

# Proyecto IMT 2118: Ciencia de Datos Geoespaciales

## Detección de blooms de algas mediante percepción remota como método de alerta temprana.

Parra A.<sup>1</sup>, Rodríguez D.<sup>1</sup>  
adolfo.parra@uc.cl - drodriguez@uc.cl  
<sup>1</sup>Pontificia Universidad Católica de Chile.

**Abstract** – En el presente trabajo se utilizan herramientas de percepción remota para desarrollar un método de alerta temprana para posibles cambios en los niveles de chl-a en cuerpos de agua continentales. En particular, el método se desarrolla para el Lago Villarrica, Chile. Para esto, se utilizaron 36 imágenes de los satélites MODIS y Sentinel 2, principalmente de épocas estivales, que abarcan el período del verano del año 2019 al verano del año 2022. Se ajustan modelos que utilizan esta información para aproximar los niveles de chl-a a través de distintos índices espectrales. En base a esto se presenta un método que permite aproximar los niveles de chl-a utilizando dichos índices, sin la necesidad de contar con las muestras in-situ de los niveles de chl-a.

**Keywords** – Percepción remota ; chl-a ; MODIS; Sentinel 2; Índices espectrales.

### 1.INTRODUCCIÓN

Los lagos desempeñan un papel fundamental en nuestro entorno, proporcionando una amplia gama de beneficios, desde el suministro de agua potable hasta la conservación de la biodiversidad, la promoción del turismo y la actividad económica. Sin embargo, la calidad del agua en estos ecosistemas lacustres puede verse amenazada por la proliferación de algas nocivas, conocidas como blooms de algas. Estos eventos pueden tener impactos negativos significativos en la salud humana, la vida acuática y las actividades económicas asociadas.

Para garantizar la salud y la sostenibilidad de los lagos, es crucial contar con sistemas de alerta temprana que permitan detectar y monitorear de manera efectiva los blooms de algas. En este contexto, las herramientas de percepción remota, como los satélites, desempeñan un papel clave por la posibilidad de proporcionar datos a gran escala y en tiempo casi real que permiten aproximar la distribución y la dinámica de la clorofila-a en los cuerpos de agua continentales.

La clorofila-a, el pigmento responsable de la fotosíntesis en las plantas acuáticas y algas, es un indicador fundamental para identificar la presencia y la concentración de algas. Cuando hay un bloom de algas en un lago, la concentración de clorofila-a aumenta significativamente. Esto se debe a que las algas están creciendo rápidamente y acumulando clorofila-a en sus células para realizar la fotosíntesis. Al medir la concentración de clorofila-a, se puede estimar la magnitud del bloom de algas y monitorear su evolución a lo largo del

tiempo.

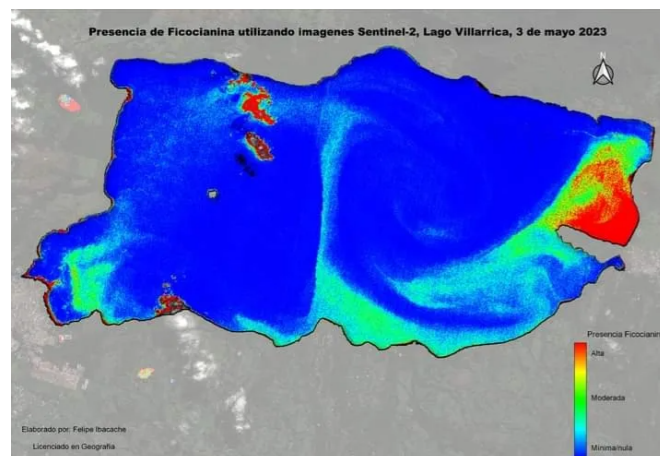


Fig. 1. Presencia de Ficocianina utilizando imágenes Sentinel-2, Lago Villarrica, 3 de Mayo 2023. Créditos: Victor Durán

La implementación de un sistema de alerta temprana basado en satélites ofrece múltiples beneficios para la gestión de los recursos hídricos y la conservación de los lagos. Permite una detección rápida de los blooms de algas, lo que facilita una respuesta inmediata y adecuada para minimizar los impactos negativos en la calidad del agua y la vida acuática. Además, proporciona información a escala regional, ayudando a identificar patrones y tendencias a largo plazo, y a tomar decisiones informadas para prevenir y controlar futuros eventos de blooms de algas.

