ERDC/EL SR-22-6

US Army Corps of Engineers®
Engineer Research and Development Center

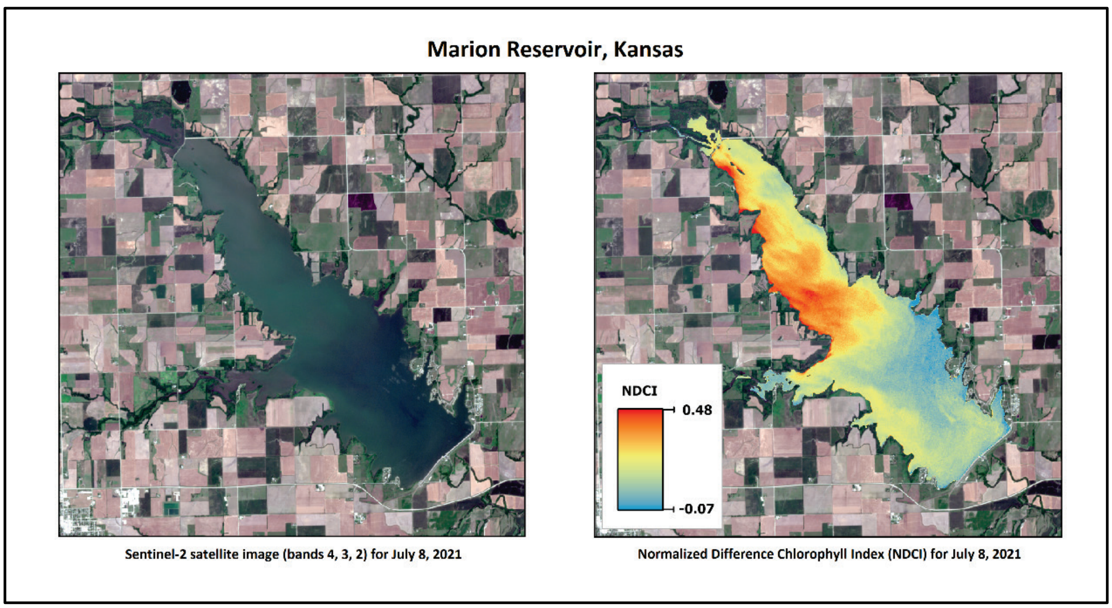
Aquatic Nuisance Species Research Program

waterquality for ArcGIS Pro Toolbox

User's Guide

Christina Saltus, Molly Reif, and Richard Johansen

September 2022



Approved for public release; distribution is unlimited.

Environmental Laboratory

Aquatic Nuisance Species Research Program

ERDC/EL TR-22-6
September 2022

waterquality for ArcGIS Pro Toolbox: User's Guide

Christina Saltus and Richard Johansen

Environmental Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199, USA

Molly Reif

Environmental Laboratory
U.S. Army Engineer Research and Development Center
Joint Airborne Lidar Bathymetry Technical Center of Expertise (JALBTCX)
Kiln, MS, 39556, USA

Final report

Prepared for Aquatic Nuisance Species Research Program US Army Engineer Research and Development Center Environmental Laboratory
Vicksburg, MS 39180-6199

Under AMSCO Code 008284
Funding Account U4377102

Abstract

Monitoring water quality of small inland lakes and reservoirs is a critical component of the US Army Corps of Engineers (USACE) water quality management plans. However, limited resources for traditional field-based monitoring of numerous lakes and reservoirs covering vast geographic areas often leads to reactional responses to harmful algal bloom (HAB) outbreaks. Satellite remote sensing methodologies using HAB indicators is a good low-cost option to traditional methods and has been proven to maximize and complement current field-based approaches while providing a synoptic view of water quality (Beck et al. 2016; Beck et al. 2017; Beck et al. 2019; Johansen et al. 2019; Mishra et al. 2019; Stumpf and Tomlinson 2007; Wang et al. 2020; Xu et al. 2019; Reif 2011). To assist USACE water quality management, we developed an Environmental Systems Research Institute (ESRI) ArcGIS Pro desktop software toolbox (*waterquality for ArcGIS Pro*) founded on the design and research established in the *waterquality* R software package (Johansen et al. 2019; Johansen 2020). The toolbox enables the detection, monitoring, and quantification of HAB indicators (chlorophyll-a, phycocyanin, and turbidity) using Sentinel-2 satellite imagery. Four tools are available: (1) automating the download of Sentinel-2 Level-2A imagery, (2) creating stacked image with options for cloud and non-water features masks, (3) applying water quality algorithms to generate relative estimations of one to three water quality parameters (chlorophyll-a, phycocyanin, and turbidity), and (4) creating linear regression graphs and statistics comparing in situ data (from field-based water sampling) to relative estimation data. This document serves as a user's guide for the *waterquality for ArcGIS Pro* toolbox and includes instructions on toolbox installation and descriptions of each tool's inputs, outputs, and troubleshooting guidance.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for the Aquatic Nuisance Species Research Program (ANSRP), “Harmful Algal Bloom indicator estimation in small inland waterbodies: Remote sensing-based software tools to assist with USACE water quality monitoring” under AMSCO Code 008284 and Funding Account U4377102. The ANSRP is sponsored by Headquarters, US Army Corps of Engineers (USACE), and is assigned to the US Army Engineer Research and Development Center (ERDC) under the purview of the Environmental Laboratory (EL), Vicksburg, MS. The ANSRP Program Manager was Mr. Michael Greer.

The work was performed by the Environmental Systems Branch of the Ecosystem Evaluation and Engineering Division, ERDC-EL. Technical peer reviews were conducted by Mr. Sam Jackson and Mr. Scott Bourne of the ERDC-EL.

At the time of publication, Mr. Mark R. Graves was the Chief of the Environmental Systems Branch; Mr. Mark D. Farr was Chief of Ecosystem Evaluation and Engineering Division; Dr. Christine VanZomeren was the Associate Technical Director for Civil Works Environmental Engineering and Sciences; Dr. Jennifer Seiter-Moser was the Technical Director for Civil Works Environmental Engineering and Sciences; and the ERDC-EL Director was Dr. Edmond J. Russo.

The Commander of ERDC was COL Christian Patterson and the Director was Dr. David W. Pittman.

1 Introduction

1.1 Background

The US Army Corps of Engineers (USACE) has a fundamental water quality mission, yet the challenge of monitoring hundreds of lakes and reservoirs covering vast geographic areas is hindered by limited resources, often leading to reactionary responses to harmful algal bloom (HAB) outbreaks. Traditional field-based water sampling methods are relied upon for water quality monitoring and are a critical component of USACE water quality management plans. Advances in remote sensing technology serve to maximize and complement current field-based approaches and to improve monitoring water quality indicators of HABs (Beck et al. 2016; Beck et al. 2017; Beck et al. 2019; Johansen et al. 2019; Mishra et al. 2019; Stumpf and Tomlinson 2007; Wang et al. 2020; Xu et al. 2019). Specifically, spatial and temporal variations of water quality may not be systematically captured across an entire waterbody using traditional monitoring methods. However, multispectral satellite imagery can provide an enhanced monitoring approach by providing a synoptic view of water quality (Reif 2011). In addition, remotely sensed imagery offers capabilities for concurrent analysis of water quality in numerous waterbodies (Xu et al. 2019). Although, satellite remote sensing techniques utilized in water quality management and monitoring cannot be used to assess toxicity directly. The technology can help reduce costs, labor, and time by focusing field-based water sampling in areas of concern and serve as an early warning approach for proactive monitoring and management of HAB outbreaks.

1.2 Objective

To address the growing need to rapidly monitor USACE water quality in small, inland waterbodies, the US Army Engineer Research and Development Center's (ERDC) Aquatic Nuisance Species Research Program (ANSRP) funded the work unit titled, "Harmful Algal Bloom indicator estimation in small inland waterbodies: Remote sensing-based software tools to assist with USACE water quality monitoring." The main objective of the research was to expand the development of remote sensing tools used to estimate potential HAB indicators: (1) chlorophyll-a (chl-a), (2) phycocyanin, a proxy for cyanobacterial or blue-green algal biomass

(BGA/PC), and (3) turbidity, focusing on small, inland waterbodies. An array of software-based tool options with decreasing complexity have been developed including the following: (1) an open-source R software package, *waterquality* (Johansen et al. 2019; Johansen et al. 2020), (2) an Environmental Systems Research Institute (ESRI) ArcGIS Pro desktop software toolbox (*waterquality for ArcGIS Pro*), and (3) an ESRI-based online web application, (*Harmful Algal Bloom [HAB] Explorer)*. The primary focus of this document is the *waterquality for ArcGIS Pro* toolbox which was founded on the design and research established in the *waterquality* R software package (Johansen et al. 2019; Johansen 2020) and builds on the extensive body of remote sensing research conducted by the USACE ERDC-EL and partners (Johansen et al. 2019; Beck et al. 2019; Johansen et al. 2018; Reif 2011).

