Collect Earth Online Module 4

Collecting Data for a Collect Earth Online Project

SERVIR Science Coordination Office
Curriculum Development Team
Micky Maganini
mrm0065@uah.edu

Prepared for “From Data to Geo-Information for Natural Resources Management” at ITC
Quartile 2 2022-2023

Table of Contents

Prerequisites and Requirements

Requirements

Before taking this course, you will need the following:

- A basic understanding of optical remote sensing
- A computer
- Connection to the Internet
- A Collect Earth Online account
- Be a member of the “ITC:NRM2” Collect Earth Online Institution
- Have Google Earth Pro downloaded on your computer
    - Click Here to Download Google Earth Pro

Prerequisites

- It is recommended that you take Collect Earth Online Module 1 prior to taking this module
    - Click here to view Collect Earth Online Module 1
- It is recommended that you take Collect Earth Online Module 2 prior to taking this module
    - Click here to view Collect Earth Online Module 2
- It is recommended that you take Collect Earth Online Module 3 prior to taking this module
    - Click here to view Collect Earth Online Module 3.

Learning Objectives

By the end of this module, you will

- Better understand how to interpret satellite imagery
- Understand how to collect data in a Collect Earth Online Project
- Understand how to import data from a Collect Earth Online Project

Best Practices for Interpreting Satellite Imagery

In order to interpret satellite imagery, we must go from the perspective of a human to the perspective of a satellite. Not only do we have to accommodate this shift in perspective, we also have to adjust to a shift in resolution, going from high definition sight to pixelated imagery. This section will cover some best practices for interpreting satellite images. The Collect Earth Online Data Collection Manual provides more in-depth guidance as to how to interpret satellite imagery. Click here to view the Collect Earth Online Data Collection Manual.

In order to accurately interpret satellite imagery, we must imagine we are high above the ground, looking straight down (i.e. perpendicular to the Earth's surface). If you have flown in a plane before, think about how the landscape looks from that perspective: people turn into miniscule dots, monstrous trees turn into green circles, and mountains become wrinkles on the Earth's surface. Even rivers look different, as most do not appear blue but rather nearly black in satellite imagery.

A good way to get accustomed to interpreting satellite imagery is to take advantage of your local knowledge. Go to your favorite imagery platform, such as Google Earth, and navigate to a place you are familiar with. See what the buildings, roads, and landmarks look like from this view.

Below we can find some examples of what different land cover classes look like from a top-down perspective. These images are provided by the Collect Earth Online Data Collection Manual.

A. Conifer Forest

B. Deciduous Forest

C. Pasture

D. Agriculture in an arid environment

E. Agriculture in a temperate environment

F. River

G. Lake

Now that you are more familiar with what different land cover classes look like in satellite imagery, let's discuss the steps you can take when interpreting an image. These tips for interpreting imagery are drawn from NASA's Earth Observatory writers and visualizers, who use them to interpret images on a daily basis. Click here to read the NASA Earth Observatory article.

1. Look for a scale
2. Look for patterns, shapes, and textures
3. Define the colors (including shadows)
4. Consider your prior knowledge

Step 1: Look for a Scale

In order to accurately interpret your imagery, you must know the size of the objects you are investigating. Does the image cover 1 kilometer or 100? Luckily, Collect Earth Online shows the scale on its map interface, providing an easy reference for scale when viewing imagery. Zooming out also gives you a better idea of the context that your assessment unit is situated in.

Step 2: Look for patterns, shapes, and textures

Shapes and patterns provide key clues about what land cover type we are seeing. Rivers stand out due to their unique meandering shapes. Farms usually have geometric shapes to them which help to distinguish them from natural vegetation. Straight lines are a good indication that this feature is human made. Textures can be very useful in distinguishing similar land cover classes from each other (e.g. grassland vs. vegetation). They are also useful to note the topography of an area, which may make certain land cover types more probable.

Step 3: Define the colors (including shadows)

Your first step here is to note if you are using true color imagery or false color imagery. If you are using false color imagery, take note of what wavelength each band corresponds to, as well as what those wavelengths can tell you about the land cover/use.