### 2. DESCRIPCIÓN DEL PROBLEMA.

Por lo general, los sistemas de alerta temprana a través de percepción remota se implementan a través de métodos que están pensados en aproximar las concentraciones de chl-a a través de índices espectrales (a veces en conjunto con otras variables como la temperatura). Dicha aproximación se hace a través de modelos de regresión; estos métodos están planteados en base a la presencia de mediciones in-situ de chl-a. Dichos métodos han mostrado ser bastante efectivos en la aproximación de chl-a, por ejemplo a través de los niveles de NDVI (Normalized Difference Vegetation Index) y GNDVI (Green Normalized Difference Vegetation Index) (Rodríguez-Lopez et al., 2020) en el caso del Lago Laja o a través de un índice espectral de la banda verde-azul (b3b2) (Bravo Herrera J. 2021) en el caso del Lago Villarrica.

El satélite MODIS-AQUA cuenta con sensores remotos especializados que pueden detectar y cuantificar los niveles

de clorofila-a a través de un algoritmo que está ajustado en base a mediciones espectrales y mediciones in-situ, brindando información detallada sobre la distribución espacial de los blooms de algas. El problema con estas mediciones es la resolución: es del orden de 5000 metros, lo que está bien para visualizar la situación general del lago, pero no permite observar de manera clara la dinámica y la estructura del fenómeno en los cuerpos de agua. Nuestro objetivo es combinar la información obtenida del satélite MODIS-AQUA con distintos índices calculados a través de Sentinel 2, cuya resolución es del orden de las decenas de metros, para ajustar un modelo que permita aproximar los niveles de concentración de chl-a en el área de estudio.

Es importante entender que la relevancia de cuantificar la relación entre los índices espectrales y los niveles de chl-a radica en que los índices que deben ser utilizados (o que pueden ser utilizados) para la aproximación de chl-a y el cómo deben ser utilizados varía de cuerpo en cuerpo de agua. Dependerá de las características de cada uno: turbiedad, nivel de nutrientes, etc.

En el presente estudio se busca entonces, como objetivo general, desarrollar un método para la aproximación de chl-a sin la necesidad de muestras in-situ previas, y utilizar dicho método para evaluar la evolución temporal de los blooms de algas; y como objetivos particulares, detectar la relación existente entre índices espectrales y en base a estas ajustar un modelo de regresión para la aproximación de los niveles de chl-a.

El área de estudio seleccionada para llevar a cabo este estudio es el Lago Villarica, ubicado en la Región de la Araucanía, Chile, con coordenadas 39°1500S 72°0500O. Las razones para elegir este cuerpo de agua son fundamentalmente 2: la existencia de estudios previos en este cuerpo de agua (Bravo Herrera J., 2021), de los cuales se obtienen marcos de referencia para el desarrollo de nuestro estudio, y la presencia de blooms de algas en el lago durante los últimos años, que lo convierte en un cuerpo de agua de interés para llevar a cabo un estudio de este tipo.

#### 4. DESCRIPCIÓN DE LOS DATOS.

A continuación se detallan los datos utilizados en este estudio:

- Se ocuparon datos respectivos a la geometría del Lago Villarica. Estos fueron obtenidos de la Dirección General de Aguas. Con estos datos se logra identificar el área de estudio y así focalizar la búsqueda del resto de datos necesarios.
- Se ocuparon imágenes satelitales de Sentinel 2, las cuales tienen una resolución entre 10 a 60 metros, dependiendo de la banda. Estos son utilizados para el cálculo de los índices considerados, los cuales se especifican en las figuras 1 y 2. El tiempo de revisita de este satélite es de 5 días, y se utilizaron las imágenes con mínima nubosidad. La disponibilidad de estas imágenes para el área de estudio es desde fines de 2018 hasta la actualidad.