1.3 Approach

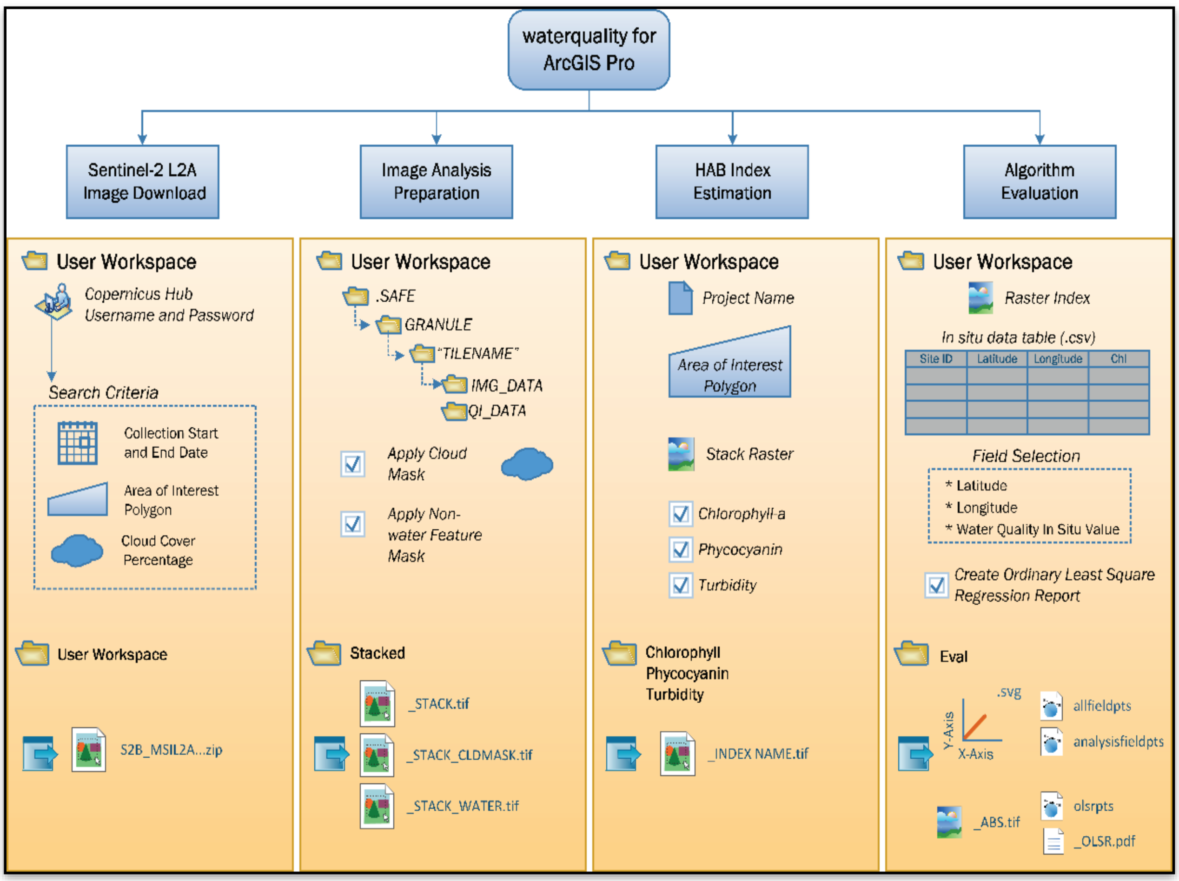
This document is intended to serve as a user's guide describing the functionality of the python-based *waterquality for ArcGIS Pro* toolbox developed for use in ESRI ArcGIS Pro desktop software. The toolbox enables the detection, monitoring, and quantification of HAB indicators (water quality parameters) using Sentinel-2 satellite imagery. It is not intended for use in predicting HAB occurrences. The use of Sentinel-2 imagery is supported by multiple research studies that have successfully applied remote sensing methodologies for a variety of aquatic and coastal applications and capitalizes on global coverage and an open-access data policy for freely available imagery (Bramich et al. 2021; Bergsma and Almar 2020; Toming et al. 2016).

Four tools are available in the *waterquality for ArcGIS Pro* toolbox (Figure 1) and are described in the following sections:

1. Sentinel-2 L2A Image Download Tool automates the download of Sentinel-2 Level-2A (Bottom-Of-Atmosphere [BOA] Reflectance) satellite imagery by accessing the European Space Agency's (ESA) Copernicus Open Access Hub via the Sentinel Application Programming Interface (API) using *sentinelsat* Python library (Valgur et al 2019).
2. Image Analysis Preparation Tool prepares Sentinel-2 satellite imagery for analysis by combining all image bands into one file for use in the following tools. Options are also provided for masking/removing cloud and non-water features from the image scene.

3. HAB Index Estimation Tool applies water quality algorithms to Sentinel-2 satellite imagery to generate a map illustrating relative estimations of one to three water quality parameters (chlorophyll-a, phycocyanin, and turbidity) in small, inland waterbodies.
4. Algorithm Evaluation automates the creation of a linear regression graph and statistics, comparing in situ data (from field-based water sampling) and relative estimation data (generated from the HAB Index Estimation Tool). Additionally, a HAB concentration map illustrating absolute water quality parameter values can be generated if the R^2 value meets or exceeds the 0.70 threshold.

Figure 1. waterquality for ArcGIS Pro toolbox.



1.4 Sentinel-2 overview

The *waterquality for ArcGIS Pro* toolbox uses imagery collected from the MultiSpectral Instrument (MSI) aboard the Copernicus Sentinel-2 satellite platform. The Sentinel-2 mission includes a constellation of two satellites (2A and 2B) operated by the ESA providing multispectral imagery at 10- to 60-m spatial resolution with a 5-10 day revisit time (https://sentinel.esa.int/web/sentinel/missions/sentinel-2). Thus, the collective advantages of Sentinel-2, including high to moderate spatial resolution, free and open access data policy, and high revisit frequency, stand to benefit and improve HAB monitoring capabilities. Furthermore, two additional satellites (2C and 2D) are planned for inclusion in the constellation in 2024-2031 to reduce the revisit time by half, thereby increasing product availability.

The ESA offers orthorectified data products collected from the MSI at two product levels: Level-1C (L1C) and Level-2A (L2A). The L1C product is a 100 x 100 km^2 tiled image with radiometric measurements provided as Top-Of-Atmosphere (TOA) reflectance values. As a result, additional processing is required to correct for atmospheric effects and to render the imagery suitable for spectral analyses. Therefore, this product level is not currently supported in the *waterquality for ArcGIS Pro* toolbox (for more information, https://sentinel.esa.int/web/sentinel/toolboxes/sentinel-2). In contrast, the L2A product is supported in the toolbox and is an atmospherically corrected image illustrating BOA surface reflectance values in which atmospheric scattering and absorption effects are removed. L2A products are derived from the L1C product and have been routinely available from the ESA since 2018. The product tiles are comprised of 13 individual spectral band images representing specified ranges of the electromagnetic spectrum (Table 1). Each band is available at one or more of the three spatial resolutions: 10, 20, and 60 m. The product tiles are stored in a Standard Archive format for Europe (.SAFE) file format containing multiple subfolders such as the IMG_DATA folder comprised of spectral band images in JPG2000 (jp2) file format. For more information on the .SAFE file structure and naming conventions, please refer to the technical guides provided on the ESA Sentinel Online website (https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-2a/product-formatting).

Table 1. Sentinel-2 spectral band description, wavelength, and spatial resolution. NIR = near infrared, SWIR=short wave infrared (Source: https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial).

Spectral Band	Description	Center Wavelength (nm)	Initial Spatial Resolution (m)	Level-2A Available Product Spatial Resolution (m)
Band 1	Coastal Aerosol	443	60	60
Band 2	Blue	490	10	10; 20; 60
Band 3	Green	560	10	10; 20; 60
Band 4	Red	665	10	10; 20; 60
Band 5	Vegetation Red Edge	705	20	20; 60
Band 6	Vegetation Red Edge	740	20	20; 60
Band 7	Vegetation Red Edge	783	20	20; 60
Band 8	NIR	842	10	10
Band 8A	Narrow NIR	865	20	20; 60
Band 9	Water Vapor	940	60	60
Band 10	SWIR-Cirrus	1375	60	NA
Band 11	SWIR	1610	20	20; 60
Band 12	SWIR	2190	20	20; 60

2 Toolbox User's Guide

The *waterquality for ArcGIS Pro* toolbox is an ArcGIS Pro desktop software geoprocessing toolbox that streamlines HAB monitoring analysis. The subsections below explain the system requirements, toolbox file structure, and installation instructions, while Chapters 3 through 6 describe individual tools within the toolbox, including the Sentinel-2 L2A Image Download tool, Image Analysis Preparation tool, HAB Index Estimation tool, and the Algorithm Evaluation tool. This guide provides a summary of each tool including descriptions of data inputs and outputs, the required file formats, and tool workflow. The output data generated from the tools contain spatial reference information enabling independent data use or used in conjunction with other geospatial layers. While advanced knowledge of ArcGIS Pro desktop software is not necessary for using the toolbox, basic knowledge of the software is beneficial.

2.1 System requirements

The *waterquality for ArcGIS Pro* toolbox was developed in ArcGIS Pro 2.7 desktop software, but specifically tested on Versions 2.7 through 2.8.1 and is based on the Python® programming language (Version 3.7). The toolbox requires a Windows 10 operating system environment, an Advanced ArcGIS Pro license with the Spatial Analyst Extension, and Python version 3.7 (ESRI 2021). ArcGIS Pro is freely available to all USACE through an Enterprise License Agreement (ELA) and can be downloaded via the ACE-IT software app portal. Python is automatically installed and available for use with ArcGIS during the default installation. However, the toolbox requires the installation of an additional Python library, *sentinelsat* (Valgur et al. 2019), for the automated download of Sentinel-2 satellite imagery via the Copernicus Open Access Hub. The installation instructions for *sentinelsat* are provided in Section 2.3.

2.2 File structure

The files comprising the *waterquality for ArcGIS Pro* toolbox are provided in a zip file, waterquality.zip, which can be downloaded from ERDC Knowledge Core (http://dx.doi.org/10.21079/11681/42240) along with the user's guide and extracted to your local computer (e.g., ACE-IT or RDE for USACE personnel). Also available for download is a subset of sample data products generated by the toolbox and provided in the zip file, SampleData.zip. The sample data allows the user to explore some hypothetical examples of data inputs and outputs. Table 2 provides a list of file names and descriptions for each file included in the zip files.

Table 2. Files in the waterquality.zip file, comprising the waterquality for ArcGIS Pro toolbox.

File Name	Description
waterquality.pyt	Principal ArcGIS Python toolbox file which manages and directs all related python code
waterquality.pyt.xml	Summary information about the toolbox
waterquality.HabStats.pyt.xml	Help documentation for the Algorithm Evaluation tool
waterquality.HIndex.pyt.xml	Help documentation for the HAB Index Estimation tool
waterquality.HS2Prep.pyt.xml	Help documentation for the Image Analysis Preparation tool
waterquality.S2Download.pyt.xml	Help documentation for the Sentinel-2 L2A Image Download tool
Scripts folder	
ClipRaster2AOI2.py	Python code, as part of the HAB Index Estimation tool, for clipping input imagery by area of interest
HAB_CalcStat.py	Python code for the Algorithm Evaluation tool
HAB_ImagePrep.py	Python code for the Image Analysis Preparation tool
HAB_sentinel2download.py	Python code for the Sentinel-2 L2A Image Download tool
HAB_Index.py	Python code for the HAB Index Estimation tool
SampleData folder	
S2A_MSIL2A_20210708T170851_N0301_R112_T14SPH_20210708T213140.zip	Sentinel-2 L2A image download in zip file format
T14SPH_20210708_20m_STACK.tif	Sentinel-2 image stack, 20-meter bands 1-8A
T14SPH_20210708_STACK_WATER.tif	Sentinel-2 image stack, 20-meter bands 1-8A with non-water feature mask
T14SPH_20210708_NDCI.tif	Normalized Difference Chlorophyll Index map product
Hypothetical_insitu_pts.csv	Hypothetical in situ sample point data table in CSV format
MarionReservoir.shp, shx, dbf, prj	Area of interest example, Marion Reservoir, Kansas
T14SPH_20210708_NDCI_HAB.gdb	Hypothetical in situ sample points as point feature layer includes allfieldpts, analysisfieldpts, and OLSRPT
T14SPH_20210708_NDCI.svg	Hypothetical linear regression graph and statistics
T14SPH_20210708_NDCIOLSR.pdf	Hypothetical ordinary least square regression report and statistics
T14SPH_20210708_NDCI_ABS.tif	Hypothetical absolute HAB concentration map product

2.3 Installation

2.3.1 Python library installation

Python is a programming language used to automate tasks through a standardized and well-established coding syntax. It is widely used for its simplicity and compatibility with multiple software applications and operating systems (Nagpal and Gabrani 2019). It also offers a large array of libraries, which are collections of modular code. *Sentinelsat* is a Python library that provides a Python API for the search and download of Sentinel-2 satellite imagery via a connection to the Copernicus Open Access Hub. Prior to using the *waterquality for ArcGIS Pro* toolbox, a one-time installation of the *sentinelsat* Python library is required through the Python Package Manager. ArcGIS Pro allows for the management of Python libraries through the Python Package Manager by consolidating all required Python libraries into distinct Python environments. Steps are listed below to for the installation process.