In a true color image, water appears as black or dark blue due to its tendency to absorb light. Sediment reflects light and colors the water. When suspended sand or mud is common in the water, it may look brown. Sunlight reflecting off the surface of the water makes the water look gray or silver. Ice is white/gray and can have a blue tinge.

Grasslands tend to be pale green, while forests are very dark green. Land used for agriculture is often much brighter in tone than natural vegetation. Plant color can also depend on the season, where spring vegetation tends to be paler than summer vegetation. Fall vegetation can be red, orange, yellow, or tan, winter vegetation is brown.

Bare ground is usually some shade of brown or tan, but can also be pink, tan, white, or black due to varying mineral content.

Urban areas are typically gray or red due to building materials.

It is important to remember that when you are investigating an image, you are looking at everything between the satellite and the ground in a single plane. This means that a white patch in your image may be a cloud, but it also could be snow, a salt flat, or sunglint on water. This is why zooming out to gather contextual information about the area and understanding the scale of your image are so important.

Step 4: Consider your Prior Knowledge

If you have local knowledge of the area you are studying, you have a significant advantage in interpreting the imagery over those who are not familiar with the area. It is for this reason that local stakeholders are often involved in the data collection process in image interpretation projects. Having local knowledge allows you to draw a connection between the satellite imagery to everyday life.

If you lack knowledge of the area of interest, a reference map or atlas can be valuable, to help give you names for the visible features.

Collecting Data in a Collect Earth Online Project

Allow Pop-ups

Collect Earth Online requires the use of browser pop-ups to function properly. Thus, if it is your first time collecting data within Collect Earth Online, you may need to allow pop-ups from the CEO site. The way to enable pop-ups varies based on the web browser you are using (if possible, use Google Chrome).

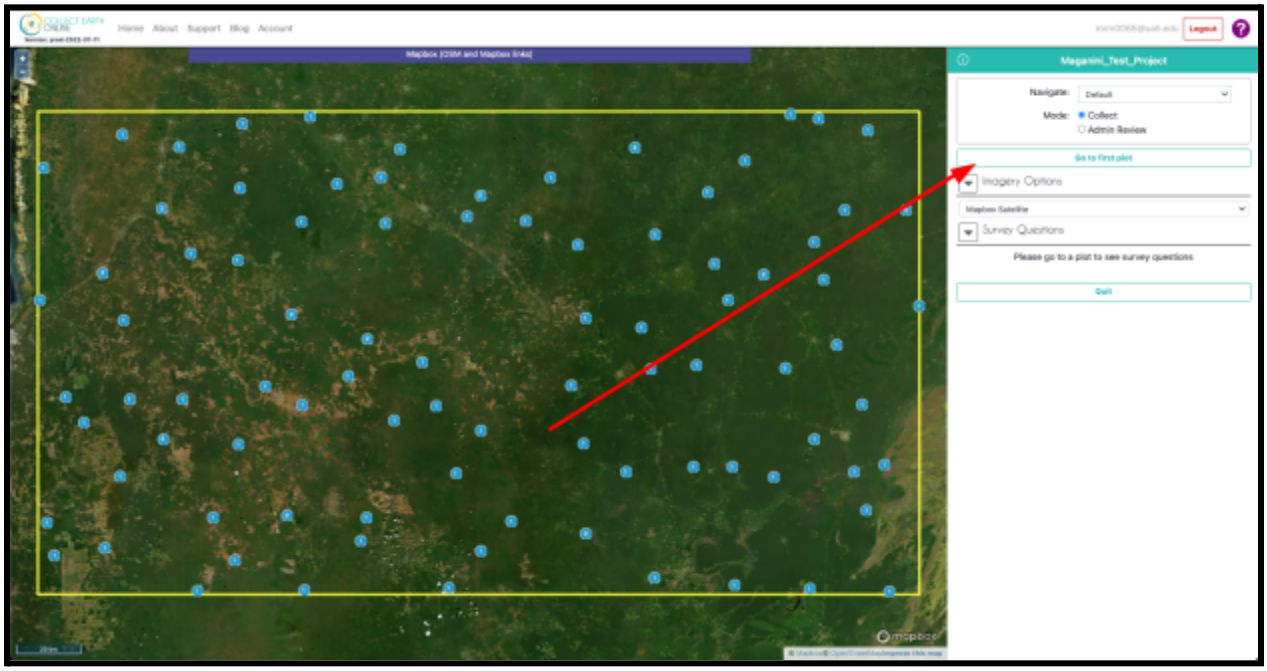
For Google Chrome: check the address bar. If it is marked with a pop-up blocked icon (as shown in the image below), then you must click on the **"popup blocked"** warning.

