



DEVELOPING A REMOTE SENSING TOOL FOR RESERVOIR VOLUME EVALUATION



João Pimenta
National Laboratory for
Civil Engineering



João N. Fernandes¹
National Laboratory for
Civil Engineering



Alberto Azevedo
National Laboratory for
Civil Engineering

ABSTRACT

The efficient management of water reservoirs is crucial for ensuring water security, flood control, and hydroelectric power generation. Accurate reservoir volume measurements are essential for optimizing these functions. Traditional methods, which rely on in-situ measurements and physical surveys, are often time-consuming, resource-intensive, and impractical in many countries due to financial constraints or complex environmental conditions. To address these challenges, this research presents the development of a novel remote sensing tool that offers an effective and globally accessible method for evaluating reservoir volumes.

This study explores the integration of satellite imagery and advanced geospatial analysis techniques to estimate reservoir volumes. Using high-resolution imagery from Sentinel-2, along with storage capacity data from official repositories and the GLOBathy dataset, the tool automatically calculates surface area and water volume at varying water levels over a user-defined time period. The tool features an intuitive interface that allows users to easily select reservoirs, date ranges, and cloud cover percentages, generating time-series data on volume. A key aspect lies in the use of machine learning algorithms, which enhance the accuracy of water surface delineation and volume estimation, even in the presence of built structures like bridges or challenging environmental conditions such as cloud cover.

The tool was tested across reservoirs of varying sizes and topographies in Portugal and California, USA. Comparative analyses showed a high degree of accuracy, with a mean absolute percentage error of 5.4% and a correlation coefficient (R^2) of 0.9 in comparison to official measurements. This cost-effective, scalable, and timely alternative offers significant improvements for reservoir volume evaluation, particularly in remote or hard-to-access regions. Ultimately, this research makes a substantial contribution to global water resource management by enhancing the resilience and sustainability of dam operations worldwide.

Keywords: Remote sensing, Reservoir volume, Satellite imagery, Machine learning, Water resources management.

¹jnfernandes@lnec.pt, Portugal

1. INTRODUCTION

Water stress is becoming a major national security issue, impacting quality of life, food and energy production, and political stability. Monitoring water levels is crucial for managing demand, flood control, and dam safety. Reservoirs have been vital since ancient times and over the centuries, dam construction and water management have advanced, with reservoirs playing key roles in flood control, hydropower, and irrigation. However, the impact of dams on rivers and water distribution is significant, with two-thirds of the world's long rivers affected by dams.

Despite extensive documentation of large dams, monitoring reservoir storage is still mostly limited to developed countries. This lack of data contributes to global water issues, as 2.2 billion people lack access to clean drinking water, and 3.5 billion lack proper sanitation, as highlighted in the 'UN World Water Development Report 2024' [1]. To address this, better monitoring and data-driven approaches are needed for efficient and sustainable water management.

The goal of this study is to develop a web-based tool that monitors the volume of reservoirs globally, especially in regions lacking data. The tool will analyse satellite images and provide information to help improve water management. By improving access to reservoir storage data, the tool aims to mitigate the effects of water stress and enhance global water security.

2. METHODOLOGY

2.1. Workflow

The tool employs a multi-stage workflow to calculate reservoir volumes. First, users define the Region of Interest (ROI) and a date range. The platform retrieves satellite images, then applies an unsupervised learning method to apply a threshold to the Normalized Difference Water Index (NDWI) to detect water pixels, excluding non-water features like bridges, which are detected using Google Places API and removed through morphological operations, such as dilation and erosion. If the images have a high percentage of cloud-covered areas, the cloud pixels areas are filtered out, and the reservoir shape is reconstructed using bathymetry data.

After determining the area, if an Area-Volume (A-V) relationship is available in the database, it is used for volume calculation. Otherwise, the user can either input the relationship or a Digital Elevation Model (DEM), and if the second option is chosen the code will use that file to estimate the A-V relationship. If none of these options are selected, global database – GLOBAthy – is used ([2]). The final step interpolates the calculated surface area to generate

a time-series of reservoir volumes. Figure 1 presents a flowchart addressing all the stages mentioned and how they are connected.