- First, clone the arcgispro-py3 Python environment. To clone the environment, open ArcGIS Pro desktop software and select the settings icon at the bottom left of the application window (Figure 2 red arrow).
- Navigate to the Python Package Manager dialog box by selecting *Python* on the menu to the left (Figure 3) and clicking the Manage Environment button located under the Project Environment.

Figure 2. ArcGIS Pro settings.

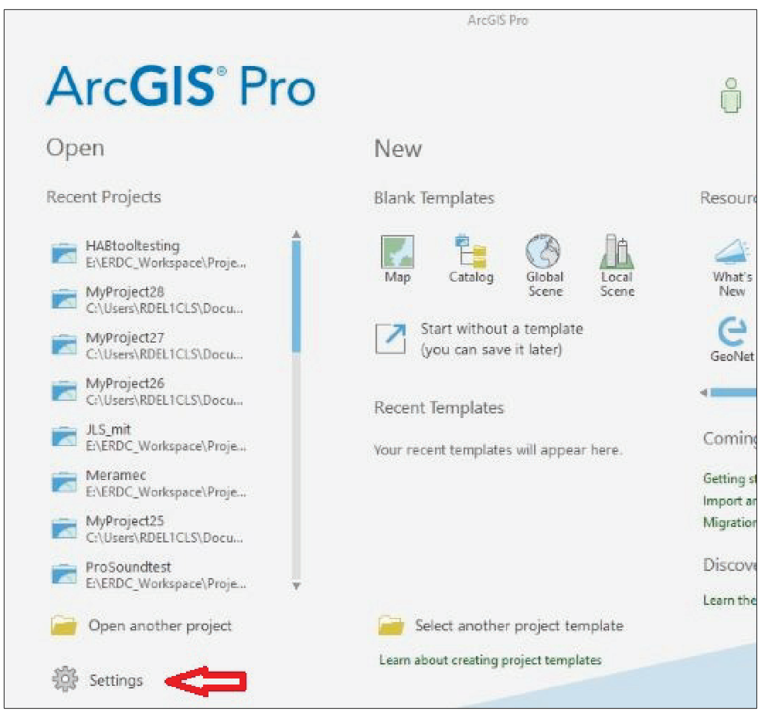
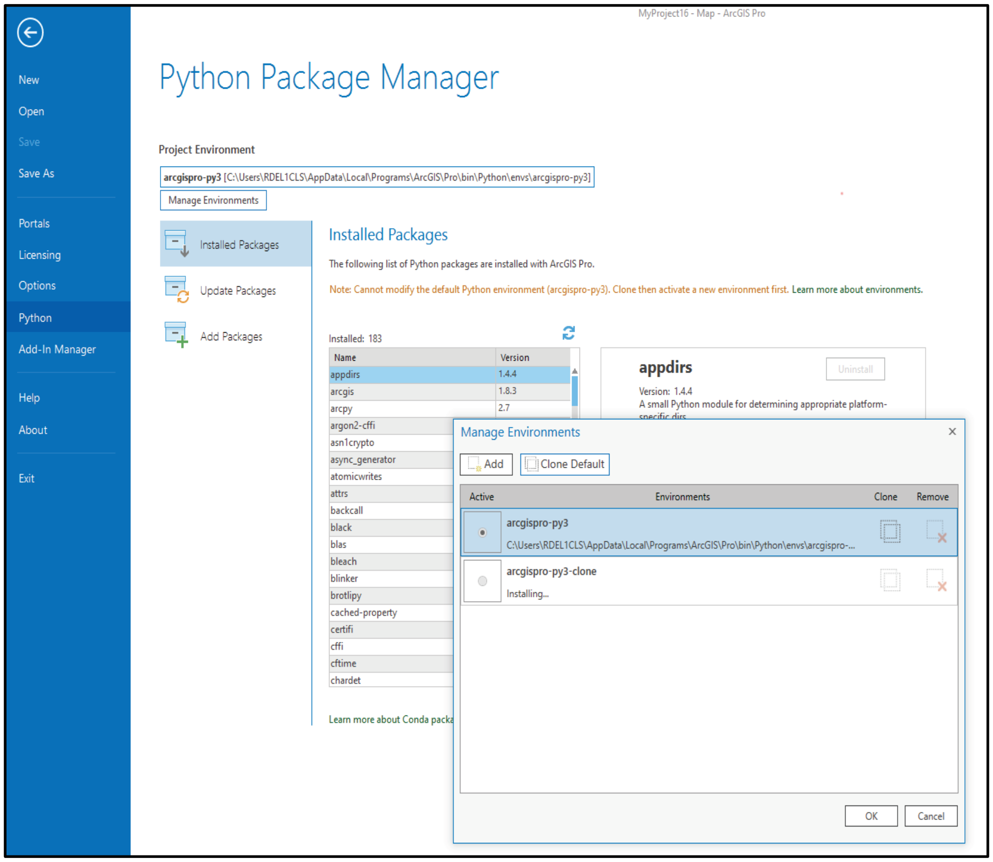


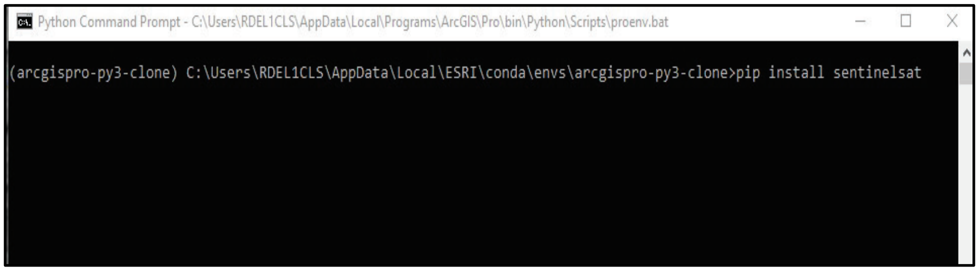
Figure 3. Python Package Manager dialog box in ArcGIS Pro desktop software.



- Next, in the Manage Environments dialog box, select the arcgispro-py3 environment radio button, then click the *Clone Default* button to clone the environment. The cloning process will take approximately 15 min.
- Once cloning is complete, exit ArcGIS Pro desktop software and restart the program. Then, return to the Manage Environment dialog box in the Python Package Manager, and select the arcgispro-py3-clone environment radio button (Figure 3). With the Python environment successfully cloned and selected as the current environment, the *sentinelsat* Python library can then be installed.

From the Windows Start menu, select the Python command prompt under the ArcGIS pull down menu. The arcgispro-py3-clone should be the environment shown in parenthesis in the command prompt window (Figure 4). If the correct environment is not shown, return to the Python Package Manager, select the cloned environment, and restart the Python Command Prompt. At the prompt, type the following: *pip install sentinelsat*.

Figure 4. Python command prompt window.

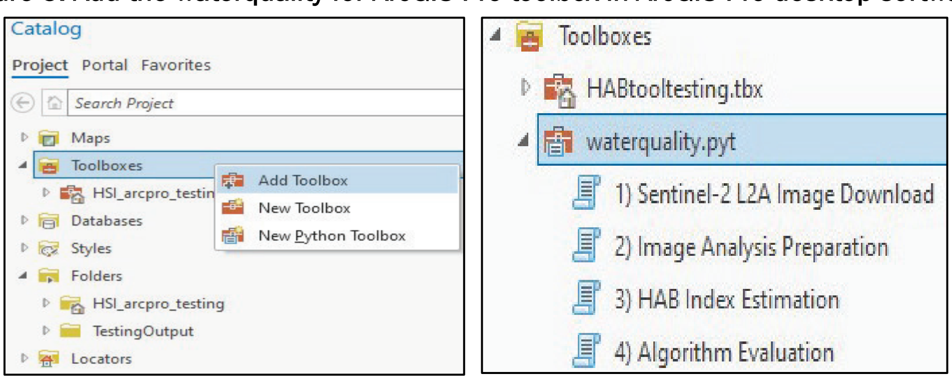


Once the library is installed successfully, close the Python command prompt window, and return to ArcGIS Pro desktop software. The *waterquality for ArcGIS Pro* toolbox is ready for installation. Instructions for toolbox installation are provided below.

2.3.2 waterquality for ArcGIS Pro Toolbox installation

Prior to installing the toolbox, verify the ArcGIS Pro license level is set to Advanced and the Spatial Analyst Extension is activated. To check your license status, open the ArcGIS Pro Settings (Figure 2) and select Licensing from the left panel menu (Figure 3). If both are set to *Yes* under the Licensed column and license has not expired, toolbox installation can begin. To install the *waterquality for ArcGIS Pro* toolbox, download the toolbox and extract the zip file, waterquality.zip, to a local folder on your computer. In the ArcGIS Pro desktop software catalog window, right click on the Toolboxes folder and select, *Add Toolbox* (Figure 5). Navigate to the local folder, select the toolbox file (waterquality.pyt), and click *OK*.

Figure 5. Add the waterquality for ArcGIS Pro toolbox in ArcGIS Pro desktop software.

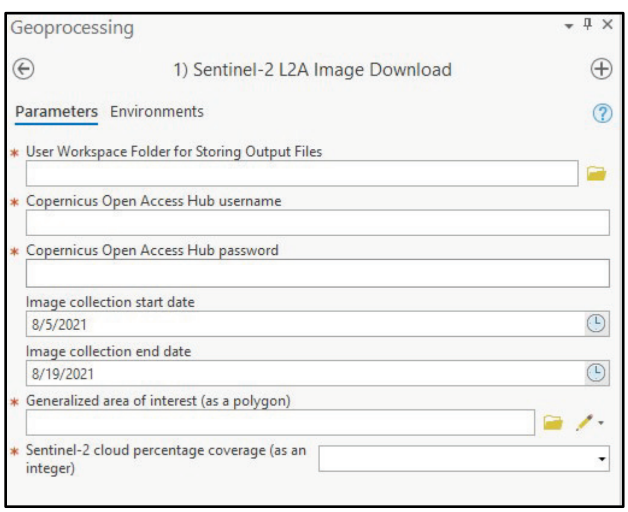


3 Sentinel-2 L2A Image Download Tool

The Sentinel-2 L2A Image Download tool automates the download of Sentinel-2 L2A (BOA reflectance; atmospherically corrected) imagery by accessing the ESA Copernicus Open Access Hub ("the Hub") via the Sentinel API (using *sentinelsat* Python library). Several parameters are used to define the search criteria for locating and downloading Sentinel imagery in the specified area of interest. These criteria include the username and password for the Hub, the start and end date time range for image collection, a generalized polygon defining an area of interest (AOI), and the percentage of acceptable cloud cover. Prior to using this tool, the user must have a Hub account (https://scihub.copernicus.eu). If the user does not have an account, instructions for creating one and requirements for username and password are provided at https://scihub.copernicus.eu/userguide/SelfRegistration.

