

¿Qué tanto se mueven nuestros ríos?: Propuesta de una nueva metodología para la caracterización de frentes erosivos y sedimentarios en las orillas de los ríos

Alexander Montalvo M.

Scipy Latam 2019
Bogotá, Colombia

09/10/2019



Presentación

- 1 Problemática**
- 2 Información secundaria**
- 3 Estado del arte**
- 4 Datos y métodos**
- 5 Resultados**
- 6 Conclusiones, Recomendaciones y Logros**



Problemática



Problemática

- Pérdidas de áreas de cultivos y/o desplazamiento de asentamientos humanos a causa de la migración de orillas.
- Visualización actual del riesgo frente a estos fenómenos.
- Exceso de generalización para la formulación de modelos estadísticos representativos.
- Herramientas para la delimitación variable de rondas hídricas en cuerpos lóticos, brechas de escala e información para su definición. *Art. 206 de la ley 1450 de 2011.*
- Delimitación cartográfica.



Información secundaria

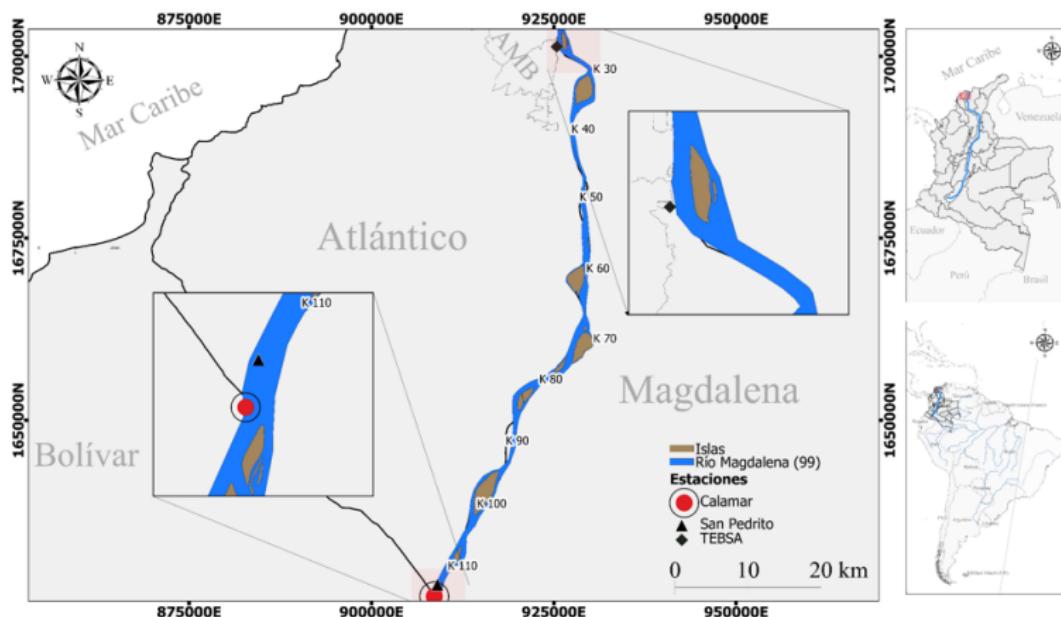


Fuentes de datos

- Se tomará como caso de estudio el río Magdalena entre el K22 y K115.
- La recopilación de los datos hidrológicos se hizo a través de las bases de datos oficiales del IDEAM y se tomó la estación de Calamar como la más representativa.
- La cuantificación de los cambios morfológicos en las orillas se encuentra asociada a la resolución espacial de los satélites.
- El dominio temporal de investigación tiene una extensión de 42 años, desde 1976 hasta 2018.
- La clasificación de eventos macroclimáticos extremos se realizó con base en la categorización publicada en el Estudio Nacional del Agua de 2010 hecha por el IDEAM.
- El NDVI se tomó como la única variable explicativa para el proceso de cambios de uso de suelo a causa de actividades antropogénicas y climáticas de las orillas en la zona de estudio.



