Physicochemical Characterisation of Pit Lakes Using Google Earth Engine: Chilean Case Study

Damián Cruz Barrientos[1], Oscar Benavente Zolezzi[2,3], Álvaro Navarrete Calvo[1], Camilo Sánchez Yáñez[1]

[1]*Universidad Mayor, Escuela de Geología, Cam. La Pirámide 5750, Huechuraba, Santiago, Chile, damian.cruz@mayor.cl.*

2SRK Consulting – Denver, GeoEnvironmental Group, 1125 17th St Ste 600, Denver, USA, obenavente@srk.com.

3SRK Consulting Spa – Chile, Cierre de Minas y Geoquímica, Av. Vitacura 2939, Piso 5, Las Condes, Santiago, Chile, obenavente@srk.com.

Abstract

This study uses satellite imagery and the HSV color model to assess pit lake water quality. Water color results from sunlight interactions influenced by dissolved ions and suspended particles. Transition metals form distinct complexes (Fe^{2+}: olive green, Cu^{2+}: blue-green, Fe^{3+}: brown-yellow), while oxyhydroxide colloids scatter specific wavelengths (Fe: red-orange, Al: white). Because pH and redox conditions affect ion complexation and mineral phase formation, HSV variations can reflect physicochemical changes. Validation at Berkeley Pit Lake showed correlations between satellite-derived HSV and known events. The method is being applied to Chilean pit lakes for environmental monitoring, aiding early detection of environmental changes.

Keywords: Mine Closure Plan, Pit Lakes Monitoring, Satellite Imagery, Acid Mine Drainage, Water Quality

Introduction

A large portion of the environmental issues linked to mining stem from the waste generated during extraction, with pit lakes being among the most challenging mining legacies to manage. These water bodies are prone to physicochemical changes due to prolonged exposure to minerals like pyrite, which can lead to acidic conditions and elevated sulfate and trace element concentrations in the pit lake and nearby water bodies. This study aims to develop a robust methodology and numerical dataset to analyze, monitor, and determine the physicochemical characteristics of pit lakes using satellite imagery.

The innovation of this research lies in its development of a cost-effective methodology for monitoring pit lakes through the color of the water in the HSV (Hue, Saturation, Value) model. Several studies have shown that the true color of the water can indicate the hydrochemical conditions of water bodies. For instance, Murphy *et al.*, (2018), and Ohsawa *et al.*, (2009) found that dissolved Fe^{2+} gives water a green olive hue due to light absorption, while native sulfur particles scatter a blue hue in volcanic crater lakes. Castellón *et al.*, (2013) observed that aluminium colloids influence water color by reflecting light in all visible wavelengths (i.e., whitish color to the water), producing the unusual sky-blue color of Celeste River in Costa Rica. Additionally, the reddish color of iron oxides, oxyhydroxides, and hydroxides is a well-known characteristic of these mineral phases, used as pigments since prehistoric times (Torrent & Barrón, 2002). These findings suggest that the physicochemical characteristics of pit lakes can be inferred from satellite imagery, as mineral speciation and ion complexation are pH and redox-dependent.

Time series of HSV values for each pit lake under study were calculated using surface reflectance data from Landsat and Sentinel-2 missions, processed via the Google Earth Engine with Python. By correlating the H, S and V of the water with physicochemical properties such as pH, redox, metal concentrations and stratification states. This methodology was validated using the publicly available dataset from Berkeley Pit Lake, Montana, USA, compiled by INAP in the Pit Lakes Database. This acidic pit lake has historical water quality records dating back to the 1980s, overlapping with publicly available satellite data from Landsat 4, 5, 7, 8, and 9, and Sentinel-2 satellites. Following this validation, the methodology was applied to Chilean pit lakes with data in situ identified through web scraping, allowing for a broader assessment of its applicability in diverse mining environments.

Method

Satellite imagery from Landsat 4, 5, 7, 8, 9, and Sentinel-2, stored in Google Earth Engine (GEE), was analysed to determine the true colour of pit lakes.

