Python-Google Earth Engine Surface Water Analysis Toolbox (PyGEE-SWToolbox)

User's Manual

Version 1.0.1

Date of Release: 03/12/2021

Collins Owusu
Nusrat Jannah Snigdha
Mackenzie T. Martin
Alfred J. Kalyanapu

Department of Civil and Environmental Engineering,
Tennessee Technological University,
Cookeville, Tennessee

Abstract

The Python-Google Earth Engine (GEE) Surface Water Analysis Toolbox provides a graphical user interface (GUI) developed using Python 3 programming language and the GEE Python API to facilitate processing and analysis of surface water using available datasets on the GEE platform, powered by Google Cloud computing. Traditional use of the GEE platform requires the user to be able to write JavaScript codes in the GEE online Code Editor to perform geospatial processing of datasets. This toolbox provides an interface for surface water analysis and visualization without the user writing code. The toolbox an "Image Processing" tab for specifying required datasets, timespan, and area of interest, an "Image Correction and Water Extraction" for cloud masking, pan sharpening, and surface water extraction, a "Spatial Analysis" tab for performing spatial analysis of extracted water, a "Plotting & Stats" tab for plotting and statistical analysis of time series of extracted water, and a "Download" tab for downloading processed satellite imagery or extracted surface water images. A main component of the toolbox is the map viewer which enables the user to visualize datasets and draw a boundary around their region of interest. The tool can be applied to any region where datasets are currently accessible. Results from the tool are subject to the quality of datasets and user verification.

Table of Contents

1 User's Agreement

By using the "Python-GEE Surface Water Analyzer Toolbox", you agree to the following:

1. You assume sole responsibility for any claims or liabilities that may arise as a result of the use of the toolbox. Your use of or reliance upon the toolbox and any third-party content and services accessed thereby is at your sole risk and discretion. You acknowledge that results from the toolbox are not a set of recommendations.
2. You acknowledge and agree that, in the event of a third-party claim that the toolbox or your possession or use of the toolbox infringes any third party's intellectual property rights, you will be responsible for the investigation, defense, settlement, and discharge of any such claim of intellectual property infringement.
3. You agree to cite the sources of the Python-GEE Surface Water Analyzer Toolbox in written reports that describe results obtained through its use.

Recommended Citation:

Owusu C, Snigdha NJ, Martin MT, Kalyanapu AJ. PyGEE-SWToolbox: A Python Jupyter Notebook Toolbox for Interactive Surface Water Mapping and Analysis Using Google Earth Engine. Sustainability. 2022; 14(5):2557. https://doi.org/10.3390/su14052557

2 Preface and Background

2.1 Preface

This manual is not intended to teach you how to conduct surface water extraction and analysis from remotely sensed imagery nor does it describe the mathematical equations for the various components of the toolbox. Instead, this manual will teach you how to use various features and capabilities of the toolbox. It is recommended to read the manual with the toolbox running. However, the manual works equally well as an occasional reference when you cannot remember exactly how to perform a certain task.

Many graduate students and professors have contributed to the development and completion of this toolbox. Each student has made valuable contributions that enhance the overall success of the program. The completion of the project was overseen by Dr. Alfred J. Kalyanapu.

2.2 Background

The Python-GEE Surface Water Analysis Toolbox was developed as part of research to reconstruct historical dynamics of surface water within wetlands under the Wetland Reserve Program (WRP) which were being monitored to assess the effectiveness of wetland restoration practices implemented by the Natural Resources Conservation Service (NRCS) in West Tennessee and West Kentucky. The intent was to utilize readily available remote sensing datasets available on the GEE and leverage the high computing power of the platform. The toolbox communicates with the GEE platform using the GEE Python API. The toolbox eliminates the need for downloading and processing large satellite imagery often required for surface water extraction and analysis.

The research is based upon work supported by the Natural Resources Conservation Service, U.S. Department of Agriculture, and The Nature Conservancy, under award number 68-5C16-17-015. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Natural Resources Conservation Service or The Nature Conservancy.

