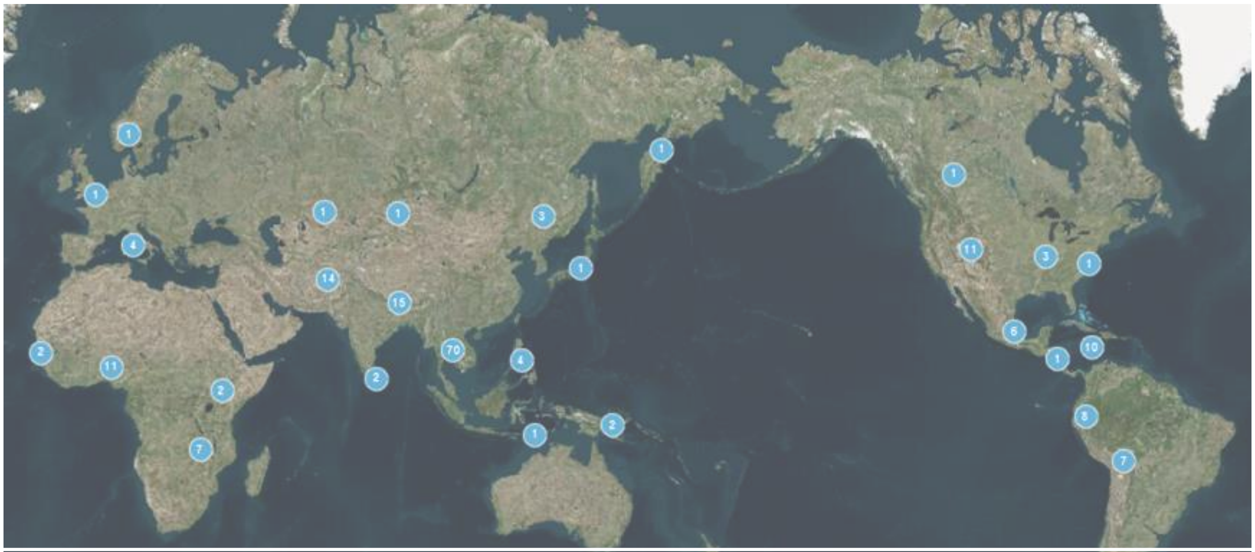


Institution & Project Creation Manual

The CEO Team Spring 2022

Table of Contents

Welcome to Collect Earth Online

Collect Earth Online, or CEO, is a free and open-source image viewing and interpretation tool, suitable for projects requiring information about land cover and/or land use. CEO enables simultaneous visual interpretations of satellite imagery, providing global coverage from MapBox and Bing Maps, a variety of satellite data sources from Google Earth Engine, and the ability to connect to your own Web Map Service (WMS) or Web Map Tile Service (WMTS). The full functionality is implemented online, no desktop installation is necessary.

CEO allows institutions to create projects and leverage their teams to collect spatial data using remote sensing imagery. Use cases include historical and near-real-time interpretation of satellite imagery and data collection for land cover/land use model validation.

Collect Earth Online (CEO) is available at https://collect.earth/.

This manual provides information for institution and project administrators, including how to set up an institution, how to set up a project, and project management suggestions. Part 1 provides a quick overview of the Collect Earth Online platform, so administrators can set up their own user accounts. Part 2 details how to set up and manage an institution in CEO. Part 3 discusses how to manage custom imagery feeds; if your institution will only be using CEO's standard imagery, you can skip this section. Part 4 discusses how to create a data collection project, Part 5 discusses project review, and Part 6 introduces the Geo-Dash, which provides additional information for uses during data collection. Part 7 explains how to manage a published project. If you are interested in setting up a CEO project for validation (e.g. of a land cover model), Part 8 discusses project creation for these specialized projects.

This manual uses a few formatting standards for ease of use. Clickable links are shown **[like this]**. The names of pages, like the **Home** page, are bolded.

The icon is used to indicate use cases. The icon is used to indicate new, interesting, and or useful functionality in CEO. Finally, CEO is a live tool that is under continuous development. Workarounds for current issues are shown in boxes like this:

This is a description of a current challenge and how to work around it.

And planned features are shown in boxes like this:

This is a description of a planned feature.

Part 1: Quick Introduction to Collect Earth Online

A. Setting up your account

1. In your browser window, navigate to https://collect.earth/. CEO supports Google Chrome, Mozilla Firefox, and Microsoft Edge.
2. Click [Login/Register] on the upper right.
3. To set up a new account, click on [Register] and follow the instructions.
4. You will need to validate your email upon registration.
5. When you have an account, login with your email and password.
6. If you forget your password, click on [Forgot your password?] and follow the instructions.

B. Website features

You can access the Home, About, Support, and Account pages from the top menu bar.

1. The Home page includes information about institutions, published projects, and a map showing locations of existing projects.
2. The About page summarizes information about CEO.
3. The Support features any currently available Collect Earth Online Manuals, Tutorials, and a Collect Earth Online Demo. This page also includes links for bug reporting and forums to ask for help.
4. The Blog page hosts interesting news about CEO and information about new features.
5. The Account page lists information such as user statistics and allows users to update their account settings.
6. Under the CEO logo on this top menu bar is the current version of CEO.
7. There is a purple ? in the top right corner of the screen. Clicking on this will bring up the help interface, which provides information about CEO’s features. These help interfaces are available for the Home page, for Data Collection (see other manual) and for Project Creation.

CEO will be adding dynamic “My Institutions/My Projects” tabs along with more detailed Account pages. This will make it easier to access your institutions and projects.

C. Report an issue and request new features

You can email support@collect.earth to report an issue or request a new feature.

Additionally, when you click on **[Support]** at the top of the webpage, there is a link to the GitHub issues page. This page is also available at: https://github.com/openforis/collect-earth-online/issues.

If you discover any of the Collect Earth Online functions are not working properly or would like to suggest an additional feature, you can use this page to log an issue or suggestion. Once logged, these messages go directly to the Collect Earth Online developer team.

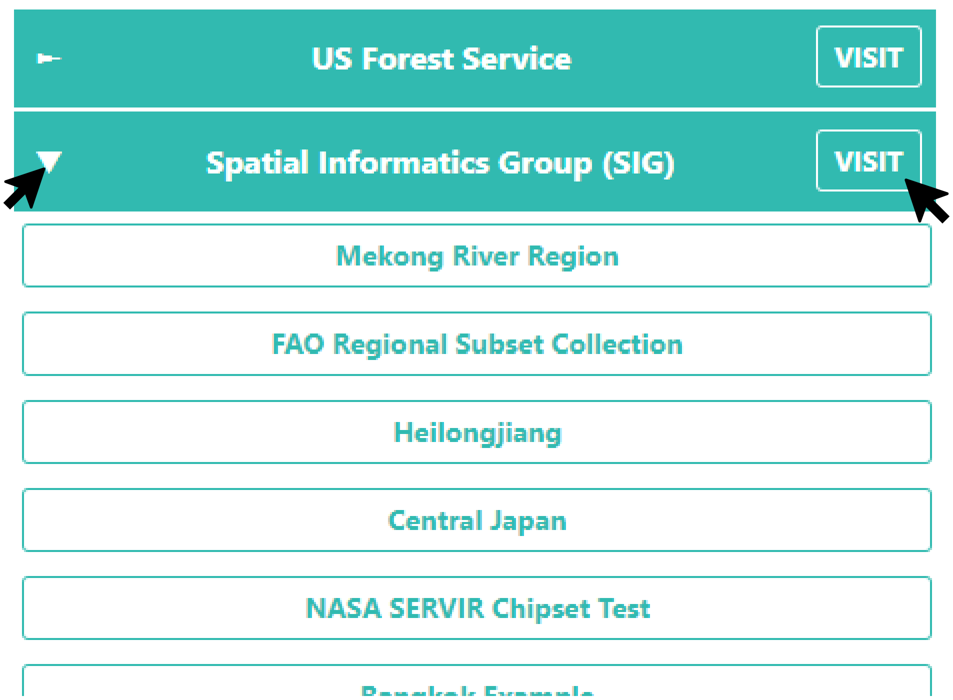
You will have to either log in or set up a GitHub account to log an issue. By logging in, the development team can contact you in case they need additional information to provide a solution to the issue or feature suggestion. Alternatively, if you are unable to register an account with GitHub, you can ask questions on the OpenForis forum, which is available here http://www.openforis.org/support.

1. When you have a GitHub account, navigate to the CEO GitHub Page.
2. To log an issue or request a new CEO feature, simply click on the green [New issue] button in the upper right-hand portion of the screen.
3. Type in a title that conveys the topic of the issue or request.
4. Then below, type in a detailed message summarizing the issue you have encountered or the additional functionality you would like to see in CEO.
5. Once you have finished providing the details, click the green [Submit new issue] button. This will log your request.

Part 2: Institution Set Up and Management

A. List of institutions & Institution pages

1. A list of all institutions is displayed on the Home page.
2. If you are logged in, this list is broken into two: Your Affiliations and Other Institutions. Your Affiliations are institutions to which you belong.
3. Clicking on an 'institutions' name expands a list of all active Projects for the Institution. Clicking on these projects will take you to the Data Collection page (see Data Collection Manual).

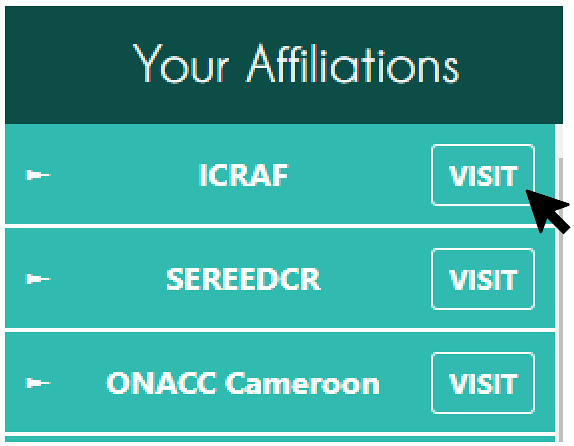


4. Next to each institution's name is a [VISIT] button. Clicking on this will open the Institution page.
5. On any Institution info page, you can find:

   i. The institution's logo, which when clicked will take you to the institution's website link,
   ii. The institution's name
   iii. A short description of the institution,
   iv. Imagery feeds available to the institution
   v. The institution's projects, and
   vi. A list of all registered users associated with the institution.

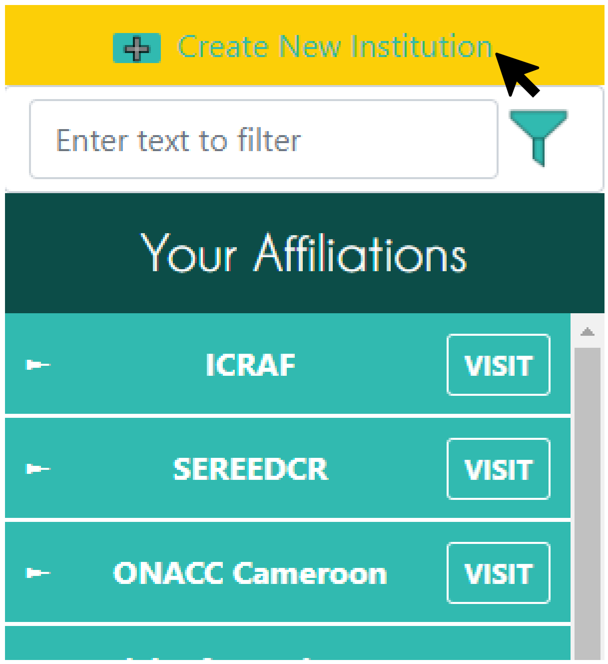
B. Request to join an institution

1. If your institution already has an account on CEO and you would like to request to be a member, start by navigating to the Institution page via the info button described in A.2-4 above.
2. The list of Users is contained under the Users tab. Click on this tab.
3. Click on the [Request Membership] button to be considered to join the institution.
4. Once you are part of an institution, the Institution panel on the Home page will show your institutions. This makes it easier to access these institutions and their projects.

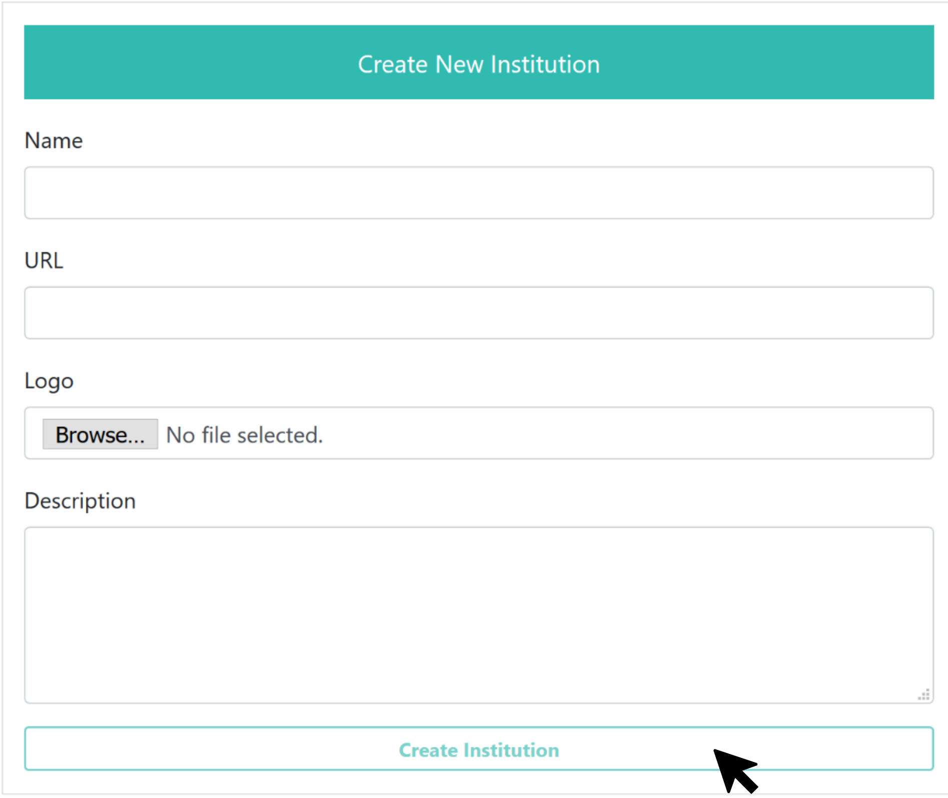


C. Creating a new institution

1. If your institution is not yet signed up for CEO, you can create a new institution. You will need to be signed into your account (click [Login/Register]).
2. When you are signed in, navigate to the Home page and click on [Create New Institution] on the left panel of the Home page.



3. Enter your institution's name, URL, and a brief description of your institution. These will display on your Institution page.
4. You can upload a logo from your computer by clicking on [Browse…] and navigating to the correct location on your computer.



5. Click [Create Institution] after you have finished entering your information.

D. Institution management

If you are an administrator for an institution, you can make changes to your institution's information after the institution is created.

1. If you want to make changes, start by navigating to your Institution's info page (see A.2-4 above).
2. You can make changes to the institution page by clicking [Edit] at the top of the page.
3. You can also delete your institution by clicking [Delete]. This action is PERMANENT, and your institution cannot be recovered afterwards. ALL PROJECTS ASSOCIATED WITH THE INSTITUTION WILL ALSO BE DELETED.

More importantly, you can also manage three aspects of your institution that allow data collection through CEO. These include the **Imagery** feeds, the data collection **Projects**, and the **Users** associated with the institution. This information is displayed in three panels on your **Institution's** info page.

4. The Imagery tab lists all available imagery and WMS Feeds. You can add new imagery feeds here as well. This is discussed in Part 3 of this manual.

5. The Projects panel lists your institution's projects, identifies projects as public or private, and allows you to create new projects. This is discussed in Part 4: Project Creation of this manual.
6. The Users panel lists your 'institutions' members and allows them to be updated. This is discussed in Section E. below.

E. Institution user (member) management

1. As an administrator, you can add a Collect Earth Online member to the institution by typing the user's email address into the box and clicking the [Add User] button. <u>Only email addresses that have already created a CEO account can be added to the institution.</u> If you are successful, a pop-up window will notify you that the user has been given the role 'member,' and the user's email address will appear in the list of users.
2. You can approve pending affiliation requests as well.
3. To change the role of a user, use the menu to the right of their email address. Available user roles are Administrator and Member. Users who are Administrators can edit any of the institution's projects.
4. To restrict the Administrator capabilities of a user to only one of the institution's projects, you can create a new institution specifically for that project. Add the user as an Administrator for the new institution.
5. You can also remove users from the institution using this drop-down menu and selecting the Remove option.

F. Other administrator privileges

1. As an administrator you can review and modify the answers for all plots, including plots labeled by other CEO users. Members of an institution can only review and modify answers for plots they themselves have labeled. To access this feature, use **Review** toggle for the **Navigate Through** dropdown menu on the Collection Page. See Part 6: Administrator review mode for more information.
2. As an administrator you can review Draft Mode projects that you or other administrators have created. This is discussed further in Part 4: Project Creation.
3. As an administrator, you can download collected data from Institution projects. See Part 5: Project Information Interface Sec. E: Export Data for a list of the data fields included in data download.

Part 3: Built-in & Adding Base Imagery Sources

A. Built-in base imagery sources

There are a handful of built-in imagery options in CEO. Each of these options has different strengths and limitations.

MapBox is an open-source mapping platform for custom designed maps. The satellite imagery uses different sources based on the zoom level and geographic availability.

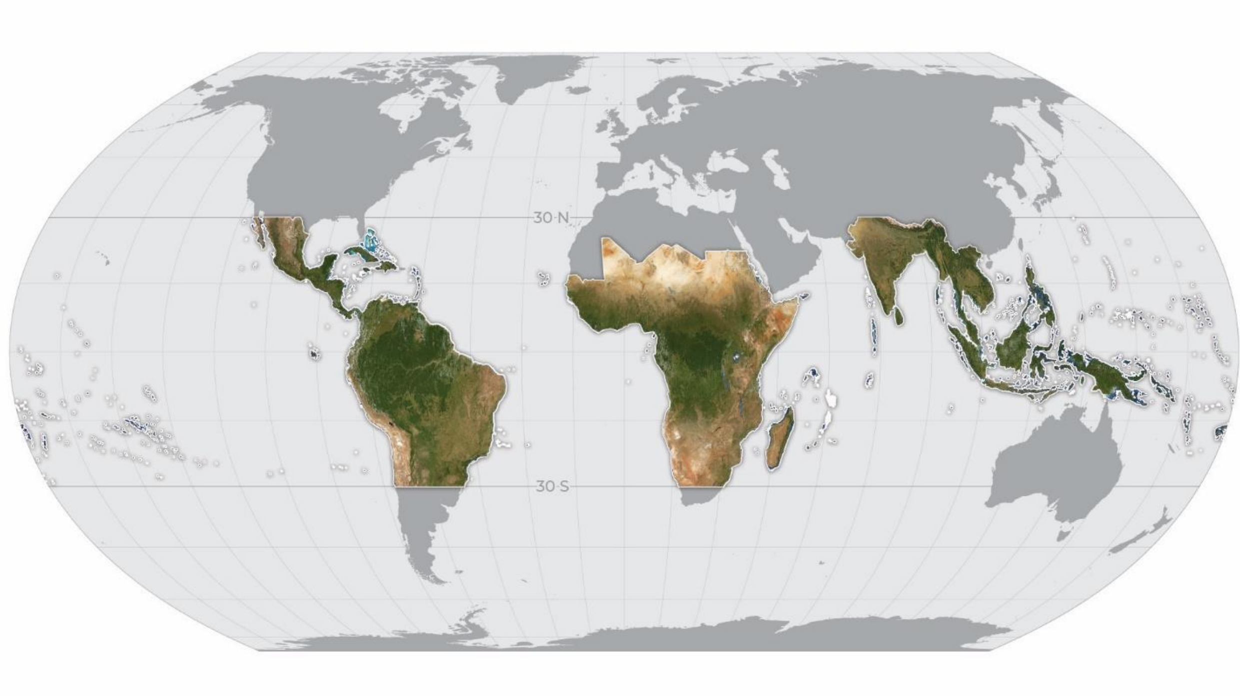
- **Zoom levels 0-8** use de-clouded data from NASA MODIS satellites.
- **Zoom levels 9-12** use NASA/USGS Landsat 5 & 7 imagery.
- **Zoom levels 13+** use a combination of open and proprietary sources, including Digital Globe for much of the world, USDA's NAIP 2011–2013 in the contiguous United States, and open aerial imagery from Denmark, Finland, and parts of Germany.

MapBox Satellite and Mapbox Satellite w/Labels are thus composite satellite imagery, where each map tile may be stitched together from imagery acquired on multiple dates. There is not necessarily a single date for an imagery tile. This is the same as, for example, Google Maps and Bing Aerial Imagery.

CEO has created a plain satellite layer as well as one with labels. More information is available here: https://www.mapbox.com/ and https://docs.mapbox.com/help/how-mapbox-works/satellite-imagery/.

> Maxar has eliminated multiple data products that were previously available on CEO, including: DigitalGlobeRecentIMagery; DigitalGlobeRecentImagery+Streets; DigitalGlobeWMSImagery; and EarthWatch.

Planet NICFI Public is imagery available through a partnership between Norway's Ministry of Climate and Environment, specifically Norway's International Climate and Forest Initiative (NICFI) and Planet. Along with other partners, they have made high-resolution (sub 5-m pixel) imagery available in the tropics. More information can be found here: https://www.planet.com/pulse/planet-ksat-and-airbus-awarded-first-ever-global-contract-to-combat-deforestation/.



B. Adding additional basemap imagery

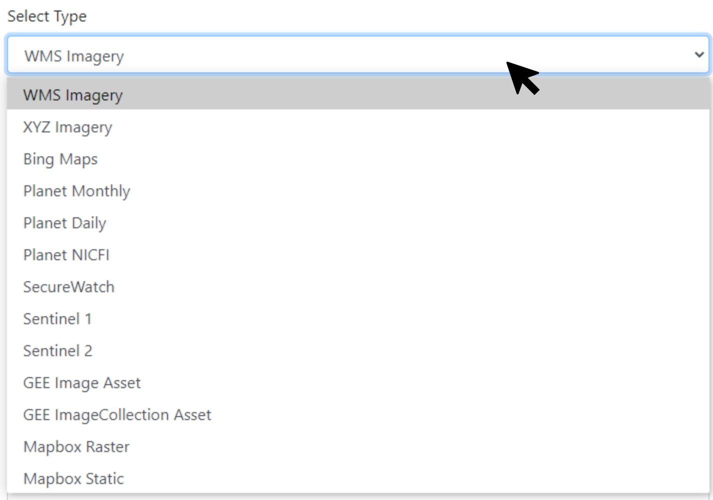
The built-in imagery is not sufficient as basemaps for some projects, including projects comparing land uses between two time periods, or projects that need imagery from specific dates or months. CEO allows institutions to add new imagery sources through their Institution page. This section discusses the different imagery types that can be added.

The instructions below assume you are starting on your Institution page (see directions in Part 2: A.2-4) and are logged in as an Administrator for your Institution.

Under the **Imagery** panel on the **Institution** page, click **[Add New Imagery]**. Under **Select** Type you will see multiple different options. Some, including WMS Imagery and XYZ Imagery, are protocols that can be applied broadly, while other options connect to specific data sources that you generally need to subscribe to.

For BingMaps, Planet Monthly, Planet Daily, Planet NICFI, SecureWatch, and the MapBox products you will first need to locate your no-authentication API key. You will be asked to provide this key in the CEO imagery setup panel.

We will now talk about how to add imagery for each of these options.



1. WMS Imagery

   i. This option can be used to access data your Institution hosts or data services that use WMS or WTMS. Imagery must be accessible using WMS or WTMS. Images stored locally on a user's computer cannot be added to a CEO project, although they can be accessed if you upload them to a WMS or WTMS (as a basemap), or as a Google Earth Engine Asset (for Geo-Dash and/or basemap).
   ii. The information needed to set up the WMS server will differ based on the specific imagery you are.
   iii. The first step of preparing to connect to your WMS with CEO is checking that your WMS is compatible with CEO.
      a) **CEO requests data in EPSG:3857**. Your WMS server is responsible for reprojecting the imagery it hosts into this requested projection.
   iv. Second, to connect your web map service (WMS) to CEO, you need to gather three pieces of information: the base URL for your WMS, the layer name that you want, and any parameters your WMS requires to serve imagery.
   v. **Title:** This will be the displayed name of the imagery.
   vi. **Attribution**: This is the displayed attribution for your imagery. It will appear at the top of the map while users perform data collection tasks.
   vii. **URL**: This should be the URL of your WMS/WTMS. It should be http:// or https://
   viii. **Layer Name**: This is the layer name from your server that you want to display.
   ix. **Params (as JSON object)**: If there are any parameters for the layer you want to specify, put them here as a JSON object. Documentation for the WMS getMapService is available online here: https://docs.geoserver.org/stable/en/user/services/wms/reference.html#getmap. These work only if they are parsed by the target WMS. Here are some examples:

```
{"TILED": true}
```

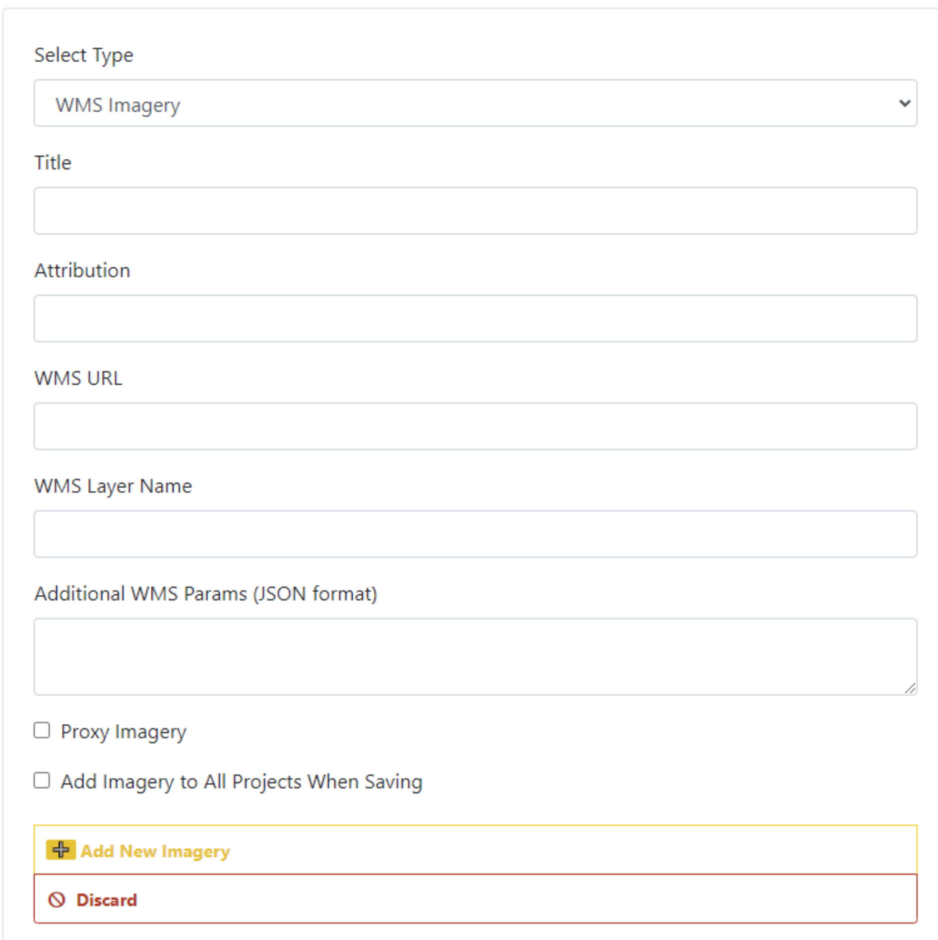
```
{"VERSION": "1.1.1", "CONNECTID": "63f634af-fc31-4d81-9505-
b62b4701f8a9", "FEATUREPROFILE": "Accuracy_Profile",
"COVERAGE_CQL_FILTER": "(acquisition_date>'2012-01-
01')AND(acquisition_date<'2012-12-31')"}
```

Quotes **MUST BE** "" ASCII/neutral/vertical quotation marks. Using “” smart/typographic/curved quotation marks will cause errors. Use only Unicode U+0022 and U+0027.

x. If you need to proxy your imagery for your WMS, please check **Proxy Imagery**. You may need this option if you need to obfuscate a key for your WMS or WTMS.

xi. If you want to add this imagery source to all of your institution’s projects, check the box next to **Add Imagery to All Projects When Saving.**

xii. When all fields are filled out, click on **[Add New Imagery]**.