The process starts with the identification and delineation of the pit lake under study over multi scenes. To do so, a two-steps method involving the uses of Normalised Difference Water Index (NDWI) and a supervised machine learning technique, Random Forest, were employed.

The first step corresponds to the identification and segmentation of the pit lakes using the NDWI histogram, calculated as (GREEN − NIR) / (GREEN + NIR), and the OTSU algorithm (Otsu, 1979). Metadata, including acquisition date, scene name, solar azimuth, cloud cover percentage, and water surface area, was extracted and stored in an Excel database for each scene. Scenes without detected water were assigned an NDWI value of 10, and each scene was validated manually by assigning 1 for correct mapping and 0 for incorrect mapping. This process is applied to each of the six satellite collections, independently.

Validated scenes (i.e., scenes where the NDWI threshold value found by the OTSU algorithm properly identifies water and non-water pixels) were then used as training cases of a Random Forest algorithm. This supervised machine learning technique uses all the available bands in each satellite collection to derive a suitable algorithm capable to distinguish water and non-water pixels. An optimum number of seven validated scenes are recommended to be used in the training process with sufficient temporal spacing among them to capture possible changes in colour and surface variations of the water body. This process is also applied to each of the six satellite collections independently.

After identifying and segmenting the pit lakes, water pixels were processed to extract HSV color values by converting RGB to HSV (Fig. 1), enabling a more detailed interpretation of lake color variations, following the method used in Murphy *et al*., (2018).

A grid of points was generated on the lake to extract HSV values and then were stored in an Excel template for inter-satellite comparisons, compiling the following dataset:

- Mean HSV values across all points
- HSV at the lake's central point (determined by Euclidean distance)
- 5th, 50th, and 95th percentile values of HSV

For a better representation of hue (0–1), a polar histogram was used, where hue values

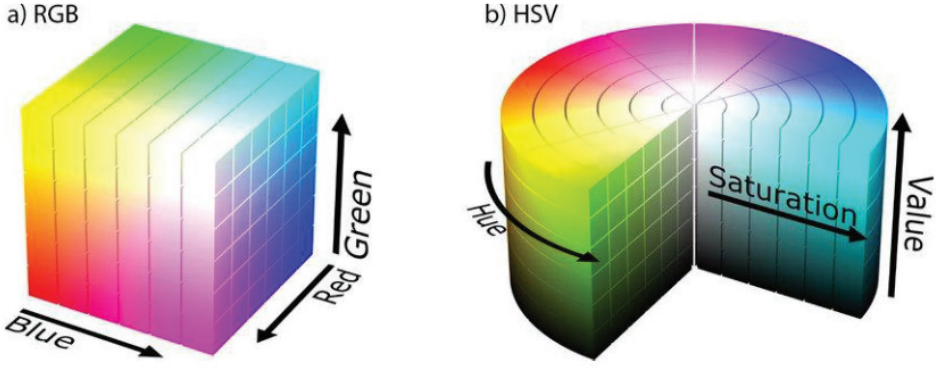

***Figure 1** RGB and HSV model color (Extracted from Saraullo* et al.*, 2019).*

Table 1 Summary of satellite images analysed and classification accuracy of NDWI and Random Forest.

Remote Sensor	Landsat 5	Landsat 7	Landsat 8	Landsat 9	Sentinel 2	Total images
Images Evaluated	159	610	52	107	235	1163
NDWI Correct (%)	26 (16.35%)	41 (6.72%)	10 (19.23%)	12 (11.21%)	103 (43.82%)	192 (16.5%)
RF Correct (%)	125 (78.61%)	599 (98.19%)	36 (69.23%)	48 (44.85%)	162 (68.93%)	970 (83.4%)

were transformed by multiplying by 360. The transformed dataset was subjected to QA/QC filters to improve data reliability and select the most representative images of each pit lake.