3 Introduction

The Python-GEE Surface Water Analyzer Toolbox is designed to extract surface water extents, estimate area, depth, volume, and time series analysis of changes in surface water. It is designed to be applicable in a wide range of geographic areas. Raster images and hydrographs can be used directly or in conjunction with other software for studies of water availability, flood monitoring, and watershed management.

3.1 Overview of Toolbox Features

This section describes the key features of the toolbox. It includes descriptions of the main components of the GUI.

3.1.1 User Interface

The user interacts with the GEE platform through the GUI which makes it easy to analyze large satellite datasets. The GUI is made up of three main components: Menu, Map Viewer, and Plotter. Figure 1 shows the GUI of the toolbox with the three main components denoted by red rectangular boxes.

Menu

The menu is made of 5 tabs for Image Processing, Water Extraction, Spatial Analysis, Plotting & Statistics, and Image Download. The image processing menu provides the user with a selection of desired satellite platforms, a definition of period of analysis, a definition of the area of interest (AOI), and a specification of cloud threshold for image filtering.

The water extraction menu gives the user options in water extraction index, thresholding method, threshold value, and color to represent surface water. The spatial analysis menu provides the user the option to compute water occurrence frequency for hydroperiod analysis and estimate water depths based on a user-defined digital elevation model (DEM). The user can compute and plot the dynamics of surface water under the plotting and statistics menu with the option to save the computed surface water areas to a user-designated file. The download and export menu provides the user the option to export satellite imagery, water maps, water occurrence frequency maps, and depth maps to Google Drive or downloads to the user’s local computer.

Map Viewer

The map viewer component provides the visualization of satellite imagery, surface water extent maps, water occurrence frequency, and depth maps. It also provides drawing tools for defining the region of interest (AOI) and measures the area (units) of the AOI. There is a toolbar for viewing, turning on and off layers, and changing settings for various layers.

Plotter

This plotter component is used to visualize time series of changes in surface water dynamics.