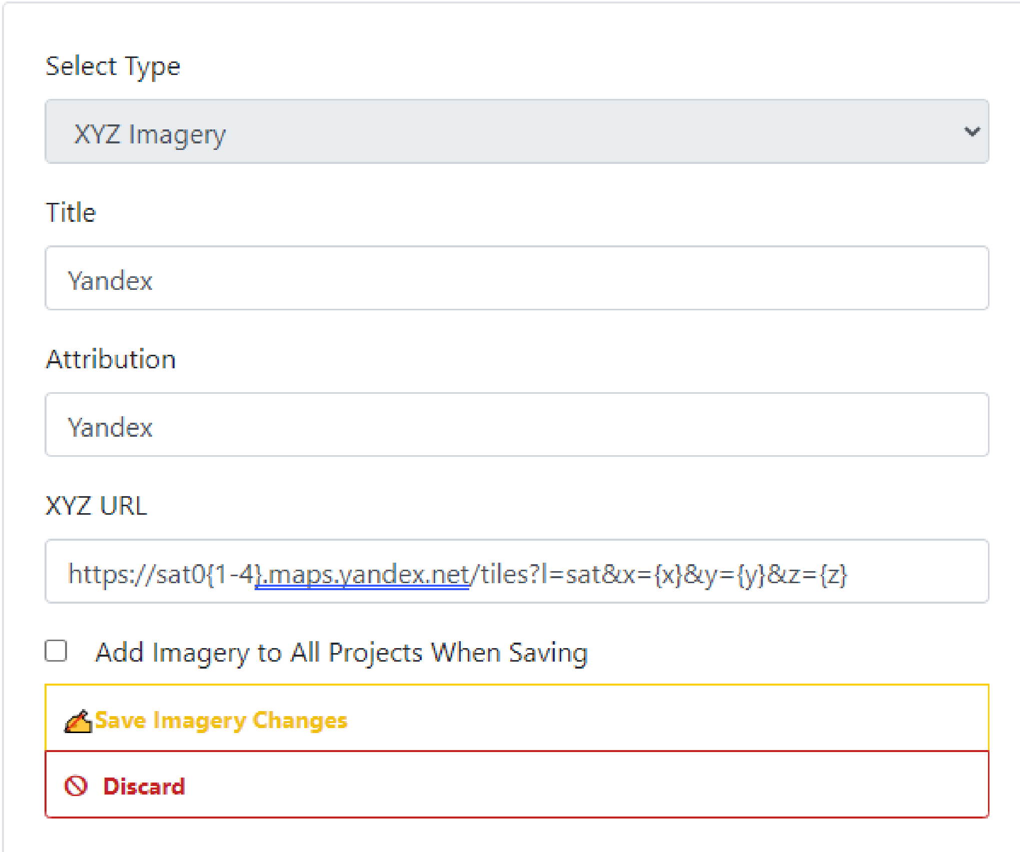

Select Type
WMS Imagery
Title
Attribution
WMS URL
WMS Layer Name
Additional WMS Params (JSON format)
☐ Proxy Imagery
☐ Add Imagery to All Projects When Saving
Add New Imagery
Discard

If you are having trouble, we have posted an extensive how-to and troubleshooting guide on the CEO blog. You can view that guide here: https://blog.collect.earth/index.php/2022/04/01/add-wms-to-ceo/.

2. XYZ Tiles

i. XYZ Tiles is a flexible way of adding imagery. Services you can now add includes any of the adaptations of Open Street Maps, Yandex, and any other service using XYZ.
ii. **Title:** This will be the displayed name of the imagery.
iii. **Attribution** for the XYZ Imagery. This will be shown to your users in the data collection pane.
iv. Add the **XYZ URL**.
   a) By default, XYZ uses the widely-used Google grid, where (x,y) (0,0) are in the top left.
   b) In general, it should have the format: <URL>?x={x}&y={y}&z={z} . For example, https://sat0{1-4}.maps.yandex.net/tiles?l=sat&x={x}&y={y}&z={z} Another example, https://mt1.google.com/vt/lyrs=y&x={x}&y={y}&z={z}
   c) You can also use grids where (x,y) (0,0) are in the bottom left. To do so, you will need to use the following format: <URL>?x={x}&y={-y}&z={z}
   d) Note that sometimes you will need to edit a provided URL to make it compatible with CEO. For example, the Open Street Map WIKI provides the following URL for the German version of Open Street Maps: https://a.tile.openstreetmap.de/${z}/${x}/${y}.png. As provided this will not work. To make it work, you will need to delete the "a." subdomain and the "$" characters.
   e) The edited URL does work in CEO: https://tile.openstreetmap.de/{z}/{x}/{y}.png

v. If you want to add this imagery source to all your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
vi. When all fields are filled out, click on **[Add New Imagery]**.



3. BingMaps

   i. This imagery option allows you to add Bing Maps with your own API key. The imagery provided by Bing Maps is composite satellite imagery. This means that each map tile is stitched together from imagery acquired on multiple dates. There is not a single date for an imagery tile. Some map tiles contain imagery collected over a multi-day window while other tiles contain imagery collected over a multi-year window. As there is not a single date for an imagery tile, CEO cannot provide the exact date of the imagery used. If you're interested in learning more, the Bing Maps API can be found here: https://docs.microsoft.com/en-us/bingmaps/rest-services/imagery/imagery-metadata. The Bing tile system uses the Mercator projection and has 23 levels of zoom (though not all levels are available in all locations). Commonly, the resolution at max zoom is about 0.3 m per pixel. For more information, see https://docs.microsoft.com/en-us/bingmaps/articles/bing-maps-tile-system.

   ii. **Title:** This will be the displayed name of the imagery.

   iii. **Imagery Id:** Only Aerial and AerialWithLabels are currently implemented. Note that the AerialWithLabels imagery uses the legacy static tile service, which is deprecated, and current data will not be refreshed. It therefore may have older imagery than the Bing Aerial dataset.

iv. **Access Token:** Your BingMaps key. For more information or to obtain your own key, see https://docs.microsoft.com/en-us/bingmaps/getting-started/bing-maps-dev-center-help/getting-a-bing-maps-key.

v. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**

vi. When all fields are filled out, click on **[Add New Imagery]**.

Directions for requesting a Bing maps key:

- To use Bing Maps imagery for your projects, you can create your own FREE Bing maps key to connect your institution's projects to your Bing Maps account. The full directions for creating a key are here: https://docs.microsoft.com/en-us/bingmaps/getting-started/bing-maps-dev-center-help/getting-a-bing-maps-key
- Visit https://www.bingmapsportal.com/ to request a Bing key or copy your existing key.
    - Sign in. You will need a Bing maps account or Microsoft account (https://docs.microsoft.com/en-us/bingmaps/getting-started/bing-maps-dev-center-help/creating-a-bing-maps-account)
    - Once you have logged in, click on **My account**, then click on **My Keys**
    - If you already have a key, click **Show key** or **Copy key**
    - If you do not have a key, click on **Click here to create a new key**.
    - Fill out the information. Application URL is optional (and I might suggest not using it) but if you do, use https://collect.earth as your Application URL
    - You will create a **Basic** key. If you need more imagery, you will need to talk to Microsoft and request an **Enterprise key** (https://www.microsoft.com/en-us/maps/create-a-bing-maps-key#enterprise)

4. Planet Monthly

i. Planet offers multiple data products (product specification here: https://support.planet.com/hc/en-us/articles/360022233473-Planet-Imagery-Product-Specifications) . This option pulls from the Planet Monthly mosaic product, which allows you to display imagery from a specific month. User help is available here: https://developers.planet.com/docs/apps/explorer//.

ii. **Title:** This will be the displayed name of the imagery.

iii. **Default Year**: The default year that will be displayed when the map loads.

iv. **Default Month**: The default month that will be displayed when the map loads. Use integer format 1-12.

v. **Access Token**: Your Planet access token. This can be accessed through your My Account page on the Planet website.

vi. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**

vii. When all fields are filled out, click on **[Add New Imagery]**.

Default Year & Default Month will let you put in any integer, positive or negative. The up and down arrow keys start at 0. Please type year in the YYYY format and month as an integer between 1-12.

5. PlanetDaily

   i. PlanetDaily is another imagery product available from Planet. It allows users to detect land use and land cover change in near real time. This data source allows you to select a start and end date, with up to daily imagery resolution. Your study area might not have full coverage every day.
   ii. **Title:** This will be the displayed name of the imagery.
   iii. **Access Token:** Your PlanetDaily API key. This can be accessed through your My Account page on the Planet website.
   iv. **Start Date:** Starting date for the imagery you are interested in; you can input the date using numeric keys or with the calendar widget on the right side.
   v. **End Date:** Ending date for the imagery you are interested in.
   vi. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
   vii. When all fields are filled out, click on **[Add New Imagery]**.

6. Planet NICFI

   i. This allows you to add your own Planet NICFI key, instead of using CEO's. Note that you will need an account from the Planet NICFI program: https://www.planet.com/nicfi/. This is separate from your 'normal' Planet account.
   ii. **Title:** This will be the displayed name of the imagery.
   iii. **Access Token:** Your PlanetDaily API key. This can be accessed through your My Account page on the Planet website.
   iv. **Default Time:** Choose the default time period of imagery to display. Note that the time periods available from NICFI are actively being changed. THESE ARE SUBJECT TO CHANGE BASED ON PLANET AND NICFI's DECISIONS.
   v. **Default Band**: Choice between Visible (RGB) and Infrared false color.
   vi. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
   vii. When all fields are filled out, click on **[Add New Imagery]**.

7. SecureWatch Imagery

i. SecureWatch is a service from Maxar focused on monitoring for new land use/land cover changes and comparing current land use/land cover with over 20 years of historic images. For more information see: https://www.digitalglobe.com/products/securewatch.
ii. https://gcs-docs.s3.amazonaws.com/Access/Miscellaneous/DevGuides/WMS/WMS_Map.htm?Highlight=key
iii. **Title:** This will be the displayed name of the imagery.
iv. **Connect ID:** This is your API key. You need to use a no-auth key here, which should be a string of letters and numbers separated by dashes.
v. **Start Date:** Starting date for the imagery you are interested in; you can input the date using numeric keys or with the calendar widget on the right side.
vi. **End Date:** Ending date for the imagery you are interested in.
vii. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
viii. When all fields are filled out, click on **[Add New Imagery]**.

8. Sentinel 1 Imagery

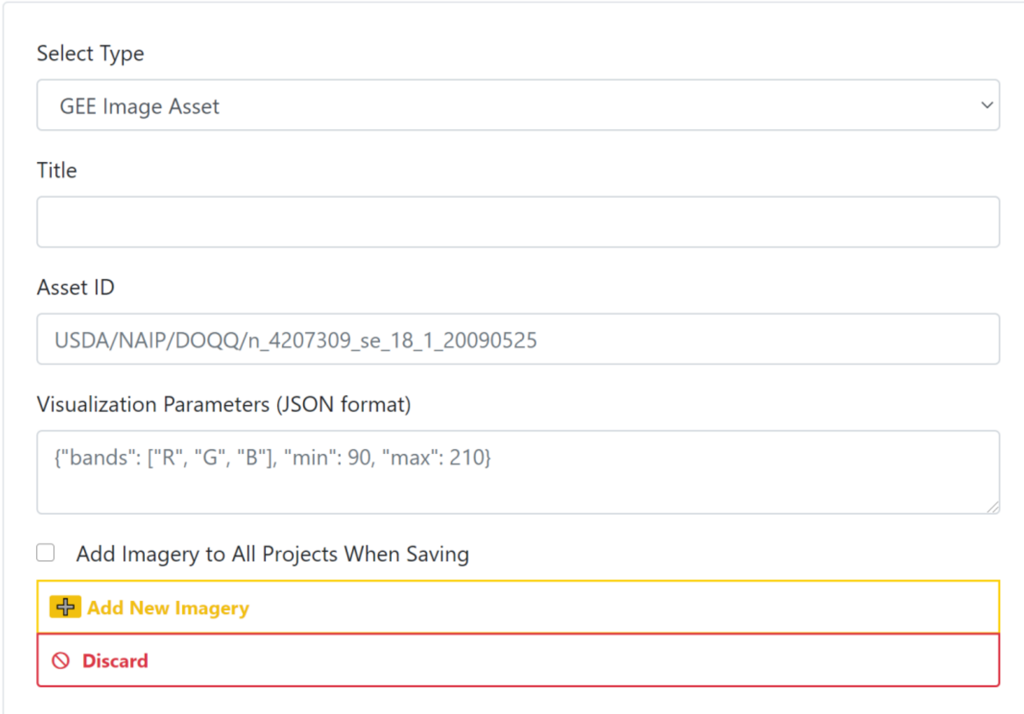
i. Sentinel 1 information is only available from April 2014 to present (Sentinel 1A launch). Sentinel data is available in CEO through GEE. If multiple images are available for the region and dates selected, the median reducer is used to produce a single image.
ii. **Title:** This will be the displayed name of the imagery.
iii. **Default Year**: The default year that will be displayed when the map loads.
iv. **Default Month**: The default month that will be displayed when the map loads. Use integer format 1-12.
v. **Band Combination**: Preset combinations of bands for most uses. VH and VV are single polarization, VH/VV and HH/HV are dual polarization. More info https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes.
vi. **Min:** Minimum value for bands that will get mapped to 0 for visualization. This can be one value for all bands, or a value for each of the three bands. This should be one number. Acceptable values for each band's minimum are the same as for Sentinel imagery available through GEE generally; see https://developers.google.com/earth-engine/datasets/catalog/sentinel. Min can be as low as -50, but 0 is frequently used.
vii. **Max:** Maximum value for bands that will get mapped to 255 for visualization. This can be one value for all bands, or a value for each of the three bands. This should be one number. Acceptable values for each band's maximum are the same as for Sentinel imagery available through GEE generally. See link above. Max can be as high as 1, but .3 is frequently used.
viii. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
ix. When all fields are filled out, click on **[Add New Imagery]**.

9. Sentinel 2 Imagery

i. Sentinel 2 imagery is available from June 2015-present. Sentinel 2 imagery is displayed in CEO from GEE. If multiple images are available for the region and dates selected, the median reducer is used to produce a single image.
ii. **Title:** This will be the displayed name of the imagery.
iii. **Default Year**: The default year that will be displayed when the map loads.
iv. **Default Month**: The default month that will be displayed when the map loads. Use integer format 1-12.
v. **Band Combination**: Select one of the options available, including True Color, False Color Infrared, False Color Urban, Agriculture, Healthy Vegetation, and Short Wave Infrared.
vi. **Min:** Minimum value for bands that will get mapped to 0 for visualization. This can be one value for all bands, or a value for each of the three bands. This should be a single number. Acceptable values for each band's minimum are the same as for Sentinel imagery available through GEE generally; see https://developers.google.com/earth-engine/datasets/catalog/sentinel For example, 0 could be used.
vii. **Max:** Maximum value for bands that will get mapped to 255 for visualization. This should be a single number. Acceptable values for each band's maximum are the same as for Sentinel imagery available through GEE generally. See link above. For example, values of 2800-4000 are frequently used.
viii. **Cloud Score**: Allowable cloud cover. Values can be 0-100.
ix. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
x. When all fields are filled out, click on **[Add New Imagery]**.

10. GEE Image Asset

i. Google Earth Engine (GEE) Assets include user's uploaded assets along with assets provided by other users and GEE. Information on Assets can be found here: https://developers.google.com/earth-engine/guides/asset_manager. More detail on uploading your own assets is below. Note that Image Asset refers to a single image (e.g. a GeoTIFF layer) while ImageCollection Asset refers to a stack of images (e.g. GeoTiff layers of the same location over different dates).
ii. **Title:** This will be the displayed name of the imagery.
iii. **Asset ID:** The Asset ID for your image asset. Will have a format similar to: USDA/NAIP/DOQQ/n_4207309_se_18_1_20090525
iv. **Visualization Parameters (JSON format)**: Any visualization parameters for your layer. For example, {"bands":["R","G","B"],"min":90,"max":210}
v. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
vi. When all fields are filled out, click on **[Add New Imagery]**.



We have created an extensive how-to and troubleshooting guide for adding GEE Image Assets and Image Collection Assets to CEO. It is available in CEO's blog here: https://blog.collect.earth/index.php/2022/04/21/connecting-gee-raster-data/. Additionally, we have directions for uploading your own image assets to GEE in the following section.

11. GEE ImageCollection Asset

i. Google Earth Engine (GEE) Assets include user's uploaded assets along with assets provided by other users and GEE. Information on Assets can be found here: https://developers.google.com/earth-engine/guides/asset_manager. More detail on uploading your own assets is below. Note that Image Asset refers to a single image (e.g. a GeoTIFF layer) while ImageCollection Asset refers to a stack of images (e.g. GeoTiff layers of the same location over different dates).
ii. **Title:** This will be the displayed name of the imagery.
iii. **Asset ID:** The Asset ID for your image asset. Will have a format similar to: LANDSAT/LC08/C01/T1_SR
iv. **Start Date**: The default start date of imagery to display.
v. **End Date**: The default end date of imagery to display.
vi. **Visualization Parameters (JSON format)**: Any visualization parameters for your layer. For example, {"bands":["B4","B3","B2"],"min":0,"max":2000}

vii. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**

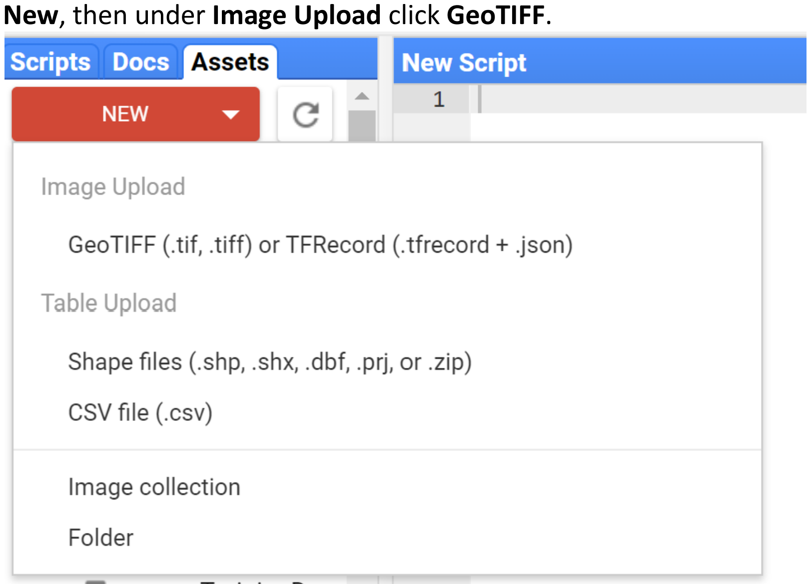
viii. When all fields are filled out, click on **[Add New Imagery]**.

ix. Note that to display the GEEImageCollection, CEO uses the "mean" reducer in Earth Engine. This takes the mean of any images in the image collection during the time period specified.

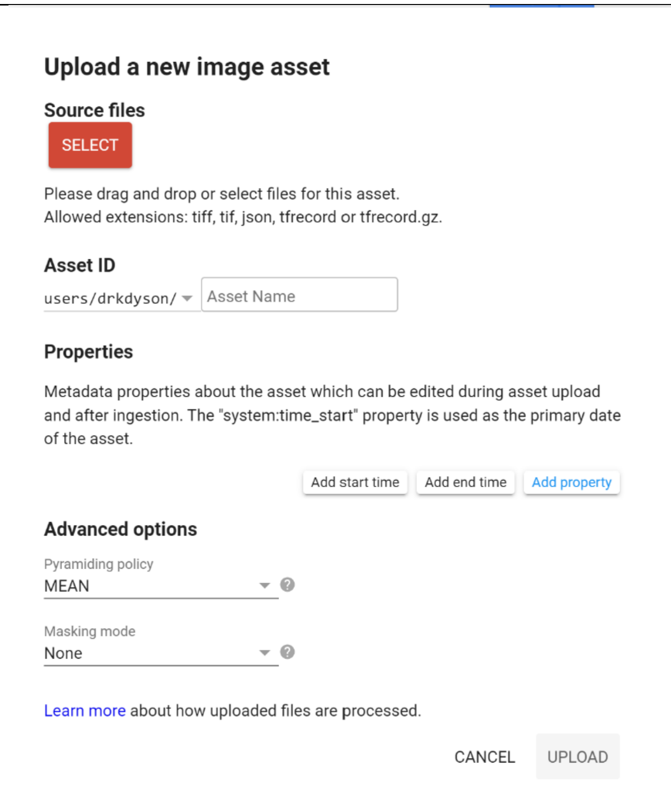
We have created an extensive how-to and troubleshooting guide for adding GEE Image Assets and Image Collection Assets to CEO. It is available in CEO's blog here: https://blog.collect.earth/index.php/2022/04/21/connecting-gee-raster-data/.

Uploading GeoTIFF images to GEE:

- Visit https://code.earthengine.google.com/
- Navigate to **Assets**
- Click **New**, then under **Image Upload** click **GeoTIFF**.



- A new window will pop up. Click **Select** and navigate to the GeoTIFF asset.
- Alter the **Asset ID** name if you would like.
- Check the default settings, e.g. if your data has a start/end time then add those.
- Click **Upload.**



- Wait for your asset to upload. When it does, click on the asset name and a new window will pop up.
- Next to ImageID there are two interlocking squares.
- Click on this to copy your ImageID to the clipboard (you'll need this to connect to CEO).
    a. You can also click on the "Bands" tab to get more information about your image. This is very useful for specifying your visualization parameters in CEO.
    b. Can also add gamma etc. information, see https://developers.google.com/earth-engine/guides/image_visualization

12. MapBox Raster

i. MapBox Raster serves raster tiles including Mapbox Satellite. For more information see:: https://docs.mapbox.com/help/glossary/raster-tiles-api/.
ii. **Title:** This will be the displayed name of the imagery.
iii. **Layer Name** is the desired layer name from MapBox.
iv. **Access Token** will be your no-auth key from MapBox. For more information, see https://docs.mapbox.com/help/glossary/raster-tiles-api/.
v. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**

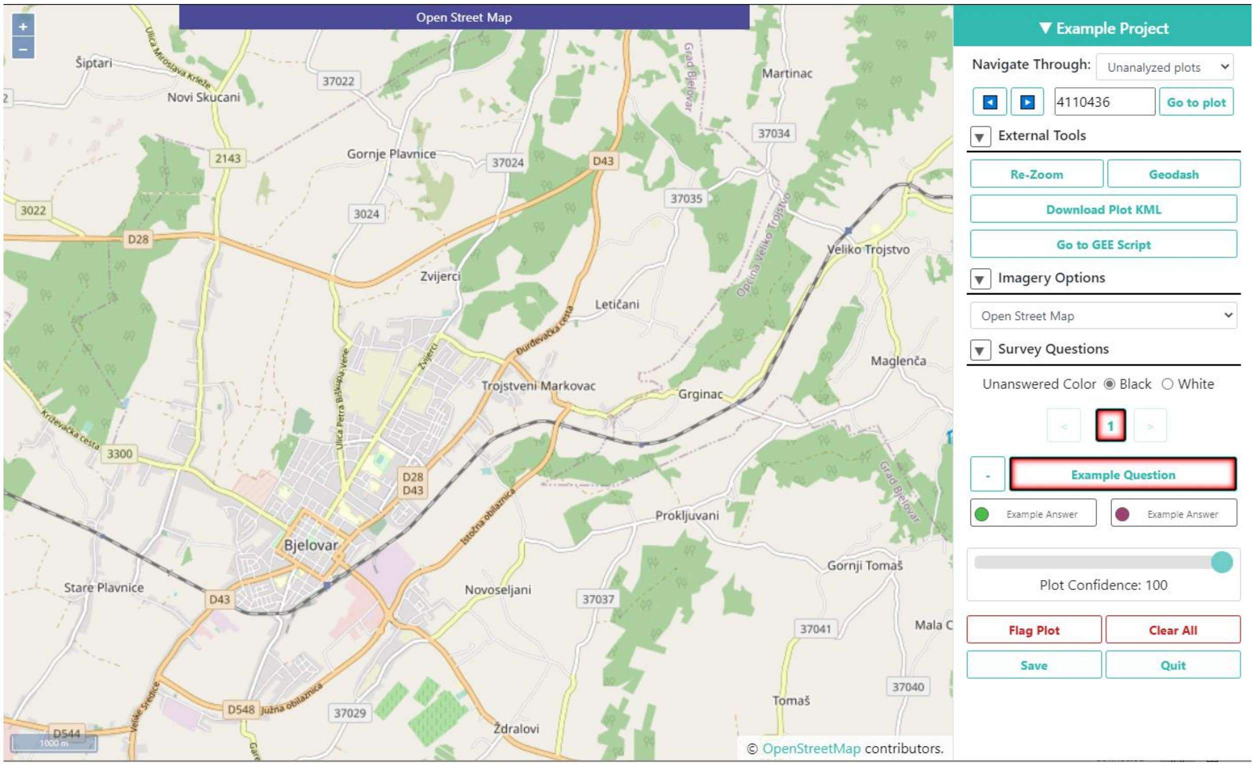
vi. When all fields are filled out, click on **[Add New Imagery]**.

13. MapBox Static

i. Mapbox Static serves raster tiles generated from a Mapbox GL-based style. This API has additional parameters that can be used to refine the results of a request. More information see: https://docs.mapbox.com/help/glossary/static-tiles-api/
ii. **Title:** This will be the displayed name of the imagery.
iii. **User Name** will be your MapBox user name.
iv. **Map Style ID** will be the id from MapBox.
v. **Access Token** will be your MapBox no-auth key. For more information see https://docs.mapbox.com/help/glossary/static-tiles-api/.
vi. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
vii. When all fields are filled out, click on **[Add New Imagery]**.

14. Open Street Maps

i. Open streets Maps is an open source mapping product. Note this provides Open Street Map's Standard Tile Layer (Standard tile layer - OpenStreetMap Wiki). This product is useful for orienting users, since it has street and place names.
ii. **Title:** This will be the displayed name of the imagery.
iii. If you want to add this imagery source to all of your institution's projects, check the box next to **Add Imagery to All Projects When Saving.**
iv. When all fields are filled out, click on **[Add New Imagery]**.