The validation process was conducted using the INAP Pit Lakes Database, which provided access to comprehensive historical data. Berkeley Pit was selected as the validation site due to its large surface area, extensive temporal data in situ, and documented studies of its evolution. Hydrochemical modelling of Berkeley Pit Lake data using PHREEQC allowed for an improved understanding of the redox conditions, mineral saturation indices and co-precipitation processes affecting water colour dynamics.

Once validated, the methodology was applied to Chilean pit lakes with available public data. To identify these lakes, a Python-based web scraping script was developed to extract relevant information from the SEA (Servicio de Evaluación Ambiental) website. The validated methodology was then applied to these sites, ensuring a systematic remote sensing approach for assessing physicochemical characteristics in pit lakes across Chile.

Results

Validation – Berkeley Pit Lake Case

The methodology was validated using satellite imagery of Berkeley Pit Lake. A total of 1163 images were found for Berkeley Pit Lake between May 1984 to September 2024, with NDWI correctly identifying water bodies in only 16.5% of the total cases. In contrast, the Random Forest algorithm achieved a significantly higher accuracy of 83.4% (Tab. 1).

To ensure an accurate representation of pit lake color, a QA/QC filtering process was applied to the 1163 images. The first filter, based on mean color variation below 35°, removed 433 images, leaving 730 valid scenes. The second filter, applying a dispersion coefficient threshold ($R > 0.95$), further reduced the dataset to 439 images. The third filter, incorporating manual area mapping with a 20% error margin, gives in a final dataset of 311 images for analysis (Tab. 2).

The evolution of HSV reveals the occurrence of four stages (Fig. 2), which are closely associated with changes in hydrochemistry over time:

- **Stage 1 (1984–1988):** Green olive hues ($H = 0.25$, $S = 0.35$, $V = 0.05$) indicate a dominant presence of Fe^{2+}, with more dissolved iron suggesting the onset of acid conditions due to groundwater contact with sulfide rocks. This phase exhibits a meromictic state, defined by the establishment of a deep chemocline.
- **Stage 2 (1988–1997):** Increase of the Fe^{3+}/Fe^{2+} ratio due to oxidation, changing the color of the water to yellow-brown ($H = 0.15$, $S = 0.35$, $V = 0.04$). This transformation is supported by an increase in the pe values of the water while maintaining an almost constant pH.
- **Stage 3 (1997–2013):** The introduction of alkaline tailings sludge into the pit lake during this stage resulted in the co-precipitation of Fe-oxyhydroxides and Fe-oxyhydroxide-sulfate. This process is

Table 2 QA/QC filtering process for satellite images used in the evaluation.

Remote Sensor	Landsat 5	Landsat 7	Landsat 8	Landsat 9	Sentinel 2	Total images
Images Evaluated	159	610	52	107	235	1163
1st QA/QC Filter	131	451	30	49	69	730
2nd QA/QC Filter	72	241	29	39	58	439
3rd QA/QC Filter	54	169	23	19	46	311

evident from the saturation indexes of Schwertmannite, Goethite, Jarosite and Ferrihydrite, and the colloidal particles of these minerals produced scattering effects that shifted the lake's hues towards red-orange and lower V values observed in this period (H = 0.07, S = 0.45, V = 0.03). Due to the alkaline tailings disposal the lake returned to a holomictic state in 2010, with a rise in pH from 2.5 to 4 and a decline in Fe concentrations from 1000 mg/L to 0.5 mg/L, accompanied by the disappearance of the chemocline. These changes were further corroborated by the vertical homogeneity of the water column, indicating the absence of chemical stratification.

- **Stage 4 (2013–Present):** The most recent stage is characterised by a marked increase in H, S, and V values (H = 0.45, S = 0.8, V = 0.1), reflecting near-complete co-precipitation of Fe-oxyhydroxides and Fe-oxyhydroxide-sulfates, and the initiation of the co-precipitation of Al-oxyhydroxides and Al-oxyhydroxide-sulfates (Jurbanite – Alunite). This transition alters the light reflectance properties of the pit lake, as Al-oxyhydroxides-sulfates scatters the lights in all the visible wavelengths (increase in S and V), enhancing the color given by both dissolved Fe^{2+} and Cu^{2+}.

