Pipeline para construir un dataset multitemporal de temperatura en superficie e índices espectrales a partir de sensores satelitales para monitoreo volcánico.

Hernán G. Rosenfeld[1,2], Guillermo P. Toyos[2,3], Mariano R. Agusto[1,2], and Juan B. Cabral[2,4,5]

[1] Instituto de Estudios Andinos "Don Pablo Groeber" (IDEAN), Universidad de Buenos Aires (UBA), Buenos Aires, Argentina
[2] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
[3] Comisión Nacional de Actividades Espaciales (CONAE), Buenos Aires, Argentina
[4] Gerencia de Vinculación Tecnológica (GVT), CONAE, Centro Espacial Teófilo Tabanera, Córdoba, Argentina
[5] Facultad de Matemática, Astronomía, Física y Computación (FaMAF), Universidad Nacional de Córdoba (UNC), Córdoba, Argentina
hrosenfeld@gl.fcen.uba.ar, gtoyos.sec@conae.gob.ar, magusto@gl.fcen.uba.ar, jbcabral@unc.edu.ar

Abstract. Este trabajo presenta un pipeline para la descarga, procesamiento y análisis de imágenes satelitales con el objetivo de construir series temporales de temperatura superficial e índices espectrales en volcanes activos. El dataset resultante permite caracterizar el comportamiento térmico a largo plazo y puede contribuir al monitoreo volcánico. Como primer paso, nos enfocamos en la constelación Landsat, debido a su extensa historia operativa y la disponibilidad pública de sus datos. El pipeline comprende la definición de regiones de interés, el acceso y descarga automatizada de imágenes Landsat 7, 8 y 9, la extracción posterior de bandas espectrales y productos de alto nivel, y finalmente la construcción de las series temporales. Implementamos este pipeline en el lago cratérico del volcán Tupungatito (Mendoza, Argentina) y generamos una serie temporal de temperaturas superficiales calibradas a partir de 837 escenas Landsat. El workflow extrae índices espectrales y utiliza metadatos de control de calidad para filtrar escenas inválidas. Dado que entre 2000 y 2025 los datos fueron adquiridos por distintos sensores Landsat, realizamos una calibración para asegurar la consistencia de la serie temporal de temperaturas superficiales máximas a lo largo de más de 25 años dentro del lago cratérico.

Keywords: series temporales, teledetección, temperatura superficial, pipeline de datos, lago cratérico volcánico

Pipeline for the construction of a multitemporal dataset of surface temperature and spectral indices from satellite sensors for volcano monitoring.

Abstract. This work presents a pipeline for downloading, processing and analyzing satellite imagery aimed at constructing time series of surface temperature and spectral indices at active volcanoes. The result dataset supports the characterization of long-term thermal behaviour and may contribute to volcano monitoring. As a first step, we focused on the Landsat constellation, given its long operational history and public data availability. The pipeline comprises the definition of regions of interest and automated access and download of Landsat 7, 8 and 9 imagery and the subsequent extraction of spectral bands and high-level products and eventually, the construction of the time series. We implemented this pipeline at the crater lake of Tupungatito volcano (Mendoza, Argentina) and generated a calibrated surface temperature time series based on 837 Landsat scenes. The workflow extracts spectral indices and uses QA metadata to filter invalid scenes. Since between 2000 and 2025 the data were acquired by different Landsat sensors, we performed a calibration to ensure the consistency of the maximum surface temperature time series across more than 25 years within the crater lake.

Keywords: time series, remote sensing, surface temperature, data pipeline, volcano crater lake

1 Motivation and Objectives

Satellite thermal remote sensing provides robust long-term monitoring capabilities and has been successfully applied to crater lake volcanoes in the Argentine-Chilean Andes, e.g. Candela-Becerra et al. (2020); Suárez-Herrera et al. (2021); Aguilera et al. (2021). However, no long-term satellite-based thermal study has yet been conducted for Tupungatito volcano, and the development of a generalizable multi-sensor workflow for this type of analysis remains an open challenge in the region.