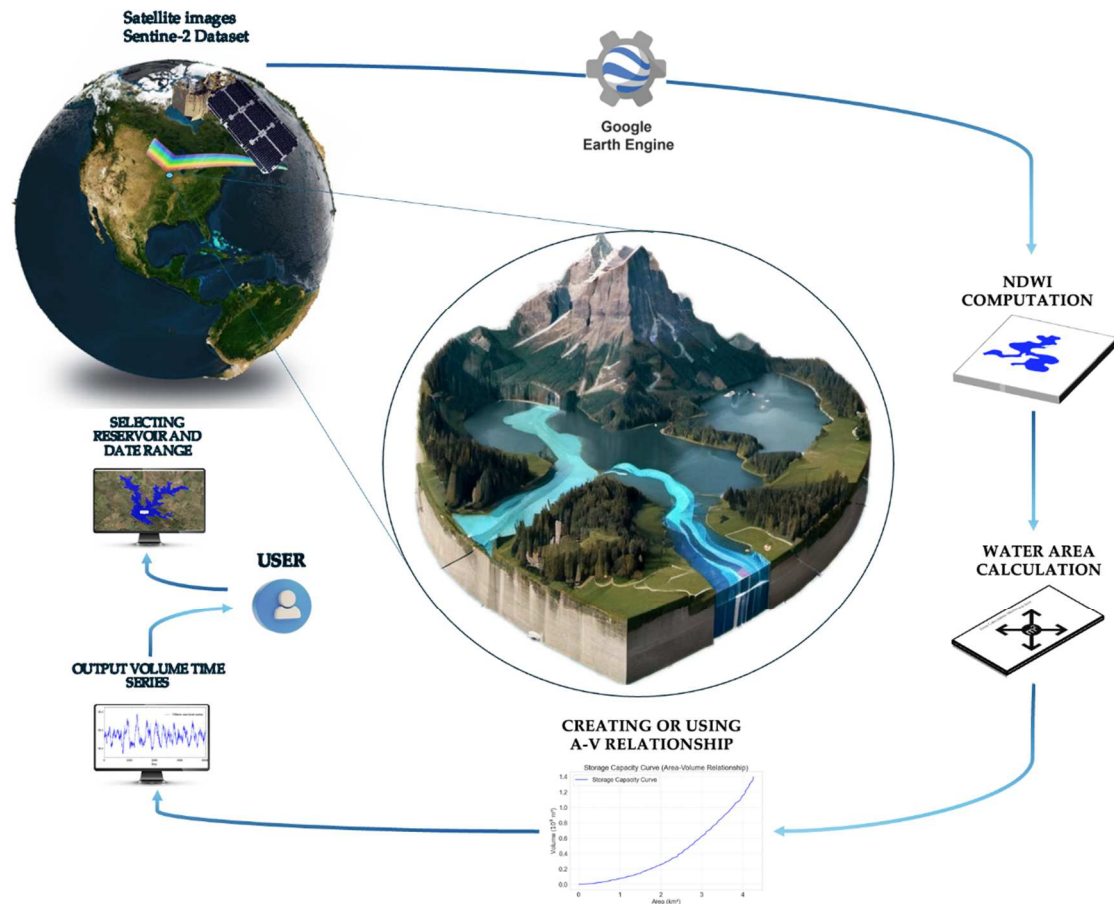


Fig. 1 — Workflow of the Tool

2.2. Google Earth Engine Platform

Google Earth Engine (GEE) is central to this study, providing an accessible platform for satellite data and computational resources. It allows for remote sensing data acquisition and processing via its Python API ([3]). Sentinel-2 imagery was chosen for its global coverage, high spatial resolution (10 m), and 5-day revisit time, making it ideal for water resources management (S2 Mission, [4]). The platform enables access to imagery based on region and cloud coverage criteria, ensuring only high-quality images are processed.

