UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE INGENIERÍA CIVIL



MONITOREO DE LA CALIDAD DEL AGUA EN RIOS URBANOS BASADO EN MACHINE LEARNING Y HERRAMIENTAS GEOESPACIALES

Por

M.C. KEVIN DAVID RODRÍGUEZ GONZÁLEZ

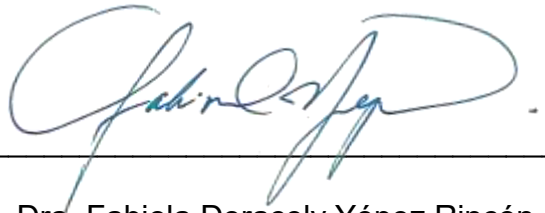
Como requisito para obtener el Grado de DOCTOR EN INGENIERÍA CON ORIENTACIÓN EN INGENIERÍA AMBIENTAL

SEPTIEMBRE, 2026

MONITOREO DE LA CALIDAD DEL AGUA EN RIOS URBANOS BASADO EN MACHINE LEARNING Y

HERRAMIENTAS GEOESPACIALES

Aprobación de la Tesis:



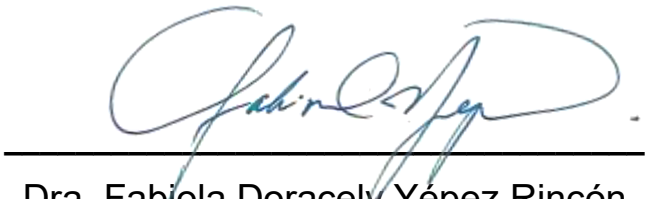
Dra. Fabiola Doracely Yépez Rincón

Jefe de la División de Estudios de Postgrado

MONITOREO DE LA CALIDAD DEL AGUA EN RIOS URBANOS BASADO EN MACHINE LEARNING Y

HERRAMIENTAS GEOESPACIALES

Dirección de tesis



Dra. Fabiola Doracely Yépez Rincón

Directora

DERECHOS RESERVADOS©

PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta Tesis está protegido, el uso de imágenes, fragmentos de videos, y demás material contenido que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde se obtuvo mencionando al autor o autores.

AGRADECIMIENTOS

TABLA DE CONTENIDO

Capítulo **Página**

I. INTRODUCCIÓN 1
1.1 Antecedentes 2
1.2 Justificación 21
2 Hipótesis 22
3 Objetivo General 23
3.1 Objetivos Específicos 23
4 METODOLOGÍA 23
4.1 Área de estudio 26
4.2 Estructura general del río Pesquería 34
4.3 Análisis de imágenes NICFI 38
4.4 Clasificación de imágenes 40
4.5 Conjunto de datos 46
4.6 Validación 48
4.7 Evaluación de precisión 54
4.8 Calidad de Agua 56
4.9 Aplicación Web 62
5 RESULTADOS Y DISCUSIONES 70
5.1 Análisis general de la cuenca del río Pesquería 72
5.2 Estructura general del río Pesquería 83
5.3 Identificación de problemáticas en el río Pesquería 88
6 Clasificación de uso de suelo 93
6.1 K-means Clustering 93
6.2 SVM 96
6.3 CART 98
6.4 RF 100
6.5 Análisis de áreas 103
6.6 Gemelo digital 3D 109
6.7 Evaluación de la precisión 111

7 Water Mask 117
7.1 Comparación de water mask 122
8 Calidad de Agua 128
8.1 Análisis de datos 128
8.2 Correlación por punto de muestreo 128
8.3 RF WQ Model 144
9 Aplicación Web 156
9.1 Plataforma Web – Pantalla principal 156
9.2 Visualización – Mapa interactivo e Índices Espectrales 158
9.3 Animación temporal Sentinel-2 y serie temporal de índices espectrales 161
9.4 Resultados de serie temporal 162
9.5 Generador de reportes PDF 163
9.6 Sección final de la página Web 164
10 LIMITACIONES Y TRABAJO A FUTURO 167
11 CONCLUSIONES 169
12 BIBLIOGRAFÍA 172
13 ANEXOS 186

LISTA DE TABLAS

Tabla **Página**

Tabla 1. Problemáticas ambientales en ríos urbanos 5
Tabla 2. Superficie a nivel estado- geología (%) 33
Tabla 3. Especificaciones NICFI 39
Tabla 4. Períodos de estiaje y lluvia por temporada 40
Tabla 5. Algoritmos de Machine Learning 45
Tabla 6. Clases de uso de suelo 45
Tabla 7. Datos iniciales 46
Tabla 8. Validación 48
Tabla 10. Ecuaciones de precisión 56
Tabla 11. Índices espectrales 56
Tabla 12. Áreas de cobertura y porcentajes de la clase agua para las 5 temporadas 107
Tabla 13. Precisión general (OA) y coeficiente de Kappa (Kc) de los modelos RF, SVM y CART para la clasificación LULC estacional (2016–2019). 112
Tabla 14. Precisión del usuario (UA), precisión del productor (PA) y puntuación F1 para la clasificación en Bosque Aleatorio (RF) por clase y temporada de cobertura del suelo (2016–2019). 113
Tabla 15. Correlación por punto de muestreo y parámetro por bandas espectrales 128
Tabla 16. Correlación por punto de muestreo y parámetro por índices espectrales 130
Tabla 17. Correlaciones promediadas 131
Tabla 18. Mayor correlación por punto y por predictor espectral 132

LISTA DE FIGURAS

Figura **Página**

Figura 1. Río urbano 3
Figura 2. Clasificación de gestión en cuencas hidrográficas 7
Figura 3. (A) Estructura física del drone. (B) Imagen en planta obtenida con el drone. (C) Imagen del río obtenida por el drone 11
Figura 4. Imagen Sentinel-2 13
Figura 5. Imagen multiespectral 15
Figura 6. Espectro electromagnético 16
Figura 7. Descripción general GEE 19
Figura 8. Metodología general. 25
Figura 9. Área de estudio y sitio de experimentación 26
Figura 10. Datos climáticos en Nuevo León y en la cuenca Pesquería 28
Figura 11. Fisiografía de México y la cuenca del río Pesquería según clasificación de INEGI 30
Figura 12. Sierra Madre Oriental y sus usos de suelo 31
Figura 13. Crecimiento poblacional de México, Nuevo León y la cuenca Pesquería 34
Figura 14. (A) Imagen compuesta NIR+GB. (B) Índice NDVI 47
Figura 15. Mapa de Voronoi 49
Figura 16. Imágenes de validación (A) Imagen Google Earth Pro. (B) Google Street View. (C) Imagen satelital Sentinel-2 50
Figura 17. Puntos de validación con Google Earth Pro 51
Figura 18. Validación 52
Figura 19. Random points 53
Figura 20. Código Random points en GEE 54
Figura 21. Flujo de desarrollo de la plataforma Web: computadora local - repositorio GitHub -Streamlit Community 62
Figura 22. Script base de prueba local de la aplicación 63
Figura 23. Modulo idiomas 65

Figura 24. Pipeline completo del modelo RF desde los datos de campo hasta la predicción en tiempo real para el usuario de la plataforma web. Los archivos. pkl serializados actúan como puente entre el entrenamiento local y la app desplegada en la nube 66
Figura 25. Import pickel serialización 67
Figura 26. Estructura de archivos del repositorio GitHub 68
Figura 27. Serie Temporal CHRIPS, LST, NDVI y T2M 74
Figura 28. Subíndices espectrales para calcular ICU 76
Figura 29. ICU cuenca Pesquería 78
Figura 30. 3D ICU 80
Figura 31. IDAU 82
Figura 32. Ortomosaico río Pesquería 84
Figura 33. Perfil del cauce (A) Perfil topográfico 1y (B) Perfil topográfico 2 86
Figura 34. Modelos digitales de elevación (DEMs) y su correspondiente modelo 3D derivado de la fotogrametría basada en UAV. Los ejemplos anteriores proceden del río Pesquería y corresponden a los perfiles topográficos de las Secciones A y B. 87
Figura 35. Problemáticas río Pesquería. (A) Nube de puntos LiDAR (B) Asentamientos irregulares cercanos a 1m del río Pesquería 89
Figura 36. Represas antiguas del río Pesquería (A) Imagen satelital de Google Earth Pro de abril de 2009 y (B) Imagen satelital de Google Earth Pro de marzo de 2016 91
Figura 37. K-means clustering 94
Figura 38. SVM 97
Figura 39. CART 99
Figura 40. Random Forest 101
Figura 41. Respuesta algoritmo RF 103
Figura 42. LULC variación de áreas en diferentes temporadas 104
Figura 43. Diferentes estaciones y clases representadas con un radar estandarizado por clase basado en la superficie estimada del algoritmo RF . 106

Figura 44. (a) Clases LULC de correlación cruzada. (b) Diversidad ecológica (Shannon, Simpson, Evennes). 108
Figura 45. Gemelo digital 3D del río Pesquería 110
Figura 46. Water Mask 2016-2019..... 118
Figura 47. Escasez de agua Water mask 119
Figura 48. Área de la clase agua 121
Figura 49. Comparación extracción agua 124
Figura 50. Error MNDWI 126
Figura 51. P_TOT vs. combinación de bandas (B3+B5)/(B1+B4) en el punto P3 134
Figura 52. Nitrógeno total vs. banda B3 (verde) en el punto..... 134
Figura 53. Nitrógeno total vs. banda B1en el punto P7..... 135
Figura 54. Nitrógeno amoniacal vs. banda B1 en el punto P6 135
Figura 55. Nitrógeno total Kjeldahl vs. B4/(B11+B12) en el punto P3..... 136
Figura 56. Nitrógeno orgánico vs. NDCI (Normalized Difference Chlorophyll Index) en el punto P1 136
Figura 57. DQO vs. FAI en el punto P1 137
Figura 58. Sólidos suspendidos totales vs. NDCI en el punto P7 137
Figura 59. Heat Map de Pearson para las 20 variables espectrales de banda frente a las 12 combinaciones de transformación evaluadas 138
Figura 60. Heat Map de Pearson para los 7 índices espectrales frente a las combinaciones de transformación evaluadas 139
Figura 61. Importancia de Features 142
Figura 63. P_TOT 147
Figura 64. N_TOT 148
Figura 65. N_Kjeldahl 148
Figura 66. Calidad de Agua 149
Figura 67. Series temporales de predicción..... 151
Figura 68. Q-Q Normalidad..... 153
Figura 69. Pantalla principal (Hero Section) de la plataforma Web Water Quality Mapping en funcionamiento 157

Figura 70. Hero section en portugués y en inglés 158
Figura 71. Módulo de visualización del área de estudio página Web 159
Figura 72. Visualización de área de estudio aplicando clasificación de uso de suelo y mapa base satelital 160
Figura 73. Sección de animación temporal Sentinel-2 (GIF) y serie temporal de índices espectrales 161
Figura 74. Serie temporal del NDWI generada automáticamente por la plataforma para el río Pesquería con 223 imágenes Sentinel-2 (enero 2019 – junio 2026) 162
Figura 75. Generación de reportes PDF para Calidad de Agua y para Índices espectrales 164
Figura 76. Interfaz Final 165
Figura 77. Código QR página web 166

LISTA DE ECUACIONES

Ecuación 1. ICU 37
Ecuación 2. IDAU 38
Ecuación 3. Datos de entrenamiento etiquetados 41
Ecuación 4. Hiperplano óptimo 41
Ecuación 5. Tarea de optimización 42
Ecuación 6. Jerarquia modelo CART 43
Ecuación 7. probabilidad RF 44
Ecuación 8. Precisión general 55
Ecuación 9. Índice Kappa 55
Ecuación 10. Yeo-Johnson 59
Ecuación 11. Matriz de correlación de Pearson 60
Ecuación 12 Normalización z-core 60

NOMENCLATURA

- SIG: Sistema de Información Geográfico
- AMM: Área Metropolitana de Monterrey
- CONABIO: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad
- CSV: Comma-Separated Values
- NICFI: Norway's International Climate and Forest Initiative
- ODS: Objetivos de Desarrollo Sostenible
- INEGI: Instituto Nacional de Estadística y Geografía
- CONAGUA: Comisión Nacional del Agua
- NDVI: Por sus siglas en inglés, Normalized Difference Vegetation Index, Índice de Vegetación de Diferencia Normalizada
- NDBI: Por sus siglas en inglés, Normalized Difference Built-up Index, Índice de Diferencia Normalizada de Áreas Construidas
- NDWI: Por sus siglas en inglés, Normalized Difference Water Index, Índice de Diferencia Normalizada de Agua
- MNDWI: Por sus siglas en inglés,Modified Normalized Difference Water Index, Índice de Diferencia Normalizada de Agua Modificado
- AWEI: Automated Water Extraction Index
- NDTI: Por sus siglas en inglés, Normalized Difference Turbidity Index, Índice de Diferencia Normalizada de Turbidez
- NDCI: Por sus siglas en inglés, Normalized Difference Chlorophyll Index, Índice de Diferencia Normalizada de Clorofila
- FAI: Por sus siglas en inglés, Floating Algae Index, Índice de Algas Flotantes
- SABI: Surface Algal Bloom Index
- CDOM: Colored (o Chromophoric) Dissolved Organic Matter
- EVI: Enhanced Vegetation Index
- JRC: Joint Research Centre
- MSI: Por sus siglas en inglés, MultiSpectral Instrument, Instrumento Multiespectral

- PTAR: Planta de Tratamiento de Aguas Residuales
- NOM: Normas Oficiales Mexicanas
- DQO: Demanda Química de Oxígeno
- SST: Sólidos Suspendidos Totales
- OD: Oxígeno Disuelto
- SINA: Sistema Nacional de Información del Agua
- REMANECA: Red Nacional de Medición de Calidad del Agua
- pH: Potencial de Hidrógeno
- GEE: Google Earth Engine
- NIR: Por sus siglas en inglés, Near-Infrared, Infrarrojo cercano
- RF: Random Forest
- CART: Por sus silgas en ingles, Classification and Regression Trees
- SVM: por sus siglas en inglés, Support Vector Machine
- LIDAR: Por sus siglas en inglés, Light Detection and Ranging, Detección de luz y rango
- LAS: Lidar Aerial Survey
- RTK: Real Time Kinematics
- LULC: Por sus siglas en ingles, Land Use Land Cover
- ANP: Área Natural Protegida
- UANL: Universidad Autónoma de Nuevo León
- ML: Machine Learning
- IA: Inteligencia Artificial
- OLI: Por sus siglas en inglés, Operational Land Imager
- RADAR: Por sus siglas en inglés, Radio Detection and Ranging
- SAR: Synthetic Aperture Radar
- NASA: National Aeronautics and Space Administration
- ESA: por sus siglas en inglés, European Spacial Agency
- USGS: Servicio Geológico de los Estados Unidos

- MODIS: Moderate Resolution Imaging Spectroradiometer
- NRT: por sus siglas en inglés, Near Real Time
- API: Application Program Interface
- OTOP: Offline Training, Online Prediction / Inference
- DL: Deep Learning
- Chl: Clorofila-a
- OD: Oxígeno disuelto
- EDA: Por sus siglas en inglés, Empirical Data Analysis
- VANT: Vehiculo Áereo No tripulado
- CHIRPS: Climate Hazards Group InfraRed Precipitation with Station data
- LST: por sus siglas en inglés, Land Surface Temperature
- GHSL: Global Human Settlement Layer
- SURB: Subíndice Urbano
- SVEG: Subíndice de Vegetación
- SHID: Subíndices hídrico
- SCLI: Subíndice climático
- ICU: Índice de Convergencia Urbana
- IDAU: Índice de Degradación Ambiental Urbana
- SIMA: Sistema Integral de Monitoreo Ambiental.
- DQO TOT: Demanda Química de Oxígeno Total
- N NH3: Nitrógeno Amoniacal
- N NO2: Nitrógeno de Nitritos
- N NO3: Nitrógeno de Nitratos
- N ORG: Nitrógeno Orgánico
- N TOT: Nitrógeno Total
- N TOTK: Nitrógeno Total Kjeldahl (suma del nitrógeno orgánico y amoniacal)
- OD: Oxígeno Disuelto

- P TOT: Fósforo Total
- pH_CAMPO: Potencial de Hidrógeno (pH) medido directamente en sitio/campo
- SDT: Sólidos Disueltos Totales
- SST: Sólidos Suspendidos Totales
- TURB: Turbidez
- IDE: Por sus siglas en inglés, Integrated Development Environment
- HTML: Por sus siglas en inglés, Hyper Text Markup Language
- CSS: Por sus siglas en inglés, Cascading Style Sheets
- RBF: Por sus siglas en inglés, Radial Basis Function
- IDW: Por sus siglas en inglés, Inverse Distance Weighting
- IWQP: Index Water Quality Points
- BWQP: Bands Water Quality Points
- RPD: Ratio of Performance to Deviation

I. INTRODUCCIÓN

Los ríos urbanos son ecosistemas cruciales para el bienestar de las personas que viven en las ciudades. Estos cuerpos de agua proporcionan una serie de servicios y beneficios esenciales, incluyendo agua potable, alimentos, energía, recreación y transporte. Además, tienen un valor cultural y estético, y pueden actuar como corredores verdes/azules que mejoran la calidad del aire y reducen la temperatura de las áreas urbanas (Herrera et al., 2016). Dada la importancia de los ríos urbanos y de la calidad del agua para garantizar la salud y el bienestar de las comunidades que viven cerca de ellos, es trascendental mantener estos ecosistemas. La calidad de los ríos urbanos puede verse afectada por la contaminación causada debido a las actividades humanas como la descarga de aguas industriales, aguas residuales sin tratar, residuos sólidos urbanos, agricultura intensiva entre otros. La contaminación del agua de los ríos puede tener graves consecuencias para la salud pública y el medio ambiente, como la propagación de enfermedades y la muerte de especies animales y vegetales, entre otros (Wan & Wang., 2021).

Con base en lo anterior, muchas de las problemáticas en los ríos urbanos son consecuencia de su cercanía con las actividades antropogénicas que suceden diariamente en cualquier ciudad. Para evaluar la contaminación de los ríos urbanos usualmente se usan mediciones in-situ que en general suelen ser muy precisas, pero requieren de mucho tiempo para la obtención de datos, además suelen ser muy costosas y proporcionan información de puntos específicos que carecen de detalles espacio temporales y de cobertura (Duan et al., 2012; Duan et al., 2013; Shelare et al., 2021). Debido a estas limitaciones, las mediciones tradicionales carecen de representatividad general de los ríos urbanos y más cuando estos son muy extensos. Por su parte, las estrategias de monitoreo de calidad ambiental usando tecnología satelital y herramientas geoespaciales son más eficaces para la obtención de datos de forma extensiva y constante, con lo cual se podrían evaluar y actualizar indicadores de calidad del agua de los ríos (Chen et al., 2021) como el indicador 6.3.2 de los ODS (UN, 2020).

Debido al alto volumen de datos que pueden obtenerse con las herramientas satelitales, se hace conveniente la aplicación de tecnologías espaciales e inteligencia artificial para procesar y analizar la calidad ambiental de ecosistemas naturales como lo son los ríos urbanos. Los modelos de ML aplicados a cuerpos de agua, son un subconjunto de la IA que permite dar seguimiento a las problemáticas de calidad del agua, a través del aprendizaje automatizado de datos, por medio de ejemplos y experiencias predefinidas reconociéndose como una tecnología potencial que jugará un papel clave para la prevención del impacto a los ecosistemas riparios (IWA, 2019; Butler et al., 2016).

Es por lo anterior, que el presente estudio busca aportar soluciones para mejorar la gestión ambiental de los ríos urbanos, a través del desarrollo de un modelo basado en ML para estimar y categorizar la calidad de agua de los ríos urbanos. Este estudio es clave para abordar los desafíos que contempla la gestión ambiental en los cuerpos receptores de agua y a su vez en el cumplimiento de los ODS.

1.1 Antecedentes

1.1.1 Ríos urbanos

Los ríos constituyen un sistema de circulación lineal, y jerarquizado que se encarga de trasladar sedimentos y fluidos vitales a través de las cuencas y sus desembocaduras, realizando complejas reacciones dinámicas, mecánicas, energéticas, químicas y bioquímicas con el propósito de dar sustento en todo su recorrido a la vida en sus diferentes formas (Machado et al., 2018). De igual manera el río se conoce como una corriente natural de agua de flujo continuo y constante, dotado de caudal y velocidad siendo parte complementaria del ciclo del agua en el planeta.

Aunado a lo anterior, los ríos urbanos de igual manera representan un margen importante en el desarrollo de las ciudades, puesto que son ecosistemas dinámicos en donde la flora y la fauna juegan un papel importante en la calidad

del medio ambiente. Estos cuerpos de agua deberían ser considerados hábitats preferenciales para la recreación de la sociedad, protección de la naturaleza y la biodiversidad, control climático y, sobre todo seguridad a la población frente a las amenazas naturales tales como inundaciones y otros efectos del cambio climático (Suzianti et al., 2023). A continuación, en la Figura 1 se evidencia la representación gráfica de un río urbano.



Figura 1. Río urbano

Fuente: Pixabay

Se conoce que las ciudades se empezaron a desarrollar cerca de los ríos debido a las principales necesidades del ser humano en este recurso y como un medio de transporte, de esta forma al momento que se expande la población aumentan las problemáticas en los ríos urbanos y de aquí la importancia del cuidado de cada uno de ellos. Gastezzi y colaboradores (2016) afirman que los ríos urbanos ofrecen muchos beneficios ecológicos dentro de una sociedad, desarrollo económico de las ciudades, disminución de ciertas problemáticas

como control de escorrentías y en especial, como mitigadores naturales de los efectos del cambio climático mitigando el efecto devastador de inundaciones, reducción de altas temperaturas o incluso como sumideros de carbono. Estos cuerpos receptores de agua deberían ser considerados hábitats preferenciales para la recreación de la sociedad, protección de la naturaleza y la biodiversidad, sin embargo, las condiciones ambientales actuales de los ríos urbanos en general no son las mejores, debido a que se encuentran altamente contaminados por la mala gestión local y la falta de conciencia en la sociedad (Xia et al., 2020). La recuperación de ríos urbanos como un espacio público, es un trabajo arduo que implica compromiso y voluntad para lograr ríos saludables y resilientes, lo que a su vez podría generar oportunidades para el desarrollo sostenible de muchas ciudades.

1.1.2 Problemáticas ambientales en ríos urbanos

Históricamente en América Latina y el Caribe se presenta muy poca importancia a la protección de los ríos urbanos y los esfuerzos por detener este deterioro no son suficientes. No obstante, muchos países cuentan con la legislación necesaria para el manejo de la contaminación del agua, aunque existe una enorme dispersión de responsabilidades para su aplicación y consecuentemente no se cumple. Por otra parte, El rápido crecimiento de las ciudades ha tenido un impacto significativo sobre los ríos que atraviesan estos entornos urbanos. En muchos casos, este desarrollo no controlado ha llevado a una degradación hidrológica y ecológica de los cursos de agua. El proceso de urbanización, si no va acompañado de una adecuada gestión ambiental, altera las condiciones naturales y el valor intrínseco de los cuerpos receptores de agua.

Este efecto se ve intensificado por otros importantes factores de impacto, como la contaminación y las alteraciones en los caudales (Quesada et al., 2021). La combinación de estos diferentes procesos de deterioro genera una problemática integral en torno a los ríos urbanos, afectando tanto sus características físicas como su calidad ambiental. La Tabla 1 a continuación

resume las principales problemáticas y tipos de contaminación que suelen afectar a estos ecosistemas fluviales insertos en los entornos urbanos de la región.

Tabla 1. Problemáticas ambientales en ríos urbanos

N°	**Tema**
1	Vertimientos directos de aguas domésticas e industriales
2	Uso masivo de fertilizantes y agroquímicos en actividades agrícolas
3	Crecimiento demográfico y expansión urbana no planificada
4	Mala gestión de residuos sólidos
5	Mayor presión humana
6	Disminución de la flora y fauna de zonas aledañas al río

Fuente: Gastezzi, (2016)

La Tabla 1 permite identificar las diversas problemáticas ambientales que afectan a la mayoría de los ríos urbanos. Estas problemáticas se relacionan principalmente con la contaminación de las aguas de los ríos, así como con los impactos sobre la flora, la fauna de estos ecosistemas fluviales, salubridad y todas las afectaciones secundarias y sociales que cause la contaminación de los ríos urbanos. De aquí radica la importancia de preservar y proteger estos valiosos ríos que atraviesan las áreas urbanas.

Aunado a lo anterior, para disminuir las problemáticas mencionadas se debe tener una buena gestión integral de los ríos urbanos, para así obtener herramientas que ayuden a la toma de decisiones y de esta forma disminuir la contaminación en estas zonas. Uno de los principales retos de los organismos gestores del agua para dar seguimiento al monitoreo y control de la contaminación de ríos urbanos es que, no se han podido establecer herramientas de monitoreo continuo para las fuentes de dichos contaminantes (Zingaro et al., 2019).

1.1.3 Gestión integral de los ríos urbanos

La gestión integral de ríos urbanos es un proceso flexible, participativo e iterativo que integra los componentes relacionados con el ciclo de aguas urbanas (suministro de agua, gestión de aguas pluviales, y gestión de residuos) con el desarrollo urbano de la ciudad y la gestión de la cuenca fluvial con la finalidad de maximizar los beneficios económicos, sociales y medioambientales de manera equitativa, de tal forma que este conjunto de actividades demuestre cómo se debe salvaguardar la calidad de los ríos urbanos (De Cabo & Marconi., 2021).

El tema de la gestión de cuencas, por otro lado, está siendo cada vez más aceptado en la región, pero a pesar de los avances logrados aún no hay un consenso generalizado sobre definiciones que precisan los objetivos de dicha gestión. La falta de claridad conceptual en la materia atenta contra el intercambio de ideas y experiencias crea cruces de funciones entre instituciones y dificulta la formulación de políticas y leyes claras sobre el tema (Martínez & Villalejo., 2018). Sin embargo, se han tenido en cuenta como base algunas temáticas principales para la gestión del recurso hídrico:

• Proteger el recurso hídrico (río).

• Prevención de acciones que aumenten la contaminación de la zona.

• Monitoreo de la calidad de agua en diferentes puntos del río.

• El mejoramiento de su calidad hidro ecológica y, por consiguiente, del “estado de salud” del ecosistema hídrico.

De esta manera el agua debe dejar de ser vista como un recurso, de manera que se le reconozcan valores funcionales de sustento para los ecosistemas, mejorar la calidad ambiental y como perspectiva para el desarrollo urbano. A continuación, en la Figura 2 se observa un ejemplo de clasificación de gestión en cuencas hidrográficas.

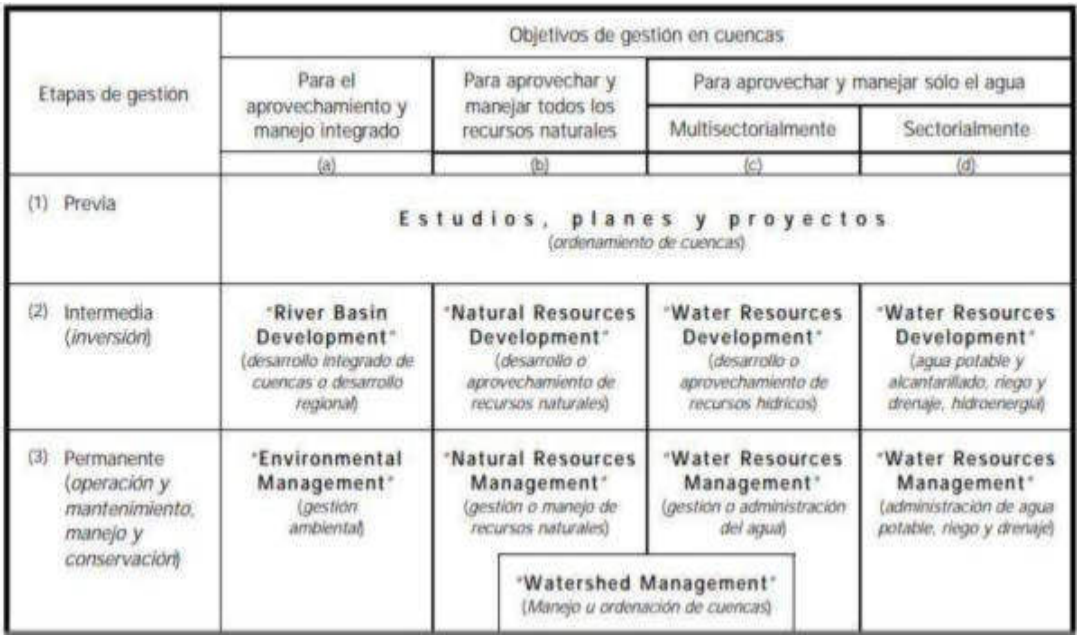


Figura 2. Clasificación de gestión en cuencas hidrográficas

Fuente: CEPAL

La Figura 2 muestra un esquema general para la gestión de cuencas en sectores urbanos o rurales, cabe resaltar que como este pueden existir diferentes alternativas para hacer una excelente clasificación y gestión de los recursos hídricos y que este esquema es muy general para todo lo que abarca una excelente gestión integral de los recursos hídricos.

Últimamente, además de la importancia de conocer todas las aplicaciones para la gestión integral de ríos urbanos, se encuentra en constante investigación aquellas tecnologías que mejoren y faciliten dicha gestión. De acuerdo con esto, se hace interesante conocer e investigar las tecnologías que puedan proporcionar una información continua y detallada de todos los procesos de calidad y monitoreo aplicados en la gestión integral de los ríos urbanos (Sun et al., 2022).

1.1.4 Sensores remotos

Los primeros satélites de observación empezaron a operar a principios de los 70s. Desde entonces el sistema científico-técnico ha puesto a disposición de la sociedad herramientas y modelos conceptuales que traducen los datos

registrados por los sensores en conocimiento útil para la toma de decisiones en múltiples áreas incluyendo las ciencias ambientales (Bauer., 2020). Los sensores remotos son sistemas empleados para recolectar información rápida y viable de un objeto, área o fenómeno a distancia. Dicha información puede ser interpretada mediante el uso de SIG con la finalidad de tener una mejor comprensión de datos.

Los sensores remotos instalados en drones han ganado gran popularidad en los últimos años por su versatilidad y eficacia. Según Herrera et al. (2018) estos dispositivos se encargan de registrar la energía electromagnética emitida o reflejada por la superficie de un objeto en distintas longitudes de ondas (según la capacidad del sensor). Los registros obtenidos por estos sensores incluyen estas longitudes desde las más cortas (gamma y radiográficas) hasta las más largas (microondas y de radiodifusión). La información va más allá del alcance del ojo humano, y es obtenida mediante la interacción de los principales elementos: sensor, objeto de estudio, atmosfera y flujo energético (Veneros et al., 2020).

Numerosos estudios han demostrado la utilidad de los sensores remotos en el ámbito ambiental. Por ejemplo, Bonansea (2019) implementó un modelo basado en sensores remotos para evaluar la calidad del agua en un embalse localizado en la provincia de Córdoba, Argentina. Asimismo, Briceño et al. (2018) por su parte, utilizaron el sensor OLI de Landsat 8 para estudiar la distribución espacial y la dinámica de los parámetros de calidad de agua del lago Vichuquén, Chile, durante el año 2016. De la misma manera Shi et al. (2019) utilizó datos satelitales de Landsat y Sentinel para clasificar el uso y la cobertura de suelo y agua en el centro urbano de Guangzhou- China. Gracias a estos sensores lograron clasificar zonas comerciales, de educación, vegetación, residenciales y cuerpos de agua en la zona urbana (ríos), obteniendo regresiones aproximadas entre el 57-80% para cada una de sus clases.

La tecnología satelital, aunque es invaluable para la accesibilidad de información, requiere un riguroso preprocesamiento debido a los desafíos recurrentes en la resolución de imágenes causados por condiciones ambientales, como la nubosidad (Osco et al., 2021). En muchos casos, esto conduce a la

omisión de ciertas imágenes. Estas situaciones podrían mitigarse mediante el empleo de sensores remotos, como el RADAR, que permiten obtener resultados independientemente de las condiciones climáticas, o mediante la captura de imágenes de mayor resolución espacial.

1.1.5 Vehículos aéreos no tripulados (VANT)

Abarcando la historia de los VANT (drones), estos se empezaron a implementar desde el siglo XIX con fines agrícolas utilizando globos para tomar fotografías y cámaras aéreas montadas en el pecho de las palomas en 1903 según Jawalekar et al., (2020). Se cree que los drones se originaron a partir de los torpedos aéreos generados en la primera guerra mundial. Actualmente los avances tecnológicos permiten que mediante sensores avanzados y tecnologías VANT se puedan obtener ortofotografías y modelos digitales de un espacio delimitado con una alta resolución, que se utilizan para los estudios geológicos, ambientales, sociales y el desarrollo de trabajos científicos. Los VANT se han consolidado como una tecnología clave para la captura de imágenes multiespectrales de alta resolución. Su principal ventaja radica en la capacidad de adaptar la planificación del vuelo a las condiciones específicas del área de estudio. Esto permite optimizar variables operativas como el horario de los sobrevuelos, la identificación de obstáculos, la altura, el solapamiento fotográfico, la gestión de la energía y la delimitación del terreno (Fernández y Gutiérrez., 2016).

Pino et al. (2019) afirma que la clasificación de los drones se da a partir de diferentes aspectos como el uso militar y civil (generalmente). Debido al tipo de control que utilizan ya sea autónomo, monitorizado, supervisado, pre-programado y controlado remotamente, así como también por la forma que presentan: multirrotores, híbridos y ala fija. Todos se encuentran en constantes mejoras tecnológicas con la finalidad de superar sus desventajas. Los VANT suelen contar con sistemas de geolocalización por satélite como el Sistema de Posicionamiento Global (GPS por sus siglas en inglés), que permite conocer su posición y controlar su estabilidad.

Dentro de las funcionalidades de los VANT, el plan de vuelo es la ruta o trayectoria que, de forma automática, debe seguir un drone, para la obtención de la información del objeto de estudio. Los drones en su mayoría, dependiendo de la investigación que realicen, pueden ser operados a través de un control remoto (manual) o de forma automática por medio de un plan de vuelo pre-establecido por softwares. Un buen resultado de una investigación con VANT depende de la buena planificación en la misión de vuelo. Esto implica la selección adecuada del sensor, las condiciones temporales para realizar el sobrevuelo, la gestión de la energía, y la definición de parámetros fotogramétricos, tales como la altura, el tiempo de vuelo, la escala, el número de fotogramas y la delimitación y seguridad de las zonas de despegue y aterrizaje (Duffy et al., 2018). A continuación, la Figura 3 ilustra la estructura física de estos equipos, así como un ejemplo de las imágenes capturadas.





Figura 3. (A) Estructura física del drone. (B) Imagen en planta obtenida con el drone. (C) Imagen del río obtenida por el drone

Fuente: Tumblr.

Por consiguiente, se han realizado diversas investigaciones implementando estas tecnologías, uno de ellos fue Elijah et al. (2018) utilizaron VANT junto con un sensor de agua que brinda información de parámetros fisicoquímicos enviando la información a una estación base y esta al ser procesada llega a las plataformas de celulares y computadoras (Parámetros fisicoquímicos del agua en tiempo real). Liu et al. (2020) por su parte, propusieron un método de monitoreo de calidad de agua por medio de drones realizando observaciones multitemporales en ríos urbanos y mostraron así los cambios de calidad de agua en el eje del tiempo utilizando teledetección multiespectral. Se han usado también drones con sensores LiDAR los cuales han logrado generar modelos de elevación, perfiles topográficos, intensidad de nube de puntos ortomosaicos de alta resolución con una precisión centimétrica la cual genera una alta confiabilidad del uso de sus datos en cualquier aplicación geoespacial regional o local (Vega et al., 2025: Rodríguez et al., 2026).

De esta manera se observa la importancia y necesidad de aplicación de estas tecnologías en los sistemas de manejo de información.

1.1.6 Imágenes satelitales

Las imágenes satelitales han sido fundamentales en la exploración y el estudio del planeta desde la era espacial. Comenzando con los primeros satélites de observación de la Tierra en la década de los 60´s, la captación de imágenes desde el espacio ha evolucionado significativamente a lo largo de los años. Los primeros satélites, como Landsat 1 lanzado en 1972, ofrecían una resolución espacial relativamente baja, con píxeles de alrededor de 80 metros (NASA, 2024). Sin embargo, a medida que la tecnología avanzó, se desarrollaron nuevos sensores y plataformas satelitales que mejoraron drásticamente la calidad y la utilidad de las imágenes.

Uno de los aspectos clave en la evolución de las imágenes satelitales ha sido el aumento de la resolución, la disminución de la nubosidad y el aumento de las bandas espectrales. Actualmente, satélites de observación de la Tierra, como los de la misión Sentinel de la ESA y los satélites Landsat operados por la NASA y el USGS, pueden capturar imágenes con resoluciones de hasta 10 y 30 metros por píxel respectivamente. Esta mejora en la resolución ha permitido un análisis más detallado y preciso de las características terrestres, facilitando aplicaciones como el monitoreo de cambios en la cobertura y uso del suelo, la evaluación de impactos ambientales, la planificación y gestión urbana, el seguimiento de actividades agrícolas y la detección de cambios en cuerpos de agua, como ríos y lagos (Carroll, 2022).

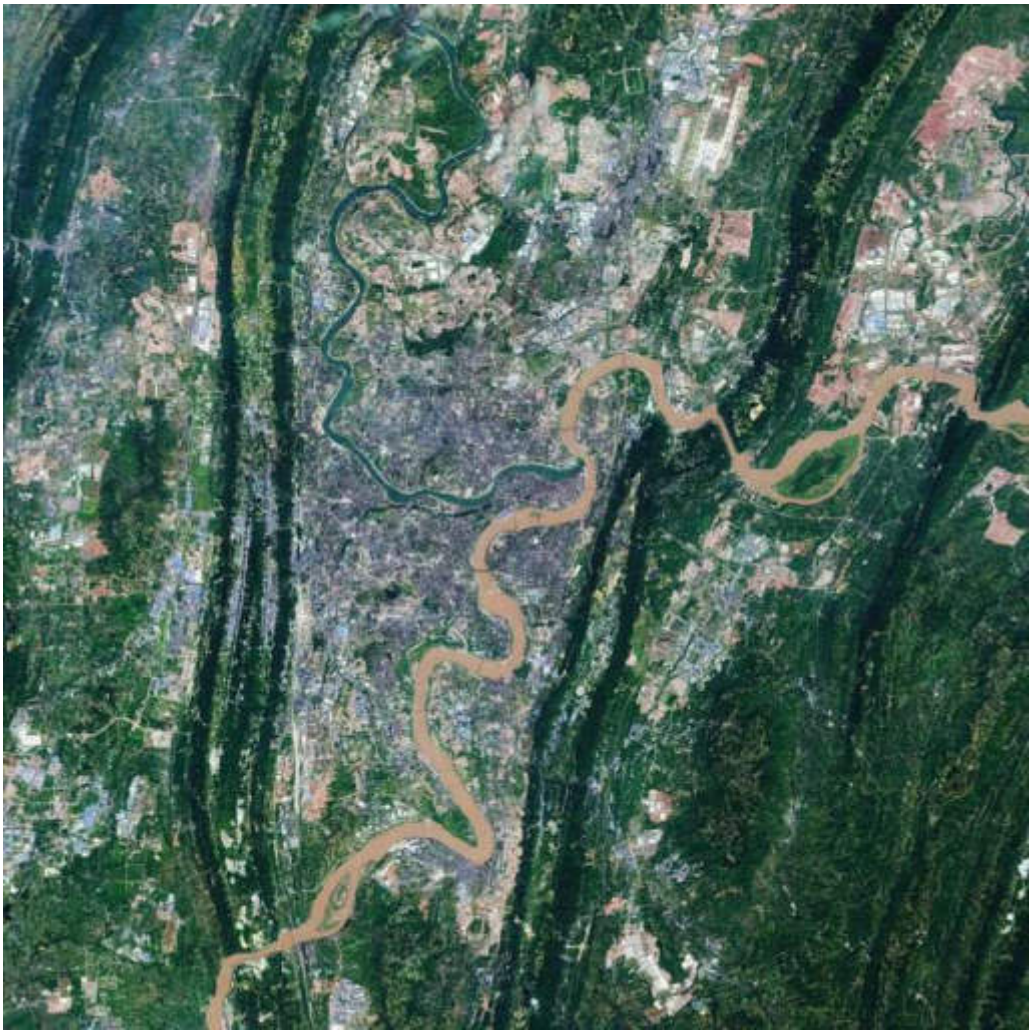


Figura 4. Imagen Sentinel-2

Fuente: ESA, río Yangtse de Chongqing, China

Algunas de las misiones de imágenes satelitales más importantes y utilizadas en la actualidad para monitorear y proteger los bosques tropicales (Vizzari, 2022) son la constelación de satélites Sentinel-2 de la ESA, los satélites Landsat (7 y 8) y la iniciativa noruega NICFI, que utiliza imágenes de alta resolución de Planet Labs. Estas misiones han sido fundamentales para avanzar en nuestra comprensión y gestión del medio ambiente, incluidos los estudios sobre ríos urbanos y otros recursos hídricos. Es importante señalar que, aunque existen otras misiones con una resolución superior que pueden proporcionar un análisis

más detallado, su costo elevado dificulta su aplicación en las áreas que más requieren este tipo de apoyo.

1.1.7 Imágenes multiespectrales

Una imagen multiespectral es aquella que captura datos espaciales registrando la reflectancia o emitancia en rangos específicos de longitud de onda a lo largo del espectro electromagnético, abarcando generalmente desde el espectro visible (aproximadamente 400 nanómetros) hasta el infrarrojo de onda corta (alrededor de 2500 nanómetros), e incluso extendiéndose al infrarrojo térmico. Para lograr esto, los sensores utilizan filtros ópticos o detectores especializados que separan la radiación entrante en distintas bandas espectrales. Comúnmente, estas adquisiciones incluyen las bandas del espectro visible (Azul, Verde y Rojo), así como regiones imperceptibles para la visión humana, tales como el Infrarrojo Cercano (NIR) y el Infrarrojo de Onda Corta (SWIR). El registro simultáneo en estas múltiples bandas permite extraer información biofísica y radiométrica detallada de las cubiertas terrestres, superando las limitaciones de los receptores biológicos RGB del ojo humano (Li & Liu., 2019).



Figura 5. Imagen multiespectral

Fuente: Flickr

En este contexto, la integración de cámaras multiespectrales proporciona imágenes de alta resolución espacial con información digital georreferenciada y de disponibilidad inmediata. Esta tecnología permite el uso de diversas cargas útiles, abarcando desde cámaras ópticas RGB y térmicas, hasta sensores multiespectrales y en algunas ocasiones hiperespectrales que amplían significativamente la capacidad de adquisición de datos ambientales (Fernández et al., 2016).

Respecto a sus aplicaciones prácticas, diversos autores han demostrado su eficacia en ecosistemas acuáticos. Por ejemplo, De Giglio et al. (2020) utilizaron VANT equipados con sensores multiespectrales abarcando las regiones ultravioleta e infrarroja para identificar residuos plásticos en entornos fluviales.

Esta metodología permitió desarrollar mapas temáticos que clasifican la cantidad, tipología y el estado de degradación de dichos contaminantes. De manera complementaria, Choo et al. (2018) implementaron técnicas de percepción remota mediante VANT para evaluar la calidad del agua en el río Nakdong, utilizando la concentración de clorofila-a como bioindicador a través del análisis de ortomosaicos y el cálculo del NDVI.

Estos antecedentes reafirman la importancia de emplear imágenes multiespectrales adquiridas por VANT para la gestión, el monitoreo y la conservación de cuerpos receptores de agua (tales como ríos urbanos, lagos y embalses) y del ecosistema en general. Para contextualizar los principios físicos detrás de estas capturas, la siguiente figura ilustra la distribución del espectro electromagnético y sus correspondientes longitudes de onda. A continuación, en la Figura 6 se observa en la siguiente Figura la representación de un espectro electromagnético y sus longitudes de onda.

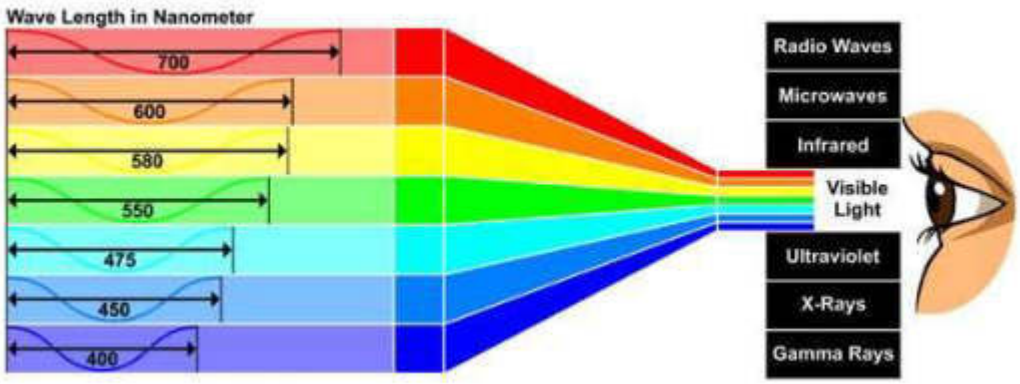


Figura 6. Espectro electromagnético

Fuente: HormonesBalance

La Figura 6 evidencia las distintas regiones del espectro electromagnético, ubicando la estrecha porción correspondiente al espectro visible. La característica fundamental de las cámaras multiespectrales radica en su capacidad para registrar la reflectancia electromagnética en bandas que trascienden las limitaciones del ojo humano. Esta amplitud de detección permite

el desarrollo de múltiples aplicaciones de análisis espacial, tales como la evaluación de la calidad del agua en ríos urbanos y el monitoreo continuo del deterioro ambiental en su entorno.

1.1.8 Google Earth Engine

En los últimos años se ha conocido una inmensa información satelital obtenida a escala global, recopilada por diferentes agencias o sensores a nivel mundial, conformando un vasto repertorio como lo son MODIS, Sentinel y Landsat como las de mayor importancia actualmente (Yang et al., 2016).

Actualmente, el principal desafío para la comunidad científica no radica en la disponibilidad de datos satelitales, sino en la capacidad computacional y técnica para procesar y analizar de manera eficiente estos grandes volúmenes de información (Wang et al., 2015; Yang et al., 2017). Históricamente, otra limitante significativa ha sido la gestión local de estos recursos; la descarga de escenas satelitales constituía un proceso ineficiente que demandaba tiempos prolongados. Sumado a esto, los equipos de cómputo convencionales frecuentemente carecían de la capacidad de almacenamiento y memoria necesaria, lo que convertía el manejo de datos geoespaciales a gran escala en una tarea compleja y computacionalmente restrictiva (Perilla & François., 2020).

Para superar estas limitantes computacionales, a finales de 2010 surgió GEE, una plataforma diseñada para facilitar análisis complejos con grandes volúmenes de datos, expandiendo así las fronteras de la ciencia en teledetección. Esta herramienta permite el procesamiento de información geoespacial directamente en la nube, eliminando la dependencia de los recursos de almacenamiento y procesamiento locales del usuario. Asimismo, el catálogo de GEE se actualiza continuamente mediante su integración directa con diversos programas satelitales, poniendo a disposición adquisiciones en tiempo real. En consecuencia, GEE representa una solución tecnológica innovadora para el manejo de datos geoespaciales masivos (Big Data), lo cual agiliza la obtención de resultados y facilita la formulación de estrategias oportunas frente a problemáticas ambientales globales (Gorelick et al., 2017).

La arquitectura de esta plataforma se contempla por medio de tres componentes principales. El primero corresponde a la infraestructura de servidores de Google, la cual otorga a los usuarios acceso a un procesamiento paralelo distribuido a través de miles de unidades centrales de procesamiento (CPUs). Esta capacidad computacional incrementa exponencialmente la velocidad de análisis frente a los equipos de cómputo individuales; en consecuencia, flujos de trabajo geoespacial que de manera convencional requerirían meses o años para completarse, ahora se ejecutan en cuestión de días, horas o incluso minutos (Moore, 2017).

El segundo elemento es el acervo de datos (Datasets). GEE contiene un gran conjunto de datos, incluyendo datos sin procesar, datos preprocesados, modelos de elevación y extensiones de productos a nivel global, nacional y regional (Zhao Q., et a, 2021). Uno de los más usados es el conjunto de datos Landsat utilizados principalmente para hacer análisis multitemporales y actualmente tienen 7 satélites multiespectrales los cuales son: Landsat 1–3 (1972–1983), Landsat-4 (1982–1993), Landsat-5 (1984–2012), Landsat-7 (1999–presente) y Landsat-8 (2013-presente) Landsat-9 (2019- presente). GEE también incluye dentro de su Datasets los satélites Sentinel desarrollados por la ESA. La colección Sentinel incluye RADAR de apertura sintética Sentinel-1 (SAR) (2014–presente), Sentinel-2 multiespectral (2015-presente), Sentinel-3 Ocean and Land Color (2016-presente) y conjuntos de datos de monitoreo troposférico en Sentinel-5P (2018-presente). Otro de los conjuntos de datos importantes en GEE son las imágenes de MODIS. MODIS presenta un gran potencial en la cartografía NRT de la superficie terrestre a nivel nacional y global, adquiriendo imágenes en 36 espectros cuyas resoluciones espaciales van desde los 250 metros hasta 1 kilómetro. Las series temporales de MODIS están disponibles en el catálogo de datos de GEE desde el año 2000 hasta la actualidad, lo que facilita el análisis temporal en todo el mundo (Amani et al., 2020).

El tercer elemento es la API, la cual consiste en una serie de comandos y funciones preestablecidas, escritos en lenguaje JAVA, que permite una

programación sencilla al desarrollar algoritmos para las investigaciones. Sin embargo, dichas funciones no pueden modificarse ni tampoco se puede acceder a su código, lo cual impide al usuario realizar modificaciones para lograr análisis alternativos a los preestablecidos (Chien et al., 2018). A continuación, se observa una imagen general de la vista de GEE y algunas de sus funciones.

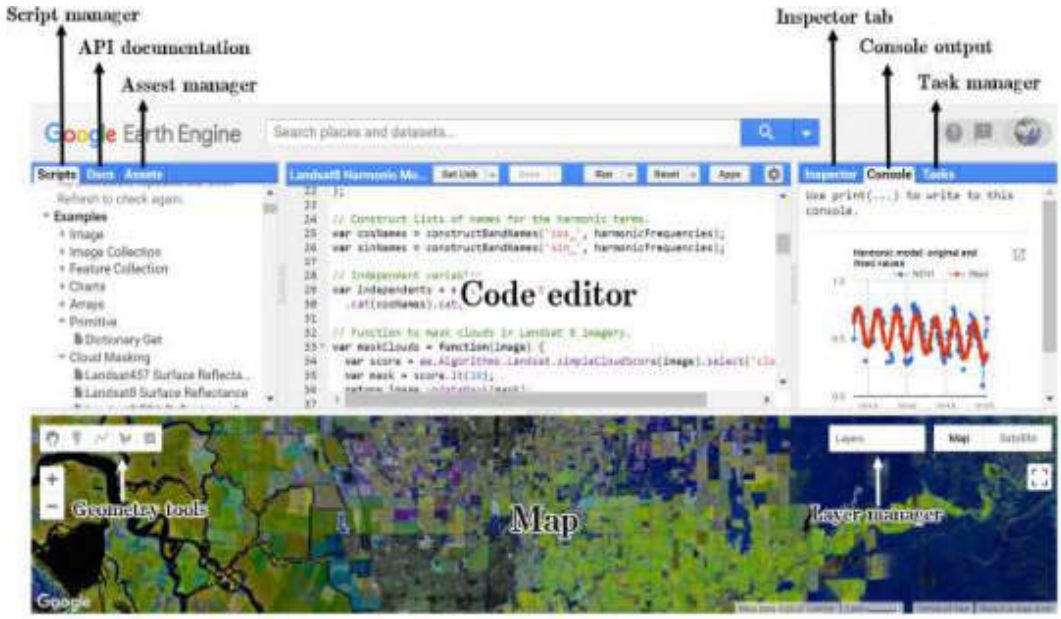


Figura 7. Descripción general GEE

Fuente: Amani, 2020

GEE ha sido utilizado para evaluar y monitorear la calidad del agua, en especial los parámetros de sólidos suspendidos totales, demostrando que pueden procesarse grandes extensiones de imágenes Sentinel-2, y extraerse valores espectrales para desarrollar modelos empíricos de la calidad del agua y visualizar los resultados usando mapas de distribución, tal es caso de la investigación realizada por Pangihutan et al., (2022) en el río Belawan, donde se aplicó el uso de GEE en épocas lluviosas y secas concluyendo su factibilidad para el mapeo de los patrones de distribución de los SST.

Por otro lado, Kwong y colaboradores (2022) utilizaron la misma técnica en el mar en Hong Kong logrando preprocesar todas las imágenes Sentinel-2, y concluyeron que la integración de imágenes satelitales, mediciones in situ y GEE

ofrecen nuevas oportunidades para proporcionar predicciones rentables y oportunas de los parámetros de calidad de agua. Asimismo, Nguyen, et al. (2022) desarrollaron un modelo para evaluar la calidad del agua de mar utilizando GEE en la bahía Rach gia en Vietnam. Por su parte Wang y colaboradores (2020) investigaron un método para la extracción de agua (ríos, lagos) utilizando imágenes Landsat en GEE. Este estudio desarrolló un entrenamiento offline y una predicción online (*OTOP,* por sus siglas en inglés) para su modelo. Determinando así, la importancia de GEE para la extracción de la clase de agua en ríos y lagos. Concluyendo que GEE es una plataforma efectiva y conveniente para el estudio de calidad de agua sin costo alguno y que no se requieren equipos costosos para almacenar y procesar el conjunto de imágenes utilizado. Demostrando de esta forma la importancia de GEE en los estudios de monitoreo de calidad del agua los cuales pueden ser apoyados por modelos que ayuden en la gestión de los ríos urbanos en todo el mundo.

1.1.9 Inteligencia artificial y Machine Learning

La IA se conoce como la capacidad de una máquina para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano, es decir la habilidad de una máquina para simular el comportamiento y actividades propias de la inteligencia humana (Gunning D et al., 2017; Hamet P & Tremblay J., 2017). Por ejemplo, la IA puede aprender de ejemplos y experiencias, reconocer objetos, tomar decisiones, resolver problemas, entre otras múltiples funciones. En la actualidad se encuentra en los servicios de correo electrónico, automatización de procesos industriales, procesamiento de lenguaje natural, etc.

Ahora bien, ML es una rama de la inteligencia artificial que permite que las máquinas aprendan por sí solas a partir del procesamiento de datos, identificando ciertos patrones con los cuales se pueden hacer predicciones. Esta rama de la inteligencia artificial usa como base la estadística y se divide en diferentes campos y puede ser utilizado en áreas de la medicina, tecnologías, marketing e ingeniería (Zhu et al., 2022; Jalife., 2022). Además, dentro de la ingeniería estas

tecnologías encuentran aplicaciones particularmente en el monitoreo de la calidad de agua en ríos urbanos (Fu et al., 2022). Esto implica la posibilidad de desarrollar un modelo predictivo para la calidad del agua en dichos cuerpos receptores de agua, mediante la integración de tecnologías satelitales, plataformas como GEE y técnicas de aprendizaje automático, como lo es ML.

Aunado a lo anterior, se han hecho investigaciones en las áreas mencionadas, como Peterson y colaboradores (2020) que desarrollaron un modelo de aprendizaje automatizado DL para la estimación de algas verdeazuladas, Chl, OD y turbidez. Este modelo se aplicó para datos entre 2013-2018 y a la vez compararon su algoritmo con otros estudios de ML con la finalidad de analizar su regresión. Dentro de sus resultados obtuvieron, para los parámetros fisicoquímicos mencionados anteriormente, regresiones entre 0.84-0.91 y a su vez elaboraron un EDA para encontrar anomalías generando así un eficiente valor agregado a la ciencia del agua mediante el uso de ML.