Since the data products are provided for download as individual tile squares, each covering approximately 100 x 100 km^2, the AOI polygon is used to select intersecting tiles and further filters the tile selection based on the criteria. Then, 12 of the most recent image tiles with the lowest cloud cover percentage (0%-35%) and availability will be downloaded. The downloaded data will be stored in the user-defined workspace as individual zip files for each tile and date. The following subsections describe the tool's inputs, outputs, and troubleshooting guidance (Figure 6).

Figure 6. Sentinel-2 L2A Image Download tool dialog box in ArcGIS Pro desktop software.



3.1 Inputs

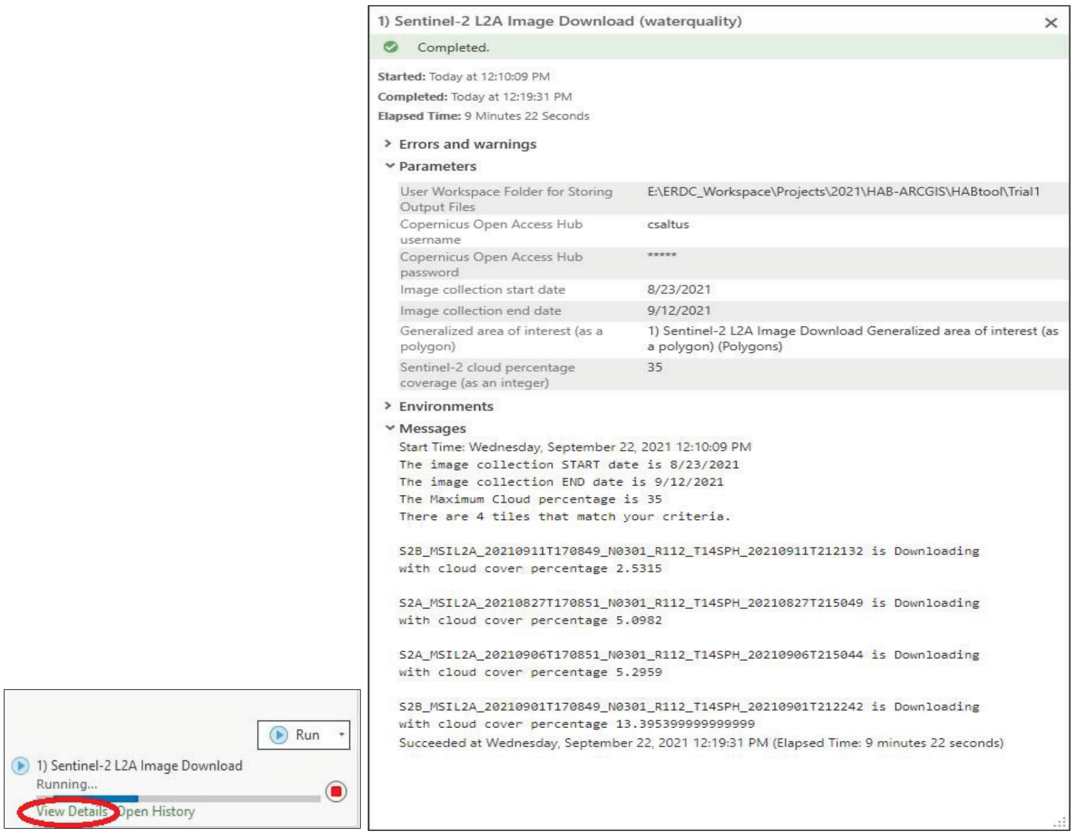
- User Workspace: Select the workspace folder where the output data will be stored, including temporary files generated by the tool (*i.e., C:\Project*). The output for each Sentinel-2 L2A tile is in a zip file format.
- Copernicus Open Access Hub Username: Enter your Copernicus Open Access Hub username. Instructions for creating a Hub account are available at https://scihub.copernicus.eu/userguide/SelfRegistration and must be established prior to running the tool.
- Copernicus Open Access Hub Password: Enter your Copernicus Open Access Hub password. Instructions for creating a Hub account are available at https://scihub.copernicus.eu/userguide/SelfRegistration and must be established prior to running the tool. Note that special character limitations apply.
- Start Date: Enter image collection START date by selecting the clock button on the right and selecting the month, day, and year of interest. The date format should be mm/dd/yyyy (i.e., 01/01/2021) and should not include a time component. The goal is to define a date range for searching Sentinel-2 imagery collection. This input represents the beginning date of the date range. The default date is two weeks prior to today's date.
- End Date: Enter image collection END date by selecting the clock button on the right and selecting the month, day, and year of interest. The date format should be mm/dd/yyyy (i.e., 01/01/2021) and should not include a time component. The goal is to define a date range for searching Sentinel-2 imagery collection. This input represents the ending date of the date range. The default date is today's date.
- AOI: Input a generalized AOI polygon by either selecting a polygon feature layer or using the draw tool to create a polygon. The polygon will be used to select the Sentinel-2 image tiles that overlap the AOI, so it does not need to be detailed, 4-5 polygon vertices are sufficient.
- Cloud Cover: Select the Sentinel-2 cloud cover percentage from the pull-down list [available coverage range is between 0-35%]. The goal for this input is to select image tiles that have limited cloud cover, thereby limiting the possibility of clouds over the area of interest.

3.2 Outputs

The tool outputs individual zip files for each image tile and date meeting the search criteria and are available for download. The zip file is saved to the user defined workspace (*i.e., C:\SampleData\S2B_MSIL2A_20210515T163839_*

N0300_R126_T15RYP_20210515T204604.zip). Informative messages concerning the tool progress are provided in the Message tab by clicking the *View Details* link at the bottom of the tool window (Figure 7). The imagery within the zip file is used as input into the Image Analysis Preparation tool. Please refer to Chapter 4 for important instructions on extracting data from the zip files.

Figure 7. Sentinel-2 L2A Image Download tool View Details dialog.



3.3 Troubleshooting guidance

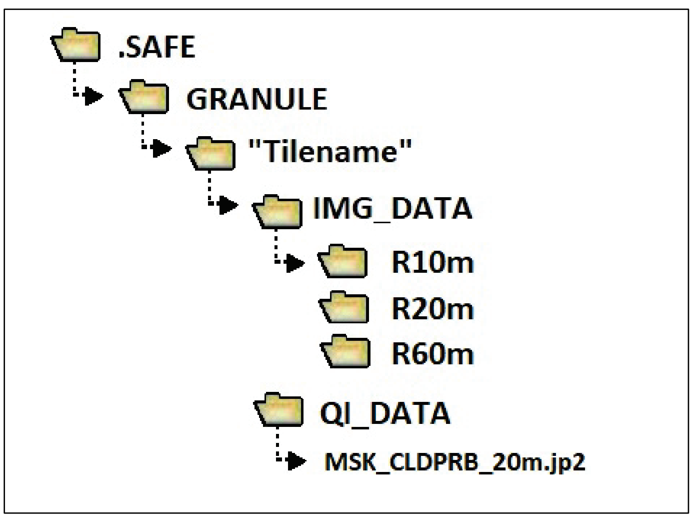
If issues occur while downloading imagery through the tool, check the search parameters entered as data may not be available for the specified time periods, or the percentage of cloud cover in the imagery may exceed the defined threshold. Error messages are also provided in the Message tab by clicking the *View Details* link at the bottom of the tool window (Figure 7). Occasionally, the Hub website may be unavailable due to scheduled maintenance or other mission-related issues. Refer to the Hub website (https://scihub.copernicus.eu) for updates on the sensor collection and image availability.

4 Image Analysis Preparation Tool

The Image Analysis Preparation tool is designed to prepare one or more Sentinel-2 image tiles downloaded using the Sentinel-2 L2A Image Download tool for the application of water quality algorithms. The tool automates the resampling of spectral band images into a multiband, 20-m resolution image by stacking bands 1-8A (Table 1). The tool also provides two options for masking cloud and land features from the image stack.

Since the spectral bands of the Sentinel-2 imagery are provided in varying spatial resolutions, it is important to standardize the output resolution, so each band is comparable. Therefore, the tool automates the creation of a multiband stacked image at 20 m spatial resolution by first accessing the individual Sentinel-2 band images within the three subfolders under the IMG_DATA folder: R10m, R20m, and R60m (10 m, 20 m, and 60 m; Table 2; Figure 8). Then, band 1 is resampled from 60 m to 20 m to match the corresponding spatial resolution of bands 2-8A. Next, each 20-m band is combined into one 8-band image stack and stored as a .tif file in the user defined workspace.

Figure 8. Sentinel-2 L2A Image file folder structure.



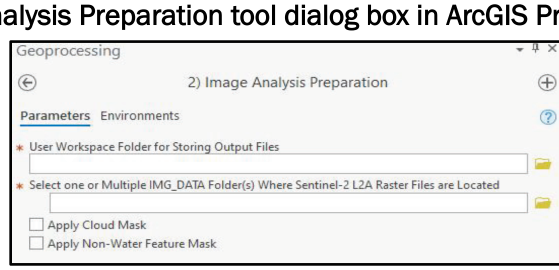
The detection and removal of cloud and land features from satellite images is an essential preprocessing step for effective HAB monitoring. Therefore, limiting these features in the image stack is advantageous. The tool offers two options for masking these features: (1) a cloud mask, and (2) a non-water feature mask.

1. The cloud mask option utilizes the cloud probability product provided with the Sentinel-2 L2A data download and applies a threshold of 45% or greater cloud probability to create a binary mask file. Then, the mask is applied to the image stack to remove the maximum amount of clouds possible and stores the output as a .tif file in the user-defined workspace. This output can then be used to extract water areas based on an ESRI polygon feature layer (i.e., detailed lake boundary) in the HAB Estimation tool. However, if a lake boundary file is not available, water areas can also be extracted by using the non-water feature mask (next option).
2. The non-water feature mask option utilizes areas identified as water within the scene classification product provided with the Sentinel-2 L2A data download and automatically extracts those features from the image stack, thereby removing all non-water features including clouds.

While the cloud probability and scene classification products are generally effective for rapid removal of undesired areas, some amount of error is expected in the products. Thus, it is possible that some clouds or non-water features are not removed, or conversely, that some water features are inadvertently removed. Therefore, certain cases may require additional methods to remove or include desired features outside the tool. Some potential alternatives are discussed in Section 4.2.