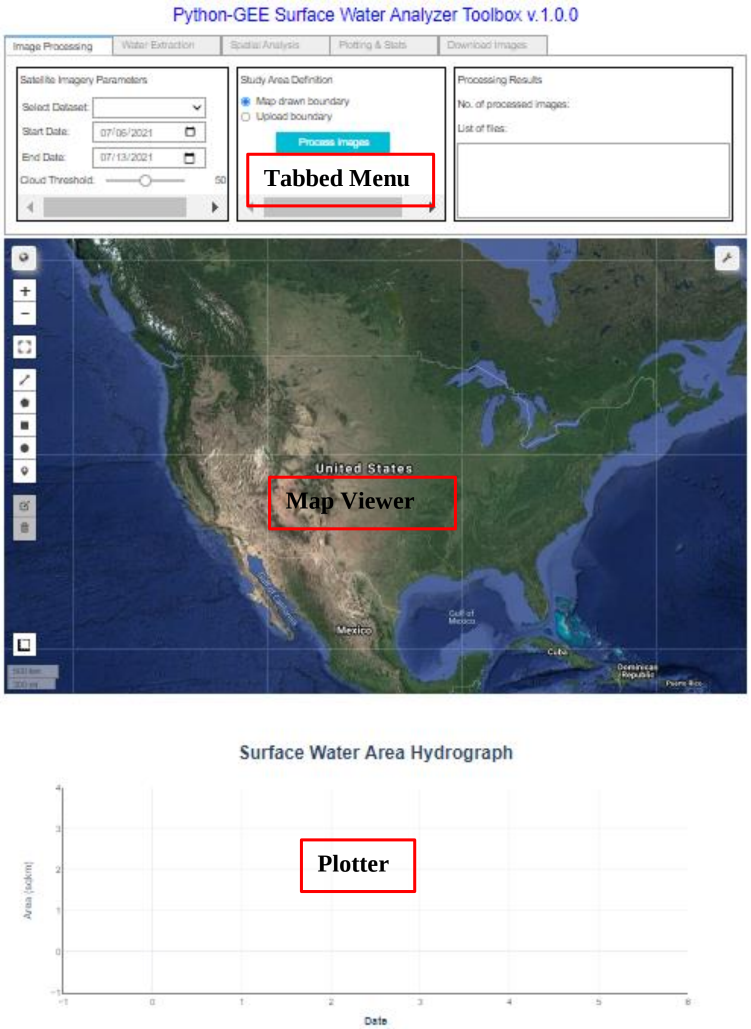


Figure 1 The main toolbox GUI with the tabbed menus at the top, map viewer in the middle, and plotter at the bottom

3.1.2 Image Retrieval and Processing

The toolbox provides the option to retrieve and process Landsat (Landsat 4-8), Sentinel-1, Sentinel-2, and the National Agriculture Imagery Program (NAIP) high-resolution satellite imagery. The Landsat and Sentinel-2 imagery is run through cloud masking algorithms which remove cloud-contaminated pixels. The Sentinel-1 imagery is processed through noise filtering using the Sigma Lee speckle filtering algorithm (Lee, 1983). The NAIP imagery is for visual analysis but can be used for accuracy checks where imagery is available.

3.1.3 Surface Water Extraction

The toolbox is equipped with five surface water extraction algorithms. The algorithms included are the Normalized Difference Water Index (NDWI) (McFeeters, 1996), Modified Normalized Difference Water Index (MNDWI) (Xu, 2006), Dynamic Surface Water Extent (DSWE) (Jones, 2019), and Automated Water Extraction Index (AWEI) (Feyisa et al., 2014). Users of the tool are encouraged to consult the above cited references for the definition and computation of the water indices.

Water extraction can be achieved by using a thresholding method that requires the user to enter a threshold value or use Otsu’s dynamic thresholding method (Otsu, 1979).

3.1.4 Spatial Analysis

A spatial water occurrence frequency map for a specified period can be computed with the toolbox based on extracted surface water maps. The frequency map can be used for hydroperiod analyses.

The toolbox can be used to generate water depth maps for extracted surface maps. The water depths are estimated from either the National Elevation Dataset (NED) (applicable to the USA only), Shuttle Radar Topography Mission (SRTM) digital elevation data, or a user-defined GEE asset.

3.1.5 Time Series Plotting

A time series of surface water extent dynamics for a given time series of satellite images can be plotted using the toolbox. The plot of the surface water dynamics will be displayed in the plotter GUI. The data generated for plotting the time series can be exported to a CSV format for use in other environments. The generated plots can be saved in image formats for reporting purposes using the available menus in the plotter GUI. Clicking a point in the plot will display the satellite image and water map of the corresponding date in the map viewer.

3.1.6 Download and Export

The toolbox provides the option to export the satellite images, water maps, water occurrence frequency maps, depth maps, and DSWE images to Google Drive or the user’s local computer by selecting a directory to store the images. The option to download the DSWE images is presented so the user may ascertain the accuracy of the toolbox-derived DSWE by comparing it to the original DSWE datasets by the USGS Earth Resources Observation (EROS) Center.

4 Installation

To use the PyGEE-SWToolbox, the user must first sign up for a Google Earth Engine account to authenticate the use of the GEE API.

To download the toolbox, navigate to the project repository, as shown:

```
https://github.com/collinsowusu/PyGEE-SWToolbox
```

Above the list of files, click the ***Code*** button to download ZIP as shown in Figure 2. Save and extract the zip file to the user's desired location. The project archive can also be downloaded from Zenodo.

