

# Predicción del transporte de sólidos en el valle alto del río Cauca a la altura de la planta Puerto Mallarino por medio de análisis de imágenes satelitales y procesamiento digital de señales)

V. D. Córdoba<sup>1</sup>, I. C. Dallos<sup>2</sup>

<sup>1</sup>Pontificia Universidad Javeriana Cali. Valle del Cauca. Colombia

<sup>2</sup>Pontificia Universidad Javeriana Cali. Valle del Cauca. Colombia

e-mail: victorcordoba@javerianacali.edu.co, 1v4nd4llos@javerianacali.edu.co

**resumen** – En el presente trabajo de grado se busca inferir la sedimentación presente en el río Cauca a la altura de la planta puerto Mallarino al Oriente de la ciudad de Cali con la información obtenida a través de la combinación de bandas espectrales presentes en imágenes satelitales de libre uso. En el estudio se contó con muestras de turbidez proporcionadas por Emcali desde el año 2013 sumado a imágenes del satélite sentinel-2 tomadas desde el 2017, las cuales fueron filtradas manualmente confirmando que la zona de estudio estuviese libre de nubes y sombras para luego tener acceso a la reflectancia de cada banda por los puntos en la zona de estudio. Con esta información se construyen modelos utilizando regresión lineal junto con los coeficientes de Pearson, Kendall y Spearman que correlacionan la firma espectral de la imagen con los valores de turbidez discriminando las combinaciones de bandas con baja correlación de Pearson. Para tener el menor error posible se crean tres modelos, uno para turbidez alta, uno para turbidez baja y otro para decidir que modelo utilizar.

**Keywords** – Puerto Mallarino – Turbidity – Regression model – Spectral Bands

## I. Introducción

La planta de Puerto Mallarino de la ciudad de Cali en Colombia toma sus aguas del río Cauca al oriente de la ciudad, la cual le llegan los sedimentos de sus ríos tributarios junto a las aguas servidas del canal sur contaminándolo y volviendo turbia el agua. Cuando el río tiene demasiada turbidez mayor a 3000 NTU se tiene que cerrar obligatoriamente la bocatoma de la planta para no generar daños en los filtros y garantizar un agua apta para el consumo humano en la salida. Un avance en las tecnologías de detección de turbidez para estos casos es lo ideal para que estas afecciones no sean graves para la planta y no se generen daños a la infraestructura de esta. Por ello, como objetivo de este documento es construir un sistema que prediga el nivel de turbidez basado en imágenes satelitales que puede incluso ser implementado siguiendo la misma metodología en otros lugares.

## II. Fundamentación Teórica

1) *Turbidez*: Se refiere a la transparencia de un líquido incoloro debido a las partículas suspendidas. Cuanto mayor sea el número de partículas en suspensión, mayor será su nivel de turbidez.

2) *Reflectancia del agua*: La reflectividad es la fracción de energía incidente reflejada por un objeto, esta cambia en función de la longitud de onda. La distribución de la reflectividad en función de la longitud de onda se denomina firma espectral, que es diferente para cada objeto en la superficie terrestre, por lo que es posible su discriminación. [1]

los cuerpos de agua terrestres tienen una reflectividad casi nula en el infrarrojo medio. Además, la reflectividad del agua depende de la profundidad del cuerpo y del sustrato del lecho. [1]

Los cuerpos de agua clara presentan una mayor reflectancia en la banda verde y azul (560 nm) a comparación de las bandas de infrarrojos (739 nm). Las aguas con presencia de turbidez por sedimentación aumentan su reflectividad hacia la banda roja (664 nm). En la figura 1 muestra cómo se comporta la reflectancia detectada por las bandas del satélite Sentinel-2 ya sea por agua clara o agua turbia debido a la sedimentación. [2]