2.3. Calculating Water Area

The process begins by retrieving satellite images based on the user-defined date range and region of interest. Next, a binary mask of the NDWI is generated using a clustering algorithm. The tool then filters out unwanted features, such as environmental interferences like clouds and shadows, as well as man-made structures like bridges. Once these elements are removed,

the reservoir's water area is calculated by determining the pixel area of the regions classified as water in the NDWI binary image. Depending on the image resolution and projection system, this area is then converted to square meters using the reducer function in GEE.

2.4. Volume Estimation

The final stage involves calculating the reservoir's volume by interpolating the inundated water area using established relationships between surface area and stored water volume. This volume can be derived through several methodologies, depending on available data and the reservoir's characteristics. The approach employs a multi-source strategy to obtain the necessary Area-Volume (A-V) relationships. If the selected reservoir does not have the relationship in the platform's database, a selection box is presented to the user to choose other options (Fig. 2).

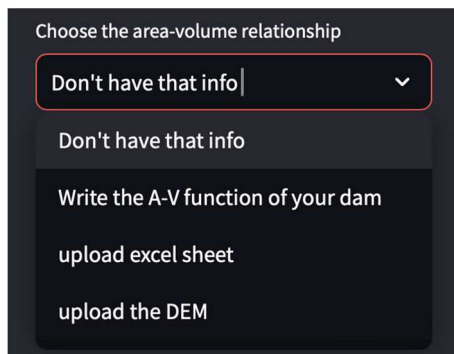


Fig. 2 – Select box to choose how to obtain the area-volume relationship

The storage capacity curve, also known as the area-storage or capacity-elevation curve, describes the relationship between the volume of water in a reservoir and its surface area or elevation. This curve is typically derived from detailed bathymetric surveys or elevation data, often maintained by local authorities. Two primary datasets for A-V relationships are the Portuguese Water Resources Information System (SNIRH) and the Global Reservoir and Dam Database (GRanD, [5]). The SNIRH provides historical hydrological data for reservoirs in Portugal, while GRanD offers a comprehensive global dataset, incorporating detailed information on reservoirs and their capacity relationships, managed by McGill University, in collaboration with several key institutions ([6]). The software retrieves these relationships and applies linear interpolation to estimate volume based on the calculated water area.

Estimating with Bathymetric Information of the Reservoir

If users have a Digital Elevation Model (DEM) (Fig. 3) but no existing A-V relationships, the tool can generate custom relationships using the DEM. This involves calculating the area and

volume at various elevation levels, providing a tailored estimate of the reservoir's storage characteristics. This method enables the software to create a specific A-V relationship for reservoirs lacking standard data. The code divides the model into different and equal elevation slices, by masking the pixels that have a specific elevation for that step, after which it defines two functions. The first function, computes the volume below a specified elevation by determining the water height above each cell and summing these values, multiplied by the cell area. The second function, creates a mask to identify inundated cells, counts them, and multiplies by the cell area to get the total inundated area. Lastly, a dictionary is initialized and created to store these calculated volumes and areas. This process systematically computes and logs the volume and area for a range of elevation levels (cf. Fig. 4), providing a comprehensive dataset for creating an Area-Volume dictionary, which will also be delivered as part of the output.