De igual manera, investigadores como Chen (2021) implementaron modelos de Machine Learning en el río Nanfei para predecir la calidad de agua en ríos urbanos de China. En su estudio utilizaron imágenes de VANT y datos caracterizados in situ para sus datos de entrenamiento y validación. Además, tuvieron en cuenta los siguientes parámetros: Clorofila, fósforo total, nitrógeno total y turbiedad obteniendo predicciones entre 60-85%.

1.2 Justificación

En México, los ríos enfrentan una vulnerabilidad significativa debido al agresivo desarrollo de las ciudades, las cuales generalmente carecen de regulaciones y supervisión adecuadas. Además, los gobiernos no cuentan con la estructura necesaria para llevar a cabo una inspección y vigilancia de los cauces. Como resultado, las principales fuentes contaminantes en los ríos urbanos están relacionadas con la presencia de asentamientos regulares e ilegales y sus vertimientos al cauce, así como la acumulación de residuos sólidos y el hecho de no contar con Plan de Gestión Integral que considere a los ríos urbanos agrava esta problemática.

En este contexto, la calidad del agua de los ríos se ha convertido en un desafío importante para los gobiernos locales y regionales, así como para los organismos encargados de la gestión del agua. La evaluación periódica de la calidad del agua de los ambientes lóticos (ríos) es esencial para garantizar la salud pública y el cuidado del medio ambiente.

Debido a las múltiples problemáticas generadas por la expansión urbana en las zonas aledañas a los ríos (también conocidas como zonas riparias o planicies de inundación), el uso de las herramientas geoespaciales (imágenes satelitales) junto con la aplicación de imágenes obtenidas con VANT y el uso de modelos de ML han demostrado ser una metodología útil y efectiva en la predicción de la calidad del agua de estos ecosistemas.

El presente estudio tiene como ***objetivo desarrollar un modelo de predicción basado en Machine Learning para estimar la calidad del agua de los ríos urbanos utilizando herramientas geoespaciales y tecnología de VANT.*** Los resultados de esta tesis tendrán implicaciones importantes para la gestión del agua en las áreas urbanas, ya que además dan cumplimiento al indicador 6.3.2. de los ODS, y se respalda a los organismos encargados de la gestión del agua con información para la toma decisiones mejor informadas y eficaces sobre la calidad del agua de los ríos urbanos del Área Metropolitana de Monterrey, por lo tanto, se espera mejore la calidad de vida de la población.

2 Hipótesis

El uso de las herramientas geoespaciales, tecnología VANT y la inteligencia artificial permitieron desarrollar un modelo de predicción que logró estimar y evaluar la calidad del agua de los ríos urbanos.

3 Objetivo General

- Desarrollar un modelo GeoAI para el monitoreo de la calidad del agua de ríos urbanos en el Noreste de México utilizando técnicas de Machine Learning y herramientas geoespaciales.

3.1 Objetivos Específicos

- Identificar y ubicar espacialmente las principales problemáticas socioambientales que influyen directa o indirectamente en la gestión de los ríos urbanos, así como sus principales fuentes de contaminación.

- Desarrollar diversas clasificaciones de uso de suelo por medio de algoritmos de Machine Learning (Random Forest, K-Means Clustering, Support Vector Machine y CART) que permitieron identificar los cambios en las temporadas de estiaje y lluvia para el período 2016-2019.

- Identificar la continuidad del río Pesquería por medio técnicas de geoprocesamiento con la finalidad de obtener la máscara de agua para las temporadas estudiadas.

- Identificar la calidad de agua del río Pesquería con la ayuda de índices espectrales y datos in situ para obtener un modelo predictivo basado en Machine Learning.

4 METODOLOGÍA

El enfoque metodológico descriptivo y el tipo de investigación están soportados en el carácter técnico e interdisciplinar del trabajo y en el uso de herramientas geoespaciales para el monitoreo de calidad de agua en los ríos

urbanos (Pesquería). En este apartado se explicará a detalle la metodología con la que se desarrolló esta investigación.

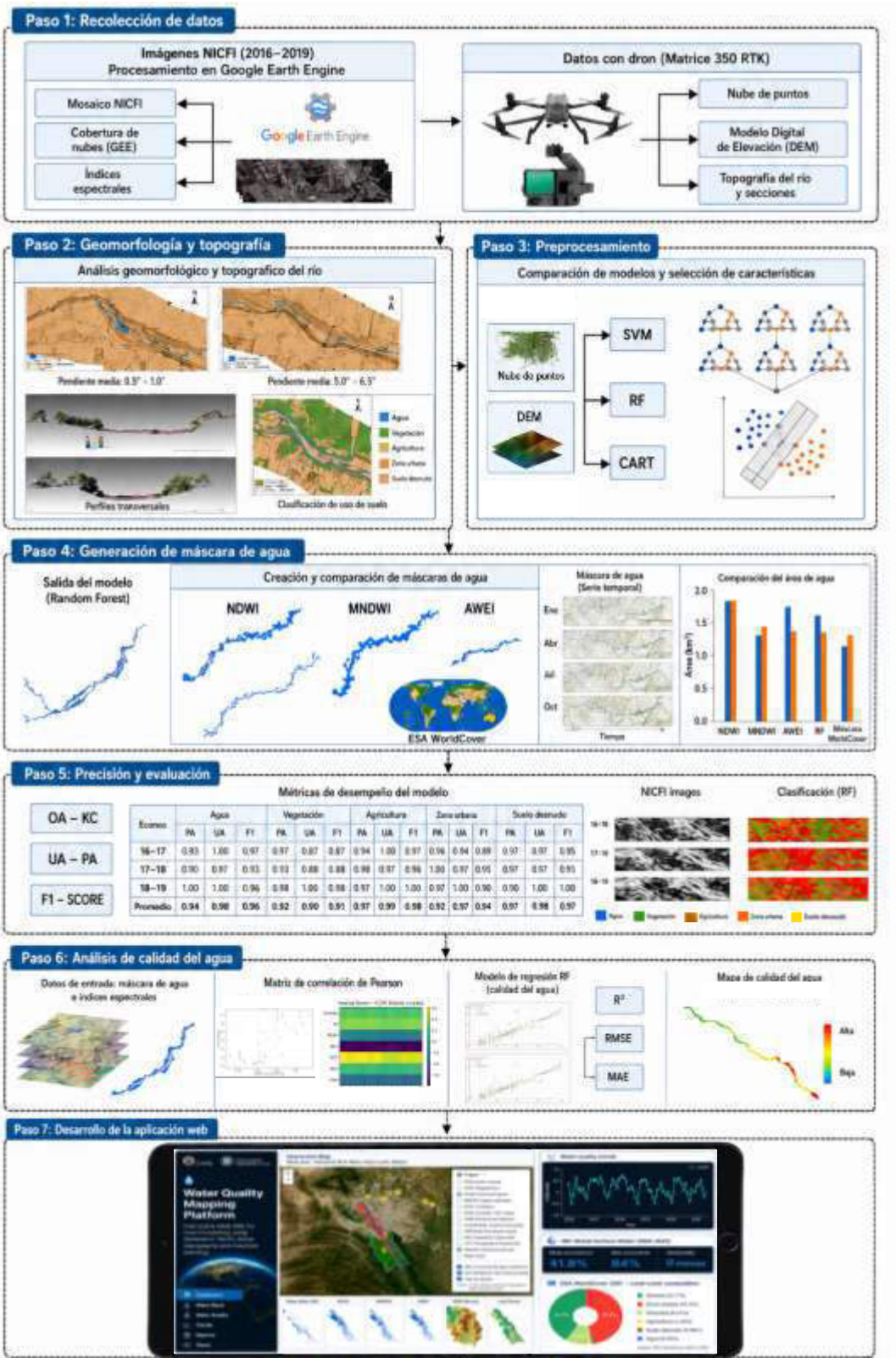


Figura 8. Metodología general.

Fuente: Elaboración propia

4.1 Área de estudio

El área de estudio general corresponde a la cuenca del río Pesquería la cual abarca alrededor de 75 km lineales del río contemplando los municipios de García, Monterrey, Escobedo, Apodaca, San Nicolás y Pesquería. También, se realizó una experimentación dentro de un área específica de 17 km lineales de poniente a oriente atravesando los Municipios de General Escobedo y Apodaca. En la siguiente Figura se observa el área de estudio de la investigación y el área seleccionada.

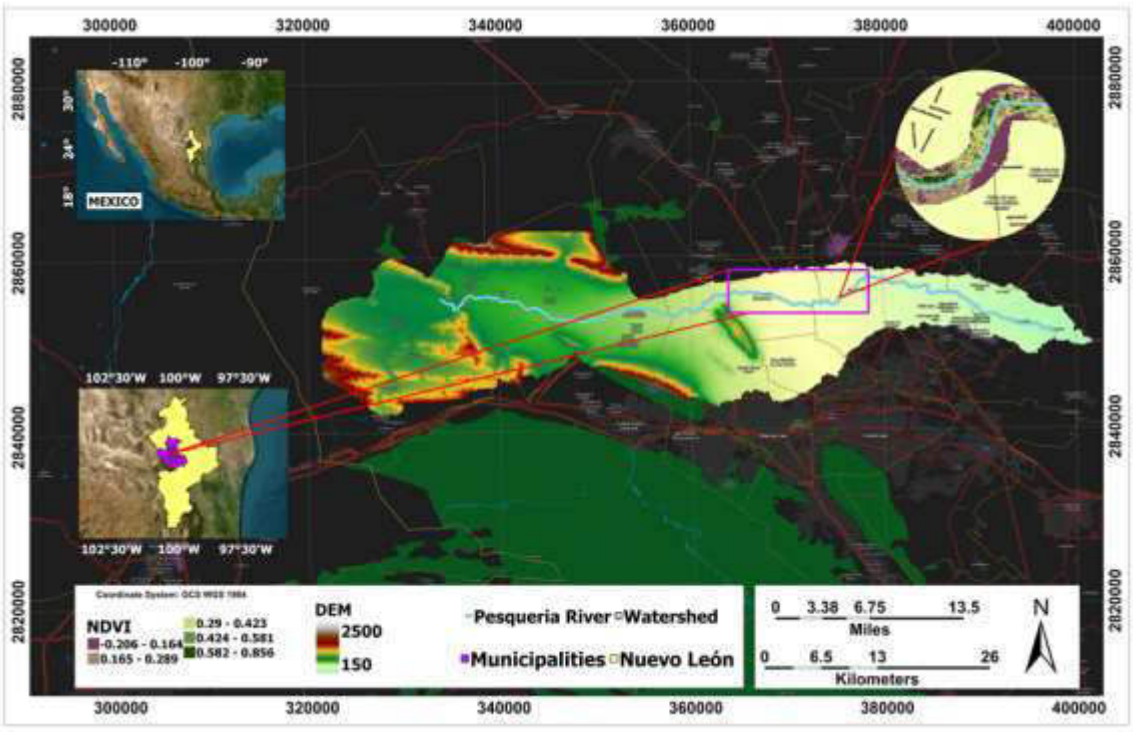


Figura 9. Área de estudio y sitio de experimentación

Fuente: Elaboración propia

En la zona mayormente urbanizada del río Pesquería se observó un patrón diverso de uso del suelo debido a los cambios que se han presentado en los últimos años. La clasificación del uso de suelo es fundamental para comprender mejor las dinámicas que afectan a este importante cauce fluvial, en este contexto

el uso de los sensores remotos se vuelve una herramienta muy valiosa para abordar este tipo de estudios a gran escala acoplados con subconjuntos de la inteligencia artificial.

4.1.1 Clima e hidrología

En zonas semiáridas, sistemas fluviales como el río Pesquería fundamentan su caudal casi exclusivamente en los eventos de precipitación, debido a su ubicación sobre sustratos rocosos o acuíferos de baja profundidad. Esta configuración geohidrológica propicia episodios frecuentes de estiaje (Mora et al., 2017). Asimismo, las propiedades fisicoquímicas del agua están determinadas principalmente por la naturaleza de las lluvias y la dinámica de los suelos superficiales (Montgomery & Buffington, 1997; Rinaldi et al., 2016). Por otro lado, la urbanización de la cuenca del Pesquería ha introducido múltiples descargas artificiales que alteran su régimen de flujo, lo cual reduce significativamente la resiliencia del cuerpo de agua y su capacidad de autodepuración frente a los contaminantes.

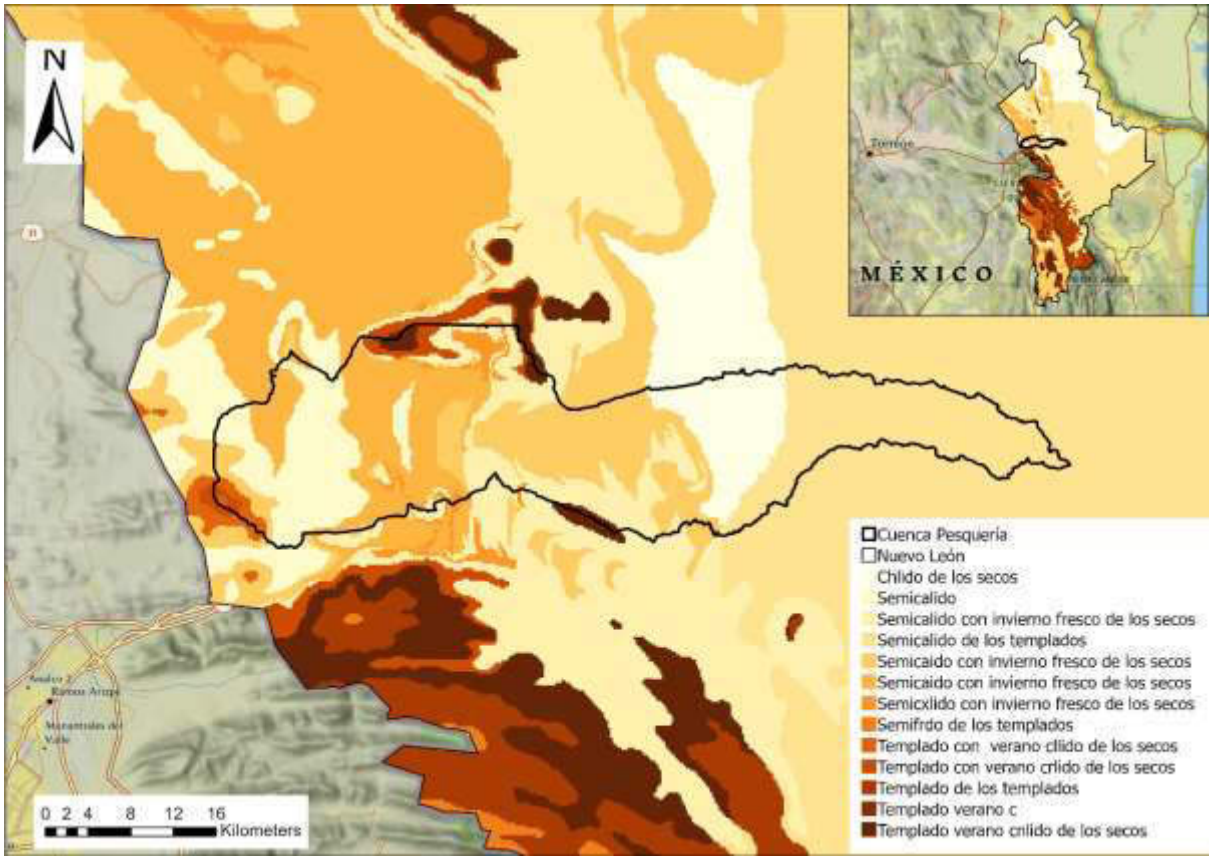


Figura 10. Datos climáticos en Nuevo León y en la cuenca Pesquería

Fuente: Elaboración propia con datos Geográficos de la carta de climas

La cuenca de estudio se caracteriza por un clima predominantemente semiárido cálido, con isotermas anuales que oscilan entre los 18 y 22 °C, una elevación media de 542 m s.n.m. y un relieve de pendiente suave (0.4 %) (Castro-López., et al., 2019). Geográficamente, se inscribe en la Región Hidrológica 24 Bravo-Conchos y forma parte de la cuenca del río San Juan; este sistema hídrico atraviesa los estados de Coahuila, Nuevo León y Tamaulipas, drenando una superficie total de 30,562 km² hasta su desembocadura en la presa El Azúcar, en Camargo, Tamaulipas (Ferriño & Burgos, 2015).

La precipitación presenta una marcada estacionalidad, con un promedio anual cercano a los 600 mm concentrados de mayo a octubre. No obstante, el área es altamente vulnerable a fenómenos hidrometeorológicos extremos.

Históricamente, en estaciones cercanas a General Escobedo se han documentado precipitaciones máximas en 24 horas de 253 mm (1988) y 280 mm (2010), destacando el impacto del huracán Gilberto (1988), el cual ocasionó desbordamientos críticos en el río Pesquería y otros afluentes del Área Metropolitana de Monterrey.

Esta variabilidad pluvial incide directamente tanto en la susceptibilidad a inundaciones como en la calidad del recurso hídrico. Si bien los picos de escorrentía superficial propician la dilución de contaminantes, simultáneamente provocan el arrastre de sólidos suspendidos y descargas urbanas hacia el cauce, alterando drásticamente el equilibrio ecológico del sistema (Yang et al., 2021). Así, el comportamiento hidroclimático de la cuenca está dictado por la dicotomía entre sequías prolongadas y lluvias torrenciales, lo que hace imperativa la implementación de evaluaciones holísticas para su gestión.

4.1.2 Relieve y Geología

La Figura 11 ilustra las provincias fisiográficas de México establecidas por la CONABIO a escala 1:1,000, 000. Esta regionalización proporciona una perspectiva macroscópica del relieve nacional, delimitando unidades espaciales provincias y subprovincias de relativa homogeneidad geomorfológica. Dichas unidades fueron definidas a partir de la integración sistemática de variables topográficas, geológicas, hidrológicas y edafológicas.

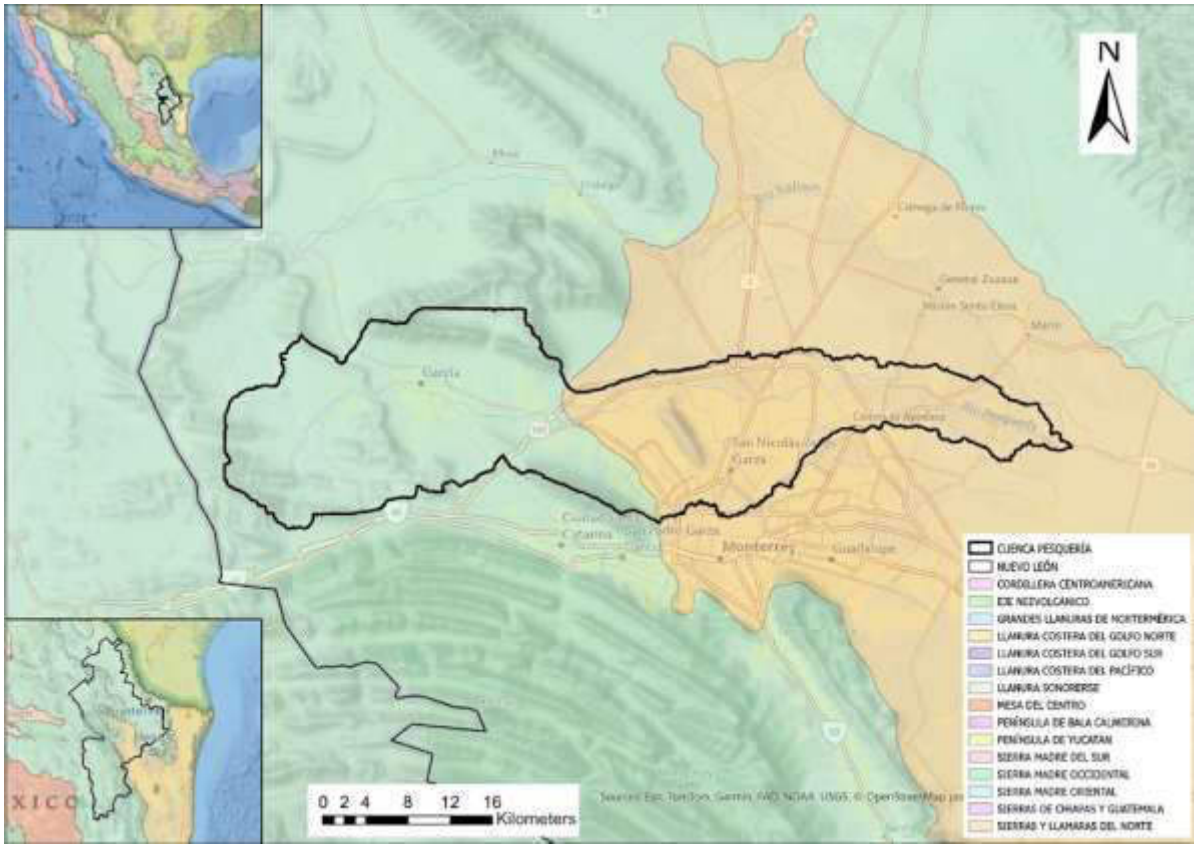


Figura 11. Fisiografía de México y la cuenca del río Pesquería según clasificación de INEGI

Fuente: Elaboración propia con datos obtenidos por INEGI

Sierra Madre Oriental

Esta provincia fisiográfica domina la porción occidental del estado de Nuevo León y alberga la secuencia estratigráfica más antigua de la entidad. Su basamento incluye esquistos precámbricos y formaciones triásicas constituidas por areniscas y alternancias de lutitas-areniscas. Durante el Jurásico Superior, la litología se caracterizó por afloramientos masivos calcáreos, evaporíticos (yesos) y asociaciones de ambos, acompañados de secuencias terrígenas. Por su parte, el Cretácico se distingue por una extensa cobertura de calizas, lutitas y facies mixtas. Posteriormente, en el Terciario, la intrusión de cuerpos ígneos de composición ácida e intermedia afectó y modificó localmente las rocas sedimentarias mesozoicas. A este periodo también corresponden los conglomerados depositados en los piedemontes serranos, a los que se superponen los depósitos aluviales del Cuaternario.

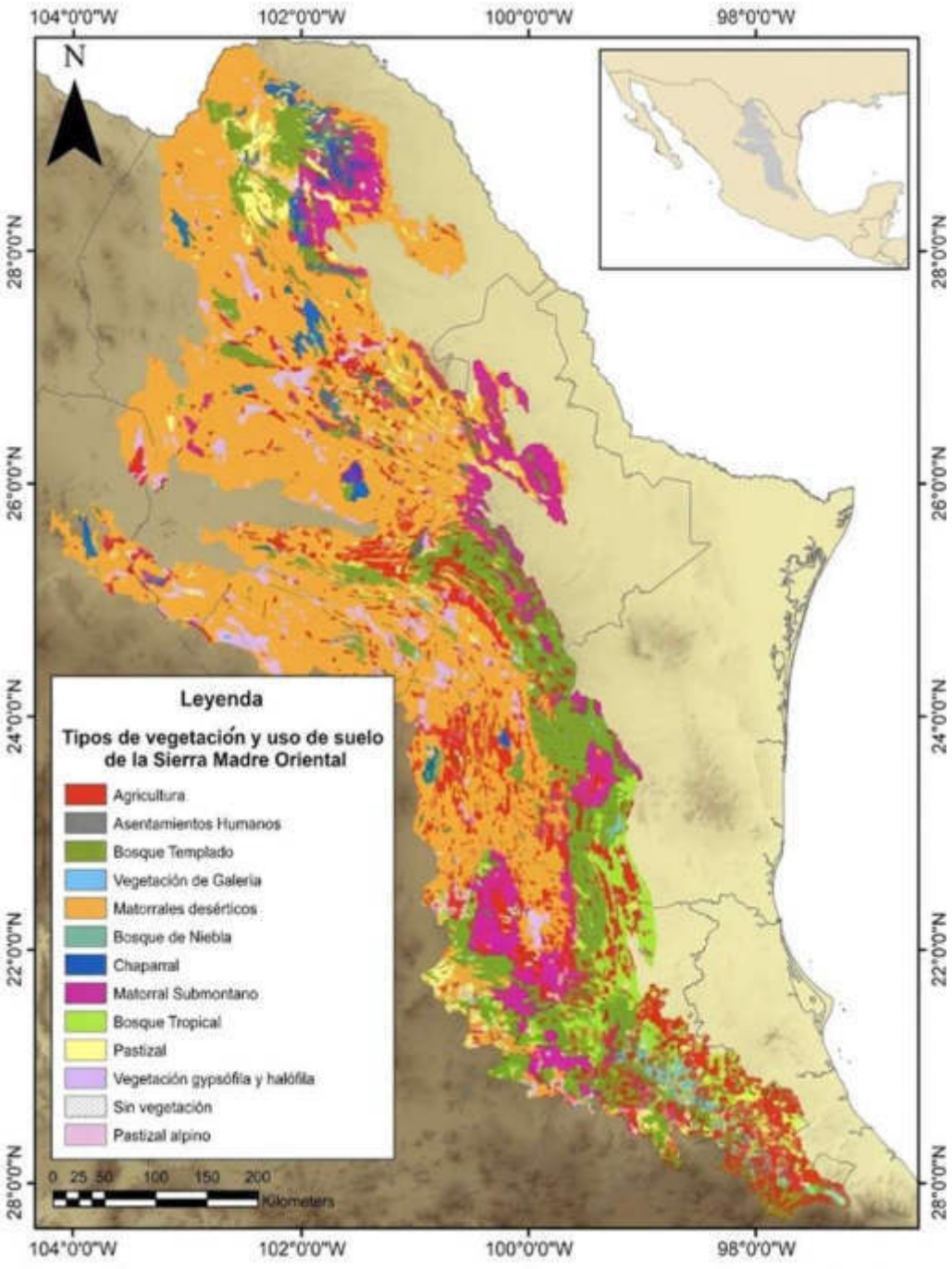


Figura 12. Sierra Madre Oriental y sus usos de suelo

Fuente: Modificado de Salinas (2018)

Desde el punto de vista tectónico, la Sierra Madre Oriental presenta una alta complejidad estructural. Su relieve está gobernado por un vasto sistema de pliegues (anticlinales y sinclinales) dislocados por fallas normales de gran magnitud y numerosas cabalgaduras (fallas inversas), las cuales han alterado drásticamente la disposición estratigráfica original (salinas. 2018).

Llanura Costera del Golfo Norte

Esta provincia fisiográfica se despliega a lo largo del margen oriental del estado de Nuevo León, configurando una franja de relieve predominantemente llano. Topográficamente, se distingue por sus bajas cotas altimétricas y un modelado suavemente ondulado, dominado por potentes secuencias sedimentarias recientes. Su basamento litológico está constituido de forma primordial por depósitos clásticos cenozoicos (Terciario y Cuaternario) tales como lutitas, areniscas, limos y gravas, originados a partir de dinámicas fluviomarinas pretéritas. Al ser materiales con un bajo grado de consolidación, suelen subyacer a horizontes edáficos modernos que presentan una alta susceptibilidad a la erosión.

Durante el Cuaternario, los procesos exógenos esculpieron extensas planicies aluviales, sistemas de terrazas y abanicos fluviales, propiciando la acumulación de suelos de alta fertilidad. Esta vocación edáfica y topográfica confiere a la provincia un carácter estratégico para el desarrollo agroindustrial, la expansión urbana y el despliegue de infraestructura.

Desde la perspectiva estructural, la llanura muestra una notable estabilidad tectónica, con escasa deformación en contraste con el relieve escarpado del occidente. No obstante, se han identificado sistemas de fallamiento menor y procesos de subsidencia localizada; estos últimos están asociados tanto a la compactación natural de sedimentos finos como a forzantes antrópicos, destacando la extracción intensiva de agua subterránea y la fracturación

hidráulica (fracking). Hidrológicamente, la provincia funciona como el principal colector del drenaje regional al captar la escorrentía superficial de la Sierra Madre Oriental. Esto constituye así una zona de transición orográfica fundamental, dentro de la cual se asienta la cuenca del río Pesquería.

Tabla 2. Superficie a nivel estado- geología (%)

Era		Período		Roca o suelo		Total de superficie (%)
Clave	**Nombre**	**Clave**	**Nombre**	**Clave**	**Nombre**	
C	Cenozoico	Q	Cuaternario	S	Sedimentaria	2.52
				Su	Suelo	31.63
		T	Terciario	Ii	Ígnea intrusiva	0.22
				S	Sedimentaria	22.95
M	Mesozoico	K	Cretácico	S	Sedimentaria	39.96
		J	Jurásico	S	Sedimentaria	1.87
			Otro			0.85

Fuente: INEGI

4.1.3 Crecimiento poblacional

De acuerdo con la Figura 13, desde la década de 1930 México ha registrado un incremento demográfico, el cual se ha desarrollado de manera paralela a un acelerado proceso de expansión urbana a nivel nacional. Esta dinámica ha propiciado que la población se consolide en zonas metropolitanas, conformadas por núcleos urbanos principales y sus respectivos municipios periféricos conurbados. En este contexto, el AMM se muestra como una de las aglomeraciones más extensas y dinámicas a nivel nacional, superada territorial y demográficamente únicamente por la Zona Metropolitana del Valle de México. En la Figura 13 se observa el crecimiento demográfico desde 1910 hasta 2020 de México, Nuevo León y la cuenca de Pesquería conjuntamente.

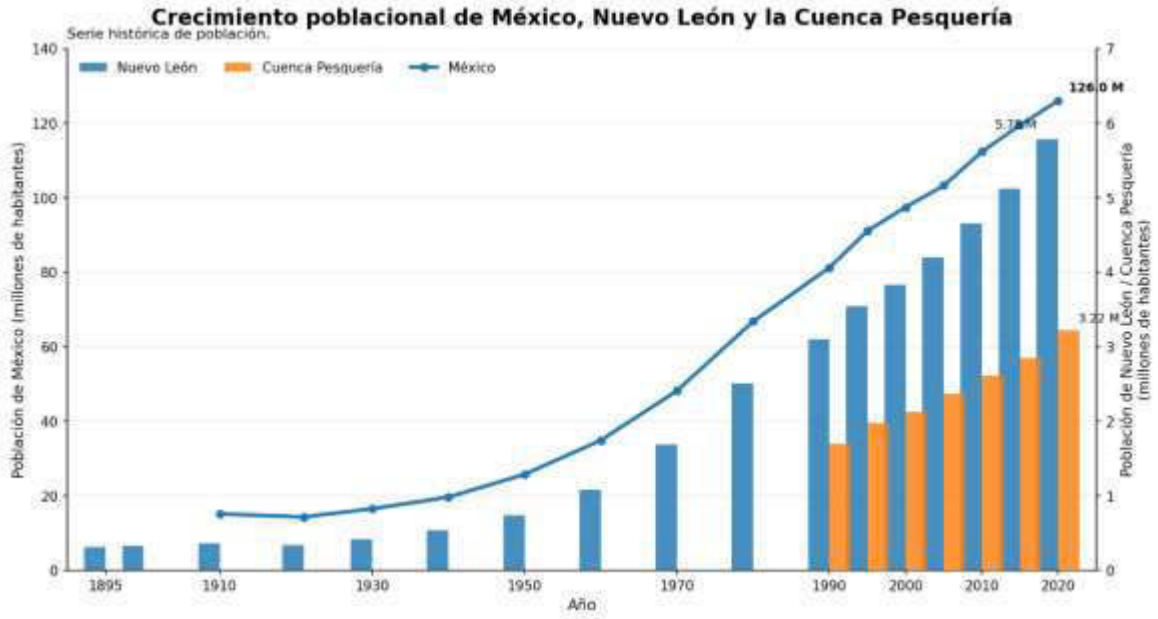


Figura 13. Crecimiento poblacional de México, Nuevo León y la cuenca Pesquería

Fuente: INEGI

4.2 Estructura general del río Pesquería

4.2.1 Levantamiento topográfico y geomorfológico mediante tecnología LiDAR

Para determinar la estructura geomorfológica y extraer los perfiles topográficos de detalle del cauce del río Pesquería, se realizó un levantamiento topográfico y fotogramétrico de alta resolución. La adquisición de datos se llevó a cabo utilizando un drone DJI Matrice 350 RTK, equipado con el sensor LiDAR Zenmuse L2. Este escáner láser aéreo, que opera con una longitud de onda de 905 nm, posee la capacidad de registrar hasta cinco retornos por pulso, logrando una precisión posicional de 2 cm (RMS 1σ) en rangos de hasta 150 m. El sistema integra de manera simultánea una cámara RGB con un sensor CMOS 4/3 de 20 megapíxeles, lo cual permitió proporcionar contexto visual y colorización fotorrealista a la nube de puntos tridimensional.

Los vuelos fueron planificados a una altitud operativa entre 90 a 100 m sobre el nivel del terreno (AGL, por sus siglas en ingles), manteniendo traslapes

frontales y laterales de entre el 75 % y el 80 %. Esta configuración de vuelo garantizó la captura de nubes de puntos de alta densidad (Rodríguez et al., 2026), almacenadas en formato LAS 1.4. El procesamiento de gabinete consistió en la segmentación de la nube de puntos y la aplicación de filtros para la remoción de la cubierta vegetal y elementos antrópicos. Esto permitió aislar los retornos correspondientes al terreno desnudo para generar Modelos Digitales de Elevación de alta precisión, los cuales se contemplaron como insumo principal para el trazado de los perfiles topográficos transversales y longitudinales del cauce. Así como también se obtuvieron ortomosaicos y modelos 3D de algunas secciones del río Pesquería y sus zonas aledañas.

4.2.2 Análisis multitemporal de problemáticas socioambientales

De forma complementaria a la caracterización geométrica, se realizó un diagnóstico visual de las presiones antrópicas sobre el sistema fluvial. Para ello, se generaron ortomosaicos de alta resolución a partir de las imágenes capturadas por el sensor RGB del VANT. Estos productos fotogramétricos recientes se contrastaron espacialmente con un registro histórico de imágenes de satélite extraídas de la plataforma GEE. La superposición y el análisis comparativo de estas fuentes ópticas permitieron identificar y cartografiar la evolución de las principales problemáticas socioambientales a lo largo del cauce, tales como asentamientos irregulares, zonas de disposición clandestina de residuos, deforestación de la vegetación ribereña y alteraciones directas en la llanura de inundación.

4.2.3 Análisis ambiental

Análisis multitemporal de covariables ambientales

Para evaluar la dinámica hidroclimática y la respuesta ecosistémica de la cuenca durante el periodo 2000-2025, se procesó una serie de covariables satelitales utilizando la plataforma GEE. Se extrajeron cuatro parámetros fundamentales: precipitación mensual-CHIRPS, LST, NDVI y temperatura del aire a 2 metros de altura (T2M).

El flujo de procesamiento espacial consistió en calcular la media espacial mensual para cada variable dentro del área de estudio. Para cuantificar la heterogeneidad territorial de estos parámetros, se extrajeron los percentiles 10 y 90 (P10-P90), delimitando así el rango de dispersión espacial intra-mensual. En el ámbito temporal, se aplicó un suavizado estadístico mediante una media móvil de 12 meses; esto permitió filtrar el ruido estacional y aislar las tendencias interanuales subyacentes. Finalmente, la evolución de estas métricas se intersectó temporalmente con forzantes climáticos clave, contrastando la respuesta de la temperatura y la vegetación frente a episodios de déficit hídrico (Sequía 2011-2013, fenómeno de la Triple La Niña y la crisis hídrica de Monterrey) y eventos de precipitación extrema asociados a ciclones tropicales (Emily, Alex y Alberto), esto con la finalidad de contemplar un análisis histórico de las principales condiciones ambientales del área de estudio.

Índices de Evaluación Urbana

A demás de lo mencionado anteriormente, se desarrolló otro análisis espacial conforme a las presiones ecológicas y urbanas en la cuenca del río Pesquería mediante un flujo de geoprocesamiento en la plataforma GEE a una resolución de 30 metros con imágenes Landsat. Esta metodología se estructuró en dos etapas secuenciales: la primera enfocada en la derivación y normalización de variables base, y la segunda centrada en el modelo para la identificación de zonas críticas explicadas a continuación.

Etapa 1: Preprocesamiento y derivación de subíndices temáticos

En la primera fase, se procesaron colecciones de datos satelitales para extraer variables crudas que logran capturar la dinámica física y espectral del territorio. Estas variables incluyeron indicadores de intensidad urbana (superficie construida mediante GHSL e Índice NDBI), salud ecológica (Índice NDVI) y estrés térmico (anomalías de Temperatura Superficial-LST).

Para estandarizar la escala de análisis, todas las variables fueron normalizadas internamente respecto a los umbrales de la cuenca (percentiles 2 y 98). A partir de estos insumos, se construyeron cuatro subíndices temáticos fundamentales: URB, Vegetación obteniendo VEG, Hidrológico como HID y CLI, los cuales sirvieron como soporte para la construcción de los índices sintéticos de la siguiente fase.

Etapa 2: Modelado de los Índices de Convergencia y Degradación Ambiental-Urbana

A partir de las variables y subíndices de la Etapa 1, se diseñó un modelo espacial basado en dos índices complementarios y metodológicamente distintos: el ICU y el IDAU.

- **ICU:** Conceptualizado para detectar los frentes de expansión activa, este índice evalúa dónde coexisten múltiples presiones territoriales. Posee una formulación geométrica, asumiendo que la falta de una sola dimensión penaliza el valor global. Esto permite identificar entornos periféricos que aún conservan complejidad ambiental, pero que están sometidos a procesos de transformación. A continuación, se presenta la ecuación contemplada.

$$ICU = (URBxVEGxHIDxCLI)^{1/4}$$

Ecuación 1. ICU

- **IDAU:** En contraste, el IDAU fue diseñado para medir el daño ecológico consumado. Su formulación es aditiva, partiendo de la premisa de que las presiones ambientales en el núcleo urbano se acumulan de manera lineal sobre el sistema:

$$IDAU = (IntURBx0.35) + (ETx0.35) + (DVx\ 0.3)$$

Ecuación 2. IDAU

Donde:

IntURB: Intensidad Urbana, ET: Estrés térmico,DV: Déficit de vegetación

La intensidad urbana se calculó ponderando la superficie construida y el NDBI (excluyendo zonas de transición); el estrés térmico se derivó de las anomalías de LST; y el déficit de vegetación se estableció como el inverso de los valores normalizados del NDVI.

4.3 Análisis de imágenes NICFI

El estudio, procesamiento y uso de imágenes satelitales en esta investigación utilizó como apoyo la plataforma GEE, la cual ha sido desarrollada para facilitar el análisis a gran escala de datos geoespaciales. Para el manejo de GEE se necesita un conocimiento en programación Javascripyt o Phyton y las siguientes consideraciones:

Manejo de Google Earth Engine.

- APIs.
- Catálogo y capas de datos.
- Satélites.
- Capas.
- Code editor

4.3.1 Imágenes satelitales

En esta investigación se usaron imágenes satelitales NICFI para la clasificación de uso y cobertura de suelo del área experimental. A diferencia de las imágenes satelitales comerciales de muy alta resolución (tabla 3) que pueden tener costos elevados, las imágenes NICFI se encuentran por medio de una licencia gratuita la cual se puede enlazar directamente a GEE para su uso continuo y avanzado. En la Tabla 3 se observan las especificaciones técnicas de este sensor.

Tabla 3. Especificaciones NICFI

N° Banda	Nombre de la banda	Resolución espacial (m)
1	Azul	4.7
2	Verde	4.7
3	Rojo	4.7
4	NIR	4.7

Fuente: Adaptado de: Google Earth Engine

Los datos de NICFI se han optimizado para reflejar la superficie estándar. Los mapas base están diseñados para aplicaciones científicas generales y son especialmente adecuados para algoritmos de clasificación de cobertura terrestre y otras aplicaciones de investigación basadas en el tiempo (Pandey et al., 2023). En esta investigación se empleó una serie temporal de datos satelitales adquiridos para el período de 2016 al 2019, con el fin de realizar un análisis de la dinámica espacio-temporal de la cobertura terrestre y el uso del suelo.

Por otra parte, estas temporalidades contemplaron imágenes captadas durante diferentes épocas fenológicas, abarcando tanto períodos de estiaje como la época de lluvias, esto permitió caracterizar las variaciones estacionales que presentaron las diferentes clases de suelo y agua en el rio Pesquería a lo largo de la zona urbana. En general se tuvieron en cuenta 5 temporadas de las cuales 3 son de estiaje y 2 de lluvias tal y como se observa en la Tabla 4.

Tabla 4. Períodos de estiaje y lluvia por temporada

Imágenes satelitales	Temporada o época	Año
NICFI	Seca	2016-2017 - (16_17)
	Lluvias	2017-2017 - (17_17)
	Seca	2017-2018 - (17_18)
	Lluvias	2018-2018 - (18_18)
	Seca	2018-2019 - (18_19)

Fuente: Elaboración propia

Las temporadas se definieron sobre la base de datos históricos de valores mensuales totales de temperatura y lluvia (promedio de 2008 a 2017) realizada por (Meléndez et al., 2020), que se obtuvieron de las estaciones del SIMA ubicadas dentro del área de estudio. De acuerdo con esto se tuvo en cuenta la estación seca (noviembre, diciembre, enero, febrero, marzo y abril) y la estación de lluvias (mayo, junio, julio, agosto, septiembre y octubre).

4.4 Clasificación de imágenes

Para la clasificación de uso de suelo se tuvieron en cuenta 4 algoritmos de ML, los cuales serán explicados a detalle a continuación.

4.4.1 Support Vector Machine

Las Máquinas de Vectores de Soporte-SVM se presentan como uno de los algoritmos estadísticos de aprendizaje automático más valiosos y recomendados por la comunidad científica para la clasificación de imágenes satelitales (Sheykhmousa et al., 2020). Esto se debe a la capacidad de analizar eficientemente las clasificaciones de uso y cobertura del suelo, con una notable aptitud para procesar comportamientos tanto lineales como no lineales. El

modelo SVM se fundamenta en un sofisticado principio de minimización del riesgo estructural, lo que lo posiciona como un algoritmo particularmente efectivo en la clasificación de imágenes multiespectrales e incluso hiperespectrales (Ouma et al., 2022).

Aunado a lo anterior, este proceso va precedido de una serie de ecuaciones las cuales se muestran a continuación. Para el entrenamiento se emplea la siguiente ecuación.

$$T = \{(x_1, y_1), (x_2, y_2)\ldots,(x_n, y_n)\}, x_1 \in R^d \text{and} y_i \in (—1, +1) \quad (1)$$

Ecuación 3. Datos de entrenamiento etiquetados

Donde x_1 representa el vector de características que describe las propiedades espectrales o espaciales de un elemento de datos, y_1 corresponde a la etiqueta de clase que identifica categóricamente dicho elemento dentro del conjunto de entrenamiento, y, x_n, y_n determinan el número total de elementos presentes en el conjunto de datos de entrenamiento utilizados para construir el modelo de clasificación para cada uno de los vectores. En la siguiente ecuación se define el hiperplano óptimo (Alkaradaghi et al., 2019).

$$w^t x+b = 0$$

Ecuación 4. Hiperplano óptimo

Donde: w es el vector de pesos (es normal o perpendicular al hiperplano y define su orientación angular), x es el vector de características (los datos o variables de entrada que el modelo está analizando), w^t indica la transpuesta del vector de pesos para realizar el producto punto con x y b es el término de sesgo, el cual determina qué tan desplazado está el hiperplano respecto al origen espacial

En el modelo de SVM, la fase de entrenamiento tiene como objetivo principal determinar el vector de pesos w y el sesgo b para trazar el hiperplano óptimo de separación, definido por la ecuación 4. El algoritmo busca clasificar los datos maximizando el margen geométrico entre las distintas clases, cuya distancia

equivale a 2/|w|. Por consiguiente, los vectores de características x que se ubican exactamente sobre los límites de este margen, satisfaciendo la condición $w^t x+b=1$, se denominan vectores de soporte, dado que son los elementos críticos que sostienen y definen la posición del hiperplano. Al resolver dicha optimización (Ec. (5)), los márgenes de separación óptimos determinados por el SVM lineal:

$$\min_{w,b,\varepsilon} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \varepsilon_i$$

Ecuación 5. Tarea de optimización

Donde: C representa el parámetro de penalización o costo (el cual regula el equilibrio entre la maximización del margen y la tolerancia a errores de clasificación), ε_i denota las variables de holgura (que permiten clasificar datos que no son linealmente separables al admitir cierto grado de error), w es el vector de pesos (normal al hiperplano de separación), y b corresponde al término de sesgo escalar.

4.4.2 CART

Los Árboles de Clasificación y Regresión-CART, es un algoritmo de ML desarrollado por Breiman y colaboradores en 1984, representa uno de los clasificadores binarios más simples basados en una estructura de árboles de decisión jerárquicos. A diferencia de las redes neuronales, CART ofrece un sistema de "caja blanca" donde las relaciones de entrada-salida pueden interpretarse de manera transparente y accesible.

Este modelo explora la estructura de datos para visualizar reglas de decisión que predicen resultados categóricos (árbol de clasificación) o continuos (árbol de regresión). En cada nodo interno, la decisión se evalúa mediante ganancia de información o entropía, comparando valores de atributos desde la raíz hasta las

hojas. El proceso de ramificación continúa hasta alcanzar un nodo terminal que representa el resultado de la clasificación (Alsabti et al., 1998).

CART presenta ventajas significativas como simplicidad interpretativa, rapidez de ejecución y alta precisión en clasificación de imágenes. Sin embargo, enfrenta limitaciones como el riesgo de sobreajuste y complejidad excesiva que puede reducir su capacidad de generalización. Estas restricciones han motivado que en múltiples dominios sea reemplazado por métodos de re-muestreo ("ensemble") que mitigan su potencial inestabilidad mediante el promedio de resultados.

Su fundamento se basa en la siguiente ecuación:

$$\sum_{i=1}^{i=n} (y_i - \hat{y}_i)^2$$

Ecuación 6. Jerarquia modelo CART

Donde: yi: Representa el valor real o verdadero del elemento i en el conjunto de datos. ŷi: Representa el valor predicho o estimado por el modelo para el elemento i.

La estructura del árbol permite una toma de decisiones jerárquica donde cada nodo representa una decisión basada en características específicas, ramificándose según condiciones predefinidas hasta alcanzar una clasificación final en los nodos terminales.

4.4.3 Random Forest

Los Árboles aleatorios-RF es una de las técnicas más aplicadas para la clasificación de LULC utilizando imágenes satelitales (Parmar et al., 2018). Este algoritmo utiliza árboles como un enfoque en conjunto, donde los clasificadores de los árboles emplean un sistema (h(x,Θk)h(x, \Theta_k)h(x,Θk), con k=1,…,nk = 1, \dots, nk=1,…,n) , siendo x un vector de entrada y Θk\Theta_kΘk vectores idénticos distribuidos aleatoriamente (Breiman et al., 2001). La función principal de los árboles es clasificar los datos según su criterio individual (voto) para

combinar los resultados y obtener un resultado final agregado. El modelo selecciona la clase resultante a través de un sistema de votación mayoritaria, donde cada árbol del conjunto emite su predicción y la clase más votada se establece como la decisión final.

El proceso comienza con la selección aleatoria de muestras para cada árbol, seguida por la construcción de árboles individuales utilizando criterios de división como el Índice Gini o la Entropía. La probabilidad RF para el entrenamiento se expresa en la siguiente ecuación.

$$\sum \sum_{j \neq i} (f(C_i, T)/|\mathrm{T}|)\,(f(C_i, T)/|\mathrm{T})$$

Ecuación 7. probabilidad RF

Los árboles de decisión convencionales suelen construirse y posteriormente someterse a procesos de poda con el objetivo de reducir la complejidad del modelo y mejorar su capacidad de generalización (Mingers, 1989). En contraste, Random Forest construye múltiples árboles hasta su máxima profundidad sin aplicar poda, utilizando muestras de entrenamiento y subconjuntos aleatorios de variables en cada árbol. Esta estrategia constituye una de las características fundamentales del clasificador Random Forest frente a los métodos tradicionales basados en árboles (Pal, 2005).

Por otra parte, en este estudio se ha llevado a cabo una categorización de las áreas aledañas al río Pesquería, así como del propio cauce, seleccionando una región de interés que abarca aproximadamente un área de 71.8 km². En esta investigación, se realizaron clasificaciones de uso del suelo utilizando GEE mediante diversos algoritmos de ML, que incluyen métodos supervisados y no supervisados, siendo tres algoritmos de clasificación supervisada: CART, SVM y RF. Estos algoritmos se entrenan con un conjunto de muestras representativas para las 5 clases: Agua, Vegetación, Zonas agrícolas, Crecimiento urbano y

Suelo (tabla 5), permitiéndoles aprender los patrones espectrales y diferenciar tipos de cobertura.

De igual manera, se desarrolló una clasificación no supervisada utilizando el algoritmo K-Means clustering. Este método agrupa los píxeles de la imagen en función de su similitud espectral sin conocimiento previo de las clases de cobertura terrestre. Los grupos resultantes se interpretan y etiquetan como diferentes tipos de cobertura. La Tabla 5 identifica los 4 algoritmos de clasificación que se implementaron en esta investigación, así como su tipo (Supervisado o No supervisado).

Tabla 5. Algoritmos de Machine Learning

Nombre	Tipo de clasificación
K- means Clustering	No supervisado
Random Forest	Supervisado
Support Vector Machine	Supervisado
CART	Supervisado

Fuente: Elaboración propia

El objetivo principal de este análisis es principalmente determinar cuál de los 4 algoritmos ofrece mayor precisión al momento de clasificar las diferentes coberturas de suelo y especialmente la continuidad de los ríos intermitentes y de poco flujo como lo es el caso del río Pesquería. A continuación, se presentan las clases utilizadas en este proyecto.

Tabla 6. Clases de uso de suelo

Clase	N° de clase	Color
Agua	1	Azul
Vegetación	2	Verde
Agricultura	3	Café
Urbano	4	Rojo
Suelo	5	Amarillo

Fuente: Elaboración propia

Cada una de estas clases formó parte fundamental de los distintos algoritmos utilizados, donde la imagen resultante se vio influenciada por los conjuntos de datos empleados en el proceso de entrenamiento y validación de dichas clases.

4.5 Conjunto de datos

Para la implementación de los algoritmos de Machine Learning, es crucial considerar conjuntos de datos de entrenamiento y validación. Estos conjuntos contribuyen significativamente a mejorar la precisión de los modelos. A continuación, se presentan los datos iniciales que se tomaron en cuenta.

Tabla 7. Datos iniciales

Datos	Imagen de alta resolución	N° de Datos
Datos de entrenamiento	NICFI	2500
Datos de validación		500
Total de datos		3000

Fuente: Elaboración propia

En la etapa de entrenamiento se utilizaron 500 datos para cada una de las temporadas de los modelos realizado, así como 100 datos de validación con un total de 600 puntos por temporada y en general 3000 puntos para los modelos contemplados en esta investigación.

Por otra parte, para la clasificación de uso y cobertura del suelo se implementó un esquema de integración de variables espectrales derivadas de imágenes satelitales de alta resolución espacial. Como fuente principal de información se utilizaron imágenes Planet NICFI tal como se mencionó anteriormente, a partir de las cuales se incorporaron las bandas espectrales disponibles y el Índice espectral NDVI. La integración de estas variables tuvo como finalidad proporcionar al algoritmo de clasificación un conjunto de predictores capaces de representar de manera diferenciada las características espectrales de las coberturas presentes en el área de estudio.

Las bandas espectrales de las imágenes NICFI aportan información asociada a la respuesta de la superficie en diferentes regiones del espectro electromagnético, permitiendo distinguir materiales y coberturas con comportamientos espectrales contrastantes. Sin embargo, el uso directo de las bandas puede presentar cierta ambigüedad entre coberturas con respuestas espectrales similares. Por esta razón, se incorporó la banda Infrarroja (Figura 12A) y el NDVI (Figura 12B) como variable derivada, debido a su capacidad para resaltar la presencia y condición relativa de la vegetación mediante la relación entre las respuestas espectrales del infrarrojo cercano y el rojo. De esta manera, el conjunto de datos empleado en la clasificación estuvo constituido por las bandas espectrales de NICFI y el NDVI, generando una representación multivariable de la superficie.

La integración se realizó mediante la combinación de las bandas originales y las variables espectrales derivadas en una misma imagen multibanda. Esta configuración permitió que el algoritmo de clasificación analizara simultáneamente la información espectral original y la información transformada mediante el índice de vegetación, en lugar de efectuar la clasificación a partir de una única banda o de un número reducido de variables.

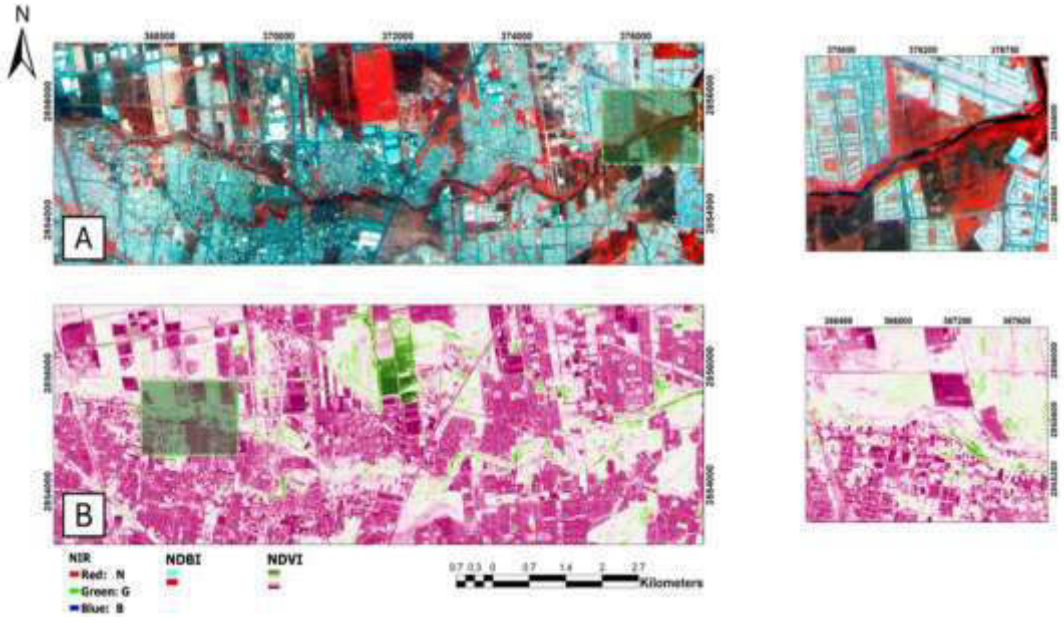


Figura 14. (A) Imagen compuesta NIR+GB. (B) Índice NDVI

Fuente: Elaboración propia con datos obtenido de imágenes NICFI y Sentinel-2

El procedimiento puede considerarse una estrategia de enriquecimiento del espacio de características, en la que cada píxel queda representado mediante un conjunto de atributos espectrales utilizados posteriormente como variables predictoras durante el proceso de clasificación.

4.6 Validación

Con la finalidad de obtener mejores resultados en los modelos de clasificación de uso de suelo en el sitio de interés dentro del AMM se hizo una validación con datos obtenidos por medio de las plataformas Google Earth Pro, Sentinel-2 y Street View. En total, se validaron 100 puntos aleatorios para cada una de las temporadas analizadas en este proyecto. Estas muestras de control permitieron evaluar la confiabilidad y la precisión de la clasificación de uso de suelo obtenida a través de los modelos aplicados en la zona urbana. Cabe resaltar que la variedad de datos en la foto interpretación hizo que el modelo fuera mucho más veraz y exacto a las condiciones reales de cada una de las temporadas estipuladas (Jeong., 2017). A continuación, en la Tabla 8 se presenta un esquema general de las imágenes utilizadas, sus datos y las temporadas en las que fueron analizadas para el proceso de validación para el presente proyecto.