Application on Chilean Pit Lakes

Currently, Chile does not have an official registry of pit lakes. In this study, a numerical

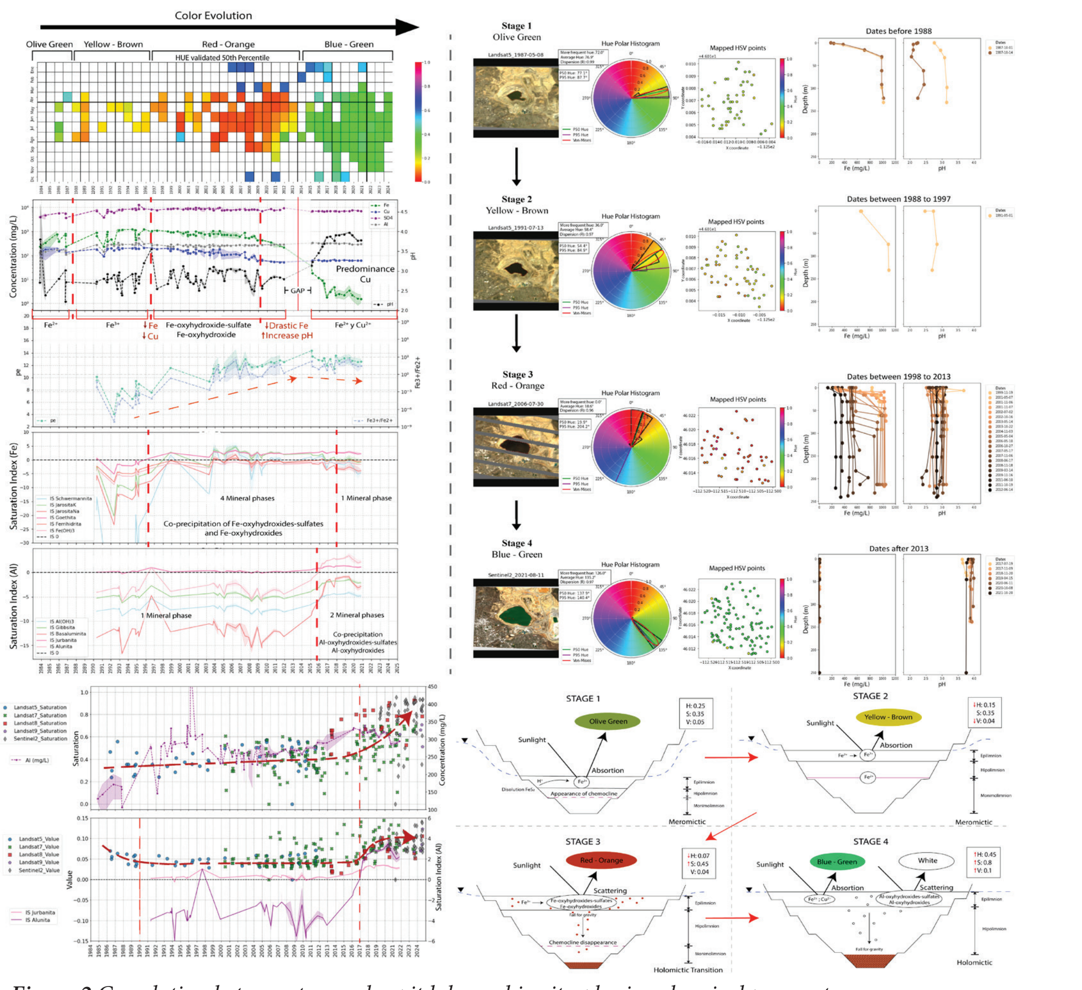

Figure 2 Correlation between true-color pit lake and in situ physico-chemical parameters.

database was established, identifying a total of 111 pit lakes across 46 mining sites. At these 46 mining sites, a web scraping approach was applied to search for available environmental data. This process identified 23 pit lakes (20.4%) with reported physicochemical parameters, allowing the methodology to be applied to those with in situ data (Tab. 3).