Zona de estudio



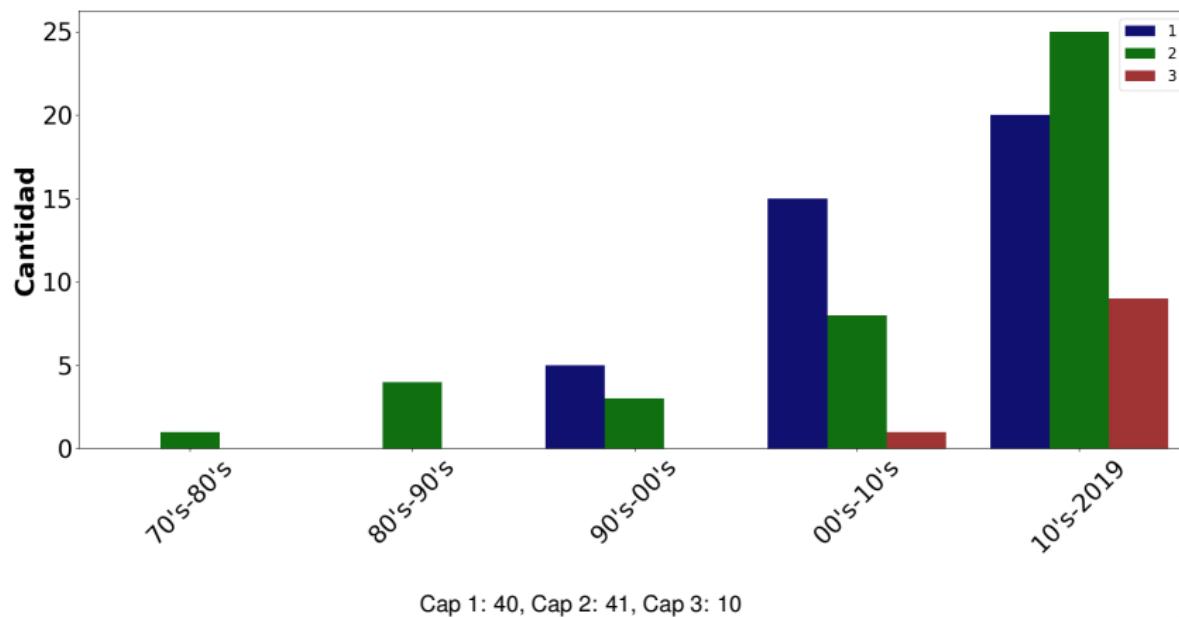
Localización de la zona de estudio donde se aplicó la metodología. Elaboración propia.



Estado del arte



Referencias



Literatura

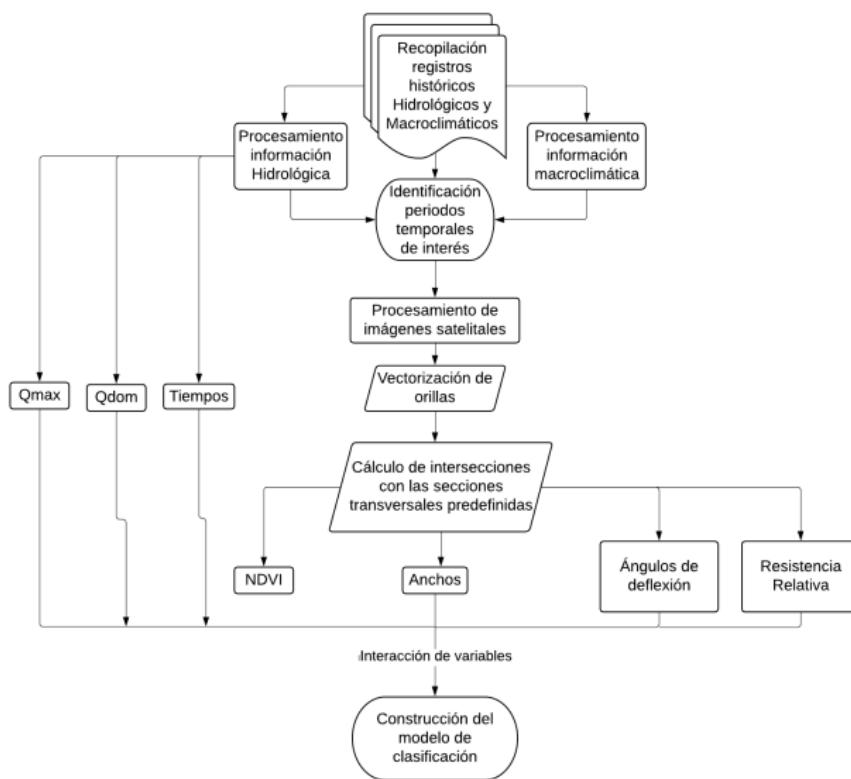
- Los eventos hidrológicos extremos son uno de los mayores desencadenantes de las alteraciones morfológicas de los ríos (Garde, 2012; Bertoldi, Zanoni y Tubino, 2010).
- La ITCZ y el ENSO canónico como los principales forzantes para el clima de la cuenca de estudio (Poveda et al., 2001; IDEAM,2018)
- Cuatro grupos de técnicas para medición de cambios: Mediciones de campo, análisis gráficos de mapas, análisis espectrales de curvaturas y análisis de parámetros morfométricos (Lawler, 1993).
- Capacidad predictiva de los modelos estadísticos (Baki y Gan, 2012), (Hossain et al., 2013).
- La vegetación ribereña **no** solamente otorga beneficios para la reducción de la migración de orillas (van Duijn,2018)
- Problema para la toma de datos en el análisis de este fenómeno debido a las brechas en las escalas temporales, espaciales y accesibilidad de la información.