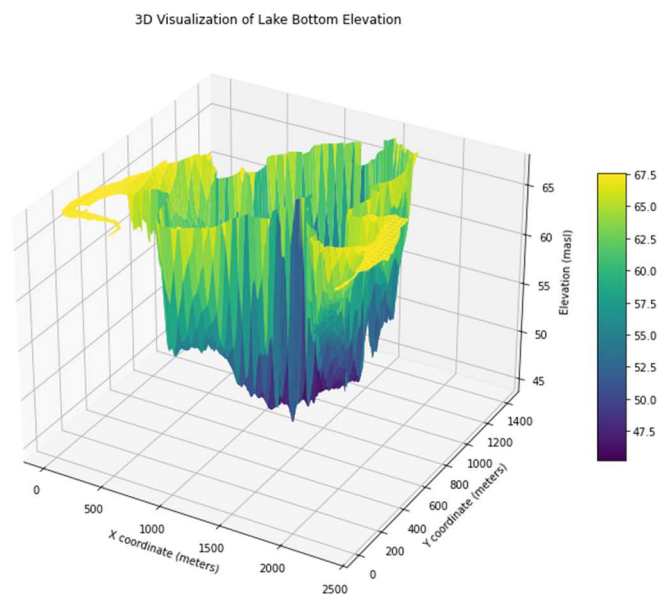


Fig. 3 – Lake bottom Elevation of the lake Patillas, Porto Rico

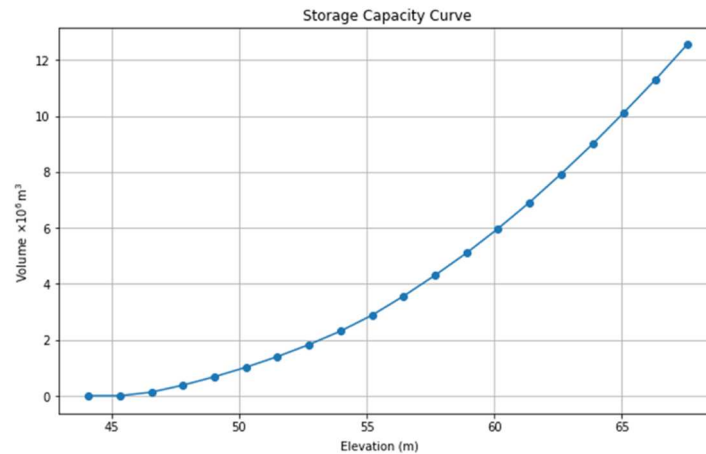


Fig. 4 – Storage capacity curve created of the lago Patillas, Puerto Rico

Using GLOBathy Data

In scenarios where neither A-V relationships nor DEM data are available, the tool defaults to the GLOBathy dataset. This comprehensive global bathymetric dataset covers over 1.4 million water bodies and integrates maximum depth estimates with water body characteristics ([2]). The GLOBathy dataset includes derived head-Area-Volume (h-A-V) relationships, ensuring that even without local data, accurate volume estimations can still be achieved by utilizing global standards.

3. RESULTS

A series of validation tests were conducted to assess the performance of the tool across its main algorithms. The objective of these tests was to verify the accuracy and reliability of the tool in processing remote sensing data for reservoir volume estimation.

The tests validated the tool's core functionality – accurate volume calculation – by establishing area-volume relationships derived from satellite imagery and comparing these estimates against measured data from various reservoirs. To comprehensively evaluate the tool's capabilities, data were used from multiple reservoirs of varying sizes and characteristics, incorporating both localized datasets, such as SNIRH and GLOBathy. Through this comparison, the accuracy of the tool is tested across different contexts, with the aim of improving its generalizability and applicability in diverse water resource management scenarios.

The diverse selection of reservoirs was chosen for testing across a minimum date range of five years. The selection criteria for these reservoirs were primarily based on the availability of publicly accessible historical data, which allowed for direct comparison between the observed

values and the volume estimates computed by the software for the same time periods. The reservoirs selected span two distinct regions: Portugal and the state of California in the United States (Fig. 5). Both regions were chosen due to their extensive networks of water bodies and the availability of high-quality historical data. In Portugal, data was sourced from the SNIRH, the national water resources information system, which provides detailed records of hydrological metrics across the country's numerous reservoirs.