- Se ocuparon imágenes satelitales de MODIS-AQUA, las cuales tienen una resolución que ronda los 5000 m. En particular, se utiliza la banda de chl-a, que es uno de los componentes fundamentales de este estudio. El tiempo de revisita de este satélite es diario. La disponibilidad de estas imágenes para el área de estudio abarca toda la disponibilidad de Sentinel 2. Según documentación, la disponibilidad es desde julio de 2002 hasta febrero de 2022.

Algoritmo	Nombre	Expresión	Referencia
Ocean Chlorophyll OC2	OC2	$10^{(0,139-2,336R+0,879R^2-0,135R^3)} - 0,07$ $R = \log_{10}(B2/B3)$	O'Reilly, 2000
Ocean Chlorophyll OC4	OC4	$10^{(0,366-3,067R+1,930R^2+0,649R^3-1,532R^4)}$ $R = \log_{10}(\max(B1, B2)/B3)$	O'Reilly, 2000
Azul-Verde	b3b1	$B3/B1$	Ha, Thao et al., 2017
	b3b2	$B3/B2$	Ha, Thao et al., 2017
Verde-Rojo	b3b4	$B3/B4$	Ha, Thao et al., 2017
Rojo-NIR	b5b4	$B5/B4$	Ha, Thao et al., 2017
	b6b4	$B6/B4$	Ha, Thao et al., 2017
	b7b4	$B7/B4$	Ha, Thao et al., 2017
	b8ab4	$B8A/B4$	Ha, Thao et al., 2017
	b8b4	$B8/B4$	Ha, Thao et al., 2017

Fig. 2. Índices incluidos por referencia de Bravo H. Javiera, 2021.

INDEX	EQUATION
Enhanced Vegetation Index (EVI)	$EVI = G \times ((nir - red)/(nir + C1 \times red - C2 \times blue + L_{cvi}))$
Normalized Ratio Vegetation Index (NRVI)	$NRVI = (red/nir - 1)/(red/nir + 1)$
Normalized Difference Water Index (NDWI)	$NDWI = (green - nir)/(green + nir)$
Normalized Difference Water Index (NDWI2)	$NDWI2 = (nir - swir2)/(nir + swir2)$
Modified Normalized Difference Water Index (MNDWI)	$MNDWI = (green - swir2)/(green + swir2)$
Green Normalized Difference Vegetation Index (GNDVI)	$GNDVI = (nir - green)/(nir + green)$
Normalized Difference Vegetation Index (NDVI)	$NDVI = (nir - red)/(nir + red)$

Fig. 3. Índices incluidos por referencia de Markogianni, V. et al, 2018.

#### 5. DESARROLLO.

Se obtienen un total de 36 imágenes de Sentinel 2 y MODIS-AQUA que abarcan desde la época estival de 2019 hasta la época estival de principios de 2022. Dichas imágenes corresponden a días de poca nubosidad en donde ambos satélites tienen una imagen del área de interés.

Dado que la intención es asociar los valores de los índices espectrales con los valores de chl-a captadas en las imágenes de MODIS-AQUA, se debe de alguna forma obtener un valor para estos índices en la misma resolución espacial del satélite MODIS-AQUA. Para esto se decidió hacer una poligonización de las imágenes de chl-a obtenidas, operación que no es costosa dada la baja resolución de dichas imágenes. Dicha poligonización se realizó por medio de gdal.