Tabla 8. Validación

Fuente de datos	N° de datos	Temporadas
Google Earth Pro	280	2016-2017,2017-2017, 2017-2018
Sentinel-2	120	2018-2018,2018,2019
Street View	100	2016-2019

Total de datos validados	**200**

Fuente: Elaboración propia

La incorporación de datos de referencia de múltiples sensores y plataformas enriqueció significativamente el proceso de validación. Al combinar la información de imágenes de satélite de alta resolución, con los detalles proporcionados por las vistas a nivel de calle de Google Street View, se logró una caracterización del entorno urbano. Este enfoque otorgó una mayor solidez y credibilidad a los resultados, asegurando que las conclusiones sobre los patrones de uso de suelo reflejaran la realidad observable en el AMM.

Por otra parte, se elaboró un mapa de Voronoi relacionado con el área de estudio y los puntos aleatorios, esto con la finalidad de evaluar la representatividad espacial de los puntos generados en función de la clasificación LULC. En la siguiente Figura 15 se observa lo comentado.



Figura 15. Mapa de Voronoi

Fuente: Elaboración propia

Evaluar la representatividad espacial de los puntos generados en función de la clasificación LULC es una técnica para garantizar que los datos muestreados capturen de manera adecuada la diversidad y proporción de las clases de uso y cobertura del suelo presentes en el área de estudio. Este análisis permite identificar si los puntos aleatorios están distribuidos de manera equitativa en todas las clases y si reflejan correctamente su proporción en el territorio. Por ejemplo, si una clase ocupa una porción significativa del área, pero está

subrepresentada en los puntos generados, los resultados de los análisis posteriores, como validación o interpolación, podrían estar sesgados.

Luego de adquirir un conjunto de datos compuesto por temporada, se seleccionaron aleatoriamente algunos de ellos junto con sus valores de respuesta para realizar una nueva validación. Esta estrategia permitió comparar los datos de respuesta generados por el modelo con diversas imágenes satelitales y de campo, añadiendo un nivel adicional de precisión y robustez a los algoritmos de aprendizaje automático empleados en la clasificación LULC.

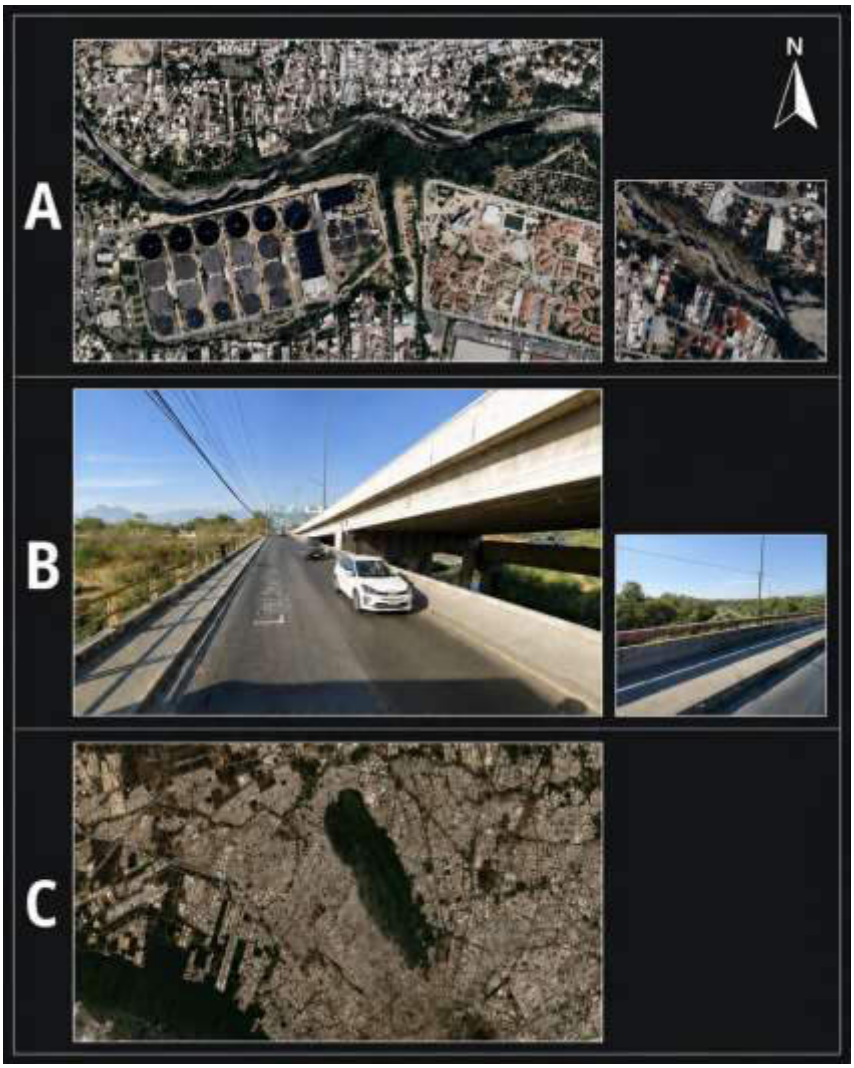


Figura 16. Imágenes de validación (A) Imagen Google Earth Pro. (B) Google Street View. (C) Imagen satelital Sentinel-2

Fuente: Elaboración propia

En cada una de estas imágenes se superpusieron los puntos aleatorios de validación (Figura 17) con el propósito de, mediante fotointerpretación, obtener una comparación de cada una de las clases presentes en los modelos presentados en este proyecto.

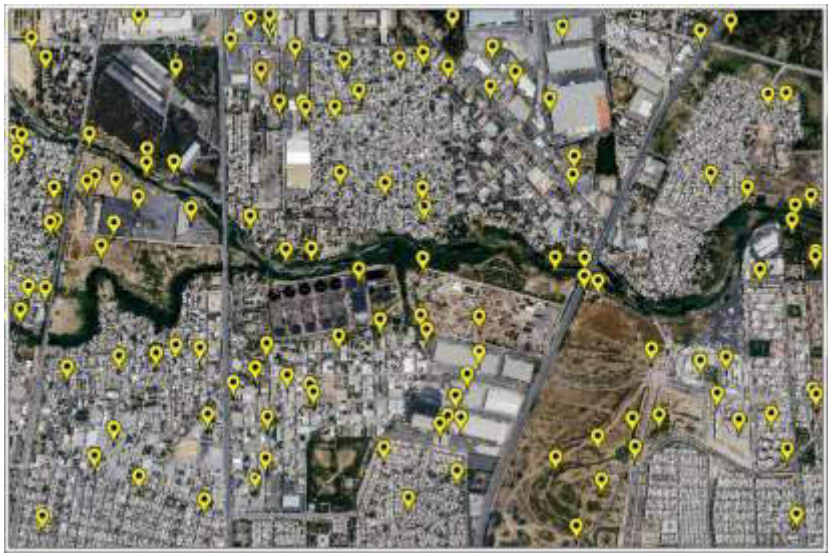


Figura 17. Puntos de validación con Google Earth Pro

Fuente: Elaboración propia

Mediante el uso de imágenes provenientes de Google Earth Pro, Street View y Sentinel-2, se generó una base de datos consistente en 100 puntos por temporada, acumulando un total de 500 datos. Estos datos se emplearon para comparar los resultados obtenidos con el primer modelo, con el objetivo de incrementar su veracidad. Este proceso no solo sirvió para evaluar el modelo inicial, sino que también permitió realizar ajustes en algunos puntos de entrenamiento y validación, mejorando así su desempeño. En este sentido, la validación desempeñó un papel crucial al refinar las predicciones del modelo inicial, aumentando su precisión y confiabilidad al incorporar una base de datos más sólida y una comparación directa con fuentes de alta resolución y referencia.



Figura 18. Validación

Fuente: Elaboración propia

Para implementar una evaluación más detallada y precisa de la clasificación de uso de suelo obtenida mediante los algoritmos de aprendizaje automático, se desarrolló un enfoque basado en Random points (puntos aleatorios). Este enfoque consistió en generar puntos distribuidos aleatoriamente sobre cada una de las imágenes satelitales correspondientes a las diferentes temporadas analizadas (2016-2019). Estos puntos aleatorios fueron considerados como muestras representativas del área de estudio, con el fin de reevaluar y mejorar la precisión de los algoritmos de clasificación. Para crear los puntos aleatorios se utilizó un código de lenguaje Javascript en GEE el cual describe la cantidad de puntos aleatorios a seleccionar tal y como se muestra en la siguiente Figura 19 (ejemplo). El código se encuentra en la metodología de este documento.

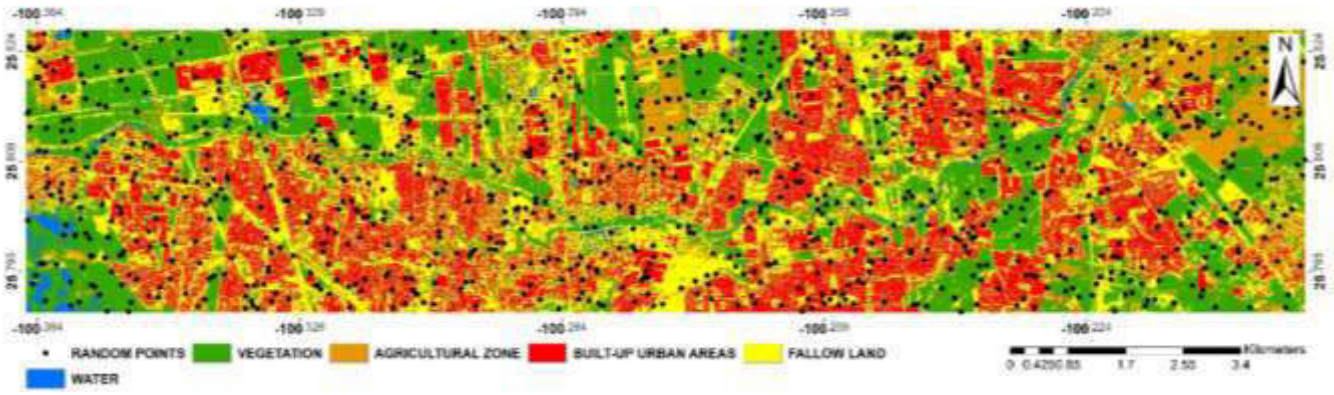


Figura 19. Random points

Fuente: Elaboración propia

Después de generar los puntos aleatorios para cada temporada, se procedió a extraer la información de la clase de uso de suelo asignada por los algoritmos de ML en cada uno de estos puntos. De esta manera, en la Figura 20 se observa el código para obtener el conjunto de 200 datos de observación (50 % validados) o muestras etiquetadas por cada respuesta del modelo.

```
//random points
var puntosSNG= ee.FeatureCollection.randomPoints(geometry,200)
Map.addLayer (puntosSNG);
Export.table.toAsset({
  collection:puntosSNG,
  description:'exportarAssets',
  assetId:'puntosSNG'
});
```

Figura 20. Código Random points en GEE

Fuente: Elaboración propia

Es importante mencionar que este proceso de extracción de datos se llevó a cabo utilizando el software ArcGIS PRO, una herramienta ampliamente utilizada en el campo de los SIG. ArcGIS PRO permitió realizar operaciones espaciales y extracción de la información sobre las imágenes clasificadas y los puntos aleatorios generados, facilitando la recopilación de la información de uso de suelo en cada uno de los puntos.

Gracias a esta metodología, se logró obtener un conjunto de datos representativo del área de estudio, con muestras distribuidas aleatoriamente y etiquetadas según la clasificación obtenida. Estos conjuntos de datos fueron importantes para poder realizar una validación de los modelos de ML elaborados anteriormente, con el fin de entrenar un nuevo modelo de clasificación supervisada y evaluar su rendimiento de manera más precisa.

4.7 Evaluación de precisión

Con el fin de evaluar el desempeño y determinar el algoritmo óptimo para la predicción de los diferentes usos de suelo y la identificación (continuidad) de ríos intermitentes y de poco flujo en la zona urbanizada del río Pesquería, se generaron matrices de confusión, se calcularon los Índices Kappa y las precisiones generales para cada uno de los modelos de clasificación supervisada. Este análisis riguroso se realizó en Python y en GEE, aplicado a las distintas imágenes satelitales correspondientes a las temporadas estiaje y de lluvias. De esta manera, fue posible obtener una cuantificación del grado de

precisión de cada modelo, permitiendo seleccionar el algoritmo de ML más adecuado para mapear con alta exactitud las diversas coberturas terrestres presentes en el área de estudio.

A continuación, se observan las ecuaciones empleadas para evaluar la precisión de los modelos.

$$\text{Precisión General} = \left(\frac{C_p}{T_p}\right) x100$$

Ecuación 8. Precisión general

Donde: C_p indica los pixeles clasificados correctactemente y T_p indica el total de píxeles empleados.

$$K=\frac{(TSxTCS)-\sum(Columna\ Total\ x\ \ Totalde\ filas)}{TS^2-\sum(Columna\ Total-\ Total\ de\ filas)}$$

Ecuación 9. Índice Kappa

Donde: TS es el número total de píxeles de referencia y TCS es el número total de píxeles clasificados correctamente.

Para la precisión general mientras más alto sea el porcentaje de precisión (cercano a 100 %), mayor será la confianza en que los resultados de la clasificación LULC reflejan fielmente la realidad en el terreno. Para el Índice Kappa un valor cercano a 1 indica un alto grado de concordancia, mientras que un valor próximo a 0 sugiere que el nivel de acuerdo no supera el que se obtendría por azar.

El cálculo de estas dos métricas estadísticas, el Índice Kappa y la precisión general, constituyen una evaluación de precisión de la calidad y solidez de los modelos de clasificación LULC implementados en este proyecto para el AMM y

específicamente el río Pesquería. Estos indicadores proporcionaron un sólido respaldo cuantitativo a las conclusiones alcanzadas sobre los patrones y dinámicas de uso y cobertura del suelo en la región.

Tabla 9. Ecuaciones de precisión

Método	Ecuación
OA	$\frac{Number\ of\ Correctly\ Classified\ Samples}{Number\ of\ Total\ Samples}\ x\ 100$
Kc	$\frac{p\ (correct\ classification) - p(classification\ probability)}{1 - p(classification\ probability)}$
UA	$\frac{Number\ of\ Correctly\ Classified\ samples\ in\ each\ class}{Number\ of\ Samples\ Classified\ to\ that\ class}$
PA	$\frac{Number\ of\ Correctly\ Classified\ samples\ in\ each\ class}{Number\ of\ Samples\ from\ Reference\ Data\ in\ each\ Class}$
F1 Score	$\frac{2\ x\ UA\ \ x\ PA}{UA + PA}$

Fuente: Elaboración propia con información obtenida de Bayable et al. (2025).

4.8 Calidad de Agua

De igual manera que con las bandas espectrales, se realizó en GEE un código automatizado para obtener los datos de los índices espectrales estudiados para cada uno de los puntos de muestreo en las diferentes fechas seleccionadas. Este código permite descargar en formato CSV los valores de los índices junto con el punto muestreado, su fecha y su coordenada geoespacial. A continuación, se presentan los índices espectrales contemplados en este estudio

Tabla 10. Índices espectrales

Índices	Nombre de banda	Implementación en Sentinel-2 (bandas MSI)
NDVI	NIR - Red	((B8 - B4) /(B8 + B4))
NDWI	Green- NIR	((B3 - B8)/(B3 + B8))
MNDWI	Green - SWIR1	((B3 – B11)/(B3 + B11))

NDTI	Red - Green	((B4 - B3)/(B4 + B3))
NDCI	Red-Edge-Red	((B5 - B4)/(B5 + B4))
FAI	NIR – Red-SWIR1	(B8 - B4 + (B11 - B4) \ t

4.8.1 Datos in situ

Para caracterizar la calidad del agua, en esta investigación se utilizaron mediciones in situ y de laboratorio provenientes de la base de datos pública de CONAGUA para el período 2016–2019 en el río Pesquería. Los parámetros considerados fueron: Color, DQO TOT, N NH3, N NO2, N NO3, N ORG, N TOT, N TOTK, OD, P TOT, pH_CAMPO, SDT, SST y TURB.

Se realizó una limpieza de la base de datos de CONAGUA por medio de Python con la finalidad de obtener datos representativos. Para esto aplicó un filtro automático de todos los parámetros medidos por la Comisión, de los cuales se retuvo únicamente los parámetros con el 80 % de datos específicos dentro de los periodos estudiados. Los parámetros que no cumplieron dicho umbral fueron excluidos de los análisis posteriores.

4.8.2 Análisis de datos

Para el análisis de datos se utilizó el entorno de programación Python 3.12 como lenguaje base. El desarrollo y ejecución del código se realizó en el IDE Spyder, aprovechando su interfaz integrada para análisis de datos. Se empleó un conjunto de bibliotecas especializadas en manejo y visualización de datos, así como en cálculos estadísticos y aprendizaje automático como lo son pandas (manipulación de datos en estructuras tipo DataFrame), numpy (operaciones numéricas de alto rendimiento sobre arreglos), matplotlib (generación de gráficos estáticos), scikit-learn (implementación de transformaciones, algoritmos de aprendizaje y utilidades de procesamiento) y scipy (funciones estadísticas y científicas de apoyo). Spyder fue base para el desarrollo del análisis de datos automatizado.

El objetivo central de este capítulo fue evaluar detalladamente todas las combinaciones posibles de transformaciones estadísticas aplicadas tanto a las variables de las bandas espectrales e índices espectrales junto con los datos muestreados de calidad de agua y así cuantificar la correlación entre ellas. Esto con la finalidad de entender el comportamiento lineal de la diversidad espectral obtenida con Sentinel-2 y los parámetros fisicoquímicos de calidad de agua en ríos urbanos de bajo flujo. Los datos y transformaciones usadas se presentan a continuación:

- **Sin transformación ("none")**: Se utiliza directamente el valor original de la variable. Esta opción sirve como línea base para comparar si las transformaciones no lineales aportan alguna mejora en la correlación.
- **Transformación logarítmica**: Consiste en aplicar el logaritmo natural a la variable después de sumar 1: ln (1 + x). Esta forma denotada *log1p* permite estabilizar la varianza y reducir la asimetría en datos positivos, a su vez evita indefiniciones en x=0. La suma de 1 asegura que incluso valores originalmente cero puedan transformarse a un número finito.
- **Transformación raíz cuadrada**: Se aplica la raíz cuadrada al valor. Al igual que el logaritmo, la raíz cuadrada es una transformación monótona que puede reducir la asimetría en distribuciones sesgadas hacia la derecha. En el contexto de calidad de agua e índices espectrales, se asume que las variables de interés son no negativas o bien se toman valores absolutos cuando corresponde.
- **Transformación de Yeo-Johnson**: Es una familia de transformaciones de potencia con parámetro Lambda que admite valores cero y negativos, propuesta por In-Kwon Yeo y Richard Johnson (2000). La transformación Yeo–Johnson extiende el enfoque de Box–Cox para que sea aplicable en toda la recta real (incluyendo datos negativos). Su definición es de forma *piecewise* dependiendo del signo del dato original.

$$\psi(y,\lambda) = \begin{cases} \dfrac{(y+1)^{\lambda} - 1}{\lambda} & y \geq 0, \quad \lambda \neq 0 \\ \ln(y+1) & y \geq 0, \quad \lambda = 0 \\ -\dfrac{(-y+1)^{2-\lambda} - 1}{2-\lambda} & y < 0, \quad \lambda \neq 2 \\ -\ln(-y+1) & y < 0, \quad \lambda = 2 \end{cases}$$

Ecuación 10. Yeo-Johnson

Donde:

ψ(y, λ): valor transformado de la observación, es decir, la nueva variable que se usará en el análisis.

y: valor original de la variable a transformar; puede ser positivo, cero o negativo (y ∈ ℝ).

λ: parámetro de la transformación, que controla la intensidad y el tipo de cambio aplicado. Se estima a partir de los datos, normalmente por máxima verosimilitud, buscando el valor que hace la distribución lo más simétrica/normal posible.

ln: logaritmo natural (base e).

El proceso consistió en aplicar cada una de las transformaciones mencionadas anteriormente a cada banda espectral utilizada de Sentinel-2, junto con las combinaciones establecidas en este estudio y para cada uno de los índices espectrales, esto con la finalidad de posteriormente identificar cual de todas las trasformaciones presentó mejores o peores correlaciones con los datos muestreados in situ.

4.8.3 Correlación de Pearson

Para determinar la relación lineal entre las bandas espectrales y cada uno de los parámetros de calidad del agua, se calculó el coeficiente de correlación de Pearson entre cada variable independiente y cada variable dependiente, definido por la ecuación:

$$rx, y = \frac{\sum_{i=1}^{n}(X_i - X^-)\,(Y_i - Y^-)}{\sqrt{\sum_{i=1}^{n}(X_i - X^-)^2}\,\sqrt{\sum_{i=1}^{n}(Y_i - Y^-)^2}}$$

Ecuación 11. Matriz de correlación de Pearson

Donde:

$r_{X,Y}$ representa la fuerza y dirección de la relación lineal entre dos variables (Benesty et al., 2009). Los resultados se visualizaron mediante gráficos de barras y una matriz de correlación tipo heatmap, permitiendo la identificación rápida de asociaciones relevantes.

4.8.4 Normalización de datos

Se implementó una normalización estándar (Z-score) para escalar los datos y evitar que las diferencias en magnitudes influyeran en los modelos. La fórmula utilizada fue:

$$Z_i = \frac{X_i - \mu x}{\sigma x}$$

Ecuación 12 Normalización z-core

Donde:

μX_es la media y σX la desviación estándar de la variable X. Este proceso asegura que cada variable tenga media cero y varianza unitaria, lo cual es especialmente útil para algoritmos sensibles a la escala (Han et al. 2012).

4.8.5 Mejora del Modelo

Winsorización

La winsorización se aplicó como un procedimiento de tratamiento estadístico orientado a disminuir la influencia de valores extremos sobre la distribución de las variables, sin eliminar las observaciones del conjunto de datos. En una primera etapa se realizó la revisión de la distribución de los datos y la identificación de valores con comportamiento extremo mediante estadísticos descriptivos y medidas de posición (Dixon, 1960). A partir de esta evaluación se

establecieron límites inferior y superior empleando los percentiles 10 y 90 en la distribución de datos utilizados (Turkey, 1962).

Los valores ubicados por debajo del límite inferior fueron sustituidos por el valor correspondiente al percentil inferior, mientras que aquellos que excedieron el límite superior fueron reemplazados por el valor definido en el percentil superior. Las observaciones comprendidas dentro de ambos límites conservaron sus valores originales. De esta forma, el procedimiento permitió reducir el peso estadístico de las observaciones extremas sobre medidas como la media, la desviación estándar y otros análisis posteriores, manteniendo simultáneamente el tamaño de la muestra y la estructura general de los registros. Finalmente, se verificaron los estadísticos descriptivos de las variables para evaluar el efecto de la transformación y comprobar que la reducción de la influencia de los valores extremos no implicara una modificación sustancial de la distribución de los datos (Blaine, 2018).

Validación Cruzada

La validación cruzada aleatoria mezcla observaciones de una misma estación entre las particiones de entrenamiento y prueba (Roberts et al., 2017). Con siete estaciones y diecinueve fechas, esa mezcla permite al modelo reconocer el nivel basal de cada sitio a partir de observaciones de otras fechas del mismo punto, lo que eleva el coeficiente de determinación sin que exista capacidad predictiva transferible. Se sustituyó por validación cruzada agrupada por fecha de muestreo, de modo que todas las observaciones de una fecha quedaran íntegramente en una sola partición. Adicionalmente se calculó la validación por exclusión de estación (leave-one-station-out), que mide la transferibilidad espacial y constituye la condición necesaria para aplicar el modelo a píxeles no muestreados (Meyer et al., 2018).

4.9 Aplicación Web

Esta sección comprende el proceso general del desarrollo de una plataforma Web interactiva para el monitoreo de la calidad del agua del río Pesquería, Nuevo León, México en su sección urbana y para el análisis espectral de cualquier área de estudio del planeta Tierra. Dicha plataforma integra el modelo predictivo de ML implementado en esta investigación, previamente entrenado con datos de campo, imágenes satelitales por medio de la unión entre la plataforma GEE y el lenguaje de programación de Python en Spyder. El proceso metodológico abarcó la unión de toda la información aplicada en Python, la serialización del modelo RF en formato portable (.pkl), la gestión del repositoriode código en GitHub, el despliegue automático de la aplicación en Streamlit Community Cloud y por último, la estructuración de los módulos y secciones, interfaces, backend y fronted de la plataforma. En la siguiente figura se resume el flujo general de desarrollo para la página Web.

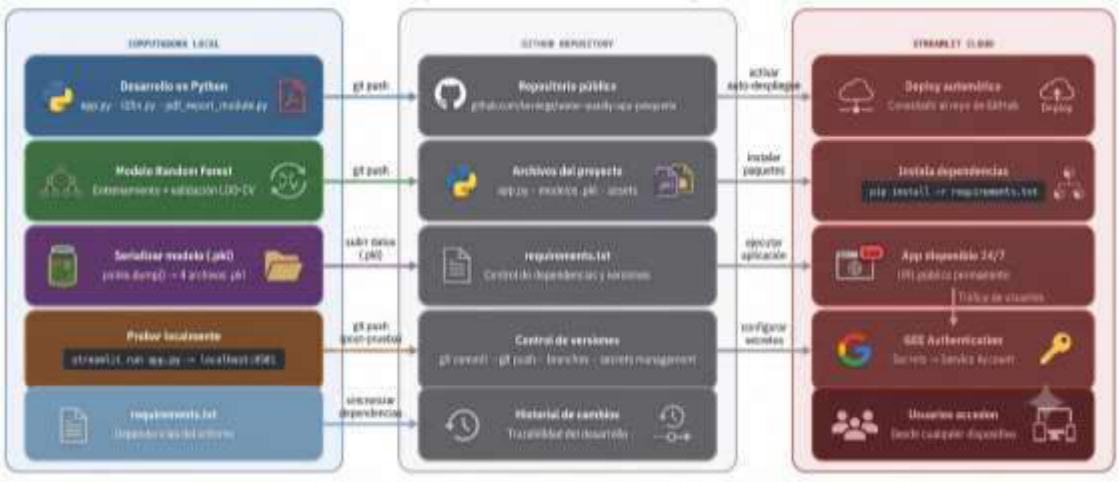


Figura 21. Flujo de desarrollo de la plataforma Web: computadora local - repositorio GitHub -Streamlit Community

Fuente: Elaboración propia.

Para el desarrollo y elaboración de esta página Web se seleccionó el framework Streamlit (versión≥1.35), una librería de código abierto para Python que permite construir aplicaciones web analíticas e interactivas directamente desde scripts de Python, sin la necesidad de conocimientos avanzados en HTML,

CSS, JavaScript o cualquier framework de mayor complejidad de backend. Para este proyecto Doctoral y para los alcances de esta investigación se seleccionó Streamlit como página Web oficial de este estudio.

4.9.1 Configuración entorno de desarrollo local

El desarrollo de este proceso se realizó en un entorno Python con las siguientes configuraciones previas al despliegue en la nube:

1. Instalación del entorno virtual en Python. Se creó un ambiente conda para aislar las dependencias del proyecto del sistema operativo.
2. Instalación de Streamlit y dependencias: todas las librerías necesarias fueron instaladas mediante pip y registradas en el archivo "requirements.txt" para garantizar la reproducibilidad del entorno.
3. Configuración de la autenticación con GEE, se generó una cuenta de servicio (Service account) en Google Cloud Platform con acceso a la API de Earth Engine. Las credenciales JSON se almacenaron localmente en el archivo ./streamlit/secrets.toml y eb los secrets de Streamlit Cloud para el entorno de producción.
4. Prueba local de la aplicación. Antes de cada subida al repositorio GitHub, la aplicación fue ejecutada localmente con el comando streamlit run app.py, que muestra la interfaz en el navegador en http://localhost:8501 para verificación

```
# Comando para ejecutar la app localmente
$ streamlit run app.py

# Estructura del archivo secrets.toml (autenticacion GEE)
[gee]
service_account = "nombre@proyecto.iam.gserviceaccount.com"
private_key     = "-----BEGIN PRIVATE KEY-----..."
```

Figura 22. Script base de prueba local de la aplicación

Fuente: Elaboración propia

4.9.2 Arquitectura del Código principal (app.py)

El archivo principal app.py centraliza todo el método de la aplicación en aproximadamente 4049 líneas de código Python, organizadas en las siguientes secciones funcionales:

- Configuración de página y estilos CSS: Definición del layout, wide, título, favicon e inyección de estilo CSS personalizados (colores, animaciones, componentes glass) mediante st.markdown().
- Gestión de estado de idioma: inicialización del estado de sesión (st.session_state) con el idioma seleccionado, cuya consulta al módulo il8n.py logra renderizar todos los textos en español, inglés y portugués junto con sus reportes y series temporales.
- Carga de assets e imágenes institucionales: lectura de logos (logo_geomatica.png, logo_uanl.png, logo_fic.png) y foto del investigador de esta tesis Doctoral en formato base64 para su integración en la interfaz.
- Funciones auxiliares: make_cmap) para paletas de color, zona_es_pesqueria() para detección automática del área de estudio, y funciones de interpolación espacial RBF/IDW.
- Sidebar interactivo: barra lateral con el flujo de trabajo guiado, carga del shapefile (el que el usuario adjunte), selectores de fecha, slider de nubes, multiselect de parámetros y selector de idioma.
- Secciones principales: renderizado secuencial de las cuatro etapas (Hero, Área de estudio, Mapas e índices espectrales, series temporales y reportes) condicionado al usuario en el flujo de trabajo.

4.9.3 Módulo de internacionalización

La plataforma fue diseñada desde el inicio con soporte multilingüe a través del módulo il8n.py, que implementa un sistema de traducción basado en diccionarios Python. Este módulo contiene todas las cadenas de texto de la interfaz organizadas por clave de traducción y disponibles en tres idiomas: Español, Inglés y Portugués. La función principal t(key) devuelve el texto correspondiente al idioma seleccionado por el usuario. El cambio de idioma es inmediato y no

requiere actualizar la página, puesto que Streamlit re-ejecuta el script completo en cada interacción del usuario, renderizando toda la interfaz en el nuevo idioma seleccionado por quien ingrese a la página web.

```
# Uso del modulo i18n en app.py
from i18n import t, IDIOMAS, get_param_label

# Ejemplo de uso — todos los textos pasan por t()
st.title(t("app_title"))
st.selectbox(t("select_parameter"), options=PARAMS)
```

Figura 23. Modulo idiomas

Fuente: Elaboración propia

4.9.4 Módulo para reportes PDF

La generación automática de reportes en formato PDF se implementó en el módulo pdf_report_module.py utilizando la librería ReportLab. Este módulo contiene tres funciones principales de generación de reportes, cada una obtenible desde la interfaz de la aplicación Web:

- Generar_pdf_fecha_unica(): genera un reporte de calidad de agua para la fecha específica, incluyendo el mapa espacial de predicciones RF, tabla de valores por punto de muestreo, comparativa contra límites NOM-001 y estadísticas descriptivas
- Generar_pdf_serie_temporal(): genera un reporte con el temporalidad 2016-2019 de kis parámetros seleccionados, gráficos de tendencia por punto de muestreo, tabla resumen estadístico e interpretación automática de los resultados.
- Generar_pdf_reporte_espectral(): genera un reporte con los mapas de los nueve índices espectrales usando sentinel-2 seleccionados por el usuarios y las series temporales de los mismos, para cualquier área geográfica del mundo.

4.9.5 Integración del Modelo RF en la plataforma Web

El modelo de aprendizaje automático es fundamental para el proceso de la aplicación Web. En la Figura 24 se observa el proceso de integración del modelo RF a la aplicación Web.

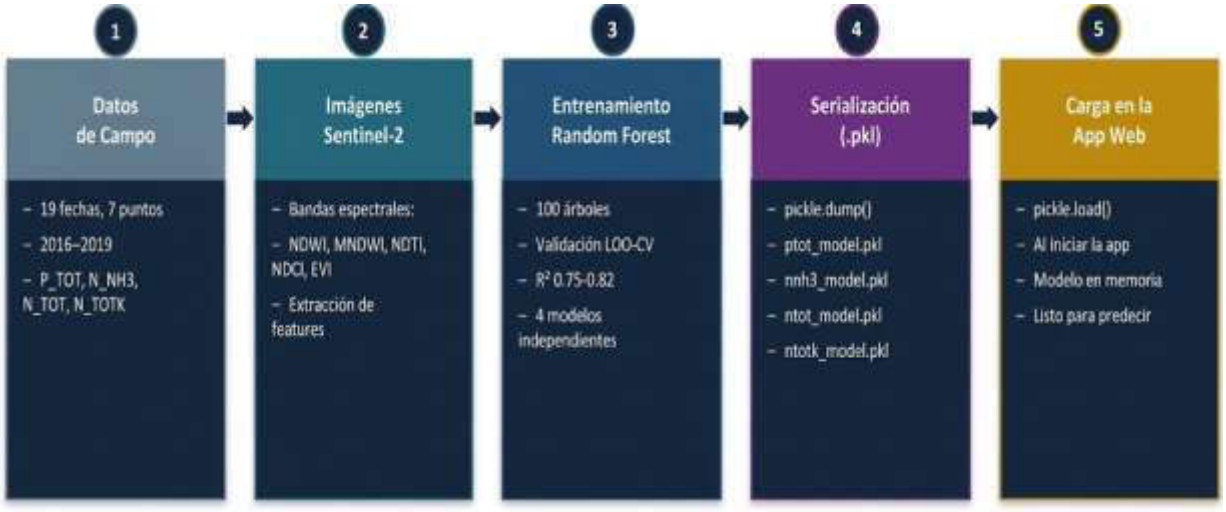


Figura 24. Pipeline completo del modelo RF desde los datos de campo hasta la predicción en tiempo real para el usuario de la plataforma web. Los archivos. pkl serializados actúan como puente entre el entrenamiento local y la app desplegada en la nube

Fuente: Elaboración propia.

4.9.6 Serialización del modelo en formato .pkl

Al momento de obtener los modelos de aprendizaje automático, fue serializado (guardado) en formato binario mediante la librería pickle de Python, que convierte el objeto del modelo scikit-learn en un archivo portable con extensión .pkl. Este proceso genera cuatro archivos independientes (uno por parámetro) que guardan el modelo entrenado, sus parámetros internos y el preprocesador de features, listos para ser cargados y utilizados para hacer predicciones automatizadas sin necesidad de re-entrenar

```
import pickle

# Guardar el modelo entrenado como archivo. pkl
with open ("ptot_model.pkl", "wb") as f:
    pickle. dump (rf_model_ptot, f)

# Resultado: archivo binario listo para subir a GitHub
# ptot_model.pkl nnh3_model.pkl ntot_model.pkl ntotk_model.pkl
```

Figura 25. Import pickel serialización

Fuente: Elaboración propia

Los archivos .pkl generados fueron subidos al repositorio de GitHub junto con el código de la aplicación. Esto es posible porque los archivos tienen un tamaño manejable dentro de los límites de almacenamiento de archivos en GitHub. Al desplegarse la app en Streamlit Cloud, estos archivos son descargados automáticamente desde el repositorio y cargados en memoria al iniciar la sesión.

4.9.7 Gestión del proyecto en GitHub

GitHub fue utilizado como plataforma principal de control de versiones, almacenamiento del proyecto y punto de conexión entre el desarrollo local y el despliegue en la nube. De esta manera, el repositorio GitHub contiene todos los archivos necesarios para ejecutar la aplicación, organizados en la matriz del repositorio tal como lo requiere Streamlit Community Cloud. La siguiente figura muestra la estructura de archivos del proyecto.

Estructura del Repositorio GitHub — water-quality-app-pesqueria

github.com/kevinrgz/water-quality-app-pesqueria

water-quality-app-pesqueria/	[RAIZ DEL REPOSITORIO]
app.py	(~4,049 lineas)
i18n.py	(Soporte ES / EN / PT)
pdf_report_module.py	(Generador de reportes PDF)
requirements.txt	(Dependencias del entorno)
ptot_model.pkl	(Modelo RF - Fosforo Total)
nnh3_model.pkl	(Modelo RF - N-Amoniaco)
ntot_model.pkl	(Modelo RF - N-Total)
ntotk_model.pkl	(Modelo RF - N-Total Kjeldahl)
logo_geomatica.png	(Logotipo institucional)
logo_uanl.png / logo_fic.png	(Logos institucionales)
photo_researcher.png	(Foto del investigador)
.streamlit/secrets.toml	(Credenciales GEE - privado)

Estructura del Repositorio GitHub — water-quality-backend-service

github.com/kevinrgz/water-quality-backend-service

water-quality-backend-service/	[RAIZ DEL REPOSITORIO]
main.py	(API con FastAPI)
routers/inference.py	(Endpoints de prediccion)
models/rf_regressor.pkl	(Modelo RF (Regresion))
requirements.txt	(Librerias principales)
dockerfile	(Definicion de contenedor)
training/train_model.py	(Script de entrenamiento offline)
assets/architecture_diagram.png	(Diagrama de microservicios)
assets/api_spec.yaml	(Especificacion OpenAPI v3.0)
config/.env.secret	(Claves de API - no commit)
deployment/deploy_to_aws.sh	(Automatizacion de despliegue)

Figura 26. Estructura de archivos del repositorio GitHub

Fuente: Elaboración propia.

4.9.8 Gestión de credenciales y archivos sensibles

Las credenciales de autenticación con GEE (Service Account y clave privada JSON) no fueron incluidas en el repositorio público por razones de seguridad. Localmente, se gestionaron mediante el archivo .Streamlit/secrets.toml (excluido del repositorio con. gitignore). En el entorno de producción (Streamlit Cloud), las mismas credenciales fueron configuradas en la sección "secrets" del panel de administración de la app, que las incluye como variables de entorno seguras en tiempo de ejecución, accesibles desde el código mediante st.secrets["gee"].

5RESULTADOS Y DISCUSIONES

PORBLEMATICAS SOCIOAMBIENTALES

5.1 Análisis general de la cuenca del río Pesquería

Para esta sección se implementó una evolución mensual de las covariables ambientales entre 2000 y 2025 con la finalidad de conocer la estructura ambiental de la cuenca. Este análisis muestra una marcada variabilidad estacional e interanual en la cuenca del río Pesquería. La precipitación estimada mediante CHIRPS presenta el comportamiento más irregular de la serie, con periodos prolongados de baja precipitación interrumpidos por pulsos de elevada magnitud. La mayor parte de los meses registra acumulados relativamente bajos, mientras que durante la temporada húmeda se observan incrementos en algunos años, superando los 150–200 mm mes^{-1}. Los eventos asociados con los ciclones tropicales destacan dentro de esta dinámica, particularmente Alex en 2010 y Alberto en 2024, cuyos periodos coinciden con aumentos abruptos de la precipitación a escala de cuenca. La amplitud del intervalo espacial entre los percentiles 10 y 90 también aumenta durante estos episodios, lo que indica que las precipitaciones intensas no se distribuyeron uniformemente en todo el territorio.

Por consiguiente, la media móvil de 12 meses permite distinguir con mayor claridad los periodos persistentes de déficit hídrico frente a las fluctuaciones mensuales. Durante la sequía de 2011–2013 se observa una reducción de la precipitación después del episodio extraordinariamente húmedo registrado en 2010. Una condición semejante se identifica durante el periodo asociado con la Triple La Niña y la crisis hídrica de Monterrey, particularmente alrededor de 2022. Aunque dentro de estos intervalos ocurrieron algunos pulsos de precipitación, estos fueron de corta duración y no modificaron inmediatamente la señal hidroclimática acumulada, evidenciando la importancia de distinguir entre eventos individuales de lluvia y la persistencia temporal de condiciones secas. Se reconoce de igual manera que esta región al ser semi árida coincide con sus condiciones y con una particularidad de bajas precipitaciones en la mayoría de los meses y muy altas precipitaciones en meses como septiembre.

Con respecto al análisis realizado con LST, este presenta un ciclo anual muy definido, con valores mínimos cercanos a 20–25 °C durante los meses fríos y máximos que frecuentemente alcanzan aproximadamente 40–45 °C durante la época cálida. La temperatura del aire a 2 m (T2M) reproduce el mismo patrón estacional, aunque con una amplitud considerablemente menor, oscilando aproximadamente entre 12 y 30 °C. Esta diferencia es consistente con el calentamiento radiativo directo de la superficie terrestre, que responde con mayor intensidad que la atmósfera a las variaciones estacionales de radiación solar y humedad superficial. A escala interanual, la media móvil de LST muestra cambios más pronunciados que la correspondiente a T2M, lo que sugiere que las condiciones de la superficie, incluyendo disponibilidad de humedad, cobertura vegetal y características del suelo desempeñan un papel importante en la respuesta térmica de la cuenca.

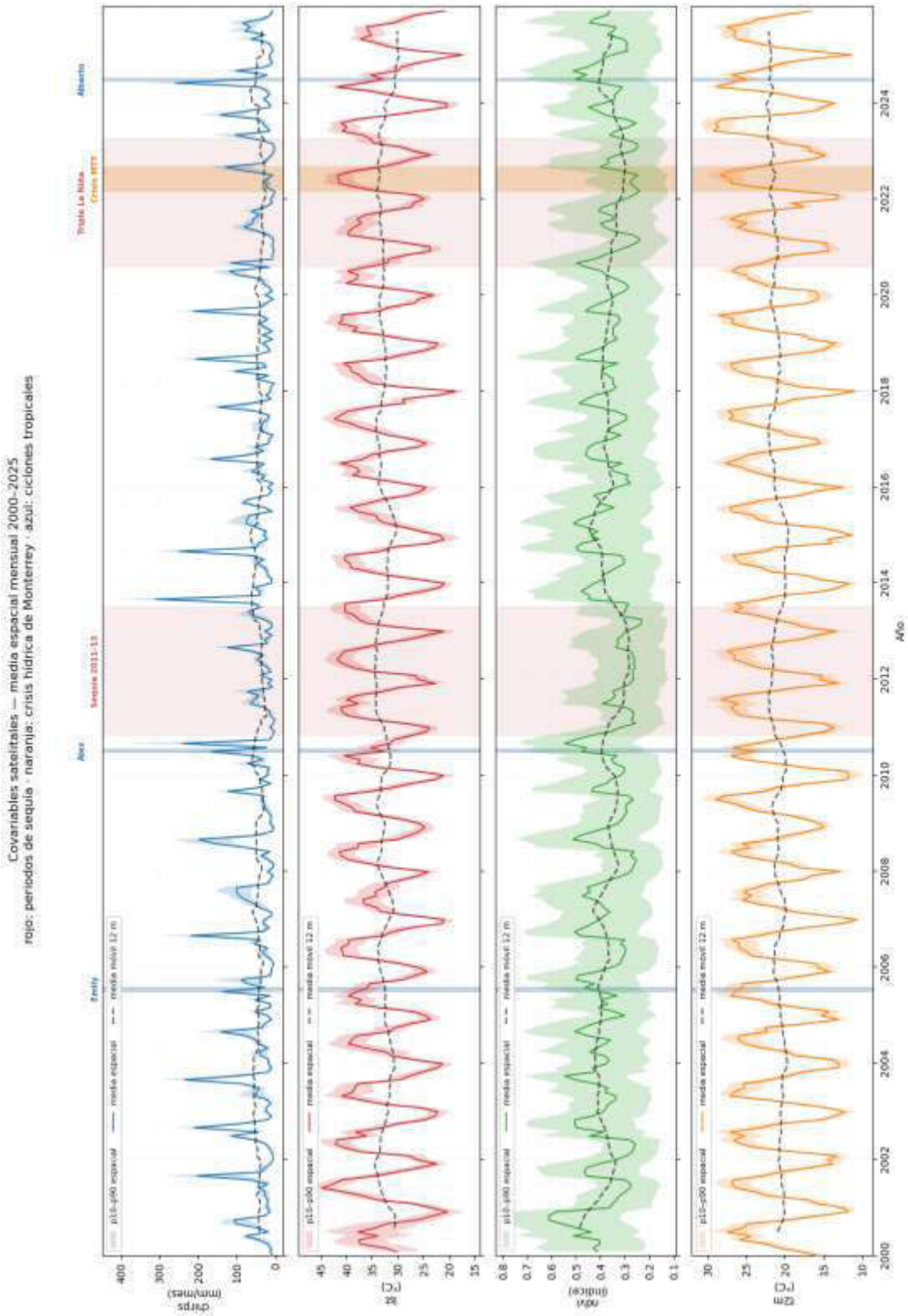


Figura 27. Serie Temporal CHRIPS, LST, NDVI y T2M

Fuente: Elaboración propia

A demás, también se estudió la serie temporal del índice NDVI desde 2000-2025 evidenciando igualmente una señal estacional, pero con una variabilidad espacial considerablemente mayor. Los valores medios se mantienen generalmente entre 0.25 y 0.55, mientras que en el intervalo P10–P90 se observa la heterogeneidad de la cobertura vegetal dentro de una cuenca donde coexisten áreas urbanizadas, suelo expuesto y superficies con distintos grados de vegetación. La media móvil de 12 meses muestra una disminución progresiva desde valores cercanos a 0.45–0.50 durante los primeros años de la serie hasta valores próximos a 0.30 durante el periodo 2011–2013. Esta reducción coincide temporalmente con condiciones de menor disponibilidad hídrica y con temperaturas superficiales elevadas, lo que refleja una disminución general de la actividad o cobertura fotosintéticamente activa. Posteriormente se observa una recuperación parcial del NDVI, seguida de una nueva disminución durante 2020–2022 y una recuperación hacia el final de la serie. Además, se observa una relación del buen estado de la vegetación con respecto a la disponibilidad hídrica de la zona.

En conjunto, las cuatro variables muestran una respuesta hidroclimática coherente. Los periodos caracterizados por menor disponibilidad de precipitación tienden a coincidir con reducciones del NDVI y mayores condiciones térmicas superficiales, mientras que los episodios de precipitación intensa generan cambios rápidos en la disponibilidad de agua, aunque la respuesta de la vegetación ocurre de manera más gradual. Esta diferencia temporal es particularmente relevante para la cuenca del río Pesquería.

5.1.1 ICU

Para comprender un poco más el medio ambiente y algunas características urbanistas que presenta el área de estudio se implementó un análisis que conlleva índices urbanos, de vegetación, climáticos e hidrológicos con la finalidad de obtener un Índice de Convergencia Urbano en la cuenca Pesquería. En la Figura 28 se observa cada uno de los sub índices calculados.

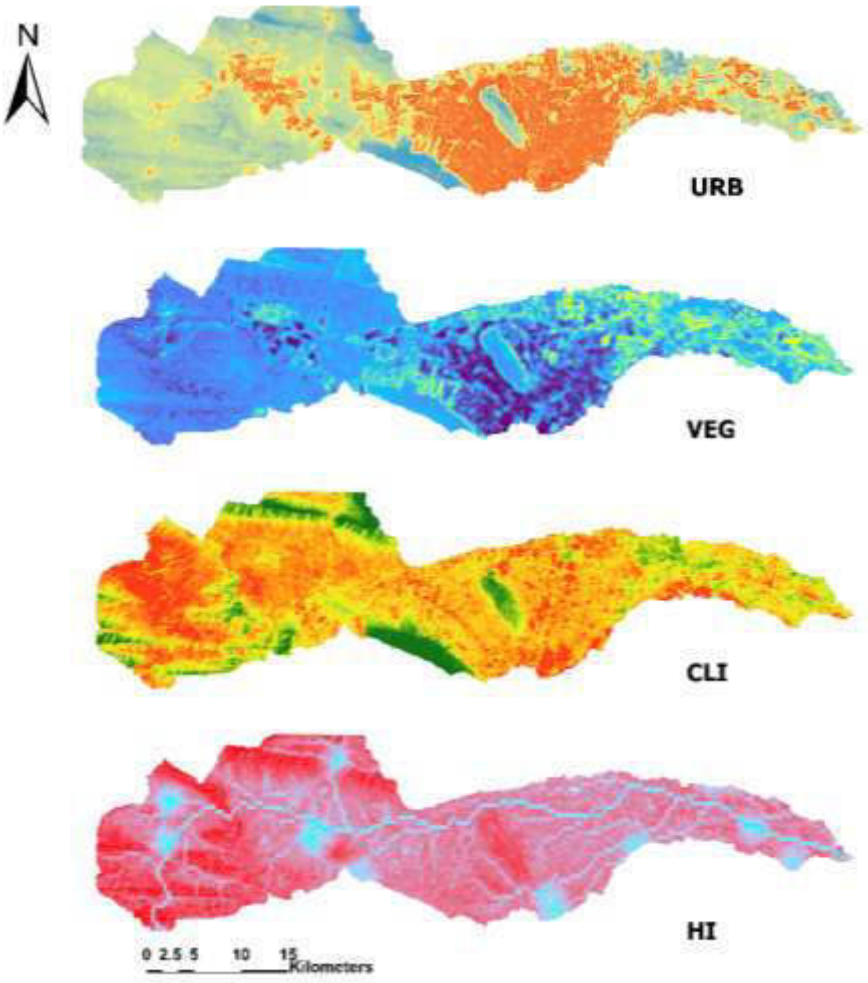


Figura 28. Subíndices espectrales para calcular ICU

Fuente: Elaboración propia

El sub índice urbano (URB) se alimentó de diversos datos urbanísticos como GHSL, NDBI, PBuilt (Dynamic World) y BordeUrb con la finalidad de identificar donde existe la mayor presencia e influencia de la urbanización y además de la heterogeneidad del entorno urbano. Los valores cercanos a 1 identificaron mayor presencia de superficie construida, mayor probabilidad de urbanización, respuesta espectral con superficies impermeables y de elevada heterogeneidad urbana. Por su parte un valor cercano a 0 corresponde a áreas con poca influencia urbana tales como vegetación, cuerpos de agua, zonas agrícolas o superficies naturales.

Con respecto al sub índice de vegetación (VEG) se aplicaron cálculos de mediana durante el periodo de estudio representando la condición vegetal relativamente típica del periodo, así como la variabilidad temporal correspondiente de la desviación estándar del NDVI el cual permitió identificar sitios donde la vegetación presentó cambios a lo largo del tiempo tales como zonas donde se pierde o se recupera vegetación, áreas con marcada estacionalidad, vegetación sometida a sequias, etc. Los valores altos de vegetación indicaron presencia vegetal acompañada de una estructura espacial y temporal heterogénea.

Por consiguiente, el sub índice hídrico (HI) conjunta la ocurrencia de agua por medio de HRC Global Surface Water, la estacionalidad con la variable “seasonality”, la proximidad al agua, TWI y MNDWI. Los valores altos indicaron generalmente zonas donde convergen la cercanía al agua, presencia recurrente de agua, condiciones topográficas favorables a acumulación y variabilidad de humedad.

Por último, el sub índice climatológico (CLI) identificó las anomalías de temperatura superficial, la variabilidad temporal de LST. Representando que en las zonas con valores altos convergen anomalías térmicas, alta variabilidad climática, zonas de precipitación irregular o de menor precipitación media. Es decir, espacialmente representa una condición relativamente más asociada con estrés térmico. Estos datos fueron de gran importancia no solo para identificar la

heterogeneidad, sino también para obtener las zonas de la cuenca Pesquería donde se identificaron las mayores convergencias del casco urbano.

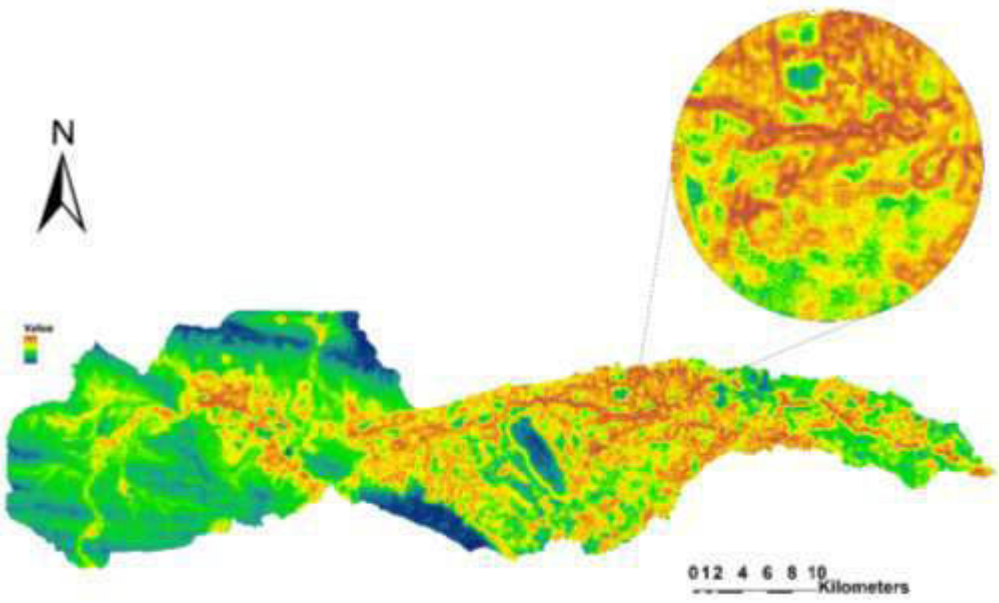


Figura 29. ICU cuenca Pesquería

Fuente: Elaboración propia

En la Figura 29 se muestra el ICU resultante para la cuenca del rio Pesquería, el cual evidencia una distribución espacial con diferencias claras a lo largo del territorio. En el sector occidental por la zona de García predominan valores bajos y medios, representados principalmente por tonalidades azul y verde, con una distribución más continua a excepción de la zona urbana. Conforme se avanza hacia la parte central, los valores aumentan y la configuración espacial se vuelve más irregular. En esta zona se concentran numerosos sectores con valores medios y altos, lo que marca un contraste evidente con la porción occidental de

la cuenca. Esta sección central corresponde a los municipios de Escobedo, Apodaca y parte de Monterrey del AMM. Hacia el oriente, aunque la extensión territorial es menor, también se mantienen valores relativamente altos en distintos tramos, pero con zonas muy bajas en relación a la alta vegetación que se encuentra por el municipio de Pesquería, Nuevo león.

La parte central de la cuenca presenta la mayor variabilidad espacial del ICU. En esta zona se observan cambios marcados entre valores altos y bajos en distancias cortas, formando parches, franjas y pequeños núcleos distribuidos de manera irregular. Este comportamiento se aprecia con mayor detalle en la ampliación de la Figura 30, donde se distinguen zonas con valores elevados rodeadas por áreas de menor intensidad. La distribución obtenida indica que las condiciones representadas por los componentes urbano, vegetación, hídrico y climático no se presentan de forma uniforme, sino que cambian de manera importante dentro de la propia cuenca. Los valores más altos del ICU se concentran principalmente en los sectores Centro-Norte, donde existe una mayor coincidencia entre las condiciones consideradas en los distintos subíndices. Por el contrario, el sector occidental presenta una menor convergencia relativa y una mayor continuidad de valores bajos.

Este resultado permitió identificar la zona de mayor Vulnerabilidad Ambiental y en la que este estudio se concentró para analizar un poco más a detalle que factores socioambientales contemplan esta convergencia y que podrían alter las condiciones naturales del río Pesquería y de sus zonas aledañas. Cabe resaltar que, un ICU alto significa que en ese punto las cuatro dimensiones superan la media simultáneamente. Es decir, un suelo en conversión sobre terreno hidrológicamente sensible, con la vegetación aún presente, pero fragmentándose.

En la Figura 30 se observa el ICU desde un punto de vista tridimensional juntos a las edificaciones en la zona central de la cuenca del río Pesquería.