For imagery options with dates, the dates you input are the default dates that the imagery will be restricted to on the collection page. However, the user will be able to change these when exploring the map as there are start & end date widgets on the collection page sidebar (there are examples in the **Data Collection Manual**). For SecureWatch, the user will also be able to choose between FeatureProfiles. Without specifying a FeatureProfile, the most recent available imagery between the start and end dates displayed on the map.

SecureWatch and the Planet products will not return imagery if the map is zoomed out too much. This results in a white map canvas being displayed at the project overview level usually. Simply click the "Go to first plot" button on the Collection page to zoom in to the plot level, and then the imagery should appear.

For SecureWatch, the date shown when data is collected will be added to the project .csv data available for download (See Part 7: E “Download your data”).

C. Adding imagery from multiple time periods

Adding multiple imagery options with different default time periods can make data collection easier for projects that compare two or more time periods to detect land use and land cover change. WMS/WMTS that you can use to create basemaps from different time points include GeoServer, Planet Monthly, Planet Daily, Secure Watch, Bing Maps, Mapbox Raster, and Mapbox Static.

1. GeoServer:

   i. For GeoServer, how to add different years of imagery depends on your server.

      a) If your different years are stored as different layers, alter the GeoServer Layer Name field when you add the second layer. Make sure your title/attribution/etc. fields are accurate for the new layer.
      b) If your server uses filtering to display imagery from different years you will need to alter the GeoServerParams field (again, making sure the information in your other fields is correct).

   ii. Once you have decided the best approach for your server, repeat the steps in Part 3 B.1 above for each time period you would like to add.
   iii. **Note that some years may not contain any imagery**, due to the sparseness of the data within the database. If no imagery for the selected time range appears, you will need to change your GeoServerParams field, possibly to change the feature profile or date ranges.

2. Planet Monthly, Planet Daily, and Planet NICFI

   i. For all Planet products, you simply need to change the time period fields to add layers with different default time periods. Users will be able to change the time period displayed during data collection; however, this is the default that will be shown first.
   ii. Be sure to change the Title field to reflect the correct default Year, Month, and Day for each new layer that you add.

3. Secure Watch

   i. For this Maxar product, you simply need to change the time period fields to add layers with different default time periods. Users will be able to change the time period displayed during data collection; however, this is the default that will be shown first.
   ii. Be sure to change the Title field to reflect the correct default Year, Month, and Day for each new layer that you add.

4. Sentinel 1 & 2

   i. Users will be able to change the default Year and Month during data collection.

ii. However, additional layers with different default years and months can be added based on user preference.
iii. Be sure to change the Title field to reflect the correct default Year, Month, and Day for each new layer that you add.

5. GEE Image Asset

   i. Different Image Assets that cover different periods of time can be added using the Imagery interface.
   ii. However, if you have multiple images of the same area over different periods of time, consider using ’sGEE’s ImageCollection functionality rather than multiple Image assets.

6. GEE ImageCollection Asset

   i. Users will be able to change the default Year and Month during data collection.
   ii. However, additional layers with different default years and months can be added based on user preference.
   iii. Be sure to change the Title field to reflect the correct default Year, Month, and Day for each new layer that you add.

D. Editing and deleting imagery

After you have added imagery you may need to change the default dates or the Visualization Parameters.

1. Navigate to your Institution page.
2. Next to the imagery you would like to edit, there is an editing hand button.

3. When you click on the editing hand button, it will take you back to the imagery creation form.
4. Edit the values as needed, using Section B above as a guide.
5. Check the Add Imagery to All Projects When Saving if you would like to add your imagery to all of the institution’s projects.
6. When you are done, click Save Imagery Changes.
7. You can delete imagery by clicking the trash can icon next to the imagery name.

There is currently no way to 'preview' what imagery will be visible in your new layer in the 'Add Imagery' workflow. There are two ways to work around this.

First, if your data source has a data viewing portal, you can use this to explore the imagery and determine what is available for the time periods you are interested in.

Second, you can add the imagery layer, then open an existing project from your institution. The imagery will be available in the dropdown menu (if you are switching between a project window and an Institution window, you may need to refresh the project window to get the new layer to appear). You can then check if the imagery is displaying correctly and go back to the Institution page to edit the imagery based on what you see.

E. Estimating imagery costs

Before setting up a project in CEO, it is important to estimate how much imagery will be used for budgetary and resource allocation. Here is a quick guide to help.

1. When is imagery used? Imagery data is used whenever there is a map on the page. This means that on CEO, all these pages can use data:

    i. Home
    ii. Data Collection
    iii. Create Project
    iv. Review Project
    v. Project Dashboard
    vi. GeoDash (specific options or modules)
    vii. TimeSync (when implemented)

On these pages, when the map first loads, imagery data is used. Every time a user zooms or pans the map window, imagery is used. The largest amount of use will probably be with Data Collection.

2. Estimating imagery use for a project:

i. Each organization that provides imagery sets their own rules for how many tiles you can download per year given the kind of account that you have with them. Therefore, it is important to estimate this before setting up a project.

ii. Additionally, services may "count" imagery against your quota differently. For example, Planet uses a rule "*Every Pixel is Charged Once*" so you can download a pixel multiple times but it is only charged once (see here: You will also need to know this. https://support.planet.com/hc/en-us/articles/360021227554--When-is-a-Planet-product-charged-against-my-quota-). You will also need to know this.

iii. To figure out how much imagery you are likely to use for a single project, count the number of plots. Then determine how many users will classify each plot (in CEO this is usually 1 user per plot currently). Next, try to factor in how often people will zoom or pan their maps for context when answering the survey questions for a plot. Multiply these numbers together.

iv. Next, decide if you are using the Geo-Dash and ask yourself how many map widgets you will display on your Geo-Dash page. Multiply that number by the number of plots to get the amount of Geo-Dash imagery you will need. Keep in mind that Geo-Dash imagery counts against our annual limit for user memory/processing in GEE, whereas the usual global layers on CEO (Bing Maps, SecureWatch, Planet) have separate annual tile-based limits. For GEE, we recommend clipping and pre-processing the imagery to image assets or imageCollection assets for the collection area. This eliminates processing on-the-fly for each user that is collecting, as Geo-Dash can just grab the pre-processed image asset.

v. Finally, add a few extra tile downloads for loading the maps while creating and reviewing the project.

vi. Once you have a sense of how many map images you will need for your project, you will then need to look up the tile counting policy for the imagery service that you are using. For example, some of them count 15 tiles as 1 unit of usage. Others use different counting rules.

3. Tip to reduce imagery use: Consider setting your default background imagery to a cheaper source—like Bing—and only switching to more expensive paid imagery when you are at the correct zoom level.

F. Synthetic Aperture Radar (SAR) data in CEO for forest degradation

Observations of backscatter variations over time in satellite SAR data can be attributed to structure and moisture. For forests, these can be linked to changes in the moisture conditions of the trees and soil as well as changes in forest structure. These are extremely useful for e.g. detecting forest degradation.

Color display of SAR data for detecting forest degradation is possible in CEO using Sentinel 1 data, WMS data, and GEE Image Assets and ImageCollection assets.

For information on detecting forest degradation through our GEE widget functionality, please see Part 4: E Project options.

For information on detecting forest degradation through our Geo-Dash widget, please see Part 6: I (Forest) Degradation tool.

For more information on SAR, please see:

1. Kellndorfer, Josef. "Using SAR Data for Mapping Deforestation and Forest Degradation." SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation. Eds. Flores, A., Herndon, K., Thapa, R., Cherrington, E. NASA. 2019. DOI: . 10.25966/68c9-gw82; available online at: https://gis1.servirglobal.net/TrainingMaterials/SAR/Ch3-Content.pdf
Also https://www.servirglobal.net/Global/Articles/Article/2674/sar-handbook-comprehensive-methodologies-for-forest-monitoring-and-biomass-estimation
2. This one pager from SERVIR & SilvaCarbon:
https://servirglobal.net/Portals/0/Documents/Articles/2019_SAR_Handbook/SAR_VegIndices_1page_new.pdf

Part 4: Project Creation

With your imagery loaded, you are ready to create a new project.

Before starting a data collection effort in CEO, make sure you have concrete goals, indicators, and metrics for your initiative. While this manual focuses on technical issues, other resources are available to help you create these goals, indicators, and metrics. Some helpful resources include:

- The *Road to Restoration: A Guide to Identifying Priorities and Indicators for Monitoring Forest and Landscape Restoration*, found at: http://www.fao.org/in-action/forest-landscape-restoration-mechanism/resources/detail/en/c/1253837/. This guide from FAO & WRI outlines steps toward setting goals, choosing indicators, and defining metrics.

If your project is focused on land use or land cover classification, you also need to have a classification scheme and an interpretation key. A classification scheme should be exhaustive, exclusive, consistent with the purpose of the work, and sufficiently descriptive. The scheme can be either single level or hierarchical and it does not need to be of uniform detail. Data collectors refer to an interpretation key—a collection of rules, imagery, and guidance—to classify land cover elements. A comprehensive interpretation key allows data collectors to produce more consistent and reliable results.

- See Chapters 2 & 3 of the Theoretical Manual from Collect Earth Online, found at https://collect.earth/downloads/CEO_Theoretical_Manual.pdf.

Designing a project is an iterative process, and you will probably need to make multiple edits to projects in CEO as you refine your goals, land use or land cover classification schemes, imagery sources, etc. These changes can be made to a project before your project is published. This means that you can create a project and collect test data in it, and edit any errors you find before publishing the project.

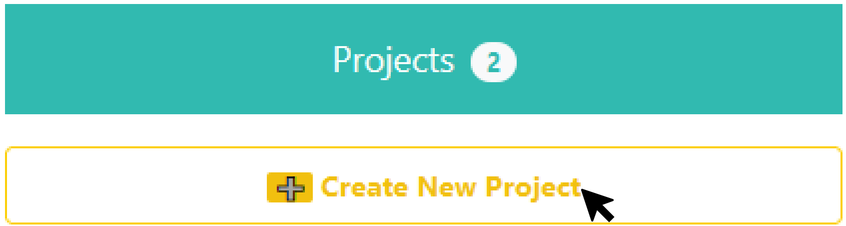
Some things can also be changed after a project is published, including imagery. For example, if you realize after you have created a project that you have forgotten to add an imagery layer, you can add it after project setup with the steps from Part 3.

The survey questions that are asked about each survey point cannot currently be changed after the project is published. Do not start data collection until you are sure that your survey questions are correct.

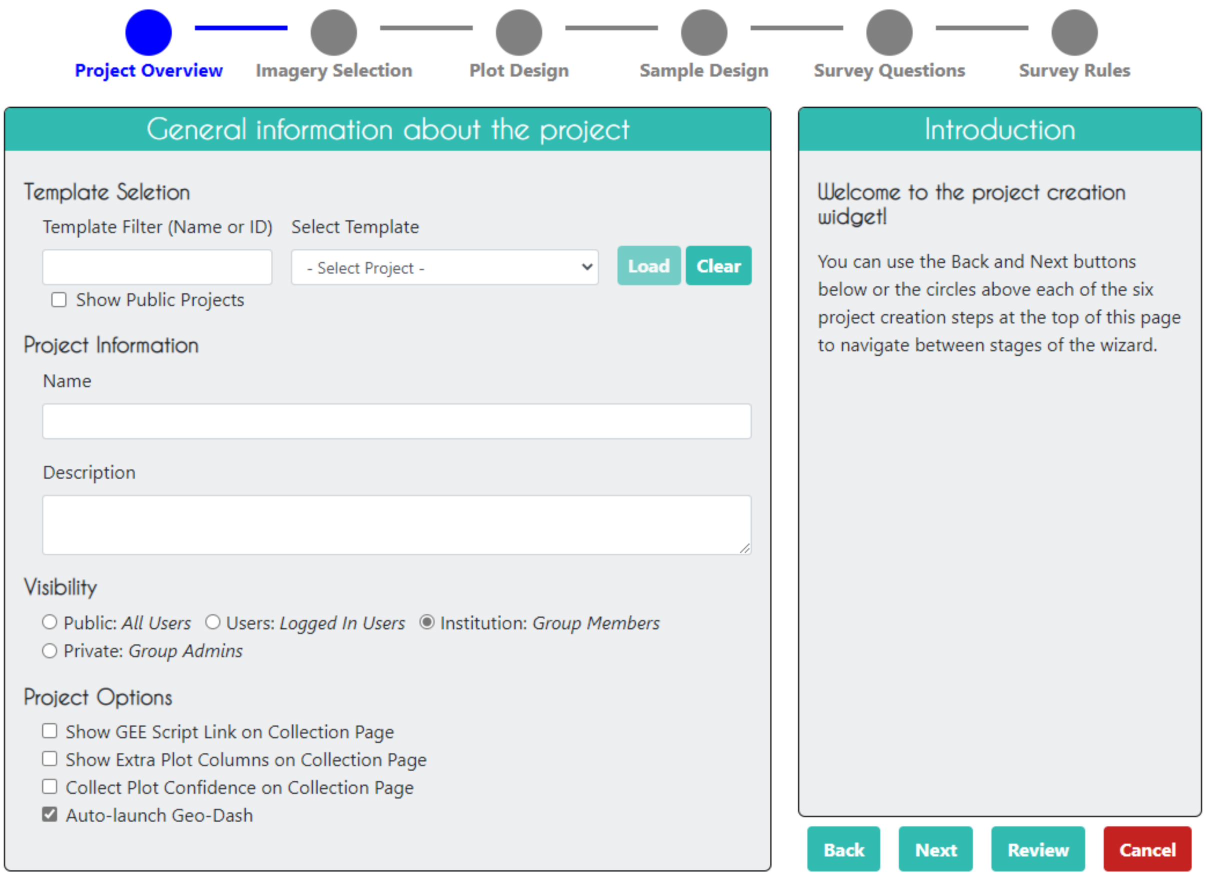
The instructions below assume you are starting on your **Institution** page (see directions in Part 2: A.2-4) and are logged in as an Administrator for your Institution.

A. Create a new project (wizard)

1. In the Projects tab on your Institution page, click [Create New Project]. This will bring you to the Create Project wizard.



2. In response to user feedback, CEO implemented a project creation wizard in the November 2020 update.
3. The wizard comprises 6 parts, each covered in one of the sections below.



B. Project Overview

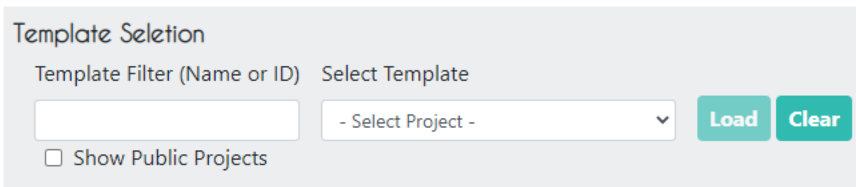
This section allows you to add general information about the project, including selecting a template (optional), project name, project description, and project options.

1. Use a project template (optional)

i. This feature is used to copy all the information—including project info, area, and sampling design—from an existing published project to a new project. This is useful if you have an existing project you want to duplicate for another year or location, or if you are iterating through project design. For a template, you may use any available published or closed project from your institution. You cannot use deleted projects. You cannot use another institution's private project but you can use another institution's public project (any website visitor or CEO user). For more on project privacy settings, see **Visibility** below.

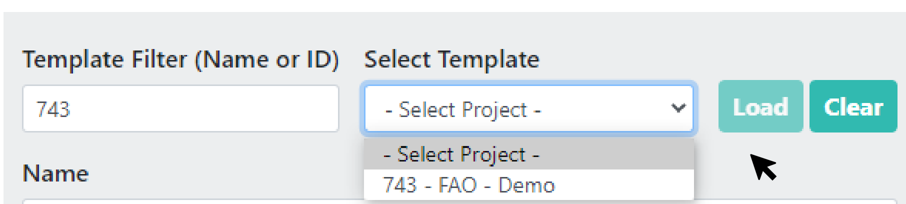
ii. If you do not want to copy another project, simply skip this section, and leave the **Select Template** field set to— **Select Project** -.

iii. Click on the **Show Public Projects** checkbox if you would like to view all public projects in CEO that you can use as a template. If you do not check this box, you will only see your institution's projects.



iv. Template Filter (Name or ID): To filter the projects, type in a keyword in an existing project's name or the Project ID Number. You can find the Project ID number by navigating to the project you want to copy and looking at the URL.

v. Then click on the dropdown menu under Select Project and click on the project's name



vi. When you find the project that you would like to use as a template, click on the project name, then click **Load** to load the template information.

vii. Click on **Clear** to clear all template information.

If you select a project under **Select Template**, and then if you clear the project selection the Imagery Preview in the next section may display only gray/white. Change the **Default Imagery** to any other source and then back to the source you are interested in and the basemap will reappear.

viii. Loading a template will create two checkboxes under **Copy Options**, **Copy Template Plots and Samples** and **Copy Template Widgets**. They are checked by default.

a) If **Copy Template Plots and Samples** is checked, the Plot Review and Sample Design sections will only display an overview of the number of plots, etc. Uncheck this box to change those parameters.
b) **Copy Template Widgets** refers to Geo-Dash options covered in Part 6: Geo-Dash Implementation.

2. Enter the project's Name and Description.

i. The **Name** should be short and will be displayed on the Home page as well as the project's Data Collection page.
ii. You should keep the **Description** short but informative. Users will see these if they click on the project's pin on the map on the home page. You will also see this when you are administering your project.
iii. If you are using a template, the **Name** and **Description** will automatically be populated. Be sure to change this to reflect your new project.

3. Select the project's Visibility.

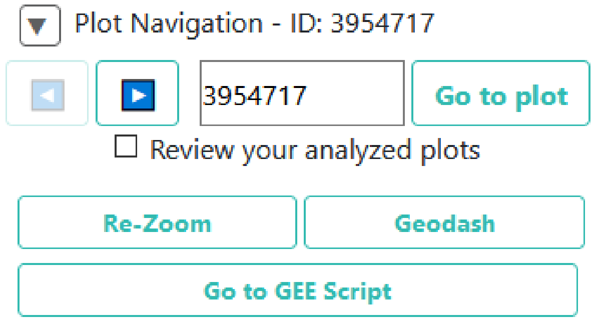
i. The Privacy Level radio button changes who can view your project, contribute to data collection, and whether admins from your institution or others creating new projects can use your project as a template.

- **Public: *All***: All users can see and contribute data to your project. Admins can use your project as a template.
- **Users: *Logged in Users***: Any user logged into CEO can see and contribute to your project. Admins can use your project as a template.
- **Institution: *Group Members***: Members of your institution can see and contribute to your project. Admins from other institutions cannot use your project as a template.
- **Private: *Group Admins***: Only your Institution's Admins can see and contribute to your project. Admins from other institutions cannot use your project as a template.

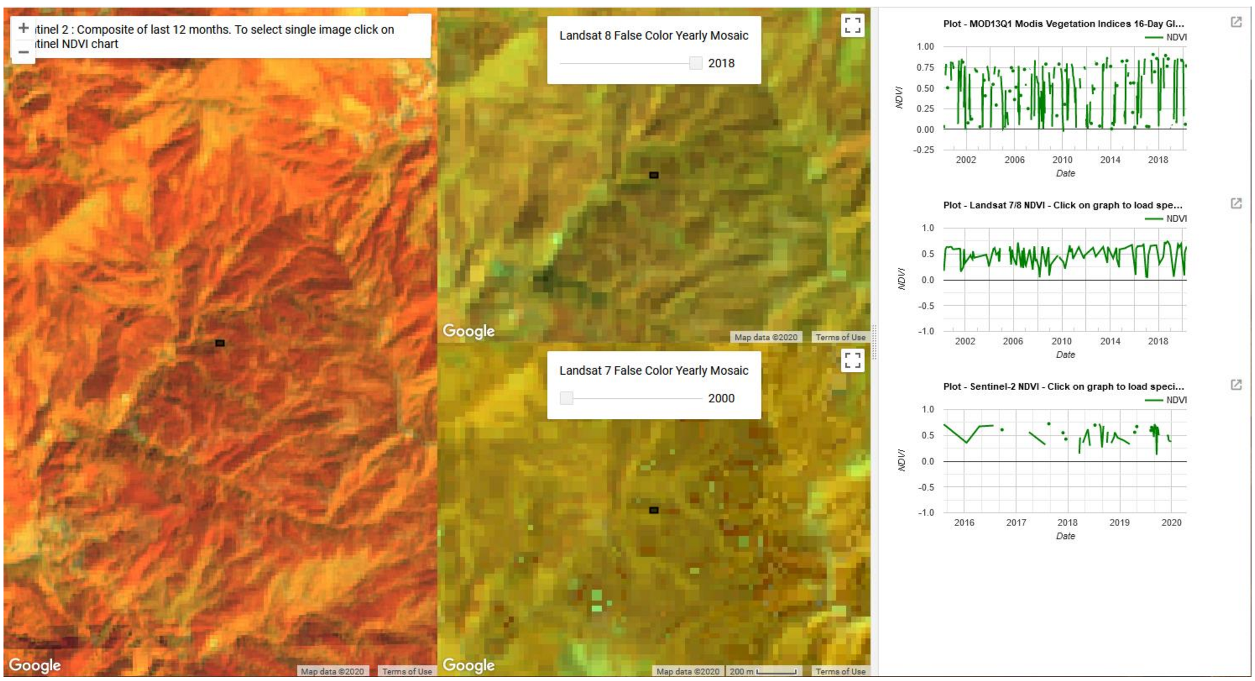
4. Select Project Options.

i. The first option is **Show GEE Script Link on the Collection page.**

a) This allows users in Data Collection to click on a button labeled Go to GEE Script.



b) This button will take them to a website with additional information about the plot. An example is shown here:



There are three panels. On the far left, there is a Sentinel 2 composite of the last 12 months. It is colored using an infrared color composite (near infrared, mid infrared, red). In the center are Landsat 8 and Landsat 7 Color Yearly mosaics, with a slider so you can choose between years. On the right are NDVI graphs of the plot from MODIS, Landsat 7/8, and Sentinel 2. For the Landsat 7/8 and Sentinel graphs, you can click a point on the graphs to load specific images in the left and center panels.

ii. The second option is **Show Extra Plot Columns on Collection Page**.

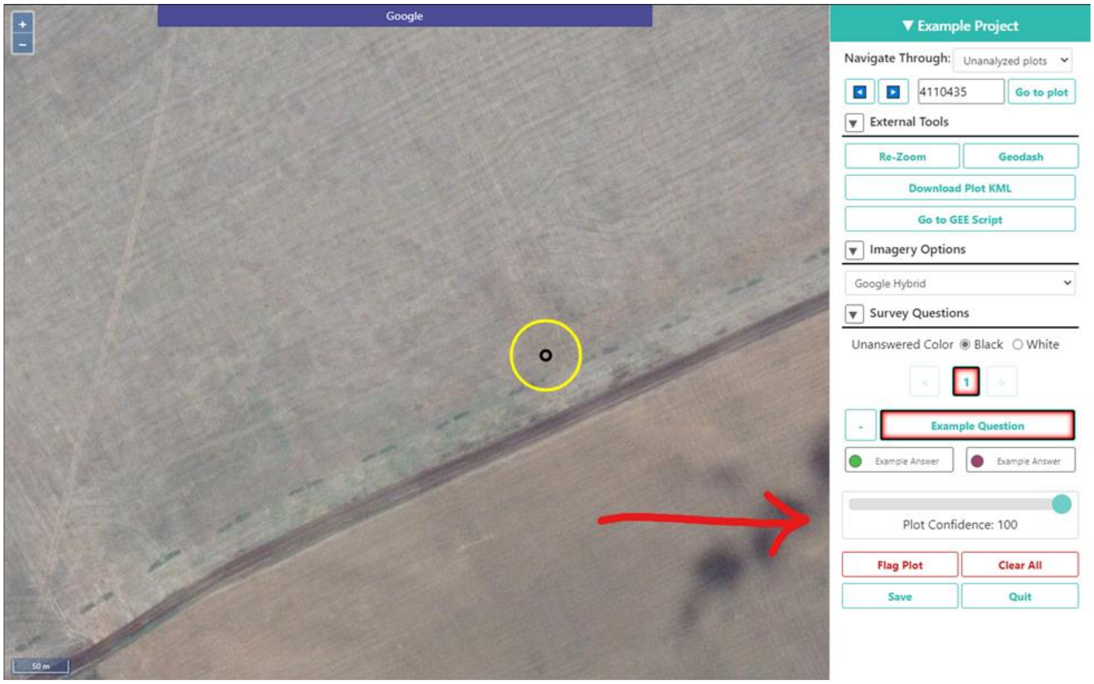
a) This option is only useful if you are using .csv or .shp files to define your **Plot Design**.

b) If you have additional columns in your .csv or .shp files, such as elevation information or land cover class, data collectors will be able to see them on the **Data Collection** page under **Plot Information**.

▼ Plot Information

- id - 132A
- plot - 132
- type - MDF
- year - 2016
- diam1 - 30
- diam2 - 31
- height - 15
- subplot - sub_plotA
- province - Xiengkhouang
- mean_diam - 30.5
- parent_index - 1
- biomass_per_mg_ha - 0.08837375
- biomass_per_kg_tree - 6.246772285

iii. The third option is **Collect Plot Confidence on Collection Page.** If you want users to be able to tell you how certain they are of their answers to the survey questions, the confidence slider will allow them to assign a value 1-100, where 100 is completely confident of their answers; this value applies to the entire plot. This info will be included in your plot and sample CSV downloads.



iv. The fourth option is **Auto-launch Geo-Dash.** This will automatically open the Geo-Dash interface in a new window or new tab when the data collector navigates to a new plot. Unchecking this option means that data collectors will need to click on the Geo-Dash icon under **External Tools** in the **Data Collection** interface.

5. Click **[Next]** when you are finished.

C. Imagery Selection

In the Imagery Selection pane, you can change the default basemap imagery and the imagery basemaps that are available to users in data collection.

1. You can change the Default Imagery, which specifies the default imagery that users will see when they begin data collection on your project.

i. You may choose any of the imagery options available to your institution.
ii. The default (public) options are MapBox Satellite, Mapbox Satellite w/Labels, and Planet NICFI Public.
iii. The Imagery Preview will display the current selection.
iv. Your users can switch between all the available imagery layers during analysis.
v. PlanetMonthly, PlanetDaily, and SecureWatch do not allow for large area data pulls, so it should not be your default basemap (users will just see a white screen).
vi. You will need to set a different default basemap and have your data collectors switch to PlanetDaily once they have zoomed in on a plot to interpret.

vii. If your project is comparing land use and cover changes between two years, select one of your focal years' WMS imagery as the default imagery here. Your users can then easily switch between this year's imagery and the other year's imagery in data collection.

viii. You will need to first set up the imagery feed for one date period under the institution imagery management panel. Refer to instructions in Part 4, Section B.

> Maxar has deprecated multiple data products previously available in CEO, including: DigitalGlobeRecentIMagery; DigitalGlobeRecentImagery+Streets; DigitalGlobeWMSImagery; and EarthWatch.

2. Public Imagery

   i. This imagery is available for all institutions. If you have a public project, all users (including those not logged in) can see the imagery.

   ii. Click the checkbox next to each imagery source you would like to have available for your project.