Optionally, the user can clone the project repository using Git Bash as shown below:

```
git clone https://github.com/collinsowusu/PyGEE-SWToolbox.git
```

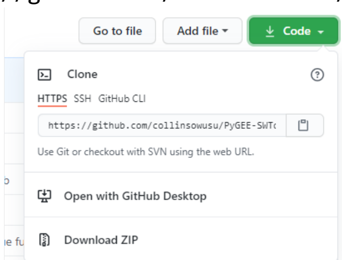


Figure 2 Downloading PyGEE-SWToolbox

It is recommended to download and install either ***Miniconda*** or ***Anaconda*** on your computer. ***Miniconda*** is recommended because it is a minimal version of Anaconda with no third-party binaries. Open the Anaconda Prompt from your programs menu after the Miniconda or Anaconda is installed.

Commented [KA1]: @Owusu, Collins (cowusu42) should this be "naconda" or "anaconda"?

Commented [KA2R1]: @Owusu, Collins (cowusu42) - It may be good to explain to the reader, why this conda environment has to be created. In otherwords, is there any other alternatives to installing/using the PyGEE-SWToolbox? Clarification on the installation (either to me or here in text), will be preferred.

Create a conda Python environment and install Jupyter notebook with the following commands (where *SWToolbox* is the name of the environment which can be changed to a user-preferred name):

```
conda create -n SWToolbox python=3.8
conda activate SWToolbox
conda install -c conda-forge notebook
```

The required packages can be found in the requirements.txt file in the project directory. These packages can be installed manually as:

```
pip install <package name>
```

Optionally, you can change to the project directory at the conda prompt using the following command:

```
cd <project directory>
```

where <project directory> represents the location of the project directory on the user's computer. For example, if the project is located in your documents as *C:\Users\user1\Documents\ PyGEE-SWToolbox*, then the command to change to the directory is:

```
cd C:\Users\user1\Documents\ PyGEE-SWToolbox
```

Install all packages using pip:

```
pip install -r requirements.txt
```

5 Working With The Toolbox

5.1 Starting Up

Open Anaconda Prompt and activate the created conda evironment if not already activated using the following command:

```
conda activate <environment name>
```

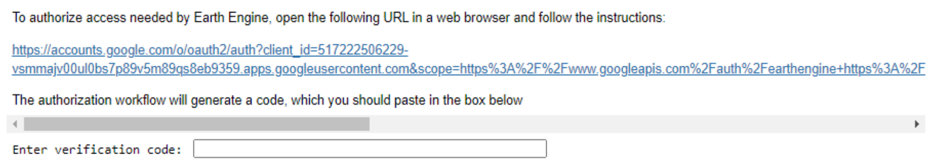
In your created conda environment, open Jupyter notebook using the following command:

```
jupyter notebook
```

With the notebook running, open the PyGEE-SWToolbox.ipynb notebook file located in the project directory.

The toolbox notebook contains two cells. The first cell contains the code to import and initialize the GEE API. The code will require the user to authenticate the user's GEE account for the first run of the toolbox. The account token will be stored on the user's computer to prevent future authentication.

Click in the first cell with the code and click on the ***Run*** button on the jupyter notebook toolbar at the top to run the code in the cell. The code can also be run by hitting the ***Ctr+Enter keys*** on the keyboard. Authentication of the GEE account will require the user to enter a verification code as shown below:



Click on the generated URL which will open in another browser. Follow the instructions to authenticate your account. Copy the verification code generated and paste it into the verification code box as shown above. Press the ***Enter key*** to complete the verification process.

Run the second cell to display the GUI of the toolbox. Follow the instructions in the subsequent sections on how to use the toolbox.

5.2 Image Retrieval and Processing

This section describes the process of retrieving and processing satellite imagery from the GEE platform using the toolbox. The toolbox requires you to set the satellite image parameters and the definition of the study area.

5.2.1 Satellite Image Parameters

Image parameters include specifying the type of satellite images to retrieve. The user selects the type of images by clicking on the dropdown menu at ***Select Dataset.*** Selecting certain datasets specifies the image visualization parameters, and enables or disables components of the toolbox that are required for processing the specified image. Currently, the toolbox allows the processing of ***Sentinel-1, Sentinel-2, Landsat 4-8***, and ***USDA NAIP*** images. Efforts are being made to include more datasets.