In this work, we present an automated workflow to retrieve and process Landsat satellite data and to construct long-term time series of surface temperature in volcanic environments. The workflow also incorporates spectral indices as auxiliary variables to characterize lake extent and identify snow cover, improving the reliability of surface temperature retrievals.

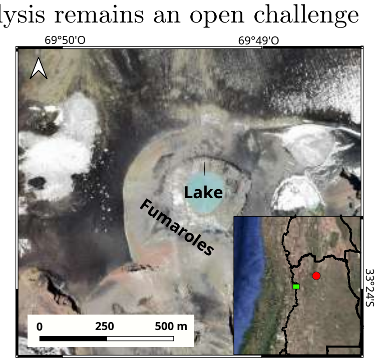

Fig. 1: Crater lake at Tupungatito volcano. Inset: location of the study area (green box) relative to Mendoza city, Argentina (red circle).

1.1 The target volcano: Tupungatito

We applied this workflow to Tupungatito volcano (on the Argentina–Chile border, at the latitude of the Argentine province of Mendoza, Figure 1), where we constructed a calibrated surface temperature time series for the crater lake spanning the 2000–2025 period. This dataset characterizes the long-term thermal behaviour of the system in a high altitude environment.

Tupungatito is ranked sixth highest-risk in Argentina according to the Argentine Volcano Monitoring Observatory of the Geological Survey of Argentina (OAVV-SEGEMAR); last erupted 1986–1987 and hosts a crater lake and active fumaroles at high altitude. Given its quiescent state and relative scarcity of prior studies (e.g. Benavente et al., 2013, 2016), it serves as benchmark for characterizing baseline thermal behavior of crater-lake volcanoes in the region.

2 Workflow and target

Figure 2 summarizes the implemented workflow, including the definition of study parameters, Landsat retrieval, processing, calibration, and time-series construction.

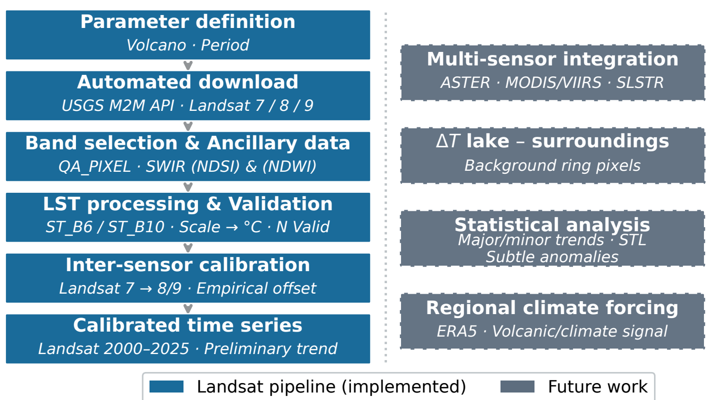

Fig. 2: Our designed Landsat-based workflow and extensions.

We accessed Landsat products from the USGS EarthExplorer platform (https://earthexplorer.usgs.gov/) via the Machine-to-Machine (M2M) application programming interface (API).

We used Landsat 7, 8, and 9 Level-2 products. The Normalized Difference Water Index (NDWI; $(NIR-SWIR)/(NIR+SWIR)$) and QA_PIXEL (Quality Assurance) metadata were used to exclude invalid observations over the lake prior to analysis (Earth Resources Observation and Science Center, 2020a, 2020b). The Normalized Difference Snow Index (NDSI; $(Green-SWIR)/(Green+SWIR)$) is also extracted by the pipeline and included in the published dataset; while not used in the present crater lake analysis, it is available for future work extending the characterization to the broader volcanic edifice.

Land surface temperature (LST) is derived from thermal infrared bands (native 60 m for L7, 100 m for L8-9; resampled to 30 m), allowing extraction of representative lake surface values. (Earth Resources Observation and Science Center, 2020a, 2020b).