3. Private Institution Imagery.

   i. This imagery will only be visible to institution members, even if you have your project set to public.

   ii. Click the checkbox next to each imagery source you would like to have available for your project.

4. Click **[Next]** when you are finished.

D. Plot Design—Background

A well-prepared sample can provide a robust estimate of the parameters of interest for the population (percent forest cover, for example). This is often the most important factor in producing a reliable inventory or accuracy assessment. The goal of a sample is to provide an unbiased estimate of some population measure (e.g. proportion of area), with the smallest variance possible, given constraints including resource availability.

The first step of the built in CEO sample design is the plot design. The second step, selecting samples within a plot, is covered in the next section (Sample (point) design). Using these two steps, many sampling designs are possible with CEO. For information on uploading your own sample as a .csv or .shp file, please see Part 4: Plot & Sample Design--CSV & SHP.

For example, suppose you are quantifying forest cover on a landscape. For this approach, sample points are used to classify land cover and are then summarized at the plot level to create an estimate of the plot's forest cover percent. Information about plots can then be used to estimate the forest cover at the landscape level and detect patterns or trends. The accuracy of your landscape level estimates will depend on how many plots you classify & how variable the landscape is, among other factors. More detailed theoretical information is available in CEO's Project Development Manual (found at https://collect.earth/downloads/CEO_Theoretical_Manual.pdf).

Also, some terminology might help for the two types of sampling available in CEO. **Simple random sampling** means that all points have an equal probability of being selected. It produces unbiased parameters. However geographic balance with small sample sizes can be difficult. Also, rare classes may not receive sufficient coverage. If you have rare classes you want to detect, we highly recommend using a stratified sampling approach. For this, you would need to create a stratified sample in SEPAL (available online at sepal.io) or using QGIS or ArcGIS and import it using the upload csv or shp (see Part 4: Part 4: Plot & Sample Design--CSV & SHP).

The second type is **systematic gridded sampling**. This is a grid of points placed over the landscape at regular intervals. This provides excellent geographic balance, but it is not possible to calculate unbiased estimates of population metric variance.

There are two main approaches for specifying an area of interest (AOI) and sampling design: 1. using CEO's built-in system and 2. Creating a sample in another program (QGIS, ArcGIS, etc.) and importing it into CEO. We will first discuss the built-in sampling design, and then discuss uploading your own sample as a .csv or .shp file.

E. Plot Design—CEO's built in system

CEO's built in system enables users to create sampling designs using an easy-to-use interface. There are two key parts, selecting the AOI and Plot Generation.

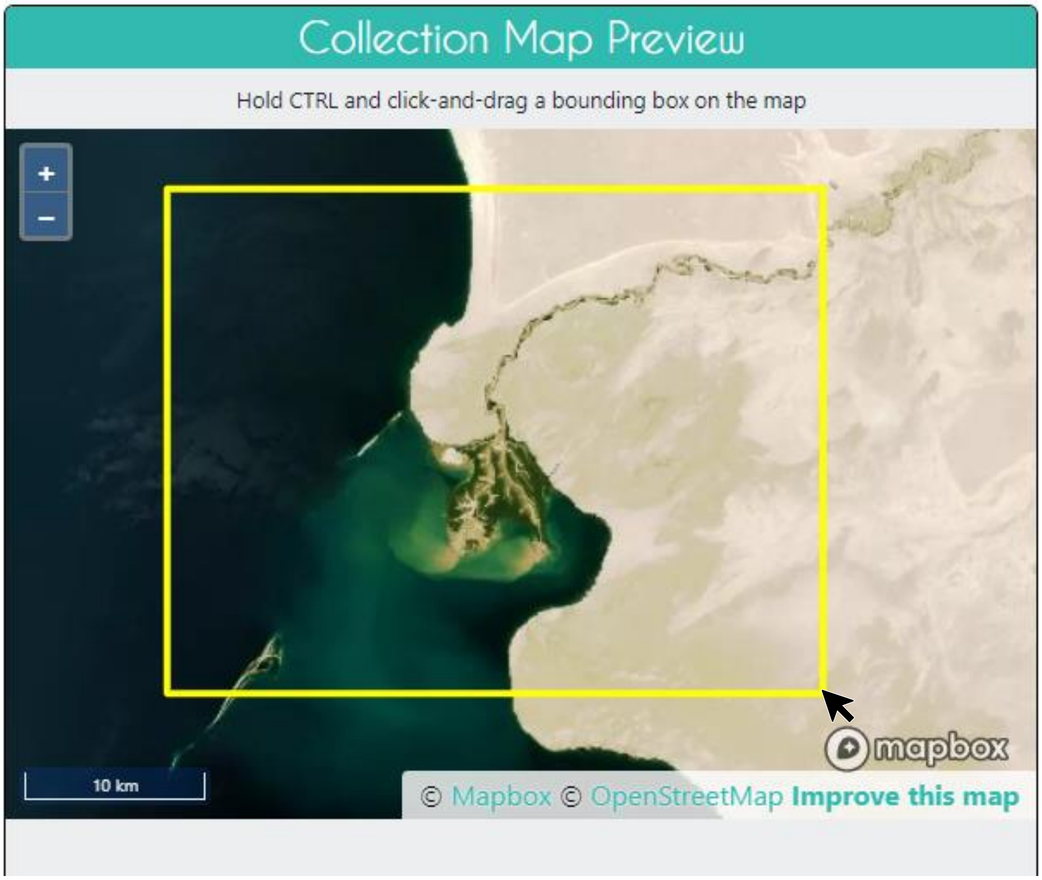
1. Select your AOI. There are three approaches.

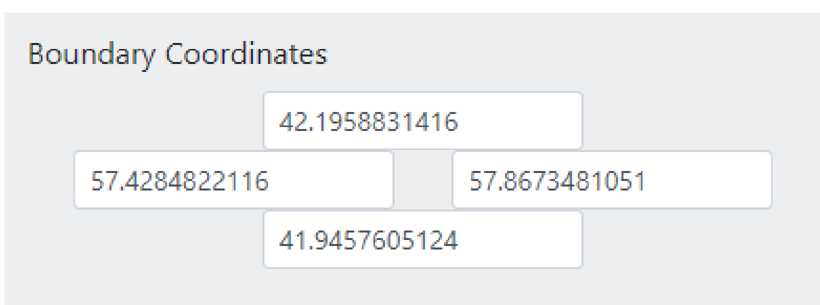
   i. The easiest way to select your project AOI is by drawing a box in the map window in the right hand pane (Collection Map Preview).

      a) Locate your area of interest by zooming in/out using the scroll wheel of your mouse, or the + and – boxes in the map window. You can pan the map by clicking on it and dragging the map window.
      b) Hold the CRTL-key (command key on a Mac) down and draw a box while keeping the left mouse key pressed down.
      c) Hold the SHIFT-key down and draw a box to zoom in.

d) The coordinate boxes will populate once the box is drawn and you let your mouse key go. Coordinates are displayed in lat/long using **WGS84 EPSG:4326**.

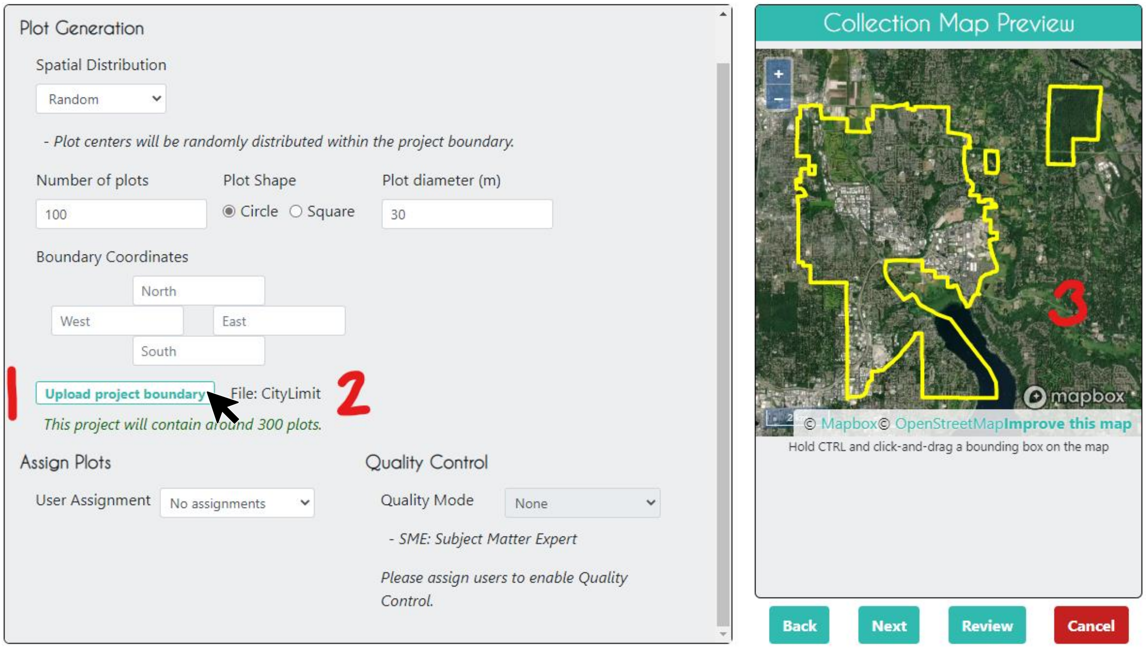
ii. You can also manually enter your Boundary Coordinates into the boxes provided.



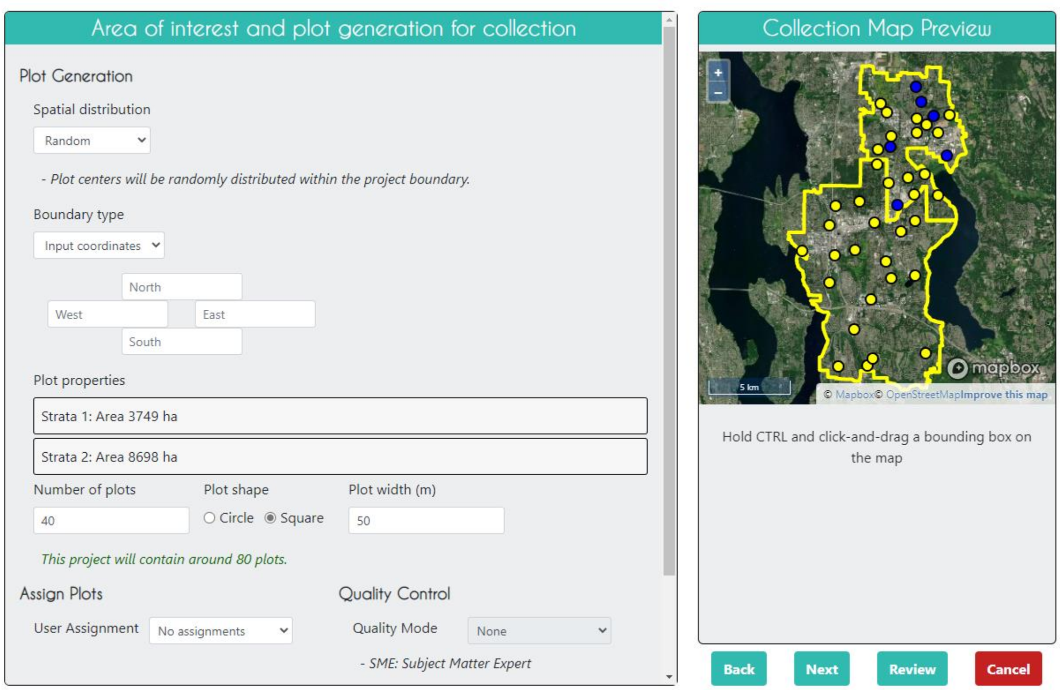


iii. The third approach is to upload a project boundary shapefile. To use this option, click on the **[Upload project boundary]** button. Navigate to your file, and click **[Open]**. Your shapefile should be a zipped folder with the requisite shapefile component pieces (.shp, .prj, etc.). Once you click open, you should see the **File:** text populate with your file name and your project boundary appear in the Collection Map Preview pane.

Please note that if you have multi-part polygons each one will be assigned the number of plots indicated. This is indicated in the green text.



You can upload shapefiles with multiple shapes for stratified sampling. Each strata will appear with its corresponding area in hectares. The number of plots will be **per strata**. This is indicated in the green text.



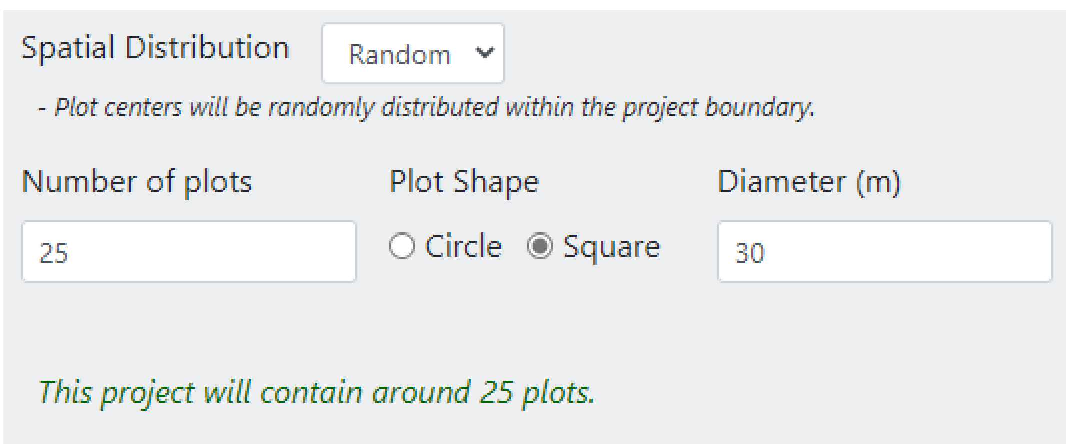
2. In the Plot Generation section, you can specify the type and number of sample plots.

   i. **Spatial Distribution** defines the distribution of the sample points. In CEO, you can specify either a random or a gridded (spatial systematic) sampling approach.

      a) Random sampling has the advantage of being extremely simple and producing unbiased parameters that are straightforward to calculate. However, geographic balance is not certain with smaller sample sizes, and rare classes may not be adequately sampled unless the sample size is large.
      b) Systematic sampling has the advantage of providing excellent geographic balance. However, it is not possible to calculate a truly unbiased estimate of the variance of population metrics when using systematic sampling. Additionally, if patterns in your landscape match up with the spacing of your systematic gridded points, you will produce a very biased estimate.
      c) If you select **Random**, you will need to provide the number of plots for the whole project.
      d) If you select **Gridded**, you will need to provide the spacing between the centers of the plots (in meters).

   ii. CEO will provide an estimate of how many plots will be generated for your project based on your sampling design.

Using CEO's sampling, the maximum number of plots for a project is 5,000. For gridded sampling, you may need to increase the space between plots to avoid exceeding 5,000 plots.



3. Plot Shape can be either a Circle or a Square.

   i. You will need to specify the **Diameter** in meters.
   ii. These sizes should be driven by the needs of your project.

iii. If they are small, your users will need to zoom out significantly to see the relevant background imagery because CEO automatically centers and zooms in to the plot's boundaries.

4. You can now choose to assign users plots to review using the User Assignment feature, and implement quality control for your plots using the **Quality Control** dropdown. See Section H: Plot Design—Quality Control Options for more.
5. Click "Next" when you are finished.

F. Sample Design—CEO's built-in system

Here we determine how many sample points are within each plot, and whether they are sampled using random sampling or gridded sampling.

1. Under Spatial Distribution:

   i. With **Random sampling** sample points will be randomly distributed within the plot boundary. You will also need to specify the **Number of Samples** per plot.
   ii. With **Gridded sampling**, sample points will be arranged on a grid within the plot boundary. You will need to specify the distance between points within the plot under **Sample resolution** (m).
   iii. With **Center** a sample point will be placed in the center of the plot; you do not need to specify anything else.
   iv. With **None,** you will not predefine any samples. This requires users to draw their own samples during collection.

2. For any of these Spatial Distributions, you can click the checkbox next to Allow users to draw their own samples to enable proactive sampling.

Proactive sampling enables data collectors to draw points, lines, and polygons directly onto the map to create their own samples. The data collector then answers questions about each shape.

Proactive sampling is useful for collecting training data to inform random forest and machine learning models. It can also increase the accuracy of land use land cover area estimates by allowing users to map the entire area of the plot instead of sample points within the plot.

Using CEO's sampling, the maximum number of sample points per plot is 200.

Using CEO's sampling, the maximum total number of sample points for the project (number of plots times the number of points/plot) is 50,000.

If you need more plots or samples, please create your sampling design in another program and upload it to CEO using the .csv or .shp file format and directions in the next section.

G. Plot & Sample Design—CSV & SHP files

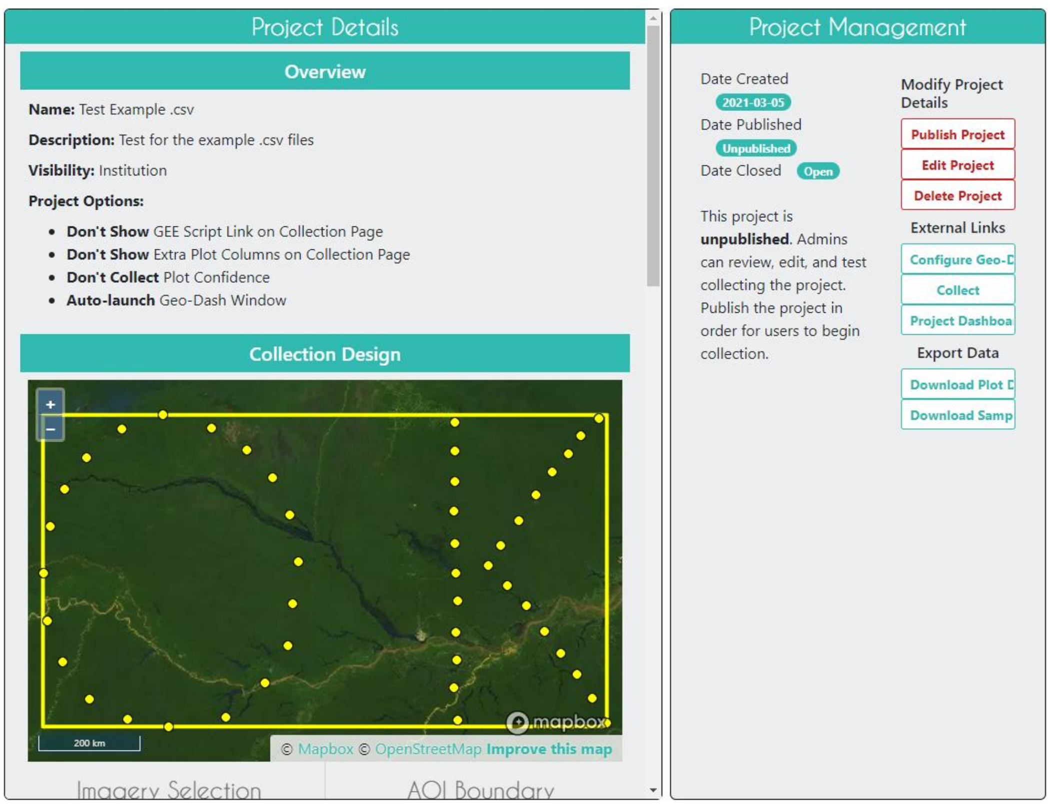
While the default sampling design will work for many users, you may want to create your own sampling design and upload it to CEO using the .csv or .shp file capability. For example, this functionality is useful when: you want to draw your sample plots from within a shape other than a rectangle (e.g. the outline of a region or country) or if you want to do any stratification in your sampling design. You can create a .csv or .shp with your desired sampling plots/points through services and applications including Sepal, ArcGIS (e.g. https://pro.arcgis.com/en/pro-app/tool-reference/data-management/create-random-points.htm), and QGIS (e.g. https://freegistutorial.com/how-to-create-random-points-inside-polygon-on-qgis/).

As when using CEO's built-in system, you can now choose to assign users plots to review using the **User Assignment** feature and implement quality control for your plots using the **Quality Control** dropdown. See Plot Design—Quality Control Options for more.

Using .csv and .shp files, the maximum number of plots is 50,000 and the total sample point limit is 350,000.

You must use WGS84 EPSG:4326 format for coordinates in both .csv and .shp files.

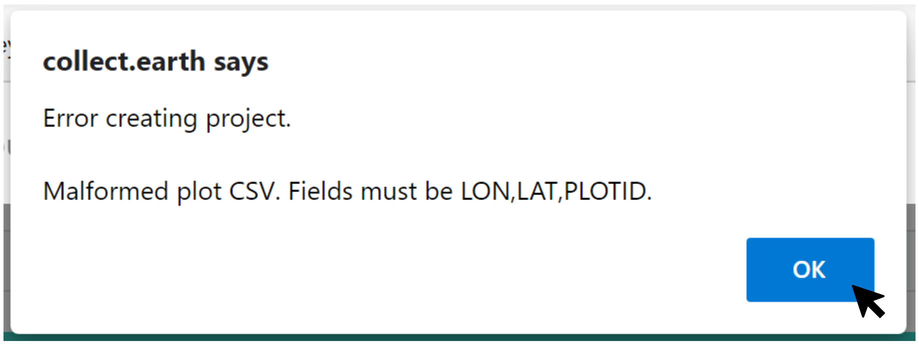
Example CSV and SHP files are available for download. When used, the plots should spell out "OK" when used to create a project.



For .csv files, specify plot centers by uploading a .csv with these columns: LON, LAT, PLOTID.

You can also upload a second file where you specify your own sample point centers by uploading a .csv with these columns: LON, LAT, PLOTID, SAMPLEID. LON and LAT can also be LONGITUDE and LATITUDE. You can have additional columns with data about your sample plots and points but they MUST come after these key fields.

If you do not specify the column names correctly (spelling or order), you will get the following error:

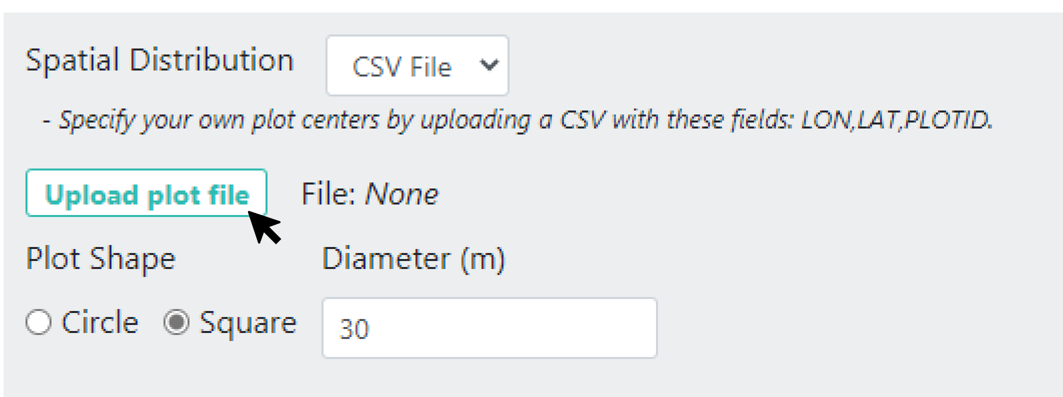


Longitude should be between -180 and 180, while latitude should be between -90 and 90. If you mix them up, you may get an error if your longitude is greater than 90 or less than -90 (when this is mixed up with latitude, it is 'above' the pole). Double check these values.

You can upload just one file for the plot centers OR two files, one for the plot centers and one for the point centers. When your .csv files are fit the above specifications, follow the directions below.

For one file with just the plot centers:

1. Under Plot Generation, select **CSV File**.
2. Then, click on **[Upload plot file]** and navigate to the .csv on your computer with your sample plot centers.

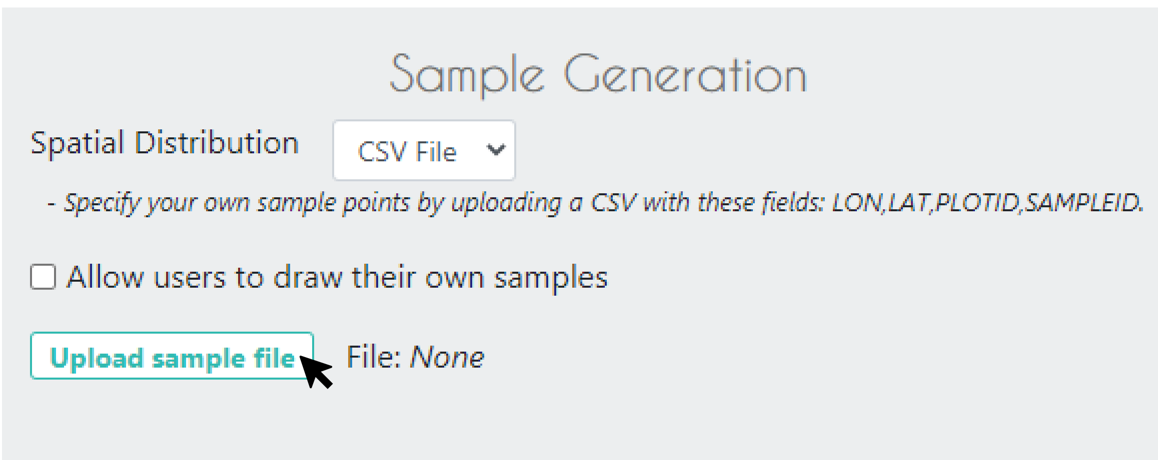


3. After you upload the file, the file name will appear next to **File:**.
4. You will need to specify the **Plot Shape** and corresponding **Diameter**.

5. You will also need to specify your **Sample (Point) Design**, as in Part 4: F Sample Plot Design above.

For two .csv files, one with plot centers and one with sample point centers,

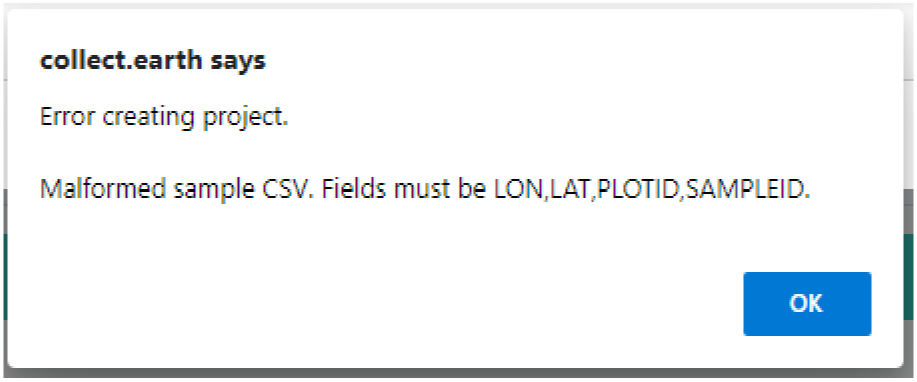
1. Follow steps 1-4 above for the Plot Design.
2. In step 4, it is important that you specify a plot size that is large enough to contain your points if they are also uploaded through a .csv or .shp. This will not be a problem if you use the built-in sample point design function.
3. Then, under Sample Design, set Spatial Distribution to CSV File.
4. Click on **[Upload sample file]** and navigate to your .csv on your computer. Click Open and the file name will appear next to **File.**
5. Note that you can also choose SHP file and upload a shape file at this point..
6. You can click the checkbox next to Allow users to draw their own samples to enable proactive sampling.