Uno de los puntos claves a considerar en este punto es que con la poligonización a través de gdal se truncarán los valores de chl-a en cada celda del raster (por ej., de 4.5 a 4). Así funciona la poligonización por defecto en gdal. Como autores no encontramos la forma de evitar que gdal haga dicha

aproximación. Una vez truncados los valores, los polígonos devueltos por gdal corresponderán en verdad a 1 polígono por valor encontrado, donde dicho polígono será la unión de los polígonos que compartían dicho valor.

Para cada imagen polygonizada, y para cada uno de sus polígonos, se necesita entonces obtener las estadísticas zonales de los índices espectrales, es decir, asociar a cada uno de esos polígonos un valor para cada índice espectral; en este caso, tomaremos la media del raster de cada índice clipeado para cada polígono.

Así, al final de este proceso, tendremos una tabla por imagen que contendrá para cada uno de sus polígonos el valor del chl-a observado y los respectivos índices. Estas tablas son unidas para obtener los datos finales.

Luego, se procedió con la limpieza de datos. Uno de los problemas de nuestro método es que algunas de las imágenes del MODIS-AQUA abarcan en gran porcentaje una región que no pertenece al lago (por su baja resolución), por lo que los valores de chl-a no son los que realmente interesan a nuestro estudio. Se debe ser cauteloso en este punto y escoger aquel rango de valores de 'chl-a' que puedan asegurar cierta densidad de valores en un rango aceptable, es decir, que podamos tener cierta certeza de que esos valores corresponden a mediciones de dentro del lago. Descartado esto, esta fase consistió principalmente en la detección de valores outliers para los índices y la remoción de valores nulos.

En la **Tabla 1** se presentan las correlaciones presentes entre los índices espectrales considerados y las mediciones remotas de chl-a luego de la limpieza de los datos.

Índice Espectral	Correlación chl-a
ndvi	-0.601
gndvi	-0.588
ndwi	0.588
mndwi	0.586
b3b4	0.363
ndwi2	0.352
b3b2	0.342
oc4	0.334
b7b4	0.278
evi	-0.266
b3b1	0.218
b8b4	0.148
b6b4	-0.096
b5b4	0.079
b8ab4	-0.056

**TABLE I**

Correlación entre los distintos índices espectrales considerados y las mediciones remotas de chl-a

Con estos datos se ajustó un modelo de regresión lineal, que entrega los resultados que se muestran en la Figura 6.