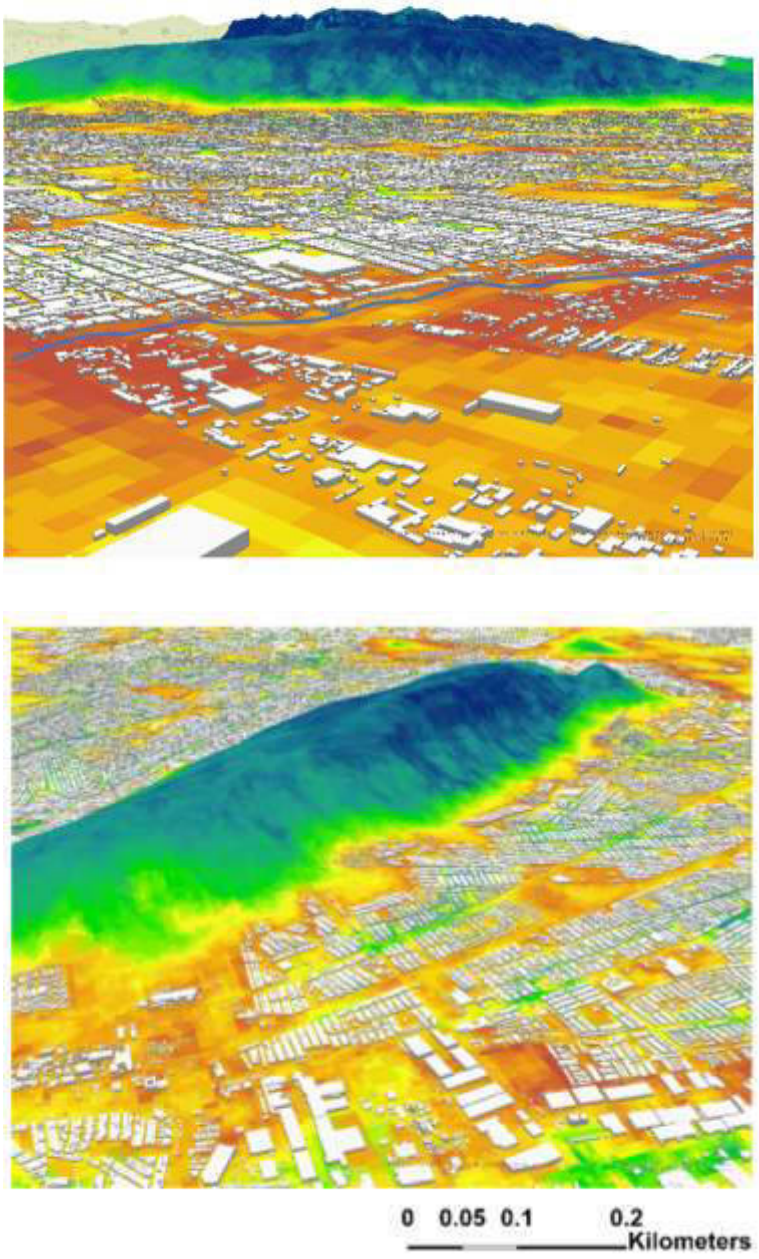


Figura 30. 3D ICU

Fuente: Elaboración propia

La representación tridimensional del ICU permite observar con mayor claridad la relación entre el relieve y la distribución espacial de los valores del índice. En

ambas perspectivas se aprecia que las zonas de mayor elevación y vegetación, asociadas principalmente con el cerro del Topo Chico, presentan valores bajos a medios del ICU, mientras que las superficies urbanizadas ubicadas en las partes bajas concentran con mayor frecuencia valores medios y altos. Este contraste muestra una diferenciación espacial marcada entre las áreas de relieve natural y las zonas donde la ocupación urbana tiene una presencia más intensa. Así como también se observa una franja roja que identifica al río Pesquería no solo por ser una zona de recurso hídrico sino también por las alteraciones ambientales que convergen en sus zonas aledañas, lo que la hace una zona de interés a ser investigada con mayor detalle.

5.1.2 IDAU

Tal como se mencionó en la metodología, el IDAU se obtiene como una suma ponderada de tres presiones principales: intensidad urbana, estrés térmico y déficit de vegetación. A nivel espectral, la radiación neta que llega a la superficie se reparte entre calor latente (evapotranspiración), calor sensible (calentamiento del aire) y almacenamiento en el sustrato. Cuando hay agua y vegetación, una parte se va como calor latente y la superficie se mantiene fresca. Cuando no la hay, ese mismo flujo se redistribuye hacía calor sensible y almacenamiento, y la temperatura superficial sube. El IDAU es, en el fondo, un proxy de hasta qué punto la superficie perdió su capacidad de disipar energía por evaporación.

En la Figura 31 se observa el IDAU calculado para la cuenca del río Pesquería.

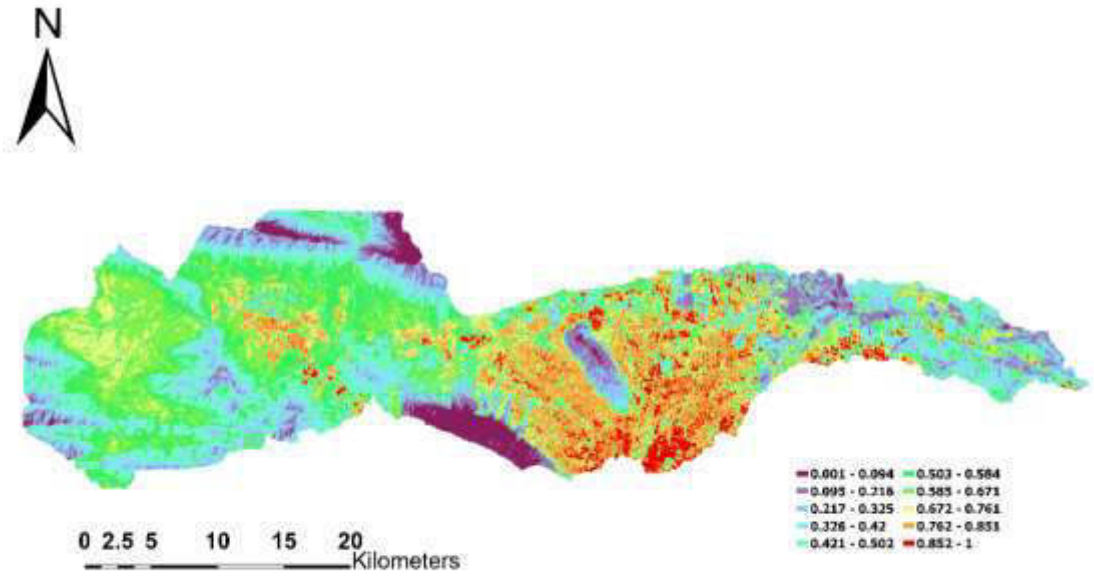


Figura 31. IDAU

Fuente: Elaboración propia

Con respecto a la distribución espacial que se observa en la Figura anterior, la zona de degradación alta claramente coincide en la zona urbanizada, pero es mayor en algunas zonas del centro de la cuenca con una tonalidad más rojiza a nivel visual. Los valores altos de IDAU indican un NDVI bajo (es decir, un déficit alto de vegetación). En otra interpretación un NDBI alto ya que las superficies construidas son secas y reflejan más la banda SWIR1 que en NIR, como el caso del agua que absorbe en SWIR, por lo tanto, cualquier superficie “deshidratada” como cemento, techos, suelo compactado eleva el índice IDAU tal como se muestra en la Figura. Además, la anomalía de LST es alta en estas mismas condiciones y claramente el GHSL es alto por su interpretación urbana. Desde un punto de vista ambiental, estos valores altos definen una impermeabilización, escorrentía rápida en lugar de infiltración, isla de calor superficial y ausencia de dosel.

Para los casos del IDAU bajo (por ejemplo, zonas de cerros) espectralmente un NDVI alto siendo un contraste rojo/NIR marcado por dosel activo con clorofila y mesófilo hidratado. NDBI bajo o negativo, porque el agua foliar absorbe en

SWIR. Además, anomalía térmica negativa, lo que constituye a una baja degradación ambiental y urbana.

5.2 Estructura general del río Pesquería

El río Pesquería perteneciente a la Región Hidrológica N°. 24 Bravo-Conchos se caracteriza por presentar escasez de agua (bajos flujos de agua) a lo largo de su cauce principal, específicamente en las zonas donde se encuentran los mayores crecimientos demográficos aledaños al río. De igual manera, los cambios de uso de suelo de la subcuenca del río Pesquería, en los últimos años, afectaron el patrón de escurrimiento natural y en la pérdida de servicios ecológicos del río urbano. La subcuenca presenta tramos secos, cauces invadidos por asentamientos humanos irregulares y actividades ilegales como porquerizas y basureros clandestinos; así como una clara evidencia de descarga ilegal de aguas residuales (Escobedo., et al 2026). A continuación, se observa la estructura general que presenta el río Pesquería.



Figura 32. Ortomosaico río Pesquería

Fuente: Elaboración propia

El río Pesquería, presenta una estructura compleja y dinámica a lo largo de cauce, desde su nacimiento en la Sierra de San Miguel hasta su desembocadura en el río San Juan. El espejo de agua del río muestra variaciones considerables tal como se muestra en la Figura 32. Presentando cambios dependiendo de la sección del río, algunas zonas de hasta 10m de ancho en promedio, pero en otras se observan anchos de lámina de agua entre 1-3m y en ocasiones hasta menos. Esto claramente puede aumentar o disminuir dependiendo de las diferentes épocas de lluvias o sequías que se presenten en el área de estudio.

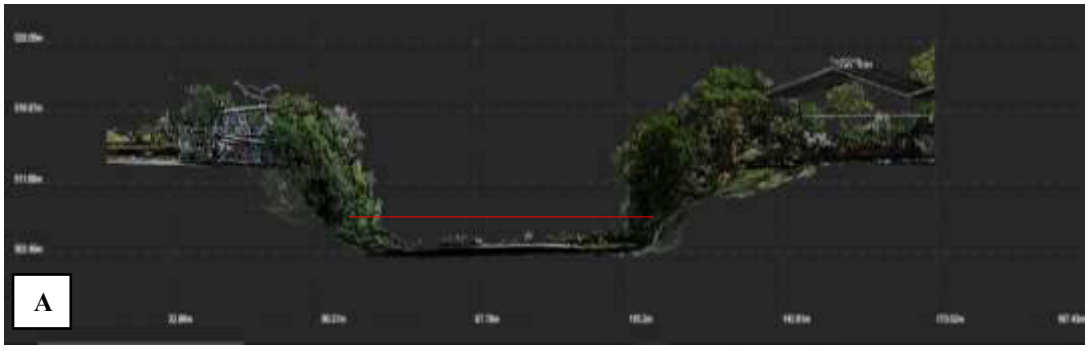
Por otra parte, el ancho del cauce tiende a ser más uniforme. En cuanto al cauce natural del río, este presenta un ancho promedio de 30-40 metros, aunque puede expandirse hasta 50-60 metros en zonas de planicie y reducirse a 20-25 metros en áreas de mayor pendiente. En las zonas modificadas por la urbanización, el cauce generalmente se mantiene entre 35-45 metros.

La geomorfología del río se caracteriza por un lecho principalmente rocoso en su curso alto, mientras que en su curso medio y bajo predominan los sedimentos aluviales, con presencia de meandros en zonas de planicie y formación de terrazas fluviales en varios tramos. El río ha sufrido modificaciones antropogénicas significativas, incluyendo erosión del cauce, puentes, cruces viales y en ocasiones es usado como deposito final de los residuos sólidos.

Hidrológicamente, el río mantiene un régimen “permanente” pero con fluctuaciones importantes en su caudal, influenciado por arroyos tributarios y aportaciones de aguas residuales tratadas. En términos ambientales, el río presenta una vegetación ribereña variable según la zona, albergando especies acuáticas nativas y sirviendo como área importante para aves migratorias. El estado de conservación del río varía considerablemente, con tramos naturales bien conservados en zonas rurales, pero mostrando degradación significativa en áreas urbanas, donde la contaminación representa un desafío importante que

requiere programas de restauración ecológica para su manejo y conservación efectiva.

Por otra parte, el análisis de la estructura general del río Pesquería se complementó con información topográfica detallada de la cuenca del río Pesquería. Gracias a los vuelos de drones equipados con sensores LiDAR, se logró generar una nube de puntos de alta resolución que permitió realizar perfiles topográficos a lo largo del cauce. Estos perfiles topográficos derivados de los datos LiDAR brindaron valiosa información sobre la profundidad y el ancho del río Pesquería en diferentes secciones. Este nivel de detalle geoespacial observado en la Figura 33 define significativamente la caracterización de las dinámicas hidrológicas de la cuenca, complementando los resultados obtenidos a partir de la clasificación de uso y cobertura del suelo.



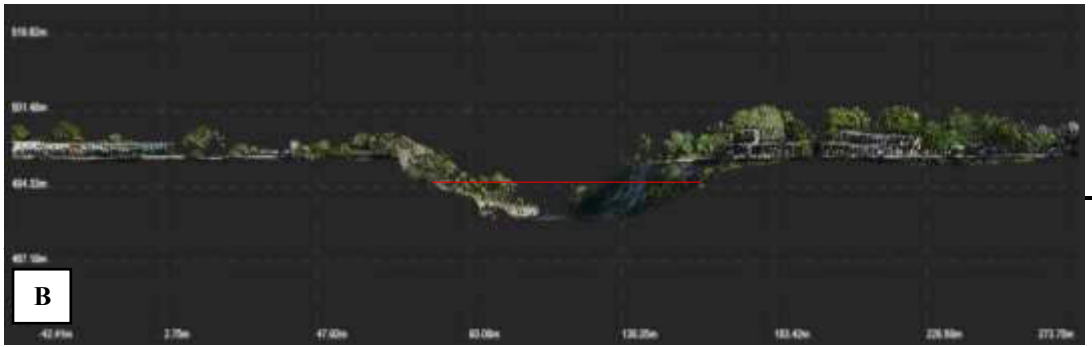


Figura 33. Perfil del cauce (A) Perfil topográfico 1y (B) Perfil topográfico 2

Fuente: Elaboración propia

Los perfiles topográficos transversales del río Pesquería, representados en las imágenes A y B de la Figura 33, revelan características morfológicas significativas de este sistema fluvial. El perfil A se caracteriza por una extensión horizontal de aproximadamente 170 metros, con elevaciones que oscilan entre los 503.4 y 528.1 m.s.n.m., exhibiendo una distintiva forma de "V" alargada que sugiere un proceso erosivo maduro. En este perfil, además de las condiciones del cauce también se logran observar asentamientos urbanos irregulares muy cerca del río Pesquería lo cual implica una problemática urbana que puede estar afectando directa o indirectamente las condiciones naturales del cuerpo receptor de agua. El cauce activo alcanza un ancho aproximado de 50 metros en su sección central, con su punto más profundo situado a una elevación de 503.4 m.s.n.m., mientras que sus márgenes presentan una considerable cobertura vegetal que contribuye a la estabilidad de las laderas.

Por otro lado, el perfil B muestra características diferentes, pero igualmente significativas, extendiéndose horizontalmente por 273.75 metros y abarcando un rango de elevación desde los 467.19 hasta los 518.62 m.s.n.m., con una morfología más amplia y asimétrica que su contraparte. En este segundo perfil, el cauce principal alcanza aproximadamente 61.08 metros de ancho, presentando terrazas fluviales bien definidas en ambos márgenes, las cuales evidencian diferentes episodios de actividad fluvial a lo largo del tiempo. La vegetación ribereña se mantiene como un elemento constante en ambos perfiles, desempeñando un papel crucial en la estabilización de las márgenes y la preservación de la estructura del cauce. Estos perfiles topográficos ilustran claramente la variabilidad morfológica del río Pesquería a lo largo de su curso, evidenciando procesos activos de erosión y sedimentación que han modelado el

paisaje fluvial actual. Donde cada sección presenta características únicas que contribuyen a la comprensión de la dinámica del río.

Por otra parte, los ortomosaicos 3D (Figura 34) presentaron varios factores ambientales críticos a lo largo del río Pesquería, incluyendo asentamientos irregulares que invaden zonas ribereñas, mala gestión de residuos sólidos y un flujo de corriente visiblemente reducido. En la siguiente Figura se observa a detalle el modelo 3D y el DEM utilizado.

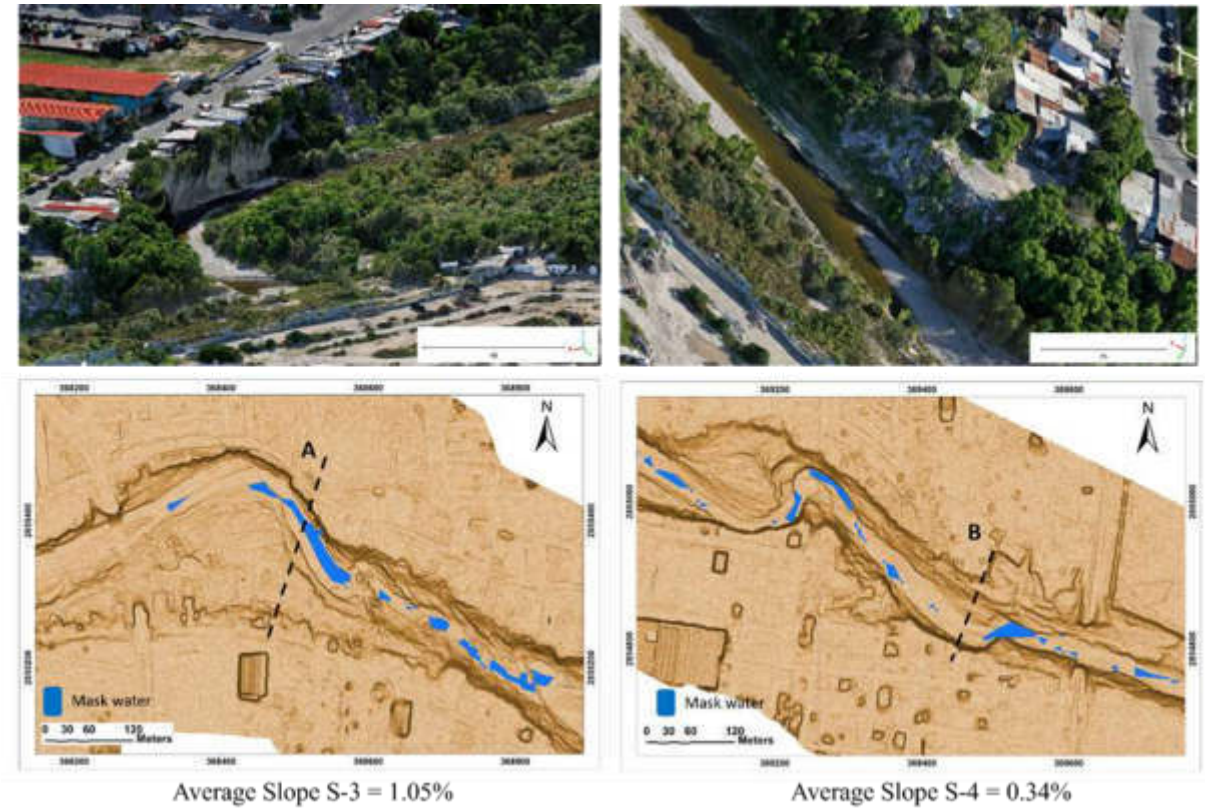


Figura 34. Modelos digitales de elevación (DEMs) y su correspondiente modelo 3D derivado de la fotogrametría basada en UAV. Los ejemplos anteriores proceden del río Pesquería y corresponden a los perfiles topográficos de las Secciones A y B.

Fuente: Elaboración propia

La Figura 34 presenta una vista combinada de los ortomosaicos y los modelos digitales de elevación, mostrando las ubicaciones de las secciones transversales A y B. Estas representaciones muestran un canal con una pendiente longitudinal suave (promedio del 0,66%), confinado en orillas empinadas y un desarrollo

limitado de llanura de inundación. En la Sección 3, la pendiente aumenta ligeramente hasta el 1,05%, generando un patrón de flujo más sinuoso. A pesar de no ser un río intrínsecamente estrecho, las imágenes confirman la casi ausencia de agua superficial durante gran parte del año.

5.3 Identificación de problemáticas en el río Pesquería

El río Pesquería, como uno de los sistemas fluviales más importantes del AMM, enfrenta una serie de desafíos críticos que han surgido principalmente como resultado de la desmedida urbanización, el desarrollo industrial y el crecimiento poblacional acelerado en su cuenca. La identificación de las problemáticas que afectan a este cuerpo de agua resulta fundamental para comprender su estado actual y desarrollar estrategias efectivas de gestión y recuperación. Este análisis cobra especial relevancia considerando que el río no solo funciona como un elemento natural esencial para el ecosistema regional, sino también como una fuente vital de recursos hídricos para diversas actividades humanas y ambientales (Castro et al., 2019; Rodríguez et al., 2024). La compleja interacción entre los procesos naturales y las actividades antropogénicas ha generado una serie de impactos significativos que requieren una evaluación detallada. Desde alteraciones en la calidad del agua hasta modificaciones en su morfología natural, el río Pesquería presenta un conjunto de problemáticas socioambientales que afectan tanto su función ecológica como su capacidad para proporcionar servicios naturales a las comunidades circundantes. La identificación y comprensión de estas problemáticas son el primer paso hacia el desarrollo de soluciones efectivas y de preservación de este importante recurso hídrico para las generaciones presentes y futuras.

Aunado a lo anterior, en este estudio y con ayuda del Departamento de Geomática, de la Facultad de Ingeniería Civil de la UANL se identificaron algunas de sus principales problemáticas con ayuda de vuelos de drones, visitas a campo y diversos análisis complementarios concernientes al río Pesquería y sus zonas

aledañas. A continuación, en la Figura 35, se observan algunas de sus principales problemáticas.

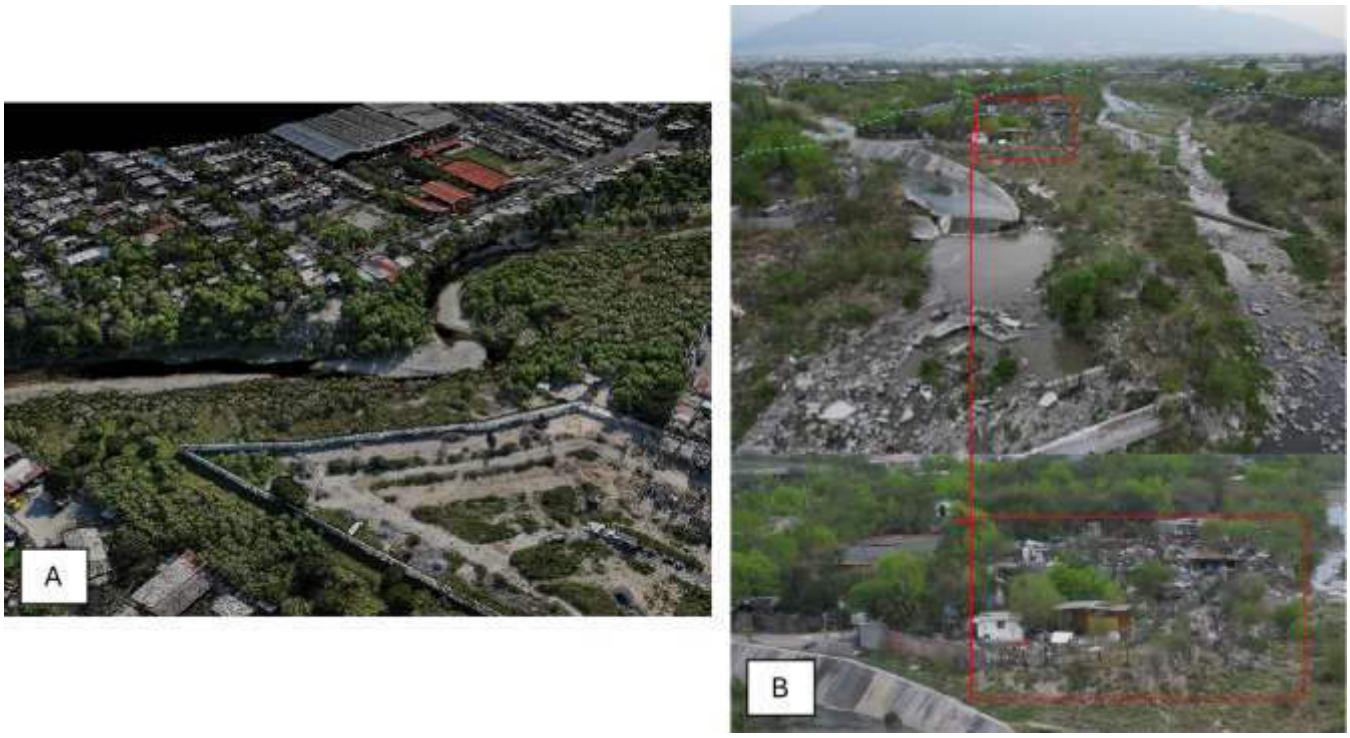


Figura 35. Problemáticas río Pesquería. (A) Nube de puntos LiDAR (B) Asentamientos irregulares cercanos a 1m del río Pesquería

Fuente: Yépez, 2024

Las Figura 35 evidencia algunas de las problemáticas que afectan al río Pesquería y su entorno. En la imagen A, mediante una representación tridimensional generada por tecnología LiDAR, se observa una problemática general tanto del cauce del río como de su zona ribereña, caracterizada por un flujo de agua notablemente reducido y un cauce estrecho. Esta condición es particularmente preocupante ya que refleja una degradación significativa del sistema fluvial, posiblemente causada por diversos factores como la sobreexplotación del recurso hídrico, alteraciones en el régimen de precipitaciones, o modificaciones antropogénicas del cauce. Esta situación amerita estudios detallados para comprender las causas subyacentes y

desarrollar estrategias de restauración que eviten la pérdida definitiva de este importante recurso hídrico.

En la imagen B se observa una problemática social igualmente significativa, que evidencia el establecimiento de asentamientos humanos irregulares en las márgenes del río. Esta situación representa un desafío social, cultural y ambiental. Por un lado, la población se expone a las a condiciones de vulnerabilidad ante posibles crecidas del río y deslizamientos de terreno, y por otro, genera impactos ambientales significativos en el ecosistema fluvial. La presencia de estas viviendas precarias frecuentemente se asocia con la ausencia de servicios básicos como drenaje y recolección de residuos, lo que deriva en la disposición directa de aguas residuales y desechos sólidos al cauce del río, demostrando los problemas de contaminación y degradación ambiental.

La relación de estas dos imágenes ilustra la compleja interrelación entre los problemas ambientales y sociales que afectan al río Pesquería. Esta situación aborda la necesidad de aplicar diversas herramientas geoespaciales que coadyuven al monitoreo tanto del río como sus zonas aledañas y a las diversas problemáticas que se mencionaron anteriormente. Con ayuda de los monitoreos se pueden desarrollar estrategias de solución efectivas a corto y mediano plazo.

Por otra parte, además de lo mencionado anteriormente, se han encontrado estructuras que alteran el flujo dinámico del río tal y como se muestra en la siguiente Figura 36 (A y B).



Figura 36. Represas antiguas del río Pesquería (A) Imagen satelital de Google Earth Pro de abril de 2009 y (B) Imagen satelital de Google Earth Pro de marzo de 2016

Fuente: Departamento de Geomática, IIC, UANL

La Figura anterior muestra una problemática significativa en el río Pesquería, donde se observan los impactos considerables en la dinámica natural del río. En la Figura 36A una antigua represa, que quedó en estado de abandono durante su construcción, representa actualmente un obstáculo artificial que interfiere directamente con el flujo natural del agua. Mientras que en la Figura 36B se observa otra represa cono una barrera transversal al cauce, ha creado una modificación no planificada en el comportamiento hidráulico del río, generando una zona de remanso aguas arriba y alterando significativamente las condiciones de flujo aguas abajo.

Esta problemática no solo afecta la continuidad del flujo de agua, sino que también puede impactar en la migración de especies acuáticas, el transporte de sedimentos, y la calidad del agua, representando así un desafío complejo que requiere atención y posibles medidas de mitigación o restauración para recuperar, en la medida de lo posible, las condiciones naturales del río.

CLASIFICACIÓN DE USO DE SUELO

6 Clasificación de uso de suelo

Como bien se sabe, el uso de suelo es un criterio socioeconómico y ambiental que hace referencia a la actividad dominante que un indicador biofísico describe los materiales que cubren el territorio a estudiar, por la necesidad de conocer la estructura y los cambios relacionados en el río Pesquería y con respecto a las zonas de mayor convergencia identificadas en el capítulo anterior, en este estudio se plantearon 4 diferentes clasificaciones de uso de suelo por medio de 4 diferentes algoritmos de ML como lo son K-means Clustering, RF, CART y SVM. Esta comparación de algoritmos de uso de suelo se hizo también para identificar la capacidad de precisión que cada uno de ellos tiene a la hora de identificar y/o mapear ríos urbanos de bajo caudal y con una alta convergencia socioambiental.

6.1 K-means Clustering

K-means es un algoritmo de clasificación no supervisada (clusterización) que agrupa objetos en k grupos basándose en sus características. El agrupamiento se realiza minimizando la suma de distancias entre cada objeto y el centroide de su grupo o cluster. Tal como se mencionó anteriormente este algoritmo es no supervisado y es por esto por lo que no necesita de datos de entrenamiento ni de validación para dar una respuesta.

La fortaleza de K-means radica en su simplicidad conceptual y eficiencia computacional (programación), lo que lo convierte en un método muy práctico para el agrupamiento de grandes conjuntos de datos. Sin embargo, tiene algunas limitaciones que se observan en las imágenes obtenidas de la clasificación por K-means clustering durante las temporadas 2016-2019 como se denota a continuación (Figura 37).

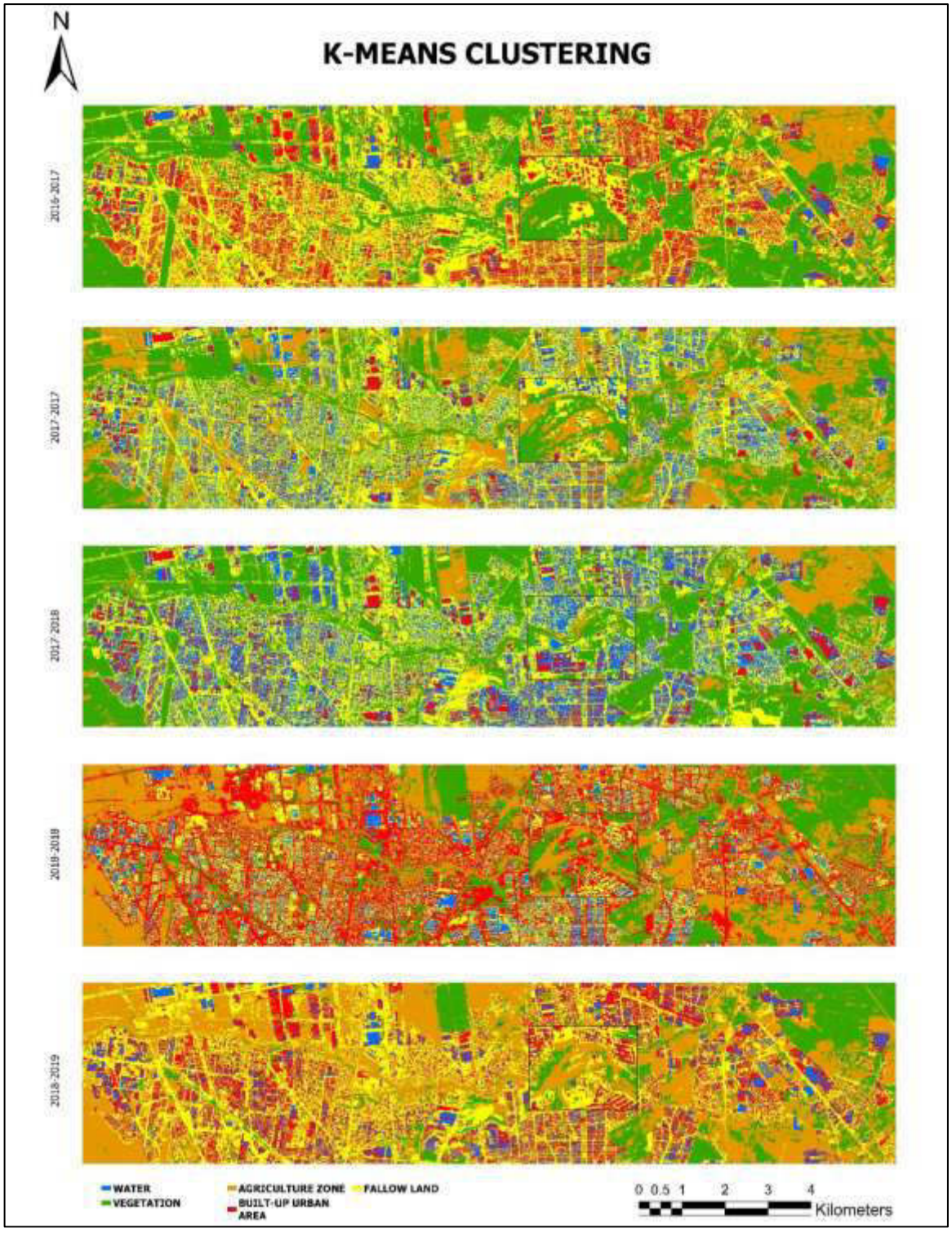


Figura 37. K-means clustering

Fuente: Elaboración propia

La Figura 37 muestra la clasificación realizada por medio del algoritmo K-means clustering en una de las secciones más urbanizadas del río Pesquería. Como se ha mencionado, al ser un algoritmo no supervisado, K-means no requiere datos de entrenamiento etiquetados previamente, sino que agrupa los píxeles en función de la similitud de sus características espectrales, sin información adicional sobre las clases presentes en la zona. Esto proporciona una vista a los subgrupos que se están clasificando sin la necesidad de agregar ninguna información relevante de la zona, lo cual ayuda a reconocer la segmentación de áreas similares y algunos patrones o errores que se puedan observar. Generalmente, la clasificación logra identificar zonas urbanas, la vegetación y algunas zonas de agricultura, por lo contrario, este algoritmo no pudo reconocer muy bien los contornos relacionados a la clase agua debido a lo angosto que es el río Pesquería. Aquí radica la complejidad e importancia de estudiar este tipo de ríos con información satelital (donde la mejor resolución es de 5m por píxel).

Por consiguiente, la clasificación identifica algunas zonas urbanas construidas, agrupándolas en una misma clase distintiva en las temporadas 2016-2017 y 2018-2019 principalmente, para los otros casos obtuvo muchos errores con contornos de clase agua. A su vez el algoritmo de K-means clustering realizado con imágenes satelitales NICFI presentó una confusión de contornos entre las clases de agricultura y vegetación a excepción de la temporada 2017-2018 donde se observa una mejor representación de estas dos clases.

No obstante, este algoritmo no pudo reconocer con precisión los contornos relacionados con la clase de agua, confundiendo algunas porciones del río con áreas urbanas o de vegetación. Esto puede deberse a que las firmas espectrales del agua y de algunas superficies artificiales o suelos desnudos son similares, lo que dificulta su separación automática. Además, una razón muy importante tiene que ver con que esta investigación está estudiando un río muy angosto, de poco flujo y en ocasiones intermitente, por lo que esto hace que un algoritmo no

supervisado no identifique con facilidad los contornos de agua. Sin embargo, el clustering no supervisado con K-means brinda una perspectiva inicial útil que puede complementarse posteriormente con información adicional o métodos supervisados para refinar los resultados.

6.2 SVM

Como bien se sabe en el ámbito del aprendizaje automático, SVM es un algoritmo de relevancia para la clasificación, particularmente notable en este estudio para la caracterización del uso del suelo. A continuación, se ilustra en la Figura 38 la respuesta del modelo SVM para las 5 temporadas analizadas en esta investigación.

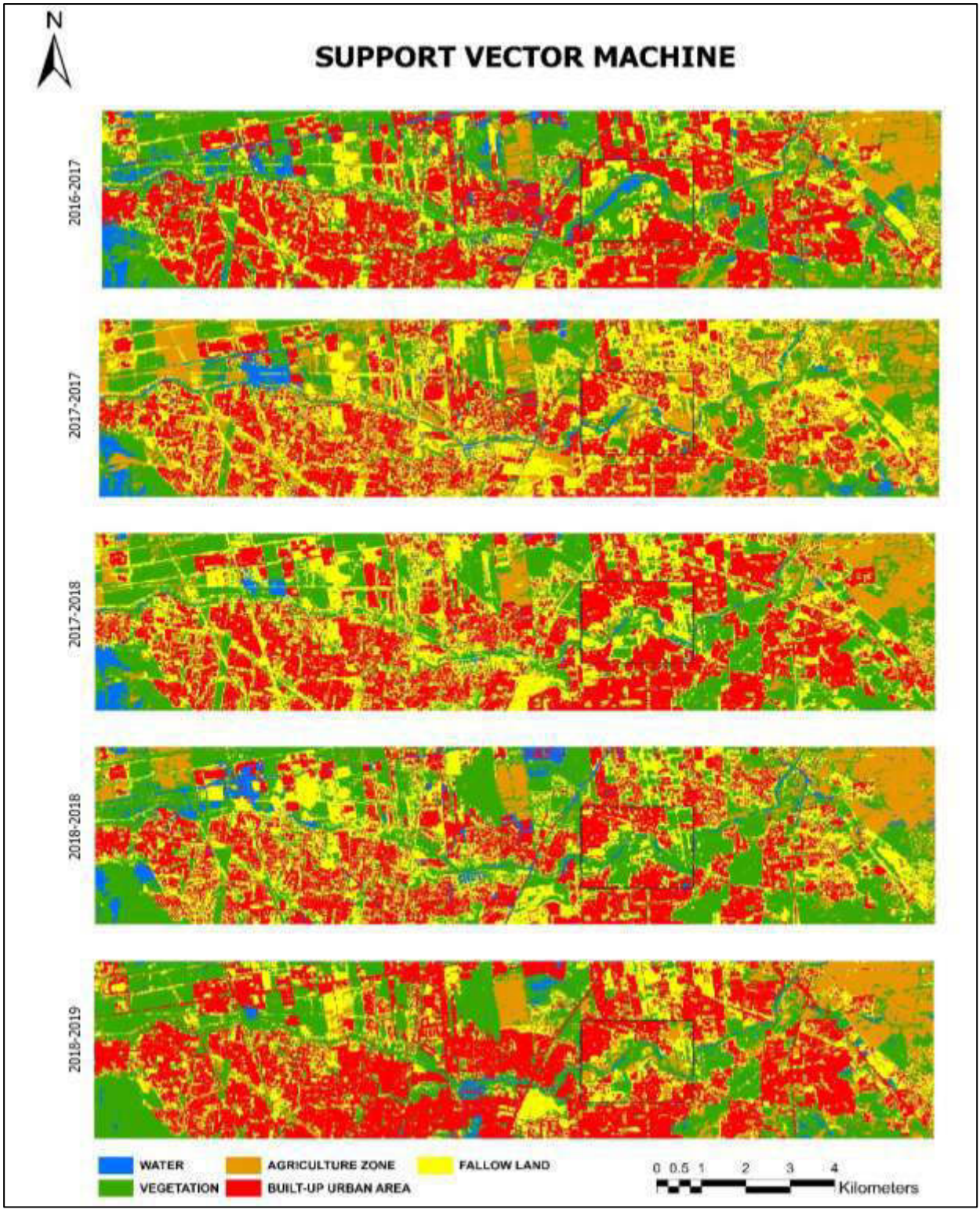


Figura 38. SVM

Fuente: Elaboración propia

La Figura 38 representa el modelo de respuesta de SVM aplicado a las temporadas 2016-2019 estudiadas. A diferencia del algoritmo no supervisado K-means, SVM es un método de clasificación supervisada, lo que implica que se le proporcionaron datos de entrenamiento y validación con imágenes satelitales NICFI etiquetadas para cada temporada. A simple vista se puede observar una mejor relación entre las diferentes clases a diferencia de los evidenciados en la Figura 37. Los contornos de crecimiento urbano presentan una buena respuesta para cada una de las temporadas en la clasificación de SVM; Por su parte, las clases de vegetación y agricultura también muestran contornos representativos de la zona de estudio. Sin embargo, se observan algunos errores relacionados con los contornos de la clase agua, los cuales se asemejan en algunas áreas a los de vegetación. Esto puede deberse a que, en ciertas zonas de vegetación, la imagen satelital presentaba muchas sombras, lo que provocó que las características espectrales de esta clase fueran similares a las del agua en esta región. Aun así, se logra observar una respuesta clara del río Pesquería, lo cual representa un desafío al utilizar imágenes satelitales de acceso gratuito y buena resolución en la actualidad. Las imágenes de respuesta lograron obtener una precisión del 94% en la temporada 18_19, representativa para las zonas estudiadas.

6.3 CART

El modelo CART, como bien se mencionó en la metodología, es un algoritmo de aprendizaje supervisado que construye un árbol de decisión a partir de un conjunto de datos de entrenamiento. Este algoritmo analiza las características de los datos y crea reglas de decisión jerárquicas en forma de árbol, lo que permite realizar para este caso clasificaciones de uso y cobertura de suelo. En la siguiente Figura 39 se observa la respuesta obtenida del algoritmo CART para las temporadas 2016-2019 en el área de estudio.

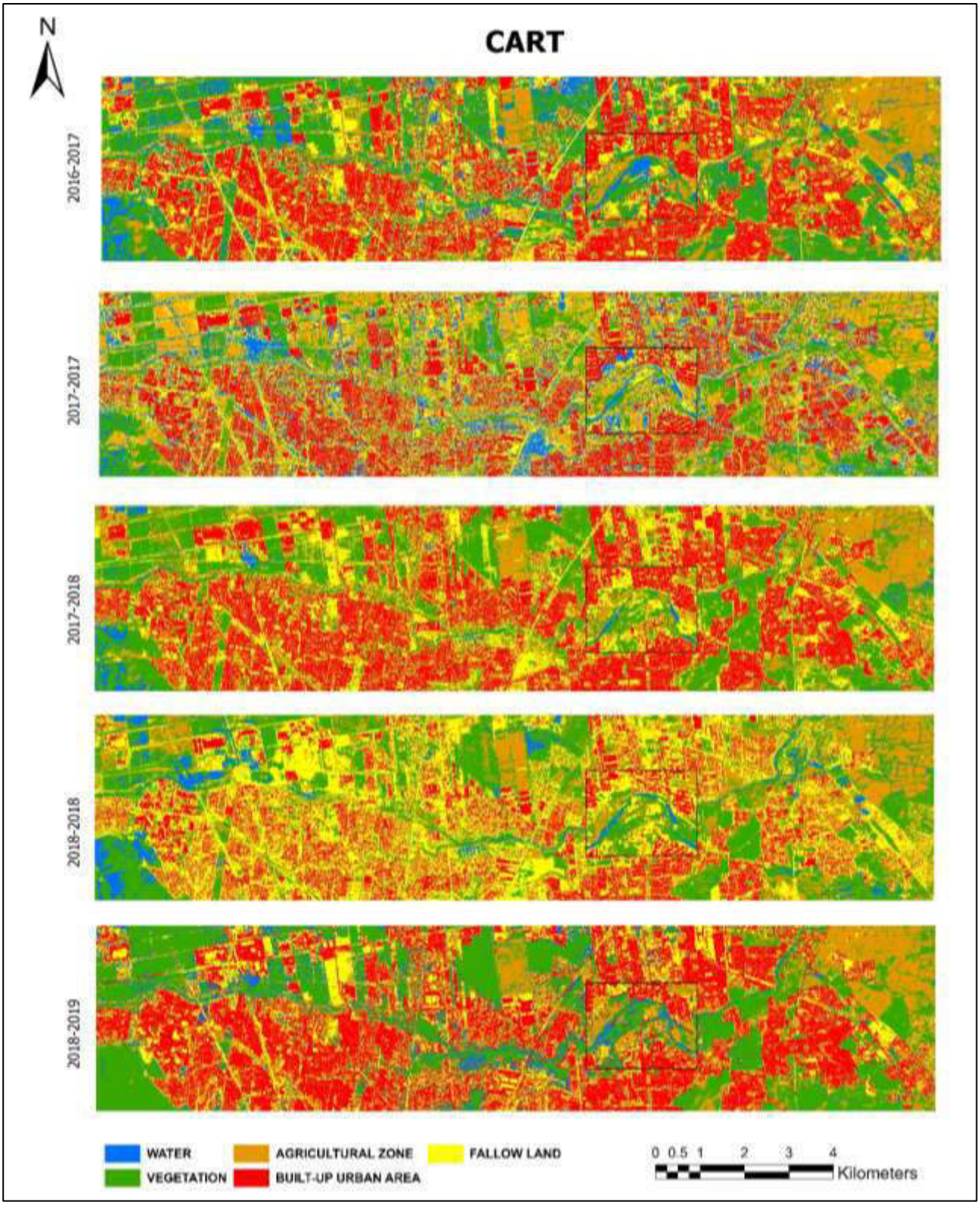


Figura 39. CART

Fuente: Elaboración propia

Como bien se ha mencionado a lo largo de este escrito GEE fue una herramienta eficaz para entrenar, validar y obtener las imágenes de respuesta de cada una de las temporadas. Para el caso del algoritmo CART, al igual que el modelo de SVM, se presentaron buenas respuestas para predecir las zonas urbanas del Área Metropolitana de Monterrey desde el año 2016-2019. Igualmente, la vegetación presentó buena respuesta, aunque la temporada de 2018-2019 es la que menores errores obtuvo a la hora de predecir la vegetación en el área de estudio.

Además, cabe mencionar que la temporada 2017-2017 presentó más errores para predecir agua en comparación con el resto de las temporadas, mientras que la temporada 2018-2019 obtuvo una mejor respuesta para la clase 1 (agua).

Luego de correr el algoritmo de clasificación en GEE. Los resultados destacaron una precisión general del 94% en la temporada 17_17, siendo igual a la mayor precisión que presentó SVM, pero entre todas las temporadas CART presentó menor consistencia entre sus clases.

6.4 RF

En este estudio, además de los algoritmos de aprendizaje automático presentados anteriormente, se evaluó otro algoritmo clave como lo es RF. Este algoritmo se basa en el uso de múltiples árboles de decisión para realizar predicciones precisas y robustas. RF fue implementado con el objetivo de obtener una representación precisa de las clases de esta investigación. La siguiente Figura 40 muestra los resultados obtenidos mediante la aplicación de este modelo en las diferentes temporadas contempladas.

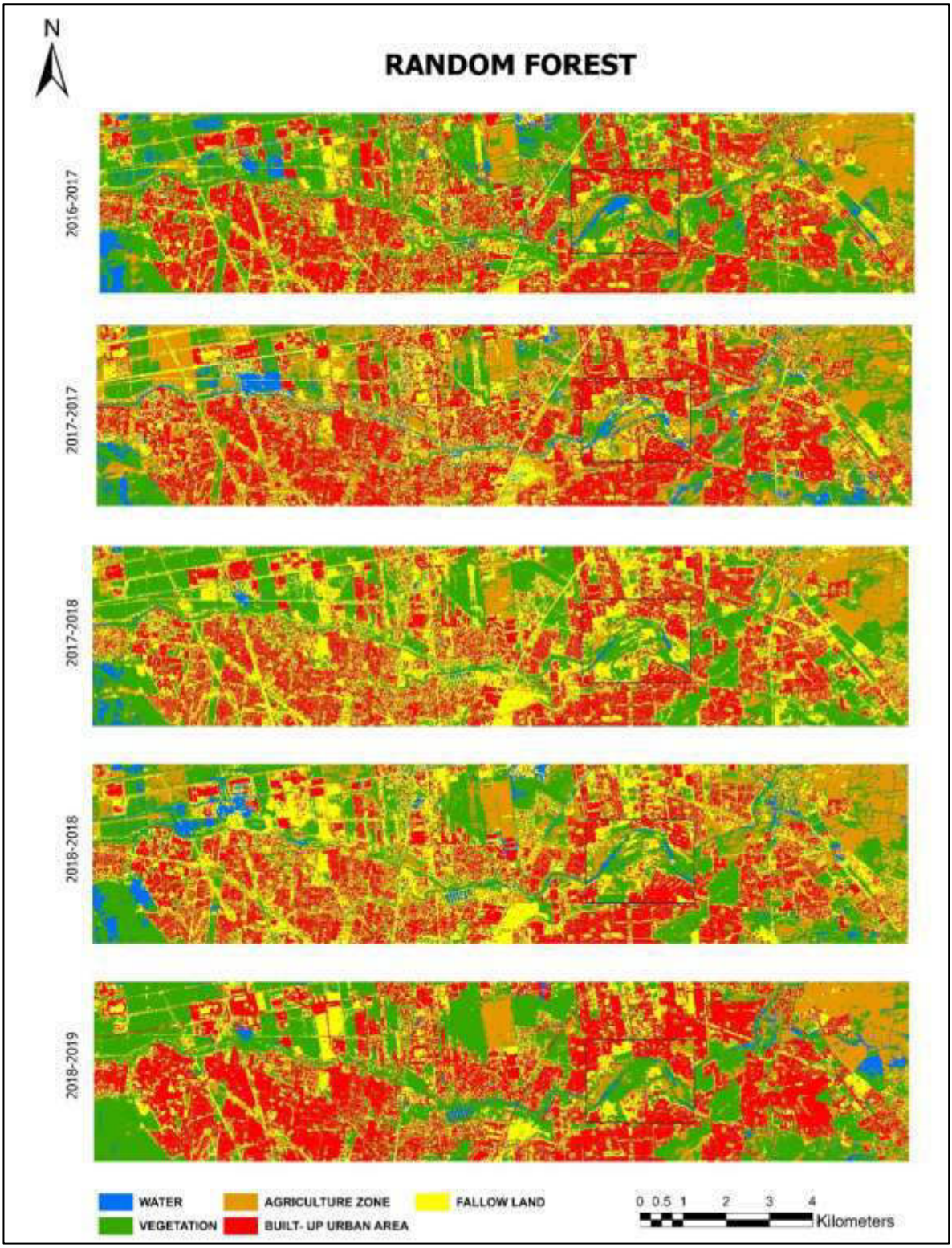


Figura 40. Random Forest

Fuente: Elaboración propia.

En la Figura 40 se observan las imágenes resultantes de la aplicación del algoritmo RF para las 5 clases de uso de suelo en las diferentes temporadas establecidas. Se puede analizar que este algoritmo presentó mejores resultados a nivel visual en comparación con las técnicas de SVM, CART y K-means clustering. Específicamente, se obtuvo una precisión general del 99%, lo cual indica que este proceso es, hasta el momento, el más apropiado para representar las diferentes clases de uso de suelo en la zona urbana del río Pesquería y de delimitar ríos de poco flujo en zonas semiáridas.

El algoritmo RF logró diferenciar eficazmente las diversas clases de uso de suelo implementadas, con excepción de algunos errores menores entre las clases de vegetación y agua, relacionados con la similitud de firmas espectrales entre los píxeles de agua y vegetación densa y también, por la cercanía entre el río angosto y los corredores biológicos en un píxel de una imagen satelital. De esta manera, el procedimiento de este modelo definió la continuidad del río Pesquería en sí con una precisión superior a los otros algoritmos evaluados previamente.

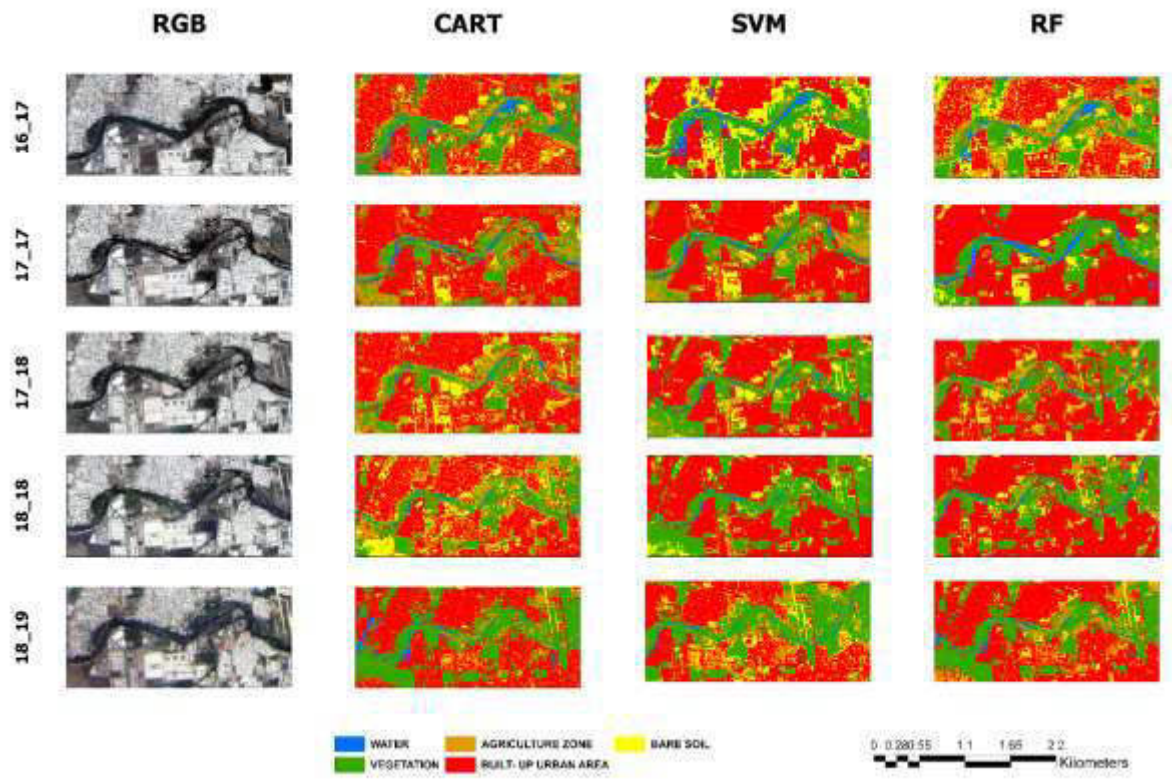


Figura 41. Respuesta algoritmo RF

Fuente: Elaboración propia

6.5 Análisis de áreas

Luego de obtener cada uno de los algoritmos de ML, se procedió a estudiar cada una de sus áreas por clase y por temporada. La Figura 42 ilustra la comparación de las áreas entre las distintas temporadas, representando cada categoría de uso del suelo y según la temporada, así como los cambios que experimentó cada clase en conjunto a lo largo de las cinco temporadas evaluadas en este proyecto.