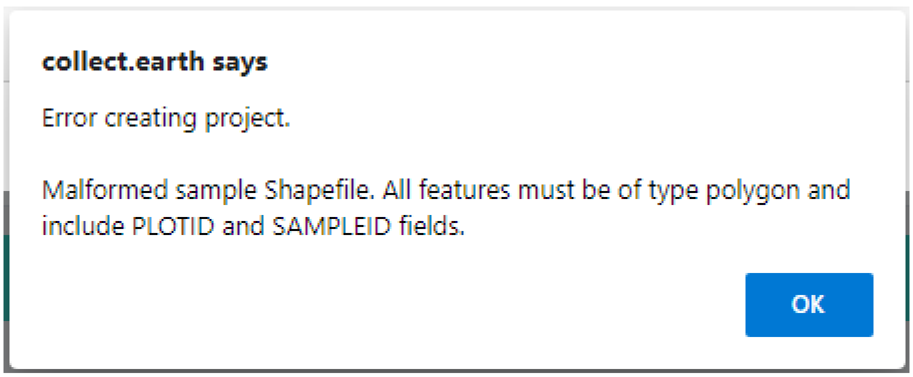


For .shp files, you can specify your own plot boundaries by uploading a zipped Shapefile (containing SHP, SHX, DBF, and PRJ files) of polygon features. Each feature must have a unique PLOTID field. LON and LAT are not required for polygons.

You will also need to upload a second file where you specify your own sample points. This can be a .csv with these columns: LON, LAT, PLOTID, SAMPLEID. LON and LAT can also be LONGITUDE and LATITUDE. It can also be a zipped .shp file (containing SHP, SHX, DBF, and PRJ files). Each feature must have PLOTID and SAMPLEID fields. Either points or polygons will work for the sample point file, though point files must include LON and LAT. As with .csv files, you can have additional fields with information about your plots and points if and only if they come after these key fields.

If you do not specify your PLOTID in the sample point .csv or .shp zip file, you will get the following errors:



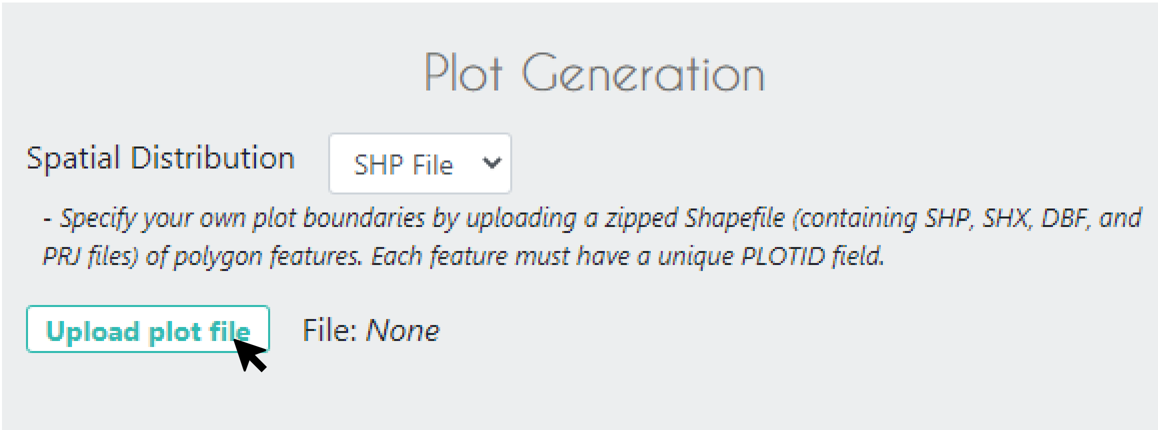


You can zip your files easily in Windows by selecting the relevant files, right clicking on one, and the clicking **[Send to] -> [Compressed (zipped) folder]**.

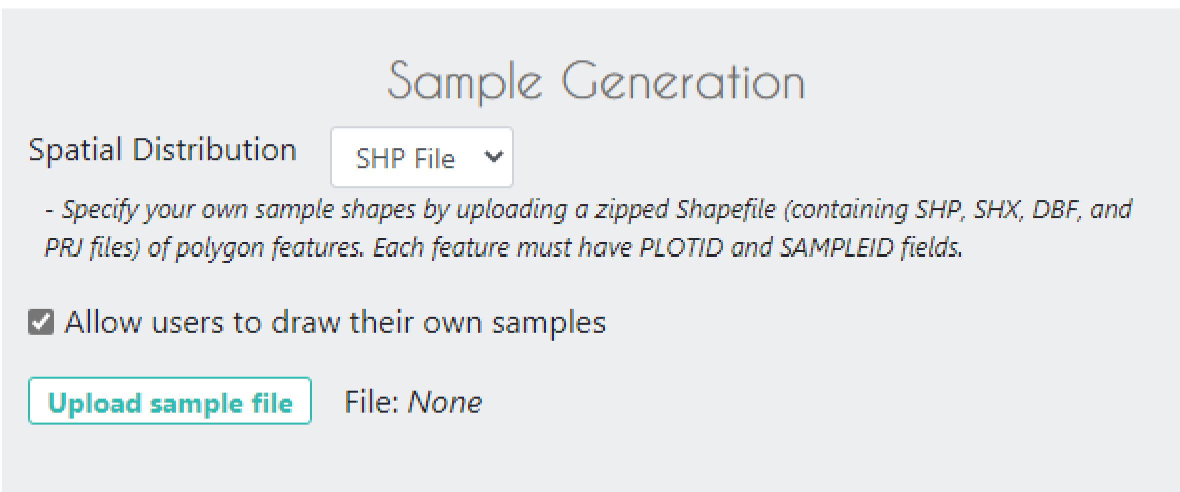
For a Mac, select the relevant files, right click on one of the files, and select **[Compress Items]** from the pop-up menu.

When your .shp files fit the above specifications, follow the directions below.

1. Under Plot Generation, set **Spatial Distribution** to **SHP File**. You must have the radio point selected before the button to upload becomes available.
2. Then, click on **[Upload plot file]** and navigate to your zipped .shp file.



3. Click [Next]. Under Sample Design, set Spatial Distribution to SHP File.
4. Click on [Upload sample file] and navigate to the zipped .shp file with your sample point points or polygons and select it.
5. This will also work with CSV File.



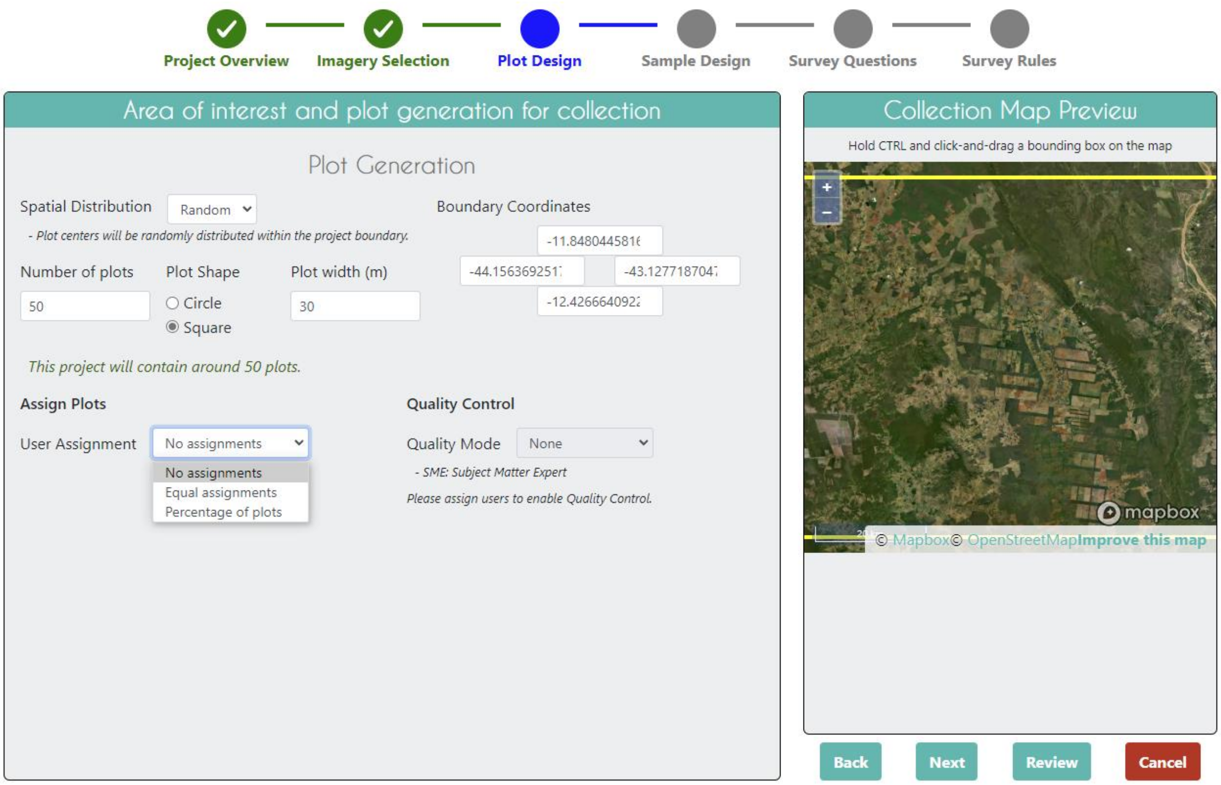
When you download your collected data, any column with extra information that were present in the uploaded .csv and .shp files will be preserved in the downloaded data (See Part 6: B). These columns can also be displayed in the Data Collection pane (see Part B Project Overview).

H. Plot Design—Quality Control Options

We have implemented automated quality assurance and quality control (QAQC) tools for projects created in CEO.

1. You can assign members of your institution to plots using the user assignment feature.
2. There are two options, Equal assignments, and Percentage of plots.

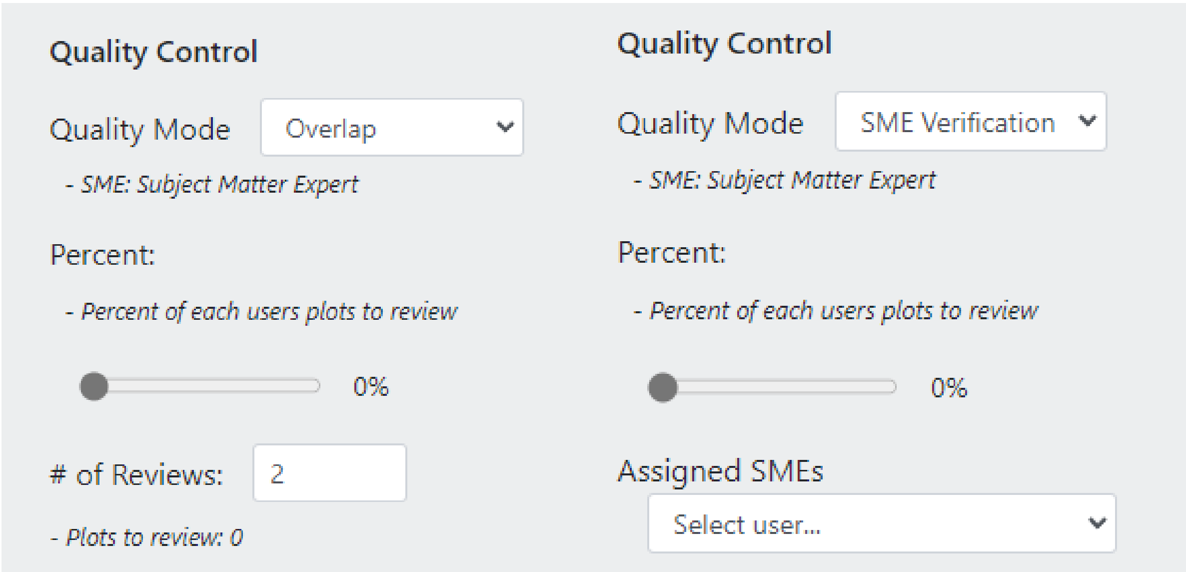
   i. **Equal assignments** splits the total number of plots evenly between users.
   ii. **Percentage of plots** allows for uneven splits. For example, one user can be assigned 50% of plots while two other users can be assigned 25% of plots respectively.



3. You can also implement quality control measures using one of three different Quality Modes. This allows multiple users to collect data for the same plot for QA/QC purposes.

   i. **None**: Quality control is not enabled by default.
   ii. **Overlap**: A certain percentage of each users' plots are reviewed by the other users.
   iii. **Subject Matter Expert Verification (SME Verification)**: A percentage of each users' plots are reviewed by one or more SMEs. For example, the SME might be someone with deep knowledge of the local landscape, or a project administrator.

4. Note that when Quality Control is enabled, the project can no longer support User Drawn samples.



5. When you implement the Quality Control, you will be able to use the QAQC Administrator Review Mode. See Part 6: Administrator Review Mode for further information.

I. Survey Questions

This is where you design the questions that your data collectors/photo interpreters will answer for each of your survey plots. Each question creates a column of data. This raw data facilitates calculating key metrics and indicators and contributes to fulfilling your project goals.

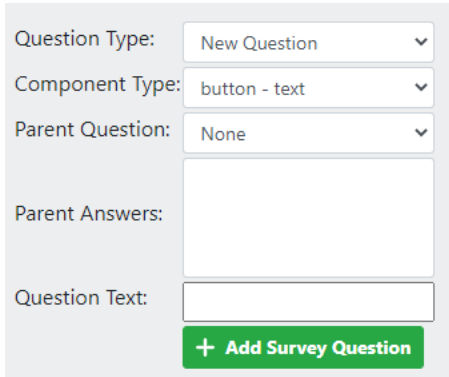
Survey Cards are the basic unit of organization. Each survey card creates a page of questions on the Data Collection interface. The basic workflow is: Create new top-level question (new survey card) → populate answers → create any child questions & answers → move to next top-level question (new survey card) & repeat until all questions have been asked.

For the **Survey Question** tab, the left panel allows you to enter questions while the right panel provides a preview of how these questions will appear to your data collectors.

We will now go into more detail about how to add a question and answers, the types of questions that can be asked and when these questions might be useful.

1. How to add questions and answers

   i. CEO provides a straightforward way to ask **multiple-choice** questions. As the most common question type used, we will use it for this example. In CEO, these questions are called “button—text” questions as in data collection they display as a button with text.



Questions of this type are useful for land use and land cover surveys, or anywhere where you want the user to choose between a limited set of mutually exclusive options.

ii. To start, type your first question into the **Question Text** box. Since it is your first question, you cannot assign a **Parent Question** or **Parent Answers** for the question.

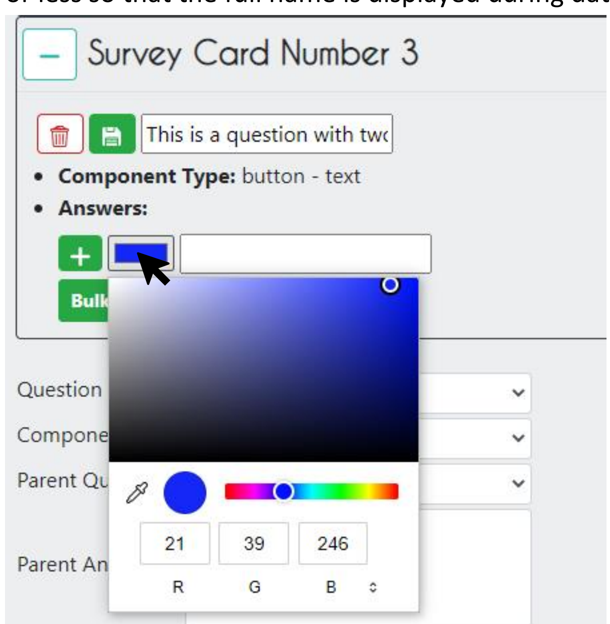
iii. Try to keep the question text below 45 characters so the whole question will be displayed during data collection.

iv. Click [+**Add Survey Question**] to create your first **Survey Card**.

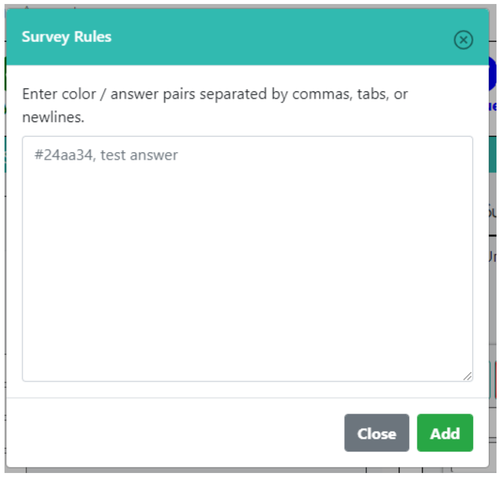
v. You can now add Answers to your question. Answers have two parts: a color and a text box.

a) Click on the **[blue rectangle]** to bring up the **Color Selector.** You can move the color selector dot, use the eyedropper tool, or type in RGB values (0-255). Click off of the color selector when you are done. The color you choose will be associated with the answer. When a data collector selects that answer, any sample points assigned that answer would also be assigned that color to display on the map.

b) You can type your answer into the **Text Box**. Try to type answers with around 15 characters or less so that the full name is displayed during data collection.



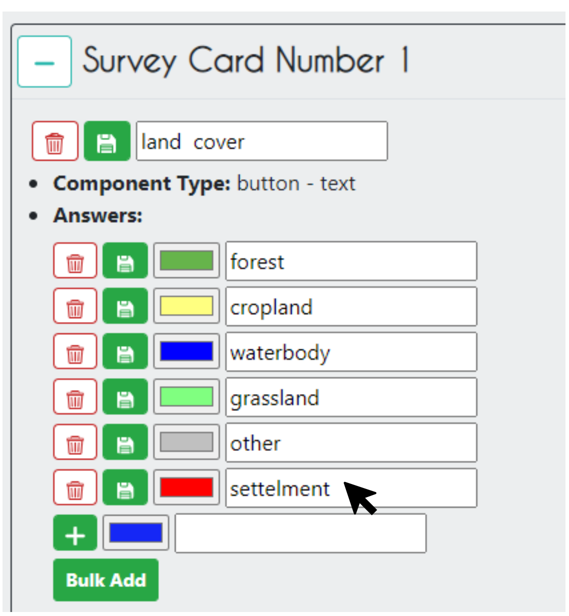
c) You can also use the **[Bulk Add]** button to add multiple color and answer pairs.

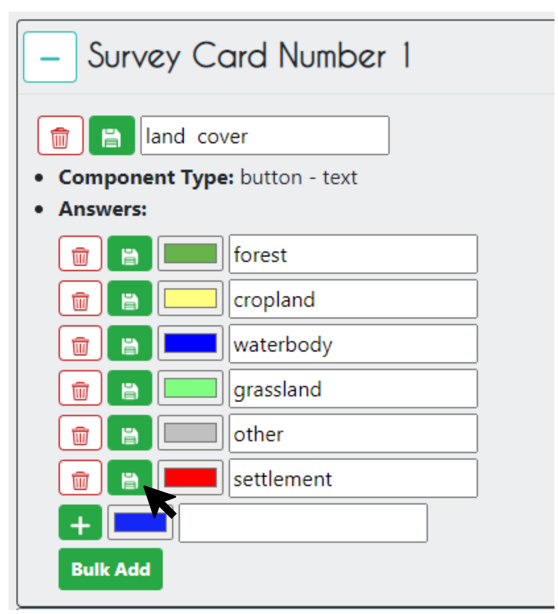


vi. Click the green [+] symbol to add the answer.
vii. Continue adding answers until all the answers to your survey question have been added.
viii. Now that you have a top-level, or parent, question with answers, you can edit and delete questions and answers, and add child questions and child questions that appear only when specific answers are chosen.

2. You can delete or edit your questions and answers after you create them.

i. To delete a question or a specific answer, click on the **Trashcan** icon next to the question or answer you would like to delete.
ii. To edit a question or an answer, first type your edits into the box, then click the **Save** icon next to the edited question or answer.



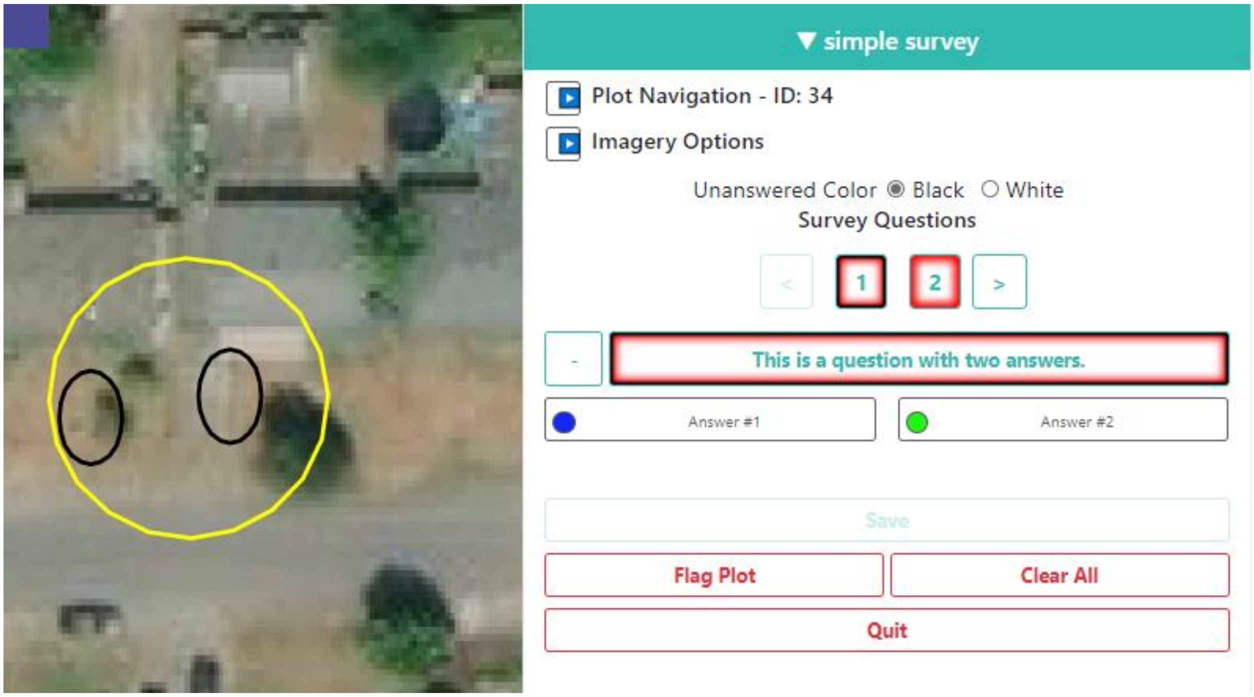


Parent & child questions, particularly child questions that have parent answers, are useful when you have broad categories and then want to refine the answer within that category.

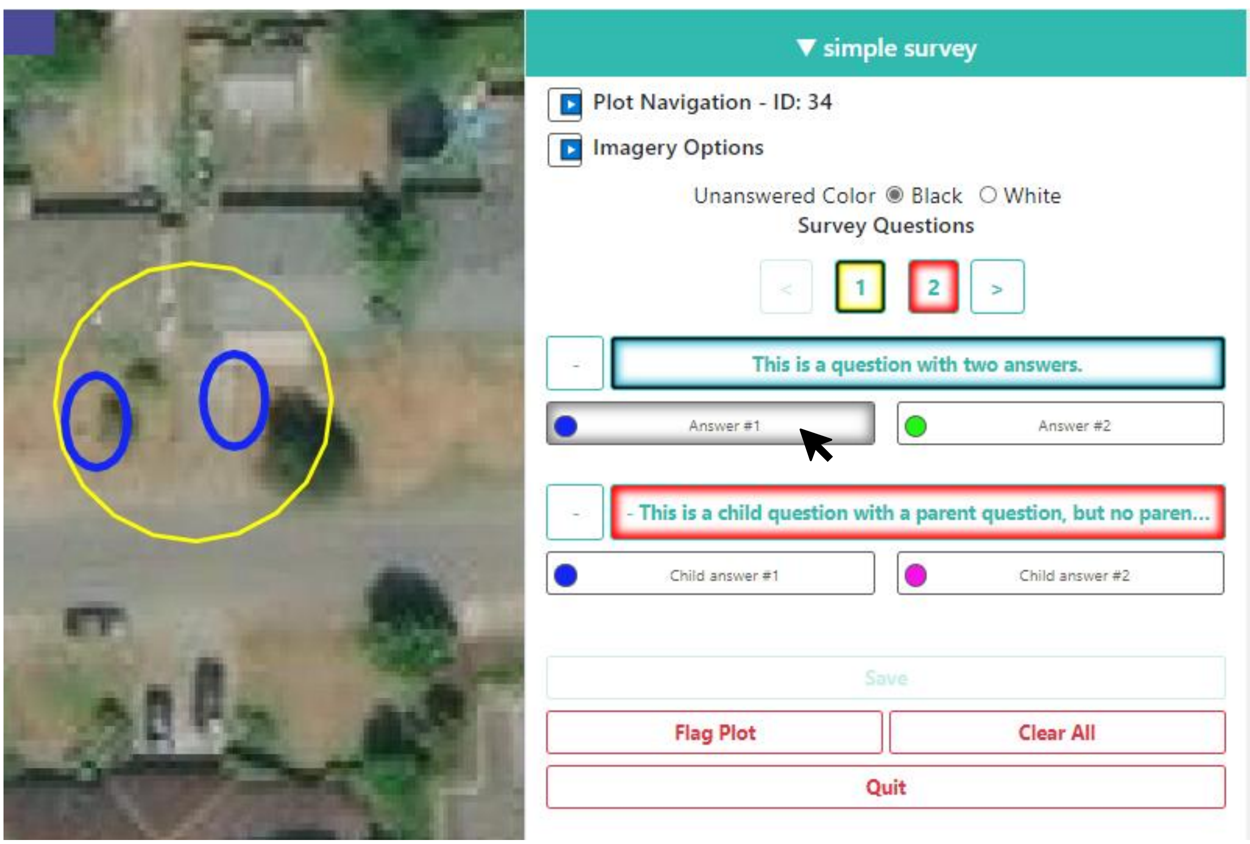
3. You can ask follow-up questions based on a user's response to further refine information about the plot. For example, if a user categorizes a plot as forest, you can follow up by asking them if it is deciduous or coniferous forest.

   i. To create a child question, next to **Parent Question** select the question you want using the drop-down menu.
   ii. You can then assign one or more **Parent Answers** through the dropdown menu. Hold down the ctrl button to select multiple answers. When one of the parent answers is chosen, the child question will appear.
   iii. If you want to have the child question appear regardless of the answer, select all of the answers for **Parent Answer**.
   iv. Once you have finished adding child questions with their answers, you can create a second survey card by setting the **Parent Question** field to **None**.
   v. You can delete a question by clicking the **Trashcan** icon. Deleting a parent question with children will delete the children questions as well.