Prior to running this tool, the zip file downloaded in the Sentinel-2 L2A Image Download tool will need to be extracted. (*i.e., C:\Project\HABWORK\S2B_MSIL2A_20210515T163839_N0300_R126_T15RYP_20210515T204604.zip*). File names generated by the ESA are long; therefore, the zip files must be extracted in a folder location where the entire path name, starting at the root directory does not exceed 200 characters or the user will receive an error. All subfolders under .SAFE folder should be extracted from the zip file. Figure 8 provides an example of the .SAFE file folder structure, while Figure 9 displays the Image Analysis Preparation tool dialog box for entry of inputs which are described in Section 4.1.

Figure 9. Image Analysis Preparation tool dialog box in ArcGIS Pro desktop software.



4.1 Inputs

- User Workspace: Select User Workspace Folder for Storing Output Files. Click the folder button and browse to the folder location for saving output files. The tool will automatically create two folders named *Stacked* and *ResampledFromTool* under the user specified workspace folder (*i.e., C:\Project\Stacked*). The latter is created for temporary storage of resampled band imagery and is deleted upon tool run completion.
- IMG_DATA: Select One or Multiple IMG_DATA Folder(s) Where Sentinel-2 L2A Raster Files are Located.

  Within the folder structure, the IMG_DATA folder is located in three subfolders below the .SAFE folder (Figure 8; i.e. C:\S2A_MSIL2A_20210708T170851_N0301_R112_T14SPH_20210708T213140\S2A_MSIL2A_20210708T170851_N0301_R112_T14SPH_20210708T213140.SAFE\GRANULE\L2A_T14SPH_A031569_20210708T171925\IMG_DATA). The files within the subfolders are used to create a multiband image stack at 20-m resolution.

  For more information on file naming conventions and folder structure for the Sentinel-2 data products, please refer to the user's guide for Sentinel-2 MSI at https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi

- Cloud Mask: A checkmark placed in the box next to *Apply Cloud Mask* will automatically apply a cloud mask to the raster stack. If the box is unchecked, a cloud mask will not be applied.

  The cloud mask is derived from the Sentinel-2 L2A cloud probability product provided in the downloaded product under the QI_Data folder with the filename MSK_CLDPRB_20m.jp2 (Figure 8). A 45% or greater cloud probability threshold from this product is used to automatically mask clouds in the image stack. Note that this method may not remove all clouds from your AOI; therefore, it is important to review all output files prior to applying water quality algorithms through the HAB Index Estimation tool. Alternative mask options are discussed in Section 4.2.

- Non-Water Feature Mask: A checkmark placed in the box next to *Apply Non-Water Feature Mask* will automatically select only water areas while masking land, clouds, and cloud shadows from the raster stack. If the box is unchecked, the water feature image will not be created. The water feature image is derived from an attribute mask on the image stack using the water features (value=6) from the Sentinel-2 L2A scene classification layer (SCL) provided with the downloaded product under the IMG_Data\R20m folder. The filename will contain the characters “SCL” (i.e., T14SPH_20210708 T170851_SCL_20m.jp2). Mask options are discussed in Section 4.2.

4.2 Mask options

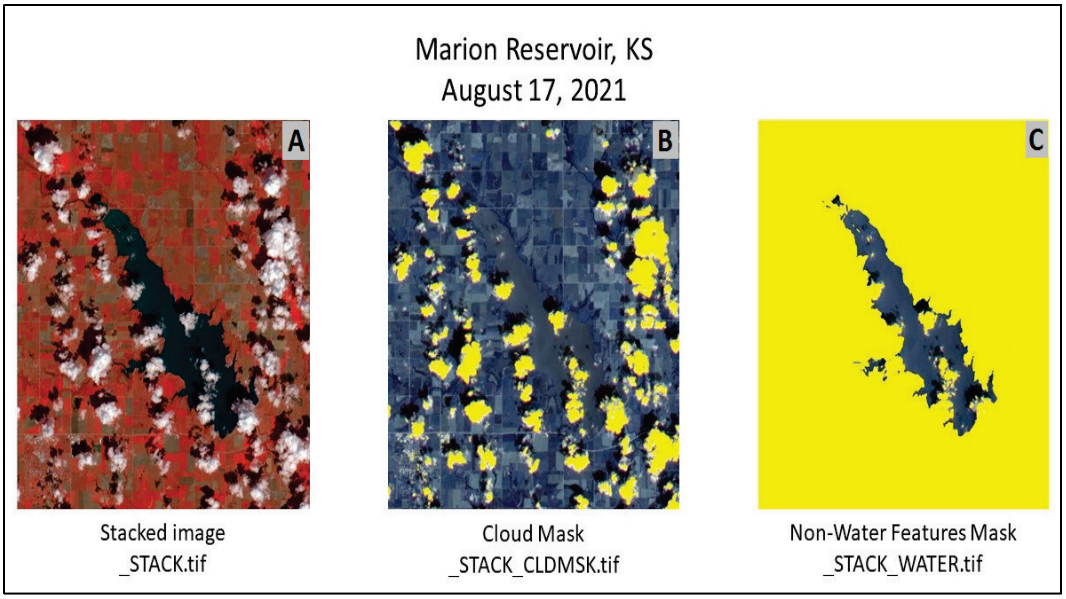
Given specific conditions, cloud and land features can be over or under classified in imagery such as the misclassification of cirrus clouds over water as water or thick algal mats (surface scum) as land. Therefore, alternative methods outside the tool may be needed for evaluating these features. The simplest alternative method for removing land from the image stack is to use a detailed polygon feature layer (.gdb or .shp) of the lake boundary to avoid mixed pixels, a pixel with a mixture of land and water, such as those found along the immediate shoreline. The HAB Index Estimation tool (Chapter 5) provides a method for extracting the water area using a feature layer or by on-screen digitizing a polygon prior to applying water quality algorithms. An alternate land mask can be derived by calculating a Normalized Difference Water Index (NDWI) on the image stack using the ArcGIS Pro *Raster Calculator* function and applying the following equation $\frac{\text{Band3}-\text{Band8A}}{\text{Band3}+\text{Band8A}}$ or $\frac{560\text{nm} - 865\text{nm}}{560\text{nm} + 865\text{nm}}$. The land mask can then be applied to the image stack using a NDWI index threshold less than or equal to zero.

Note that these methods are an attempt to remove the maximum amount of non-water features from the image stack, yet artifacts are possible in certain cases. Therefore, it is important to review all output files prior to using the masked image stack in the HAB Index Estimation tool.

4.3 Outputs

The primary output from the tool is the image stack (Figure 10). It is saved in the workspace directory under the *Stacked* folder with the prefix of the tile name and date and the suffix "_STACK.tif" (*i.e., C:\SampleData\Stacked\T14SPH_20210708_20m_STACK.tif*). Two optional output files may be generated if the checkboxes are selected. If the cloud mask option is selected, the output file will be saved into the workspace directory under the *Stacked* folder with the suffix "_ STACK_CLDMSK.tif " (*i.e.,* *C:\SampleData\Stacked\ T14SPH_20210708_20m_ STACK_CLDMSK.tif*). If the non-water feature mask option is selected, the output file will be saved into the workspace directory under the *Stacked* folder with the suffix "_STACK_WATER.tif" (*i.e., C:\UserWorkspace\Stacked\ T14SPH_20210708_20m_STACK_WATER.tif*).

Figure 10. Sentinel-2 imagery of Marion Reservoir, Kansas for August 17, 2021. A.) Sentinel-2 stacked image -bands 1-8A with no mask applied, B.) Cloud Mask – yellow areas represent the removal of clouds from the imagery, C.) Non-Water Features Mask – yellow areas represent all features (including clouds) removed from the imagery with only water areas remaining.



4.4 Troubleshooting guidance

If an error message is received for the IMG_DATA folder input, ensure the entire path name to the IMG_DATA folder, starting at the root directory (*i.e., E:*), does not exceed 200 characters. For example, this path name contains 181 characters and is acceptable input: *E:\S2B_MSIL2A_20210515T163839_N0300_R126_T15RYP_20210515T204604\S2B_MSIL2A_20210515T163839_N0300_R126_T15RYP_20210515T204604.SAFE\GRANULE\L2A_T15RYP_A021888_20210515T164437\IMG_DATA.* Whereas the following pathname contains 217 characters and is not acceptable input: *E:\HABWORK/MYPROJECTNAME/MYPROJECTAREA/S2B_MSIL2A_20210515T163839_N0300_R126_T15RYP_20210515T204604\S2B_MSIL2A_20210515T163839_N0300_R126_T15RYP_20210515T204604.SAFE\GRANULE\L2A_T15RYP_A021888_20210515T164437\IMG_DATA.* To shorten the path name, extract the “Tilename” folder and all subfolders as shown in Figure 8. In the example path above, the “Tilename” is *L2A_T15RYP_A021888_20210515T164437*.

If "QI_DATA\MSK_CLDPRB_20m.jp2 does not exist or is not supported," confirm all subfolders under the "Tilename" folder have been extracted from the zip file. Please refer to Figure 8 for the required subfolders. The IMG_DATA and QI_DATA folder are required extractions.

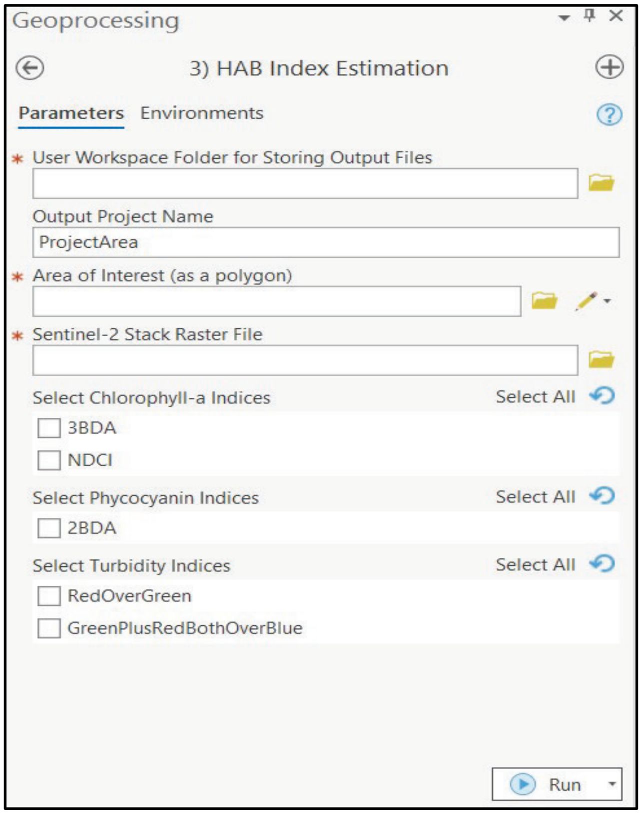
5 Harmful Algal Bloom (HAB) Index Estimation Tool

This tool was developed to generate relative estimations of water quality via spectral indices for the detection, quantification, and monitoring of HABs. The indices, or water quality algorithms, are applied to Sentinel-2 imagery prepared through the Image Analysis Preparation tool (i.e., image stack with non-water features masked). The algorithms available in the tool were selected based on their performance for detecting HABs using three water quality parameters: chlorophyll-a, phycocyanin, and turbidity and are further described in Beck et al. (2016, 2017, and 2019). It is important to note that the algorithms are designed to estimate surface and near-surface phytoplankton biomass but do not differentiate between algal species. Five algorithm options are provided with the tool (Table 3; Figure 11). The outputs generated from the application of the algorithms are represented as relative water quality index values (unitless), or estimations, and not the actual concentrations of chlorophyll-a, phycocyanin, or turbidity.