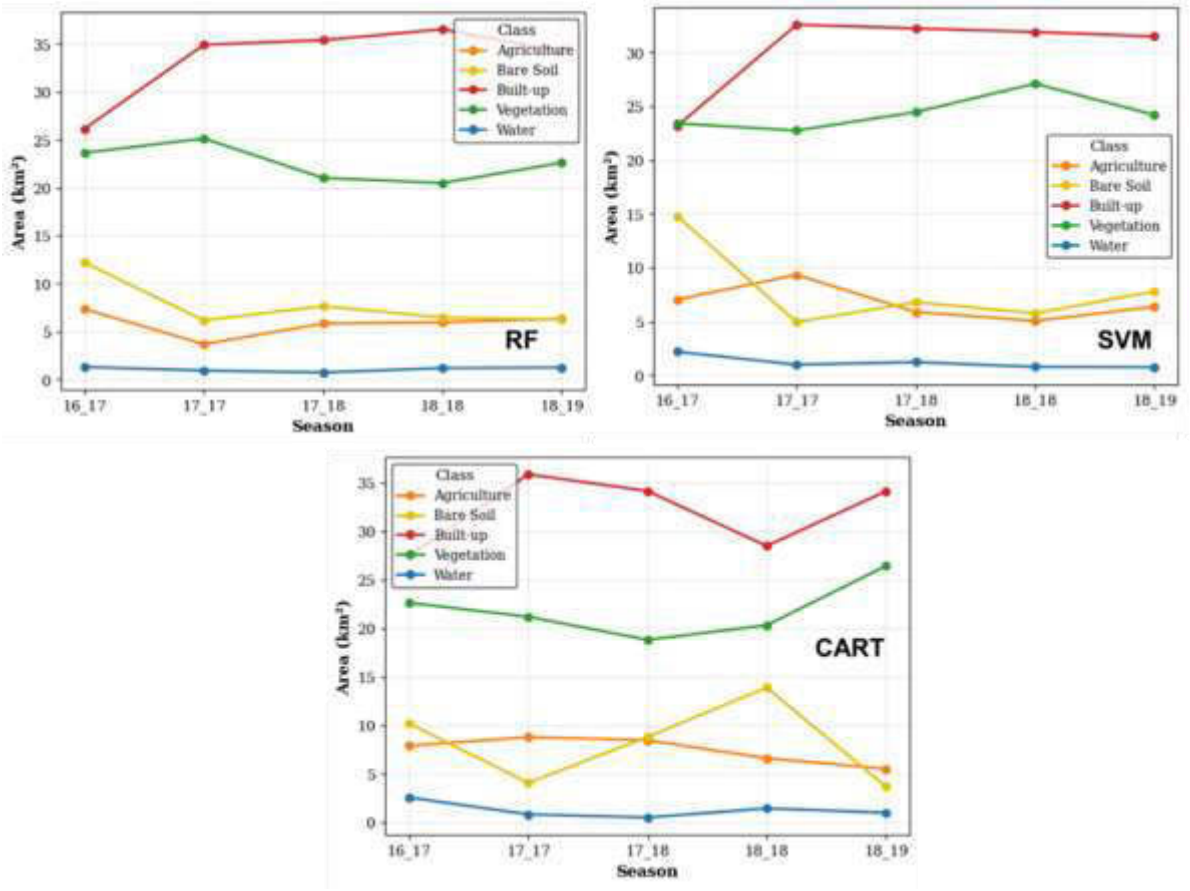


Figura 42. LULC variación de áreas en diferentes temporadas

Fuente: Elaboración propia

RF describe un incremento continuo de la superficie urbana el cual coincide con los análisis urbanísticos y de convergencia realizados en esta misma investigación. Su área urbana presenta un incremento que pasa de 26.2 km² en 16_17 a 35.0 km² en 17_17 y continúa ascendiendo de manera gradual hasta 36.8 km² al final de la serie. Este aumento de aproximadamente 10.6 km² ocurre simultáneamente con la contracción de la clase suelo, que cae de 12.2 a 6.1 km² en las dos primeras temporadas y se estabiliza en cerca de 6.3 km², y con la disminución de agricultura de 7.3 a 3.7 km² en el mismo intervalo, con una recuperación parcial posterior hasta 6.3 km². La vegetación presenta un comportamiento no monotónico: alcanza su máximo de 25.2 km² en 17_17, desciende a 20.5 km² en 18_18 y repunta nuevamente a 22.6 km² en la última temporada. El agua permanece entre 0.7 y 1.4 km², sin tendencia definida.

SVM por su parte, muestra un análisis similar a RF, pero con algunas atenuaciones considerables. La clase de urbano tiene una ascendencia similar de 23.4 km² a 32.6 km² entre 16_17 y 17_17 respectivamente, y a partir de ahí decrece levemente hasta 31.5 km², de modo que la superficie construida se estabiliza en lugar de seguir creciendo. La diferencia entre SVM y RF en la estimación final de esta clase es de aproximadamente 5.3 km², cifra que puede tomarse como una banda de incertidumbre operativa para la variable de mayor interés del estudio. La clase suelo evidencia uno de los valores más bajos del análisis de suelo en los 3 modelos, 14.8 km² en 16_17, y cae a 5.0 km² en la temporada siguiente. La vegetación es más estable en SVM que en RF y alcanza 27.1 km² en 18_18, su valor más alto en los tres algoritmos. El agua vuelve a ser la clase de menor variación absoluta, con un rango de 0.7 a 2.2 km².

CART resultó siendo el algoritmo que presentó la mayor diferencia de área entre los 3 clasificadores. La clase urbana aumenta de manera similar en las primeras temporadas, pero desciende de 36.0 km² en 17_17 hasta un mínimo de 28.5 km² en 18_18 y vuelve a 34.2 km² en 18_19, una oscilación de 5.7 km² generando varias incertidumbres por sus cambios. Lo anterior no se observa en RF ni en SVM. El comportamiento de la clase suelo es todavía más errático, con un mínimo de 4.1 km² en 17_17, un pico de 13.9 km² en 18_18 y una caída a 3.8 km² en la temporada final. Ese pico coincide temporalmente con el mínimo de urbano, por lo que se infiere podría ser un intercambio directo de píxeles entre ambas clases más que un cambio real en la cobertura. La vegetación sigue una trayectoria descendente hasta 18.9 km² en 17_18 y luego asciende a 26.5 km², el valor más alto que registra esta clase en el conjunto de resultados.

De manera general, el análisis de clasificación de cobertura y uso de suelo durante las diferentes temporadas analizadas revela una tendencia hacia la urbanización descontrolada. Los datos muestran un claro aumento de dicha clase, que han ido ocupando terrenos que anteriormente eran suelo desnudo o áreas naturales. Esta expansión urbana se ha producido no solo en las zonas céntricas, sino también en las áreas aledañas al río, poniendo en riesgo estos

valiosos ecosistemas ribereños. En la Figura 43 se observa más a detalle esta idea, en donde se ilustra un análisis tipo radar de las clases resultantes del algoritmo de RF.

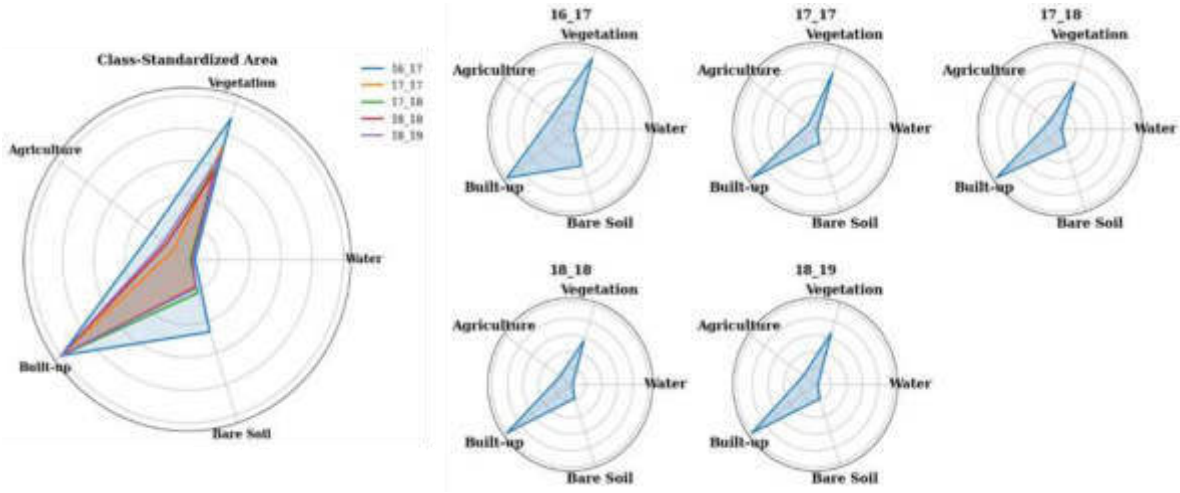


Figura 43. Diferentes estaciones y clases representadas con un radar estandarizado por clase basado en la superficie estimada del algoritmo RF

Fuente: Elaboración propia

Este análisis radar confirma y realza lo mencionado anteriormente ya que el radar se alarga hacia la clase urbano en las cinco temporadas. Agua y suelo permanecen próximos al centro en toda la serie temporal, con proyecciones radiales bajas y estables. La clase agricultura ocupa posiciones intermedias, salvo en 17_18, donde su vértice se extiende de forma notoria, aunque sigue siendo a baja escala. La clase vegetación al igual que las demás clases presenta un incremento, pero no es tan consistente como el de la clase de urbano durante todas las temporadas.

Por otra parte, se estudió a nivel especificó la clase de agua, por lo que en la Tabla 12 se muestran que las coberturas de agua alcanzaron un máximo de apenas 1.82% en la temporada con mayor presencia hídrica (2016-2017), y un mínimo de 1% en la temporada 2017-2018 en relación con el área de estudio.

Tabla 11. Áreas de cobertura y porcentajes de la clase agua para las 5 temporadas

Temporadas	Época	Área km² (sin sesgo filtrado)	Porcentaje (%)	Área km² (con sesgo filtrado)	Porcentaje (%)
2016-2017	De lluvia	1.29	1.82	0.12	0.17
2017-2017	De estiaje	0.92	1.30	0.4	0.57
2017-2018	De lluvia	0.72	1.01	0.18	0.25
2018-2018	De estiaje	1.17	1.66	0.3	0.43
2018-2019	De lluvia	1.23	1.75	0.21	0.3

Fuente: Elaboración propia

En la Tabla 12 las variaciones identifican como corresponde una mayor cantidad de agua en épocas de lluvia que en épocas de estiaje. Estas cifras evidencian la escasez de recursos hídricos en la región estudiada. Es importante destacar que las variaciones observadas en la cobertura de agua no solo están relacionadas con las fluctuaciones estacionales, sino que también están vinculadas a factores socioeconómicos y ambientales. Lo anterior representa un desafío para el estudio de este tipo de ríos, convirtiéndose en una temática área importante para investigar no solo en el estado de Nuevo León, México sino también a nivel mundial. Lo cual representa una ventana de oportunidad para el uso de imágenes satelitales de observación terrestre. El área sin sesgo se obtuvo luego de varias corridas del modelo hasta obtener el más optimo para la clase agua, además se eliminaron manualmente alguno pequeños polígonos que notoriamente no eran clase agua.

Por otra parte, la Figura 44-a muestra un análisis complementario de la correlación cruzada entre las clases LULC, a traves de la matriz que identifica las interacciones estructurales entre las clases del área de estudio. Se muestra una fuerte correlación negativa en el tipo de zonas urbanas y el suelo desnudo (r = −0,94), lo que indica una sustitución directa del sustrato sin vegetación a las

superficies urbanas. También es evidente una relación importante entre clase urbana y agricultura (r = −0,60), y entre clase urbana y vegetación (r = −0,43), reflejando cómo la urbanización ejerce una presión creciente y constante sobre áreas vegetadas y que aún no son productivas.

En contraste, la Figura 44-a muestra correlaciones positivas entre las clases de suelo y agricultura (r = 0,67) y entre las clases agrícola y de agua (r = 0,58). Esto sugiere que existe una concurrencia espacial entre estas combinaciones de clases, que muchas persisten antes de la conversión de la clase urbana. Ambos patrones confirman que el crecimiento urbano en las zonas periurbanas es el principal motor de la reorganización observada en el corredor del río Pesquería a lo largo del periodo de estudio, es decir, este fenómeno reduce progresivamente la heterogeneidad de las clases, comprometiendo el equilibrio ecológico del entorno natural del área de investigación.

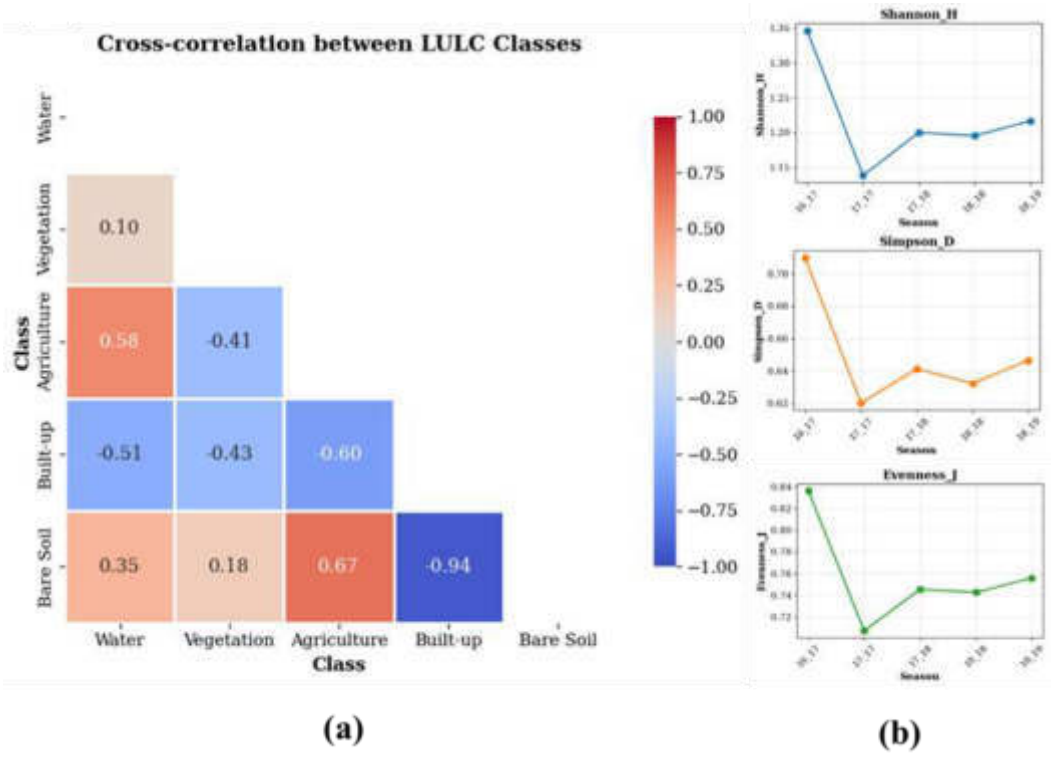


Figura 44. (a) Clases LULC de correlación cruzada. (b) Diversidad ecológica (Shannon, Simpson, Evennes).

Fuente: Elaboración propia

En la Figura 44b, se utilizaron los índices de diversidad ecológica aplicados en este caso a LULC para entender la diversidad y heterogeneidad del paisaje entre las clases estudiadas. El índice de Shannon, refleja, al igual que los otros índices, una disminución en la temporada 17-17 asociada a la distribución desigual de clases, debido al aumento significativo de una clase dominante. Shannon suele ser sensible a las clases minoritarias, lo que lo hace eficiente al estudiar cuerpos de agua intermitentes como el río Pesquería.

Según su análisis, Simpson indica que hay una mayor probabilidad de que dos píxeles seleccionados al azar pertenezcan a la misma clase, dando mayor peso a lo anterior. En este caso, la clase dominante es la urbana, por lo que hay una expansión de las masas (aunque en este estudio solo se evalúa la expansión horizontal). El índice de uniformidad muestra un patrón similar en las 5 estaciones, indicando que el cambio presentado en 17_17 se debe a un desequilibrio en las proporciones de cada clase, haciendo que el paisaje esté menos equilibrado en esa estación.

6.6 Gemelo digital 3D

La incorporación de una visualización tridimensional permite ampliar la interpretación de la clasificación de uso y cobertura del suelo obtenida mediante RF en esta investigación, al pasar de una representación bidimensional de las coberturas a una caracterización espacial integrada del corredor fluvial. En el caso del río Pesquería, esta diferencia es particularmente importante debido a la reducida anchura del flujo de agua, la presencia de superficies aledañas al río y la discontinuidad de la lámina de agua. La información tridimensional permitió representar simultáneamente la convergencia entre el relieve, el cauce, las edificaciones junto con el modelo de inundación con un periodo de retorno TR500 en el área de estudio. En la Figura 45 se observa el modelo 3D de una sección del río Pesquería y sus zonas aledañas



Figura 45. Gemelo digital 3D del río Pesquería

Fuente: Elaboración propia

Bajo esta visualización, la integración de la clasificación RF con la información tridimensional se dio pie a la representación de un gemelo digital del entorno fluvial urbano. La geometría tridimensional permitió identificar zonas de escasa vegetación relacionada con asentamientos humanos irregulares en zonas de riesgo por inundación, representando una mejor visualización hacia la toma de decisiones y futura construcción de políticas que disminuyan los riesgos en las terrazas de los ríos urbanos del país.

Desde la perspectiva ambiental, el resultado evidencia que el río Pesquería se encuentra inmerso en una matriz urbana altamente intervenida, con sectores donde la vegetación asociada al cauce se encuentra fragmentada y donde los asentamientos humanos irregulares alcanzan prácticamente el borde del río. La identificación de estas condiciones, permite localizar espacialmente zonas donde convergen pérdida de cobertura vegetal, presión urbana, reducción del espacio ribereño y exposición de la población y su infraestructura a inundaciones.

Lo anterior proporciona una base más adecuada para evaluar la problemática ambiental del río, ya que el gemelo digital puede integrar posteriormente información hidrológica, calidad del agua, riesgo de inundación y cambios de cobertura. En este sentido, la aportación principal no se limita a generar una visualización tridimensional, sino a estructurar un modelo espacial capaz de relacionar las características físicas del cauce con las condiciones ambientales y urbanas que determinan su vulnerabilidad. Se espera que esta también sea la base para futuros proyectos que apliquen estas tecnologías geoespaciales en zonas tan fragmentadas como lo es el río Pesquería.

6.7 Evaluación de la precisión

Se evaluaron los modelos de aprendizaje automático CART, SVM y RF, para todas las estaciones, según se indica en la Tabla 13, utilizando el coeficiente de Kappa y la precisión general. Los resultados indican un rendimiento consistente en todas las temporadas, evaluado individualmente.

CART registró una precisión mínima del 85% con un coeficiente Kappa de 0,81 y una precisión máxima del 94% con un Kappa de 0,924 durante las temporadas 17-18. En contraste, el modelo SVM alcanzó precisiones mínimas del 86% con un Kappa de 0,823 y exactitudes máximas del 94% con un coeficiente de Kappa

de 0,923. Mientras tanto, el algoritmo RF superó a los otros modelos a nivel general comparando durante las cinco estaciones evaluadas, alcanzando precisiones de clasificación del 99% y un coeficiente Kappa de 0,987. La menor precisión reportada para RF fue del 93%, con un Kappa de 0,911 durante la temporada 18-18. RF fue el algoritmo más consistente para todas las estaciones. En general, RF mantuvo consistentemente una alta precisión en la clasificación, superando el 90% en las cinco temporadas.

Tabla 12. Precisión general (OA) y coeficiente de Kappa (Kc) de los modelos RF, SVM y CART para la clasificación LULC estacional (2016–2019).

	Seasons	OA	Kc
RF	16_17	94	92.3
	17_17	98	97.4
	17_18	96	94.9
	18_18	93	91.1
	18_19	99	98.7
SVM	16_17	90	87.3
	17_17	86	82.3
	17_18	92	89.9
	18_18	92	89.8
	18_19	94	92.3
CART	16_17	87	83.6
	17_17	85	81
	17_18	94	92.4
	18_18	93	91.1
	18_19	91	88.5

Además, la Tabla 14 presenta las métricas de precisión del usuario, exactitud del productor y puntuación F1 calculadas para el modelo RF, seleccionadas para

registrar el mejor rendimiento en todas las temporadas analizadas. La evaluación de estas métricas para cada clase nos permite valorar con mayor detalle la robustez y fiabilidad del modelo. Los resultados obtenidos confirman el buen rendimiento de RF y respaldan la eficacia de la metodología propuesta para el mapeo de máscaras de agua en ríos estrechos.

Tabla 13. Precisión del usuario (UA), precisión del productor (PA) y puntuación F1 para la clasificación en Bosque Aleatorio (RF) por clase y temporada de cobertura del suelo (2016–2019).

	Agua			**Vegetación**			**Agricultura**			**Urbano**			**Suelo**		
Seasons	**PA**	**UA**	**F1**	**PA**	**UA**	**F1**	**PA**	**UA**	**F1**	**PA**	**UA**	**F1**	**PA**	**UA**	**F1**
16_17	0.93	1.00	0.97	0.91	0.96	0.94	1.00	0.95	0.98	1.00	0.89	0.94	0.81	0.93	0.87
17_17	1.00	1.00	1.00	0.96	1.00	0.98	1.00	1.00	1.00	1.00	0.96	0.98	0.94	0.94	0.94
17_18	0.93	0.93	0.93	0.95	0.91	0.93	0.95	1.00	0.97	1.00	0.96	0.98	0.94	1.00	0.97
18_18	0.94	0.88	0.91	0.95	0.91	0.93	0.94	1.00	0.97	0.93	0.93	0.93	0.88	0.94	0.91
18_19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	0.98	1.00	0.94	0.97

La clasificación de cobertura y uso del suelo mediante algoritmos de aprendizaje automático ha mostrado resultados consistentemente altos en este estudio en un ambiente semiárido, algunos investigadores han empleado estos mismo algoritmos bajo diferentes condiciones tales como Martínez de León et al. (2026) quienes compararon el rendimiento de SVM, RF, GBT y CART integrando datos espectrales de Sentinel-2A, índices topográficos e índices espectrales en un Área Natural Protegida de la Reserva de la Biosfera "El Cielo" (Tamaulipas, México). Sus hallazgos mostraron que SVM y GBT obtuvieron la mayor fiabilidad

global temática con 88% y 86% de precisión, respectivamente, destacando que el 46% de las comparaciones pareadas por clase presentaron diferencias estadísticamente significativas.

Aldiansyah y Saputra (2022) evaluaron CART, RF y SVM utilizando imágenes Landsat-8 y Sentinel-2 en la cuenca del río Wanggu (Wanggu Watershed, Indonesia). El estudio concluyó que el algoritmo RF superó a CART y SVM en ambos sensores. Con Landsat-8, las precisiones globales (OA) promedios fueron: RF (94.74%), CART (92.03%) y SVM (83.54%); mientras que con Sentinel-2 las precisiones alcanzadas fueron RF (96.15%), SVM (93.34%) y CART (93.14%).

Ouma et al. (2022) realizaron una evaluación multitemporal y multisensores (1984–2020) comparando RF, SVM, GTB y CART en el Área de Planificación de Gran Gaborone (GGPA, Botsuana). Al evaluar la precisión global promedio de las ocho épocas del estudio, RF (87.8%) y SVM (87.5%) registraron el desempeño más alto y estable, seguidos por GTB (86.4%) y CART (85.3%). En términos de precisión promedio por clase, RF logró un 95.9%, SVM un 95.8%, GTB un 95.6% y CART un 95.1%, con coeficientes Kappa promedio de 0.85 (RF y SVM), 0.84 (GTB) y 0.82 (CART)

Por su parte, Garzón Barrero et al. (2023) analizaron el desempeño de SVM frente a Random Trees (RT/RF) utilizando un modelo de segmentación basada en objetos (OBIA) e imágenes Landsat-9 OLI-2 en el Área Metropolitana de Barranquilla (Colombia). Los autores demostraron que, en conjuntos de datos desequilibrados (que replican las proporciones reales del paisaje), SVM superó consistentemente a RF, alcanzando una precisión general de 86% con el 100% de muestras y 84% con el 60% de muestras, en comparación con el 81% y 78% de RF en los mismos escenarios. Sin embargo, en muestras estrictamente equilibradas de 1250 píxeles, RF alcanzó un 68% de precisión frente a 64% de SVM.

A demás, Rahman et al. (2026) evaluaron RF, CART y SVM procesados en GEE con imágenes Landsat-8 para el distrito de Moulvibazar (Bangladesh). Los

resultados ubicaron a RF como el modelo superior con un 96.9% de precisión general y un índice Kappa de 0.96, seguido por CARTcon un 93.4% de precisión general (Kappa de 0.91) y SVM con un 81.9% de precisión (Kappa de 0.76). Jasim et al. (2025) investigaron la precisión de RF, SVM, ML y K-Nearest Neighbors (KNN) utilizando imágenes Sentinel-2A en la provincia de Karbala (Irak). El clasificador RF registró los valores más altos con un 93.1% de precisión general y un Coeficiente Kappa de 0.90, seguido de SVM (91.2% de OA, Kappa 0.88), ML (86.2% de OA, Kappa 0.81) y KNN (82.5% de OA, Kappa 0.76). Rawat et al. (2024) desarrollaron una comparación estadística entre RF, SVM para imágenes Landsat-5, Landsat-8 y Sentinel-2A en el distrito de Chennai (India). La precisión general a lo largo de diversos periodos osciló entre 80% y 89% para RF (con Kappa de 0.72 a 0.84), entre 79% y 87% y para SVM (con Kappa de 0.67 a 0.80). Esta investigación logró obtener una mayor precisión general que las investigaciones anteriormente mencionadas para el algoritmo RF.

WATER MASK

7 Water Mask

El presente análisis se basó en las extracciones de agua realizadas sobre el río Pesquería mediante clasificación LULC utilizando el algoritmo RF (algoritmo de mayor precisión en este trabajo de investigación), centrándose específicamente en la clase 1 (Agua) para las temporadas y estaciones 2016-2017, 2017-2017, 2017-2018, 2018-2018 y 2018-2019. A cada una de las imágenes clasificadas se le hizo un respectivo procesamiento de limpieza de errores y sesgos del modelo.

Esta extracción de agua realizada por medio de una clasificación LULC buscó presentar una mejor respuesta del espejo de agua a diferencia de los índices de agua usados convencionalmente. En la Figura 46 se observa la máscara de agua resultante para cada una de las temporadas estudiadas.

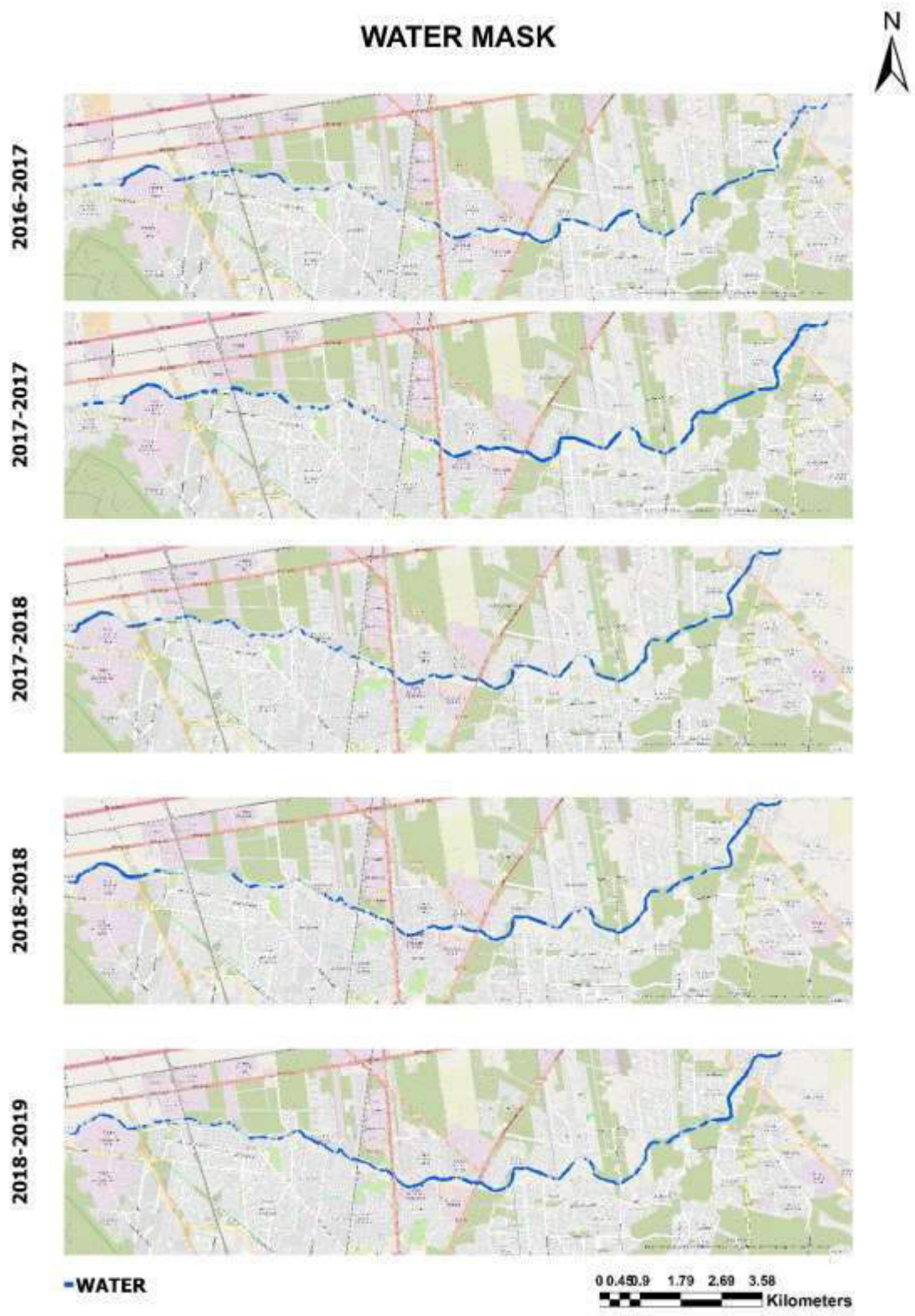


Figura 46. Water Mask 2016-2019

Fuente: Elaboración propia

De manera general, se observa una delimitación clara de la dinámica del espejo de agua en el área de estudio. Sin embargo, al analizar los periodos de 2016-2017, 2017-2017 y 2017-2018, se identifica un patrón espacial recurrente en el sector oeste caracterizado por la ausencia de agua superficial. Este comportamiento indica una vulnerabilidad crítica en dicha sección del cauce, lo que demanda atención prioritaria para su manejo y conservación. La Figura 47 expone estas variaciones a mayor detalle.

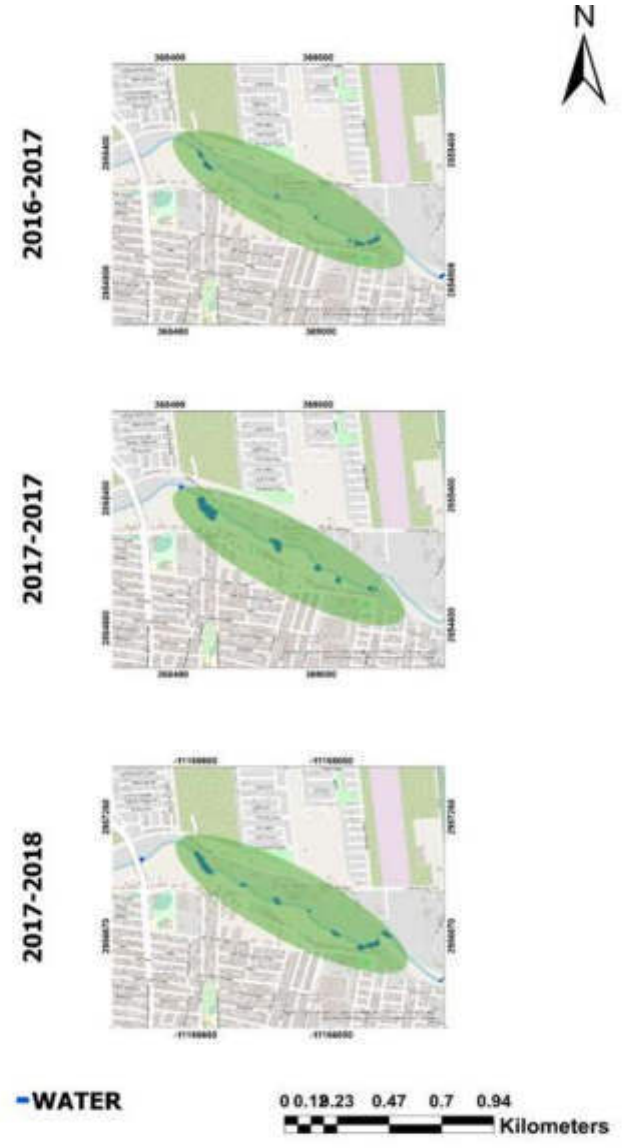


Figura 47. Escasez de agua Water mask

Fuente: Elaboración propia

La Figura anterior refleja una disminución en los contornos identificados como cuerpos de agua mediante la clasificación LULC. Este fenómeno puede atribuirse a dos factores principales.

El primer factor está relacionado con limitaciones en la capacidad del modelo de clasificación de uso de suelo y cobertura terrestre para identificar correctamente todos los contornos de agua. Esto podría deberse a un bajo caudal del río Pesquería en ciertas temporadas, lo que dificulta la detección precisa del agua, especialmente en áreas donde los niveles hídricos son fluctuantes o marginales, lo cual es poco probable ya que en otras zonas si se logró obtener información de la lámina de agua.

El segundo factor obedece a una problemática ambiental subyacente: el severo estrés hídrico de la región. Como se destaca en la zona sombreada de la Figura 47, identificada como el sector de mayor criticidad en el área de estudio, se registró un déficit persistente de la lámina de agua tanto en la temporada de lluvias como en la de estiaje. Este comportamiento valida los resultados del análisis espacial y subraya la urgencia de implementar acciones para la restauración de este corredor fluvial. Cabe resaltar que la pérdida de continuidad en el flujo hídrico trasciende el impacto social, representando una grave amenaza para la conectividad y viabilidad de los ecosistemas locales. Por ello, se espera que el uso de estas geo tecnologías proporcione una base metodológica sólida para el diseño de futuros proyectos de rehabilitación ambiental y desarrollo social.

Es de gran importancia combinar las herramientas tecnológicas, como el análisis LULC, con datos complementarios y observaciones in situ, tal como se hizo en este estudio, para lograr una comprensión más holística del problema. Además, se deben diseñar e implementar estrategias orientadas a la sostenibilidad, como la mejora de prácticas de manejo del agua, la restauración de ecosistemas asociados al río y la promoción de la participación comunitaria en la gestión de los recursos hídricos.

A continuación, se observan los datos evidenciados del water mask como se muestra en la Figura 48.

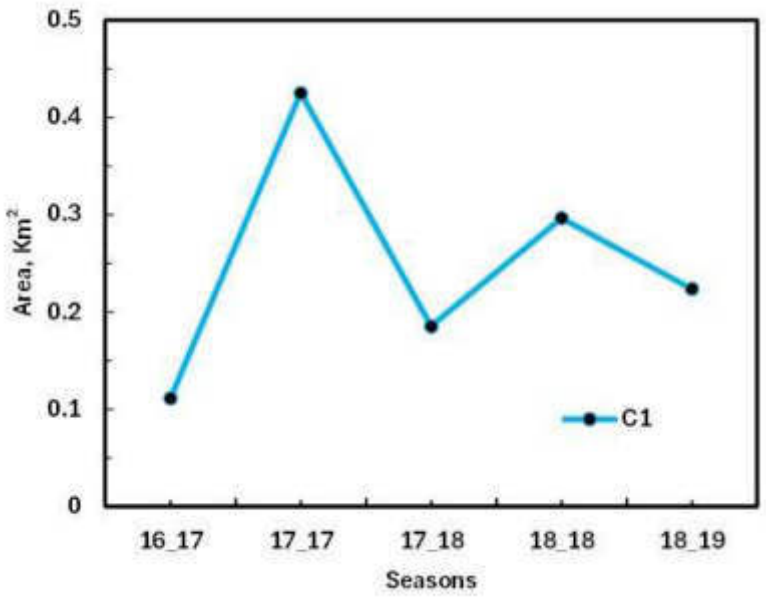


Figura 48. Área de la clase agua

Fuente: Elaboración propia

La Figura 48 muestra la variación del área identificada como agua (clase C1) en el río Pesquería a lo largo de cinco temporadas consecutivas: 2016-2017, 2017-2017, 2017-2018, 2018-2018 y 2018-2019. En la temporada 2016-2017, se identifica un aumento considerable en el water mask detectada, alcanzando un pico cercano a los 0.4 km² para el sitio de interés (como se observa en la Tabla 12). Sin embargo, en la siguiente temporada (2017-2017), esta área experimenta una disminución abrupta, cayendo por debajo de 0.2 km², lo que podría estar relacionado con una reducción en el caudal del río o con la influencia de factores externos que afectaron la visibilidad del agua en las imágenes clasificadas.

En la temporada 2017-2018, el water mask vuelve a aumentar a aproximadamente 0.3 km², lo que podría indicar un aumento temporal en el flujo del río o una mejora en la clasificación de los datos. No obstante, en la temporada

2018-2019, el área vuelve a disminuir ligeramente a poco más de 0.2 km², sugiriendo una posible tendencia hacia la reducción del recurso hídrico.

Esta variabilidad en los datos puede ser consecuencia de diversos factores, como las fluctuaciones estacionales y climáticas que afectan el caudal del río en algunas zonas del río ya que como se vio en las figuras anteriores algunas zonas no fueron afectadas por las estaciones. Para interpretar con mayor certeza estos resultados, sería necesario realizar un monitoreo constante, validando los datos a través de mediciones directas en el terreno. Además, se recomienda implementar estrategias de gestión sostenible que permitan mitigar los posibles efectos negativos sobre el ecosistema fluvial y asegurar el abastecimiento de agua en el futuro.

7.1 Comparación de water mask

En el estudio de la dinámica de los cuerpos de agua, especialmente en ríos con características particulares como el flujo estrecho representativo en los últimos años del río Pesquería, es crucial seleccionar métodos de extracción de agua que sean precisos y adecuados a las condiciones del entorno. Para este propósito, se han utilizado diversas técnicas de clasificación y análisis espectral. En este caso, se compararon tres enfoques diferentes para la extracción de agua a lo largo de cinco temporadas (2016-2019). La clasificación LULC con la clase 1 asignada al agua, el índice NDWI, el Índice MNDWI, el modelo AWEI y el ESA world cover (clase agua). Cada uno de estos métodos presenta ventajas y limitaciones dependiendo de factores como la resolución de las imágenes y las características específicas del río.

La Figura 49 ilustra la comparación entre estos métodos, mostrando los resultados obtenidos en cada temporada y proporcionando un análisis visual de su efectividad y precisión.

Esta comparación proporciona una visión clara sobre las ventajas y limitaciones de cada uno de los enfoques en la extracción de cuerpos de agua en un entorno fluvial.

En el caso de la clasificación LULC, la extracción de agua a través de la clase 1 muestra buenos resultados en general, especialmente en las temporadas donde los niveles de agua son más evidentes. La clasificación identifica de manera precisa las áreas cubiertas por agua, sin embargo, su efectividad depende de la resolución espacial y temporal de las imágenes utilizadas. Este método presenta una ventaja en términos de interpretar visualmente los patrones de uso del suelo y agua a lo largo del tiempo, y en términos generales se tuvo una mejor respuesta que los índices NDWI y MNDWI.

El índice NDWI, por otro lado, tuvo dificultades para realizar una extracción efectiva del agua debido a las características del río Pesquería, específicamente su bajo y angosto flujo. El NDWI, al ser un índice que depende de las diferencias espectrales en las longitudes de onda del infrarrojo cercano y del verde, no pudo capturar correctamente las variaciones de agua en las áreas más pequeñas o con menor contraste, lo que resalta la necesidad de imágenes satelitales de mayor resolución espacial para mejorar su desempeño o en su debido caso este índice está más enfocado a estudiar ríos de mayor anchura y flujo. Esta limitación es clara en las imágenes obtenidas, donde el NDWI no logró extraer adecuadamente las áreas de agua (verificadas) del río en algunas temporadas.

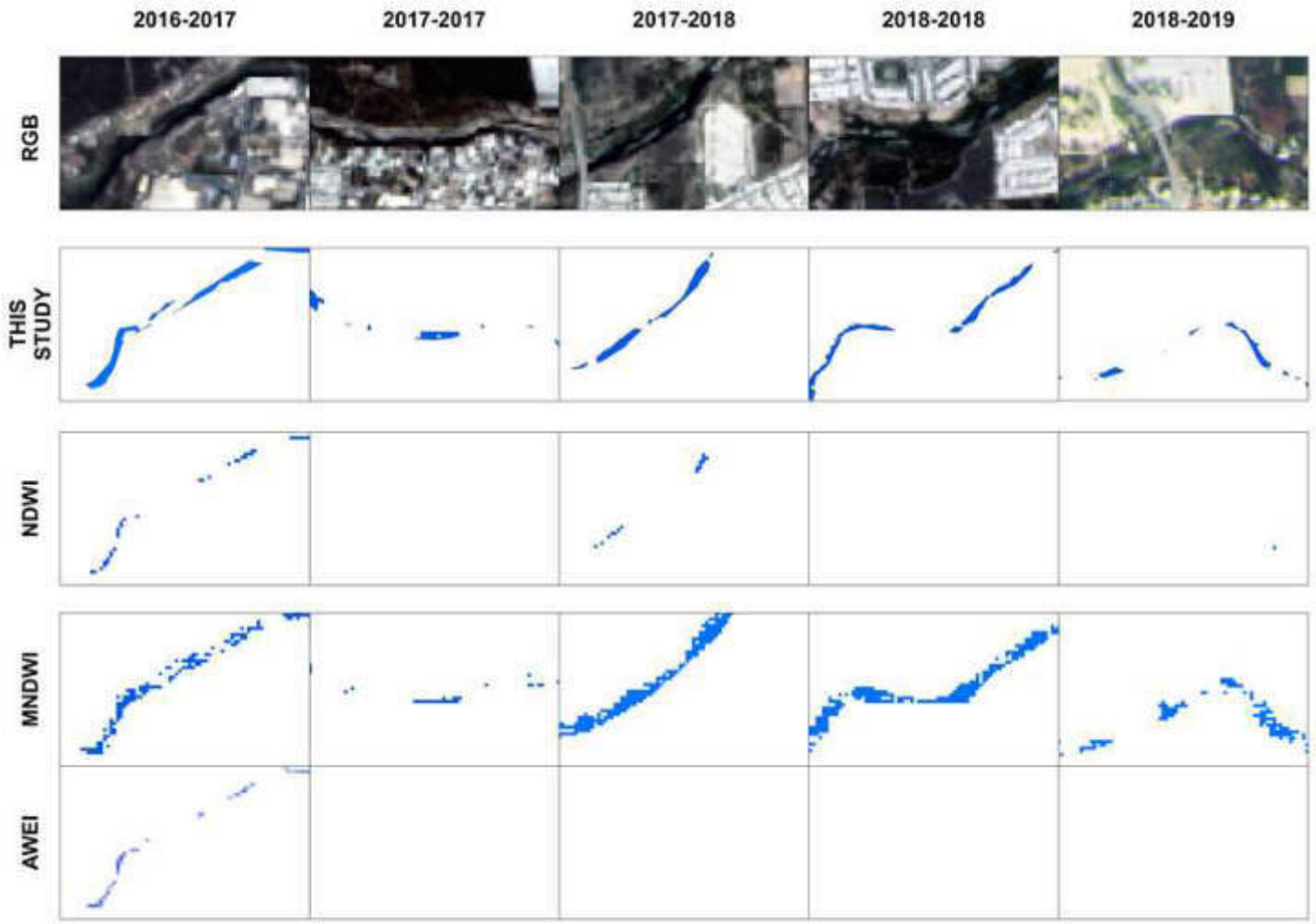


Figura 49. Comparación extracción agua

Fuente: Elaboración propia

Por otro lado, el índice MNDWI mostró resultados más favorables y similares a la clasificación LULC en términos de la extracción de cuerpos de agua. El MNDWI, que se basa en la diferencia entre las bandas del verde y del infrarrojo medio, presenta una mayor sensibilidad para detectar agua en áreas urbanas o de vegetación densa. Sin embargo, a pesar de sus buenos resultados generales, el MNDWI también presentó algunos errores al clasificar como agua ciertas estructuras, como puentes que cruzan el río Pesquería, un detalle que puede generar confusión en la interpretación de los datos. Este error es visualmente evidente en las imágenes, donde los puentes, que no deberían ser clasificados como cuerpos de agua, aparecen resaltados en azul, lo que indica la necesidad de ajustar el índice o aplicar técnicas de corrección para mejorar la precisión.

El método AWEI por su parte solo pudo representar una respuesta espectral en la temporada 16_17, en las otras temporadas no se logró obtener una respuesta del río Pesquería.

Además, de manera general se hizo una verificación usando el modelo de la ESA World Cover con un filtro para la clase de agua (80) la cual tampoco logró identificar al menos un pixel de agua para el río Pesquería en ninguna de las temporadas.

En terminos generales, los métodos de extracción analizados muestran limitaciones en función de las características del río Pesquería y la resolución de las imágenes satelitales utilizadas. La clasificación LULC con la clase 1 mostró consistencia y precisión, mientras que el NDWI requirió una mayor resolución para mejorar su desempeño. Por su parte, el MNDWI, aunque efectivo, presentó cierta confusión en su respuesta espectral entre estructuras (puentes) y agua. Estos hallazgos subrayan la importancia de elegir el método de extracción adecuado en función de las condiciones del área de estudio y la calidad de los datos disponibles, además de resaltar la necesidad de realizar ajustes y validaciones para mejorar la precisión en estudios de monitoreo hídrico.



Figura 50. Error MNDWI

Fuente: Elaboración propia

La Figura anterior destaca de manera clara el error observado en la extracción de agua utilizando el índice MNDWI, un problema que no se presentó con la clasificación LULC. Este contraste subraya la importancia y la robustez de la clasificación LULC, la cual no solo se diseñó para el monitoreo de uso de suelo, zonas urbanas y vegetación, sino que también demuestra ser una herramienta eficaz para el estudio de cuerpos de agua con características específicas, como los ríos de bajo flujo y cauce estrecho, tal como es el caso del río Pesquería. Este análisis resaltó la precisión de la clasificación LULC en contextos donde otros métodos pueden tener limitaciones, particularmente cuando se trata de flujos complejos y difíciles de identificar.

CALIDAD DE AGUA

8 Calidad de Agua

8.1 Análisis de datos

Debido a la complejidad de estudiar ríos urbanos de bajo caudal con imágenes satelitales Sentinel-2, en este estudio se procedió a hacer dos tipos de análisis de datos con la finalidad de identificar los varios comportamientos espectrales con los cambios de calidad de agua. Uno de ellos fue estudiar la correlación entre diversas bandas espectrales e índices espectrales por cada punto de muestreo de campo individual y el otro fue un análisis holístico identificando las correlaciones en general para los 7 puntos de muestreos en campo.

8.2 Correlación por punto de muestreo

En esta sección se evaluó la correlación lineal entre la reflectancia superficial de Sentinel-2 y 13 parámetros fisicoquímicos de calidad de agua (DQO, P total, pH, oxígeno disuelto, turbiedad, SST, SDT, color verdadero, N-amoniacal, N-nitritos, N-nitratos, N-orgánico, N-total y N-total Kjeldahl), por cada punto de muestreo (P1-P7) sobre el río Pesquería. En la Tabla 15 se observan las relaciones de los parámetros fisicoquímicos con las Bands Water Quality Points (BWQP) por cada punto de muestreo.

Tabla 14. Correlación por punto de muestreo y parámetro por bandas espectrales

Parámetro	P1	P2	P3	P4	P5	P6	P7
COLOR_VER	-0.624	0.204	0.642	-0.433	-0.616	-0.546	-0.572
DQO_TOT	0.569	-0.517	-0.731	-0.511	-0.641	-0.561	-0.571
N_NH3	0.712	0.312	0.773	-0.565	-0.656	-0.723	-0.717
N_NO2	0.748	-0.650	-0.578	-0.531	-0.468	-0.756	-0.697
N_NO3	0.416	-0.607	-0.470	-0.199	-0.372	-0.412	-0.411

Parámetro	P1	P2	P3	P4	P5	P6	P7
N_ORG	-0.554	0.407	0.601	0.446	0.405	0.341	-0.332
N_TOT	0.368	-0.656	0.715	-0.649	-0.762	-0.837	-0.821
N_TOTK	0.601	0.348	0.798	-0.560	-0.729	-0.709	-0.748
OD_mg_L	-0.457	-0.568	-0.283	0.236	0.305	0.372	-0.336
P_TOT	0.525	0.439	-0.887	-0.645	-0.802	-0.777	-0.683
SDT	0.557	-0.450	-0.406	-0.187	0.249	-0.290	-0.465
SST	0.515	0.411	0.696	-0.345	-0.424	-0.396	0.558
TURBIEDAD	0.501	0.512	0.671	0.236	-0.364	-0.247	0.390
pH_CAMPO	-0.383	0.399	0.316	0.394	0.317	0.172	0.346

Se logra identificar de manera muy rápida como las mejores correlaciones se encuentran en los parámetros N_NH3, N_NO2, N_TOT, N_TOTK y P_TOT. Dando un indicio a que estos parámetros presentan un mejor comportamiento lineal y a su vez una menor colinealidad en comparación a los otros parámetros fisicoquímicos. Además, se puede decir que en P2 se presentaron las menores correlaciones, aun así, los nitrógenos obtuvieron las mejores relaciones de ese punto en cuestión.

Por otro lado, se realizaron las correlaciones de Pearson usando los index water quality points (IWQP) mencionados en esta investigación para cada uno de los puntos de muestreo y los parámetros fisicoquímicos mencionados anteriormente. En la Tabla 16 se observan lo resultados de las mejores correlaciones por punto.

Tabla 15. Correlación por punto de muestreo y parámetro por índices espectrales

Parámetro	P1	P2	P3	P4	P5	P6	P7
COLOR_VER	-0.505	0.260	-0.447	-0.149	-0.347	0.301	-0.213
DQO_TOT	-0.751	-0.570	0.448	-0.504	0.595	-0.670	-0.658
N_NH3	-0.569	0.479	0.574	0.668	0.435	-0.273	-0.425
N_NO2	0.663	-0.558	-0.480	-0.390	-0.334	-0.291	-0.474
N_NO3	-0.202	-0.502	-0.402	0.491	-0.504	0.499	0.387
N_ORG	-0.789	-0.380	0.370	-0.507	-0.502	0.383	0.316
N_TOT	0.300	-0.651	0.587	0.613	0.428	-0.473	-0.511
N_TOTK	-0.424	0.295	0.578	0.637	0.487	-0.313	-0.407
OD_mg_L	-0.228	-0.426	-0.556	0.472	0.651	-0.513	-0.357
P_TOT	-0.526	0.623	-0.478	0.363	-0.504	-0.383	0.418
SDT	0.368	0.553	-0.536	0.251	-0.290	-0.527	-0.408
SST	-0.390	-0.255	0.594	-0.520	-0.518	0.358	-0.704
TURBIEDAD	-0.318	0.425	0.592	-0.437	-0.383	0.255	0.370
pH_CAMPO	0.366	-0.492	-0.546	-0.443	0.399	-0.181	-0.389

A diferencia de la Tabla 15, no se encuentran buenas correlaciones en los 7 puntos de muestreo para cada uno de los índices espectrales usados. Sin embargo, las mejores correlaciones siguieron estando en los nitrógenos y algunas muy alejadas en la DQO_TOT y SST para el P1 y P7 respectivamente.

Cabe resaltar que esto es solo un indicio de los resultados de correlación por punto de muestreo, es importante también poder analizarlos en conjunto para así tener un comportamiento más holístico de los cambios en las diferentes temporadas estudiadas.

8.2.1 Análisis de correlaciones

Para determinar las mejores predicciones para BWQP y IWQP de cada uno de los parámetros fisicoquímicos se promedió dicha correlación a través de los siete puntos de muestreo. Estos resultados demuestran una disminución de la linealidad dando a entender la importancia de trabajar todos los puntos unidos a diferencia de individual. Aunque, de igual manera los mayores valores de relación lineal siguen siendo los nitrógenos y fósforos, específicamente N_NO2, N_TOT, N_TOTK y P_TOT. En la Tabla 17 se observan los parámetros con cada uno de sus correlaciones promediadas y el mejor predictor.

Tabla 16. Correlaciones promediadas

Parámetro	|r| medio BWQP (bandas)	|r| medio IWQP (índices)	Mejor predictor
COLOR_VER	0.519	0.317	**BANDAS**
DQO_TOT	0.586	0.599	**ÍNDICES**
N_NH3	0.637	0.489	**BANDAS**
N_NO2	0.633	0.456	**BANDAS**
N_NO3	0.412	0.426	**ÍNDICES**
N_ORG	0.441	0.464	**ÍNDICES**
N_TOT	0.687	0.509	**BANDAS**
N_TOTK	0.642	0.449	**BANDAS**
OD_mg_L	0.365	0.458	**ÍNDICES**

Parámetro	|r| medio BWQP (bandas)	|r| medio IWQP (índices)	Mejor predictor
P_TOT	0.680	0.471	**BANDAS**
SDT	0.372	0.419	**ÍNDICES**
SST	0.478	0.477	**BANDAS**
TURBIEDAD	0.417	0.397	**BANDAS**
pH_CAMPO	0.333	0.402	**ÍNDICES**

De manera específica, en la Tabla 18 se observan las correlaciones más fuertes obtenidas para cada uno de los puntos. Dando como primer análisis los mejores predictores espectrales para cada parámetro, pero exclusivamente por punto de muestreo.

Tabla 17. Mayor correlación por punto y por predictor espectral

Tipo	Punto	Parámetro	Predictor espectral	r Pearson	R^2
BANDAS	P3	**P_TOT**	B3+B5/B1+B4	**-0.887**	0.786
BANDAS	P6	**N_TOT**	B3	**-0.837**	0.700
BANDAS	P7	**N_TOT**	B1	**-0.821**	0.674
BANDAS	P5	**P_TOT**	B1	**-0.802**	0.643
BANDAS	P3	**N_TOTK**	B4/(B11+B12)	**0.798**	0.637
ÍNDICES	P1	**N_ORG**	NDCI	**-0.789**	0.622
BANDAS	P6	**P_TOT**	B3	**-0.777**	0.604
BANDAS	P3	**N_NH3**	B4/(B11+B12)	**0.773**	0.597

Tipo	Punto	Parámetro	Predictor espectral	r Pearson	R^2
BANDAS	P5	**N_TOT**	B1	**-0.762**	0.580
BANDAS	P6	**N_NO2**	B3	**-0.756**	0.572
ÍNDICES	P1	**DQO_TOT**	FAI	**-0.751**	0.564
BANDAS	P7	**N_TOTK**	B1	**-0.748**	0.560
BANDAS	P1	**N_NO2**	B3+B5/B1+B4	**0.748**	0.559
BANDAS	P3	**DQO_TOT**	B3+B5/B1+B4	**-0.731**	0.535
ÍNDICES	P7	**SST**	NDCI	**-0.704**	0.496

8.2.2 Dispersogramas y Heatmaps

Se seleccionaron los dispersogramas y los Heatmaps de correlación de Pearson más representativos de la investigación. Sin embargo, no son los únicos los demás serán enviados al anexo de este documento.