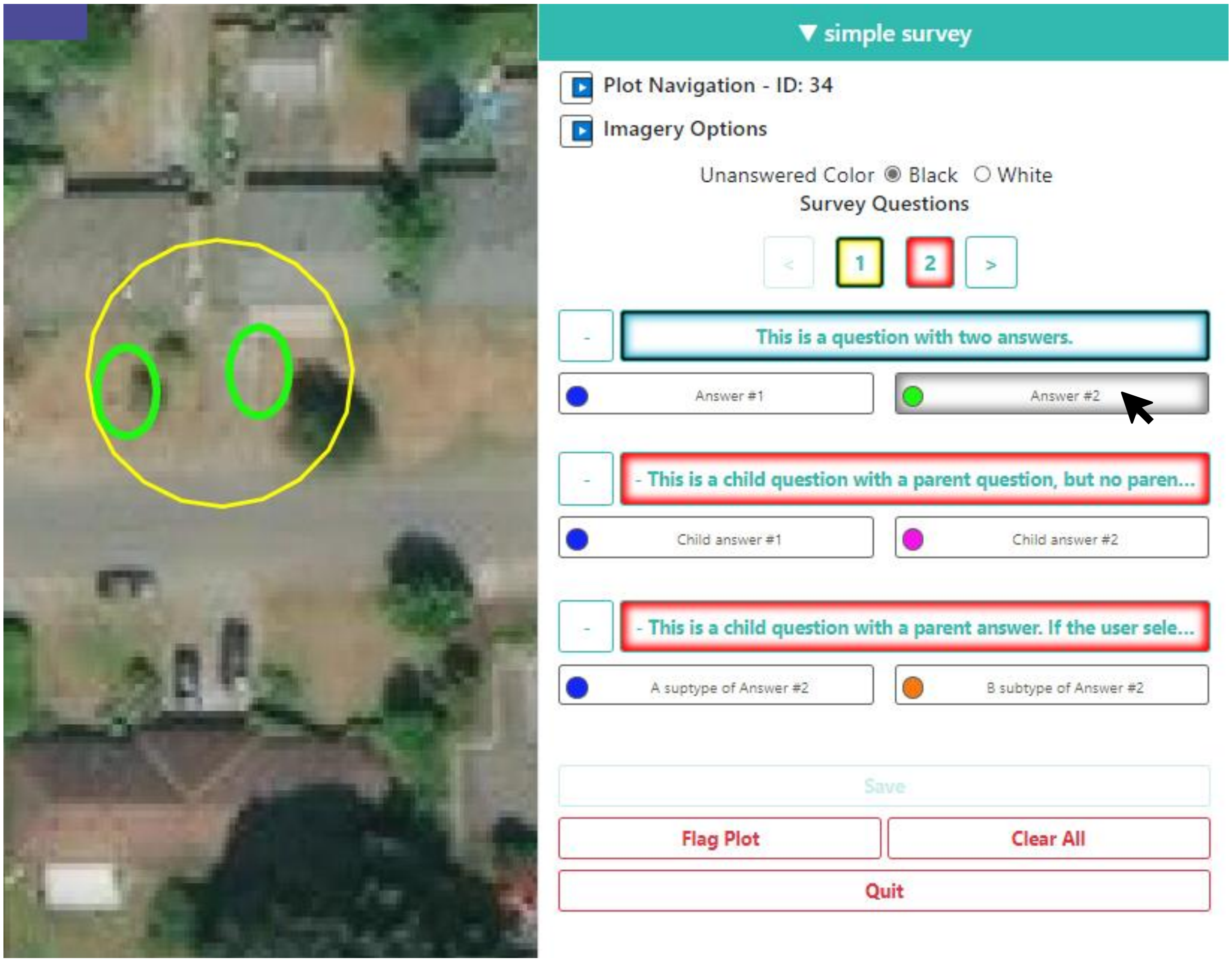
And here is how a survey card with parent/child questions appears on the data collection page initially:



And after selecting “Answer #1”:



And after selecting “Answer #2”:



The child question with no parent answer will appear when either answer is selected. The child question with “Answer #2” as the parent answer only appears when “Answer #2” is selected.

You can also see that these questions were too long, and the user cannot read the whole question. Try to keep your question below 45 characters.

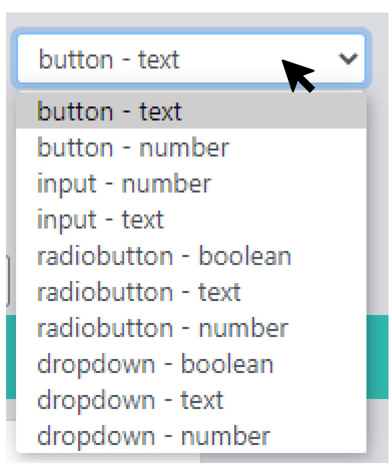
You can create extensive lists of land cover classification options for data collection using this approach.

4. Manipulating the survey cards.

   i. You can collapse a survey card by clicking the [-] symbol in the upper left.
   ii. You can change the order of the Survey Cards by clicking the blue up & down arrows in the top right.

5. Here are the other types of questions that can be asked and when these questions might be useful:

   i. Now we will explore the other **Component Type** options.



   ii. These **Component Types** include combinations of four question types and three data types.

The four question types are:

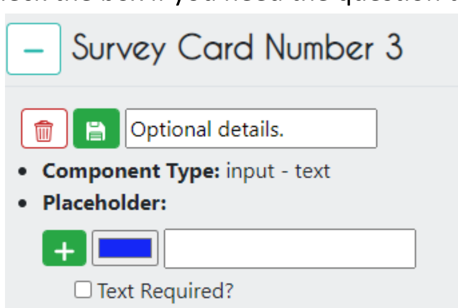
- **Button**: This creates clickable buttons, allowing users to select one out of many answers for each sample point.
- **Input:** Allows users to enter answers in the box provided. The answer text provided by the project creator becomes the default answer.
- **Radiobutton:** This creates radio buttons, allowing users to select one out of many answers for each sample point.
- **Dropdown:** Allows users to select from a list of answers.

The three data types allowed are:

- **Boolean:** Use this when you have two options for a question (yes/no).
- **Text:** Use this when you have multiple options which are text strings. They may include letters, numbers, or symbols.
- **Number:** Use this when you have multiple options that are numbers, which do not contain letters or symbols.

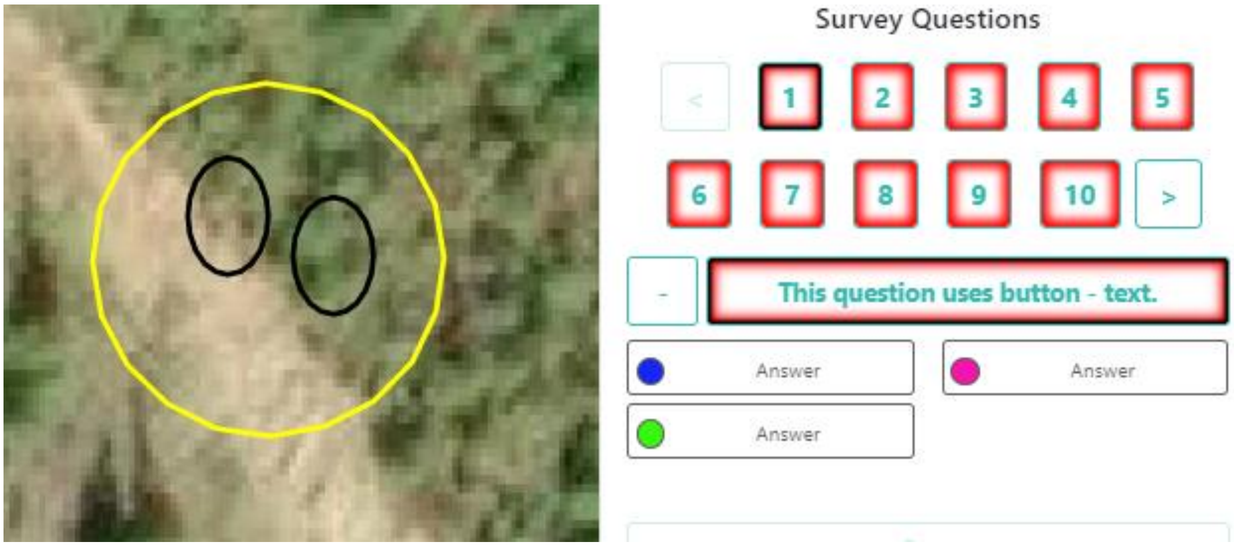
iii. For the Input type only, you can choose whether to require an answer during data collection.

a) Leave this box unchecked if you don't want users to always collect the information. This is useful for adding optional details about a plot.
b) Be sure to check the box if you need the question to be answered.

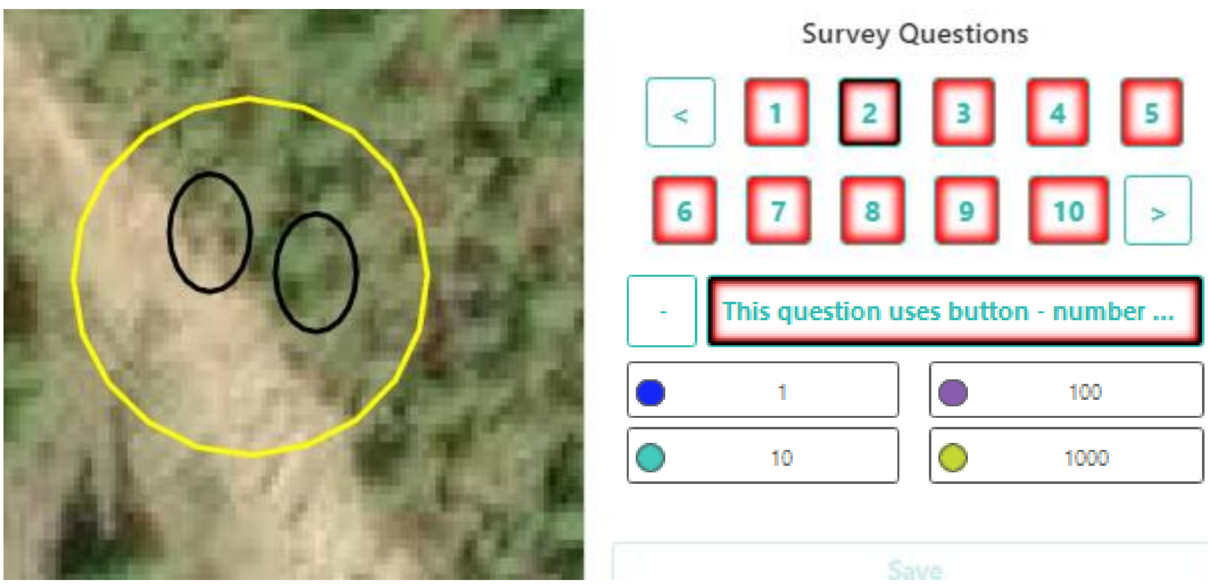


Following are examples of how each question type listed under **Component** Type appears in the **Data Collection** screen, and notes on when each type might be useful.

Button—text:



Button—number:

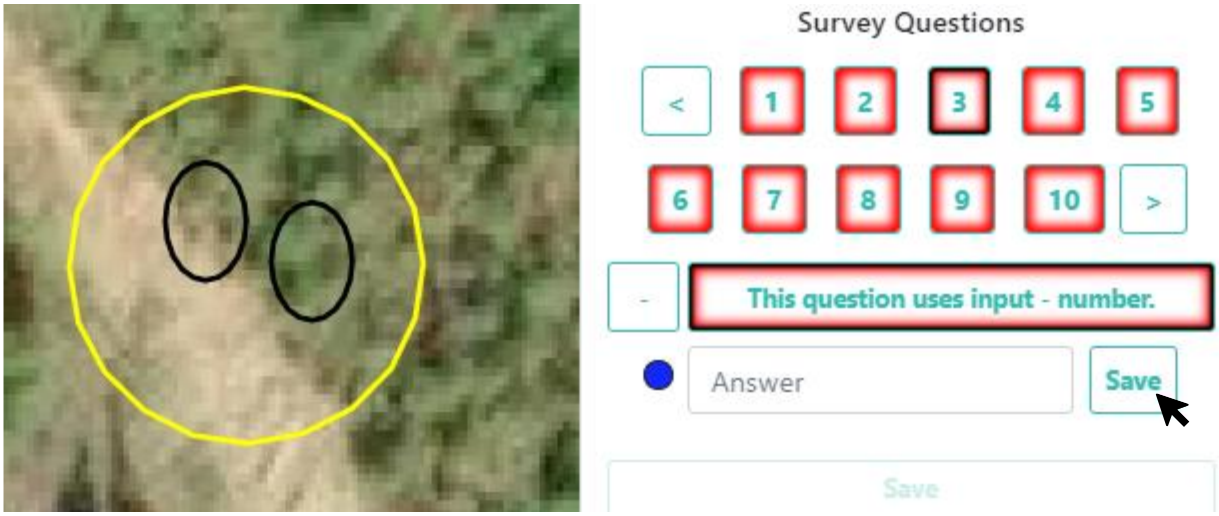


Button—text and **Button—number** are useful when you want the user to classify each sample point as one of many different options. Different sample points can be assigned different answers. These are a great option for land use and land cover questions.

Try to use answer names with 15 characters or less so that the full name is displayed during analysis.

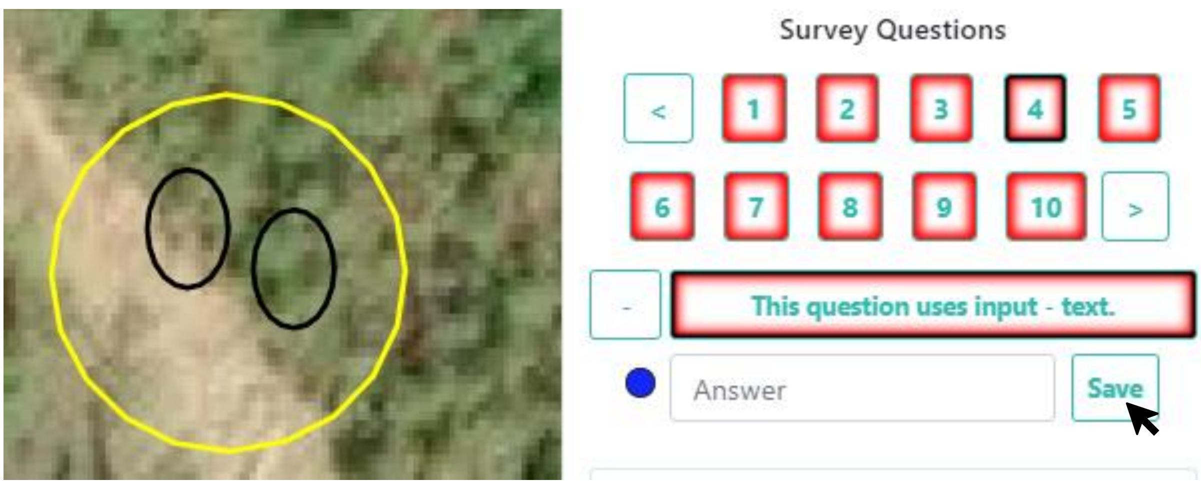
The **Button—number** option, coupled with numeric answers, allows you to implement numeric-only rules that will not work on **Button—text** elements.

Input—number:



Users can input integers, decimals, negative and positive numbers, and the letter “e” for scientific notation. Decimals must use “.” and not “,”. What the survey creator types in the “Answer” field will become the default text in the input box. Note the user must click “Save” to input the answer.

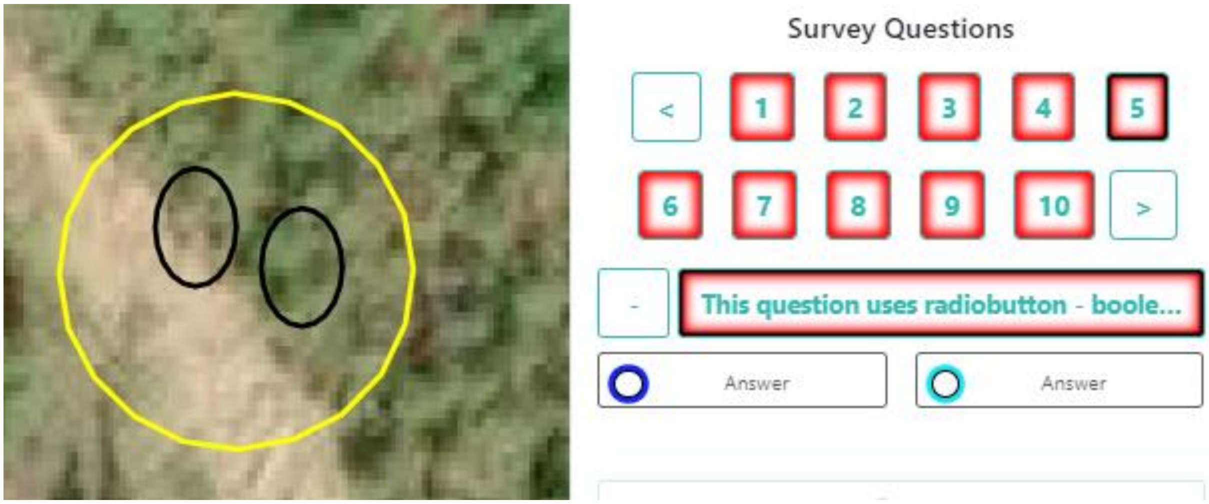
Input—text:



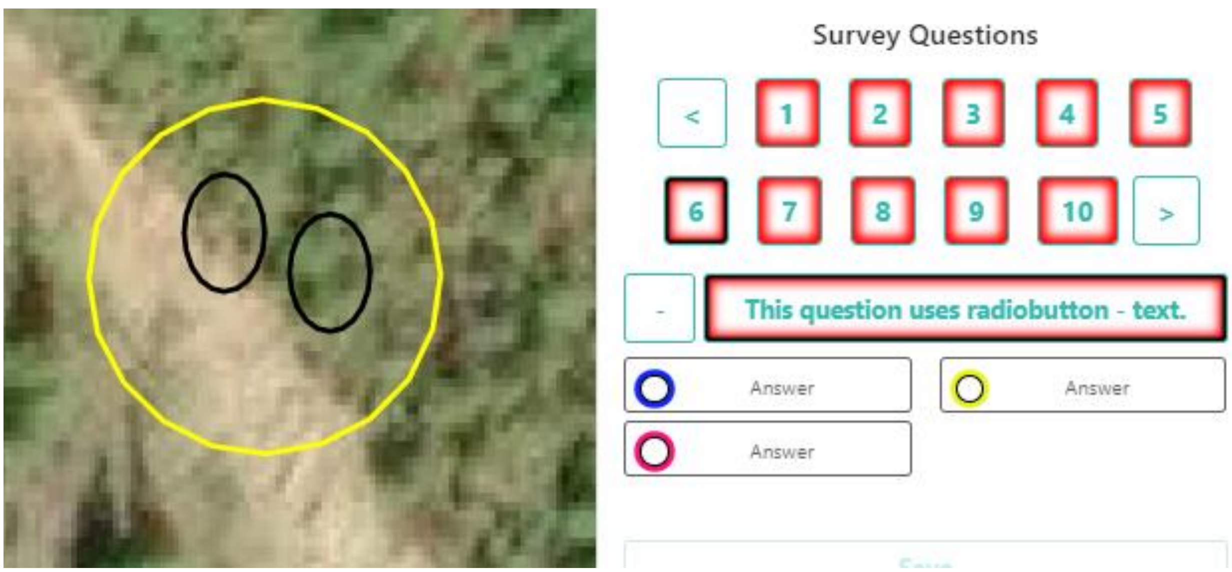
Users can input any character. What the survey creator types in the “Answer” field will become the default text in the input box.

Input—number and **Input—text** are useful when you want the user to provide custom input for each plot. For example, you might ask local participants to identify agricultural crops. Answers can be long (500+ characters).

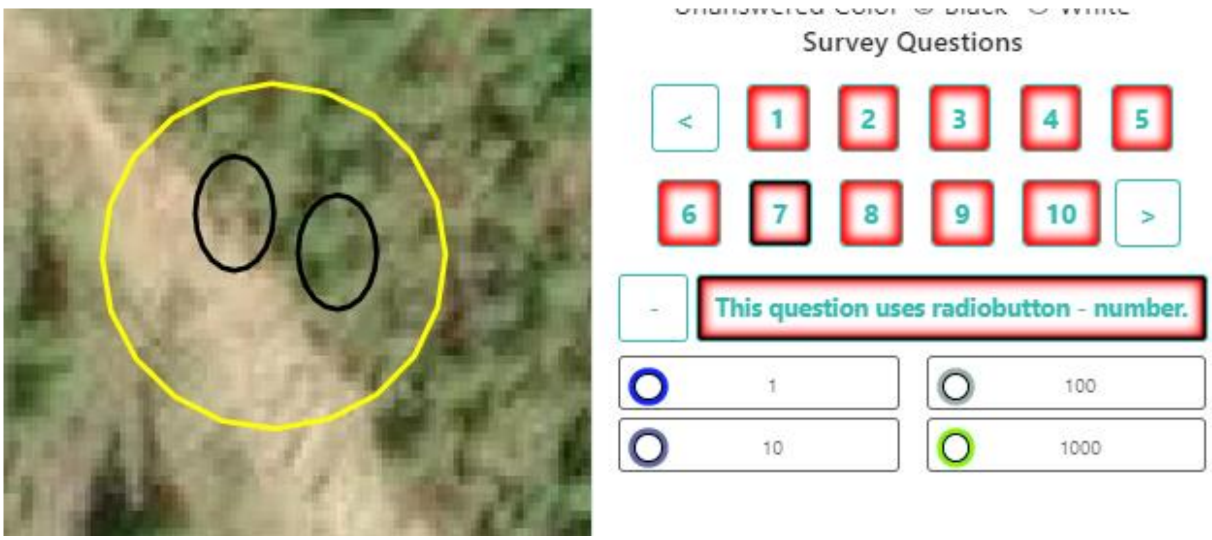
Radiobutton—boolean:



Radiobutton—text:



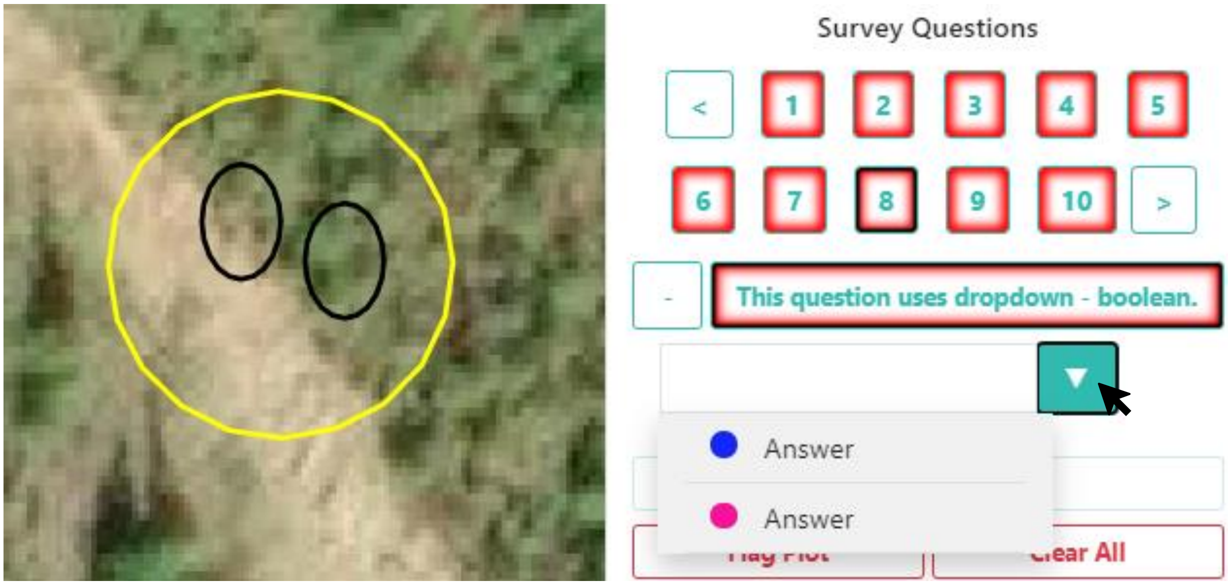
Radiobutton—number:



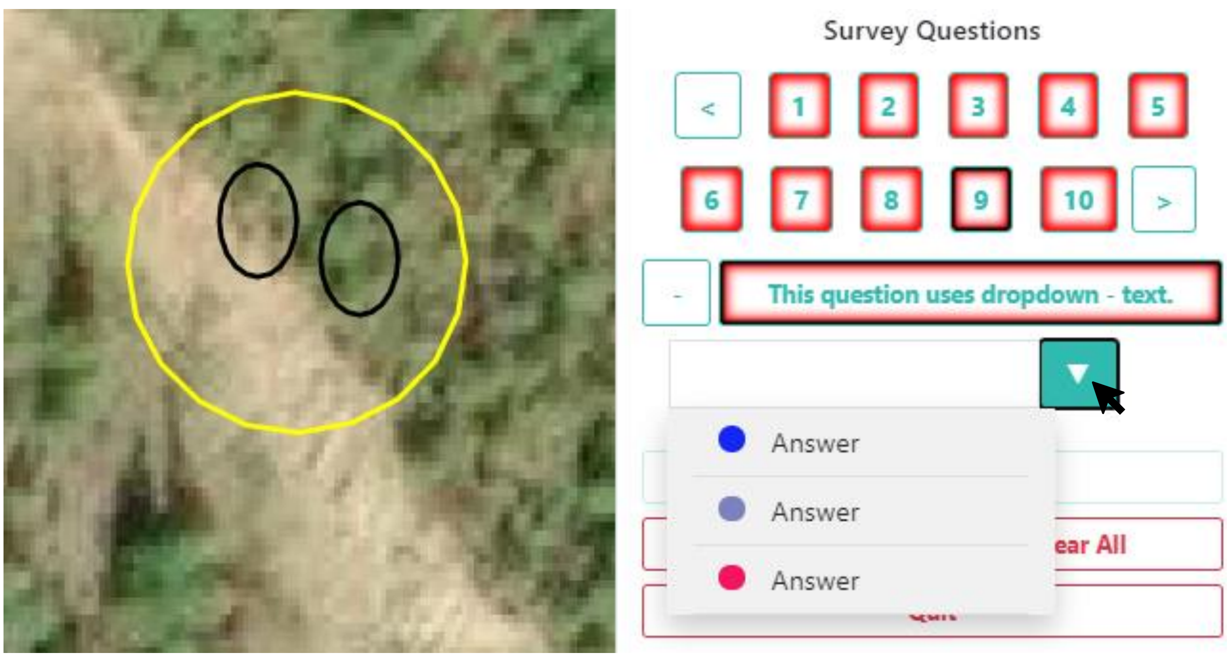
Radiobuttons are functionally much the same as Buttons, but with a different aesthetic and the added Boolean option. You can choose different answers for each sample point. Try to use names with around 15 characters or less so that the full name is displayed during analysis.

Radiobuttons are useful when you want your users to choose one answer for each of your sample points from multiple options. They work well for land use and land cover questions. The Boolean option also works well for areas you want to classify as yes/no, e.g. Forested or Not Forested.