Table 3. Select water quality algorithms for estimating chlorophyll-a, phycocyanin, and turbidity (described in Beck et al. 2016, 2017, and 2019).

Water Quality Algorithm	Water Quality Parameter	Sentinel-2 Band Calculation Wavelengths in Nanometers (nm)	Reference
3BDA	Chlorophyll-a	$\frac{\left(\frac{1}{665\text{nm}}\right)}{\left(\frac{1}{705\text{nm}}\right)} * 740\text{nm}$	Gitelson et al 2003
NDCI	Chlorophyll-a	$\frac{(705\text{nm} - 665\text{nm})}{(705\text{nm} + 665\text{nm})}$	Mishra and Mishra 2012
2BDA	Phycocyanin	$\frac{705\text{ nm}}{665\text{ nm}}$	Wynne et al 2008
RedOverGreen	Turbidity	$\frac{665\text{ nm}}{560\text{ nm}}$	Bowers and Binding 2006
GreenPlusRedBothOverBlue	Turbidity	$\frac{560\text{ nm} + 665\text{ nm}}{490\text{ nm}}$	Frohn and Autrey 2009

Figure 11. HAB Index Estimation tool dialog box in ArcGIS Pro desktop software.



5.1 Inputs

- User Workspace: Select User Workspace Folder for Storing Output Files. Click the folder button and browse to the folder location for storage of output files. The tool will create three subfolders based on the water quality parameter selected (Chlorophyll, Phycocyanin, and Turbidity) under the user specified workspace folder (*i.e., C:\Project\Phycocyanin*).
- Project Name: Enter the name of the project. This name will be added as a prefix to the file geodatabase where temporary data, such as the AOI polygon, will be stored in the user workspace. (*ex. project name = NewArea; resulting filename = C:\Project\NewAreaHABprj.gdb*)
- AOI: The AOI represents the spatial extent of the water body area provided as a polygon feature (i.e., shapefile or geodatabase) or manually digitized on-screen by clicking the pencil icon tool and selecting polygon option. The tool uses this polygon to clip the input image stack. If a predefined feature layer is provided, it must have a defined spatial reference. All polygon features in the dataset will be used unless selected in the ArcGIS Pro map view prior to running the tool. A suggested source for lake boundaries is the National

Hydrography Dataset waterbody polygon layer which can be obtained through the USGS ScienceBase catalog (https://www.sciencebase.gov/catalog).

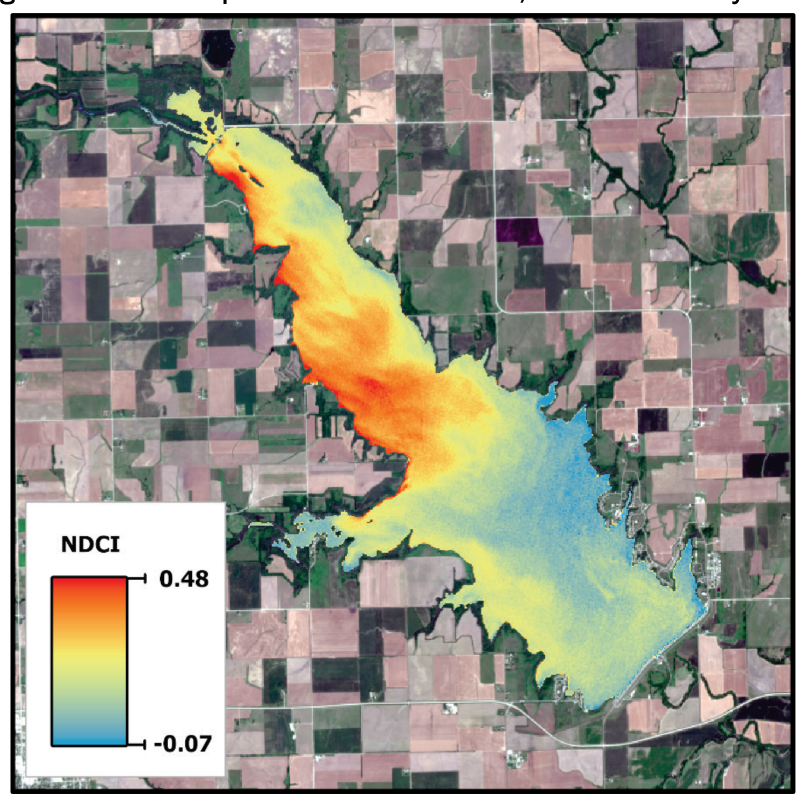
- Raster Stack: Select the Sentinel-2 Raster Stack File. This raster is an image stack created in the Image Analysis Preparation tool as detailed in Chapter 4. Three potential raster choices are available: the STACK, STACK_CLDMSK, or the STACK_WATER raster. (*i.e., C:\SampleData\Stacked\ T14SPH_20210708_20m_STACK.tif*). If the input raster is not visible in the folder browser window, refresh the dialog box (see Section 5.3).
- Indices: Five water quality indices are available for selection through the tool. The indices are grouped by three water quality parameters (chlorophyll-a, phycocyanin, and turbidity) with each parameter containing at least one index option. Of the five checkboxes, one or multiple checkboxes may be chosen by adding a check mark next to the desired index. Please refer to Table 3 for details regarding each index including the algorithm used to calculate the index and corresponding scientific literature.

5.2 Outputs

Each index selected is saved in a .tif file format in the workspace directory under the appropriate water quality folder name (Chlorophyll, Phycocyanin, and Turbidity) with the filename comprised of the tile name and image collection date as prefix and the index name as the suffix. (*i.e., C:\SampleData\Chlorophyll\ T14SPH_20210708_NDCI.tif*). The output is provided as a relative index where higher index values represent greater parameter (chlorophyll-a, phycocyanin, turbidity) estimations while lower index values represent lower parameter estimations.

Figure 12 illustrates an example output of relative index estimations of chlorophyll-a developed by applying the Normalized Difference Chlorophyll Index (NDCI; Mishra and Mishra 2012) to Sentinel-2 L2A imagery acquired on 8 July 2021 for the Marion Reservoir in Kansas. In general, NDCI values can range from -1 to 1. At the time the NDCI output in Figure 12 was generated, it was processed using Sentinel-2 L2A imagery with Processing Baseline 3.01. Generally, values less than 0 represented areas without algae (blue), while values from 0 to ~0.2 indicated relatively low chlorophyll-a concentration (blue to yellow), and values greater than 0.2 showed relative increases in chlorophyll-a concentration (yellow to red), in which higher values may be indicative of a HAB and can be used to help prioritize field sampling (Figure 12). The values are displayed as a continuous stretched color ramp which is helpful for visualizing subtle patterns and shifts in chlorophyll-a concentrations. Figure 12 shows the NDCI output displayed using the *Prediction* color scheme (blue-yellow-red) in ArcGIS Pro. Note that while the NDCI does not differentiate algal species or directly assess toxicity, it is commonly used to estimate surface and near-surface phytoplankton biomass, a water quality indicator of HABs (Mishra and Mishra 2012). Furthermore, it is important to note that future iterations of the Sentinel L2A Processing Baseline may affect trends in index values. Thus, the general interpretation expressed for the output in Figure 12 may change based on how the imagery is processed, including atmospheric correction, radiometric correction, scaling factors, etc. Since each algorithm has different index value ranges, it is suggested that users explore the visualization of water quality conditions through ArcGIS Pro symbolization such as Classify-Natural Breaks. Additionally, the most thorough way to interpret index values is through examining relationships between index values and associated in situ data values (from field-based sampling) using linear regression analysis (discussed in Chapter 6). More information about the algorithms and related output can be found in Beck et al. (2016, 2017, and 2019) and Johansen et al. (2019).

Figure 12. NDCI output for Marion Reservoir, Kansas for 8 July 2021.



5.3 Troubleshooting guidance

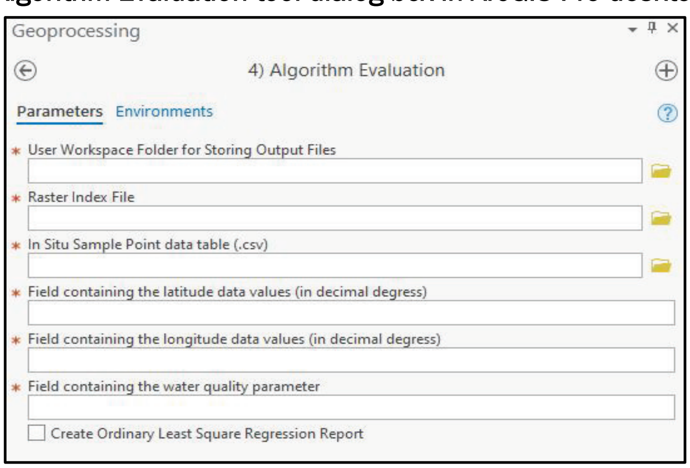
The AOI polygon must fall (at least partly) within the extent of the image, or the user will receive an error prior to running the tool.

If the image stack does not appear as a selectable choice when adding it through the folder browser window or the *Add Data* tool in ArcGIS Pro, click the refresh button (clockwise open circle arrow located to the right of the directory path) to refresh the dialog box window.

6 Algorithm Evaluation Tool

The Algorithm Evaluation tool evaluates the relationship between relative water quality index values and associated in situ data values (from field-based sampling) using linear regression analysis. The two main inputs into the tool include (Figure 13): (1) one water quality index raster, generated as output in the HAB Index Estimation tool, and (2) a data table in comma delimited format (CSV format) of in situ sample data points including the latitude, longitude, and water quality values for each sample point for a corresponding parameter (chlorophyll-a, phycocyanin, or turbidity). First, the tool converts the data table into a point layer using the spatial coordinates, then filters the layer by the spatial extent of the relative water quality index raster, and finally saves it to a point layer (*analysisfieldpts*) within a file geodatabase. The filtered point layer is then used to extract values from the relative water quality index raster and append those values to the attribute table in the point layer. Next, a linear regression analysis and/or an optional ordinary least squares (OLS) regression analysis is applied using the in-situ sample and relative water quality index data fields within the point layer to generate a linear regression graph and statistics (i.e., R^2 and p-value). The tool output can then be used to gain insight into patterns or relationships in the data. Additionally, a HAB-concentration map product is created by converting relative index values to absolute values based on the linear regression analysis if the R-squared value meets or exceeds the 0.70 threshold. Due to the dynamic nature of water conditions, in situ sample data is best used in statistical comparison with relative index values when the field data are collected in close temporal proximity to the source image date and time.

