

## 1. Introducción y contexto

En esta sección presentarás la importancia de estudiar la susceptibilidad a deslizamientos y avenidas torrenciales en los Andes del norte de Colombia ( $\geq 500$  msnm). Explica la relevancia geológica y social: la región andina concentra población, infraestructura vital y ecosistemas frágiles, mientras que la combinación de fuertes pendientes y lluvias intensas incrementa el riesgo de eventos geohidrológicos extremos. Describe brevemente el estado del arte: estudios previos de zonificación de riesgo en escalas locales, limitaciones al extrapolar modelos a toda la cordillera, y la oportunidad de integrar grandes volúmenes de datos satelitales (DEM, CHIRPS) y técnicas de Machine Learning para una evaluación más robusta y reproducible.

---

## 2. Objetivos

- **Objetivo general:** Generar mapas de susceptibilidad a deslizamientos y avenidas torrenciales en los Andes del norte de Colombia ( $\geq 500$  msnm) mediante el uso de parámetros morfométricos derivados de un DEM de alta resolución, series históricas de precipitación (CHIRPS 1981–2024) e inventarios de eventos.
- **Objetivos específicos:**
  1. Delimitar automáticamente subcuencas de escorrentía adecuadas para el análisis (áreas entre 10 y 50 km<sup>2</sup>).
  2. Calcular un conjunto extensivo de variables morfométricas (pendiente, curvaturas, índices de drenaje, tiempo de concentración, integral hipsométrica, entre otros) para cada subcuenca.
  3. Aplicar funciones discriminantes establecidas para clasificar la susceptibilidad a avenidas torrenciales, incorporando además información litológica para diferenciar tipos de flujo.
  4. Reunir y depurar un inventario georreferenciado de deslizamientos históricos, y entrenar modelos de Machine Learning (Random Forest y XGBoost) para predecir la susceptibilidad a deslizamientos.
  5. Validar ambos modelos mediante metodologías robustas (validación cruzada, métricas ROC/AUC, análisis de falsos positivos/negativos) y cuantificar la incertidumbre espacial.

6. Documentar todo el flujo de trabajo con scripts automatizados y producir un informe técnico exhaustivo y mapas listos para aplicación en planificación territorial y gestión de riesgos.
- 

### **3. Área de estudio**

Describe detalladamente cómo se definirá el dominio:

- Fuente y procesamiento del DEM (SRTM 30 m), reproyección a sistema local (MAGNA-Sirgas), recorte a la región andina con elevaciones  $\geq 500$  msnm.
  - Justificación del umbral de 500 msnm: relación con climatología de montaña y ocurrencia histórica de procesos geomorfológicos extremos.
  - Descripción geomorfológica, litológica y climatológica de la región: cadenas montañosas, principales formaciones rocosas, gradientes de precipitación y zonas de uso del suelo.
- 

### **4. Datos y preprocesamiento**

#### **4.1 DEM y análisis hidrológico**

Explica cómo se usará el DEM para generar la dirección y acumulación de flujo. Detalla cada paso: llenado de hundimientos, cálculo de dirección de flujo (D8), generación de red de drenaje y extracción de cuencas. Describe la lógica de elección de subcuencas de 10–50 km<sup>2</sup> para equilibrar resolución y capacidad de cómputo.

#### **4.2 Precipitación CHIRPS**

Describe la adquisición de la serie temporal CHIRPS (1981–2024), reproyección y remuestreo al mismo grid del DEM. Explica el cálculo de variables pluviométricas: lluvia máxima diaria, acumulada mensual, número de días con precipitación  $> 50$  mm, y anomalías respecto a la media histórica.

#### **4.3 Inventario de deslizamientos**

Documenta las fuentes: IDEAM, publicaciones académicas, portales de datos abiertos. Describe el proceso de depuración: eliminación de duplicados, verificación de coordenadas, segmentación por fecha y magnitud.

#### **4.4 Datos complementarios**

Incluye mapas de litología (IGAC), cobertura de suelo, fallas geológicas y, si disponible, curvas de nivel temblorosa (VS30) para análisis sísmico que pueda influir en estabilidad de taludes.