The ***Start Date*** and ***End Date*** of the period of analysis can be specified by entering the date in the provided box or clicking on the calendar symbol and selecting a date. Entering the date requires the date to be in the same format as displayed in the toolbox. The default end date is set to the user's current date and the default start date is set to 7 days before the current date.

The ***Cloud Threshold*** slider allows the user to specify the threshold (%) of cloud cover permitted by the user with the default value being 50%. Setting a cloud cover threshold tells the toolbox to retrieve only images with cloud cover that is equal to or less than the specified threshold. **Note:** the cloud threshold slider is deactivated when the Sentinel-1 image type is selected because these images have no cloud cover. Figure 3 shows the satellite image parameter specification box.

5.2.2 Study Area Definition

The toolbox provides two options to define the region of interest or study area. The first option "**Map drawn boundary**" tells the toolbox to use a boundary drawn by the user in the ***Map viewer***. The boundary can be drawn by the ***polygon***, ***rectangle,*** or a ***circle*** drawing tool in the map viewer as shown in Figure 4.

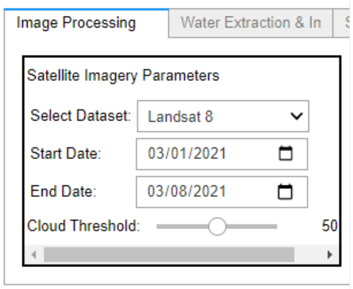


Figure 3 Satellite image parameter specification box

When the ***polygon*** drawing tool is selected, the user continuously clicks on the map with the mouse to create ***points*** to define a polygon boundary. You must click on the ***starting point*** to close the boundary. The ***rectangle*** or ***circle*** drawing tool requires the user to click on the map with the mouse to define the starting point, and drag it to the desired location then release the mouse to finish drawing the boundary. The map viewer displays the ***area*** of the boundary while drawing if the rectangle or circle tool is selected.

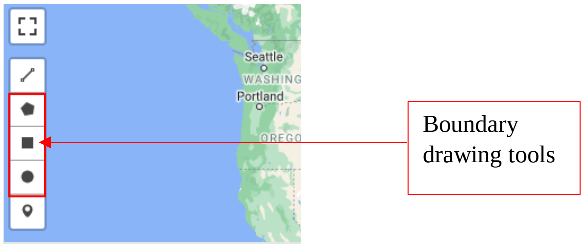
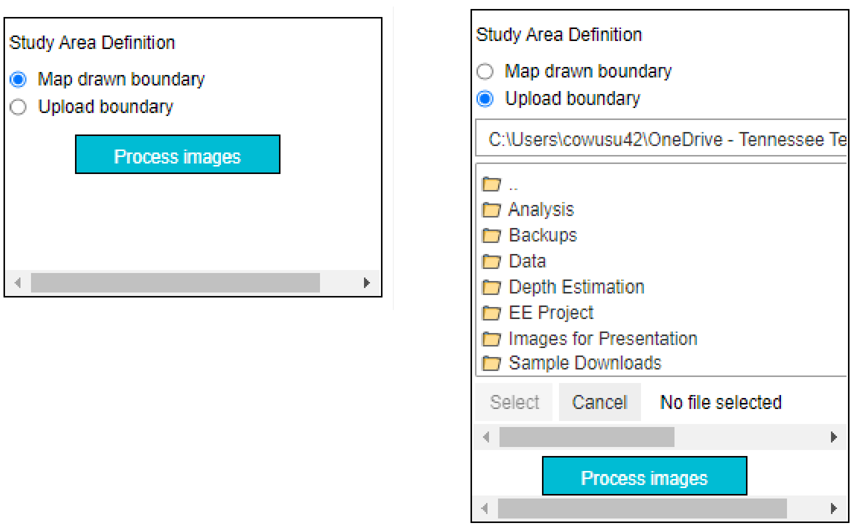


Figure 4 Boundary drawing tools

The second option is ***Upload Shapefile***. This option allows the user to select an ESRI type of shapefile on their local computer which will be converted to a GEE boundary for analysis. Note: the shapefile must be in the ***GCS_WGS_1984*** projection.