Datos y métodos



Esquema General



Resolución y fuentes de datos

Eventos extremos

- **Macroclimatología (BoM y NOAA):** 0% → frecuencia mensual.
- **Hidrológicos (IDEAM):** 1.14%, 0.46%, 5.53% → frecuencia diaria
- Análisis exploratorio de datos.
- Análisis de correlación cruzada.
- Identificación de periodos extremos.

SIG

- Imágenes satelitales → Landsat 5,7 y 8. Recolección, ortorectificación y reducción de colecciones representativas de imágenes satelitales por periodo temporal.
- Aplicación de filtros para selección de colecciones de imágenes → temporal, geográfico, cobertura de nubes.
- Aplicación de la metodología propuesta por Donchyts et al.(2016) para la plataforma Aquamonitor.
- Reclasificación de raster mediante CART e identificación de puntos de intersección entre orillas y secciones transversales → 2Km.



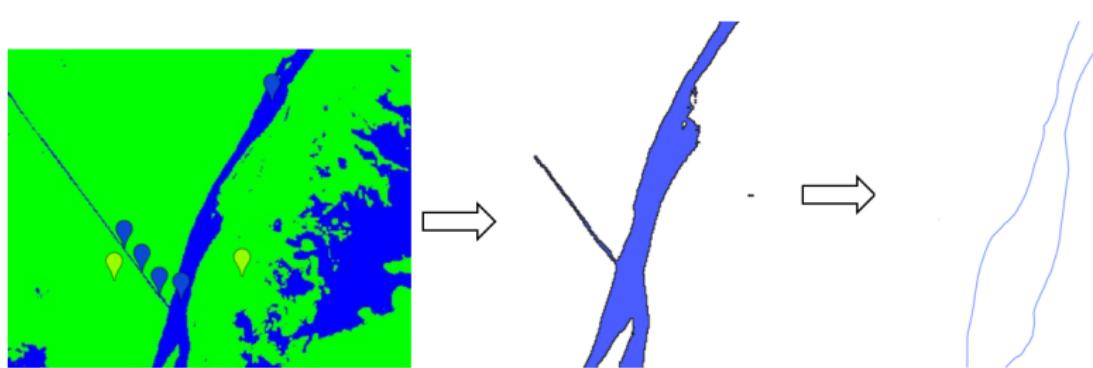
Imágenes satelitales

$$NDVI = \frac{NIR - RED}{NIR + RED}$$
$$NDWI = \frac{NIR - SWIR}{NIR + SWIR}$$

Año	Satélite	Resolución	Red	NIR	SWIR
1976			Origen bases de datos del LEHF		
1984, 1988, 1991, 1996, 1998	Landsat 5				
2001, 2004, 2010, 2012, 2013	Landsat 7	30 m	B3	B4	B5
2014 a 2018	Landsat 8		B4	B5	B6



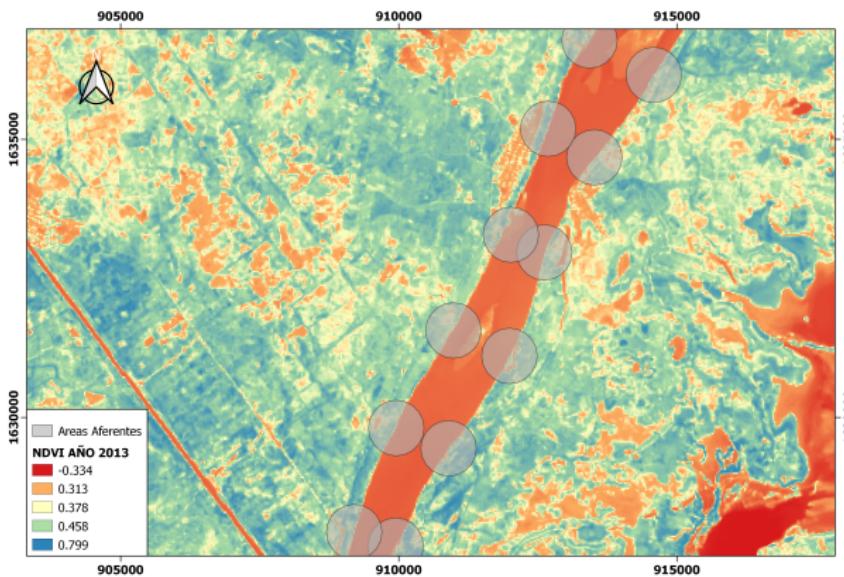
Imágenes satelitales



Proceso de vectorización de las orillas de la zona de estudio a partir de imágenes satelitales reclasificadas a través de CART en la plataforma GEE. Elaboración propia