The analyzed pit lakes predominantly exhibit sulfate-calcium water types, with pH values ranging from acidic to neutral-alkaline (1–9), electrical conductivity (500 –140000 µS/cm), and concentrations of SO_4 (40–185000 mg/L), Cu (0.001–4000 mg/L), Fe (0.001–4300 mg/L), and Al (0.005–7800 mg/L).

Most pit lakes in Chile have an olive green hue (H = 0.3), indicating the presence of dissolved Fe^{2+}. In some cases, on limited dates, the color changes to blue-green, suggesting the presence of dissolved Cu^{2+}, typically associated with neutral to alkaline pH conditions (Fig. 3).

In two specific cases (Mina Sur and Quebrada Blanca) the color of the pit lake was initially olive green, then changed to a yellow-brown and finally to a red-orange. Given the red-orange color, this indicates the presence of colloids Fe hydroxides, suggesting that the pH conditions are acidic (Fig. 3).

The application of this method made it possible to determine the time of occurrence of the pit lake (i.e., “Cerro Blanco”), its seasonal occurrence being only between September and February. It also made it possible to determine backfill times in different pit lakes (i.e., “Andacollo”).

One of the limitations of the application of this methodology in Chilean pit lakes is the presence of shadows projected on the lake. This effect is particularly pronounced in winter due to the steep slopes of the Andean range and the low solar azimuth, which complicates the accurate segmentation of water bodies, as observed in the case of “El Soldado” (Fig. 4).

Another limitation is the low surface area of the pit lakes, which results in poor segmentation using random forest, due to the low number of pixels representing the water body. Therefore, for a reliable representation of the data, a minimum threshold of 10 pixels is recommended, resulting in an area of approximately 9000 m^2 in Landsat images and

Table 2 Summary of data collected through web scraping of pit lakes in Chile.

Mining Site	Type of deposit	Number of samples	Data type	No. of pit lakes with data
Andacollo	Copper Porphyry	4	Tables	4
Candelaria	IOCG	108	Database	1
Cerro Blanco	Calcium Carbonate (Others)	1	Database	1
Cerro Colorado	Copper Porphyry	12	Modeling	2
Chuquicamata	Copper Porphyry	4	Database	1
Collahuasi	Copper Porphyry	39	Database	1
El Salvador	Copper Porphyry	16	Writings and Laboratory Certificates	3
El Soldado	Stratabound	3	Laboratory Certificates	1
La Coipa	Porphyry / Epithermal Au	4	Modeling	1
Los Bronces	Copper Porphyry	4	Database	1
Mina Sur	Copper Porphyry	9	Database	1
Quebrada Blanca	Copper Porphyry	54	Database	3
Santo Domingo	Stratabound	2	Tables	2
Zaldívar	Copper Porphyry	3	Database	1
Total	**5**	**260**	**5**	**23**