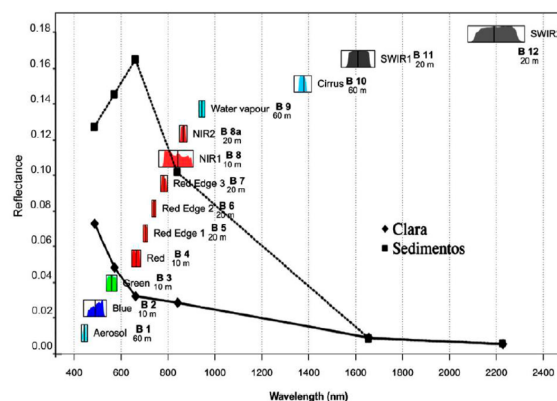


Figura 1. Firma espectral agua sedimentada vs agua clara [1] y [3]

3) *Satélites de observación*: Existen varios satélites de observación; LandSat-7 [4], LandSat-8 [5], Sentinel-2 [6], Superview [7], Pleiades 1 [8] y

KompSat [9] cada uno tiene diferentes rangos en las bandas espectrales y períodos de revisita. Los satélites LandSat-7, LandSat-8 y Sentinel 2 son de libre acceso y por tanto representan una gran ventaja para la realización de este proyecto. Seguido con su amplio espectro de bandas. El factor decisivo para la elección del satélite Sentinel-2 para el proyecto, se debe a su período de revisita (cada 5 días) y a la calidad de imágenes; el LandSat-8 tiene un período de revisión de 16 días y el satélite LandSat-7 presenta líneas blancas en las imágenes (errores de reflexión electromagnética). [10]

4) *Bandas espectrales*: Una banda espectral se refiere a la parte del espectro electromagnético (EM) que es captada por un sensor, el EM se divide en varias bandas que se diferencian entre sí por la frecuencia de la onda reflejada. Cada objeto recibe todo el EM, pero refleja solo su firma espectral. De esta forma, es posible diferenciar qué tipo de objeto se captura en una imagen. Actualmente existen 13 bandas espectrales para el satélite Sentinel 2 que describen en general los materiales u objetos que tienen una firma espectral en cada rango: [11]

- **1. Aerosol costero (433,9-453,9 nm)**: Refleja azules y púrpuras. Es menos absorbido por el agua y es más sensible a las nubes, el humo y la bruma lo que permite observar el color del océano y las aguas costeras.
- **2. Azul (464,1-529,1 nm)**: Ideal para obtener imágenes de aguas profundas, reconocer: humo y nubes.
- **3. Verde (542,5-577,5 nm)**: la banda verde es útil para evaluar el vigor de plantas y vegetación en mares, bosques y zonas urbanas.
- **4. Roja (649,5-679,5 nm)**: Los suelos tropicales, entornos construidos y características geológicas generalmente se reflejan en esta banda. Algunas aplicaciones pueden ser el análisis del tipo de suelo y área construida.
- **5. Borde rojo (703,9-782,5 nm)**: Está entre la banda roja y el infrarrojo cercano. Se utiliza para analizar la salud de las plantas. Las plantas absorben la luz roja y reflejan el infrarrojo. Hay tres sub-bandas en esta banda y cuanto mayor es la longitud de onda, menos se ve afectada por las condiciones atmosféricas (bandas 5, 6 y 7).
- **8. Infrarrojo cercano 1 - NIR (777.6-892.6nm)**: Una de las formas para clasificar vegetación saludable. Separar clases como el agua y la vegetación es más fácil en esta región. Las plantas reflejan energía en esta banda, mientras que el agua la absorbe. Se emplea para evaluar cambios sutiles en las plantas de un área. Por otro lado, Narrow NIR es una banda más estrecha encapsulada en este mismo rango (bandas 8 y 8b).
- **9. Vapor de agua (935-955nm)**: Similar al infrarrojo cercano. La principal diferencia es

que esta banda se ve menos afectada por las condiciones atmosféricas.