With the Upload shapefile option selected, the user must click on the ***Select*** button which opens up the file selection menu. From there, the user can navigate to the desired location on their local computer and select study area shapefile. Figure 5a displays the shapefile *select* button highlighted in red and Figure 5b shows the folder navigation pane.



(a) (b)

Figure 5 Study area definition

5.2.3 Processing

Clicking on the ***Process images*** button (See Figure 5) will retrieve satellite images based on the specified image parameters and the defined study area boundary. The images will be clipped to the defined boundary. The results box will display the number of images processed and a list of the images. The retrieved images will be added to the map viewer for visualization.

5.3 Surface Water Extraction

The ***Water Extraction*** menu as shown in Figure 6 is used to define water extraction parameters such as water extraction index, thresholding, and color. The water index menu provides the options to select user-preferred water extraction index from the available indices. The available indices are described in section 3.1.3. Select a water index by clicking on the ***Water Index dropdown*** widget.

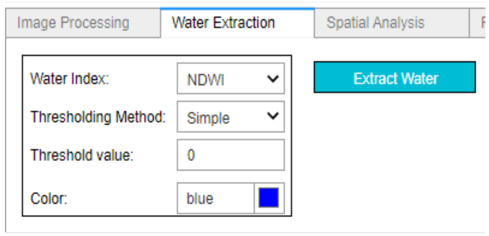


Figure 6 Water extraction menu

The user then selects the water thresholding method using the ***Thresholding Method dropdown*** widget. Available methods are the ***Simple*** and ***Otsu***. Selecting the simple method requires the user to enter a threshold value. The threshold value can be entered into the ***Threshold value text box*** or by clicking the spin up and down arrows to change the value. Clicking the spin arrows increases or decreases the value by ***0.05***. The entered threshold value is applied to all the satellite images retrieved to extract water.

Using the ***Otsu thresholding*** method deactivates the threshold value text box. The optimal threshold required to separate water and non-water pixels are dynamically computed for every image in the satellite image collection. Otsu's method is a histogram-based thresholding approach that computes a threshold that maximizes the inter-class variance between two classes (Otsu, 1979). The toolbox also provides the option to select a ***color*** for display in the map viewer. The color is selected by clicking on the ***color picker*** and selecting the desired color.

Clicking on the ***Extract Water button*** to extract surface water from all satellite images using the selected water index and thresholding method and display the maximum water extent map in the map viewer. The extract water button is ***disabled*** until satellite images for the area of interest have been retrieved and processed. Note: The toolbox will display the maximum water extent image for all images in the study period. The user is encouraged to plot the time series of water changes to see the dynamics in surface water over the study period.

5.4 Spatial Analysis

The Spatial Analysis menu provides two options to compute **Water Frequency Map** and **Water Depth Map**. Water Frequency and Depth Map contain two dropdown menus for specifying an elevation dataset for water depth estimation as shown in Figure 7. Note: The buttons and dropdown menus are disabled until the water extraction process is complete.

Commented [MN3]: See my comment in 3.1.3. Providing recommendations for surface water extraction indicies would improve the manual. For example, are some indicies more accurate? If the user is looking for shallow water, is one index better then another?

Commented [MN4]: How does the thresholding method affect the water extraction?

Commented [OC(5R4]: Requires user familiarity with the water indices.It defers from place to place and depends on image quality too

Commented [MN6]: I noticed that if you have too long of a time period, it takes a long time to process images (no brainer). But if there was some sort of progress bar that might be nice.

Commented [OC(7R6]: Will look into showing a progress

Commented [MN8]: Is there a way to cancel processing? It seems like it kind of gets "stuck" if someone who is unfamiliar with the tool is playing around with it.