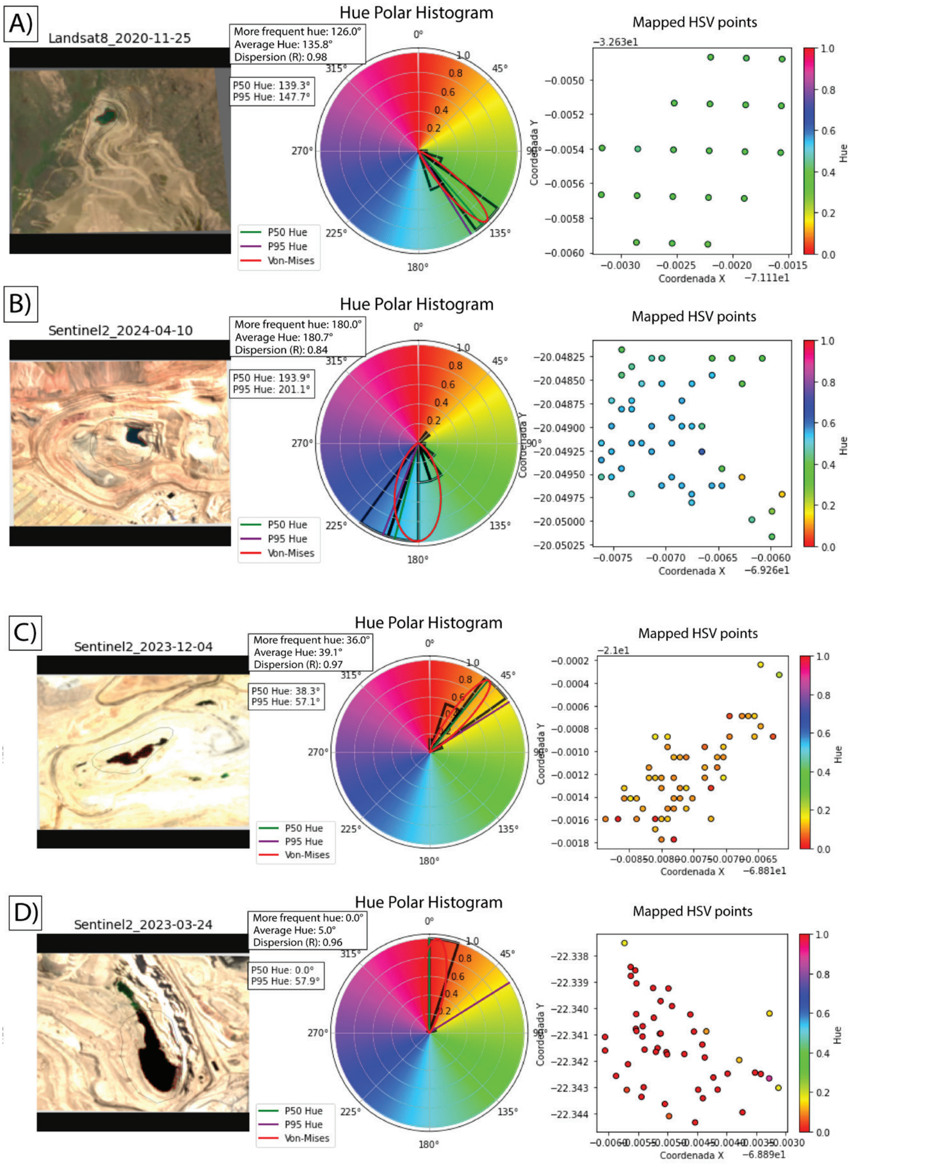

Figure 3 *Examples of olive green, blue-green, yellow-brown and red-orange colors. A) El Soldado, B) Cerro Colorado, C) Quebrada Blanca, D) Mina Sur.*

1000 m² in Sentinel-2 images. This limitation arises from the low spatial resolution of the different satellites.

Conclusions

This study emphasizes the use of remote sensing for the physicochemical characterisation of pit lakes, utilizing true color analysis using the HSV model and the Random Forest machine learning algorithm. The use of QA/QC filters improves the segmentation of pit lakes in satellite images, certifying a correct color representation. The validation was performed at Berkeley Pit Lake, as it contains extensive historical data, has many studies, and has a large area. It also demonstrated that physicochemical characteristics can be inferred from the absorption of sunlight by dissolved chemical elements such as Fe^{2+} (Olive Green), Cu^{2+} (Blue-Green), and Fe^{3+} (Yellow-Brown). Mineral colloids such as