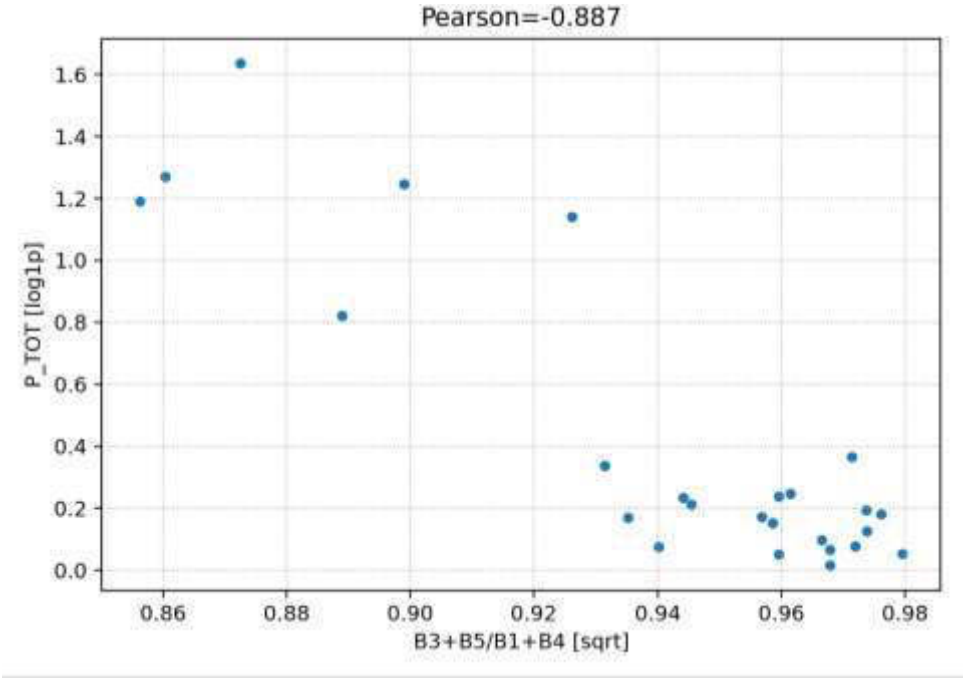


Figura 51. P_TOT vs. combinación de bandas (B3+B5)/(B1+B4) en el punto P3

Fuente: Elaboración propia

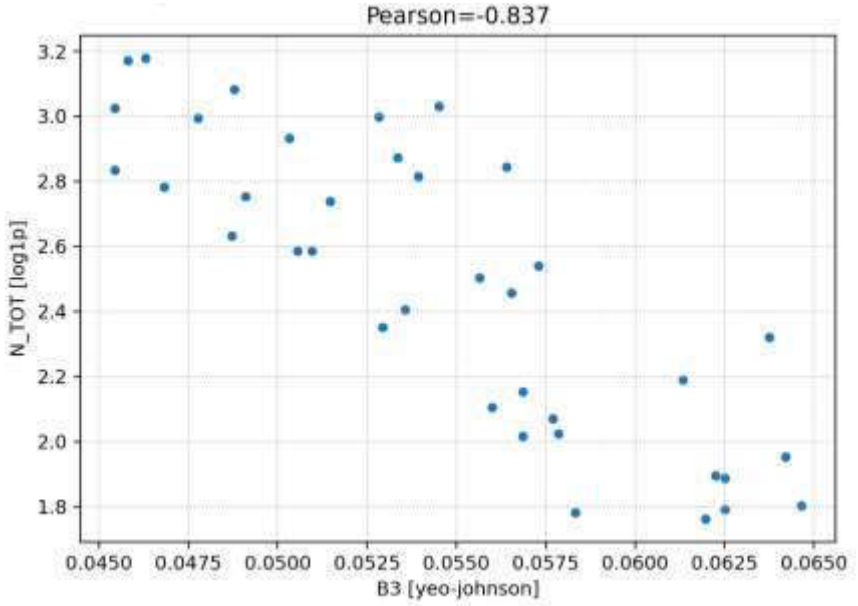


Figura 52. Nitrógeno total vs. banda B3 (verde) en el punto

Fuente: Elaboración propia

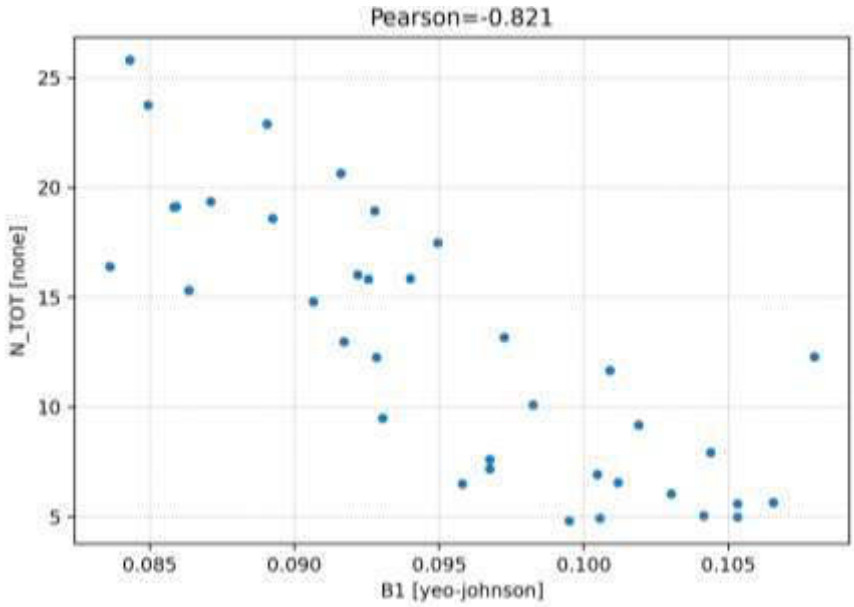


Figura 53. Nitrógeno total vs. banda B1en el punto P7

Fuente: Elaboración propia

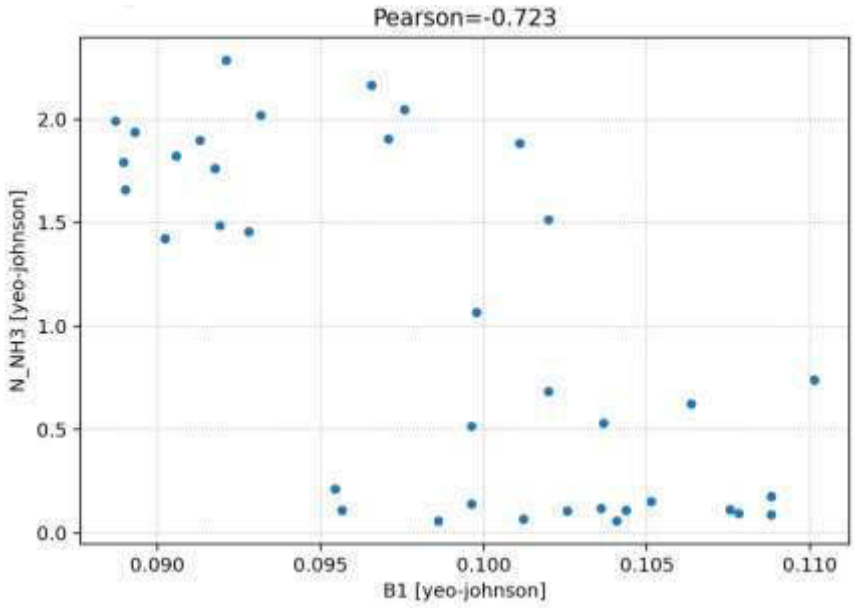


Figura 54. Nitrógeno amoniacal vs. banda B1 en el punto P6

Fuente: Elaboración propia

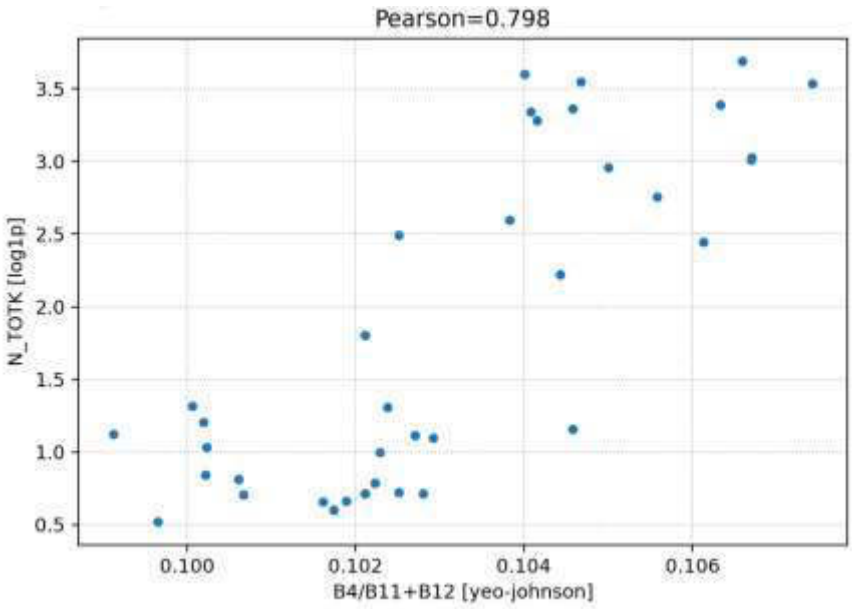


Figura 55. Nitrógeno total Kjeldahl vs. B4/(B11+B12) en el punto P3

Fuente: Elaboración propia

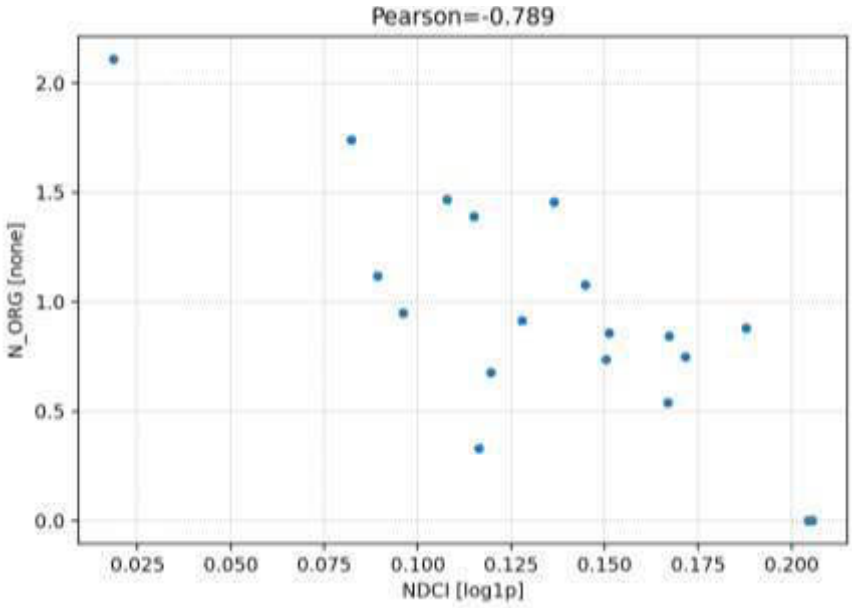


Figura 56. Nitrógeno orgánico vs. NDCI (Normalized Difference Chlorophyll Index) en el punto P1

Fuente: Elaboración propia

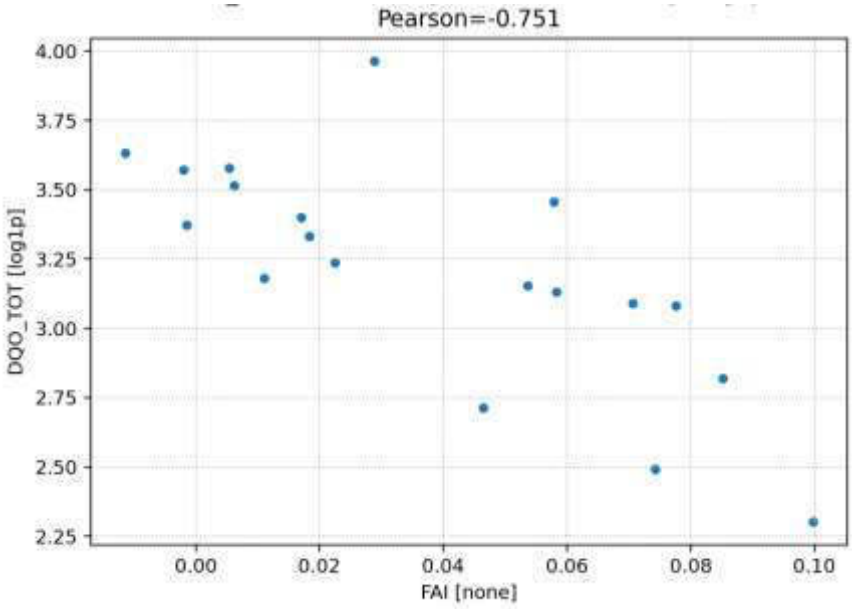


Figura 57. DQO vs. FAI en el punto P1

Fuente: Elaboración propia

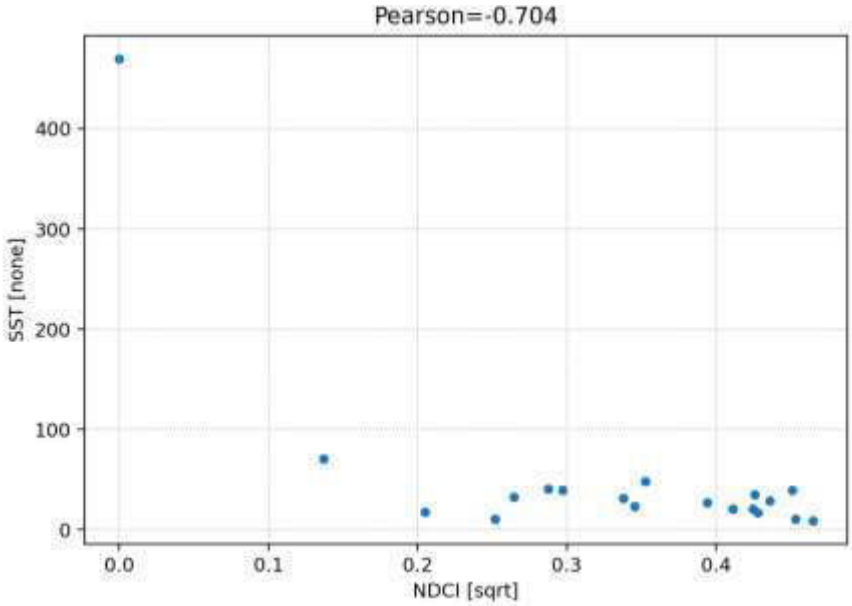


Figura 58. Sólidos suspendidos totales vs. NDCI en el punto P7

Fuente: Elaboración propia

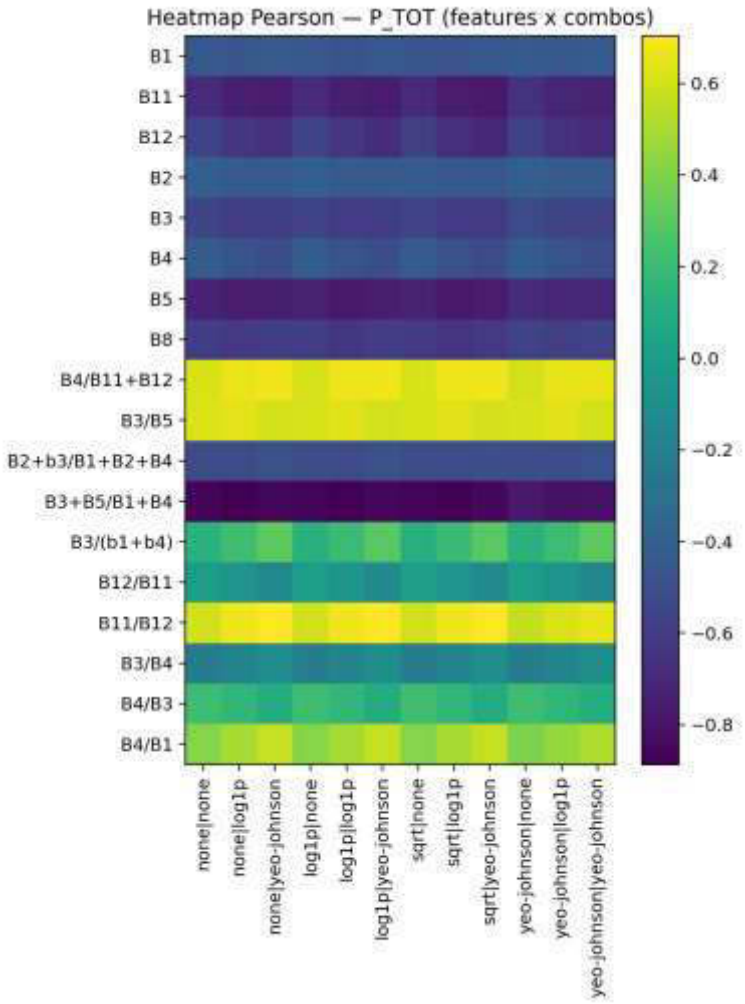


Figura 59. Heat Map de Pearson para las 20 variables espectrales de banda frente a las 12 combinaciones de transformación evaluadas

Fuente: Elaboración propia

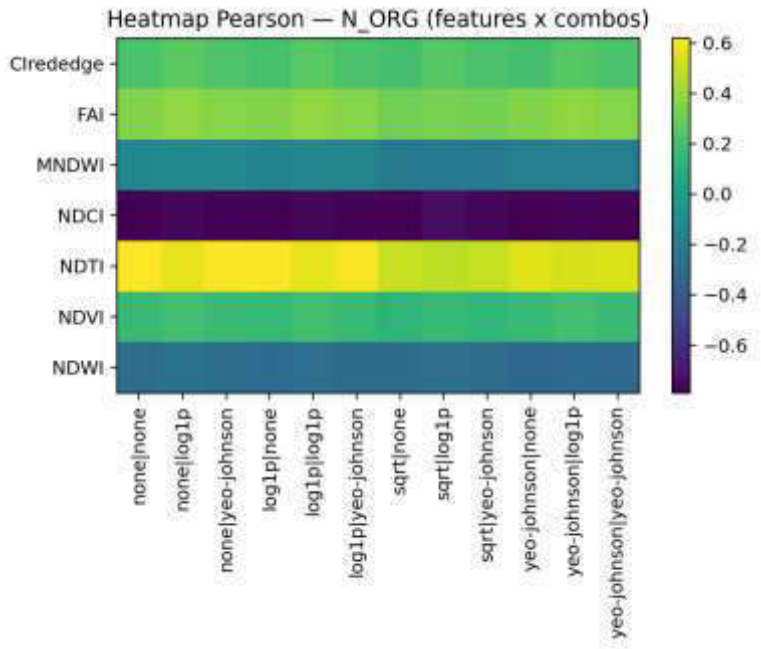


Figura 60. Heat Map de Pearson para los 7 índices espectrales frente a las combinaciones de transformación evaluadas

Fuente: Elaboración propia

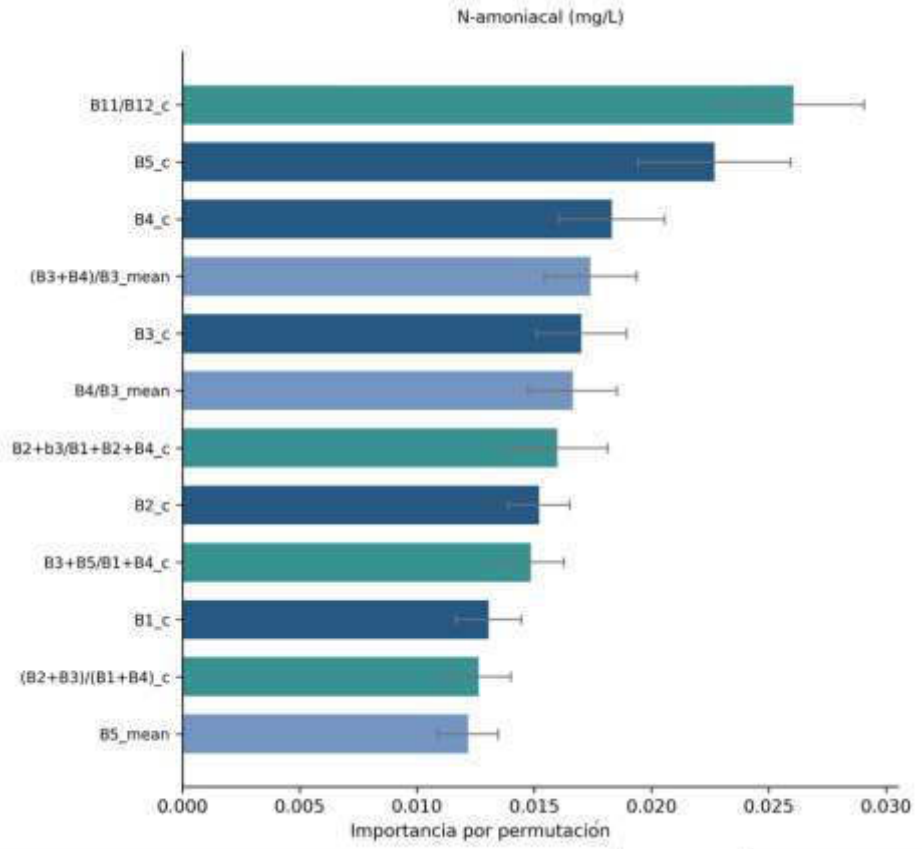
Los resultados de los scatter generados para estimar parámetros de calidad del agua a partir de bandas satelitales muestran que fue necesario aplicar transformaciones estadísticas para lograr los mejores ajustes. El uso de métodos como Yeo-Johnson, transformaciones logarítmicas (Log1p) y raíces cuadradas, tanto en las variables independientes como en las dependientes, refleja la naturaleza asimétrica y no lineal de estos datos medioambientales. En la mayoría de los casos, las relaciones encontradas son inversas, con coeficientes de correlación de Pearson negativos respecto a los valores medidos en campo.

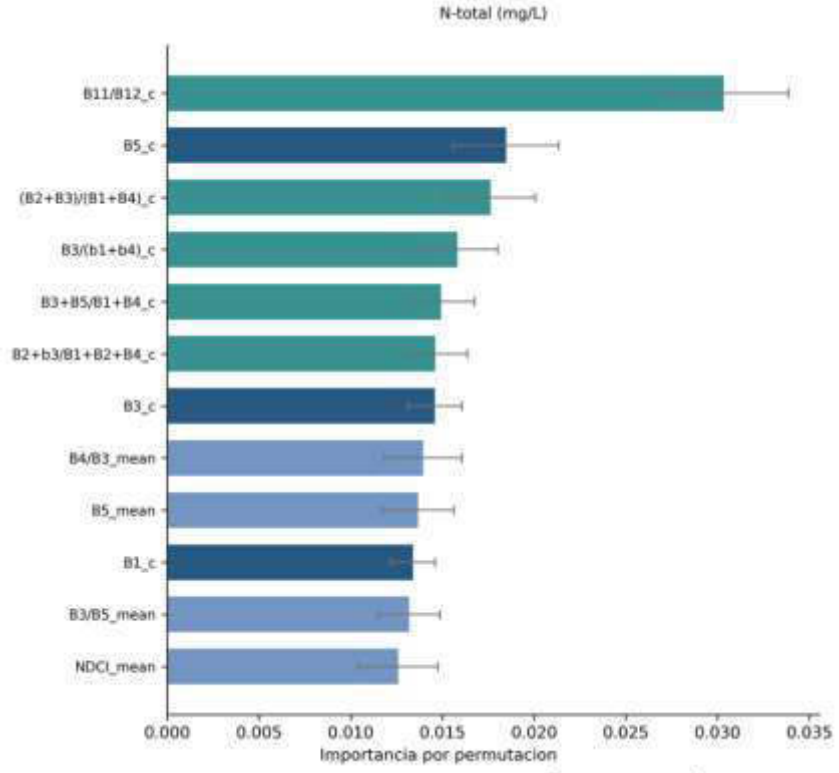
El fósforo total presenta el mejor ajuste predictivo, con un coeficiente de correlación de -0.887. Este resultado se obtuvo mediante un cociente de sumas de bandas transformado por raíz cuadrada y una transformación logarítmica en la variable dependiente. La gráfica muestra una tendencia descendente muy clara, lo que sugiere una alta viabilidad para modelar este parámetro mediante teledetección. En un nivel de desempeño similar, el nitrógeno total responde favorablemente al uso de bandas individuales transformadas con el método Yeo-

Johnson, obteniendo correlaciones de -0.837 y -0.821 con las bandas B3 y B1, respectivamente. La banda B3 muestra una menor dispersión de los datos en la parte central de la distribución, lo que podría traducirse en una mayor estabilidad del modelo.

Parámetros como el nitrógeno total Kjeldahl, el nitrógeno orgánico y la demanda química de oxígeno registran correlaciones con valores absolutos entre 0.75 y 0.79. El nitrógeno total Kjeldahl es el único que muestra una relación directa positiva, con un Pearson de 0.798. Sin embargo, los datos de estos tres parámetros presentan una mayor dispersión general. La forma en que se agrupan los puntos en las gráficas muestra la influencia de factores externos no registrados directamente, como diferencias entre los sitios de monitoreo o las fechas de muestreo. Además, el tamaño reducido de la muestra hace que los valores estadísticos sean muy sensibles a cada observación.

Por ende, en la Figura 61 se observa el nivel de importancia de las bandas espectrales utilizadas para cada parámetro fisicoquímico estudiado.


N-amoniacal (mg/L)
B11/B12_c
B5_c
B4_c
(B3+B4)/B3_mean
B3_c
B4/B3_mean
B2+b3/B1+B2+B4_c
B2_c
B3+B5/B1+B4_c
B1_c
(B2+B3)/(B1+B4)_c
B5_mean
0.000
0.005
0.010
0.015
0.020
0.025
0.030
Importancia por permutación


N-total (mg/L)
B11/B12_c
B5_c
(B2+B3)/(B1+B4)_c
B3/(b1+b4)_c
B3+B5/B1+B4_c
B2+b3/B1+B2+B4_c
B3_c
B4/B3_mean
B5_mean
B1_c
B3/B5_mean
NDCI_mean
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
Importancia por permutacion

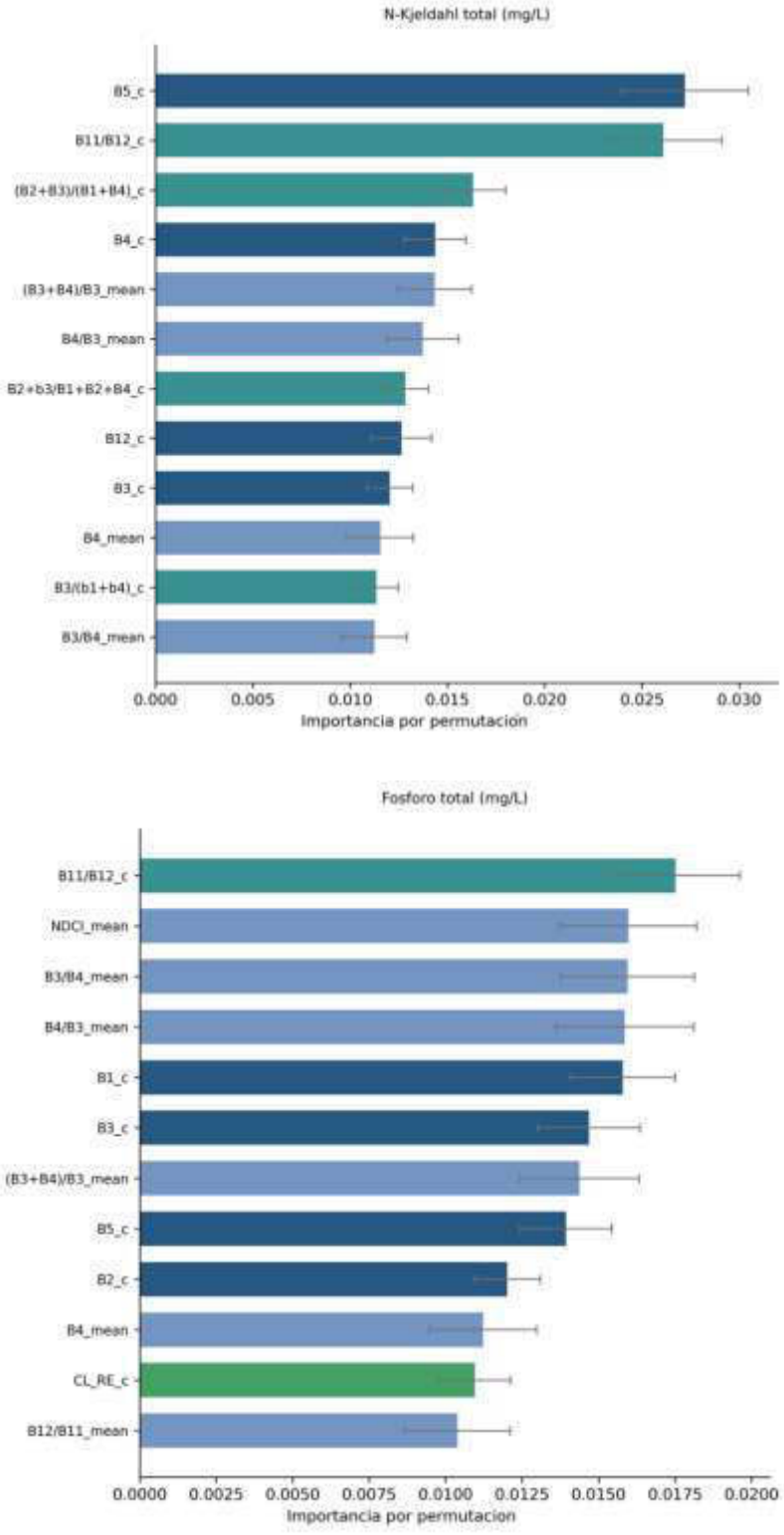


Figura 61. Importancia de Features

Fuente: Elaboración propia

La Figura 61 presenta la importancia de las variables obtenida mediante permutación en la prueba del modelo RF desarrollados para los parámetros de calidad del agua del río Pesquería. La comparación permite distinguir la contribución de cuatro grupos de predictores: la banda centrada, asociada con la variación temporal; el índice original centrado; el índice nuevo centrado, correspondiente a las combinaciones propuestas a partir de NDWI, MNDWI y otros índices espectrales; y la media espectral por punto, que representa el nivel espectral basal de cada sitio. En términos generales, se observa que la importancia no se concentra exclusivamente en un único tipo de variable, sino que cambia de acuerdo con el parámetro analizado.

Los resultados de importancia de variables muestran un comportamiento particularmente favorable para los compuestos nitrogenados y el fósforo total, en contraste con el resto de los parámetros evaluados. En N-$N0_2$, las variables con mayor contribución alcanzan valores cercanos a 0.045, seguidas por un segundo predictor alrededor de 0.040. Una tendencia similar se observa en NO_3, donde los dos primeros predictores presentan importancias próximas a 0.045 y 0.040, respectivamente. La concentración de la importancia predictiva en un conjunto reducido de variables sugiere que el modelo RF identifica una firma espectral consistente para modelar la variabilidad de estos compuestos. En el caso del Nitrógeno Orgánico (N-Orgánico), si bien la brecha de importancia entre los predictores es más estrecha, las variables principales conservan pesos significativos (oscilando entre 0.035 y 0.045). Esto evidencia que los datos presentan una respuesta espectral lo suficientemente sensible para discriminar gradientes en la concentración.

Para N-Total Kjeldahl, las dos variables principales alcanzan aproximadamente 0.028 y 0.026, mientras que el resto se mantiene por debajo de 0.020. En N-Total se presenta una situación semejante, aunque con una separación más marcada: el predictor principal alcanza cerca de 0.030, mientras que las demás variables presentan contribuciones considerablemente menores. En el caso de N-Amoniaco, la importancia máxima es menor, cercana a 0.025,

pero todavía se mantiene una diferencia apreciable respecto de varios de los predictores restantes. En conjunto, estos resultados muestran que la respuesta espectral desarrollada en este estudio presenta mayor capacidad informativa para los parámetros asociados con las diferentes formas del nitrógeno, particularmente para nitritos y nitratos. Esta consistencia entre diferentes fracciones nitrogenadas resulta relevante, debido a que sugiere que los predictores espectrales están capturando variaciones asociadas con el estado químico del agua y no únicamente el comportamiento de un parámetro aislado.

El Fósforo Total por su parte, presenta un comportamiento diferente, aunque también con una señal identificable. Las variables principales alcanzan aproximadamente 0.015–0.016, mientras que las siguientes mantienen valores muy próximos entre sí. La importancia se encuentra, por tanto, más distribuida entre los predictores y no existe un único indicador que domine claramente el modelo. Aun así, la presencia de varias variables con contribuciones similares sugiere que la información necesaria para la estimación del fósforo se encuentra distribuida entre diferentes componentes espectrales. En comparación con los compuestos nitrogenados, la magnitud de la importancia es menor, por lo que la respuesta predictiva para fósforo debe considerarse más moderada, aunque todavía presenta información espectral aprovechable. En los demás parámetros de calidad del agua, la importancia de las variables resulta menos consistente. Aunque DQO Total, pH, Oxígeno Disuelto y SST presentan algunos predictores con valores relativamente altos, estos resultados no mantienen el mismo patrón observado para los compuestos nitrogenados.

8.3 RF WQ Model

8.3.1 Mejora del modelo

Debido a la falta de datos de muestreo y a la periodicidad de las imágenes satelitales, es necesario hacer algunos cambios de análisis de datos para poder realizar un modelo de ML bien estructurado. Al momento de trabajar con todos

los datos en conjuntos se presentaban algunas limitaciones.Uno de los principales problemas de las versiones anteriores del modelo fue la marcada heterogeneidad espacial del río Pesquería. Como cada punto de muestreo presentaba un nivel basal de contaminación distinto, entrenar el modelo con todos los datos crudos daba lugar a la Paradoja de Simpson. Al analizar los datos en conjunto, el algoritmo ajustaba una regla global que explicara las diferencias entre sitios, y esa regla ocultaba los cambios temporales dentro de cada sitio. Para corregirlo se reestructuró la arquitectura del modelo con tres capas que separan la dinámica temporal de la espacial.

El modelo preliminar se construyó sobre 132 pares imagen-muestreo correspondientes a siete estaciones y diecinueve fechas de muestreo entre 2016 y 2019. El vector de predictores incluyó 27 variables espectrales centradas por sitio, sus respectivas medias por estación y el identificador de la estación, para un total de 55 predictores. La variable respuesta se transformó con log (1+y) en once de los catorce parámetros. Para atenuar el efecto de la variación de las imágenes satelitales se aplicó winsorización, un procedimiento de estadística robusta introducido por W. J. Dixon (1960) y John W. Tukey (1962). La winsorización sustituye los valores superiores a un umbral por el valor del propio umbral, sin eliminar la observación.

Dentro de los resultados obtenidos, los parámetros evaluados dan resultados muy dispares, y el patrón se explica por las propiedades ópticas de cada parámetro fisicoquímico. Los coeficientes de determinación bajo validación cruzada agrupada por fecha obteniendo valores negativos a 0.756. Los mejores resultados corresponden a los compuestos nitrogenados y a los parámetros ópticamente activos. El nitrógeno Kjeldahl obtuvo el mejor resultado del conjunto.

El modelo seleccionado ExtraTrees con y winsorización en el percentil 80 alcanzó $R^2 = 0.756$, RMSE de 4.45 mg/L y MAE de 3.35 mg/L sobre más de 42 observaciones. El RPD, cociente entre la desviación estándar de los valores observados y el error de predicción, fue de 2.05. Ese valor lo ubica en la categoría

de bueno de la escala convencional de quimiometría y lo hace el único parámetro del estudio por encima del umbral de 2.0. La pendiente de la regresión entre observado y predicho fue 0.97, indistinguible de la unidad, lo que descarta atenuación de regresión: el modelo cubre el rango completo de variación sin comprimir las predicciones hacia el centro de la distribución.

En consecuencia, el modelo para el Nitrógeno Total presentó métricas de rendimiento ligeramente inferiores, aunque mantuvo un nivel de consistencia estadística aceptable. Durante el preprocesamiento de este parámetro, se excluyó el 9.8 % de los registros originales, conservando un conjunto final de 46 pares de observaciones sobre los cuales se ejecutó el procedimiento de winsorización El modelo presentó un $R^2 = 0.669$, RMSE de 4.12 mg/L y RPD de 1.76, dentro de la categoría aceptable. La pendiente fue exactamente 1.00. El nitrógeno amoniacal por su parte presentó un desempeño de: $R^2 = 0.687$ y RMSE de 3.95 mg/L con winsorización en el percentil 80, sobre un rango de 0.06 a 19.33 mg/L.

Entre los parámetros ópticamente activos, la turbidez llegó a un $R^2 = 0.450$ con RMSE de 6.12 NTU en el rango calibrado de 2.4 a 27.5 NTU. El color verdadero obtuvo 0.424 de precisión y los SST de 0.395. Esto quiere decir que este modelo no logró relacionar los valores espectrales obtenidos de la imagen Sentinel-2 con los parámetros fisicoquímicos turbiedad, color verdadero y SST. Curiosamente estos parámetros son muy reportados en investigaciones relacionadas a modelos de calidad de agua tales como Miller (2022) quien obtuvo del 92% para turbudez en humedales. Rico-Serrano (2026) por su parte obtuvo una precision del 75% para los SST y Amaro (2024) presentó una precisión del 55% en el Lago Buhi. Cabe resaltar que este estudio eleva la complejidad en comparación a los anteriores puesto que al ser un río angosto se puede estar alterando las bandas visibles del espectro y también que la heterogeneidad del río en base a sus cambios tan abruptos implica un sistema complejo de estimar. Además, es por lo que se deberían obtener más datos de muestreos para disminuir esta heterogeneidad que presenta el río Pesquería y a su vez emplear modelos que

tengan en cuenta estas variaciones. Lastimosamente, para el río Pesquería no existe una gran cantidad de datos necesaria como para aplicar un modelo de ML en optimas condiciones.

Por otra parte, los últimos cinco parámetros no alcanzaron capacidad predictiva. Los sólidos disueltos totales obtuvieron una precisión de $R^2 = 0.191$, los nitratos 0.143, el oxígeno disuelto 0.085, el nitrógeno orgánico cerca de cero y el pH un valor negativo. Ninguno tiene un mecanismo físico que lo vincule con la reflectancia superficial. Amado-Álvarez et al. (2019) reportaron lo mismo en embalses del río Conchos, dentro de la misma cuenca: de oxígeno disuelto, pH, nitratos, sólidos disueltos y turbidez, solo la turbidez obtuvo un modelo aceptable. Sin embargo, no es comparable a este estudio por la gran diferencia entre el ancho de lámina de agua del río Pesquería con el de las presas.

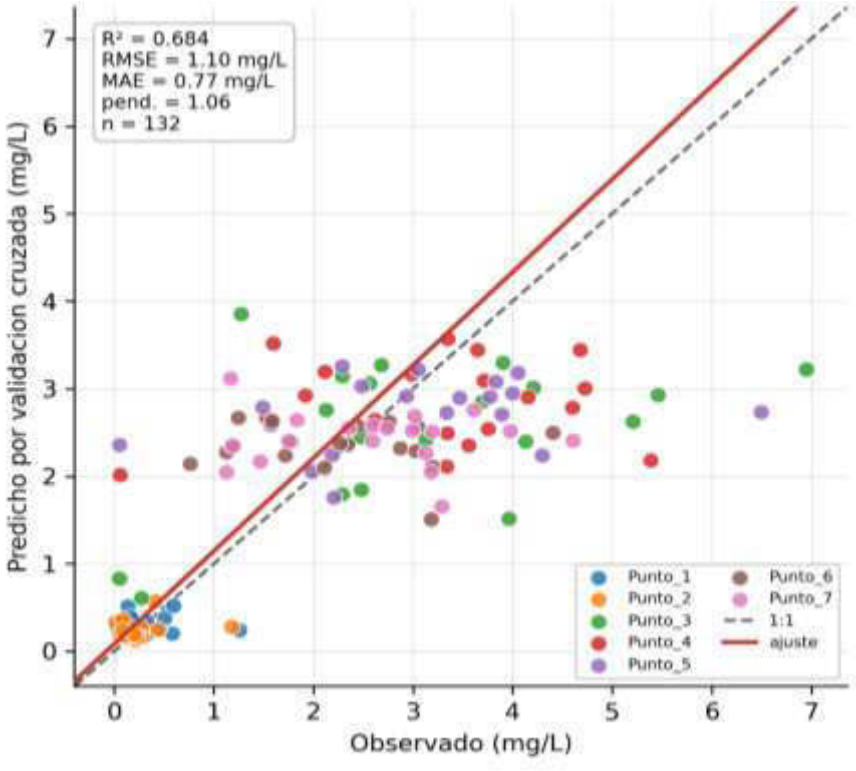


Figura 62. P_TOT

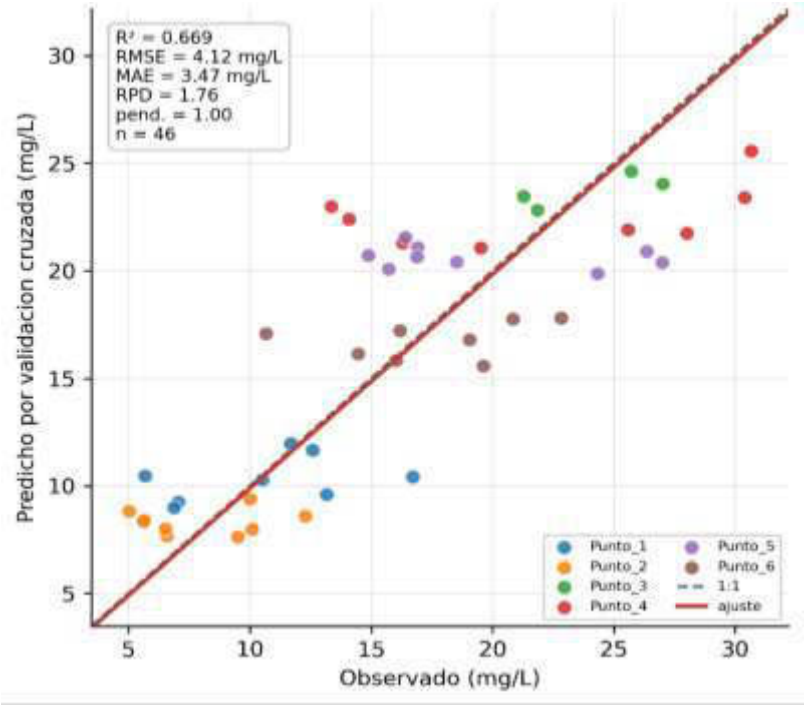


Figura 63. N_TOT

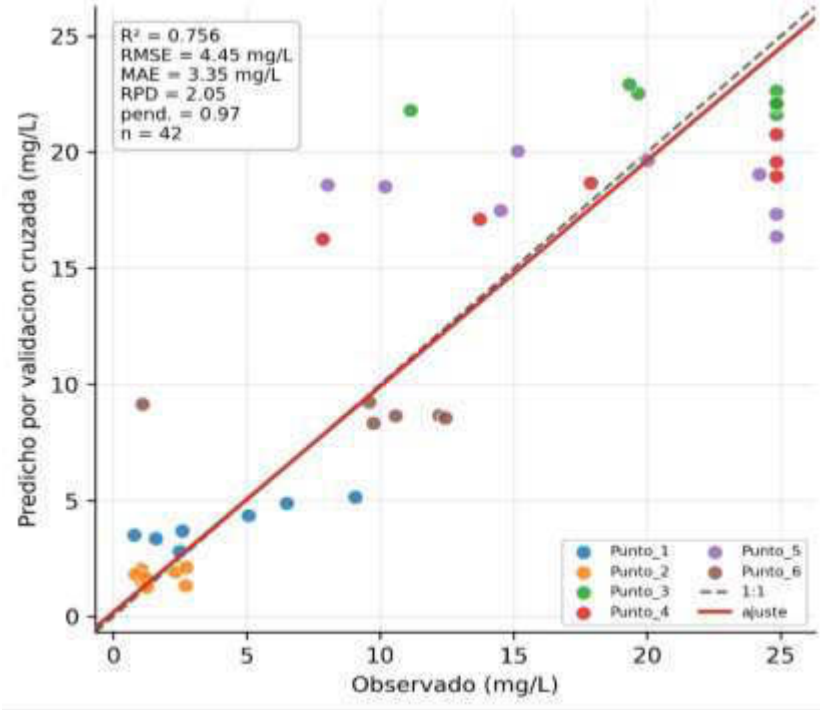


Figura 64. N_Kjeldahl

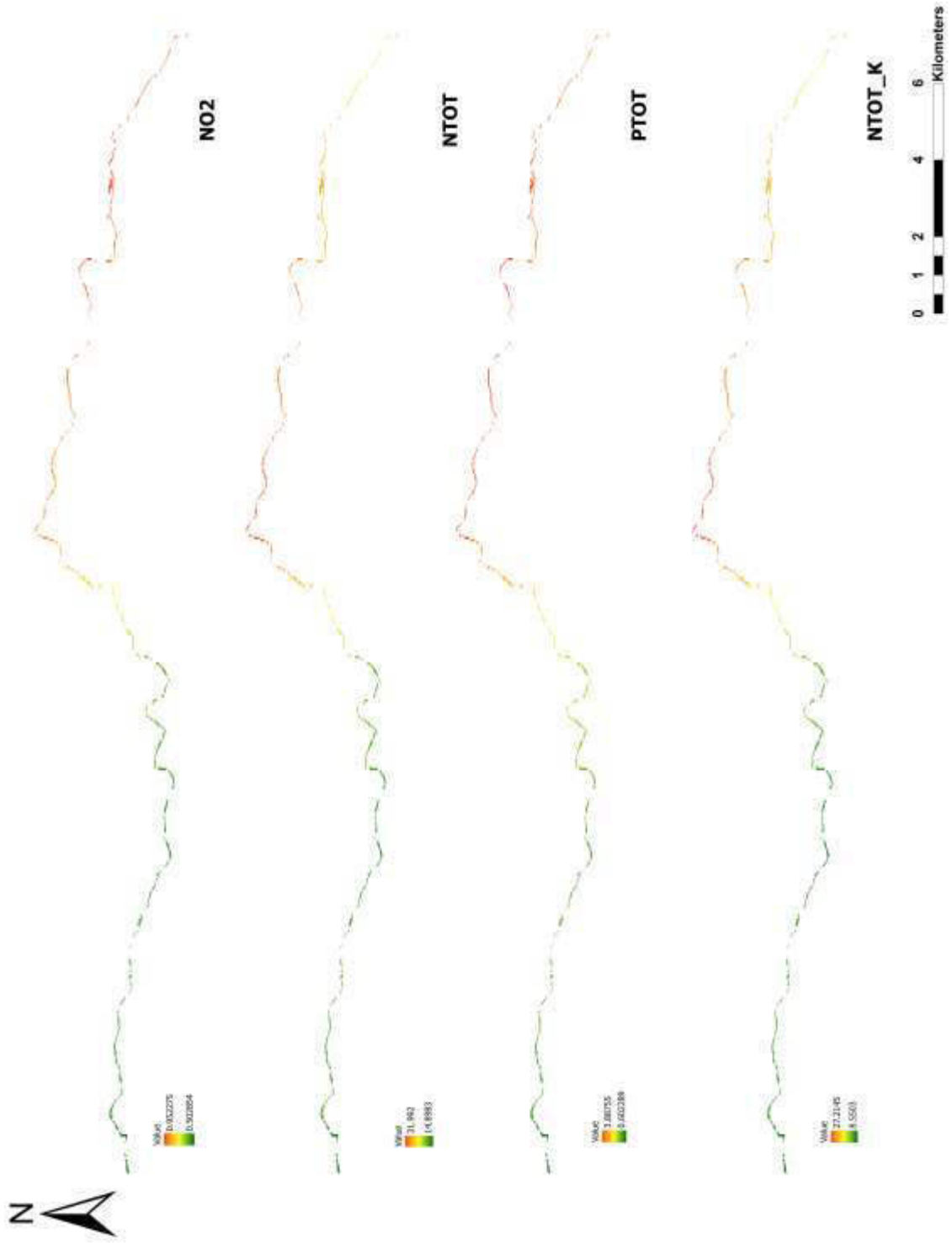


Figura 65. Calidad de Agua

Fuente: Elaboración propia

Las series temporales del periodo 2016–2019 muestran dos comportamientos claramente diferenciados entre los ocho parámetros representados, y la separación coincide con la transformación aplicada a la variable respuesta.



Figura 66. Series temporales de predicción

Fuente: Elaboración propia

De acuerdo con la Figura 66 la comparación entre las series observadas y las estimadas por los modelos muestra que la mejor correspondencia temporal se obtuvo para los parámetros asociados con nutrientes, particularmente para las diferentes formas de nitrógeno y Fósforo Total. En estos casos, las series predichas reproducen con mayor claridad la dirección y magnitud relativa de las variaciones observadas entre las fechas de muestreo. El comportamiento es especialmente evidente en N-Total, donde los incrementos y descensos de la concentración son seguidos de manera más consistente por las predicciones, incluyendo los cambios registrados durante los periodos de mayor concentración. Aunque existen diferencias puntuales entre ambos valores, el modelo conserva la estructura temporal de la serie, lo que representa una respuesta favorable para este parámetro. Nitrógeno total, Fósforo total ajustados sin transformación o con logaritmo presentan predicciones que siguen la trayectoria de los valores observados. El nitrógeno total las curvas predichas reproducen tanto el nivel basal de cada estación como las oscilaciones interanuales dentro del rango de 5 a 40 mg/L. En Fósforo total la correspondencia se mantiene entre 0 y 7 mg/L, y en pH las predicciones cubren el intervalo observado de 6.9 a 8.4, si bien con desfases en los picos.

Un comportamiento semejante se observa en N-Amoniaco, donde las predicciones consiguen representar buena parte de las fluctuaciones observadas, principalmente durante los incrementos registrados hacia la parte final del periodo analizado. La capacidad para reproducir estos cambios resulta relevante, ya que no solamente se obtiene una aproximación a los valores individuales, sino que se mantiene parcialmente la dinámica temporal del nutriente. En Fósforo Total se aprecia igualmente una correspondencia favorable: los valores predichos que siguen de manera general los principales máximos y mínimos de la serie observada, aunque con una menor amplitud en algunos eventos. Esta respuesta coincide con el análisis de importancia de variables presentado anteriormente, donde los predictores asociados con los nutrientes mostraron una contribución más consistente dentro de los modelos.

En contraste, los resultados obtenidos para DQO Total, Oxígeno Disuelto, Turbidez, SST y pH presentan una correspondencia considerablemente menor entre las series observadas y predichas. En DQO Total y Turbidez, por ejemplo, se observan eventos de concentración elevada que no son reproducidos adecuadamente por el modelo, particularmente los picos extremos presentes en determinados muestreos. Esto genera diferencias importantes entre las trayectorias observadas y estimadas. El caso de SST es todavía más evidente, debido a la presencia de un incremento puntual de gran magnitud que el modelo no logra representar. Por su parte, Oxígeno Disuelto y pH muestran una dispersión considerable entre observaciones y predicciones, sin una coincidencia temporal suficientemente estable a lo largo de la serie.

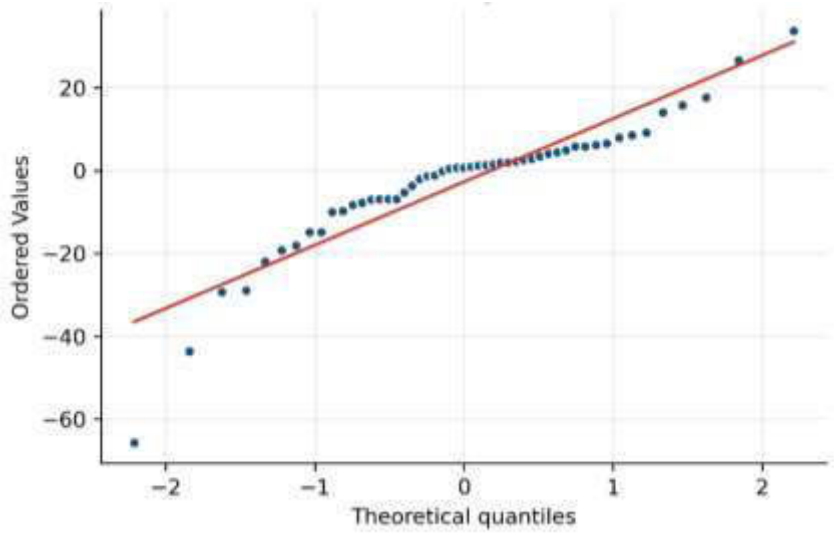


Figura 67. Q-Q Normalidad

Fuente: Elaboración propia

La Figura anterior presenta un gráfico Q-Q (Quantile–Quantile) que permite evaluar visualmente el supuesto de normalidad de los valores analizados. La distribución presenta una aproximación razonable a la linealidad en la zona central, donde la mayor parte de los puntos se concentra alrededor de la recta de

referencia. Esto indica que los valores centrales mantienen un comportamiento relativamente próximo al esperado bajo una distribución normal. Sin embargo, la desviación aumenta claramente hacia las zonas bajas de la distribución, particularmente en el extremo inferior, donde se observan valores considerablemente alejados de la recta. En el extremo superior también existe una separación, aunque de menor magnitud.

APLICACIÓN WEB

9 Aplicación Web

En este capítulo se observan los resultados obtenidos luego de desarrollar la plataforma Web Water Quality Mapping disponible públicamente en https://waterqualitygeomaticauanl.streamlit.app/. En esta sección, cada una de las interfaces que se pueden encontrar en la página Web, en donde cualquier usuario puede hacer uso de ella de manera gratuita para cada uno de los análisis que requiera.

9.1 Plataforma Web – Pantalla principal

La Figura 64 muestra la pantalla de inicio (Hero section) de la plataforma Web tal como se observa publicada en Streamlit Community Cloud. La interfaz presenta un diseño visual oscuro de alta calidad con fondo animado tipo “aurora boreal” y un globo terráqueo 3D giratorio, además de elementos que contextualizan visualmente el análisis de datos geoespaciales y satelitales. Se observan también los logos institucionales (UANL-FIC-Departamento de Geomática) en la barra superior, seguido del título principal “Water Quality Mapping” y la descripción de las aplicaciones que tiene la página Web. En la esquina superior derecha se presenta el indicador de estado de conexión con GEE, lo que avala el uso de la información geoespacial en esta plataforma y confirma que la autentificación con la API satelital esta activa.



Figura 68. Pantalla principal (Hero Section) de la plataforma Web Water Quality Mapping en funcionamiento

Fuente: Elaboración propia

Como se mencionó anteriormente, esta plataforma está disponible hasta la fecha y sin costo para cualquier usuario con acceso a internet, sin requerir instalación de software adicional. Además, el acceso multilingüe (español, inglés y portugués) se configura desde la barra lateral mediante un selector desplegable que adapta instantáneamente toda la interfaz al idioma elegido.



Figura 69. Hero section en portugués y en inglés

Fuente: Elaboración propia

9.2 Visualización – Mapa interactivo e Índices Espectrales

En la Figura 70 se muestra la sección de previsualización del área de estudio que tiene la página Web. Esta previsualización aparece luego de que el usuario o la persona que ingrese a la página adjunte su área de estudio, la cual tiene que ser un shapefile en formato. Zip. En la parte izquierda de la figura se observa la barra lateral de la aplicación con el flujo de trabajo guiado (4 pasos), la sección

de idioma con un menú desplegable con las diferentes opciones explicadas anteriormente, y la confirmación de carga del .Zip que en este caso es una sección del río Pesquería. El aviso en color verde identifica si la sección explica si es favorable generar un reporte de calidad de agua o si solo necesitas un reporte o análisis de índices espectrales.



Figura 70. Módulo de visualización del área de estudio página Web

Fuente: Elaboración propia

El mapa interactivo central muestra el área de estudio utilizada (río Pesquería) trazado con un color cyan sobre el fondo de mapa base, un fondo que puede variar según la preferencia de quien lo use ya se cómo imagen satelital o mapa base. Asimismo, para los datos de calidad de agua se observan los puntos en los que se tuvieron los muestreos de campo usados para entrenar este modelo de RF. Además, el panel de capas en la esquina superior derecha ofrece 11 opciones de visualización: RGB (Color natural), NDVI, NDWI, MNDWI, NDTI (Turbidez), NDCI (Clorofila Red Edge), SABI (Floraciones Algales), CDOM (Materia Orgánica Disuelta), AWEI (Extracción Agua), EVI (Vegetación Mejorado), Satélite de referencia, Mapa base, JRC Ocurrencia de Agua y ESA WorldCover 2021(Figura 71).



