

Welcome to InfeRes: A Python package for inferring reservoir water surface area, level and storage volume

InfeRes is a python package that is designed to help automatic extraction of reservoir characteristics (water surface area, level, and storage-volume) time-series by taking leverage of the Google Earth Engine data collection (Landsat series, Sentinel-2), and high resolution DEM (30m). It built on top of GDAL, Scikit-Learn, NumPy and Matplotlib, and other popular python packages. **InfeRes** is developed with a novel algorithm which helps inferring reservoir characteristics even from the partially cloudy images. **InfeRes** can be applied to monitor water surface area in any reservoir or waterbody; whereas, storage-volume can be obtained for the large reservoirs (storage $\geq 0.1\text{m}^3$) listed in the GRanD database.

Components of InfeRes

1. Data download
 - Using standalone python environment (`DataDownload_GEE.py`)
 - Using web browser-based python environment (`DataDownload_GEE_GoogleColab.py`)
2. Data processing
 - Main python module (`main.py`)
 - Python module to create reservoir's Area-Elevation-Storage curves (`CURVE.py`)
 - Python module for pre-processing of satellite images (`PREPROCESSING.py`)
 - Python module to estimate reservoir's area and storage time-series (`WSA.py`)

Folder structure

Download the **Python_codes** folder containing the **InfeRes** package, and unzip it inside any directory. For instance, after unzip InfeRes path

should be *D:/Python_codes/InfeRes/*. Another folder 'Reservoirs' (path *D:/Python_codes/Reservoirs/*), where your satellite data will be downloaded. We have provided the example data for AyunHa reservoir for a quick setup and testing InfeRes as a case study.

NOTE: Please unzip all compressed folders before running InfeRes. The folder directories and their paths should be as follows after the unzip:

1. **Folder containing InfeRes modules:** *D:/Python_codes/InfeRes/*
2. **Folder containing reference GRanD curves:** *D:/Python_codes/InfeRes/GRAND_Curves/*
3. **Folder containing all reservoir's data:** *D:/Python_codes/Reservoirs/*
4. **Folder containing data for (say) AyunHa reservoir:**
D:/Python_codes/Reservoirs/AyunHa/
5. **Folder containing raw satellite images for (say) AyunHa reservoir:** *D:/Python_codes/Reservoirs/AyunHa/AyunHa_RawData/*
6. **Folder containing supplementary satellite data for (say) AyunHa reservoir:** *D:/Python_codes/Reservoirs/AyunHa/AyunHa_Supporting/*

Dependencies

- Python version-3.8 and above (we used Anaconda3, which is an open-source distribution of the Python)
- Python standard library (os, numpy, pandas, matplotlib, csv)
- Python advanced library (ee, osgeo, rasterio, sklearn.cluster, scipy.ndimage, skimage.morphology)

Installation

- Install the latest version of Anaconda (download [here](#)).

To create the conda environment with python=3.10, for instance, use:

```
(base) C:/User/UserName/conda create -n environment_name python=3.10
```

To activate the conda environment, use:

```
(base) C:/User/UserName/conda activate environment_name
```

- Install all libraries within the built environment (following steps are recommended).
 - i) conda install -c conda-forge **gdal=3.9.0** (assuming 3.9.0 is the latest version of GDAL)
 - ii) conda install -c conda-forge **rasterio**
 - iii) conda install -c conda-forge **spyder**

- iv) conda install -c conda-forge **earthengine-api**
- v) Similarly install all the other libraries
- Open spyder and load all the InfeRes modules (i.e. `DataDownload_GEE.py`, `main.py`, `CURVE.py`, `PREPROCESSING.py`, and `WSA.py`)

Usage Instructions

1. `DataDownload_GEE.py`

`DataDownload_GEE.py` is the first step towards running **InfeRes**. `DataDownload_GEE.py` will download the satellite images and store them in the Google Drive. Therefore, make sure you have sufficient space in your cloud storage (Google Drive in this case) before running `DataDownload_GEE.py`. Please also note that the downloading will take time to finish, which depends on the size of satellite image, downloading speed, and the number of images ordered. Therefore, one should first run `DataDownload_GEE.py` standalone, and wait until all the orders are successfully downloaded before running the other modules of InfeRes.

Inputs required (variable name):

- Name of the reservoir (`res_name`) = AyunHa
- Year of commission (`res_built_year`) = 1997
- Bounding box (`boundary`) = [108.155, 13.700, 108.300, 13.575]. Where, (108.155, 13.700) and (108.300, 13.575) are the (longitude, latitude) of top-left and bottom-right points of the bounding box.

The data will be downloaded inside *D:/Python_codes/Reservoirs/AyunHa/* in two different folders.

- Raw satellite data (Normalized Difference Water Index or NDWI in this case) will be at *D:/Python_codes/Reservoirs/AyunHa/AyunHa_RawData/*.
- Supplementary data (DEM, Water frequency, Maximum reservoir extent in this case) will be at *D:/Python_codes/Reservoirs/AyunHa/AyunHa_Supporting/*.