- **10. Cirro (1358,5 -1388,5 nm)** Se usa en la detección de nubes de cirros. El suelo es apenas visible en esta banda y las nubes de gran altitud que no son visibles en otras bandas se detectan y seleccionan.
- **11. Infrarrojos de onda corta 1 - SWIR-1 (1568.7-1658.7nm)**: Puede ayudar a discriminar entre suelos secos y húmedos, hacer clasificación de suelos y penetrar nubes delgadas, incluso humo y neblina mejor que las bandas visibles.
- **12. Infrarrojos de onda corta 2 - SWIR-2 (2112.4-2292.4nm)**: Se utiliza principalmente para obtener imágenes de tipos de suelo, características geológicas y minerales. También es sensible a la vegetación y a las variaciones de humedad del suelo. La nieve, el hielo y las nubes característicos aparecen en tonos más oscuros.

Las bandas espectrales permiten inferir información al operar entre ellas para obtener nueva información. Herramientas como Google Earth Engine [12] o LandViewer [13] ofrecen el obtener y tratar imágenes satelitales además de acceder a la información de las bandas espectrales.

5) *Datos de turbidez Emcali*: Se tienen datos de turbidez desde 2013 hasta noviembre del 2020 tomados por EMCALI<sup>1</sup> en 2 sitios, Puerto Mallarino y Milán. Milán se encuentra a 3 horas río arriba de Puerto Mallarino. La turbidez se toma cada minuto por lo que para ajustar la ventana de tiempo a los datos de reflectancia retornados por las imágenes satelitales se toma el promedio diario de turbidez.

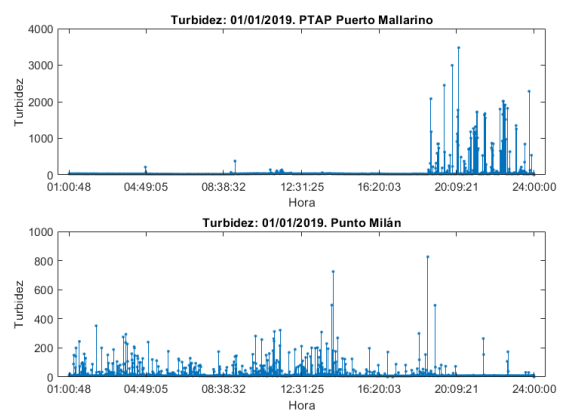


Figura 2. Turbidez Puerto Mallarino vs Milán.

En la Figura 2 los datos entre ambos puntos presentan inconsistencias en un mismo día. Para Milán hay una variabilidad considerable durante un mismo día, que no está representada en Puerto Mallarino. Esta situación en el punto de muestreo Milán puede ser debido a las aguas residuales del canal sur que son depositadas en esta sección del río lo que causa una

<sup>1</sup>Empresa Municipales de Cali E.I.C.E. E.S.P.

mayor densidad de partículas de suspensión, por esto se prefiere excluir los datos del punto de nuestro Milán.

6) *Regresión*: Para iniciar el modelo matemático de este trabajo, se usan los coeficientes de correlación como punto de referencia.

Los coeficientes de correlación utilizados son el coeficiente de Kendall, el coeficiente de Spearman y el coeficiente de Pearson de cada banda con respecto al dato real de turbidez los cuales indican una medida basada en rangos para los dos primeros y una medida lineal respectivamente [14]. Las fórmulas correspondientes se indican en las ecuaciones (1 - 2 - 3).

$$Rho_{Pearson}(x,y) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (1)$$

$$Rho_{Spearman}(x,y) = 1 - \frac{6 \sum d_i^2}{n^3 - n} \quad (2)$$

$$Rho_{Kendall}(x,y) = \frac{2K}{n(n-1)} \quad (3)$$

$\sigma_{xy}$  - Covarianza de x,y.

$\sigma_x, \sigma_y$ - Desviación típica de x y y respectivamente.

$d_i$  - Vector de diferencias entre x y y.

$n$  - Longitud de datos.

$K$  - Diferencia entre la suma del conteo de todos los mayores a  $Y(i)$  para cada  $Y(j > i)$  y la suma del conteo de todos los menores a  $Y(i)$  para cada  $Y(j > i)$ .

Para reducir la complejidad de estas ecuaciones, se utiliza la función "corr" de MATLAB [14] la cual retorna 2 valores: el Rho que es el valor de la correlación como tal y el Pval que es una medida de que tan confiable es este Rho. Un Pval cercano a 0 indica que el Rho entregado es un valor confiable mientras que un Pval cercano a 1 indica todo lo contrario.