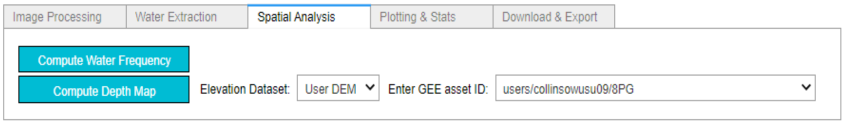


Figure 7 Spatial analysis menu

5.4.1 Water Frequency Map

Click on the **Compute Water Frequency** button to compute the water occurrence frequency from the extracted surface water mask layers for each image in the study period specified. The occurrence frequency is represented by values between 0 and 100%. The percent frequency at any pixel location represents the relative proportion of images that displayed open water at that location. 100% represent permanent open water and 0% represent areas that are rarely inundated.

The water frequency map is displayed in the map viewer together with a legend after the water occurrence frequency computation is complete.

5.4.2 Water Depth Map

To compute the water depth maps, the user first has to select the type of ***elevation dataset*** to be used. The options are the National Elevation Dataset (NED) (applicable to the USA only), Shuttle Radar Topography Mission (SRTM) digital elevation data, or a User-DEM. The default elevation dataset is the NED. Selecting the User-DEM option will show up another dropdown menu with a list of all raster images that the user has uploaded onto their GEE account as assets. Click to select the DEM to be used from the list of assets. Click on the ***Compute Depth Map*** button to compute estimate water depths. Note: the toolbox will display the maximum water depth map. Plot water depth time series for a location to see the dynamics in depth over time.

5.5 Plotting and Statistics

The plotting and statistics menu is made up of four sections; surface water area computation, depth hydrograph plotting, summary statistics, and data saving. Figure 8 shows the plotting and statistics menu.

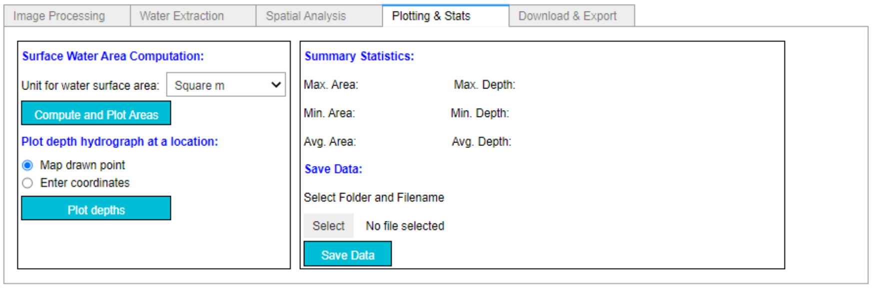


Figure 8 Plotting and Stats menu

5.5.1 Plotting Surface Water Area Hydrographs

To plot the surface water area hydrograph, the user first selects the unit for computing the surface water area. The options for the units are ***square meters***, ***square kilometers***, ***hectares***, and ***acres*** with default units in squares meters. Select the desired units by clicking on the dropdown menu. Click on the ***Compute and Plot Areas*** button to compute the surface water area for each image in the image collection and plot the time series. Figure 9 illustrates a sample hydrograph plot. ***Clicking a point*** in the hydrograph will filter and display the water map and satellite image for the corresponding date in the map viewer.

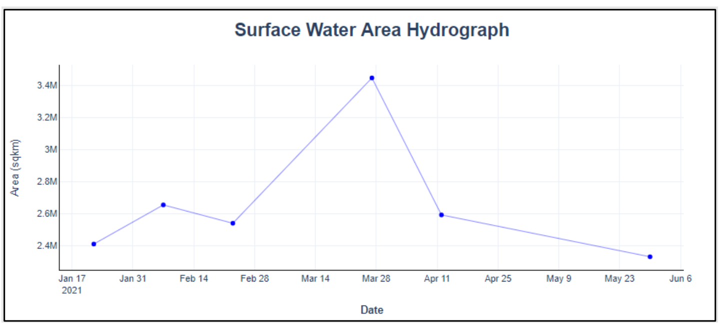


Figure 9 Sample hydrograph

5.5.2 Plotting Depth Hydrographs

Plotting a depth hydrograph requires the user to indicate the location or the point at which the depth hydrograph is required. The location is specified by two options:

Map drawn point: This option allows the user to draw a point using the map viewer tools at the location desired. Note: drawing the point with the polygon, rectangle, or circle tool instead of using the point marker will result in the mean depth of the specified area.