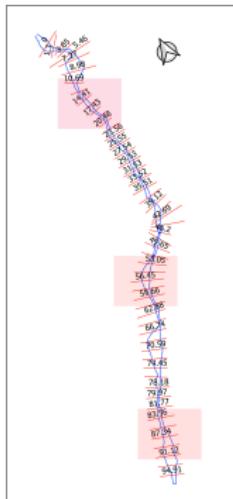
Imágenes satelitales



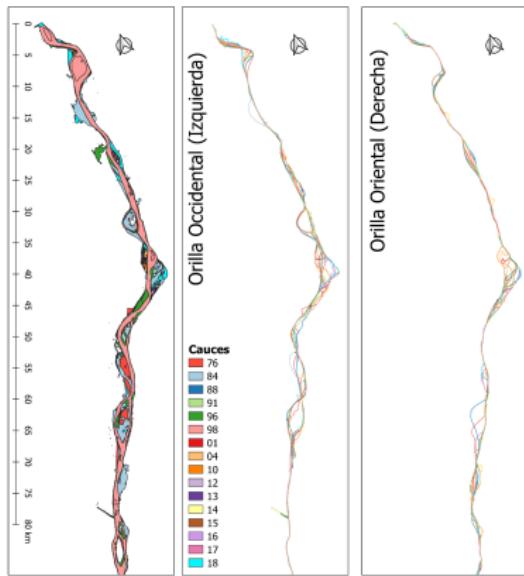
Cálculo del NDVI de las orillas mediante zonas aferentes (500 m) de los puntos de intersección encontrados. Los datos corresponden al año 2013. Elaboración propia.



Imágenes satelitales



Secciones transversales y puntos de intersección. Elaboración propia.



Cauce (polígonos) y orillas (líneas) para los años analizados. Elaboración propia.

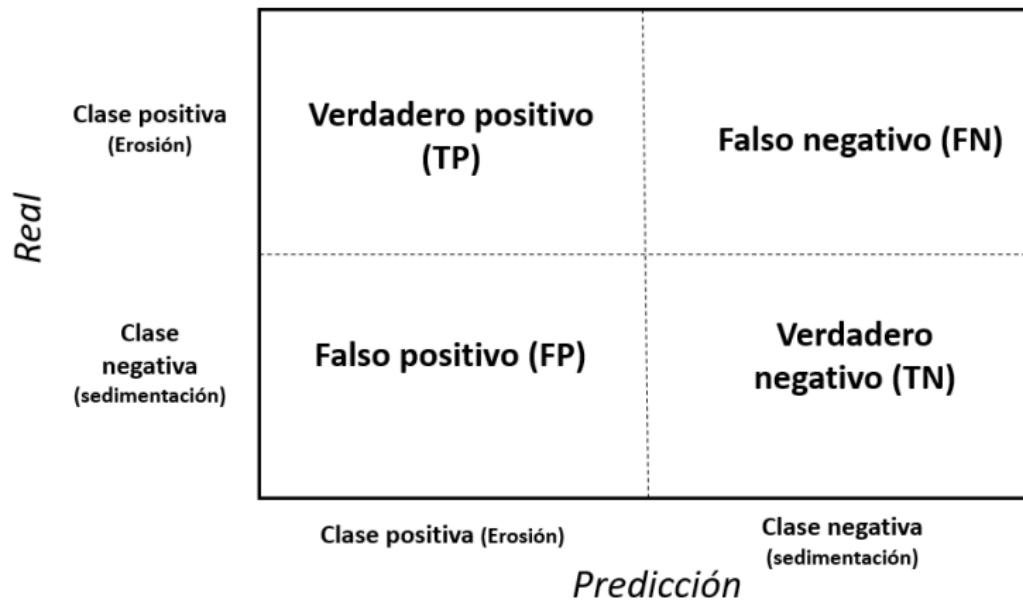


Integración de variables

- Modelo base de aprendizaje supervisado Árboles de clasificación → Bagging → **Random Forest**
- Error de generalización
 - Overfitting → Alta varianza
 - Underfitting → Alto sesgo (bias)
- Optimización de Hiperparámetros.
 - Máx. profundidad.
 - Cant. mínima de muestras.
 - Cant. máxima de features.
- Eficiencia del modelo
 - $Exactitud = \frac{TP+TN}{TP+TN+FP+FN}$
 - $Precision = \frac{TP}{TP+FP}$
 - $Sensibilidad = \frac{TP}{TP+FN}$
 - $F1 Score = 2 \frac{Precision \cdot Sensibilidad}{Precision + Sensibilidad}$
 - AUC ROC