Por facilidad de nomenclatura:

- RhoP significa Rho-Pearson. PvalP significa Pval - Pearson.
- RhoS significa Rho-Spearman. PvalS significa Pval - Spearman.
- RhoK significa Rho-Kendall. PvalK significa Pval - Kendall.

### III. Resultados

El resultado de combinar la gama de bandas espectrales que posee el satélite Sentinel 2 al rededor de la planta de Puerto Mallarino como se puede apreciar en la Figura 3, con el objetivo de obtener el mayor coeficiente de correlación de Pearson da como resultado las ecuaciones (4 - 5 - 6).

$Relacion_{Bajo} =$

$$\frac{2,5 * B_2}{B_5} + \frac{3 * B_{12}}{B_6} + \frac{2 * B_{11}}{B_5} + \frac{1,5 * B_5}{B_7} + \frac{2 * B_3}{B_4} \quad (4)$$

$$Relacion_{Alto} = \frac{0,75 * B_2}{B_5} + \frac{0,75 * B_9}{B_2} + \frac{3 * B_4}{B_5} \quad (5)$$

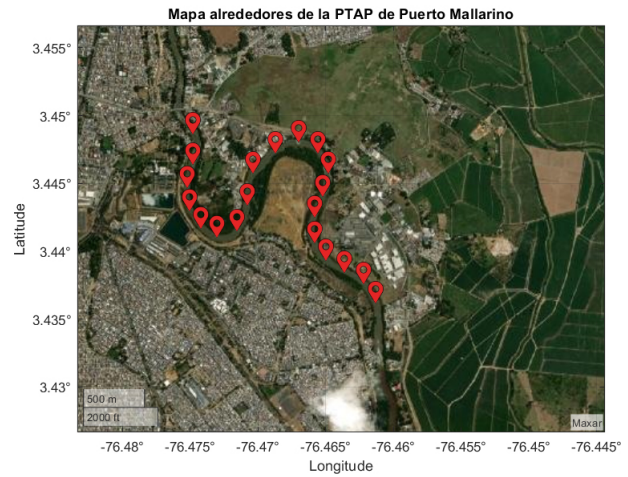


Figura 3. Puntos de referencia en el río Cauca.

$$Relacion_{Cambio} = \frac{1,75 * B_2}{B_5} + \frac{1,75 * B_6}{B_{8A}} + \frac{1,5 * B_{12}}{B_5} \quad (6)$$

$Relacion_{Bajo}$  - Combina bandas espectrales para valores bajos de turbidez.

$Relacion_{Alto}$  - Combina bandas espectrales para valores altos de turbidez.

$Relacion_{Cambio}$  - Indica si un valor de turbidez puede ser alto o bajo para su respectivo uso de relación.

En la tabla I se consignan los datos de correlación para cada relación de turbidez encontrada.

#### Cuadro I

Indices de correlación en modelo doble.

Tipo	R. Bajo	R. Alto	R. cambio
Rho - Pearson	0.7658	0.88021	0.9153
Pval - Pearson	0.0060	0.00035	0
Rho - Spearman	0.8546	0.9	0.9132
Pval - Spearman	0.0016	0.00023	0
Rho - Kendall	0.6727	0.74545	0.7345
Pval - Kendall	0.0031	0.00076	0

Un valor de turbidez se considera alto cuando es mayor a 150 NTU y un valor de turbidez bajo cuando es menor o igual a 150 NTU. Esto se da al analizar la Figura 4 donde se observa que luego de 150 NTU los valores comienzan a tener alta dispersión.