Enter coordinates: This option allows the user to enter the coordinates of the location desired. The coordinates should be entered as longitude first followed by the latitude separated by a comma (e.g. long,lat). The coordinates should be in decimal degrees in the WGS84 Geodetic coordinate system (EPSG:4326). A point will be drawn in the map viewer based on the coordinates entered by the user.

After specifying a point, the user then clicks on the ***Plot depths*** button to generate a water depth hydrograph for the specified location. ***Click a point*** in the depth hydrograph to filter and display the water depth map of the corresponding date in the map viewer. Note: failure to indicate a point of interest will plot the mean depth for the whole study area for each depth map in the generated depth maps.

5.5.3 Summary Statistics

The summary statistics section of the GUI displays the maximum, minimum, and average surface water area and water depths. The surface water area statistics are the statistics from the generated time series of surface water area for the study period. This is similar to the water depth statistics except that it represents only the data for the point location specified and not the entire study area.

5.5.4 Save Data

The toolbox provides the option to save the time-series data of either the surface water areas or water depths to CSV files depending on which data is displayed in the plotter.

To save the data, click on the **Select button** to open the file browser. Browse to the directory where you want to save the data. Enter the ***filename with the file extension*** (.csv) in the output filename box. Click the Select button again to close the file browser. Click the Cancel button if you don't want to save the data. Note: without the file extension, the select button will be disabled. The Select button label changes to ***Change*** which can be clicked if you want to change the filename. The file browser is as shown in Figure 10.

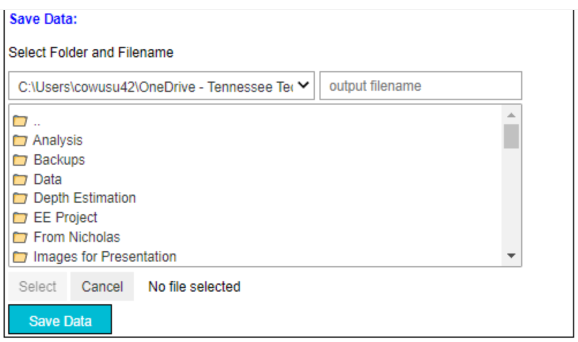


Figure 10 Save data menu

Click the Save Data button to save the data to the filename and directory specified in the file browser.

5.6 Image Download and Export

Satellite data retrieved from the GEE servers and computed water and depth maps can be exported to Google Drive or downloaded to the user's local disk. This is achieved through the ***Download & Export*** menu.

Select from the list of files that can be downloaded. Select download location; Google Drive or Local disk. The ***Google Drive*** option will require the user to enter a folder name which will be used to create a directory in the user's Google Drive account to contain the exported images. The ***Local Disk*** option requires the user to select a folder on the user's computer using the available file browser.

6 References

Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, *140*, 23–35. https://doi.org/10.1016/j.rse.2013.08.029

Jones, J. W. (2019). Improved automated detection of subpixel-scale inundation-revised Dynamic Surface Water Extent (DSWE) partial surface water tests. *Remote Sensing*, *11*(4). https://doi.org/10.3390/rs11040374

Lee, J.-S. (1983). A simple speckle smoothing algorithm for synthetic aperture radar images. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(1), 85–89. https://doi.org/10.1109/TSMC.1983.6313036

McFEETERS, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, *27*(14), 3025–3033. https://doi.org/10.1080/01431160600589179