---

## **5. Cálculo de variables morfométricas**

Presenta con detalle las más de cuarenta variables que derivarás:

- Parámetros de forma de cuenca (área, perímetro, factor de forma, circularidad).
- Parámetros de pendiente y relieve (pendiente media, máxima, desviación estándar, relieve relativo, integral hipsométrica).
- Parámetros de drenaje (densidad, frecuencia, índice de Melton, constante de mantenimiento, tiempos de concentración).
- Índices de humedad y escorrentía (TWI, SPI).
- Curvaturas (planar, perfil).

Explica brevemente la relevancia geológica de cada uno para procesos de deslizamiento y avenidas torrenciales.

---

## **6. Modelación de avenidas torrenciales**

Describe la metodología basada en las funciones discriminantes de Aristizábal et al. (2021). Explica el origen de cada variable ( $Nu$ ,  $Rw$ ,  $C$ ,  $H$ ,  $M$ ), la ecuación matemática, y la interpretación de los umbrales  $A$  y  $T$ . Detalla cómo se incorpora la litología para refinar la clasificación en flujos de escombros canalizados versus inundaciones de sedimentos finos. Finalmente, describe cómo se generará el mapa continuo de susceptibilidad y la capa de tipo de avenida.

---

## **7. Modelación de deslizamientos**

### **7.1 Preparación de datos para Machine Learning**

Detalla la construcción de la tabla de entrenamiento: cada registro corresponde a una celda de terreno o subcuenca, con sus variables morfométricas y pluviométricas, más la etiqueta binaria (evento/no evento). Explica el balanceo de clases (submuestreo, SMOTE) si la ocurrencia de deslizamientos es rara.

## **7.2 Selección y entrenamiento de modelos**

Describe por qué elegir Random Forest y XGBoost: robustez a variables correlacionadas, interpretación de importancia de variables. Explica el proceso de optimización de hiperparámetros (grid search o Bayesian optimization), validación cruzada a 5 folds y métrica principal (AUC-ROC).

## **7.3 Evaluación y análisis de resultados**

Explica cómo interpretar la curva ROC, la matriz de confusión y las métricas de precisión, sensibilidad y especificidad. Detalla el análisis espacial de falsos positivos y negativos para entender sesgos geográficos.

---

## **8. Validación conjunta y análisis de incertidumbre**

Describe un enfoque para combinar la incertidumbre de ambos modelos: generación de mapas de probabilidad continua, identificación de áreas con alta variabilidad entre folds, y cómo estas zonas deben ser tratadas con precaución en aplicaciones prácticas.

---

## **9. Resultados esperados**

### **9.1 Mapas de susceptibilidad**

- Mapa de susceptibilidad a avenidas torrenciales, con clasificación en categorías (baja, media, alta).
- Mapa de susceptibilidad a deslizamientos, con probabilidades continuas y categorías definidas según cuantiles.

### **9.2 Análisis de variables**

- Gráficos de importancia de variables para cada modelo, destacando qué parámetros morfométricos y pluviométricos controlan mejor los procesos.
- Discusión sobre similitudes y diferencias en drivers de deslizamientos versus torrenciales.

### **9.3 Recomendaciones para gestión del riesgo**

- Identificación de corredores críticos donde ambos riesgos se superponen.
- Sugerencias de monitoreo y medidas de mitigación (reforestación, estructuras de retención, alertas tempranas).

---

## 10. Documentación y reproducibilidad

Explica que todo el código en Python (con librerías GDAL, richdem, scikit-learn, pandas) y los scripts de GEE estarán disponibles en un repositorio público de GitHub, con instrucciones paso a paso para replicar el análisis.

---

## 11. Cronograma narrative

Durante las ocho semanas, trabajarás de manera intensiva siguiendo este ritmo:

- **Semanas 1–2:** obtención y preprocesamiento de datos. Generación de cuencas y cálculo de variables morfométricas y pluviométricas.
- **Semana 3:** modelación de avenidas torrenciales, clasificación y refinamiento litológico.
- **Semanas 4–5:** preparación de datos para deslizamientos, entrenamiento y ajuste de modelos de Machine Learning.
- **Semana 6:** validación cruzada de ambos modelos, análisis de incertidumbre y mapeo de probabilidades.
- **Semana 7:** elaboración de mapas finales en QGIS y gráficos de importancia de variables en Python.
- **Semana 8:** redacción intensiva del informe, revisión de coherencia, preparación de la presentación y entrega de todos los productos.