

Laboratorio 4

Análisis de modelos usando datos geospaciales.

INSTRUCCIONES

Ejercicio.

Los lagos Atitlán y Amatitlán son cuerpos de agua de gran importancia ecológica, económica y cultural en Guatemala. Sin embargo, en las últimas décadas ambos han mostrado signos alarmantes de deterioro ambiental, especialmente por la proliferación de **cianobacterias** — microorganismos que pueden formar floraciones tóxicas cuando las condiciones del agua son propicias (alta temperatura, nutrientes, estancamiento, etc.).

El monitoreo constante de estas floraciones es esencial para gestionar riesgos a la salud pública, el turismo y los ecosistemas acuáticos. Sin embargo, realizar muestreos físicos frecuentes puede ser costoso, limitado en alcance espacial, y logísticamente complejo.

Aquí es donde entra en juego la **observación de la Tierra mediante satélites**. En particular, la misión **Sentinel-2** del programa Copernicus proporciona imágenes multiespectrales de alta resolución que permiten detectar cambios en la vegetación, calidad del agua y floraciones algales mediante índices espectrales.

Además, herramientas como **Sentinel Hub** y su API nos permiten acceder a scripts personalizados (como el de detección de cianobacteria) y procesar las imágenes directamente desde la nube, sin necesidad de descargar grandes volúmenes de datos.

DESCRIPCIÓN DE LOS DATOS

Se te proporciona en este laboratorio las coordenadas de cada uno de los lagos y su respectivo geojson. Debes contar con tu usuario y contraseña para ingresar al copernicus browser. Estaremos usando scripts personalizados que pueden encontrarse en <https://custom-scripts.sentinel-hub.com>.

Coordenadas:

```
lago_atitlan = {
  "west": -91.326256,
  "east": -91.07151,
  "south": 14.5948,
  "north": 14.750979
}
```

```
lago_amatitlan = {
  "west": -90.638065,
  "east": -90.512924,
  "south": 14.412347,
  "north": 14.493799
}
```

EJERCICIOS

1. Establezca una conexión con la api de sentinel 2, puede usar para eso el módulo openeo.
2. Descarga los .tif de cada uno de los lagos usando para eso las coordenadas o el geojson provisto.
3. Selecciona un período de observación: **mínimo 6 meses**. Tenga en cuenta que serán útiles imágenes con menos nubosidad.
4. Aplica el script de **detección de cianobacteria** de Sentinel Hub (por ejemplo, Cyano Detection Script) para cada imagen.
 - 4.1. Puede calcularlo en copernicus browser y descargar los resultados o puede calcularlo usando para eso las bandas que incluyó la descarga de cada una de las imágenes.
5. Conviértalo en arreglos de numpy para analizarlos
6. Haga un análisis temporal para eso:
 - 6.1. Calcula un **índice promedio de cianobacteria por lago y por fecha**.
 - 6.2. Visualiza la evolución temporal en un gráfico de línea por lago.
 - 6.3. Identifica posibles picos de floración y fechas críticas.
7. Haga un análisis espacial para esto:
 - 7.1. Mapea la distribución de cianobacteria dentro de cada lago usando matplotlib o folium (genera un mapa interactivo)
 - 7.2. Crea mapas comparativos entre diferentes fechas para cada lago.
8. Haga una correlación de los índices (NDVI y NDWI) con la presencia de la cianobacteria. Explique sus hallazgos
9. Análisis de los lagos y comparación entre ellos:
 - 9.1. Haga un análisis de la proliferación de la cianobacteria en el período estudiado para cada uno de los lagos.
 - 9.2. Compara la intensidad y la frecuencia de floraciones entre ambos lagos.
 - 9.3. Evalúa posibles diferencias con respecto a las causas de la proliferación de la cianobacteria y discútalas (geografía, uso del suelo, presión urbana, temperatura del agua, del ambiente, clima, etc.).

Luego, para Cada Lago:

10. Utilice una serie temporal para predecir el índice de cianobacteria.
11. Haga un modelo que permita determinar si un punto en específico en el lago tiene cianobacteria o no.
12. Haga un modelo híbrido donde primero prediga el índice de cianobacteria usando series temporales y luego, utilice el índice predicho junto con otras características para clasificar si un área estará contaminada o no. Siéntase libre de añadir variables de otros conjuntos de datos que le permitan complementar el índice (temperatura, índice de urbanización, etc.).
13. Use los modelos para predecir y analice los resultados que está obteniendo.
14. Muestre en un mapa, los resultados de las proyecciones.

dataset con imagen, índice de vegetación, ...
Las fechas que no tenemos, llenarlas (promedio o algo)

REFERENCIAS

En estos dos sitios, hay datos de precipitaciones y temperatura de los lagos que pudieran usar:

- https://weatherspark.com/y/11701/Average-Weather-in-Amatitl%C3%A1n-Guatemala-Year-Round#google_vignette
- <https://weatherspark.com/compare/y/11701~11135/Comparison-of-the-Average-Weather-in-Amatitl%C3%A1n-and-Santiago-Atitl%C3%A1n>

EVALUACIÓN

NOTA: La evaluación de cada integrante del grupo será de acuerdo con sus contribuciones al trabajo grupal, es necesario versionar.

(15 puntos) Serie temporal:

- Se hace un modelo de serie de tiempo para predecir el avance de la cianobacteria en cada uno de los lagos.

(10 puntos) Modelo de clasificación

- Se elabora un modelo de clasificación para cada uno de los lagos para determinar si en un punto x habrá cianobacteria o no.

(20 puntos) Modelo híbrido

- Se combinan ambas técnicas para elaborar modelos híbridos para cada uno de los lagos.

(20 puntos) Análisis de los modelos

- Se discuten los resultados de las predicciones y la utilidad de los modelos seleccionados. Se dan recomendaciones, para el mejor uso de los modelos. Analice si necesita descargar más datos para mejorar el desempeño de estos.

(20 puntos) Visualización

- Se visualizan los resultados de los modelos en un mapa.

FECHAS DE ENTREGA

- 14 de agosto de 2025 17:20:
 - .1. Avances. Modelos de series temporales.
- 17 de agosto de 2025 23:59
 - .1. Ejercicios Completos.

NOTA: Para poder tener nota completa debe entregar las asignaciones en el tiempo adecuado. No se calificará el avance del laboratorio si no fue entregado en tiempo, aunque esté en el repositorio.

MATERIAL A ENTREGAR

- Archivo .pdf con el informe que contenga, los resultados de los análisis y las explicaciones. Si incluye las explicaciones y el análisis en un jupyter notebook o markdown no es necesario entregar archivo .pdf.
- Link de Google drive donde trabajó el grupo.
- Script de R (.r o .rmd) o de Python que utilizó para hacer su análisis exploratorio y predicciones.
- Link del repositorio usado para versionar el código.