Luego se obtienen 3 relaciones de linealización para las relaciones obtenidas que se convierten en el resultado final:

$$Turbidez_{Baja} = -93,329 * Relacion_{Bajo} + 955,12 \quad (7)$$

$$Turbidez_{Alta} = -686,74 * Relacion_{Alto} + 2880,9 \quad (8)$$

$$Turbidez_{Cambio} = -513,86 * Relacion_{Cambio} + 2324,4 \quad (9)$$

El objetivo de utilizar la relación  $Relacion_{Cambio}$  y

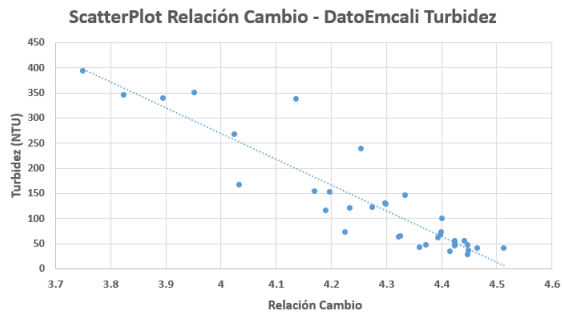


Figura 4. Relación cambio vs Datos de turbidez reales

### Cuadro II

Datos de error Modelo de Turbidez

Métrica	Modelo Turbidez
RMSE	43.1432
$R^2$ Ordinary	0.84065
$R^2$ Adjusted	0.83582
Error promedio porcentual	23.98 %

posteriormente  $Turbidez_{Cambio}$  como su nombre indica es para cambiar de un modelo de turbidez a otro, pues tener un modelo doble a un modelo unico mejora el error promedio porcentual un 21.81 %.

En la Figura 5 se compara el resultado de la predicción con los datos reales de turbidez utilizados para la construcción del modelo.

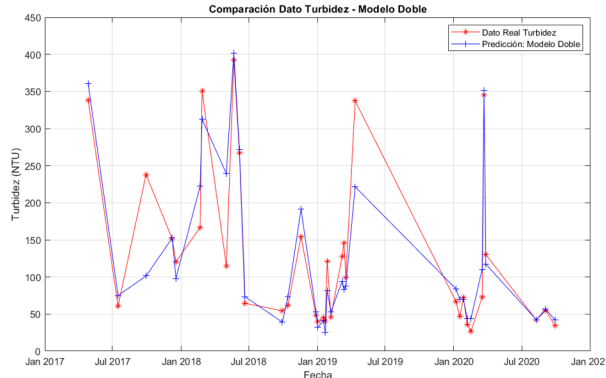


Figura 5. Comparación predicción de turbidez con dato real de turbidez.

En la tabla II se consigan las métricas utilizadas para la evaluación del modelo.

### IV. Discusión y Conclusiones

La Figura 6 indica la validación del modelo calculando la turbidez y comparandolo con los datos reales proporcionados por EMCALI hasta el año 2019.

Las métricas utilizadas para la validación del modelo se consignan en la Tabla III.

El modelo final de este trabajo se creó utilizando las imágenes satelitales del satélite Sentinel-2 del 2017 a noviembre del 2020 debido a que es en este periodo de tiempo que se tienen los datos reales de turbidez de la PTAP de Puerto Mallarino.

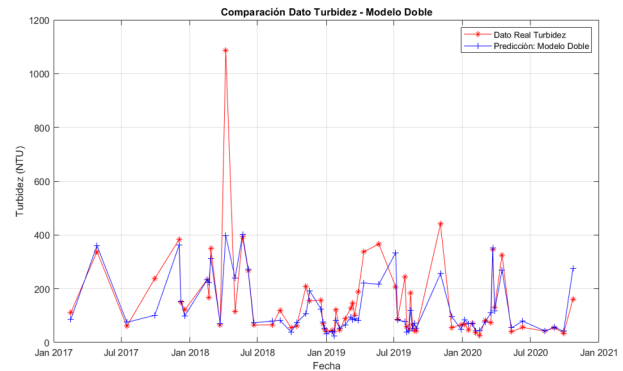


Figura 6. Validación del modelo de turbidez con los datos reales de turbidez.

### Cuadro III

Datos de error para la validación del modelo.