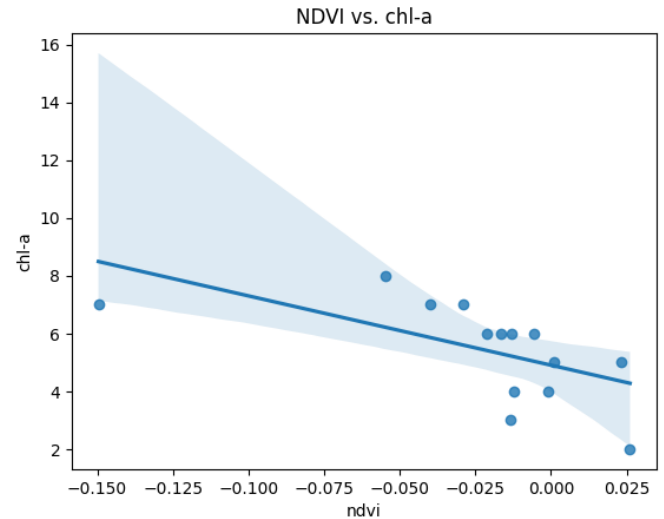


Fig. 4. NDVI vs. chl-a

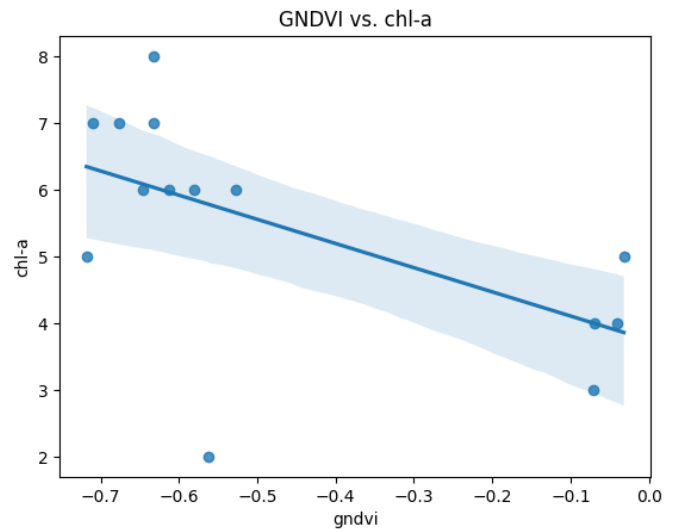


Fig. 5. GNDVI vs. chl-a

REGRESSION				
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES				
-----				
Data set	:	unknown		
Weights matrix	:	None		
Dependent Variable	:	chl-a		
Mean dependent var	:	5.4286		
S.D. dependent var	:	1.6968		
R-squared	:	0.5474		
Adjusted R-squared	:	0.4651		
Sum squared residual	:	16.942		
Sigma-square	:	1.540		
S.E. of regression	:	1.241		
Sigma-square ML	:	1.210		
S.E. of regression ML	:	1.1001		
-----				
Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	3.7298849	0.6688208	5.5768074	0.0001660
gndvi	-2.7724848	1.3059774	-2.1229194	0.0572792
ndvi	-18.6954084	8.4512242	-2.2121539	0.0490336
-----				

Fig. 6. Resultados de la regresión lineal ajustada en base a los índices GNDVI y NDVI para aproximar la concentración de chl-a.

## 6. RESULTADOS Y CONCLUSIONES.

Si bien los resultados de la regresión no son los mejores, si parece ser una posibilidad para la estimación de los niveles de chl-a. Dada la poca cantidad de datos que tenemos a nuestra disposición, no fue viable hacer validación basada en estos datos. Una posible mejora a este método es la utilización de algún satélite que tenga una disponibilidad histórica mayor a Sentinel 2.

Se utiliza el modelo obtenido para obtener una predicción de los niveles de chl-a basada en los índices espectrales de GNDVI y NDVI. Esta predicción tendrá una resolución de 10m. A modo ilustrativo, se realiza en particular la predicción para el día 03 de mayo de 2023, día para el que se cuenta con una imagen en alta resolución del bloom de algas publicada por la NASA. Compararemos los resultados de la aproximación de los niveles de chl-a hecha por el modelo con el aspecto visual del Lago Villarica en ese día.

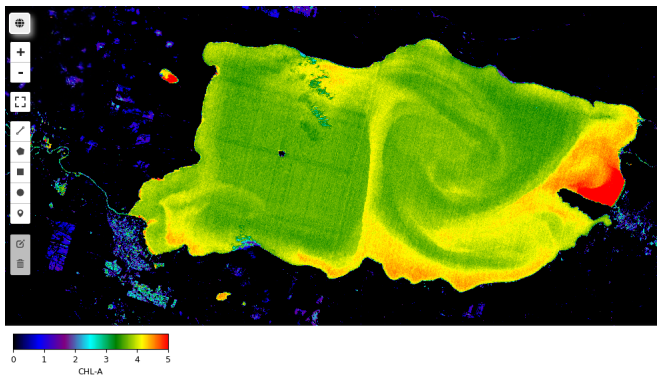


Fig. 7. Aproximación de la concentración de chl-a hecha por el modelo de regresión para el día 3 de mayo de 2023.

La siguiente imagen fue publicada por la NASA respecto del bloom de algas en el Lago Villarica; la imagen corresponde al día 3 de mayo de 2023.



Fig. 8. Imagen del Lago Villarica del día 3 de mayo de 2023, publicada por la NASA [8].