Figura 71. Visualización de área de estudio aplicando clasificación de uso de suelo y mapa base satelital

Fuente: Elaboración propia

En consecuencia, en la barra de estado inferior del mapa se muestra la información de la imagen Sentinel-2 seleccionada, la cantidad de imágenes encontradas, cobertura de nubes, rango temporal y la fecha de muestreo de campo en caso de ser necesaria. Esta previsualización permite obtener información rápida de que tan bien están las imágenes y los datos que se van a descargar, por lo que es importante tenerlos en cuenta para no descargar una imagen con mucha nubosidad o información sesgada por la falta de datos en el rango de las fechas utilizadas.

9.3 Animación temporal Sentinel-2 y serie temporal de índices espectrales

En la Figura 73 se observa la sección de análisis temporal de la página Web. Luego de seleccionar los índices espectrales a analizar o los parámetros de calidad de agua contemplados por el modelo se tiene la opción de descargarlos en formato .TIFF (solo índices espectrales) para posteriores análisis que se requieran hacer en software complementarios como Arcgis o Qgis. Asimismo, en la barra lateral que se observa en la figura se logra identificar el rango de búsqueda de las imágenes Sentinel-2 desde 2015 hasta la actualidad, el filtro de nubes configurado por porcentajes (%) y los parámetros fisicoquímicos activados para su análisis.

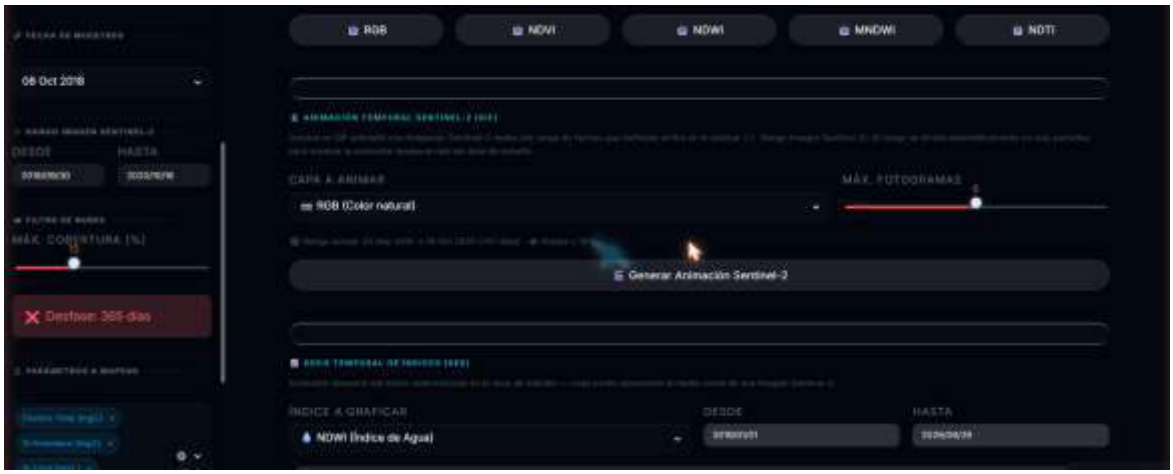


Figura 72. Sección de animación temporal Sentinel-2 (GIF) y serie temporal de índices espectrales

Fuente: Elaboración propia

Por consiguiente, la sección principal muestra la generación de animación temporal GIF, en donde el usuario puede seleccionar la capa para hacer una animación temporal en formato GIF y el número máximo de fotogramas. Debajo se presenta la sección de serie temporal de índices GEE, donde el usuario selecciona el índice a graficar y el rango temporal. Estos campos son de gran

importancia para poder obtener un análisis visual de los cambios de infraestructura, vegetación, ocurrencia de agua, uso de suelo según las capas que se deseen estudiar o analizar.

9.4 Resultados de serie temporal

Esta página Web no solo muestra visualmente los cambios temporales de las diferentes imágenes estudiadas, sino que también puede brindar un análisis de serie temporal de los diferentes índices espectrales que contiene la página Web para cualquier área de estudio en el planeta. Para esta investigación se observa una serie temporal del índice NDWI cubriendo el periodo de 2019-2026 con un total de 223 imágenes analizadas con una cobertura de nubes inferior al 15%. La grafica interactiva muestra la evolución del índice promedio zonal del área de estudio junto con la línea de tendencia calculada por regresión lineal. Sumado a un análisis de tendencia realizado con Mann- Kendall.



Figura 73. Serie temporal del NDWI generada automáticamente por la plataforma para el río Pesquería con 223 imágenes Sentinel-2 (enero 2019 – junio 2026)

Fuente: Elaboración propia

Los iconos que se encuentran en la parte superior derecha permiten hacer “zoom in” o “zoom out” del gráfico, descargar en formato PNG dicho gráfico y hacer un “full screen” para una mejor visualización. Además, el análisis muestra valores mínimos, máximos y promedios junto con los datos estadísticos de Mann-Kendall siendo para este caso una tendencia no significativa, pero claramente dependerá de cada análisis de estudio y de las fechas o índices que se necesiten contemplar. Lo más importante es que quien use esta aplicación Web puede previsualizar si su tendencia es significativa y luego puede descargar los datos en formato CSV tal como se observa en la parte inferior de la Figura 74.

9.5 Generador de reportes PDF

Esta sección presenta la generación de reportes PDF automáticos en una de las últimas interfaces que presenta la página Web realizada en esta investigación. Se observa claramente la distinción entre los tipos de reportes disponibles, en lo que se presenta a la izquierda el Reporte de Calidad del Agua (para estos casos el shapefile tiene que ser un water mask) disponible exclusivamente para el río Pesquería (al menos que el modelo se alimente con datos locales de Calidad del Agua) con el modelo RF validado. Este reporte se genera con un click en el botón “Generar Reporte de Calidad del Agua”.

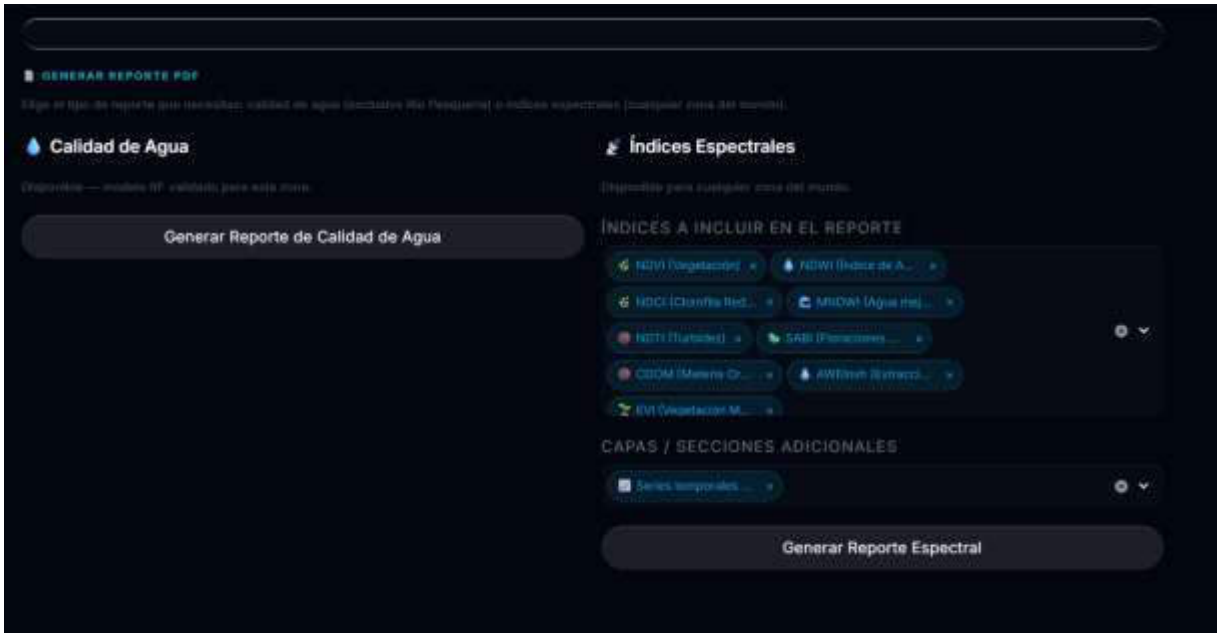


Figura 74. Generación de reportes PDF para Calidad de Agua y para Índices espectrales

Fuente: Elaboración propia

En la zona de la parte derecha de la Figura 74 se encuentra la generación del reporte de Índices espectrales, disponible para cualquier área geográfica del mundo, donde quien use la página Web puede seleccionar de manera personalizada los Índices espectrales a incluir en su reporte (NDVI, NDWI, NDCI, MNDWI, NDTI, SABI, CDOM, AWEI, EVI) y las secciones adicionales como series temporales. Ambos reportes se descargan directamente desde la interfaz en formato PDF.

9.6 Sección final de la página Web

Para la última sección de esta página Web se le da importancia al modelo aplicado en esta investigación, en donde los datos de Calidad de Agua en el río pesquería, se generó un reporte de serie temporal en PDF y los datos del autor principal (tesista Doctoral) y los coautores (directores de esta investigación). El reporte de serie temporal genera un documento PDF con la temporalidad de 2016-2019 aplicada en este estudio, con los parámetros que obtuvieron una precisión mayor 0.6, así mismo se incluyen gráficos de tendencia, tabla resumen estadístico e interpretación automática de los resultados y conclusiones. Para

este reporte igualmente se debe subir el .Zip de la máscara de agua implementada en esta investigación.

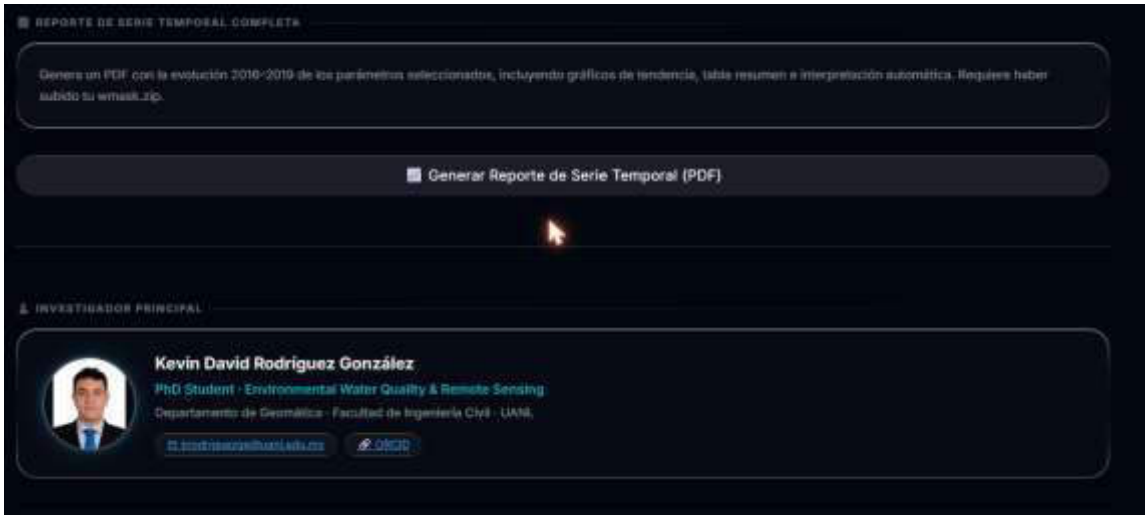


Figura 75. Interfaz Final

Fuente: Elaboración propia

En la Figura 75 se encuentra la interfaz del grupo de investigadores que hizo la página Web con su nombre completo (p.e. Kevin David Rodríguez González), la afiliación (PhD Student Environmental Water Quality & Remote Sensing, Departamento de Geomática · Facultad de Ingeniería Civil UANL), el correo institucional (krodriguezge@uanl.edu.mx) y el enlace al perfil ORCID.

En la Figura 77 se puede escanear el código QR donde se encuentra la página Web creada en esta investigación.



Figura 76. Código QR página web

Fuente: Elaborado en QR.io

10 LIMITACIONES Y TRABAJO A FUTURO

Aunque este estudio demuestra un rendimiento sólido, sigue teniendo limitaciones de flujo de trabajo como las siguientes:

Los eventos de bajo caudal ocurren de forma continua en los ecosistemas riparios con condiciones climáticas áridas, y aun con adquisiciones frecuentes de imágenes satelitales. Las rápidas fluctuaciones hidrológicas pueden mostrar altibajos en los pulsos de agua entre capturas de imágenes satelitales, lo que provoca inconsistencias al implementar colecciones de características (Feature Collection) en plataformas digitales (por ejemplo, GEE).

La heterogeneidad de las respuestas espectrales en ríos estrechos y de bajo caudal, junto con la velocidad y consistencia de los cambios impulsados por la urbanización, supone un reto para los clasificadores. Existen mezclas de subpíxeles debido a la proximidad de los elementos dentro de los ríos, como modificaciones en los canales, terraplenes artificiales, patrones de depósito de sedimentos, vegetación ribereña, sombras o incluso la Calidad del Agua.

Las condiciones de bajo caudal del río limitan la disponibilidad de datos de imágenes satelitales que representan la presencia de máscaras de agua a lo largo de los años. Además, las mediciones de caudal *in situ* dependen de los proyectos disponibles en la zona, que suelen ser escasos. Aunque Google Street View tiene algunas validaciones cuando la cobertura lo permite. Esta escasez de información fiable sobre el terreno también complica la adopción de arquitecturas de aprendizaje profundo, que normalmente requieren conjuntos de datos sustancialmente más grandes y diversos para lograr alta precisión y generalización robusta.

Este trabajo sigue evolucionando integrando imágenes de incrustación a corto plazo (p.e. ortomosaicos VANT), lo que podría ofrecer una mejor respuesta al desafío de clasificar las máscaras de agua, gracias a la mejora en la resolución y accesibilidad de la información. Además, este trabajo ya se complementa con información complementarioa proveniente de otros proyectos de investigación como el proyecto GIRRIO que se enfoca en las zonas ribereñas de los ríos

urbanos del AMM. Este tipo de información integra imágenes de alta resolución lo que proporciona mayor certeza sobre las inconsistencias entre el agua y otros elementos dentro de la planicie aluvial. Como trabajo futuro, se pretende incorporar IoT para obtener información más actualizada, con conectividad en tiempo real y automatización de procesos.

11 CONCLUSIONES

De acuerdo a lo presentado en este documento se logra concluir que el río Pesquería contempla una serie de complicaciones socioambientales que alteran las condiciones naturales del sitio y del entorno cercano. A nivel cuenca, el ICU presentó una configuración territorial reticular estructurada sobre la red hidrográfica del río Pesquería. Esta morfología obedece a la formulación multiplicativa del Índice, donde la zona de mayor convergencia se presentó en la zona Central-Norte de la cuenca, que coincide en la sección más conflictiva de problemáticas socioambientales del afluente.

La clasificación LULC se realizó para el área de interés, es decir la zona más afectada río Pesquería en el noreste de México. El objetivo fue identificar las máscaras de agua de los ríos urbanos a lo largo de cinco periodos, abarcando tanto la estación seca como la de lluvias. La principal fuente de datos para la implementación de modelos y los sitios de entrenamiento fueron las imágenes del satélite NICFI.

Este estudio aplicó con éxito tres algoritmos de aprendizaje automático para clasificar ríos de bajo caudal: RF, SVM y CART El algoritmo RF demostró la mayor precisión y fue el más consistente durante el periodo del estudio. RF alcanzó un 99% de precisión durante la temporada 18-19. RF mantuvo una precisión más consistente en las cinco estaciones

Los Índices espectrales convencionales NDWI, MNDWI y AWEI no fuerpn adecuados para obtener máscaras de agua en este tipo de río y, con imágenes de resolución media (interpretación visual), este estudio ilustra el potencial de las imágenes de alta resolución de NICFI para la monitorización efectiva de estos cuerpos de agua poco estudiados.

Uno de los principales logros del estudio es la máscara de agua para ríos urbanos estrechos y de bajo caudal derivada de las imágenes satelitales NICFI, que también son de libre acceso. La principal contribución es la metodología

resultante, que representa un paso importante en el monitoreo de ríos urbanos estrechos y de bajo caudal, aprovechando imágenes y plataformas innovadoras como GEE, y además validar los resultados es factible utilizando imágenes de Google Earth Pro, Sentinel-2 y Street View las cuales son también de acceso libre.

GeoAI permite la visualización del río como una representación digital, y el aspecto más importante es que esta información apoya la toma de decisiones, la planificación urbana y la gestión ambiental de las ciudades. Los resultados de esta investigación abordan los retos de monitorizar ríos con caudales intermitentes. Finalmente, se concluye que es necesario aumentar la concienciación entre la sociedad y los gobiernos locales y promover el monitoreo continuo de estos cuerpos de agua altamente vulnerables, que requieren esfuerzos específicos de conservación y recuperación.

Por consiguiente, los estudios de Calidad de Agua empleando imágenes satelitales en ríos de tan poco flujo como lo es el río Pesquería presentaron una gran limitación por parte de la respuesta espectral. En este estudio solo los parámetros como NTOT, PTOT y NKjeldahl presentaron buenas precisiones. El río Pesquería es un cauce con alta heterogeneidad por lo cual su monitoreo necesita ser más conciso y cada vez más actualizado. Por lo que se recomienda una mayor convergencia de fechas entre los datos de muestreo y los satelitales, además de usar imágenes satelitales de alta resolución para un mejor análisis. Las imágenes satelitales L1C no lograron obtener buenos resultados en esta área de estudio debido a las pocas imágenes que hay en el periodo estudiado y la alta nubosidad.

Los estudios que emplean imágenes satelitales para analizar ríos urbanos intermitentes, de escaso flujo y angostos, representan un desafío actual debido a la complejidad de estos cuerpos de agua y a su difícil acceso en entornos urbanos densamente construidos. Sin embargo, este proyecto tiene el potencial de generar conciencia y promover la necesidad de monitorear estos ríos, que son

de vital importancia para los ecosistemas urbanos. Se espera que esta concientización conduzca a acciones concretas destinadas a la recuperación de estos ecosistemas vulnerables y valiosos.

Por último, Water Quality Mapping es una aplicación Web científica desarrollada en esta investigación, que permite monitorear la Calidad del Agua del río Pesquería mediante teledetección satelital. Utiliza imágenes satelitales de Google Earth Engine procesadas con un modelo de Machine Learning (RF) para predecir parámetros fisicoquímicos como fósforo total, nitrógeno amoniacal y nitrógeno Kjeldahl,

Aunado a lo anterior, en esta página Web de acceso gratuito se pueden realizar análisis geoespaciales a nivel mundial logrando calcular nueve diferentes índices espectrales (NDVI, NDWI, MNDWI, NDCI,), clasificación de uso de suelo de le ESA World Cover junto con un índice de riesgo de contaminación MCDA. Cualquier persona (incluso sin formación geoespacial) que ingrese a la página puede hacer uso, análisis, descarga y obtención gratuita de todos esos procesamientos junto con un reporte en PDF que explica a detalle la información utilizada y un análisis de serie de tiempo. Su importancia radica en que transforma datos satelitales complejos en información accionable para la gestión ambiental, permitiendo detectar zonas de contaminación, evaluar tendencias temporales con pruebas estadísticas como Mann-Kendall, y verificar el cumplimiento con la norma NOM-001-SEMARNAT-1996. Al ser una herramienta Web accesible y gratuita, tiene el potencial de apoyar a autoridades ambientales, investigadores y tomadores de decisión en la protección de un recurso hídrico crítico para Nuevo León y para cualquier zona del planeta tierra, representando un avance significativo en el monitoreo ambiental inteligente a escala regional y global.

12 BIBLIOGRAFÍA

•Aldiansyah, S., & Saputra, R. A. (2022). Comparison of machine learning algorithms for land use and land cover analysis using Google Earth Engine (Case study: Wanggu Watershed). International Journal of Remote Sensing and Earth Sciences, 19(2). https://doi.org/10.30536/ijreses.v19i2.13716.

•Alsabti, K., Ranka, S., & Singh, V. (1998). CLOUDS: A decision tree classifier for large datasets. *Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98),* New York, USA, 2-8.

•Amado-Álvarez, J. P., Rubiños Panta, J. E., Gavi Reyes, F., y colaboradores (2019). Análisis multiespectral para la estimación de la turbidez como indicador de la calidad del agua en embalses del estado de Chihuahua, México. Revista Geográfica de América Central, núm. 62.

• Amani m, Ghorbanian A, Ahmadi S, Kakooei M, Moghimi A, Mirmazloumi M, Hamed S, Mahdavi S, Ghahremanloo M, Parsian S, Wu Q, Brisco N. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, VOL. 13, 2020.

•Amaro, K. A. N., Cortez, L. M. G., Medina, J. M., and Blanco, A. C.: MAPPING AND ASSESSMENT OF THE SPATIAL AND TEMPORAL DISTRIBUTION OF TURBIDITY IN LAKE BUHI FROM SENTINEL-2 IMAGES USING GEOGRAPHICALLY WEIGHTED REGRESSION AND NORMALIZED DIFFERENCE TURBIDITY INDEX, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-4/W9-2024, 29–36, https://doi.org/10.5194/isprs-archives-XLVIII-4-W9-2024-29-2024, 2024.

• Atlas de Peligros y Riesgo San Nicolas de los Garza 2023. Disponible el 20 de Noviembre de 2024 en: https://sn.gob.mx/wp-content/uploads/2024/01/ATLAS-SN-2023-1_compressed.pdf.

• Bauer M. Remote Sensing of Environment: History, Philosophy, Approach and Contributions, 1969 –2019. Remote Sensing of Environment Volume 237, February 2020, 111522.

• Berrio M, Mosquera T, Alzate V. Uso de drones para el análisis de imágenes multiespectrales en agricultura de precisión. @limentech ciencia y tecnología alimentaria. Volumen 13, No. 1, p.28-40, 2015.

• Blaine, B. E. (2018). Winsorizing. En B. Frey (Ed.), The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation (pp. 1817-1818). SAGE Publications.

• Bonansea M., Ledesma M., Rodríguez C., Pinotti C. Using new remote sensing satellites for assessing water quality in a reservoir. Hydrological sciences journal 2019, VOL. 64, NO. 1, 34–44.

•Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). *Classification and Regression Trees.* Belmont, Calif.: Wadsworth International Group.

•Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5-32 https://doi.org/10.1023/A:1010933404324.

• Briceño, I.; Pérez, W.; San Miguel, D.; Ramos, S. 2018. Determinación de calidad de agua en el lago Vichuquén, con imágenes de satélite Landsat 8, sensor OLI, año 2016, Chile. Revista de teledetección, 52: 67-78.

•Butler, D., Ward, S., Sweetapple, C., Astaraie-Imani, M., Diao, K., Farmani, R., Fu, G., 2016. Reliable, resilient and sustainable water management: the Safe & SuRe approach. Global Challenges 1 (1), 63–77.

• Carroll I. Imágenes Satelitales, la importancia de una nueva lectura. Universidad de los Andes, Facultad de Ingeniería. Revisado el 06 de noviembre de 2024

en: https://revistacontacto.uniandes.edu.co/contacto-25-cadena-aeroespacial/imagenes-satelitales-la-importancia-de-una-nueva-lectura/

• Castro-López, D., Guerra-Cobián, V., & Prat, N. (2019). The role of riparian vegetation in the evaluation of ecosystem health: The case of semiarid conditions in Northern Mexico. River Research and Applications, 35(1), 48–59. https://doi.org/10.1002/rra.3383.

• Chen B, Mu X, Chen P, Wang B, Choi J, Park H, Xu S, Wu Y, Yang H. Machine Learning-based inversion of water quality parameters in typical reach of the urban river by UAV multispectral data. Ecological Indicators 133 (2021) 108434.

•Chen P, Wang B, Wu Y, Wang Q, Huang Z, Wang C.Urban river water quality monitoring based on self-optimizing Machine Learning method using multi-source remote sensing data. Ecological Indicators 146 (2023) 109750.

• Chien C, Chang M, Syun M, Wang k. Flood Prevention and Emergency Response System Powered by Google Earth Engine. Remote sensing, 2018, 10(8), 1283.

• Choo, Y., Kang, G., Kim, D. et al. A study on the evaluation of water-bloom using image processing. Environ Sci Pollut Res 25, 36775–36780 (2018).

• De Cabo L & Marconi P. Estrategias de remediación para las cuencas de dos ríos urbanos de llanura Matanza-Riachuelo y reconquista. VAZQUEZ MAZZINI EDITORES. Buenos Aires -Argentina (2021).

• De Giglio M, Dubbini M, Cortesi I, Maraviglia M, Parisi E, Tucci G. Plastics waste identification in river ecosystems by multispectral proximal sensing: a preliminary methodology study, water enviromental Journal. Volume35, Issue 2. (2020) Pages 569-579.

• Dixon, W. J. (1960). Simplified Estimation from Censored Normal Samples. The Annals of Mathematical Statistics, 31(2), 385–391.

•Duan, H., Ma, R., Hu, C., 2012. Evaluation of remote sensing algorithms for cyanobacterial pigment retrievals during spring bloom formation in several lakes of east China. Remote Sens. Environ. 126, 126–135.

•Duan, W. K., T. B. He, P. Luo, D. Nover, and Y. Yamashiki. 2013. “Spatial and Temporal Trends in Estimates of Nutrient and Suspended Sediment Loads in the Ishikari River Japan.” Science of the Total Environment 461: 499–508 .2013.05.022.

• Duffy, J.P.; Cunliffe, A.M.; DeBell, L.; Sandbrook, C.; Wich, S.A.; Shutler, J.D.; Myers-Smith, I.H.; Varela, M.R.; Anderson, K. Location, location, location: considerations when using lightweight drones in challenging environments. Remote sensing in ecology and conservation, 4(1): 7-19 (2018). A Concept Paper on Smart River Monitoring System for Sustainability in River.

• Elhag M, Gitas I, Othman A, Bhrawi J, Gikas P. Assessment of Water Quality Parameters Using Temporal Remote Sensing Spectral Reflectance in Arid Environments, Saudi Arabia. Water 2019, 11, 556.

• Elijah O, Rahman T, Leow C, Yeen H, Sarijari M, Aris A, Salleh J, Chua T. International journal of integrated engineering vol. 10 No. 7 (2018) 130-139.

• Escobedo Tamez, A., Yépez-Rincón, F. D., Antonio Durán, Y. Z., Pérez Martínez, M., Rodríguez González, K. D., Treviño Carreón, J., Perez, D. C., and Abrego, C. J.: Analysis of socio-environmental problems in the La Silla River, Monterrey, México: An approach with UAV geospatial data (LiDAR and RGB), ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-3/W4-2025, 127–134, https://doi.org/10.5194/isprs-annals-X-3-W4-2025-127-2026, 2026.

• European Space Agency. "European Space Agency." https://www.esa.int/. Accedido el 16 de octubre de 2023.

• Feng, M.; Sexton, J.O.; Channan, S.; Townshend, J.R. A global, high-resolution (30-m) inland waterbody dataset for 2000: first results of atopographic-

spectral classification algorithm. International Journal of Digital Earth, 9(2) 2016, 113- 133.

• Fernández-Lozano, J.; Gutiérrez-Alonso, G. 2016. Aplicaciones geológicas de los drones. Revista de la sociedad geológicas de España, 29(1): 89-105.

• Ferriño, A. L. & Burgos, D. (2015). Delimitación de zonas federales y áreas de amortiguamiento en ríos afectados por el crecimiento urbano como estrategia para prevenir inundaciones. Caso de estudio: cuenca del río Pesquería (México). Epistemus, Ciencia Tecnología y Salud, (19).

• Fu G, Jin Y, Sun S, Yuan Z, Butler D. The role of deep Learning in urban water management: a Critical review. Water Research 223 (2022) 118973.

• Garzón Barrero, J., Sánchez Pineda, N. E., & Londoño Pinilla, D. F. (2023). Evaluación comparativa de los algoritmos de aprendizaje automático Support Vector Machine y Random Forest: efectos del tamaño del conjunto de entrenamiento. Ciencia e Ingeniería Neogranadina, 33(2), 131–148. https://doi.org/10.18359/rcin.6996.

• Gastezzi P, Alvarado V, Perez G. La importancia de los ríos como corredores interurbanos, Biocenosis. Vol. 31 (1-2) 2017.

• Gorelick etal., 2017. Google Earth Engi¬ne: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27. https:// doi.org/10.1016/j.rse.2017.06.031

• Gunning D, Stefik M, Choi J, Miller T, Stumpf F, Zhong G, Explainability is essential for users to effectively understand, trust, and manage powerful artificial intelligence applications. Since Robottics 4, eaay7120 (2019).

• Hamet P & Tremblay J. Artificial intelligence in Medicine. Metabolism Volume 69, Supplement, April 2017, Pages S36-S40.

• Herrera S, Los ríos urbanos en el ecosistema ciudad. Clavigero N°9 (2016).Jawalekar S.B, Shelare S.D, Development and performance analysis

of low cost combined harvester for rabicrops, Agricultural Engineering International:CIGR Journal 22 (1) (2020) 197–201.

• IWA, 2019. Digital Water: industry leaders chart the transformation journey. https://iw a-network.org/wp-content/uploads/2019/06/IWA_2019_Digital_Water_Report.pdf.

• Jalife S. Qué es la inteligencia artificial. Centro México Digital., Ciudad de México- México (2022). Disponible el 22/03/2023.

•Jasim, L. K., Hasan, R. H., & amp; Ibrahim, O. A. (2025). Evaluation of the accuracy of machine learning classifiers and spectral indices in land cover classification. Engineering, Technology & Applied Science Research, 15(3), 22548–22552. https://doi.org/10.48084/etasr.10406.

• K. Alkaradaghi, S.S. Ali, N. Al-Ansari, J. Laue, Land use classification and change detection using multi-temporal Landsat imagery in sulaimaniyah governorate, Iraq, in: H.M. El-Askary, S. Lee, E. Heggy, B. Pradhan (Eds.), Advances in Remote Sensing and Geo Informatics Applications, Springer International Publishing, Cham, 2019, pp. 117–120.

• Kharuf S, Hernández L, Orozco R, Aday O, Delgado I. Análisis de imágenes multiespectrales adquiridas con vehículos aéreos no tripulados, EAC vol.39 no.2. 2018.

• Kwong I, Wong F, Fung T. Automatic Mapping and monitoring of marine water quality parameters in Hong kong using sentinel-2 image time-series and Google Earth Engine Cloud Computing. Original Research, Volume 9, 2022.

• Lacaux, J.P. Tourre, Y.M., Vignolles, C., Ndione, J.A., Lafaye, M. (2007). Classification of ponds from high-spatial resolution remote sensing: Application to Rift Valley Fever epidemics in Senegal. Remote Sens. Environ.,106, 66–74 (in English).

• Landsat—Earth observation satellites. (2015-3081). (2016). Reston, VA: U.S. Geological Survey Retrieved from http://pubs.er.usgs.gov/publication/fs20153081.

•Lee, Su Jeong, Myoung-Hwan Ahn, and Sung-Rae Chung. 2017. "Atmospheric Profile Retrieval Algorithm for Next Generation Geostationary Satellite of Korea and Its Application to the Advanced Himawari Imager" Remote Sensing 9, no. 12: 1294.

• Li J & Liu Z. Multispectral Transforms Using Convolution Neural Networks for Remote Sensing Multispectral Image Compression. Remote Sens. 2019, 11(7), 759.

• Liu, C., Zhou, X., Zhou, Y., and Akbar, A.: Multi-temporal monitoring of urban river water quality using uav-borne multi-spectral remote sensing, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B3-2020, 1469–1475.

• Louhaichi, M., Borman, M.M., Johnson, D.E., 2001. Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. Geocarto International, Volume 16(1), pp. 65–70.

• Ma, Y., Wang, L., Liu, P., & Ranjan, R. (2015). Towards building a data-intensive index for big data compu¬ting - A case study of Remote Sensing data processing. Information Sciences, 319, 171–188.

•Martínez de León, N., González Gutierrez, I., Ramírez Campanur, X. C., Rocandio Rodríguez, M., & Medina Puente, A. (2026). Algoritmos de aprendizaje automático para el mapeo de coberturas en un Área Natural Protegida. Revista Mexicana de Ciencias Forestales, 17(94), 28–53. https://doi.org/10.29298/rmcf.v17i94.1580

• Martinez Y & Villalejo V. La gestión integrada de los recursos hídricos: una necesidad de estos tiempos. riha vol.39 no.1 La Habana ene.-abr. 2018.

• Machado J., Mantovani P., Gomes R., Márcia R., Demanboro A., Bettine S. Degradação ambiental e qualidade da água em nascentes de rios urbanos. Soc. nat. 30 (1) • Jan-Apr 2018.

• Melendez E. Mariposas diurnas (lepidoptera: papilionoidea) como indicadoras de la calidad ambiental en el área metropolitana de Monterrey, México. Tesis Doctoral. Universidad Autónoma de Nuevo León – Facultad de Ciencias Forestales (2020).

• Melicé J, Servain J. The tropical Atlantic meridional SST gradient index and its relationships with the SOI, NAO and Southern Ocean. Climate Dynamics 20, 447–464 (2003).

• Meyer, H., Reudenbach, C., Wöllauer, S., & Bendix, J. (2018). Importance of spatial predictor variable selection in machine learning applications of move persistence models. Ecological Modelling, 383, 1-13.

• Miller, M. E. (2022). Análisis de la Hidroconectividad del Continuo Fluvial-Marino del Bajo Río Doce (ES, Brasil) por Medio de Geotecnologías* (Tesis de doctorado en Oceanografía Ambiental). Universidade Federal do Espírito Santo, Vitória, Brasil.

•Mingers, J. (1989). An empirical comparison of selection measures for decision-tree induction. Machine Learning, 3(4), 319–342. DOI: 10.1007/BF00116837

• Montgomery DR, Buffington JM (1997) Channel-reach morphology in mountain drainage basins. Geological Society of America Bulletin 109(5): 596-611.

• Moore, R. (Directora de Google Earth, Earth Engine y Earth Outreach). (2017). Accompanying presentation Earth Engine Users' Summit 2017 (Videoconferencia). Recuperado de https://youtu.be/5yy1EwtZmhE

• Mora, A., Rosales-Lagarde, L., Hernández-Antonio, A., & Mahlknecht, J. (2017). Hydrogeochemistry of groundwater supplied to the city of Monterrey, Mexico.

Procedia Earth and Planetary Science, *17*. https://doi.org/10.1016/j.proeps.2016.12.090

• NASA. Landsat Science. Landsat. Revisado el 06 de Noviembre de 2024 en : https://landsat.gsfc.nasa.gov/satellites/landsat-1/

• Nguyen Q, Anh V, Hong N, Thu H. Modelling seawater Quality of Rach Gia Bay of Vietnam, using Sentinel-2 imagery processed in the Google Earth Engine.

• NICFI. Norway´s International Climate and Forest Initiative. Standing with the World´s tropical forests. Disponible en: https://www.nicfi.no/. 16 /10. 2022.

• Osco L., Marcato J., Marques A., De Castro L., Fatholahi S., Silva J., Matsubara E., Pistori H., Gonçalves W., Li J. ıinternational Journal of Applied Earth Observations and Geoinformation. 102 (2021) 102456.

• Ouma, B. Nkwae, D. Moalafhi, P. Odirile, B. Parida, G. Anderson, J. Qi, Comparison of machine learning classifiers for MULTITEMPORAL and multisensor mapping of urban LULC features, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLIII-B3-2022 (2022) 681–689,

• Pandey P, Kington J, Kanwar A, Simmon R, Abraham L. PLANET basemaps for nicfi data program addendum to basemaps product specification. Norwgian Ministry of climate and environment. V2. 2023

• Pangihutan m, Barita J, Frenly A, Duva G, Handoco E, Basten W. TSS Analysis (Total suspended soil) usin GEE (Google Earth Engine) Cloud Technology in Belawan waters. Journal of Applied Geospacial Information Vol6. N°2 2022.

•Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26(1), 217–222. https://doi.org/10.1080/01431160412331269698

• Parmar A, Rakesh Katariya, Vatsal Patel. International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018, 2019, Volume 26 ISBN: 978-3-030-03145-9

• Peterson K, Sagan B, Sloan J. Deep Learning based water quality estimation and anomaly detection using Landsat 8 Sentinel 2 virtual constellation and cloud computing. Taylor & Francis. 2020, VOL. 57, NO. 4, 510–525.

• Perilla & Francois Google Earth Engine (GEE): una poderosa herramienta que vincula el potencial de los datos masivos y la eficacia del procesamiento en la nube. Instituto de Geografía en la UNAM., Núm. 101. (2020)

• Pino, V.E. Los drones una herramienta para una agricultura eficiente: un futuro de alta tecnología. IDESIA, 37(1): (2019) 75-84. Quintero D. Elementos de recuperación de ríos urbanos. Caso de estudio: Río Torca (Tramo I). Contrato para el Desarrollo de Actividades Científicas y Tecnológicas. Fondo editorial: secretaria Distrital de Medio Ambiente, 2011.

• Quesada A., Castro J., Feoli S. Geomorphology, land use, and environmental impacts in a densely populated urban catchment of Costa Rica. Journal of South American Earth Sciences.Volume 112, Parte 1, 103560.

•Rahman, M. M., Nuhash, M. M. R., Ray, S., & Reza, M. S. (2026). Evaluation of machine learning algorithms in land use/land cover change detection using remote sensing and GIS for Moulvibazar, Bangladesh. Turkish Journal of Engineering, 10(3), 964–976. https://doi.org/10.31127/tuje.1823517.

• Ramos B y Peláez D. La teledetección y los sistemas de información geográfica para el manejo de tierras. Rev Cie Téc Agr vol.27 no.1 San José de las Lajas ene.-mar. 2018.

•Rawat, K. S., Kumar, S., & Garg, N. (2024). Statistical comparison of simple and machine learning based land use and land cover classification algorithms: A case study. Journal of Water Management Modeling, 32, Art. H524. https://doi.org/10.14796/JWMM.H524.

•Rico-Serrano, L. A. (2026). Determinación de sólidos suspendidos totales en cuerpos de agua por medio de imágenes satelitales y su correlación con datos in situ. Boletín de Ciencias de la Tierra, (59), 85-95

• Rinaldi, M., Gurnell, A. M., del Tánago, M. G., Bussettini, M., & Hendriks, D. (2016). Classification of river morphology and hydrology to support management and restoration. Aquatic Sciences, 78(1), 17–33. https://doi.org/10.1007/s00027-015-0438-z.

• Roberts, D. R., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913–929.

•Rodríguez González, K. D., Arista Cázares, L. E., and Yépez Rincón, F. D.: Spatiotemporal land use land cover (LULC) change analysis of urban narrow river using Google Earth Engine and Machine learning algorithms in Monterrey, Mexico, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-3-2024, 371–375, https://doi.org/10.5194/isprs-annals-X-3-2024-371-2024, 2024.

•Rodríguez González, K. D., Vega-Aguilar, A., Yépez-Rincón, F. D., Treviño Carreón, J., Velastegui-Montoya, A., Freire-Granda, D., and Chuizaca Espinoza, I. A.: High-resolution LiDAR and thermal UAV data for 3D analysis of urban vegetation structure and its cooling effect in San Nicolás, Mexico, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIX-B3-2026, 421–429, https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-421-2026, 2026.

•Rodríguez González, K. D., Yépez-Rincón, F. D., Escobedo Tamez, A., and Vega-Aguilar, A.: UAV Remote Sensing and Deep Learning for Automatic Car Detection and Parking Occupancy Analysis: The Case of UANL Stadium, Mexico, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-3/W3-2025, 85–92, https://doi.org/10.5194/isprs-annals-X-3-W3-2025-85-2026, 2026.

• Shelare S, Aglawe K, Waghmare S, Belkhode P. Advances in water sample collections with a drone – A review. Materials Today: Proceedings Volume 47, Part 14, 2021, Pages 4490-4494.

• Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi and S. Homayouni, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 6308-6325, 2020, doi: 10.1109/JSTARS.2020.3026724.

•Sun X, Zhang Y, Shi K, Zhang Y, Li N, Wang W, Huang X, Qin B. Monitoring water quality using proximal remote sensing technology. Science of the Total Environment 803(2022) 149805.

• Suzianti N., Kamaludin N., Abdullah M., Awang D., Ahmad M., Abdul N., Baizura T. Water Science and Engineering. Article in press. Disponible en SinceDirect.

• Tukey, J. W. (1962). The Future of Data Analysis. The Annals of Mathematical Statistics, 33(1), 1–67.

• United Nations. Data drive un-water integrated monitoring initiative for SDG 6. SDG indicator 6.3.2 (2020). Disponible el 25 de noviembre 2021 en: SDG 6.3.2 Home - SDG 6.3.2 - Sitio global (unep.org).

• Vega Aguilar, A., Yépez-Rincón, F. D., Rodríguez Gonzalez, K. D., Vázquez Rodríguez, D. A., Ferriño Fierro, A. L., and Guerra Cobián, V. H.: UAV for flooding risk assessment in a new Dam Construction in Northeast Mexico, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-2/W2-2025, 189–197, https://doi.org/10.5194/isprs-annals-X-2-W2-2025-189-2025, 2025.

• Veneros, J, García, L, Morales, E, Gómez V, Torres M, & López F. Aplicación de sensores remotos para el análisis de cobertura vegetal y cuerpos de agua. Idesia (Arica), 38(4), (2020) 99-107.

• Vizzari M. PlanetScope, Sentinel-2, and Sentinel-1 Data Integration for Object-Based Land Cover Classification in Google Earth Engine. *Remote Sens* 2022, 14(11), 2628.

• Wan L & Wang H. Control of urban river water pollution is studied based on SMS. Evironmental technology & innovation 22 (2021) 101468.

• Weier, J. and Herring, D. (2000) Measuring Vegetation (NDVI & EVI). NASA Earth Observatory, Washington DC

• Xia C., Liu G., Meng Y., Wng Z., Zhang X. Impact of human activities on urban river system and its implication for water-environment risks: an isotope-based investigation in Chengdu, China. Human and Ecological Risk Assessment: An International Journal. Volume 27, 2021 - Issue 5.

• Xu, H. "Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery." International Journal of Remote Sensing 27, No. 14 (2006): 3025-3033." (ESRI, 2018).

• Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2017). Big Data and cloud computing: innovation opportu¬nities and challenges. International Journal of Digital Earth, 10(1), 13–53.

• Yang, W., John, V. O., Zhao, X., Lu, H., & Knapp, K. R. (2016). Satellite climate data records: Development, applications, and societal benefits. Remote Sensing, 8(4), 1–24. https://doi.org/10.3390/rs8040331.

• Yang L, Li J, Zhou K, Feng P, Dong L (2021) The effects of surface pollution on urban river water quality under rainfall events in Wuqing district, Tianjin, China. Journal of Cleaner Production 293:126136.

•Yépez Rincón, F. D., & Aguilar Cruz, R. (2024). El río urbano como infraestructura resiliente ante el cambio climático. La Noria Digital, (18), 24–30.

• Zha, Y., Gao, J., and Ni, S., “Use of normalized difference builtup index in automatically mapping urban areas from TM imagery.” International Journal of Remote Sensing, Vol. 24, No. 3, (2003), 583–594.

• Zhao Q, Yu L, Li X, Peng D, Zhang Y, Gong P. Progress and Trends in the Application of Google Earth and Google Earth Engine. Remote Sensing, 2021, 13(18), 3778 (2021).

• Zhu M, Wang J, Yang X, Zhang Y, Zhang L, Ren H, Wu B, Ye L. A review of application of Machine Learning in water quality evaluation. Eco-Environment & Health 1 (2022) 107–116

• Zingaro, M., Refice, A., Giachetta, E., D’Addabbo, A., Lovergine, F., De Pasquale, V., Pepe, G., Brandolini, P., Cevasco, A., & Capolongo, D. (2019). Sediment mobility and connectivity in a catchment: A new mapping approach. Science of the Total Environment, 672, 763–775.

13 APENDICES

Con el fin de proporcionar una comprensión completa y respaldar la información presentada en este documento, se incluyen varios anexos que complementan y enriquecen los temas abordados. Estos anexos están diseñados para ofrecer detalles adicionales, datos específicos, o material de referencia que puede resultar útil para el lector interesado en profundizar en aspectos particulares del tema. A continuación, se presenta una breve descripción de cada apéndice.

APENDICE A

(Imagen Infrarroja)





APENDICE B

(ORTOMOSAICO SECCIÓN)



APENDICE C

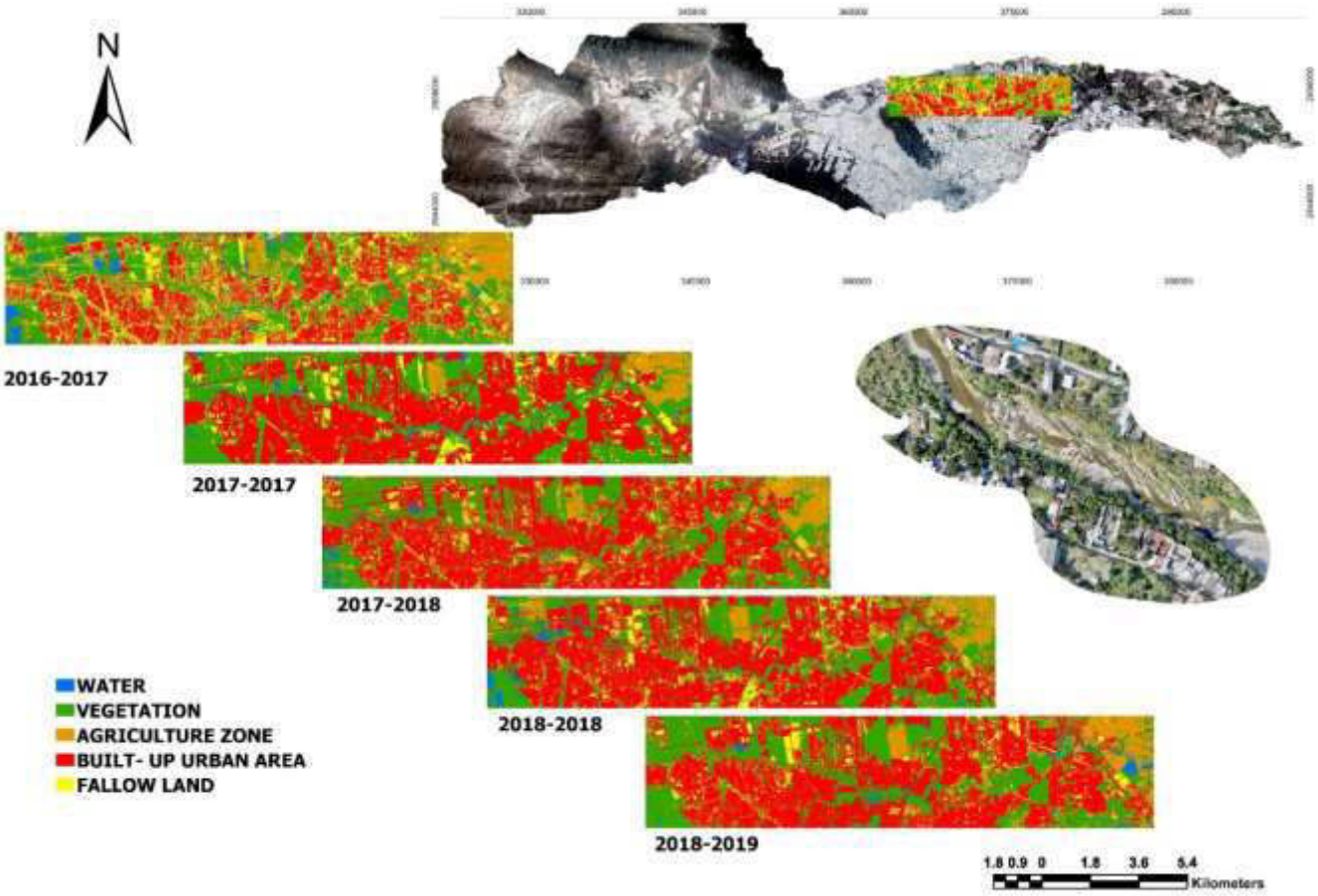
(OPEN BUILDINGS)





APENDICE D

(LULC)


N
2016-2017
2017-2017
2017-2018
2018-2018
2018-2019
WATER
VEGETATION
AGRICULTURE ZONE
BUILT- UP URBAN AREA
FALLOW LAND
1.8 0.9 0 1.8 3.6 5.4
Kilometers


Observado vs Predicho (CV) - RF v3 | Río Pesquería

APENDICE E
TABLAS

Precisiones de cada modelo de clasificación LULC

Classifier	Season	Class	Support	PA	UA	F1
RF	16_17	Class 1	15	0.933333	1	0.965517
RF	16_17	Class 2	24	0.916667	0.956522	0.93617
RF	16_17	Class 3	20	1	0.952381	0.97561
RF	16_17	Class 4	25	1	0.892857	0.943396
RF	16_17	Class 5	16	0.8125	0.928571	0.866667
RF	16_17	Macro average	100	0.9325	0.946066	0.937472
RF	17_17	Class 1	15	1	1	1
RF	17_17	Class 2	24	0.958333	1	0.978723
RF	17_17	Class 3	20	1	1	1
RF	17_17	Class 4	25	1	0.961538	0.980392
RF	17_17	Class 5	16	0.9375	0.9375	0.9375
RF	17_17	Macro average	100	0.979167	0.979808	0.979323
RF	17_18	Class 1	15	0.933333	0.933333	0.933333
RF	17_18	Class 2	21	0.952381	0.909091	0.930233
RF	17_18	Class 3	20	0.95	1	0.974359
RF	17_18	Class 4	27	1	0.964286	0.981818
RF	17_18	Class 5	17	0.941176	1	0.969697
RF	17_18	Macro average	100	0.955378	0.961342	0.957888
RF	18_18	Class 1	16	0.9375	0.882353	0.909091
RF	18_18	Class 2	22	0.954545	0.913043	0.933333
RF	18_18	Class 3	17	0.941176	1	0.969697
RF	18_18	Class 4	28	0.928571	0.928571	0.928571
RF	18_18	Class 5	17	0.882353	0.9375	0.909091
RF	18_18	Macro average	100	0.928829	0.932294	0.929957
RF	18_19	Class 1	17	1	1	1
RF	18_19	Class 2	20	1	1	1
RF	18_19	Class 3	16	1	1	1
RF	18_19	Class 4	31	0.967742	1	0.983607
RF	18_19	Class 5	16	1	0.941176	0.969697
RF	18_19	Macro average	100	0.993548	0.988235	0.990661
CART	16_17	Class 1	15	0.866667	0.722222	0.787879
CART	16_17	Class 2	24	0.75	0.9	0.818182

CART	16_17	Class 3	20	1	0.909091	0.952381
CART	16_17	Class 4	25	1	0.925926	0.961538
CART	16_17	Class 5	16	0.6875	0.846154	0.758621
CART	16_17	Macro average	100	0.860833	0.860679	0.85572
CART	17_17	Class 1	15	0.8	0.705882	0.75
CART	17_17	Class 2	24	0.791667	0.791667	0.791667
CART	17_17	Class 3	20	1	1	1
CART	17_17	Class 4	25	0.96	0.96	0.96
CART	17_17	Class 5	16	0.625	0.714286	0.666667
CART	17_17	Macro average	100	0.835333	0.834367	0.833667
CART	17_18	Class 1	15	0.933333	0.933333	0.933333
CART	17_18	Class 2	21	0.857143	0.947368	0.9
CART	17_18	Class 3	20	1	0.909091	0.952381
CART	17_18	Class 4	27	1	0.931034	0.964286
CART	17_18	Class 5	17	0.882353	1	0.9375
CART	17_18	Macro average	100	0.934566	0.944165	0.9375
CART	18_18	Class 1	16	0.9375	0.9375	0.9375
CART	18_18	Class 2	22	0.954545	0.875	0.913043
CART	18_18	Class 3	17	0.882353	1	0.9375
CART	18_18	Class 4	28	0.964286	0.931034	0.947368
CART	18_18	Class 5	17	0.882353	0.9375	0.909091
CART	18_18	Macro average	100	0.924207	0.936207	0.928901
CART	18_19	Class 1	17	0.941176	0.941176	0.941176
CART	18_19	Class 2	20	0.95	0.791667	0.863636
CART	18_19	Class 3	16	0.875	0.823529	0.848485
CART	18_19	Class 4	31	0.935484	1	0.966667
CART	18_19	Class 5	16	0.8125	1	0.896552
CART	18_19	Macro average	100	0.902832	0.911275	0.903303
SVM	16_17	Class 1	15	0.866667	0.764706	0.8125
SVM	16_17	Class 2	24	0.958333	0.884615	0.92
SVM	16_17	Class 3	20	1	0.952381	0.97561
SVM	16_17	Class 4	25	0.92	0.958333	0.938776
SVM	16_17	Class 5	16	0.6875	0.916667	0.785714
SVM	16_17	Macro average	100	0.8865	0.89534	0.88652
SVM	17_17	Class 1	15	0.933333	0.666667	0.777778
SVM	17_17	Class 2	24	0.75	1	0.857143
SVM	17_17	Class 3	20	1	1	1
SVM	17_17	Class 4	25	0.96	0.888889	0.923077

SVM	17_17	Class 5	16	0.625	0.714286	0.666667
SVM	17_17	Macro average	100	0.853667	0.853968	0.844933
SVM	17_18	Class 1	15	0.8	0.857143	0.827586
SVM	17_18	Class 2	21	0.904762	0.904762	0.904762
SVM	17_18	Class 3	20	1	1	1
SVM	17_18	Class 4	27	0.925926	0.961538	0.943396
SVM	17_18	Class 5	17	0.941176	0.842105	0.888889
SVM	17_18	Macro average	100	0.914373	0.91311	0.912927
SVM	18_18	Class 1	16	0.875	0.823529	0.848485
SVM	18_18	Class 2	22	0.909091	0.833333	0.869565
SVM	18_18	Class 3	17	0.941176	1	0.969697
SVM	18_18	Class 4	28	0.964286	1	0.981818
SVM	18_18	Class 5	17	0.882353	0.9375	0.909091
SVM	18_18	Macro average	100	0.914381	0.918873	0.915731
SVM	18_19	Class 1	17	0.823529	0.933333	0.875
SVM	18_19	Class 2	20	0.95	0.826087	0.883721
SVM	18_19	Class 3	16	0.9375	1	0.967742
SVM	18_19	Class 4	31	0.967742	1	0.983607
SVM	18_19	Class 5	16	1	0.941176	0.969697
SVM	18_19	Macro average	100	0.935754	0.940119	0.935953

.