Given the 16-day revisit time of each Landsat satellite, we combined observations from Landsat 7, 8, and 9 to increase temporal sampling, producing a dataset spanning over 25 years increasing the annual number of usable observations from 18 to 61 scenes and a median interval of 8 days between valid observations.

Because differences in sensor characteristics introduce inconsistencies in retrieved surface temperature, inter-sensor LST bias was estimated between Landsat 7 and Landsat 8–9 using 93 near-contemporaneous scene pairs acquired within 24 hours, as Landsat 8 and 9 share comparable sensor specifications and simultaneous acquisitions are unavailable. Scenes with maximum lake LST below or equal to 0 °C were excluded as a quality-control step before constructing the main 837-scene dataset, using a pragmatic threshold for filtering based on a 0 °C proxy. The median LST difference was -3.7 °C (SD = 5.0 °C), and this offset was applied to Landsat 7 data to produce a consistent multitemporal series. Kernel Density Estimate (KDE) plots of maximum LST before and after correction show improved inter-sensor agreement, reflected by lower absolute skewness and excess kurtosis approaching zero after calibration (Figure 3a).

The processing workflow is implemented in Python; the codebase is under active development and will be considered for public release in a future version.

3 Preliminary Results

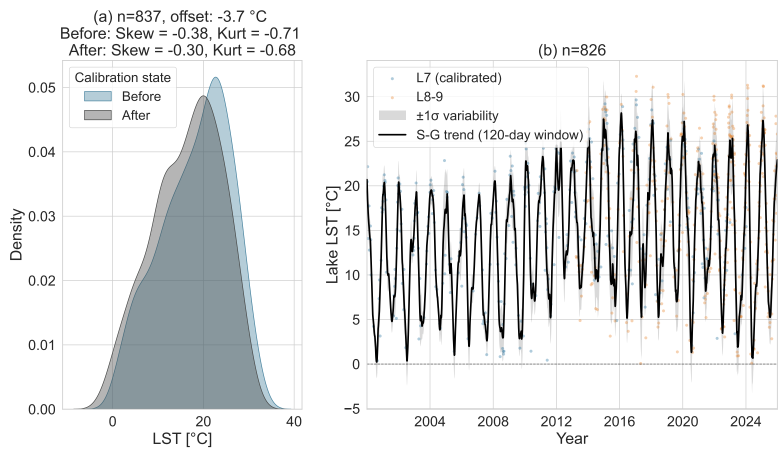

Fig. 3: (a) KDE plot of maximum LST for all 837 valid scenes, before and after inter-sensor calibration (offset: $-3.7\,^{\circ}$C, see Section 2) (b) Calibrated LST time series of Tupungatito volcano crater lake (2000–2025, $n = 826$), with Savitzky–Golay trend (120-day window) and $\pm 1\sigma$ variability band.

Although the inter-sensor correction offset produced some calibrated LST values below 0 °C, these observations were excluded from the final temporal series as a conservative quality-control step, resulting in 826 valid samples. They were nevertheless retained in the KDE comparison to preserve distributional consistency between pre- and post-calibration datasets.

To characterize the long-term thermal behaviour of the system, we applied a Savitzky-Golay (S-G) smoothing filter with a window of 120 days and polynomial order 3 to the daily-resampled calibrated time series, using nearest-neighbour interpolation to handle irregular sampling.The data show clear seasonal variability, likely driven by atmospheric forcing, with temperature peaks during summer and minima in winter, while remaining overall stable within a limited temperature range (Figure 3b). A subtle trend appears primarily in the Landsat 7 segment, possibly affected by residual orbital drift, and requires further analysis after seasonal decomposition.

The 0 °C screening threshold is only a proxy, particularly in an acidic hydrothermal lake, and will be refined in future work.