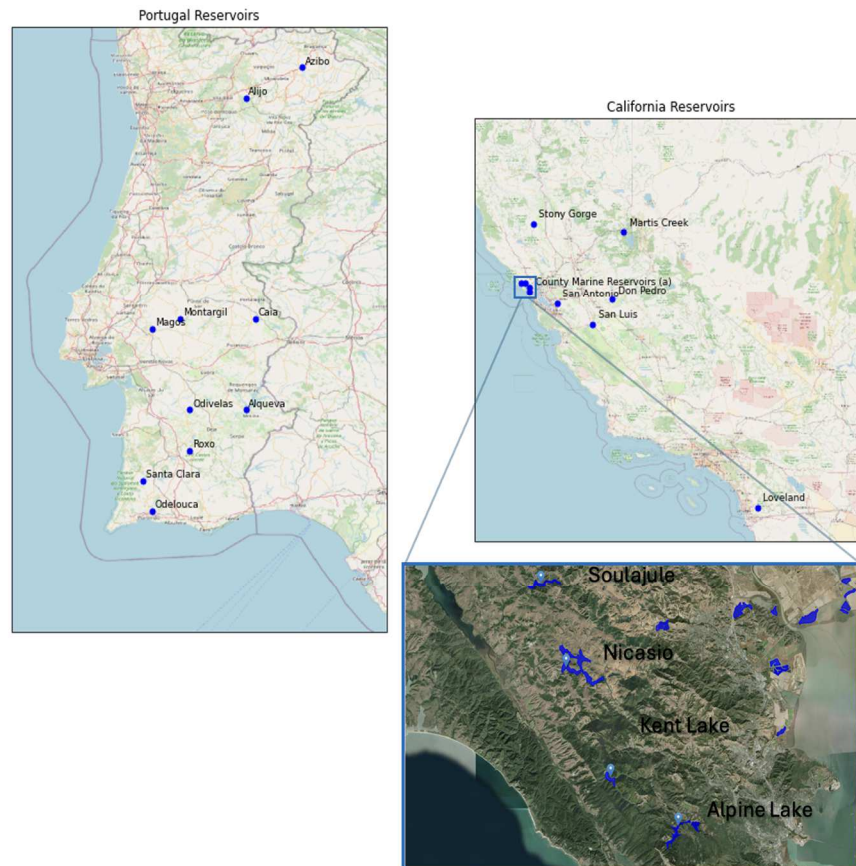


Fig. 5 – Location of the reservoirs tested

In California, data was retrieved from the California Data Exchange Center (CDEC), a centralized repository for hydrological and meteorological information in the state. This validation not only tests the software's performance across different geographical and hydrological conditions but also demonstrates its versatility in using data from different sources and environments. These reservoirs vary in type, use, geometry, and bathymetry together and represent a significant amount of water storage corresponding to a total volume of 13 402 hm³ and an inundated water area of 141 744 ha.

The tool prioritizes relationships stored in its internal database, giving preference to official repositories due to their higher reliability and accuracy. Then, to ensure a thorough evaluation, its performance was tested under default conditions, simulating its typical operation by comparing strictly its output with the official data for each reservoir. Figure 6 illustrates this

process, where the red line in each chart represents the platform's automatic output for six of the reservoirs tested, adapting dynamically to the available data source, whether from SNIRH or GLOBathy.

Ultimately, the accuracy and reliability of the software in calculating reservoir volumes are quantitatively assessed using key performance metrics, including the Coefficient of Determination (R^2), Mean Absolute Percentage Error (MAPE), and Mean Bias Deviation (Bias). These metrics provide a robust evaluation of the software's effectiveness under different scenarios and are presented in Table 1, offering a detailed assessment of its performance.