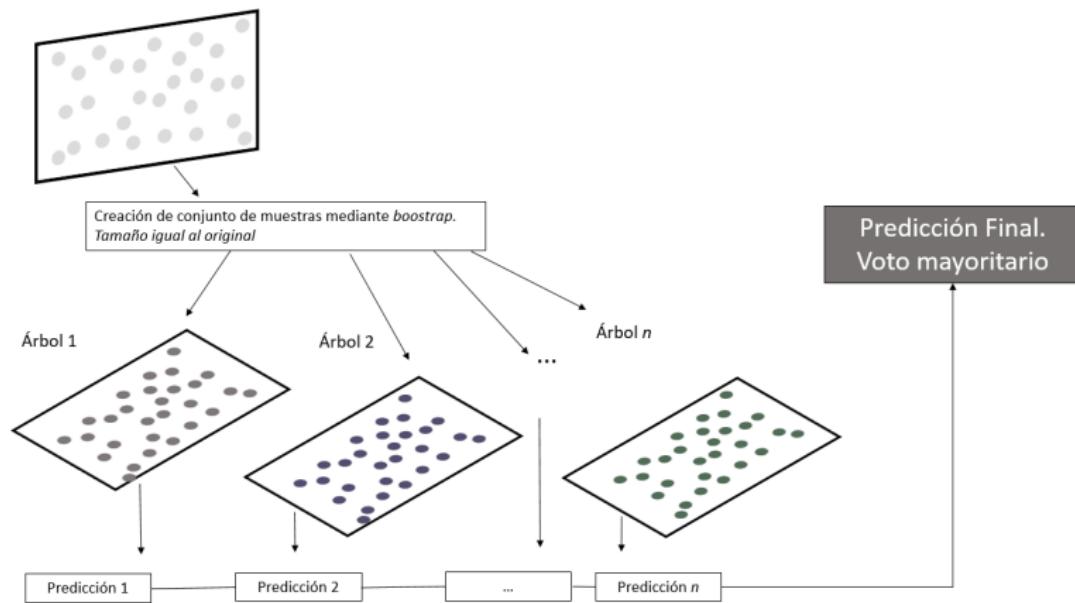
Integración de variables



Matriz de confusión para la definición de los métricos utilizados para evaluar el desempeño del modelo. Elaboración propia.



Integración de variables



Esquema de funcionamiento de la técnica de Random Forest para la tarea de clasificación. Elaboración propia.

