial por deciles mejora la \*\*interpretabilidad\*\* y reduce el costo de publicación. Para \*\*escalabilidad\*\*, resulta clave: (1) adoptar \*\*100 m\*\* como resolución operativa inicial (con pruebas puntuales a 30 m donde sea crítico), (2) \*\*recortar\*\* todo insumo a buffer de red vial de INVÍAS, (3) \*\*segmentar\*\* por regiones/extentos y \*\*tilear\*\* para paralelizar cómputo y servir estáticos.

Recomendaciones operativas: definir una \*\*política de actualización\*\* (p. ej., mensual para dinámicas de GEE; anual o bienal para estáticas), habilitar \*\*mantenimiento nocturno\*\* para corridas completas, instrumentar \*\*limpieza automática\*\* de temporales y versiones antiguas, y mantener un \*\*checklist\*\* por capa (origen, método, resolución, NoData, reproyección, distancia, normalización). En infraestructura, negociar con INVÍAS \*\*nodo de pruebas\*\* y medir tiempos reales antes de cerrar la arquitectura. Documentar todo en \*\*wiki\*\* (pipeline, parámetros, versiones) y mantener \*\*demo técnica\*\* corta en cada comité.