2. `DataDownload_GEE_GoogleColab.py`

`DataDownload_GEE_GoogleColab.py` is an alternative of `DataDownload_GEE.py`, which runs of web browser-based python environment such as Google Colab. It also takes the same set of inputs (i.e. Name of the reservoir, Year of commission, and Bounding box). However, in this case the data will be downloaded in next in your Google Drive, so the downloading path will be *D:/Python_codes/AyunHa_RawData/*

and *D:/Python_codes/AyunHa_Supporting/* for raw satellite data and supplementary data, respectively.

Please note that you need to maintain the folder structure as *D:/Python_codes/Reservoirs/AyunHa/AyunHa_RawData/* and *D:/Python_codes/Reservoirs/AyunHa/AyunHa_Supporting/* before running the InfeRes modules. Therefore, you need to move the data to the correct folder arrangement once the downloading is completed.

3. PREPROCESSING.py

PREPROCESSING.py performs the following tasks:

- Creating the reservoir isolation raster (binary map of reservoir maximum extent).
- Creating reservoir isolation for DEM (masked DEM)
- Reprojecting and resizing (or clipping) the satellite images including DEM, water extent, and frequency raster.
- Creating a collection of relatively good quality (less cloud cover) satellite images.

Inputs required (variable name):

- Name of the reservoir (res_name) = AyunHa
- Year of commission (res_built_year) = 1997
- Maximum water level in meter (max_wl) = 211
- A point coordinates on the reservoir (point) = [108.232, 13.638]
- Reservoir's bounding box coordinates (boundary) = [108.155, 13.700, 108.300, 13.575]

4. CURVE.py

CURVE.py creates the Area-Elevation-Storage relationship for a reservoir.

Inputs required (variable name):

- a. If reservoir has built before the acquisition of DEM (i.e. year 2000, as we are using SRTM DEM):
 - Name of the reservoir (res_name) = AyunHa
 - Identification number of the reservoir in the GRanD v1.3 database (grandID) = 7153
 - Maximum water level in meter (max_wl) = 211
 - A point coordinates on the reservoir (point) = [108.232, 13.638]

- Reservoir's bounding box coordinates (boundary) = [108.155, 13.700, 108.300, 13.575]
 - Reservoir's design capacity (i.e. maximum capacity) = 253 (this value can be extracted from the GRanD database)
- b. If reservoir has built after the acquisition of DEM (i.e. year 2000, as we are using SRTM DEM):
- Name of the reservoir (res_name) = AyunHa
 - Maximum water level in meter (max_wl) = 211

5. WSA.py

WSA.py estimates the area and storage time-series from the pre-processed time satellite images, which only takes input as the name of the reservoir.

Inputs required (variable name):

- Name of the reservoir (res_name) = AyunHa

How to Run?

Step 1. Run either **DataDownload_GEE_GoogleColab.py** or **DataDownload_GEE.py** standalone, and let the data download finish (i.e. Satellite NDWI images, Maximum water extent, Water frequency, and DEM).

Step 2. (Assuming you already have all the required datasets) Open Spyder and locate the directory to the **InfeRes**, and load the modules **main.py**, **PREPROCESSING.py**, **CURVE.py**, and **WSA.py**.

Step 3. Configure **main.py**

- Modify the path of InfeRes directory (i.e. **parent_directory**)
- Prepare the input file (i.e. **inputs_InfeRes_185res.csv**)

inputs_InfeRes_185res.csv contains:

- Name of the reservoir (res_name) = AyunHa
- Year of commission (res_built_year) = 1997
- Maximum water level in meter (max_wl) = 211
- GRanD ID = 7153 (if GRanD ID is not available, assign 0)
- A point coordinates on the reservoir (point) = [108.232, 13.638]
- Reservoir's bounding box coordinates (boundary) = [108.155, 13.700, 108.300, 13.575]

- Reservoir’s design capacity = 253 (assign 0, if reservoir has built after 2000)
- Run the `main.py`

NOTE: `main.py` calls other modules in a sequential order (`PREPROCESSING.py` -> `CURVE.py` -> `WSA.py`) to get the desired outputs (i.e. reservoir’s area, level, and storage in this case).

Outputs

The outputs will be saved in a folder called *’Outputs’* in the same directory where your input data are kept.

InfeRes will generate the following outputs:

- Area-Elevation-Storage relationship (**Curve.csv**)
- List of images used for estimation of storage (**Image_List.csv**)
- Table containing the scene-based (landsat and Sentinel) reservoir area and storage (**WSA.csv**)
- Updated table containing scene-based reservoir area in km^2 , water level in m, and storage in million m^3 (**WSA_updated.csv**)
- Intermediate raster images
- Intermediate figures (inside a separate folder called *JPG_files*)

Acknowledgement

We have acquired the reference GRAND_Curves (reservoir’s reconstructed bathymetry) from Hao et al., (2024), which is available for the list of reservoirs in the Global Reservoir and Dam (GRanD) Database (Lehner et al., 2011).

- Hao, Z., Chen, F., Jia, X., Cai, X., Yang, C., Du, Y., & Ling, F. (2024). GRDL: A New Global Reservoir Area-Storage-Depth Data Set Derived Through Deep Learning-Based Bathymetry Reconstruction. *Water Resources Research*, 60(1), e2023WR035781.
- Lehner, B., C. Reidy Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J.C. Robertson, R. Rodel, N. Sindorf, and D. Wisser. 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.