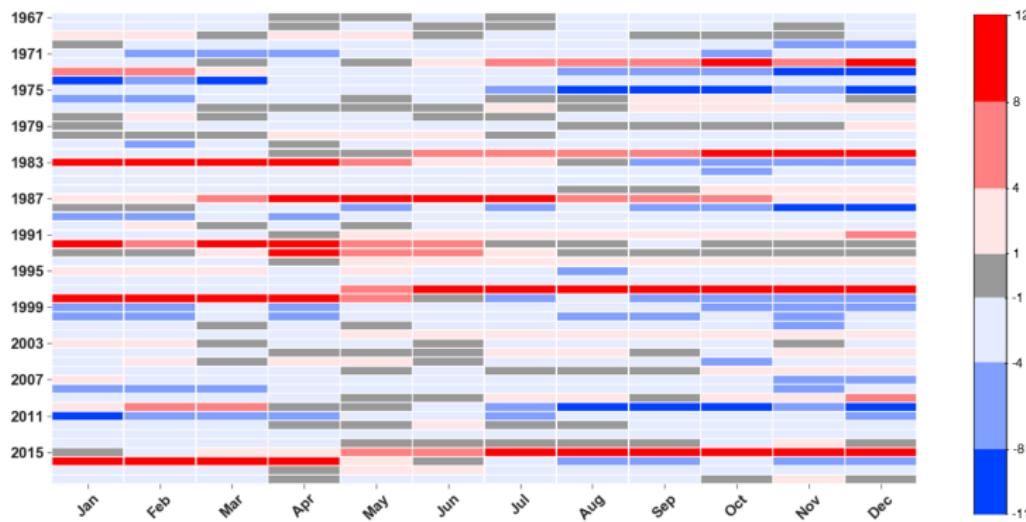


Resultados



Eventos hidrológicos extremos

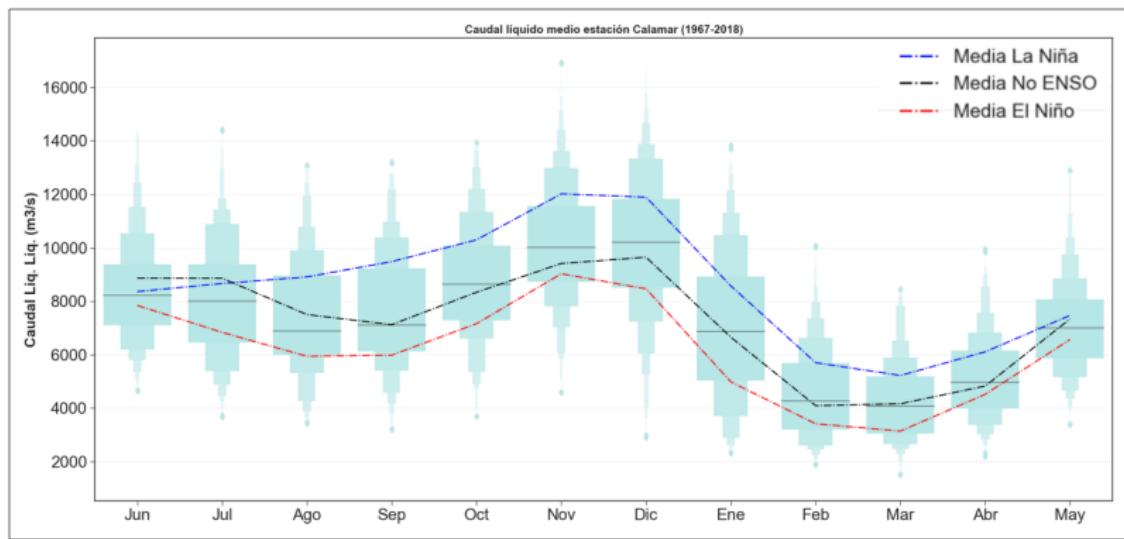
- Índices macroclimáticos ENSO: ONI, SOI, MEI y TSM.



Clasificación de meses extremos. Se destaca la existencia de un rezago de 2 meses entre los índices macroclimáticos (MEI y TSM) y las variables hidrológicas. Elaboración propia.



Eventos hidrológicos extremos

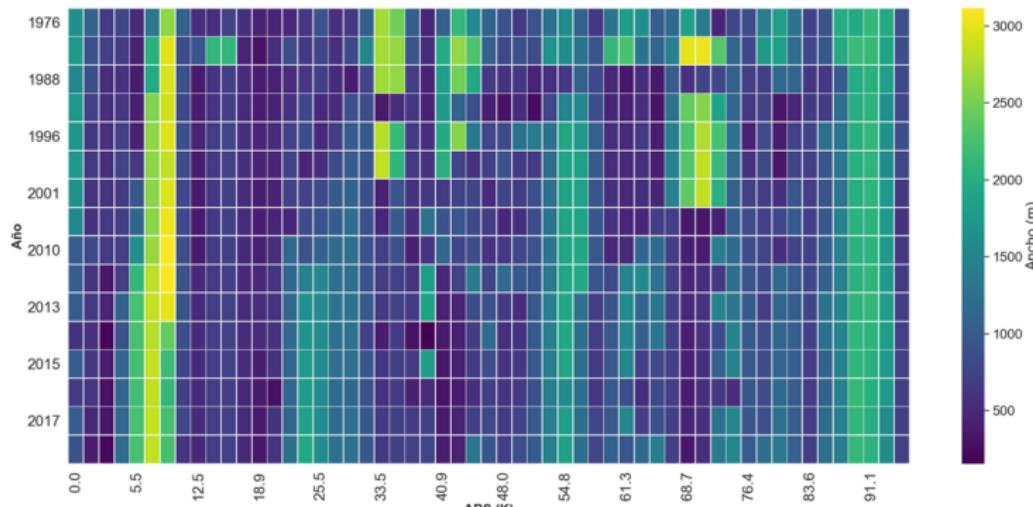


Boxplot caudal líquido para la estación Calamar (1967-2019) teniendo en cuenta año el año hidrológico en la región Caribe de Colombia. Elaboración propia



Morfología

- **Anchos:** La sección más ancha presenta una longitud de 2.7 Km en promedio, mientras que la más angosta de 418 m.
 - K68-K80 → Mayor tendencia a la sedimentación
- El año que tuvo la mayor mediana de anchos fue el de 1984 y el de menor fue el 1991.

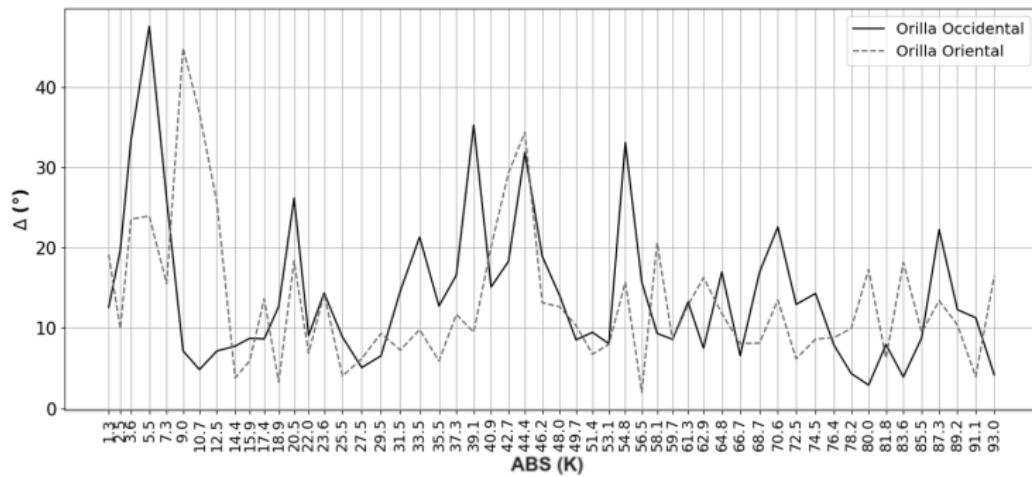


Anchos secciones de estudio para los años finales seleccionados. Elaboración propia.