Figure 13. Algorithm Evaluation tool dialog box in ArcGIS Pro desktop software.



6.1 Inputs

- User Workspace: Select User Workspace Folder for Storing Output Files. Click the folder button and browse to the folder location for storage of output files. The tool will automatically create an output folder named *Eval* under the user specified workspace folder. (*i.e., C:\SampleData \Eval*).
- Index File: Select the Raster Index File created in HAB Index Estimation tool (*i.e., C:\SampleData\Chlorophyll\T14SPH_20210708_NDCI.tif*). Refer to Section 6.3 if there are issues selecting the layer.
- In Situ Table: Select the In Situ Sample Point Data Table (in CSV format). This data table represents in situ sample values for at least one parameter: chlorophyll-a, phycocyanin, and turbidity. This table is used to create a point feature layer called *allfieldpts*. Then, the points from this layer that overlap with the image extent of the relative water quality index raster are saved into the feature layer *analysisfieldpts*, which is used to extract values from the relative water quality index raster to run linear regression statistics between the two corresponding values. Prior to running the tool, the data table must first be formatted into a CSV tabular file format (Figure 14). The first row of the table must contain column heading names and at least three fields representing the latitude (in decimal degrees), longitude (in decimal degrees), and in situ sample values, all represented as floating point numbers. Multiple fields may exist per water quality parameter. In addition, at least three sample points with water quality data (three rows) are required; however, depending on the variability of the water quality conditions and size of the waterbody, at least seven well distributed points are preferred for more accurate results.

Figure 14. Example in situ data table formatting.

	A	B	C	D
1	SiteID	lat	lon	Chl
2	0	38.437	-97.158	0.244
3	1	38.447	-97.159	0.183
4	2	38.444	-97.154	0.231
5	3	38.442	-97.154	0.194
6	4	38.431	-97.146	0.225
7	5	38.429	-97.153	0.295
8	6	38.418	-97.146	0.271
9	7	38.419	-97.138	0.276
10	8	38.424	-97.137	0.238
11	9	38.425	-97.145	0.248
12	10	38.435	-97.150	
13	11	38.365		
	12	[illegible]		

- Latitude Field: Field containing the latitude data values (in decimal degrees). From the pull-down list, select the field containing the latitude data values in decimal degrees (i.e., lat). The value list is derived from the column heading names in the field sample point CSV file.

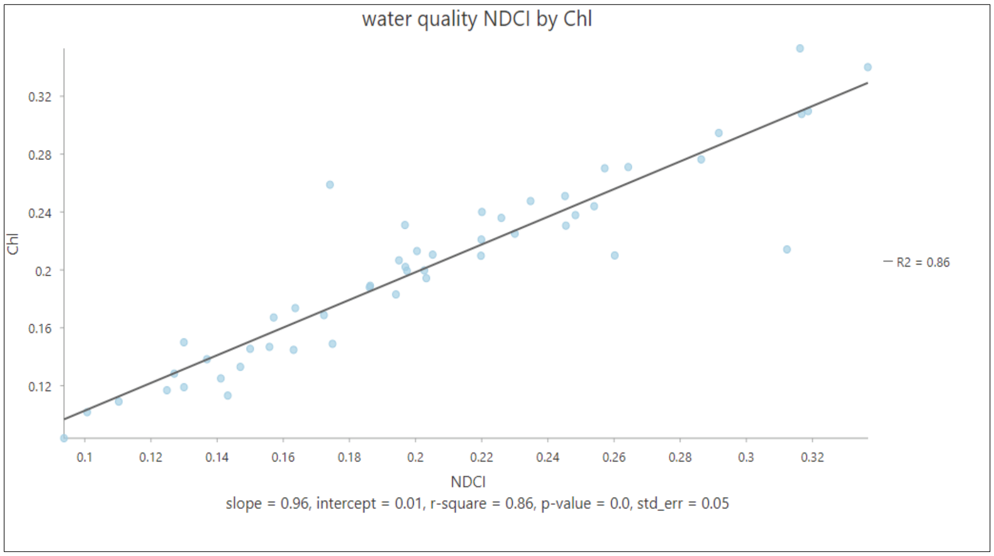
- Longitude Field: Field containing the longitude data values (in decimal degrees). From the pull-down list, select the field containing the longitude data values in decimal degrees (i.e., lon). The value list is derived from the column heading names in the field sample point CSV file.
- Water Quality Field: Field containing the water quality parameter data. From the pull-down list, select the field containing the in situ sample values (i.e., Chl). The value list is derived from the column heading names in the field sample point CSV file.
- OLS Regression: Create OLS report checkbox. If a checkmark is placed in the box, a .pdf file is created containing a summary of the OLS results.

6.2 Outputs

Three standard output files, one conditional file, and two optional output files are available through the tool and are stored in the *Eval* folder under the user workspace (Figure 1).

- Standard Output: Of the three standard outputs, the first two are point feature layers (*allfieldpts* and *analysisfieldpts*) included in a file geodatabase with the filename comprised of the tile name, image collection date, and index name as prefix and "*_HAB*" as the suffix (*i.e., C:\SampleData\Eval\T14SPH_20210708_NDCI_HAB.gdb*). The *allfieldpts* layer represents all possible sample points listed in the field sample point data table (CSV), while the *analysisfieldpts* layer represents only those points from the *allfieldpts* layer that overlap with the relative water quality index raster layer. The attribute fields within each layer contain the following fields: a unique point id (ptid), latitude, longitude, and in situ water quality parameter value. The ptid field and data values are automatically generated by the tool to ensure that each point has a unique integer identifier. The *analysisfieldpts* has an additional field for index values extracted from the relative water quality index raster. Therefore, this table can also be exported for use in other statistical packages. The third standard output is a linear regression graph displaying the relative water quality index and in situ sample value relationship including the slope, intercept, R^2, p-value, and standard error statistics in a scalable vector graphics (.svg) file format easily viewed in an internet browser (Figure 15). The filename of the .svg includes the tile name and image collection date as prefix and the index as the suffix (i.e., C:\ SampleData \Eval\T14SPH_20210708_NDCI.svg).

Figure 15. Linear regression graph comparing NDCI relative index values to hypothetical (synthetic) chlorophyll-a sample values (in situ); sample .svg graph output.



- Conditional Output: The conditional output file is a HAB concentration map product (raster file) which uses the in situ sample and relative water quality index value relationship to convert relative HAB estimations into absolute HAB concentrations. The creation of this product is dependent on the R^2 value being greater than or equal to 0.70. The product generated is stored under the *Eval* folder with the filename comprised of the tile name, image collection date, and index as prefix and the "_ABS" as the suffix (*i.e., C:\ SampleData \Eval\ T14SPH_20210708_NDCI _ABS.tif*).
- Optional Output: The two optional output files are derived from the OLS regression analysis. The first file is a report summarizing the OLS results including model variable statistics, diagnostics, variable distributions and relationships, standardized residuals histogram, and residual versus predicted graphs in a .pdf format (https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/ordinary-least-squares.htm). The filename of the .pdf will include the tile name, image collection date, and index as prefix and "_OLSR" as the suffix (i.e., C:\Project\Eval\ T14SPH_20210708_NDCIOLSR.pdf). The second file is a point feature layer called "OLSRPT" stored in the same file geodatabase as the standard output (*i.e., C:\ SampleData\Eval\T14SPH_20210708_NDCI_HAB.gdb\OLSRPT).* The layer includes

the following attribute fields: unique point id (ptid), in situ sample value, relative index value, estimated, residual, and standard residual values.

If statistical comparisons do not show good agreement, it is likely due to changes that occurred between the date/time of imagery and in situ data collections. Thus, it does not necessarily mean that the results of the relative water quality index are poor or not reflective of conditions at the time of image collection. While there is no 'golden rule' for the optimal time difference between imagery and in situ data collections, the more temporal coincidence between the collections, the more appropriate the data are for statistical comparison. Several factors contribute to the dynamic nature of waterbodies, and thus, these factors may affect the temporal proximity in which in situ and imagery data can be appropriately compared. Thus, experimentation with available in situ data and comparison to relative water quality index values will yield a greater understanding to identify the appropriate timeframe for data comparison at the local or regional level.

6.3 Troubleshooting guidance

If the algorithm raster image does not appear as a selectable choice when adding it either through the folder browser window or the *Add Data* tool in ArcGIS Pro desktop software, click the refresh button (clockwise open circle arrow located to the right of the directory path) to refresh the dialog box window.

Ensure all fields in the CSV table are correctly formatted according to the guidance in Section 6.1 and Figure 14. In addition, confirm all fields selected as input into the tool from the CSV table include accurate values and are not null or left blank.

If the results do not appear as expected, examine the in situ point locations within geodatabase layer *analysisfieldspts* against the image stack to ensure that each point does not coincide with a non-water pixel (i.e., land, clouds, etc.). The OLSR point layer and .pdf file can also assist with identifying potential outlier values which may be skewing the results.

The user will receive an error message if less than three sample points are present in the in situ data table. At least three sample points are required to generate a linear regression; however, more samples are needed to determine a statistical relationship.

7 Conclusion

The primary purpose of the *waterquality for ArcGIS Pro* toolbox and this software user guide document is to enable water operations managers across USACE, as well as others, to detect, monitor, and quantify water quality indicators of HABs in small, inland waterbodies using Sentinel-2 satellite imagery. It is part of a series of software tools developed under research and development supported by ERDC's Aquatic Nuisance Species Research Program (ANSRP) with the broader goal of making remote sensing technology more accessible to a variety of managers, researchers, and stakeholders responsible for and interested in improving water quality monitoring. This broader goal is also shared by USACE partners, including the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), the U.S. Geological Survey (USGS), and the U.S. Environmental Protection Agency (EPA), who have developed complementary technologies for HAB monitoring. More specifically, the Cyanobacteria Assessment Network (CyAN) is a multi-agency project which primarily utilizes the ESA's Sentinel-3 Ocean Land Colour Instrument (OLCI) with a spatial resolution of 300 m for regional assessment (Schaeffer et al. 2015). Thus, the software tools represent a range of options that are not only geographically complementary, but also maximize specific advantages associated with each sensor in which different spectral bands and spatial resolutions offer a balanced variety of data products to address a range of monitoring needs.