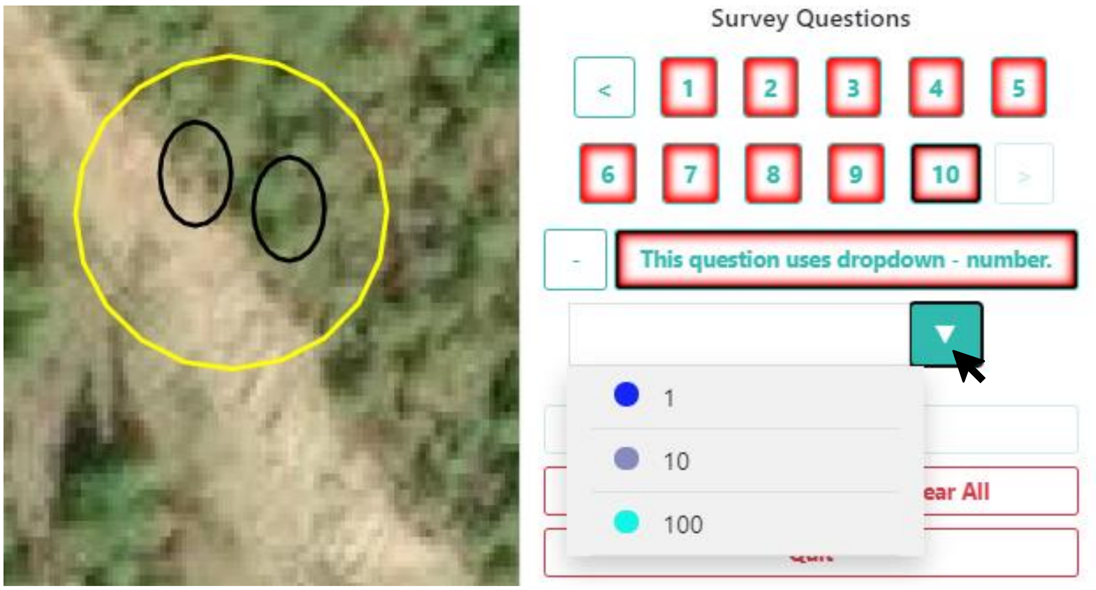
Dropdown—boolean:



Dropdown—text:



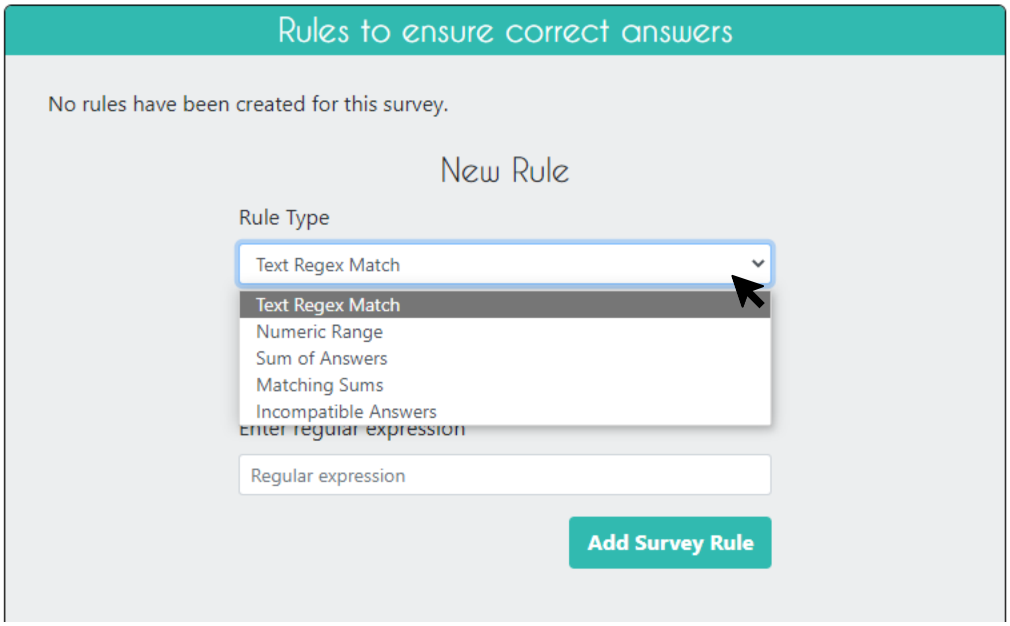
Dropdown—number:



The **Dropdown** questions function similarly to the **Button** and **Radiobutton** options, but with the dropdown menu aesthetic instead of the button aesthetic. As with the other options, you can assign each sample point a different answer, though it is more difficult from a user perspective. This option may encourage assigning only one answer to all the sample points. Overall, choose the option that will be easiest for your users to understand.

J. Survey Rules

Survey rules help ensure users collect logical and correct answers. If you do not want to add any rules, you can just click **Next**.



Rule Types include:

- **Text Regex Match**: This rule applies only to **Input—text** questions & their answers. It allows you to verify if the entered value fits, using regular expressions.
  However, unless you have a specific reason to use the **Input—text** question type, consider using **Button—text** or **Radiobutton—text** instead. These options are easier for users and will always provide exact text.
  This rule uses the JavaScript RegExp function, documentation for writing a regular expression can be found here: https://developer.mozilla.org/en-US/docs/Web/JavaScript/Guide/Regular_Expressions.

- **Numeric Range:** This rule applies to **Input—number** questions & their answers. With this rule, you can verify that the numeric input falls within a predefined range.

For example, if you are asking about the proportion of points in the plot that contain trees, you could constrain the answers between 0 and 1.

- **Sum of Answers:** This rule applies to any **number** type questions & their answers. You select multiple questions (2 or more) and specify what the questions should sum to.

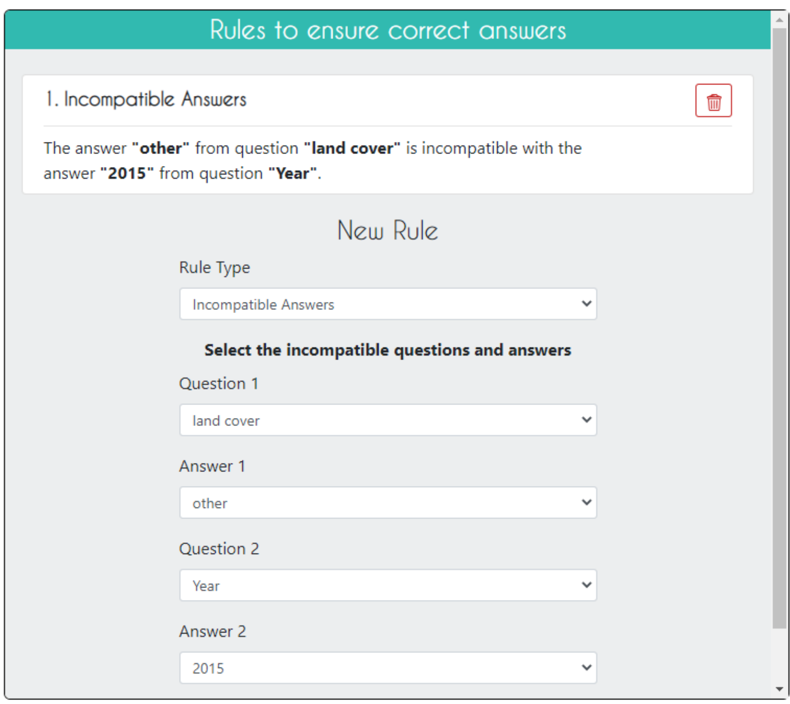
For example, this is helpful if you have multiple questions asking about percent of land cover, where the sum should be 100%.

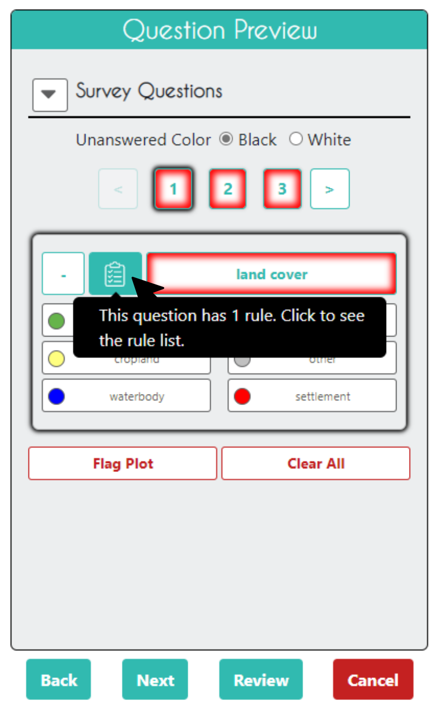
- **Matching sums:** This rule applies to any **number** type questions & their answers. With this rule, you specify two sets of multiple questions (2 or more) that should have equal sums.

- **Incompatible answers:** This rule can apply to any type of question. It allows the user to define incompatible sets of answers.

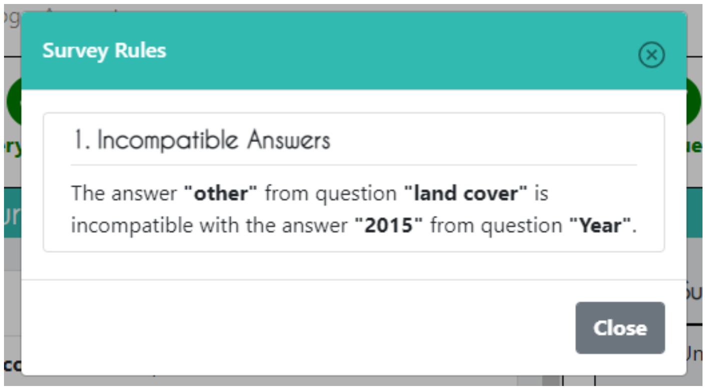
For example, if the answer to one question is land cover = “Water”, the answer to another question could not be land use = “Industrial”.

You can preview how the rules will look for users in the Question Preview pane by mousing over and clicking on the **Rule** icon (clipboard).

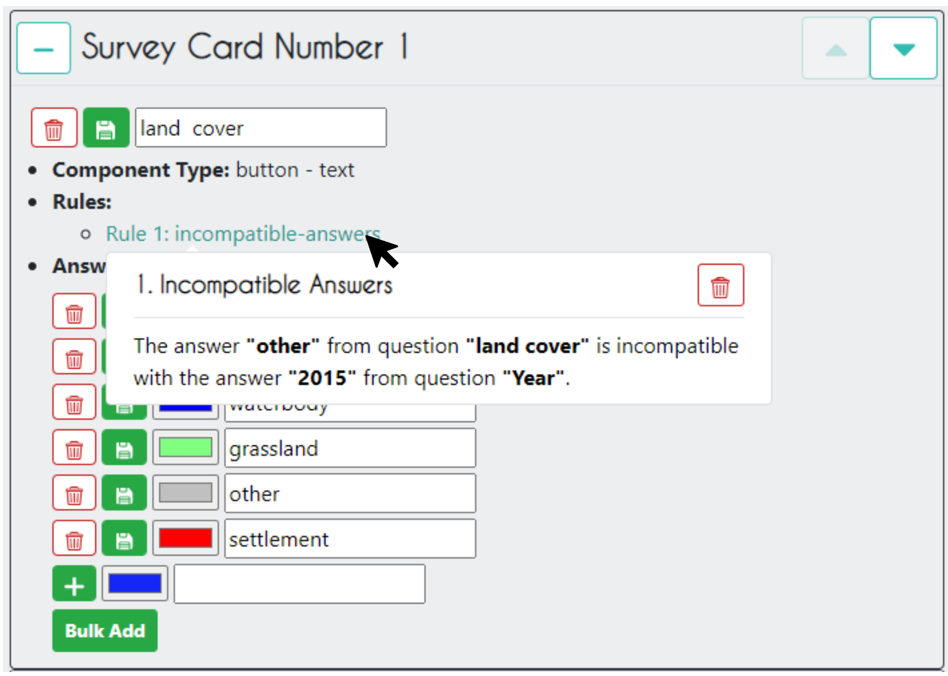




Clicking on the **Rule** icon will show a list of rules.



Navigating back to the Survey Questions tab will also show when questions have Rules associated with them. Mousing over the Rule will display the rule text.



Data collectors will be able to view any rules you implement by mousing over the rules icon.

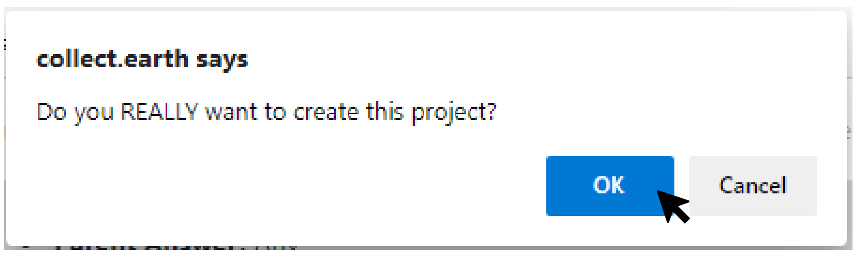
They will see an error if they try to enter an answer that conflicts with the rules:

K. Review Project

1. Once the project set up is complete, click on **[Review]** in the bottom right-hand corner.
2. You will now see an overview of your project details.
3. Scroll through to check that everything is correct.

   i. If everything is correct, click **Create Project**. You will need to accept CEO's terms and conditions.
   ii. If you notice a mistake, click **Continue Editing** to fix it.

4. After you click Create Project, a pop-up window will ask 'Do you REALLY want to create this project?' Click OK.



5. After you Create the project, you will be taken to the Project Information panel. This will be discussed further in Part 5: Project Information Interface.

Part 5: Project Information Interface

If you are navigating back to the **Project Information** (not visiting right after project creation), you can reach this page by finding your Institution on the **Home** page left panel. Click the **[Edit]** icon ( ) next to the project name to get to the **Project Information/Review Project** page.

The Project Information Interface has two components: **Project Details** on the left and **Project Management** on the right**.**

A. Project Details

The **Project Detail** pane includes an overview of your project, including sample collection design, imagery selection, the AOI boundary, and survey questions. The **Collection Design** will now show the location of a subset of your plots (a maximum number can be displayed). The **Plot Design** and **Sample Design** sections will show a summary of the choices you made or the .csv and .shp files you uploaded. **Survey Questions** shows all the **Survey Cards** you created, along with the corresponding **Component Type, Answers, and Rules.**

B. Project Management

The **Project Management** pane includes project publication information and important links to manage your project. First, examine the **Date Created**, **Date Published**, and **Date Closed** for your project. If your project is not published yet, the **Date Published** will not be present. There is also text describing your project status:

- Draft Mode projects: This project is in draft mode. Admins can review, edit, and test collecting the project. Publish the project in order for users to begin collection. Any data collected in a Draft Mode project will not be retained.
- Published projects: This project is published. Users can begin collecting. Limited changes to the project details can be made. Close the project to prevent anymore updates.
- Closed projects: This project is closed. The project is closed to all changes. Reopen the project for additional collection.

C. Modify Project Details

Next, examine the important links for your project:

1. **Publish Project**: Clicking on this will publish your project and allow users to collect data. Note that if a project is not "Published" only the admins of the institution can see it, not the members.
2. **Close Project**: If you have a published project, you can close it and stop data collection by clicking this button.

3. **Edit Project**: This button will take you back to the project creation wizard (See Part 4).

   i. When your project is in DRAFT MODE, you can change all aspects of your project.
   ii. When your project is PUBLISHED, you can change the **Project Name, Description, and Privacy Level**.

4. Delete Project: This will permanently delete your project.

If you accidentally publish a faulty project, you can use it as the template for a revised project and not lose all your work.

D. External Links

1. Configure Geo-Dash

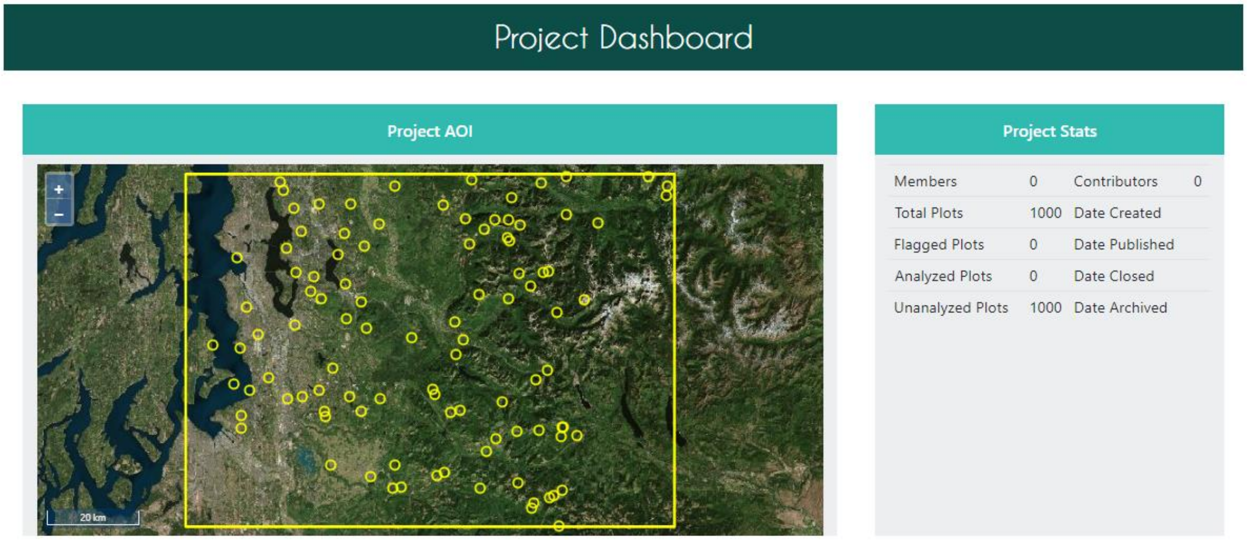
   i. Clicking on **[Configure Geo-Dash]** will open the Geo-Dash configuration interface.
   ii. For more detail, please see Part 6: Geo-Dash Implementation below.

2. Collect

   i. This will take you to your project data collection interface so you can start collecting data right away.

3. Project Dashboard

   i. If you click on **[Project Dashboard]** you will see an overview of the AOI, along with the number of members, contributors, total plots, flagged plots, analyzed plots, unanalyzed plots, and the date the project was created, published and closed.
   ii. Admins do not count towards the project's member count.



E. Export Data

1. There are two data download options: Download Plot Data, which downloads your data summarized by plot, and Download Sample Data, which downloads your raw data, with information for each point within each plot as its own row. Both are downloaded in .csv, which can be opened in programs like Microsoft Excel or imported into data analysis software.

2. Data downloaded from CEO will be in WGS84 EPSG:4326 format.

3. **Download Plot Data**: Downloaded .csv data from Download Plot Data will have the following columns:

    i. **PLOT_ID:** the CEO-assigned unique sample plot number or the user provided Plot ID (for .csv and .shp files).
    ii. **CENTER_LON** and **CENTER_LAT** are the geographic coordinates of the center of your sample plots.
    iii. **SIZE_M** and **SHAPE** describe the size in meters and the shape (circle or square) of the sample plot.
    iv. **SAMPLE_POINTS** indicates the number of samples in each plot.
    v. **EMAIL** is the user id (email address) of the person that classified the plot.
    vi. **FLAGGED**: This will be FALSE for plots where data was collected and for plots where data has not been collected yet. It will be TRUE when a user has flagged the quality of the background map as not good enough to analyze the samples (e.g. due to clouds or poor image resolution).

        a) Plots can either be flagged or saved by a user, but not both.
        b) If a user flags the plot, then goes back and enters data and saves the plot, the plot will not be flagged.

vii. **FLAGGED_REASON**: A user provided reason for flagging the plot.
viii. **CONFIDENCE**: Collected user confidence.
ix. **COLLECTION_TIME**: date and time the user collected the plot data. Time in UTC.
x. **ANALYSIS_DURATION**: Amount of time, in seconds, the user spent analyzing the plot.
xi. **COMMON_SECUREWATCH_DATE**: Most common SecureWatch date used.
xii. **TOTAL_SECUREWATCH_DATE**S: Number of SecureWatch dates used.
xiii. If you used a .csv or .shp file for plot design, any additional data columns you uploaded will be preserved in the .csv download. They will be preceded by PL_(column name).
xiv. All the following columns will have information about each of the survey questions broken down by answer. They are labeled **QUESTION TEXT:ANSWER TEXT**. For example, LULC:Built Surface would indicate that "LULC" was the question and "Built Surface" was the answer.
xv. In the plot summary download, these are quantified as percent (max 100) of the sample points in the plot that were assigned that answer.

   a) For example, suppose you have four sample points within your plot and two answers (e.g. land cover class) to choose from.
   b) If one sample point is assigned to one answer and the other three points to the second answer, the data when downloaded will say '25' for the first answer and '75' for the second answer.
   c) For proactive sampling, percentages are based on sample count, NOT AREA.

4. Download QA/QC
5. **Download Sample Data**: If you would like your plot data analyzed differently, the Download Sample Data option is a better fit. Downloaded .csv data from Download Sample Data will have the following columns:

i. **PLOT_ID**: the CEO-assigned unique sample plot number or the user provided Plot ID (for .csv and .shp files).
ii. **SAMPLE_ID**: the CEO-assigned unique sample point number or the user provided Sample ID (for .csv and .shp files).
iii. **LON** and **LAT** are the geographic coordinates of the center of your sample points.
iv. **EMAIL:** is the user id (email address) of the person that classified the plot.
v. **FLAGGED**: This will be FALSE for plots where data was collected and for plots where data has not been collected yet. It will be TRUE when a user has flagged the quality of the background map as not good enough to analyze the samples (e.g. due to clouds or poor image resolution).
vi. **COLLECTION_TIME**: The date and time when the user classified the plot. Time zone is UTC.
vii. **ANALYSIS_DURATION**: Time in seconds that the user took to analyze the plot.
viii. **IMAGERY_TITLE**: Name of the Imagery layer that the user had selected while analyzing the plot.
ix. **IMAGERY_ATTRIBUTIONS**: Any attribution for the imagery used.

If multiple imagery sources were used, only the name of the last imagery layer used will be recorded. There is no way to know all the layers used by the user, e.g. if you want to check that two years of imagery were used.

x. If you used a .csv or .shp file for sample plot design, any additional data columns you uploaded will be preserved in the .csv download. They will be preceded by PL_(column name).

xi. If you used a .csv or .shp file for sample point design, any additional data columns you uploaded will be preserved in the .csv download. They will be preceded by SMPL_(column name).

xii. All the following columns will have information about each of the survey questions. They will be labeled **QUESTION TEXT**, where question text is the literal text of the question.

Note that imagery dates are not available as many of the imagery sources are composite. This means that each map tile is stitched together from imagery acquired on multiple dates. There is not a single date for an imagery tile.

If you are using SecureWatch imagery, you will have four additional columns:

xiii. **IMAGERYDATESECUREWATCH** will have a value for any samples which were classified while a specific date was selected from the imagery date dropdown.

xiv. **IMAGERYSTARTDATESECUREWATCH**, **IMAGERYENDDATESECUREWATCH**, and **FEATUREPROFILESECUREWATCH** will have values for any samples which were classified while a date range and feature profile were selected.

xv. Note that imagery properties are associated with samples (not plots) because users are free to change these properties while classifying samples. Thus, any given plot may have some of its samples classified with one map image and other samples classified with a different map image.

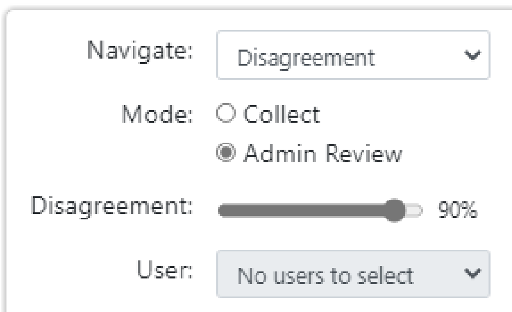
Part 6: Administrator Review Mode

Using the Data Collection interface, Administrators can review their projects.

1. For projects with QA/QC enabled, administrators can use the Review Mode to Navigate through plots with Disagreement. This navigation feature allows administrators to choose the level of disagreement between questions on the fly, then allow them to navigate between plots that have this level of disagreement or more.

   Disagreement is provided as a percentage and represents how far apart users' responses for the plot are from one another. If you have multiple samples per plot, the disagreement per sample is calculated and then averaged.

   We suggest that you start with a disagreement percentage of around 50% and adjust to your needs.



2. Once data for the project has been collected using a quality control approach, Administrators can press the "View Disagreements" button to bring up a new window that compares every user's answers for the plot's questions.

   This navigation mode is particularly useful for situations where administrators are identifying difficult plots to identify, difficult classes to identify, or are holding a group meeting to work through the identification of plots with high disagreement.

3. In addition, administrators can Navigate through unanalyzed plots, all analyzed plots, flagged plots, low confidence plots, and individual different users' plots.

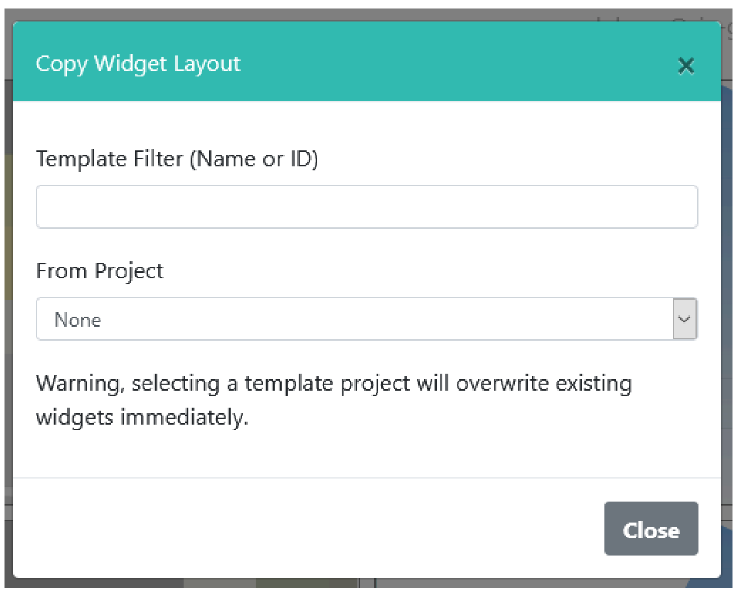
Part 7: Geo-Dash Implementation

Geo-Dash is a dashboard that opens in a second window when users begin to analyze sample plots. Geo-Dash provides users with additional information to help them interpret the imagery and better classify sample points and plots. The Geo-Dash tab can be customized to show information such as NDVI time series, additional imagery, and digital elevation data.

You can configure your **Geo-Dash** by clicking on **[Configure Geo-Dash]** from the **Project Information/ Review Project** page**.**

This will bring up the **Geo-Dash layout screen**.

[Copy Layout] will allow you to copy the Geo-Dash from another project. This will delete any Geo-Dash you have associated with the project!



You can add individual Geo-Dash widgets with the **[Add Widget].** To find out what each of the widgets do, clicking on **[Geo-Dash Help]** will open the **Geo-Dash Help Center**. We will also discuss this functionality below.

CEO uses Google Earth Engine to process many of the Geo-Dash widgets. Therefore, the information about Image Visualization for GEE is also useful here, specifically the min and max descriptions: [https://developers.google.com/earth-engine/image_visualization.](https://developers.google.com/earth-engine/image_visualization)

A. Common Indices and Data Sources

Some of the widgets proide access to common indices and data sources. These include NDVI, EVI, EVI2, NDMI, and NDWI, along with Landsat and Sentinel imagery. NDVI, EVI, EVI 2, NDMI, and NDWI indices are available in both image overlays and time series graphs.