Morfología

- Los mayores ángulos están en la zona de la Isla Cabica y en segundo lugar en el sector de Guáimaro. En términos generales la orilla occidental presenta ángulos mayores que la oriental, con medias globales de 13.7 y 12.9 grados respectivamente

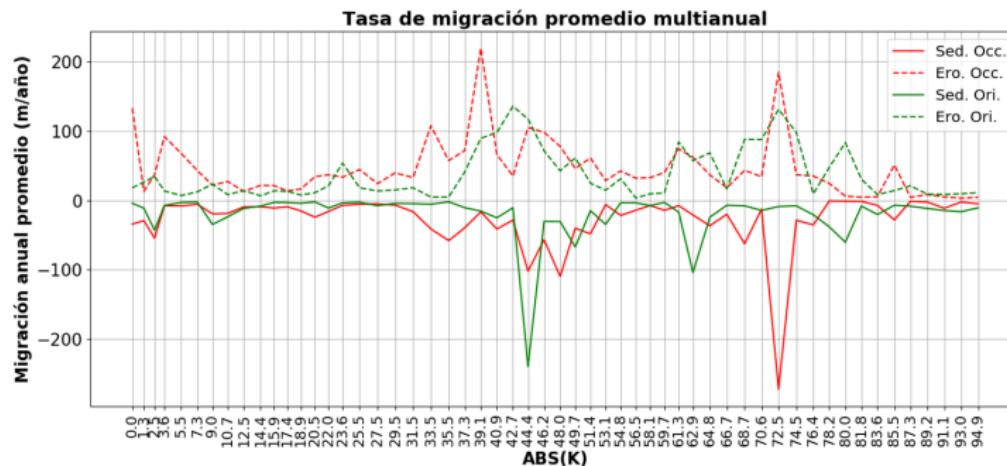


Ángulos promedio multitemporal de las secciones analizadas para el río Magdalena para ambas orillas en el tramo de estudio. Elaboración propia.



Tasas de migración lateral

- La orilla occidental presenta un mayor movimiento que la oriental en casi todo el dominio espacial analizado.
- Las mayores tasas de erosión anual promedio son de aproximadamente 220 m/año y de sedimentación son de 273 m/año.
- Orilla oriental → 12 m/año y 21 m/año (Sed-Ero)
- Orilla occidental → 16 m/año y 27 m/año (Sed-Ero)



Tasas de migración anual promedio para ambas orillas. Elaboración propia.



Aprendizaje Supervisado

■ Grupo de variables:

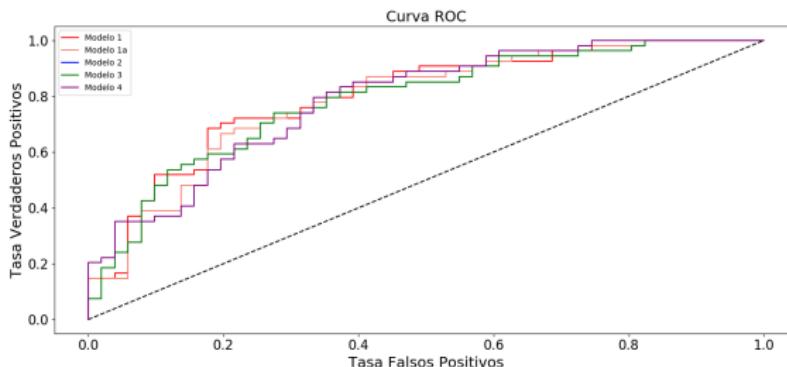
- Hidrológicas: t, t75, tdom, Qmaxdom, Qmaxmin, Probmax.
- Hídricas: Qmax/W, NDVImedian.
- Morfológicas: fid, Angle, W, resis.