El modelo parece captar bien la estructura y dinámica del fenómeno de bloom de algas, un resultado bastante esperanzador tomando en cuenta que la relación encontrada

entre los índices espectrales utilizados y los niveles de chl-a no es trivial.

El modelo tiende a sobreestimar los valores de chl-a. Tiene un valor base que no parece ser representativo de la realidad. Sin embargo, sí parece ser capaz de captar los incrementos sobre un nivel normal de chl-a.

El modelo fue utilizado para observar la evolución de dos eventos de blooms de algas en el Lago Villarica, que ocurrieron en Enero y Mayo de 2023. En ambos casos, antes del fenómeno se puede observar un color verde parejo indicativo del valor base que capta nuestro método. Cuando ocurre el fenómeno de bloom de algas, nuestro método es capaz de identificar tanto la estructura como la dinámica de este fenómeno, aunque se desconoce la precisión con que los valores de chl-a son aproximados.

En conclusión, el método propuesto permite analizar la relación entre distintos índices espectrales calculados a través de Sentinel 2 y la medición remota de chl-a a través de MODIS-AQUA, con el fin de poder establecer un modelo de regresión lineal para aproximación de los niveles de chl-a sin la necesidad de muestras in-situ anteriores.

Este método podría ser utilizado para explorar la situación de lagos de difícil acceso o en situaciones donde simplemente no se cuentan con los recursos necesarios para acceder a mediciones in-situ de chl-a (como en nuestro caso).

Algunos aspectos dudosos del método planteado es si acaso la correlación captada en los índices espectrales respecto de las concentraciones de chl-a obtenidas por percepción remota tiene que ver con el algoritmo que utiliza el satélite MODIS-AQUA para la obtención de la banda de clorofila-a.

También, cabe la posibilidad de que el satélite MODIS-AQUA sea eficiente aproximando los niveles de chl-a en el Lago Villarica particularmente, pero no lo sea tanto en lagos de otras características. Queda pendiente verificar si la replicación del método sobre un lago en condiciones totalmente distintas ofrece resultados favorables.

De cualquier forma, el método planteado parece ser una buena herramienta exploratoria para verificar la situación de cuerpos de agua respecto de la presencia de algas, sobretodo teniendo en cuenta de que la utilización de uno u otro índice para explorar dicha situación no es siempre clara.

## REFERENCES

- [1] Bravo Herrera, J. (2021). Diagnóstico de la condición trófica del Lago Villarica utilizando imágenes satelitales. Disponible en <https://repositorio.uchile.cl/handle/2250/181628>
- [2] Llamam a prevenir exposición con "bloom" de algas en Lago Villarica. (2022). Radio Bio Bio. Recuperado de <https://www.biobiochile.cl/noticias/nacional/region-de-la-araucania/2020/01/25/llamam-a-prevenir-exposicion->

con-bloom-de-algas-en-lago-villarrica.shtml

- [3] Lien Rodríguez-López, Iongel Duran-Llaser, Lisdelys González-Rodríguez, Rodrigo Abarca-del-Rio, Rolando Cárdenas, Oscar Parra, Rebeca Martínez-Retureta, Roberto Urrutia, Spectral analysis using LANDSAT images to monitor the chlorophyll-a concentration in Lake Laja in Chile, *Ecological Informatics*, Volume 60, 2020, 101183, ISSN 1574-9541, <https://doi.org/10.1016/j.ecoinf.2020.101183>.
- [4] Markogianni, V., Kalivas, D., Petropoulos, G., Dimitriou, E. (2018). An Appraisal of the Potential of Landsat 8 in Estimating Chlorophyll-a, Ammonium Concentrations and Other Water Quality Indicators. *Remote Sensing*, 10(7), 1018. <https://doi.org/10.3390/rs10071018>
- [5] Algae in the Andes. (2023). <https://earthobservatory.nasa.gov/images/151363/algae-in-the-andes>