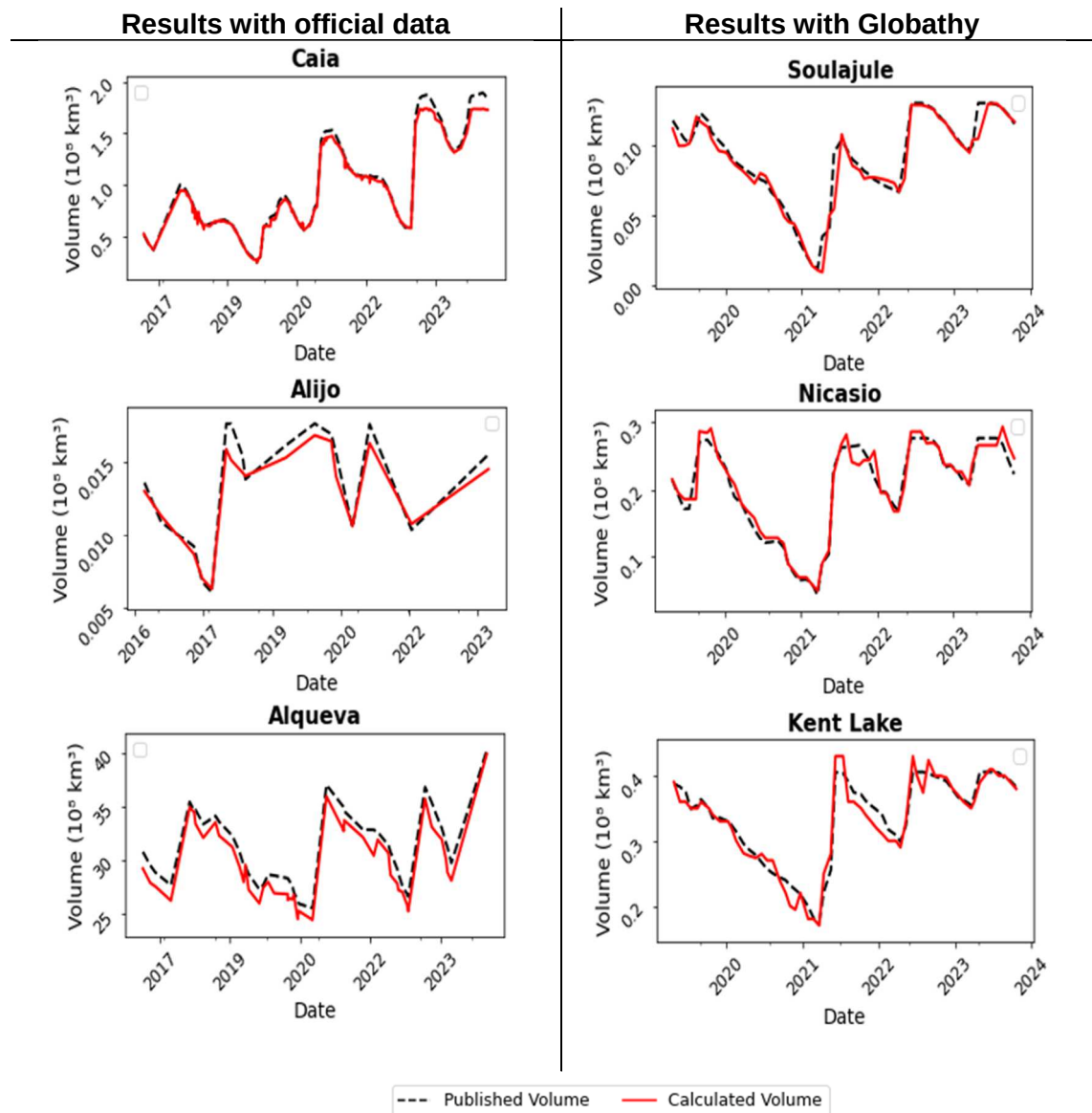


Fig. 6 – Actual vs Predicted volume for 6 reservoirs in Portugal and California, USA

The performance metrics indicate a strong predictive capability of the model. The tests yield an average Mean Absolute Percentage Error (MAPE) of 5.1%, an average Coefficient of Determination (R^2) of 0.9, and a Bias of 0.0. The low MAPE reflects that the model's predictions, on average, deviate by only 5.11% from the observed values, suggesting a high

level of accuracy in the estimations. The R^2 value of 0.9 indicates that 90% of the variance in the observed data is explained by the model, demonstrating its robustness and reliability in capturing the underlying data patterns.