Modelo	Features	Objetivo
I	t, t75, tdom, qmaxdom, probmax, fid, angle, W, NDVImedian, resis, qmaxW	Conocer la importancia de todas las variables y su desempeño en conjunto para la clasificación del fenómeno.
IA	t, t75, tdom, qmaxdom, probmax, fid, angle, resis, qmaxW	Establecer el desempeño del modelo sin la variable representativa de la vegetación (NDVI) y sin el ancho.
II	t, tdom, fid, angle, resis, qmaxW, NDVImedian	Comparar la diferencia de desempeño del modelo al solamente tener una variable que haga referencia al caudal máximo y al tiempo de excedencia de caudales.
III	t, tdom, fid, angle, qmaxW, resis	Establecer comparativa con el modelo II al no tener la variable de NDVI incluida.
IV	t, tdom, fid, angle, qmaxW	Establecer comparativa con todos los modelos al no tener las variables de resistencia relativa y de vegetación incluidas.

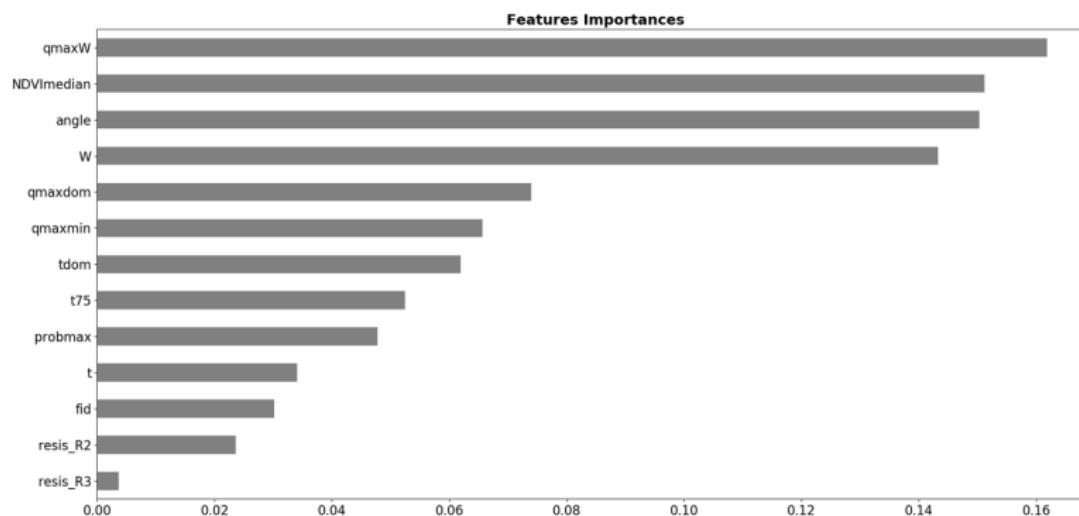


Aprendizaje Supervisado

Modelo	ROC AUC	Exactitud	Precisión	Sensibilidad	Medición F1	Promedio
I	0.80	0.74	0.75	0.74	0.74	0.754
IA	0.78	0.74	0.75	0.74	0.74	0.750
II	0.78	0.73	0.74	0.73	0.73	0.742
III	0.78	0.72	0.74	0.72	0.72	0.736
IV	0.77	0.69	0.71	0.7	0.69	0.712



Aprendizaje Supervisado



Importancia de cada variable independiente (feature importance) para la capacidad de distinción entre clases del modelo I. Elaboración propia



Aprendizaje Supervisado: Ejemplo de árbol



Conclusiones, Recomendaciones y Logros



Conclusiones

- La metodología propuesta para la clasificación de puntos erosivos o sedimentarios en las orillas puede ser aplicada en cualquier río rectilíneo o meandrífico que cuente con disponibilidad de imágenes satelitales.
- Esta metodología se convierte en una nueva herramienta para los objetivos de calibración y validación de modelos numéricos que consideren cambios morfológicos de orillas.
- Incluir variables hidrológicas, morfológicas, hidráulicas y referentes a la vegetación ribereña, dan un aporte significativo en la correcta predicción de frentes erosivos y sedimentarios.
- Esta metodología también puede ser aplicada para las fases de diseño de obras de protección de orillas y de nagevabilidad.



Recomendaciones y futuras investigaciones

- Establecer comparativa de resultados de mediano y largo plazo entre esta metodología y softwares de modelación hidrosedimentológica.
- Trabajar con registros hidrológicos no menores de 20 años y de preferencia con una frecuencia de muestreo diaria.
- Investigar la aplicabilidad de la metodología en otros tipos de ríos con una mayor presencia de islas para así establecer modificaciones que mejoren su alcance.
- Detallar la relación de cada variable a manera de pares para así establecer mejores configuraciones y proponer otras combinación de técnicas de aprendizaje supervisado para mejorar el desempeño.
- Buscar la posibilidad de utilizar imágenes satelitales con una mejor resolución. Teniendo en cuenta la combinación de bandas respectiva (de ser necesaria) para la correcta identificación de cuerpos de agua.
- Comparar los features importances resultantes de este estudio a través de técnicas locales de importancia como LIME o Shapley.



¡Muchas gracias!