The ESA's growing suite of sensors reflects a remote sensing technology trend with an overall increasing number of satellite sensors offering improvements in spatial, spectral, and temporal resolutions. Such improvements are leading to more options for improved HAB detection and monitoring across large geographic areas at project-specific scales using readily available satellite imagery and software tools. The *waterquality for ArcGIS Pro* toolbox likewise reflects this trend and is flexible for incorporating new sensors, algorithms, output, and analytical approaches. As remote sensing technology continues to rapidly evolve, so too will the software tools used to exploit the rich content provided by the increasing amounts of high spatial and spectral resolution imagery. The variety of software tools should attempt to reach a broad user base with a wide range of skillsets and monitoring needs. To that end, web-based applications, such as the *Harmful Algal Bloom [HAB] Explorer* (https://arcportal-ucop-corps.usace.army.mil/hab) developed in partnership between

ERDC and ESRI reflect the need for non-desktop options and serves as a cursory screening tool to quickly identify potential areas of concern that may require additional monitoring. Future software tool development will require an iteratively adaptive approach to keep pace with the rapid advances in satellite sensor technology which will only continue to improve, resulting in many exciting new opportunities to enhance HAB monitoring, and ultimately empowering managers to take a more proactive stance in HAB management.

References

Beck, R., S. Zhan, H. Liu, S. Tong, B. Yang, M. Xu, A. Ye, Y. Huang, S. Shu, Q. Wu, S. Wang, K. Berling, A. Murray, E. Emery, M. Reif, J. Harwood, J. Young, C. Nietch, D. Macke, M. Martin, G. Stillings, R. Stump, and H. Su. 2016. "Comparison of satellite reflectance algorithms for estimating chlorophyll-a in a temperate reservoir using coincident hyperspectral aircraft imagery and dense coincident surface observations." *Remote Sensing of Environment*, 178: 15-30, https://doi.org/10.1016/j.rse.2016.03.002

Beck, R., M. Xu, S. Zhan, H. Liu, R. A. Johansen, S. Tong, B. Yang, S. Shu, Q. Wu, S. Wang, K. Berling, A. Murray, E. Emery, M. Reif, J. Harwood, J. Young, M. Martin, G. Stillings, R. Stumpf, H. Su, Z. Ye, and Y. Huang. 2017. "Comparison of satellite reflectance algorithms for estimating phycocyanin values and cyanobacterial total biovolume in a temperate reservoir using coincident hyperspectral aircraft imagery and dense coincident surface observations." *Remote Sensing*, 9(6):538. https://doi.org/10.3390/rs9060538

Beck, R., M. Xu, S. Zhan, R. A. Johansen, H. Liu, S. Tong, B. Yang, S. Shu, Q. Wu, S. Wang, K. Berling, A. Murray, E. Emery, M. Reif, J. Harwood, J. Young, C. Nietch, D. Macke, M. Martin, G. Stillings, R. Stump, H. Su, Z. Ye, and Y. Huang. 2019. "Comparison of satellite reflectance algorithms for estimating turbidity and cyanobacterial concentrations in productive freshwaters using hyperspectral aircraft imagery and dense coincident surface observations." *Journal of Great Lakes Research*, 45(3): 413-433, https://doi.org/10.1016/j.jglr.2018.09.001

Bergsma, E., and R. Almar. 2020. "Coastal coverage of ESA' Sentinel 2 mission." *Adv. Space Res*. 65, 2636-2644. https://doi.org/10.1016/j.asr.2020.03.001

Bramich, J., C. J. S. Bolch, and A. Fischer. 2021. "Improved red-edge chlorophyll-a detection for Sentinel 2." *Ecol. Indic.* , 120: 106876–106885. https://doi.org/10.1016/j.ecolind.2020.106876

Bowers, D. G., and C. E. Binding. 2006. "The optical properties of mineral suspended particles: A review and synthesis." *Estuarine Coastal and Shelf Science* 67 (1–2): 219–230. https://doi.org/10.1016/j.ecss.2005.11.010

Environmental Systems Research Institute (ESRI). 2021. *ArcGIS Desktop*. Redlands, CA.

Frohn, R. C., and B. C. Autrey. 2009. *Water quality assessment in the Ohio River using new indices for turbidity and chlorophyll-a with Landsat-7 Imagery*. Draft Internal Report, U.S. Environmental Protection Agency. https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NERL&dirEntryId=163848

Gitelson, A. A., U. Gritz, and M. N. Merzlyak. 2003. "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves." *J. Plant Phys*., 160, 271-282. https://doi.org/10.1078/0176-1617-00887

Johansen, R. 2018. *waterquality_workflow: A case study and workflow for detecting and quantifying cyanobacterial harmful algal blooms (CHABs) from Sentinel-2 imagery*. Version (v.0.2.) Zenodo. http://doi.org/10.5281/zenodo.2003619

Johansen, R., M. Reif, E. Emery, J. Nowosad, R. Beck, M. Xu, and H. Liu. 2019. *waterquality: An open-source R package for the detection and quantification of cyanobacterial harmful algal blooms and water quality*. ERDC/EL TR-19-20; DOI: 10.21079/11681/35053 http://dx.doi.org/10.21079/11681/35053

Johansen, R., J. Nowosad, M. Reif, and E. Emery. 2020. *waterquality: Satellite Derived Water Quality Detection Algorithms*. R package version 0.2.6. https://CRAN.R-Project.org/package=waterquality

Mishra, S., and D. Mishra. 2012. "Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbin productive waters." *Remote Sensing of Environment* 117: 394-406. https://dx.doi.org/10.1088/1748-9326/9/11/114003

Mishra, S., R. Stumpf, B. Schaeffer, P. Werdell, K. Loftin, and A. Meredith. 2019. "Measurement of cyanobacterial bloom magnitude using satellite remote sensing." *Scientific Reports*. Nature Publishing Group, London, Uk, 9:18310, (2019). https://doi.org/10.1038/s41598-019-54453-y

Nagpal, A., and G. Gabrani. 2019. "Python for data analytics, scientific and technical applications," 2019 Amity International Conference on Artificial Intelligence (AICAI). pp. 140-145, https://doi.org/10.1109/AICAI.2019.8701341

Reif, M. 2011. *Remote sensing for inland water quality monitoring: A US Army Corps of Engineers perspective*. ERDC/EL TR-11–13. Vicksburg, MS: US Army Engineer Research and Development Center.

Schaeffer, B., K. Loftin, R. Stumpf, and J. Werdell. 2015. "Agencies collaborate, develop a cyanobacteria assessment network. EOS, Transactions, American Geophysical Union." *American Geophysical Union*, Washington, DC, 96:1-9.

Stumpf, R. P., and M. C. Tomlinson. 2007. Remote sensing of harmful algal blooms. In: Miller R.L., Del Castillo C.E., Mckee B.A. (eds) *Remote Sensing of Coastal Aquatic Environments*. *Remote Sensing and Digital Image Processing*, vol 7. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-3100-7_12

Toming, K., T. Kutser, A. Laas, M. Sepp, B. Paavel, and T. Nõges. 2016. "First experiences in mapping lake water quality parameters with Sentinel-2 MSI Imagery." *Remote Sensing*. 2016; 8(8):640. https://doi.org/10.3390/rs8080640

Valgur, M., K. Jonas, L. Delucchi, G. Baier, Malte, unnic, S. Staniewicz, L. Kioi kinyanjui, v. Bahr, P. Salembier, M. Ber, G. Keller, D. Salmeida, C. Castro, and A. Raspopov. 2019. *sentinelsat/ sentinelsat*: v0.13. Zenodo. https://doi.org/10.5281/zenodo.2629555

Wang, L., X. Min, L. Yang, L. Hongxing, R. Beck, M. Reif, E. Emery, J. Young, and Q. Wu. 2020. "Mapping freshwater chlorophyll-a concentrations at a regional scale integrating multi-sensor satellite observations with Google Earth Engine" Remote Sensing 12(20): 3278. https://doi.org/10.3390/rs12203278

Wynne, T. T., R. P. Stumpf, M. C. Tomlinson, R. A. Warner, P. A. Tester, and J. Dyble. 2008. "Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes." *Int. J. Remote Sens.*, 29: 3665–3672. https://doi.org/10.1080/01431160802007640

Xu, M., H. Lui, R. Beck, M. Reif, E. Emery, and J. Young. 2019. *Regional analysis of lake and reservoir water quality with multispectral satellite remote sensing images*. ERDC/EL TR 19-19. Vicksburg, MS: US Army Engineer Research and Development Center. http://dx.doi.org/10.21079/11681/34933

Unit Conversion Factors

Multiply	By	To Obtain
Feet (ft)	0.3048	Meters (m)
Meters (m)	3.2808	Feet (ft)
Meters (m)	1e+9	Nanometers (nm)
Square-kilometers (km^2)	247.105	Acres (ac)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
September 2022

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE

waterquality for ArcGIS Pro Toolbox: User's Guide

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Christina Saltus, Molly Reif, and Richard Johansen

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Environmental Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road, Vicksburg, MS 39180-6199, USA

Environmental Laboratory
U.S. Army Engineer Research and Development Center
Joint Airborne Lidar Bathymetry Technical Center of Expertise (JALBTCX)
Kiln, MS, 39556, USA

8. PERFORMING ORGANIZATION REPORT NUMBER

ERDC/EL SR-22-6

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Aquatic Nuisance Species Research Program
US Army Engineer Research and Development Center
Environmental Laboratory
Vicksburg, MS 39180-6199

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

AMSCO Code 008284; Funding Account U4377102

14. ABSTRACT

Monitoring water quality of small inland lakes and reservoirs is a critical component of the US Army Corps of Engineers (USACE) water quality management plans. However, limited resources for traditional field-based monitoring of numerous lakes and reservoirs covering vast geographic areas often leads to reactional responses to harmful algal bloom (HAB) outbreaks. Satellite remote sensing methodologies using HAB indicators is a good low-cost option to traditional methods and has been proven to maximize and complement current field-based approaches while providing a synoptic view of water quality (Beck et al. 2016; Beck et al. 2017; Beck et al. 2019; Johansen et al. 2019; Mishra et al. 2019; Stumpf and Tomlinson 2007; Wang et al. 2020; Xu et al. 2019; Reif 2011). To assist USACE water quality management, we developed an Environmental Systems Research Institute (ESRI) ArcGIS Pro desktop software toolbox (*waterquality for ArcGIS Pro*) founded on the design and research established in the *waterquality* R software package (Johansen et al. 2019; Johansen 2020). The toolbox enables the detection, monitoring, and quantification of HAB indicators (chlorophyll-a, phycocyanin, and turbidity) using Sentinel-2 satellite imagery. Four tools are available: (1) automating the download of Sentinel-2 Level-2A imagery, (2) creating stacked image with options for cloud and non-water features masks, (3) applying water quality algorithms to generate relative estimations of one to three water quality parameters (chlorophyll-a, phycocyanin, and turbidity), and (4) creating linear regression graphs and statistics comparing in situ data (from field-based water sampling) to relative estimation data. This document serves as a user's guide for the *waterquality for ArcGIS Pro* toolbox and includes instructions on toolbox installation and descriptions of each tool's inputs, outputs, and troubleshooting guidance.

15. SUBJECT TERMS

Algal blooms – Monitoring
Remote-sensing images
Reservoirs – Water quality
Water quality management
Geospatial data
Geographic information systems

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER *(include area code)***
Unclassified	Unclassified	Unclassified	SAR	43	