The normalized difference vegetation index (NDVI) is used to determine if the cell contains live green vegetation. In map based representations, dead plants and inanimate objects are represented as red, while live healthy plants are represented as green. In numerical representations (e.g. time series graphs), values below 0 represent dead plants or inanimate objects, 0-0.33 represents unhealthy plants, 0.33-0.66 represents moderately healthy plants, and 0.66-1 represents very healthy plants. For more information, see e.g. https://www.usgs.gov/land-resources/nli/landsat/landsat-normalized-difference-vegetation-index

The enhanced vegetation index (EVI) and two-band EVI (EVI 2) are optimized vegetation indexs. They are designed to have higher sensitivity in high biomass regions e.g. along the equator, correct for canopy background signals, and reduce atmospheric influence on index values. In doing so, these indices addresses some of the key limitations of NDVI, however EVI requires more data to calculate and therefore has its own limitations. EVI 2 is in development and can be calculated just from red and near infrared bands. As with NDVI, red is used to represent dead plants/inanimate objects and green to represent healthy plants. The index varies between 0-1, with 0 representing dead plants and 1 representing very healthy plants. For more information on EVI see e.g. https://www.usgs.gov/land-resources/nli/landsat/landsat-enhanced-vegetation-index.

The normalized difference moisture index (NDMI) is used to determine the water content of vegetation. NDMI can be used for drought monitoring and for determining fuel loads (combustability) for wildfire hazard assessments. Values near -1 indicate plants with low moisture while values near 1 indicate plants with high moisture. More information on NDMI can be found at e.g. https://www.usgs.gov/land-resources/nli/landsat/normalized-difference-moisture-index.

The normalized difference water index (NDWI) is also related to plant water content and plant water stress. It can be used to map water bodies, determine crop health and for wildfire risk analysis. Values near -1 indicate low water content and vegetation cover and values near 1 indicate high water content and vegetation cover. More information on NDWI can be found at e.g. https://edo.jrc.ec.europa.eu/documents/factsheets/factsheet_ndwi.pdf.

Data from Landsat satellites 5, 7, 8 and Sentinel 2 are also available. For more information about these datasets, see: Landsat: https://www.usgs.gov/land-resources/nli/landsat and Sentinel 2: https://sentinel.esa.int/web/sentinel/missions/sentinel-2. When more than one image is available for the date range selected for Landsat, the simpleComposite function in GEE is used to create simple cloud-free composite. It uses the median of the least cloudy pixels. Sentinel 2 data is also reduced using the median of least cloudy pixels.

B. Widget Formatting Notes

1. Titles cannot contain special characters.
2. Image parameters use JSON. Quotes used for image parameters in widgets should be vertical (" not “”). There should be no spaces in the image parameters.
3. Lists can be formatted either as "1,2,3" or as [1,2,3] for numbers and "B3,B2,B1" or ["B3","B2","B1"] for bands.
4. When displaying three bands (e.g. B3, B2, B1 in RGB), GEE can use either one or three values for min and max.
5. For the Date Ranges, if the end date is longer than the period of record, only the available data will be displayed.

C. Move & Resize widgets

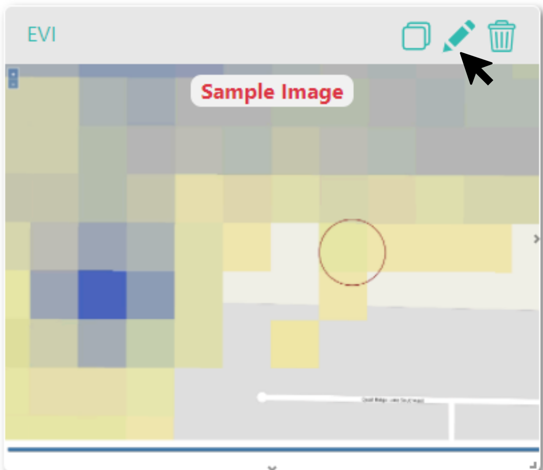
1. Widgets can be manipulated on the Geo-Dash Widget Layout Editor in multiple ways:

   i. Drag and drop
   ii. Resize by dragging from the bottom right corner
   iii. Widgets will auto align

2. Widgets are updated in real time
3. View rendered results in the Geo-Dash window
4. For a gif illustrating these movements, see https://collect.earth/geo-dash/geo-dash-help and click To Move and Resize Widgets.

D. Duplicate, Edit, and Delete Widgets

1. You can copy a widget once it has been created by clicking on the duplicate icon in the upper right hand corner of the widget.
2. You can edit widgets once they are created by clicking on the pencil icon in the upper right hand corner of the widget.
3. To delete a widget, click on the trashcan in the upper right-hand corner.



E. (Forest) Degradation tool

The Degradation tool provides time series information on forest degradation from selective logging, fire, and other large and small disturbances. The Normalized Difference Fraction Index (NDFI) allows for enhanced detection of forest canopy damage from multiple sources, including selective logging and forest fires.

It is calculated with this equation:

$$\mathrm{NDFI} = \frac{\mathrm{GV}_{\mathrm{Shade}} - (\mathrm{NPV} + \mathrm{Soil})}{\mathrm{GV}_{\mathrm{Shade}} + \mathrm{NPV} + \mathrm{Soil}}$$ where

$$\mathrm{GV}_{\mathrm{Shade}} = \frac{\mathrm{GV}}{100 - \mathrm{Shade}}$$

More detail can be found in: Souza, C. M., Roberts, D. A., & Cochrane, M. A. (2005). Combining spectral and spatial information to map canopy damage from selective logging and forest fires. *Remote Sensing of Environment*, 98(2), 329-343.

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select Degradation Tool in the Type drop down.
3. Give the widget a Title.
4. Choose the Basemap source from dropdown. This will be the basemap for the widget and other data will be layered on top.
5. Choose which Band to graph. Available options include SWIR1, NIR, Red, Green Blue, SWIR2, and NDFI. NDFI allows for enhanced detection of forest canopy damage from multiple sources, including selective logging and forest fires.
6. Select the Date Range you would like. This tool can draw from Landsat 4 (July 1982-December 1993), Landsat 5 (March 1984-January 2013), Landsat 7 (April 1999 – current), and Landsat 8 (February 2013 – current) based on the date range you select. Sentinel information is only available from April 2014 onward (Sentinel 1A launch).
7. Click Create.
8. Reposition and resize to your liking. This tool will have two panels, so be sure to give it ample vertical space.

F. Add a Dual Imagery Widget

The dual imagery widget combines two different image collection widgets so the user can compare the images side by side.

Show example of the widget to illustrate

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select Dual Imagery in the Widget Type drop down.
3. Give the widget a Title. We suggest an informative title such as {Data} {Date range} e.g. NDVI 2001 or NDVI 2001 – 2002.
4. Choose Basemap source from dropdown. This will be the basemap for the widget and other data will be layered on top.
5. You will now select the Imagery type for the Bottom and Top images. You can toggle between them by clicking the [Next Image] and [Previous Image] buttons.

   i. Select the Imagery Type for each image. Options include Image Asset, Image Collection Asset, and Preloaded Imagery Collection.
   ii. For more information about Image Assets, see the section titled “Add an Image Asset Widget.”
   iii. For more information about Image Collection Assets, see the section titled “Add an Image Collection Asset Widget.”

iv. For more information about Preloaded Imagery Collections, see the section titled “Add a Preloaded Image Collection Widget.”

a) **NDVI**, **EVI**, **EVI 2**, **NDMI**, and **NDWI** are preconfigured with the correct Landsat bands and image parameters. NDVI merges :Landsat 4 thru Landsat 8 and Sentinel 2.

b) **Landsat 5, Landsat 7, Landsat 8**, and **Sentinel 2** are partially configured leaving you the option to adjust the bands, min, max, and cloud score. Available bands for each are:

- Landsat 5 - B1, B2, B3, B4, B5, B6, B7, BQA
- Landsat 7 - B1, B2, B3, B4, B5, B6_VCID_1, B6_VCID_2, B7, B8, BQA
- Landsat 8 - B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, BQA
- Sentinel 2 - B1, B2, B3, B4, B5, B6, B7, B8, B8a, B9, B10, B11, B12, QA10, QA20, QA60

c) For information on the min, max, and cloud score see Part 6: B “Add an Image Collection Widget”. Min and Max are the values for the bands, while cloud score is a max percent of clouds in the image.

6. Select the Date range you would like, if applicable. If the end date is longer than the period of record, only the available data will be displayed.
7. Click **[Create]**.
8. Reposition and resize to your liking.

G. Institution Imagery

This widget allows you to simply show one of your institutional imageries. With this you can select

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select Institution Imagery in the Type drop down.
3. Give the widget a Title.
4. Select a Basemap. This has all of your institutional imagery available in the dropdown menu.
5. Click **[Create]**.
6. Reposition and resize to your liking.

> Note that the Public Planet NICFI will not work as expected. Please use your own Planet NICFI key to set up institutional imagery that specifies the specific time step. See “Part 3: B Adding additional basemap imagery” for more information.

H. Add an Image Asset Widget

Both the image asset widget and the image collection asset widget allow users to connect to GEE data. This includes publicly available data through GEE as well as image assets from their own user accounts. There is a video tutorial for creating GEE image assets available online here: https://youtu.be/l57IhmduVBQ. Some example GEE scripts to create Image Asset in GEE:

- NDVI: https://github.com/SERVIR/gee-scripts/blob/master/GEE_NDVI_Clip_to_plot_Collection.js#L6
- Landsat8: https://code.earthengine.google.com/109485ad0712818dba270b767efbc540

7. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
8. Select Image Asset in the Type drop down.
9. Give the widget a Title.
10. Choose the desired Basemap source from dropdown. This will be the basemap for the widget and other data will be layered on top.
11. Enter the GEE Image Asset ID. For example, COPERNICUS/S2/20170420T184921_20170420T190203_T10SEG . This can be any image asset in Earth Engine, including assets from your own account. If you are using an asset from your own account, it must be shared publicly.
12. You can view the Available Bands for your image by clicking on the **[Refresh]** button.
13. Enter Image Parameters for the asset in JSON format. These follow the image visualization function in GEE (more info: https://developers.google.com/earth-engine/guides/image_visualization?hl=en):

```
{"bands":"CO_column_number_density,H2O_column_number_density,cloud_height","min":"0","max":"0.5"}
```

```
{"bands":"NDVI","max":"1","min":"-1","palette":"c9c0bf,435ebf,eee8aa,006400"}
```

    Note that there are no spaces, and all color values must be hex. The quotes around the numbers are optional, e.g. "max":0.5 will also work. As in GEE, you may also use brackets around the palette values, however if you use brackets each individual number must be quoted.

14. Click Create.
15. Reposition and resize to your liking.

More on SRTM Digital Elevation Data

Previously in CEO you could add digital elevation data directly. This is now handled with the image asset widget. The SRTM Digital Elevation Data 30m widget pulls data from the Earth Engine Data Catalog. More information on this dataset can be found here: https://developers.google.com/earth-engine/datasets/catalog/USGS_SRTMGL1_003. The SRTM is an image asset USGS/SRTMGL1_003.

I. Add an Image Collection Asset Widget

An Image Collection is a stack or time series of images. There is a video tutorial for creating GEE Image Collection assets available online here: https://youtu.be/7eIvltgDbXw.

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select **Image Collection Asset** in the Type drop down.
3. Give the widget a **Title**.
4. Choose Basemap source from dropdown. This will be the basemap for the widget and other data will be layered on top.
5. Enter the Google Earth Engine Image Collection Asset ID, for example LANDSAT/LC8_L1T_TOA. This can be any image collection in Earth Engine, including assets from your own account. If you are using an asset from your own account, it must be shared publicly.
6. You can view the Available Bands for your image by clicking on the **[Refresh]** button.
7. Select a Collection Reducer. Reducers are methods for collapsing multiple images in the image collection into a single image for the Geo-Dash to display. You can learn more about reducers here https://developers.google.com/earth-engine/guides/reducers_intro.
8. Enter Image Parameters for the asset in JSON format. These follow the image visualization function in GEE (more info: https://developers.google.com/earth-engine/guides/image_visualization?hl=en):

   i. {"bands":"B4,B5,B3","min":"10,0,10","max":"120,90,70"}
   ii. Again, do not use any spaces.

9. Select the Date Range.
10. Click [Create].
11. Reposition and resize to your liking.

More on Landsat, one of the Image Assets you can add:

For better performance when using a GEE collection, we recommend that you preprocess and clip the data to your AOI in GEE before connecting it to Geo-Dash. You can find a video tutorial on this process in the Geo-Dash Help Center under "Video Tutorial to Create GEE Image Assets."

The most common strings for Landsat are listed below. More detailed information can be found on the Google Earth Engine help pages; available composites and the covered time frame are found here: https://explorer.earthengine.google.com/#search/tag%3A32day.

For short change intervals, test the 8 day NDVI composite images; for longer change intervals, explore the 32 day composites.

You might notice in the 8 day composites unexpectedly low NDVI values for the season. This can be caused by cloud cover. If your sample sites are in an area with persistent cloud cover, choose the 32 day composite.

Satellite	Type	Start date	Image collection ID
Landsat 8	NDVI, 32 day	2013-04-07	`LANDSAT/LC8_L1T_32DAY_NDVI`
Landsat 8	NDVI, 8 day	2013-04-07	`LANDSAT/LC8_L1T_8DAY_NDVI`
Landsat 7	NDVI, 32 day	1999-01-01	`LANDSAT/LE7_L1T_32DAY_NDVI`
Landsat 7	NDVI, 8 day	1999-01-01	`LANDSAT/LE7_L1T_8DAY_NDVI`
Landsat 8	NDWI, 32 day	2013-04-07	`LANDSAT/LC8_L1T_32DAY_NDWI`
Landsat 8	NDWI, 8 day	2013-04-07	`LANDSAT/LC8_L1T_8DAY_NDWI`
Landsat 7	NDWI, 32 day	1999-01-01	`LANDSAT/LE7_L1T_32DAY_NDWI`
Landsat 7	NDWI, 8 day	1999-01-01	`LANDSAT/LE7_L1T_8DAY_NDWI`

The image collection ID for Sentinel-2 is COPERNICUS/S2 (2015-06-23 to present) and for Sentinel-1 is COPERNICUS/S1_GRD (2015-10-03 to present).

J. Add a Polygon Compare Widget

The polygon compare widget allows you to display polygons in Geo-Dash, so that e.g. you can compare them with your plot outlines. This uses a featureCollection in Google Earth Engine as input. You can use any feature collection, if it has a field that matches the PLOTID for your CEO project. The polygon is extracted from the featureCollection and turned into an image to display as WMS tiles in the widget.

We suggest you use the shapefile (.shp) option when creating your CEO project.

IMPORTANT: Your featureCollection needs to be shared as public so that CEO can access it.

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select Polygon Compare in the Type drop down.
3. Give the widget a Title.
4. Choose Basemap source from dropdown. This will be the basemap for the widget and other data will be layered on top.
5. Enter the Google Earth Engine Feature Collection asset ID. This can be any feature collection in Earth Engine, including assets from your own account. If you are using an asset from your own account, it must be shared publicly.
6. Enter the field that matches the PLOTID for the project. In the compare featureCollection widget, you will need to match PLOTIDs. The PLOTID column for your featureCollection can be named anything, just type in the proper column name into the Field to match PLOTID box.
7. Enter Image Parameters in JSON format to change the color of the polygon displayed. This uses the GEE feature and feature collection visualization function.
8. Click **[Create].**

K. Add a Preloaded Image Collection Widget

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select **Preloaded Image Collections** in the Widget Type dropdown menu.
3. Give the widget a Title.
4. Choose your Basemap source from the dropdown menu. This will become the basemap over which the other data is overlaid.
5. Select the Data Type:

   i. **NDVI, EVI, EVI 2, NDMI**, and **NDWI** are preconfigured with the correct Landsat bands and image parameters. NDVI combines Landsat 4 through Landsat 8 and Sentinel 2. If you select one of these, you will just need to add a **Date Range**. The index displayed will be the mean for the selected date range.

   ii. **Landsat 5, Landsat 7, Landsat 8, and Sentinel 2** are partially configured leaving you the option to adjust the bands, min, max, and cloud score. Available bands for each are:

a) Landsat 5 - B1, B2, B3, B4, B5, B6, B7, BQA
b) Landsat 7 - B1, B2, B3, B4, B5, B6_VCID_1, B6_VCID_2, B7, B8, BQA
c) Landsat 8 - B1, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, BQA
d) Sentinel 2 - B1, B2, B3, B4, B5, B6, B7, B8, B8a, B9, B10, B11, B12, QA10, QA20, QA60
e) Min, Max, and Cloud Score will be determined by the user's needs. Min and Max are the values for the bands, while cloud score is a max percent of clouds in the image. Example values include:
   (i) Landsat5:
   bands: B4,B5,B3,
   min: 0.05,
   max: 0.45,
   Cloud Score: 90
   (ii) Landsat7:
   bands: B4,B5,B3,
   min: 0.05,
   max: 0.45,
   Cloud Score: 90
   (iii) Landsat8:
   bands: B5,B6,B4,
   min: 0.04,
   max: 0.5,
   Cloud Score: 90
   (iv) Sentinel 2:
   bands: B8,B4,B3,
   min: 450,
   max: 3000,
   Cloud Score: 10
f) The pieces that are preconfigured include:
   (i) landsat8: imageCollection: LANDSAT/LC08/C01/T1_RT; simpleComposite: 50
   (ii) landsat7: imageCollection: LANDSAT/LE07/C01/T1; simpleComposite: 60
   (iii) landsat5: imageCollection: LANDSAT/LT05/C01/T1; simpleComposite: 50
   (iv) sentinel 2: imageCollection: COPERNICUS/S2

6. Select Date Range. If the end date is longer than the period of record, only the available data will be displayed.
7. Click **[Create].**
8. Reposition and resize to your liking.

L. Add a Statistics Widget

Statistics provide additional information, including population, area, and elevation about the current plot area that is being analyzed. The population data is from the CiESIB Gridded Population of the World v4 dataset and the elevation data is from the USGS EROS Global 30-Arc Second DEM.

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select **Statistics** in the type drop down.
3. Give the widget a Title.
4. Click [Create].
5. Reposition and resize to your liking.

M. Add a Time Series Graph Widget

Time Series Graphs, particularly for the vegetation indices, can help users differentiate between different types of vegetation based on seasonal patterns.

1. Click **[Add Widget]** in the upper right hand of the Geo-Dash layout screen.
2. Select **Time Series Graph** in the Type drop down.
3. Give the widget a Title.
4. Select the Data Type:

    i. **NDVI, EVI, EVI 2, NDMI**, and **NDWI** are preconfigured with the correct Landsat band calculations and image parameters. NDVI merges Landsat-4 thru Landsat 8 and Sentinel 2. For more information on these, please see the section titled “Add a Preloaded Image Collection Widget.”

    ii. **Custom** allows you to graph an GEE image collection. An image collection from Google Earth Engine can be added here. You simply need to know the image name, the band you would like graphed, and how you would like the graph reduced. For example:

        a) GEE Image Collection - COPERNICUS/S5P/OFFL/L3_CO.
        b) Band to graph - CO_column_number_density.
        c) Select the Reducer in the dropdown.

5. Select a **Date Range**. If the end date is longer than the period of record, only the available data will be displayed.
6. Click [Create].
7. Reposition and resize to your liking.

Part 8: Creating a Project for Validation

Creating a project in CEO to validate a spatial model—for example, a land cover map or a forest change map—presents some unique sampling and project design issues. In this section, we provide a very brief overview these issues both for single and for ensemble change maps.

Many resources provide a more in-depth review of this subject. Consider looking at:

- CEO's Theoretical Manual, available here: https://collect.earth/downloads/CEO_Theoretical_Manual.pdf
- the FAO's *Map Accuracy Assessment and Area Estimation: A Practical Guide* available online here: http://www.fao.org/3/a-i5601e.pdf
- Documentation in AREA², available online here: https://area2.readthedocs.io/en/latest/background.html
- A great overview on confusion matrixes here: http://spatial-analyst.net/ILWIS/htm/ilwismen/confusion_matrix.htm

A. Accuracy assessment

Accuracy assessments allow producers of spatial models to understand how accurate they are, that is, how well the model fits. This information is important for many reasons, including improving accuracy and determining if the product is suitable for management use.

For single models, such as a single land cover classification map, a confusion matrix is used to calculate the accuracy of spatial models. The purpose of this error analysis is to quantify two key metrics for each land cover class as well as the overall classification:

1. Given that a pixel is of a (land cover class, etc.), what is the chance that it was correctly classified as that (land cover class, etc.)?
2. Given that a pixel has been classified as a (land cover class, etc.), what is the chance that it belongs to that (land cover class, etc.)?

There are two pieces of information needed to answer these questions: the 'true' land cover class assignment, which is done by a human, and the 'as classified' land cover class assignment, which is output by the model. An example of the end goal (the confusion matrix) is shown below for a two class system.

	Actual: Forest	Actual: Not Forest
Predicted: Forest	29 (true positives)	22 (false positives)
Predicted Not Forest	12 (false negatives)	50 (true negatives)

For ensemble models, the approach is dependent on the type of ensemble model. For ensemble models where you have combined the results of the different models into a single model (e.g. through averaging), you can use an approach like that used for single models.

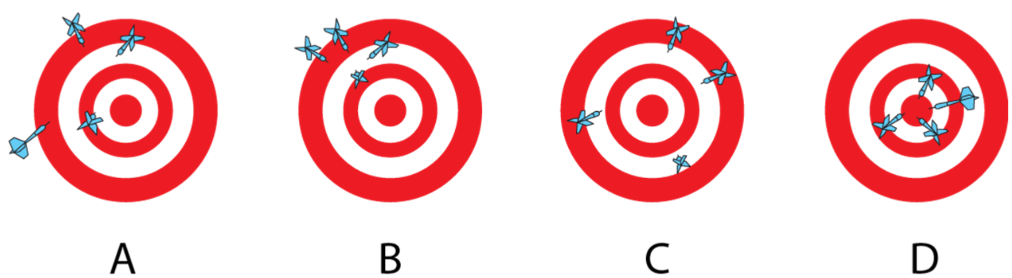
Where you have multiple different maps, you may need a more complex approach. For example, if you have two model outputs you may choose to stratify based on the values of both outputs. We will not discuss more complex approaches in depth here.

B. Validation goals & required sample size

Precision & accuracy goals for a project help us decide on the sample size and sample design.

Precision is how close your measurements or estimates are to one another. This is sometimes called dispersion. Precision describes random errors & variability.

Accuracy is how close your measurements or estimates are to the true answer. Accuracy describes systematic errors—including observational error. In statistics this is sometimes called bias instead of accuracy. Accuracy is sometimes hard to evaluate as the “true” value might not be known—but with Land Use & Land Cover analysis usually it can be.



Precise?	No	Yes	No	Yes
Accurate?	No	No	Yes	Yes

In this example, A is neither precise nor accurate. B is precise, but not accurate. C is accurate, but not precise, and D is both precise and accurate. This is of course where we would like to be ideally.

Accuracy and precision influence the necessary sample size. Detecting small changes in outcome (e.g. 4% decrease in forest cover) with high certainty requires large sample sizes.

When you are determining your sample size, you will want to consider the primary goals of the project, along with the resources available for data collection.

There are many different formulas for determining sample sizes. Which one you want to use is based in part on your sampling method (simple random sample, stratified random sample, etc.).

Two examples follow:

From SEPAL.io:

Formula to calculate the overall sample size

The equation below calculates an adequate overall sample size for stratified random sampling that can then be distributed among the different strata.

- N is number of units in the area of interest (number of overall pixels if the spatial unit is a pixel, number of polygons if the spatial unit is a polygon)
- S(O) is the standard error of the estimated overall accuracy that we would like to achieve
- Wi is the mapped proportion of area of class i
- Si is the standard deviation of stratum i.

$$n = \frac{(\sum W_i S_i)^2}{[S(\widehat{O})]^2 + (1/N)\sum W_i {S_i}^2} \approx \left(\frac{\sum W_i S_i}{S(\widehat{O})}\right)^2$$

From the CEO Technical Manual:

Expected accuracy of the product (P_0)

Precision of detecting differences from this accuracy (minimum detectable difference, δ)

Tolerance of Type I error (alpha, α)

Tolerance of Type II error (beta, ß)

$$n' = \left[\frac{Z_\alpha\sqrt{P_0(1-P_0)} + Z_\beta\sqrt{P_1(1-P_1)}}{\delta}\right]^2$$

Equation 1. The sample size determination equation.

The following equation can then be applied:

$$n = \frac{n'}{4}\left(1 + \frac{2}{n'\delta}\right)^2$$

Equation 2.The continuity correction.

Area Estimation & Accuracy Assessment (AREA[2]) also has helper scripts for determining sample size, including assigning area-based weights. Instructions can be found here: https://coded.readthedocs.io/en/latest/sample.html.

If you have access, the following paper also has a good approach specifically tailored to classification: Foody, G. M. Sample size determination for image classification accuracy assessment and comparison. Proc. 8th Int. Symp. Spat. Accuracy Assess. Nat. Resour. Environ. Sci. 30, 154–162 (2008). https://doi.org/10.1080/01431160903130937

C. Sampling approach

For validating a single model or an ensemble model that has been averaged/combined, stratified sampling is the preferred method. This is to ensure you can accurately and precisely determine how well your classification performs for each of your predicted groups (land cover classes, forest degradation, etc.). If you use random or gridded sampling, you might not sample enough points in rare classes to be able to tell if they are well estimated.

Strata used for stratification must be exhaustive--they include the entire study area--and they must be exclusive--a sample unit can't belong to more than one strata.

CEO cannot currently create stratified sampling designs. However, you have multiple good options to create a stratified sample outside of CEO.

1. You can use SEPAL.IO to create a stratified sample based on how accurate you think your model is for each stratum and a raster of your model output.
2. You will need to sign up to use it by visiting sepal.io and clicking on **[Sign Up]**.
3. From the apps, find the **Stratified Area Estimator-Design** tool.
4. From there, follow SEPAL's directions.
5. You can also run this on your own machine; see the instructions here: https://github.com/openforis/accuracy-assessment

You can also use QGIS or ArcGIS for creating a stratified sample. Many resources are available online, including:

- This blog post for QGIS: https://pvanb.wordpress.com/2012/12/03/stratified-random-sampling-in-qgis/ Note you will need to convert your classification raster to a vector layer.
- This description of the Sampling Design Tool in ArcGIS: https://www.arcgis.com/home/item.html?id=28f08ca526ae44e8ac107a2a0d5f50e3

Once you have created a stratified sample, you can create a project in CEO and follow the directions in Part 4: H Uploading CSV & SHP files to upload your sampling points into CEO.

After collecting data on all of your sampling points, you can download the data (see Part 7: E Download your data) and calculate your confusion matrix.

D. Alternative: Using TimeSync

TimeSync is an application that allows researchers and managers to characterize and quantify disturbance and landscape change by facilitating plot-level interpretation of Landsat time series stacks of imagery (a plot is commonly one Landsat pixel).

TimeSync can be used to validate map data products, especially where estimating change is a primary objective. Also, TimeSync can facilitate production of independent estimates of change and disturbance rates from remotely sensed imagery.

For more information on TimeSync please see: http://timesync.forestry.oregonstate.edu/. A tutorial on how to use TimeSync is available at: http://timesync.forestry.oregonstate.edu/tutorial.html.