Moreover, the Bias of 0.0 suggests the absence of systematic errors in the model's predictions, implying that the model neither consistently overestimates nor underestimates the actual values. This balance is crucial for ensuring the credibility and generalizability of the model's outputs across different datasets.

Table 1 – Summary table of the tools' performance towards the reservoirs tested

Reservoir	Maximum Area (ha)	Maximum Volume (10^3 hm^3)	A-V relationship source	Date range tested	Performance Metrics		
					R^2	MAPE	BIAS
Caia	1 970	203 000	SNIRH	8 years	0,988	2,404	-0,002
Alijó	18	1 760	SNIRH	8 years	0,924	5,546	0,000
Alqueva	25 000	4 150 000	SNIRH	8 years	0,833	4,081	-0,124
Kent Lake	350	40 582	GLOBathy	5 years	0,943	3,917	0,000
Soulajule Reservoir	730	13 198	GLOBathy	5 years	0,932	4,5	0,079
Nicasio Reservoir	3 170	85 800	GLOBathy	5 years	0,957	5,112	0,000

4. CONCLUSIONS

This research successfully developed a remote sensing tool that significantly improves reservoir volume evaluation, addressing key challenges in water resource management. By utilizing Sentinel-2 satellite imagery and machine learning algorithms, the tool offers a cost-effective alternative to traditional methods, providing accurate calculations of reservoir surface area and volume. Tested across reservoirs in Portugal and California, the tool demonstrated a mean absolute percentage error of 5.1% and a strong correlation coefficient of 0.9 compared to conventional measurements.

Although the prototype shows promising results, areas for improvement remain. These include updating storage capacity curves, improving satellite image resolution, and expanding the database with more local reservoir data. Future enhancements may incorporate additional data sources, such as radar imagery and telemetry, and improve the tool's adaptability to handle complex environmental factors like sedimentation. Machine learning capabilities could also enhance forecasting of reservoir volumes under various climate scenarios, providing greater value for water management stakeholders.

In conclusion, this software represents a significant advancement in remote sensing for water resource management, with the potential to greatly enhance the sustainability and resilience of global reservoir operations. Continued refinement and integration of new data sources will be crucial for maximizing its utility and accuracy.

The tool is available in <https://github.com/joao862/BLU>.

5. REFERENCES

- [1] Nations, U. (2024). The United Nations World Water Development Report 2024: Water for Prosperity and Peace. UN. Available at: <https://www.unwater.org/publications/un-world-water-development-report-2024>
- [2] Khazaei, B.; Read, L. K.; Casali, M.; Sampson K. M.; Yates D. N. (2022) GLOBathy, the global lakes bathymetry dataset. Scientific Data, 9(36)
- [3] <https://developers.google.com/earth-engine/tutorials/community/intro-to-python-api> - An Intro to the Earth Engine Python API | Google Earth Engine. Google Developers.
- [4] https://www.esa.int/Enabling_Support/Operations/Sentinel-2_operations - Sentinel-2 operations, European Space Agency.
- [5] Lehner, B.; Liermann, C.; Revenga, C.; Vörösmarty, C.; Fekete, B.; Crouzet, P.; Döll, P.; Endejan, M.; Frenken, K.; Magome, J.; Nilsson, C.; Robertson, J.; Rodel, R.; Sindorf, N.; Wisser, D. (2011). High-Resolution Mapping of the World's Reservoirs and Dams for Sustainable River-Flow Management. *Frontiers in Ecology and the Environment* 9 (9): 494-502.
- [6] Messenger, M. L.; Lehner, B.; Grill, G.; Nedeva, I.; Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7, article number 13603.