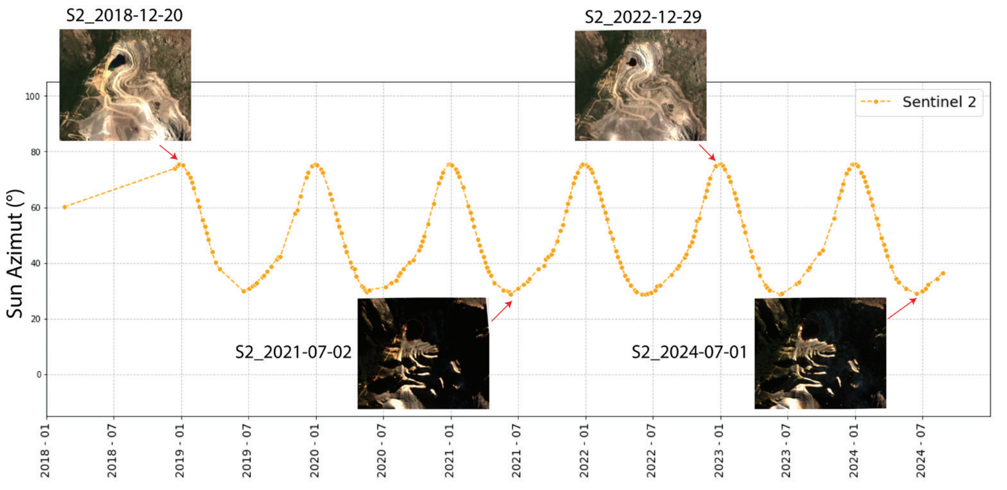

***Figure 4** Presence of shadows that affected the segmentation of the pit lake by the azimuth of the Sun in "El Soldado".*

Fe and Al oxyhydroxide-sulfates also affect the color spectrum, as Fe-oxyhydroxide-sulfates disperse the orange-red color, while Al-oxyhydroxide-sulfate colloids influence the color of water by reflecting light at all visible wavelengths, increasing S and V in the HSV model.

This methodology has been successfully applied in Chilean pit lakes, although its main limitation lies in the spatial resolution of Landsat and Sentinel images, particularly for smaller pit lakes. To improve monitoring, the use of hyperspectral drones is recommended to obtain higher resolution images, allowing for more accurate analysis of surface waters. Continuous monitoring with the HSV model offers a cost-effective alternative to in situ methods, facilitating long-term environmental assessment and improving pit lake management.

Acknowledgements

Thanks to INAP for granting access to the Pit Lake Database, which was essential for the validation of this methodology.

References

Castellón, E., Martínez, M., Madrigal-Carballo, S., Arias, M. L., Vargas, W. E., & Chavarría, M. (2013). Scattering of Light by Colloidal Aluminosilicate Particles Produces the Unusual Sky-Blue Color of Río Celeste (Tenorio Volcano Complex, Costa Rica). PLoS ONE, 8(9). https://doi.org/10.1371/journal.pone.0075165

Murphy, S., Wright, R., & Rouwet, D. (2018). Color and temperature of the crater lakes at Kelimutu volcano through time. *Bulletin of Volcanology*, 80(1). https://doi.org/10.1007/s00445-017-1172-2

Ohsawa, S., Saito, T., Yoshikawa, S., Mawatari, H., Yamada, M., Amita, K., Takamatsu, N., Sudo, Y., & Kagiyama, T. (2009). Color change of lake water at the active crater lake of Aso volcano, Yudamari, Japan: Is it in response to change in water quality induced by volcanic activity? *Limnology*, 11(3), 207–215. https://doi.org/10.1007/s10201-009-0304-6

Otsu, N. (1979). *A threshold selection method from gray-level histograms. Automatica*, 11(285-296), 23–27.

Saraullo, F. A. V., Larosa, F. S., Ghignone, R. A., & Lanzillotta, L. (2019). *Diseño, implementación y ensayo de un lector de colores de bajo costo para personas ciegas y disminuidas visuales*.

Torrent, J., & Barrón, V. (2002). *Diffuse reflectance spectroscopy of Iron Oxides*.