4 Conclusions

This work introduces an automated pipeline for constructing LST time series at remote volcanoes using the Landsat constellation, integrating over 25 years of data (2000–2025) and increasing observation frequency from 18 to 61 annual scenes. Inter-sensor calibration improved the statistical consistency of the LST distribution, as evidenced by reduced absolute skewness and excess kurtosis approaching zero after correction. Preliminary results suggest that Tupungatito crater lake exhibits stable long-term thermal behavior dominated by seasonal variability. A subtle trend in the Landsat 7 segment may reflect residual orbital drift and requires further analysis after seasonal decomposition. Future work will refine the calibration methodology, analyze subtle trends and anomalies, and integrate additional sensors (ASTER, MODIS/VIIRS, SLSTR) and ERA5 reanalysis to improve climate-signal filtering.

Data Availability Statement. Processed datasets and Python dependencies are available in the GitHub repository:
https://github.com/hernogr/supplementary-material-jaiio-2026

Acknowledgments. This work is part of the PhD project of HGR, supported by a CONICET doctoral fellowship. The authors acknowledge the Instituto de Estudios Andinos 'Don Pablo Groeber' (IDEAN, UBA-CONICET) and the United States Geological Survey (USGS) for institutional support and Landsat data, respectively. During the preparation of this work, the authors used Claude (Anthropic) to improve the language and readability of the manuscript. After using this tool, the authors reviewed and edited the content as necessary and take full responsibility for the content of the publication.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

Aguilera, F., Caro, J., & Layana, S. (2021). The evolution of Peteroa volcano (Chile–Argentina) crater lakes between 1984 and 2020 based on Landsat and planet labs imagery analysis. *Frontiers in Earth Science*, *9*, 722056. https://doi.org/10.3389/feart.2021.722056

Benavente, O., Tassi, F., Aguilera, F., Gutiérrez, F., Agusto, M., Caselli, A., Saltori, O., Vaselli, O., & Aguirre, M. (2013). Origin of fumarolic fluids from Tupungatito volcano (Central Chile): Interplay between magmatic, hydrothermal, and atmospheric gases. *Bulletin of Volcanology*, *75*(8), 746. https://doi.org/10.1007/s00445-013-0746-x

Benavente, O., Tassi, F., Reich, M., Aguilera, F., Capecchiacci, F., Gutiérrez, F., Vaselli, O., & Rizzo, A. (2016). Chemical and isotopic features of cold and thermal fluids discharged in the Southern Volcanic Zone between 32.5°s and 36°s: Insights into the physical and chemical processes controlling fluid geochemistry in geothermal systems of Central Chile. *Journal of Volcanology and Geothermal Research*, *320*, 85–104. https://doi.org/10.1016/j.chemgeo.2015.11.010

Candela-Becerra, L. J., Toyos, G., Suárez-Herrera, C. A., Castro-Godoy, S., & Agusto, M. (2020). Thermal evolution of the crater lake of Copahue volcano with ASTER during the last quiescence period between 2000 and 2012 eruptions. *Journal of Volcanology and Geothermal Research*, *392*, 106752. https://doi.org/10.1016/j.jvolgeores.2019.106752

Earth Resources Observation and Science Center. (2020a). *Landsat 7 Enhanced Thematic Mapper Plus Level-2, Collection 2* (dataset). dataset. U.S. Geological Survey. https://doi.org/10.5066/P9C7I13B

Earth Resources Observation and Science Center. (2020b). *Landsat 8-9 Operational Land Imager / Thermal Infrared Sensor Level-2, Collection 2* (dataset). dataset. U.S. Geological Survey. https://doi.org/10.5066/P9OGBGM6

Suárez-Herrera, C. A., Toyos, G., Candela-Becerra, L. J., & Agusto, M. (2021). Analysis of thermal anomalies at Copahue volcano between October 2011 and the December 2012 eruption with MODIS. *Journal of South American Earth Sciences*, *110*, 103310. https://doi.org/10.1016/j.jsames.2021.103310