Datos Campo Calidad de Agua

N_NH3	N_NO2	N_NO3	N_ORG	N_TOT	N_TOTK	n_images	nombre	range_end
0.133896	0.098527	12.085357	1.45656	13.77434	1.590456	2	Punto_1	3/27/2016
0.256978	0.173554	9.055294	0.949352	10.435178	1.20633	2	Punto_1	5/13/2016
0.888774	0.198768	4.103905	1.118552	6.309999	2.007326	4	Punto_1	6/17/2016
0.123326	0.086101	2.998253	1.740189	4.947869	1.863515	2	Punto_1	7/24/2016
0.216185	0.150218	4.92391	2.108327	7.39864	2.324512	2	Punto_1	8/26/2016
0.14086	0.232997	6.843417	0.878977	8.096251	1.019837	2	Punto_1	10/5/2016
3.39953	0.60629	3.56664	0.74838	8.32084	4.14791	3	Punto_1	3/25/2017
0.48537	0.19026	7.348	0.85719	8.88082	1.34256	3	Punto_1	5/5/2017
0.15085	0.16031	6.07035	0.9147	7.29621	1.06555	2	Punto_1	6/16/2017
1.00143	0.35631	2.84107	1.46841	5.66722	2.46984	2	Punto_1	7/28/2017
0.695851	0.162819	3.48283	1.077743	5.419243	1.773594	1	Punto_1	9/8/2017
0.91	0.52822	4.91822	0.67626	7.0327	1.58626	4	Punto_1	10/19/2017
0.944684	0.64234	10.824501	0.73616	13.147685	1.680844	5	Punto_1	3/11/2018
0.772011	0.552796	5.535266	0	6.860073	0.772011	5	Punto_1	4/13/2018
0.398571	0.358772	7.175577	0.84161	8.77453	1.240181	4	Punto_1	5/27/2018
2.051996	0.7567	6.7114	0.53867	16.7	2.590666	8	Punto_1	7/9/2018
3.664455	0.768876	4.67184	1.39154	10.496711	5.055995	3	Punto_1	11/8/2018
6.171222	0.808721	5.276764	0.32917	12.585877	6.500392	6	Punto_1	12/13/2018
9.062008	0.207039	2.40375	0	11.672797	9.062008	4	Punto_1	2/14/2019
0.061787	0.049126	10.952402	0.621923	11.685238	0.68371	2	Punto_2	3/27/2016
0.168234	0.071222	8.056467	0.895375	9.191298	1.063609	2	Punto_2	5/13/2016
0.948572	0.184817	3.81023	1.105956	6.049575	2.054528	4	Punto_2	6/17/2016
0.120119	0.083802	2.994051	1.872156	5.070128	1.992275	2	Punto_2	7/24/2016
0.127784	0.103917	5.501363	1.1993	6.932364	1.327084	2	Punto_2	10/5/2016
0.72282	0.40623	4.72282	1.35326	7.20513	2.07608	3	Punto_2	3/25/2017
0.15467	0.05397	6.61939	0.78821	7.61624	0.94288	3	Punto_2	5/5/2017
0.10123	0.13892	6.85791	0.83024	7.9283	0.93147	2	Punto_2	6/16/2017
1.22264	0.24176	2.70273	1.4834	5.65053	2.70604	2	Punto_2	7/28/2017
0.754261	0.117181	3.01984	1.053111	4.944393	1.807372	1	Punto_2	9/8/2017
0.11965	0.09627	5.43767	0.92546	6.57905	1.04511	4	Punto_2	10/19/2017
0.071094	0.102048	8.963197	0.95428	10.090619	1.025374	5	Punto_2	3/11/2018
0.061228	0.045523	5.419332	0.98384	6.509923	1.045068	5	Punto_2	4/13/2018
0.094463	0.075861	4.102574	0.73058	5.003478	0.825043	4	Punto_2	5/27/2018
0.197819	0.004	4.4136	1.00847	5.6	1.2	8	Punto_2	7/9/2018
0.118227	0.465243	7.771686	1.14211	9.497266	1.260337	3	Punto_2	11/8/2018
0.242949	0.789775	9.157088	2.09054	12.280352	2.333489	6	Punto_2	12/13/2018
2.736161	0.986876	6.275272	0	9.998309	2.736161	4	Punto_2	2/14/2019
33.822792	0.332239	2.756518	0	36.911549	33.822792	2	Punto_3	3/27/2016

21.106628	0.065095	0.188468	6.784106	28.144297	27.890734	2	Punto_3	5/13/2016
0.906302	0.188474	3.639453	1.277882	6.012111	2.184184	4	Punto_3	6/17/2016
0.112017	0.091011	3.742595	1.765071	5.710694	1.877088	2	Punto_3	7/24/2016
12.115935	1.035021	2.048602	2.612333	17.811891	14.728268	2	Punto_3	8/26/2016
11.762973	0.227063	6.724592	0.644497	19.359125	12.40747	2	Punto_3	10/5/2016
11.19543	3.4161	3.11592	7.05136	24.77881	18.24679	3	Punto_3	3/25/2017
18.33103	3.15498	2.82137	10.33882	34.6462	28.66985	3	Punto_3	5/5/2017
3.44682	0.36367	16.38479	1.61786	21.81314	5.06468	2	Punto_3	6/16/2017
7.98699	1.25613	9.45815	3.14036	21.84163	11.12735	2	Punto_3	7/28/2017
5.001161	0.181926	10.424929	3.211798	18.819814	8.212959	1	Punto_3	9/8/2017
28.04398	0.02495	0.3277	5.27636	33.67299	33.32034	4	Punto_3	10/19/2017
9.003569	0.183853	10.536886	1.54898	21.273288	10.552549	5	Punto_3	3/11/2018
9.744357	0.401679	6.948036	9.9171	27.011172	19.661457	5	Punto_3	4/13/2018
19.330729	0.567854	5.809398	0	25.707981	19.330729	4	Punto_3	5/27/2018
31.665638	0.004	0.0537	7.89142	39	39	8	Punto_3	7/9/2018
25.579832	0.348746	5.809213	0	31.737791	25.579832	3	Punto_3	11/8/2018
34.918854	1.119425	2.868285	0.62711	39.533674	35.545964	6	Punto_3	12/13/2018
27.288797	0.774076	4.441631	0	32.504504	27.288797	4	Punto_3	2/14/2019
29.457109	0.502059	4.080553	1.548567	35.588288	31.005676	2	Punto_4	3/27/2016
11.436772	0.328579	1.948792	7.739169	21.453312	19.175941	2	Punto_4	5/13/2016
22.645325	0.007742	0.033297	3.740265	26.426629	26.38559	4	Punto_4	6/17/2016
0.164782	0.08513	3.319942	2.090625	5.660479	2.255407	2	Punto_4	7/24/2016
14.931744	0.136449	0.569382	0.973957	16.611532	15.905701	2	Punto_4	8/26/2016
7.084556	2.853639	7.098924	0	17.037119	7.084556	2	Punto_4	10/5/2016
8.41787	0.78734	3.66339	1.57115	14.43975	9.98902	3	Punto_4	3/25/2017
13.83431	0.01794	0.78435	1.99194	16.62854	15.82625	3	Punto_4	5/5/2017
1.50024	1.8198	11.47205	1.51841	16.3105	3.01865	2	Punto_4	6/16/2017
7.83237	0.65286	7.81105	0.0003	16.29658	7.83267	2	Punto_4	7/28/2017
10.50805	0.942357	5.850658	0	17.301065	10.50805	1	Punto_4	9/8/2017
17.87655	0.34267	1.29709	0	19.51631	17.87655	4	Punto_4	10/19/2017
6.746825	0.49299	4.979726	1.12468	13.344221	7.871505	5	Punto_4	3/11/2018
13.7227	0.061352	0.305581	0	14.089633	13.7227	5	Punto_4	4/13/2018
15.390627	0.842682	2.591877	6.74681	25.571996	22.137437	4	Punto_4	5/27/2018
26.565269	0.004	1.977	0	28	26	8	Punto_4	7/9/2018
27.072903	0.7318	4.028619	0	32.131091	27.072903	3	Punto_4	11/8/2018
27.682676	0.421523	1.720422	0.55456	30.379181	28.237236	6	Punto_4	12/13/2018
25.456268	0.66825	4.52956	0	30.654078	25.456268	4	Punto_4	2/14/2019
28.026876	1.314005	0.163843	6.467761	35.972485	34.494637	2	Punto_5	3/27/2016
14.862432	0.234132	3.602477	2.272923	20.971964	17.135355	2	Punto_5	5/13/2016
17.49527	0.146223	1.089237	0	18.73073	17.49527	4	Punto_5	6/17/2016
0.19897	0.08864	2.966255	2.035227	5.289092	2.234197	2	Punto_5	7/24/2016
14.812707	0.143649	0.691685	0	15.648041	14.812707	2	Punto_5	8/26/2016
6.516456	2.86385	6.462073	0	15.842379	6.516456	2	Punto_5	10/5/2016

6.23719	0.6559	3.135	1.31815	11.34624	7.55534	3	Punto_5	3/25/2017
14.24505	0.12554	2.12672	2.63938	19.13669	16.88443	3	Punto_5	5/5/2017
1.54379	1.61935	10.55741	0	13.72055	1.54379	2	Punto_5	6/16/2017
8.02846	0.64702	7.73176	0	16.40724	8.02846	2	Punto_5	7/28/2017
10.173388	0.962283	5.76855	0	16.904221	10.173388	1	Punto_5	9/8/2017
0.19135	0.16527	1.55421	14.96537	16.8762	15.15672	4	Punto_5	10/19/2017
8.631162	0.441545	6.630691	0	15.703398	8.631162	5	Punto_5	3/11/2018
11.604465	0.045613	0.317231	2.90545	14.872759	14.509915	5	Punto_5	4/13/2018
13.311944	0.676751	4.010695	0.51107	18.51046	13.823014	4	Punto_5	5/27/2018
25.12188	0.004	2.3555	0	27	25	8	Punto_5	7/9/2018
26.254547	1.015246	3.855048	0	31.124841	26.254547	3	Punto_5	11/8/2018
24.194824	0.036104	0.097868	0	24.328796	24.194824	6	Punto_5	12/13/2018
20.013908	0.670299	5.660401	0	26.344608	20.013908	4	Punto_5	2/14/2019
19.997015	1.297839	2.604687	0.02	23	20	2	Punto_6	3/27/2016
12.823109	0.833293	1.463541	1.553491	16.673434	14.3766	2	Punto_6	5/13/2016
9.968126	2.635171	2.659753	0.448252	15.711302	10.416378	4	Punto_6	6/17/2016
15.931958	0.965731	1.644522	1.178067	19.720278	17.110025	2	Punto_6	7/24/2016
4.587569	2.302267	4.555181	0.834661	12.279678	5.42223	2	Punto_6	8/26/2016
4.319666	2.500402	6.797928	1.531306	15.149302	5.850972	2	Punto_6	10/5/2016
8.62469	2.59189	7.17197	0.58988	18.97843	9.21457	3	Punto_6	3/25/2017
9.855	0.50732	4.79187	2.6082	17.76239	12.4632	3	Punto_6	5/5/2017
2.31913	0.53997	6.87615	1.49493	11.23018	3.81406	2	Punto_6	6/16/2017
9.583384	0.515076	6.085307	0	16.183767	9.583384	2	Punto_6	7/28/2017
4.838984	0.99535	7.068262	0	12.902596	4.838984	1	Punto_6	9/8/2017
12.21145	0.42605	1.81807	0	14.45557	12.21145	4	Punto_6	10/19/2017
7.73362	1.07662	5.54031	1.66855	16.0191	9.40217	5	Punto_6	3/11/2018
8.160942	1.079561	8.779614	1.58969	19.609807	9.750632	5	Punto_6	4/13/2018
5.061073	1.323424	6.792492	5.87459	19.051579	10.935663	4	Punto_6	5/27/2018
1.090153	2.440548	7.132111	0.00873	10.671542	1.098883	8	Punto_6	7/9/2018
6.473399	0.912149	6.489912	0.8185	14.69396	7.291899	3	Punto_6	11/8/2018
10.57575	2.020838	8.237898	0	20.834486	10.57575	6	Punto_6	12/13/2018
11.621893	1.145836	9.239569	0.81002	22.817318	12.431913	4	Punto_6	2/14/2019
14.849945	0.82093	4.825092	2.412151	22.908118	17.262096	2	Punto_7	3/27/2016
16.912712	0.568569	1.966688	1.196588	20.644557	18.1093	2	Punto_7	5/13/2016
9.490211	2.669083	2.652653	1.222863	16.03481	10.713074	4	Punto_7	6/17/2016
15.037431	1.008143	1.568934	1.319472	18.93398	16.356903	2	Punto_7	7/24/2016
6.297402	1.500846	3.588249	3.429294	14.815791	9.726696	2	Punto_7	8/26/2016
6.064168	2.448197	5.482851	1.316983	15.312199	7.381151	2	Punto_7	10/5/2016
7.5083	1.82569	8.12196	1.66112	19.11707	9.16942	3	Punto_7	3/25/2017
10.27019	0.96859	6.7081	1.42492	19.3718	11.69511	3	Punto_7	5/5/2017
2.34038	0.53587	11.29299	1.69275	15.86199	4.03313	2	Punto_7	6/16/2017
9.464705	0.495049	7.529853	0	17.489607	9.464705	2	Punto_7	7/28/2017
5.173765	0.98079	6.837186	0	12.991741	5.173765	1	Punto_7	9/8/2017

13.66911	0.41285	1.635	0.11518	15.83214	13.78429	4	Punto_7	10/19/2017
10.65776	0.77923	5.16277	1.99533	18.59509	12.65309	5	Punto_7	3/11/2018
7.601915	1.054627	7.925232	2.54933	19.131104	10.151245	5	Punto_7	4/13/2018
4.751857	1.367082	6.188564	0	12.307503	4.751857	4	Punto_7	5/27/2018
2.25908	1.89624	8.058075	0.97578	13.189175	3.23486	8	Punto_7	7/9/2018
6.121209	0.920699	6.401384	2.97504	16.418332	9.096249	3	Punto_7	11/8/2018
16.790316	1.407999	7.604623	0	25.802938	16.790316	6	Punto_7	12/13/2018
12.027733	1.133372	10.590473	0	23.751578	12.027733	4	Punto_7	2/14/2019

Bandas Espectrales

B1	B11	B12	B2	B3	B4	B5	B8	nombre
0.1303	0.1115	0.0673	0.093	0.0711	0.0608	0.0842	0.0629	Punto_1
0.1373	0.1235	0.0743	0.0969	0.0865	0.0749	0.089	0.105	Punto_1
0.13770001	0.12255	0.06815	0.10235	0.08945	0.0728	0.08575	0.11295	Punto_1
0.1381	0.1216	0.062	0.1078	0.0924	0.0707	0.0825	0.1209	Punto_1
0.1373	0.1183	0.0599	0.0942	0.0786	0.0673	0.0755	0.1085	Punto_1
0.1269	0.0985	0.0474	0.088	0.073	0.0554	0.0683	0.0533	Punto_1
0.1252	0.1087	0.064	0.0975	0.0776	0.0709	0.0885	0.0735	Punto_1
0.12865	0.11145	0.05925	0.0973	0.0811	0.06985	0.08165	0.0933	Punto_1
0.14255001	0.14094999	0.08765	0.1105	0.09755	0.08615	0.09995	0.1103	Punto_1
0.1451	0.14515001	0.08155	0.10915	0.0973	0.08365	0.09835	0.10135	Punto_1
0.1424	0.17910001	0.119	0.1006	0.088	0.0819	0.1116	0.1137	Punto_1
0.1426	0.146	0.1023	0.10875	0.0975	0.08555	0.10345	0.09855	Punto_1
0.1364	0.1165	0.076	0.1039	0.0907	0.079	0.097	0.1139	Punto_1
0.1328	0.1165	0.0786	0.0991	0.0906	0.0761	0.097	0.1139	Punto_1
0.15305001	0.1329	0.0767	0.11565	0.0982	0.07975	0.10405	0.13265	Punto_1
0.16235	0.1663	0.09725	0.1407	0.13545001	0.13150001	0.1401	0.14665	Punto_1
0.1296	0.1736	0.0928	0.0977	0.0902	0.0707	0.0967	0.2397	Punto_1
0.1284	0.16060001	0.0931	0.0967	0.0896	0.0692	0.09765	0.2158	Punto_1
0.12485	0.126	0.07435	0.09065	0.07965	0.06155	0.08455	0.14285	Punto_1
0.1559	0.16930001	0.1165	0.1127	0.1013	0.0871	0.122	0.1556	Punto_2
0.1584	0.2033	0.1238	0.1291	0.1274	0.1166	0.1302	0.25650001	Punto_2
0.1613	0.19320001	0.1141	0.13155	0.12955	0.1109	0.1285	0.25350002	Punto_2
0.16419999	0.1831	0.1044	0.134	0.13169999	0.1052	0.1268	0.25049999	Punto_2
0.1524	0.1917	0.106	0.1209	0.1194	0.1047	0.1244	0.255	Punto_2
0.1548	0.15350001	0.0723	0.1095	0.1027	0.0708	0.0984	0.236	Punto_2
0.14579999	0.1463	0.0805	0.1047	0.0969	0.0674	0.1028	0.162	Punto_2
0.14579999	0.1463	0.0765	0.1047	0.0998	0.0765	0.1057	0.213	Punto_2
0.16485	0.18775	0.13495	0.12575001	0.1167	0.10905	0.13405	0.21185	Punto_2
0.1706	0.19284999	0.1346	0.12995	0.12055	0.1158	0.12989999	0.20015	Punto_2
0.155	0.16410001	0.088	0.1089	0.105	0.0791	0.1137	0.18539999	Punto_2

0.1566	0.1417	0.0799	0.11255	0.10325	0.08045	0.10865	0.1609	Punto_2
0.14935	0.1278	0.08425	0.1049	0.0892	0.06805	0.10295	0.11125	Punto_2
0.14354999	0.13945	0.0878	0.10715	0.0998	0.07905	0.1123	0.17225	Punto_2
0.1673	0.1705	0.0921	0.1292	0.1217	0.0894	0.1208	0.27680001	Punto_2
0.1673	0.1705	0.0855	0.1292	0.1217	0.0897	0.1208	0.27680001	Punto_2
0.1372	0.1283	0.0658	0.0972	0.0873	0.0621	0.0856	0.14910001	Punto_2
0.13675	0.12385	0.06505	0.09475	0.08055	0.05915	0.08525	0.13644999	Punto_2
0.1383	0.1054	0.0624	0.0935	0.07255	0.0522	0.0798	0.0887	Punto_3
0.1432	0.1252	0.0844	0.1074	0.0915	0.0871	0.0838	0.1344	Punto_3
0.13950001	0.1542	0.0897	0.122	0.1137	0.1006	0.1074	0.2397	Punto_3
0.13950001	0.1542	0.0897	0.122	0.1137	0.1006	0.1074	0.2397	Punto_3
0.1363	0.1565	0.0841	0.1165	0.1113	0.1084	0.1015	0.2273	Punto_3
0.13940001	0.1149	0.0556	0.1069	0.0934	0.0664	0.0817	0.1706	Punto_3
0.1355	0.1065	0.0566	0.1014	0.0852	0.0697	0.0867	0.0949	Punto_3
0.1355	0.1121	0.0566	0.1066	0.1012	0.0823	0.0927	0.20469999	Punto_3
0.15885	0.18054999	0.13240001	0.1397	0.12625	0.1089	0.1192	0.24770001	Punto_3
0.16275001	0.17685001	0.12925001	0.13940001	0.1231	0.10915	0.11955	0.22759999	Punto_3
0.1319	0.0983	0.0487	0.0954	0.0786	0.0599	0.0764	0.0864	Punto_3
0.14315	0.094	0.05005	0.11165	0.09915	0.0818	0.09835	0.1363	Punto_3
0.13935	0.0836	0.05065	0.10255	0.0795	0.06405	0.08125	0.0794	Punto_3
0.13355	0.0861	0.04845	0.09805	0.0821	0.0687	0.08145	0.07875	Punto_3
0.1689	0.1146	0.0617	0.1363	0.1202	0.0899	0.1018	0.2095	Punto_3
0.1567	0.10945	0.06195	0.12684999	0.1152	0.08625	0.10125	0.1927	Punto_3
0.1268	0.0857	0.0462	0.0912	0.0701	0.0526	0.0627	0.0564	Punto_3
0.12885	0.0872	0.0508	0.0916	0.07015	0.05445	0.06615	0.0587	Punto_3
0.13045	0.07895	0.0477	0.09095	0.0696	0.05025	0.0629	0.08625	Punto_3
0.14290001	0.16249999	0.1124	0.1021	0.0851	0.0721	0.115	0.15459999	Punto_4
0.1489	0.22570001	0.1525	0.1147	0.1102	0.0959	0.1533	0.2608	Punto_4
0.1515	0.2282	0.145	0.1181	0.11235	0.09395	0.1476	0.25935	Punto_4
0.1541	0.2307	0.1375	0.1215	0.1145	0.092	0.1419	0.2579	Punto_4
0.1497	0.22939999	0.1427	0.1132	0.1073	0.0915	0.1478	0.2552	Punto_4
0.1462	0.1751	0.0849	0.1198	0.1018	0.0882	0.12710001	0.183	Punto_4
0.13500001	0.14	0.0874	0.0936	0.0761	0.0606	0.0946	0.13079999	Punto_4
0.141	0.19769999	0.13249999	0.1003	0.1	0.0793	0.14740001	0.19769999	Punto_4
0.14595	0.22085001	0.1362	0.10655	0.10525	0.0829	0.14575	0.27410001	Punto_4
0.1496	0.2147	0.13335	0.10695	0.10395	0.07885	0.13965	0.25740001	Punto_4
0.1275	0.1331	0.0752	0.0893	0.077	0.0538	0.0884	0.1266	Punto_4
0.13935	0.1371	0.077	0.10025	0.08755	0.06485	0.0918	0.137	Punto_4
0.14435001	0.0932	0.0553	0.09585	0.0718	0.05455	0.07875	0.0766	Punto_4
0.13785	0.12549999	0.0765	0.09445	0.08475	0.06435	0.08895	0.12725	Punto_4
0.1551	0.15854999	0.09405	0.11775	0.1095	0.0849	0.1159	0.20514999	Punto_4
0.15099999	0.17715	0.10205	0.111	0.10735	0.0786	0.12265	0.25260001	Punto_4
0.124	0.0848	0.0383	0.0911	0.0695	0.0556	0.055	0.0669	Punto_4

0.1268	0.0827	0.04375	0.08935	0.06775	0.0527	0.06475	0.06945	Punto_4
0.13320001	0.09725	0.0558	0.09055	0.06585	0.0499	0.0722	0.0802	Punto_4
0.13590001	0.1416	0.0879	0.0998	0.0793	0.0687	0.0879	0.1142	Punto_5
0.1374	0.1916	0.109	0.1038	0.0906	0.0693	0.1073	0.206	Punto_5
0.14125	0.1927	0.1082	0.1088	0.10045	0.07875	0.10935	0.21960001	Punto_5
0.1451	0.1938	0.1074	0.1138	0.1103	0.0882	0.1114	0.2332	Punto_5
0.13079999	0.1945	0.1018	0.0993	0.0948	0.0787	0.1063	0.21420001	Punto_5
0.13150001	0.14120001	0.0641	0.1	0.0847	0.0618	0.0901	0.1839	Punto_5
0.12899999	0.1286	0.0709	0.0904	0.0804	0.0543	0.0827	0.1304	Punto_5
0.13070001	0.15880001	0.091	0.0944	0.08675	0.06455	0.10335	0.20395	Punto_5
0.1372	0.1866	0.10905	0.1087	0.10345	0.0861	0.11715	0.2349	Punto_5
0.13855	0.1912	0.1085	0.11495	0.11035	0.0925	0.11765	0.23045	Punto_5
0.1239	0.17030001	0.091	0.1176	0.1082	0.0956	0.1039	0.2159	Punto_5
0.1239	0.14399999	0.091	0.1176	0.1082	0.0956	0.1039	0.21089999	Punto_5
0.1322	0.07645	0.0422	0.0894	0.0614	0.0446	0.06045	0.06425	Punto_5
0.12645	0.1018	0.0578	0.0894	0.07325	0.0557	0.07545	0.11335	Punto_5
0.1521	0.1706	0.0937	0.1197	0.1116	0.0826	0.1069	0.25619999	Punto_5
0.1446	0.16965	0.0885	0.11615	0.11065	0.08485	0.1082	0.26185	Punto_5
0.1202	0.0741	0.0355	0.0874	0.0696	0.0477	0.0597	0.0994	Punto_5
0.12235	0.07655	0.03655	0.0858	0.0622	0.0458	0.0582	0.08235	Punto_5
0.1213	0.0788	0.0412	0.0823	0.0571	0.0374	0.0522	0.0726	Punto_5
0.12980001	0.1292	0.0712	0.0904	0.0697	0.057	0.0891	0.079	Punto_6
0.14129999	0.19419999	0.1128	0.1007	0.0886	0.0737	0.1135	0.1521	Punto_6
0.1402	0.1779	0.0967	0.1029	0.09035	0.0705	0.10675	0.1542	Punto_6
0.1391	0.16159999	0.0806	0.1051	0.0921	0.0673	0.1	0.15629999	Punto_6
0.1312	0.1706	0.0848	0.0922	0.0817	0.0615	0.1018	0.22130001	Punto_6
0.126	0.1095	0.0486	0.0851	0.071	0.0434	0.0756	0.075	Punto_6
0.1267	0.0969	0.0655	0.0922	0.0733	0.0578	0.0752	0.0558	Punto_6
0.12815	0.149	0.0842	0.0926	0.07995	0.0601	0.09345	0.13315	Punto_6
0.14614999	0.183	0.1004	0.10615	0.09575	0.0712	0.1139	0.16635001	Punto_6
0.14910001	0.1697	0.09285	0.1104	0.09825	0.07275	0.1087	0.15710001	Punto_6
0.1294	0.0919	0.0458	0.089	0.0757	0.0514	0.0736	0.0544	Punto_6
0.13195001	0.08965	0.04175	0.09795	0.0831	0.0648	0.0735	0.05575	Punto_6
0.1291	0.0792	0.0468	0.0926	0.0677	0.0527	0.0665	0.0454	Punto_6
0.1235	0.0807	0.0468	0.0855	0.0677	0.0508	0.0701	0.0561	Punto_6
0.1511	0.0692	0.0403	0.1132	0.087	0.0614	0.066	0.0797	Punto_6
0.1511	0.14210001	0.0711	0.1132	0.0987	0.0688	0.1011	0.1123	Punto_6
0.1236	0.0606	0.0349	0.0886	0.0767	0.0599	0.0668	0.0483	Punto_6
0.12415	0.064	0.0373	0.0903	0.0759	0.0598	0.06685	0.04655	Punto_6
0.123	0.06625	0.0418	0.08935	0.06855	0.0508	0.06405	0.0419	Punto_6
0.1284	0.1236	0.0659	0.0879	0.0675	0.05	0.0708	0.1138	Punto_7
0.134	0.1823	0.0792	0.0978	0.0894	0.0638	0.095	0.22930001	Punto_7
0.1353	0.1732	0.07765	0.10075	0.09305	0.06575	0.09295	0.24805	Punto_7

0.1366	0.16410001	0.0761	0.1037	0.0967	0.0677	0.0909	0.26679999	Punto_7
0.1319	0.16580001	0.0788	0.0906	0.0855	0.0571	0.0869	0.2606	Punto_7
0.1227	0.117	0.0441	0.0863	0.0703	0.0475	0.0656	0.14929999	Punto_7
0.1216	0.1071	0.0537	0.0863	0.0707	0.0466	0.0691	0.0957	Punto_7
0.12425	0.14120001	0.0676	0.08965	0.08145	0.05685	0.0808	0.2049	Punto_7
0.13940001	0.1741	0.0861	0.1105	0.10935	0.08535	0.0989	0.24895001	Punto_7
0.14155	0.16785	0.08035	0.1167	0.11435	0.09385	0.10025	0.2395	Punto_7
0.13420001	0.12	0.0625	0.0992	0.0861	0.0725	0.0819	0.0923	Punto_7
0.13609999	0.12	0.0688	0.1154	0.10855	0.09205	0.0904	0.13474999	Punto_7
0.12880001	0.0943	0.0498	0.0872	0.0631	0.0437	0.063	0.0697	Punto_7
0.1218	0.0955	0.0511	0.0927	0.0742	0.0609	0.0714	0.0923	Punto_7
0.1745	0.19374999	0.13420001	0.13625	0.1349	0.12585001	0.14035	0.196	Punto_7
0.14695001	0.14625	0.07315	0.1228	0.1161	0.09715	0.09905	0.21795	Punto_7
0.1171	0.0681	0.0327	0.0836	0.0671	0.0452	0.0564	0.0764	Punto_7
0.1185	0.07785	0.0379	0.0832	0.06305	0.0454	0.05605	0.07345	Punto_7
0.1198	0.09995	0.05265	0.0813	0.05635	0.0394	0.0569	0.0667	Punto_7

APENDICE F
CODIGOS

Serie temporal análisis Ambiental

```
# -*- coding: utf-8 -*-
"""
Created on Sun Aug 16 14:10:02 2026
@author: Kevin Rodriguez
"""
import os
import numpy as np
import pandas as pd
import rasterio
import matplotlib.pyplot as plt
import matplotlib.dates as mdates
# ---- RUTA ----#
DIR = r"C:\Users\geomatica\Desktop\K\TESIS_FINAL\drive-download-20260816T190548Z-1-001"
AÑO_INI, AÑO_FIN = 2000, 2025
VARIABLES = ["chirps", "lst", "ndvi", "t2m"]
UNIDADES = {"chirps": "mm/mes", "lst": "°C", "ndvi": "índice", "t2m": "°C"}
COLORES  = {"chirps": "#1f77b4", "lst": "#d62728",
            "ndvi": "#2ca02c", "t2m": "#ff7f0e"}
RANGOS   = {"chirps": (0, 450), "lst": (10, 50),
            "ndvi": (0.05, 0.85), "t2m": (5, 35)}
# Eventos de referencia (inicio, fin, etiqueta, tipo)
EVENTOS = [
    ("2005-07-01", "2005-07-31", "Emily",          "ciclon"),
    ("2010-06-25", "2010-07-15", "Alex",           "ciclon"),
    ("2010-11-01", "2013-06-30", "Sequía 2011–13", "sequia"),
    ("2020-08-01", "2023-03-31", "Triple La Niña", "sequia"),
    ("2022-03-01", "2022-08-31", "Crisis MTY",     "crisis"),
    ("2024-06-15", "2024-07-05", "Alberto",        "ciclon"),
]
ESTILO = {
```

```
    "sequia": dict(color="#c94c4c", alpha=0.10),
    "crisis": dict(color="#e07b00", alpha=0.22),
    "ciclon": dict(color="#2b6cb0", alpha=0.28),
}
def extraer_series(dir_tif, año_ini, año_fin):
    """Media, p10 y p90 espaciales por mes, para cada variable."""
    filas = []
    for var in VARIABLES:
        for año in range(año_ini, año_fin + 1):
            path = os.path.join(dir_tif, f"{var}_{año}.tif")
            if not os.path.exists(path):
                continue
            with rasterio.open(path) as src:
                nod = src.nodata
                for mes in range(1, min(src.count, 12) + 1):
                    a = src.read(mes).astype("float32")
                    if nod is not None and not np.isnan(nod):
                        a[a == nod] = np.nan
                    v = a[~np.isnan(a)]
                    filas.append({
                        "var": var, "año": año, "mes": mes,
                        "media": float(v.mean()) if v.size else np.nan,
                        "p10": float(np.percentile(v, 10)) if v.size else np.nan,
                        "p90": float(np.percentile(v, 90)) if v.size else np.nan,
                        "n_validos": int(v.size),
                        "frac_nan": float(np.isnan(a).mean())})
    df = pd.DataFrame(filas)
    if df.empty:
        print(f"[X] No se encontraron archivos en {dir_tif}")
        return df
    df["fecha"] = pd.to_datetime(dict(year=df.año, month=df.mes, day=1))
    return df.sort_values(["var", "fecha"]).reset_index(drop=True)
```

```
def marcar_eventos(ax, etiquetas=False):
    """Sombrear los eventos históricos sobre un eje."""
    # alternar la altura de las etiquetas para que no se encimen
    # cuando dos eventos caen cerca en el tiempo
    niveles = [1.02, 1.09]
    for k, (ini, fin, txt, tipo) in enumerate(EVENTOS):
        ax.axvspan(pd.Timestamp(ini), pd.Timestamp(fin), **ESTILO[tipo], zorder=0)
        if etiquetas:
            medio = pd.Timestamp(ini) + (pd.Timestamp(fin) - pd.Timestamp(ini)) / 2
            ax.annotate(txt, xy=(medio, niveles[k % 2]),
                        xycoords=("data", "axes fraction"),
                        ha="center", va="bottom", fontsize=7.5,
                        color=ESTILO[tipo]["color"], weight="bold")
def graficar(df, salida="serie_completa_2000_2025.png"):
    fig, axes = plt.subplots(len(VARIABLES), 1, figsize=(17, 12), sharex=True)

    for i, (ax, var) in enumerate(zip(axes, VARIABLES)):
        d = df[df["var"] == var]
        if d.empty:
            continue
        marcar_eventos(ax, etiquetas=(i == 0))
        c = COLORES[var]
        ax.fill_between(d.fecha, d.p10, d.p90, color=c, alpha=0.20,
                        lw=0, label="p10–p90 espacial", zorder=2)
        ax.plot(d.fecha, d.media, color=c, lw=1.3, zorder=3,
                label="media espacial")
        # media móvil de 12 meses
        suave = d.set_index("fecha").media.rolling(12, center=True).mean()
        ax.plot(suave.index, suave.values, color="black", lw=1.6, ls="--",
                alpha=0.65, zorder=4, label="media móvil 12 m")
        vacios = d[d.n_validos == 0]
        for f in vacios.fecha:
```

```
            ax.axvline(f, color="red", ls=":", lw=1, alpha=0.6, zorder=1)
        ax.set_ylabel(f"{var}\n({UNIDADES[var]})", fontsize=10)
        ax.grid(alpha=0.20, ls=":")
        ax.legend(fontsize=8, loc="upper left", ncol=3, framealpha=0.9)
        ax.set_xlim(pd.Timestamp(f"{AÑO_INI}-01-01"),
                    pd.Timestamp(f"{AÑO_FIN}-12-31"))
    axes[-1].xaxis.set_major_locator(mdates.YearLocator(2))
    axes[-1].xaxis.set_major_formatter(mdates.DateFormatter("%Y"))
    axes[-1].set_xlabel("Año")
    fig.suptitle(
        f"Covariables satelitales — media espacial mensual {AÑO_INI}–{AÑO_FIN}\n"
        "rojo: periodos de sequía · naranja: crisis hídrica de Monterrey · "
        "azul: ciclones tropicales",
        fontsize=12, y=0.985)
    plt.tight_layout(rect=[0, 0, 1, 0.96])
    plt.savefig(salida, dpi=150, bbox_inches="tight")
    print(f"\n[OK] {salida}")
    plt.show()
def anomalias(df, salida="anomalias_estandarizadas.png"):
    """
    Anomalías estandarizadas por mes calendario
    """
    fig, axes = plt.subplots(len(VARIABLES), 1, figsize=(17, 10), sharex=True)
    for i, (ax, var) in enumerate(zip(axes, VARIABLES)):
        d = df[df["var"] == var].copy()
        if d.empty:
            continue
        # z-score respecto al mismo mes calendario de toda la serie
        d["z"] = d.groupby("mes").media.transform(
            lambda s: (s - s.mean()) / s.std(ddof=0))
        marcar_eventos(ax, etiquetas=(i == 0))
        ax.axhline(0, color="black", lw=0.8, alpha=0.5)
```

```
        ax.fill_between(d.fecha, 0, d.z, where=d.z >= 0,
                        color="#2166ac", alpha=0.65, lw=0, zorder=3)
        ax.fill_between(d.fecha, 0, d.z, where=d.z < 0,
                        color="#b2182b", alpha=0.65, lw=0, zorder=3)
        suave = d.set_index("fecha").z.rolling(12, center=True).mean()
        ax.plot(suave.index, suave.values, color="black", lw=1.6, zorder=4)
        ax.set_ylabel(f"{var}\n(anomalía σ)", fontsize=10)
        ax.set_ylim(-3.5, 3.5)
        ax.grid(alpha=0.20, ls=":")
        ax.set_xlim(pd.Timestamp(f"{AÑO_INI}-01-01"),
                    pd.Timestamp(f"{AÑO_FIN}-12-31"))
    axes[-1].xaxis.set_major_locator(mdates.YearLocator(2))
    axes[-1].xaxis.set_major_formatter(mdates.DateFormatter("%Y"))
    axes[-1].set_xlabel("Año")
    fig.suptitle(
        "Anomalías estandarizadas por mes calendario\n"
        "(azul: por encima de lo normal · rojo: por debajo · "
        "línea negra: media móvil 12 meses)",
        fontsize=12, y=0.985)
    plt.tight_layout(rect=[0, 0, 1, 0.95])
    plt.savefig(salida, dpi=150, bbox_inches="tight")
    print(f"[OK] {salida}")
    plt.show()
def resumen_anual(df, salida="resumen_anual.png"):
    """Barras anuales: los años secos y húmedos deben saltar a la vista."""
    fig, axes = plt.subplots(2, 2, figsize=(15, 7))
    for ax, var in zip(axes.ravel(), VARIABLES):
        d = df[df["var"] == var]
        if d.empty:
            continue
        if var == "chirps":
            anual = d.groupby("año").media.sum()      # acumulado anual
            etiq = "mm/año"
```

```
        else:
            anual = d.groupby("año").media.mean()
            etiq = UNIDADES[var]
        media = anual.mean()
        colores = ["#b2182b" if v < media else "#2166ac" for v in anual.values]
        ax.bar(anual.index, anual.values, color=colores, alpha=0.85)
        ax.axhline(media, color="black", ls="--", lw=1,
                   label=f"media = {media:.1f}")
        ax.set_title(f"{var} ({etiq})", fontsize=10)
        ax.grid(axis="y", alpha=0.25, ls=":")
        ax.legend(fontsize=8)
        ax.tick_params(axis="x", rotation=45, labelsize=8)
    plt.suptitle("Resumen anual — rojo: bajo la media · azul: sobre la media",
                 fontsize=12)
    plt.tight_layout()
    plt.savefig(salida, dpi=150, bbox_inches="tight")
    print(f"[OK] {salida}")
    plt.show()
def diagnostico(df):
    print("\n" + "=" * 70)
    print("DIAGNÓSTICO")
    print("=" * 70)
    for var in VARIABLES:
        d = df[df["var"] == var]
        if d.empty:
            print(f"\n{var}: sin datos")
            continue
        lo, hi = RANGOS[var]
        m = d.media.dropna()
        print(f"\n{var}")
        print(f"  meses con dato : {len(m)} de {(AÑO_FIN-AÑO_INI+1)*12}")
        print(f"  media global   : {m.mean():.2f} {UNIDADES[var]}")
        print(f"  rango          : {m.min():.2f} a {m.max():.2f} "
```

```
            f"(esperado {lo}–{hi})")
        print(f"  meses sin datos: {(d.n_validos == 0).sum()}")
    ch = df[df["var"] == "chirps"]
    if not ch.empty:
        anual = ch.groupby("año").media.sum().sort_values()
        print("\n--- Precipitación anual acumulada (más seco -> más húmedo) ---")
        for a, v in anual.items():
            marca = ""
            if a in (2011, 2012, 2013):
                marca = "  <- sequía 2011-13"
            elif a in (2021, 2022):
                marca = "  <- crisis Monterrey"
            elif a == 2010:
                marca = "  <- huracán Alex"
            print(f"  {a}: {v:7.1f} mm{marca}")

        secos = list(anual.head(6).index)
        print(f"\n  6 años más secos: {sorted(secos)}")
        esperados = {2011, 2012, 2013, 2021, 2022}
        coincide = esperados & set(secos)
        print(f"  De los esperados {sorted(esperados)}, aparecen: {sorted(coincide)}")
        if len(coincide) >= 3:
            print("  [OK] Los datos reproducen las sequías documentadas.")
        else:
            print("  [!] Pocas coincidencias: revisa la agregación mensual.")
if __name__ == "__main__":
    print(f"Leyendo {DIR} ...")
    df = extraer_series(DIR, AÑO_INI, AÑO_FIN)
    if not df.empty:
        df.to_csv("series_covariables_2000_2025.csv", index=False)
        print(f"[OK] Tabla: series_covariables_2000_2025.csv ({len(df)} filas)")
        diagnostico(df)
        graficar(df)
```

```
anomalias(df)
resumen_anual(df)
```

K-means clustering

```
// Crear composición de imágenes Sentinel 2
var img = ee.ImageCollection("projects/planet-nicfi/assets/basemaps/americas")
                .filterDate('2018-11-01', '2019-04-30')
                //.filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 20))
                .filterBounds(geometry)
                .map(function(image){return image.clip(geometry)})
                .median();

// Centrar el mapa en el área de estudio
Map.centerObject(geometry);

// Añadir imagen al mapa
Map.addLayer (img, {
  max: 4000,
  min: 0.0,
  gamma: 1.0,
  bands: ['R','G','B']},
  'Imagen Sentinel 2');

// Crear datos de entrenamiento
var training = img.sample({
  region: geometry,
  scale: 5,
  numPixels: 1000
});

// Aplicar algoritmo k-means y "entrenarlo"
var kmeans = ee.Clusterer.wekaKMeans(5).train(training);
```

```
// Agrupar los resultados del algoritmo
var resultado =  img.cluster(kmeans);

// Añadir imagen clasificada al mapa con colores aleatorios
Map.addLayer(resultado.randomVisualizer(), {}, 'Clasificación K-means');

// Exportar imagen clasificada a Drive
Export.image.toDrive({
  image: resultado,
  description: 'kmeans_img',
  scale: 5,
  region: geometry
});
```

Random Forest

```
//// 1) IMAGEN SATELITAL ////

// Crear composición de imágenes Sentinel 2
var img = ee.ImageCollection("projects/planet-nicfi/assets/basemaps/americas")
                .filterDate('2017-11-01', '2018-04-30')
                //.filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 20))
                .filterBounds(geometry)
                .map(function(image){return image.clip(geometry)})
                .median();

// Centrar el mapa en el área de estudio
Map.centerObject(geometry);

//Configura el mapa base para que se muestre como satélite.
Map.setOptions('SATELLITE');
```

```
// Añadir imagen al mapa
Map.addLayer (img, {
  max: 4000,
  min: 0.0,
  gamma: 1.0,
  bands: ['R','G','B']},
  'Imagen NICFI');

//-------------------------------------------------------------------------------------------------
----
//// 2) CLASIFICACIÓN ////

// Paleta de colores
var paleta = [
  "#112ed6", // Agua
  "#4f8400", // Vegetacion
  "#e47940", // Agricultura
  "#cd381c", // Urbano
  "#fffe46"]; // Suelo

// Seleccionar bandas
var bands = img.bandNames();



// Unir los datos de entrenamiento de las distintas clases
var                                    training_data                                    =
Agua.merge(vegetacion).merge(agricultura).merge(urbano).merge(suelo);

// "Muestrear las regiones"
```

```
var rf_training = img.select(bands).sampleRegions(
  {collection: training_data,
    properties: ["land_class"],
    scale: 5
  });

/*
//muestrar regiones con puntos gps de otra parte

var ptos= img.select(bands).sampleregions(
{ collection: puntos,                                    puntos: shape o csv de datos. "id clase" columna de puntos
  properties: ["id_clase"],
  scale: 4.7
)}. randomcolumn();



//seleccionar porcentaje para separar los datos
var split = 0.7

var training_data= ptos.filter(ee.Filter.lt("random", split));
var validation_data= ptos.filter(ee.Filter.gte("random", split));

// Entrenar al clasificador (100 arboles)
var rf = ee.Classifier.smileRandomForest(300).train({
  features: training_data,
  classProperty: "id_clase",
  inputProperties: bands
});
```

```
// Obtener la imagen clasificada
var img_clas = img.select(bands).classify(rf);

// Añadir la imagen clasificada al mapa
Map.addLayer(img_clas, {min:1, max:5, palette: paleta}, "LULC area de estudio");

*/


// Entrenar al clasificador
var rf = ee.Classifier.smileRandomForest(100).train({
  features: rf_training,
  classProperty: "land_class",
  inputProperties: bands
});

// Obtener la imagen clasificada
var img_clas = img.select(bands).classify(rf);

// Añadir la imagen clasificada al mapa
Map.addLayer(img_clas, {min:1, max:5, palette: paleta}, "LULC area de estudio");

// Exportar la imagen clasificada a Drive
Export.image.toDrive({
  image: img_clas,
  description: 'RF(3)_2018_2019',
  scale: 5,
  folder: "GEE",
  region: geometry});
```

```
 //Agua20
 //vegetacion 15
 //agricultura
 //urbano
 //suelo


 /* ee.Geometry.Polygon(
       [[[-100.36503572808192, 25.82683533519227],
         [-100.36503572808192, 25.78943681366401],
         [-100.19543428765223, 25.78943681366401],
         [-100.19543428765223, 25.82683533519227]]], null, false);
*/
// Unir los datos de validación de las distintas clases
var validation_data = Valagua.merge(Valvegetacion).merge(Valagricultura).merge(Valurbano).merge(Valsuelo);


var validation = img.sampleRegions({
  collection: validation_data,
  properties: ["land_class"],
  scale: 5,
  tileScale: 16
});


var validated = validation.classify(rf);


// Obetener matriz de confusión
var mc = validated.errorMatrix("land_class", "classification");
```

```
//Obtener matriz de confusión
print("Matriz de Confusión", mc);
print("Precisión General: ", mc.accuracy());
print("Coeficiente Kappa: ", mc.kappa());

// Calcular Producer's Accuracy (PA) y User's Accuracy (UA)
var pa = mc.producersAccuracy();
var ua = mc.consumersAccuracy();

// Imprimir los resultados en la consola
print('Producer´s Accuracy:', mc.producersAccuracy());
print('User´s Accuracy:', mc.consumersAccuracy());
//random points 1000
//random points 1000
var puntosSNG= ee.FeatureCollection.randomPoints(geometry,1000)
Map.addLayer (puntosSNG);
Export.table.toAsset({
  collection:puntosSNG,
  description:'exportarAssets',
  assetId:'puntosSNG'
});


//random points drive
Export.table.toDrive({
  collection: puntosSNG,
  folder: "estudio_zmm_regionTemp",
  description: 'exportarDriveCSV',
  fileNamePrefix: "puntosSNG",
  fileFormat: "CSV"
```

```
})

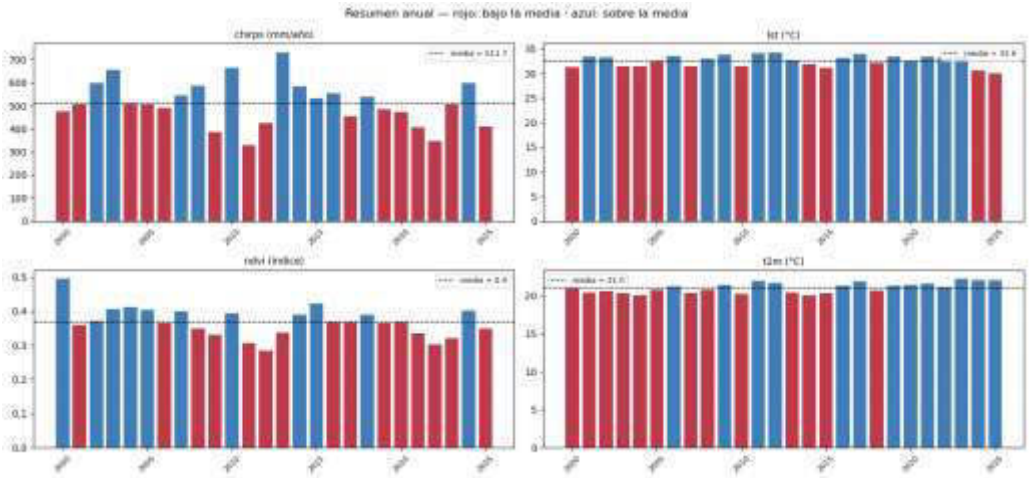
Export.table.toDrive({
  collection:puntosSNG,
  folder: "estudio_zmm_regionTemp",
  description:'exportarDriveSHP',
  fileNamePrefix: "puntosSNG",
  fileFormat: "SHP"
})
 ;
 // Cargar DEM
var dem = ee.Image("USGS/SRTMGL1_003").clip(geometry);

// Obtener pendiente
// (la tasa máxima de cambio de valor del DEM de un pixel con respecto a sus vecinos)
var pendiente =  ee.Terrain.slope(dem);

// Obtener aspecto (orientación de laderas)
// (puede ser pensado como la dirección de la pendiente)
var aspecto =  ee.Terrain.aspect(dem);

// Añadir aspecto al mapa
Map.addLayer(aspecto, {min: 0,
  max: 360,
  palette: [
    '3ae237', 'b5e22e', 'd6e21f', 'fff705', 'ffd611', 'ffb613', 'ff8b13',
    'ff6e08', 'ff500d', 'ff0000', 'de0101', 'c21301', '0602ff', '235cb1',
    '307ef3', '269db1', '30c8e2', '32d3ef', '3be285', '3ff38f', '86e26f'
  ]},"Aspecto");
```

```
// Añadir pendiente al mapa
Map.addLayer(pendiente, {
  min: 0,
  max: 90,
  palette: [
    '3ae237', 'b5e22e', 'd6e21f', 'fff705', 'ffd611', 'ffb613', 'ff8b13',
    'ff6e08', 'ff500d', 'ff0000', 'de0101', 'c21301', '0602ff', '235cb1',
    '307ef3', '269db1', '30c8e2', '32d3ef', '3be285', '3ff38f', '86e26f'
  ]},"Pendiente");
```



APENDICE G
REPORTE PDF APP WEB

SPECTRAL INDICES REPORT

Sentinel-2 Analysis — RGB, Vegetation, Water, and Turbidity

Search period: 30 Sep 2018 → 11 Oct 2019
Image date: L2A_2018 · **Clouds:** 0.0%
Polygon area: 322.39 km² · **Indices:** NDVI, NDWI, NDCI
Generated: 16/09/2026



Source: Copernicus Sentinel-2 SR Harmonized, via Google Earth Engine.

Report automatically generated from Sentinel-2 imagery (Copernicus), processed in Google Earth Engine. Valid for any study area on the planet.

Designed by Kevin Rodríguez González
Geomatics Department · UANL · FIC

1. Introduction

Spectral indices derived from Sentinel-2 imagery (ESA Copernicus) make it possible to characterize biophysical and optical properties of surface water in a spatially continuous and repeatable way. This report presents the results of the multivariate analysis of spectral indices computed in Google Earth Engine (GEE) for the study area defined by the shapefile uploaded to the Water Quality Mapping platform. The indices cover water quality (NDCI, SABI, CDOM, NDTI), surface water presence (NDWI, MNDWI, AWEInsh), vegetation (NDVI, EVI) and surface temperature (LST). The temporal analysis includes the Mann-Kendall test to detect significant trends in the historical series.

2. Methodology

Component	Detail
Sensor	Sentinel-2 MSI (ESA Copernicus), 10 m / 20 m spatial resolution
GEE collection	COPERNICUS/S2_SR_HARMONIZED (surface reflectance)
Compositing	Median composite with per-pixel QA60 cloud mask
Spatial clip	Exact clip to the uploaded shapefile polygon (not rectangular)
Statistics	reduceRegion — mean, std. dev., min, max, 50th and 90th percentiles
Trends	Mann-Kendall (Kendall's τ) + least-squares linear regression
Platform	Google Earth Engine · Python · Streamlit · UANL Geomatics

Spectral Index Formulas

Index	Formula	Reference
NDVI	(B8−B4)/(B8+B4)	Rouse et al., 1974
NDWI	(B3−B8)/(B3+B8)	McFeeters, 1996
MNDWI	(B3−B11)/(B3+B11)	Xu, 2006
NDTI	(B11−B3)/(B11+B3)	Lacaux et al., 2007
NDCI	(B5−B4)/(B5+B4)	Mishra & Mishra, 2012
SABI	(B8−B4)/(B2+B3)	Alawadi, 2010
CDOM	B3/B4	Kirk, 1994
AWEInsh	4(B3−B11) −0.25B8 −2.75B12	Feyisa et al., 2014
EVI	2.5(B8−B4)/(B8+6B4−7.5B2+1)	Huete et al., 2002
LST	Landsat 8/9 ST_B10 (TsHARP 10m via S2)	Malakar et al., 2018

3. Study Area

Coordinates: Longitude -100.44160° – -100.16740° · Latitude 25.48210° – 25.79650°
Estimated area: 322.39 km
Spatial resolution: 10 m (visible/NIR bands) · 20 m (SWIR/Red-Edge)

4. Interpretive Summary

This analysis spectrally characterizes the study area using a Sentinel-2 SR image, calculating standardized indices for vegetation, water bodies, and relative turbidity. **NDVI**: area mean value = 0.403 — moderate vegetation (dense grassland, developing crops, shrubland). **NDWI**: area mean value = -0.378 — dense vegetation or dry soil, no water signal. **NDCI**: area mean value = 0.221 — high chlorophyll-a concentration — possible algal bloom. These indices are dimensionless (range −1 to 1) and are calculated directly from surface reflectance, requiring no local calibration or field data — they are applicable to any ecosystem or region in the world.

5. Zonal Statistics by Index

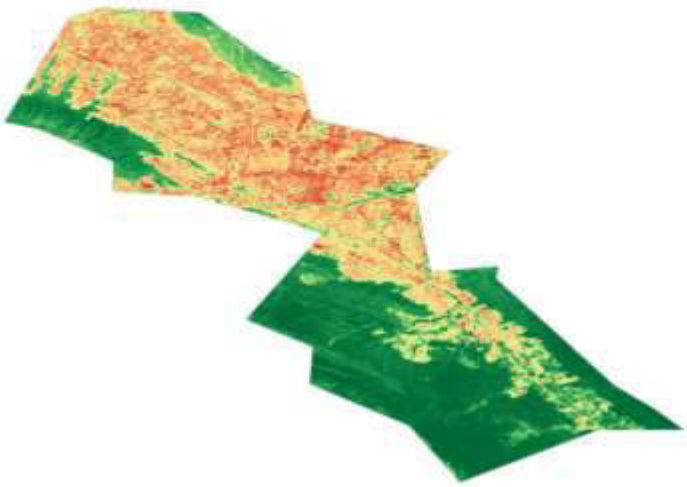
Index	Mean	Std. Dev.	Min	Max	P50 (median)
NDVI	0.403	0.299	-0.474	0.890	0.379
NDWI	-0.378	0.232	-0.794	0.618	-0.379
NDCI	0.221	0.174	-0.197	0.614	0.166

Statistics calculated over the entire uploaded polygon area (reduceRegion, 20 m/pixel). σ = standard deviation; P50 = median.

6. Spectral Maps by Index

Each map shows the spatial distribution of the corresponding index over the study area, along with its mean value and associated interpretive classification.

NDVI



Zone mean value: **0.403** (σ=0.299) — moderate vegetation (dense grassland, developing crops, shrubland)

NDWI



Zone mean value: **-0.378** (σ=0.232) — dense vegetation or dry soil, no water signal

NDCI



Zone mean value: **0.221** (σ=0.174) — high chlorophyll-a concentration — possible algal bloom

7. Time Series (Google Earth Engine)

The time series were extracted from Google Earth Engine by computing the zonal mean of each index over the study area for every available Sentinel-2 image. The temporal evolution, the linear trend line and the Mann-Kendall test result are shown to assess the statistical significance of the trend ($\alpha = 0.05$).

Trend Summary (Mann-Kendall)

Index	N images	Minimum	Maximum	Mean	MK trend
NDVI	24	0.2837	0.4846	0.4076	↓ τ=-0.72 (sig.)
NDWI	24	-0.4638	-0.2834	-0.3957	↑ τ=0.74 (sig.)
NDCI	24	0.1583	0.2737	0.2238	↓ τ=-0.59 (sig.)






NDCI — Time series (zonal mean)
Mann-Kendall: ↓ Decreasing (p<0.05 ✓)
Zonal mean
Linear trend
NDCI
0.26
0.24
0.22
0.20
0.18
0.16
Oct 18
Nov 18
Dec 18
Dec 18
Jan 19
Apr 19
Apr 19
Jun 19
Jul 19
Jul 19
Aug 19
Aug 19
Aug 19

8. Applications and Recommendations

These spectral indices are rapid diagnostic tools applicable to environmental monitoring, precision agriculture, watershed management, urban planning, and land cover change studies. For quantitative analyses requiring physical units (mg/L, NTU, etc.), it is recommended to complement with field data and, if applicable, a site-specific calibration model, similar to the one developed for the Pesquería River on this same platform.

Cite as: Rodríguez González, K.D. (2026). Water Quality & Spectral Indices Mapping Tool. Universidad Autónoma de Nuevo León, Faculty of Civil Engineering, Department of Geomatics. https://waterqualitygeomaticauanl.streamlit.app/