Métrica	Modelo Doble
RMSE	102.27
$R^2$ Ordinary	0.624
$R^2$ Adjusted	0.618
Error promedio porcentual	27.99 %

- De 262 imágenes satelitales consultadas, se eliminaron 50 por no tener contraparte con datos reales de turbidez de EMCALI.
- De 212 imágenes satelitales con dato real de turbidez, solo se utilizaron 66 para la validación del modelo, ya que tenían baja nubosidad.
- De 66 imágenes satelitales con dato real de turbidez y baja nubosidad, solo se usaron 35 imágenes satelitales para la construcción del modelo ya que presentaban la nubosidad más baja.

Para mejorar el modelo en este aspecto, se pueden agregar a la base de datos las imágenes y los datos de turbidez del año 2021 ya que no fueron utilizados para el modelo debido a que no se compartieron por parte de EMCALI. Así mismo es de resaltar que la alta nubosidad que presenta la región dificulta un poco la recolección adecuada de datos.

La poca cantidad de datos utilizados para la construcción del modelo (35 en total) es un factor que juega en contra pues para el modelo de turbidez bajo se tienen 25 datos mientras que para el modelo de turbidez alto se tienen 10 datos lo cual hace que para valores altos de turbidez el modelo tenga fallas de predicción como se observa en la Figura III. Para valores altos de turbidez el modelo tiende a desmejorar porque para la

### Referencias

- [1] Palacios, L. Pellat, F. Oropeza-Mota, J. Figueroa, Benjamín y Ménez, Mario y Ortiz, Carlos y Exebio-García, Adolfo, "Clasificador genérico de objetos en imágenes ETM +", *Agrociencia*, vol. 40, no. 5, pp. 613–626, 2006.

- 
- [2] Universidad de Murcia, “Apuntes de clase para Fundamentos físicos de la teledetección”, *Licenciatura en Geografía*, 2005 - 2006.
- [3] M. Immitzer, F. Vuolo, C. Atzberger, “First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe», *Remote Sensing*, 2016.
- [4] Earth Observing System(EOS), “Landsat 7”, Disponible en: <https://eos.com/find-satellite/landsat-7/>(Consultado: 2020/10/28), (Pagina web).
- [5] Earth Observing System (EOS), “Landsat 8”, Disponible en: <https://eos.com/find-satellite/landsat-8/>(Consultado: 2020/10/28), (Pagina web).
- [6] Earth Observing System(EOS), “Sentinel-2”, Disponible en: <https://eos.com/find-satellite/sentinel-2/>(Consultado: 2020/10/28), (Pagina web).
- [7] Earth Observing System (EOS), “SuperView1”, Disponible en: <https://eos.com/find-satellite/superview-1/>(Consultado: 2020/10/28), (Pagina web).
- [8] Earth Observing System(EOS), “Pléiades”, Disponible en: <https://eos.com/find-satellite/pleiades-1/>(Consultado: 2020/10/28), (Pagina web).
- [9] Earth Observing System (EOS), “KOMPSAT 3/3A”, Disponible en: <https://eos.com/find-satellite/kompsat-3-3a/>(Consultado: 2020/10/28), (Pagina web).
- [10] Earth Observing System(EOS), “Sistema de observación de la Tierra”, Disponible en: <https://eos.com/es/>(Consultado: 2020/10/18), (En línea).
- [11] GISGeography, “Spectral Signature Cheatsheet - Spectral Bands in Remote Sensin”, Disponible en: <https://gisgeography.com/spectral-signature/>(Consultado: 2020/10/18), (En línea).
- [12] Google Earth Engine, “Platform”, Disponible en: <https://earthengine.google.com/platform/>(Consultado: 2020/10/30), (En línea).
- [13] Land Viewer, “My LandViewer”, Disponible en: [https://eos.com/products/landviewer/user-guide/my\\_landviewer/](https://eos.com/products/landviewer/user-guide/my_landviewer/)(Consultado: 2020/10/29), (En línea).
- [14] Matlab R2021b, “Corr: Linear or rank correlation References”, Disponible en: [https://la.mathworks.com/help/stats/corr.html#mw\\_1b19e0d5-7906-4577-a0a5-b20311da7faf](https://la.mathworks.com/help/stats/corr.html#mw_1b19e0d5-7906-4577-a0a5-b20311da7faf)(Consultado: 2021/03/15).