In the popup window that appears, click**"Always allow pop-ups from https://collect.earth/"** . Then click **"Done"**.

Navigating to the Data Collection page in your Project

By this point, you should have a published Collect Earth Online Project. Our first step is to navigate our institution's ("ITC: NRM2") homepage. Click here to visit our Institution homepage. After visiting this link, make sure you are logged in by clicking on the **"Login/Register"** button in the upper right corner. Your screen will now appear similar to the screenshot shown below (except you will see more projects underneath the "Maganini_Test_Project" Project). Click on your project.

After clicking on your project, your screen will appear similar to the screenshot shown below. Click the **"Go to first plot"** button towards the upper right of the screen, as indicated by the arrow in the image below.



After clicking this button, your screen will appear similar to the screen shown below, unless you have enabled Geo-Dash widgets. If you have enabled Geo-Dash Widgets, you will see the Geo-Dash screen, which is opened in a separate tab. Click the tab to the left of the GeoDash screen in your browser window and you will now see a screen similar to the image shown below. We are now ready to collect data!

Classifying Imagery

Pictured above is the image you will see when you are collecting data. The yellow box in the image above represents your assessment unit. The black dots that appear within the box represent your sampling points. You will classify each of these points based on your classification schema. In the bottom left of the map interface, you can see the scale of the imagery. The number indicated by the red arrow shows the plotid of the plot we are investigating (which in this case is 1 as it is the first plot we are investigating). Below that, we can see the external tools we have made available for this project. Clicking the "Re-Zoom" button will automatically fit your window to the yellow square. Clicking the "Geo-Dash" button will bring up a separate window that displays the Geo-Dash widgets you have selected for this project.

Clicking the **"Download Plot KML"** button will download a KML file containing the location of this assessment unit. Clicking that file will open Google Earth Pro (if you have it installed) and will automatically take you to the plot's location. Google Earth Pro is useful because it offers high to very high resolution timestamped optical data. To learn how to download and interact with Google Earth Pro, click here to view the supplementary material for CEO Module 4: Google Earth Pro Basics. Underneath the text that says **"Imagery Options"**, you will see the imagery source that is currently displayed in CEO, which in this case is the Mapbox Satellite. To select a different imagery source, open the dropdown menu by clicking where it currently says **"Mapbox Satellite"**, then click the dataset you would like to view, as shown in the image below. You may need to click the **"Update Imagery"** button after selecting your new dataset.

To interpret a sampling location, click on one of the black circles within the assessment unit. The circle you clicked will now be colored blue. Then, in the bottom right corner of your screen, select one of the available land cover/use classes. This location will then be highlighted as the color that you assigned to that class during project creation. As you can see in the image below, the sampling location in the upper left corner of the assessment unit is highlighted red because this pixel was classified as "Impervious Surface" (which is, of course, incorrect). Repeat this process by clicking the other sampling locations within your assessment unit.

If multiple pixels within your assessment unit are of the same land cover class, you can use the “bulk-select” mode in Collect Earth Online to classify multiple pixels at once. To use bulk select mode, press CMD + click + drag on Mac or ctrl + click + drag on PC to draw a rectangle on your screen. All of the sampling locations that fall within the box you drew will be highlighted in blue. Then click one of the available land cover classes to classify these points. You can also enter multi select mode by holding shift and clicking multiple sampling locations, then clicking a land cover class to classify the highlighted locations as that land cover class.

Confidence Slider

Below your survey questions, you will see the confidence slider. The confidence slider allows you to make a note of how confident you were of your classification. This feature is useful in group efforts where plots with low confidence can be reviewed as a team. Shift the slider by clicking the teal circle and dragging right to raise or left to lower the confidence. Place the slider at 100 if you are completely confident in your assessment. The Confidence Slider is indicated by the arrow in the image below.

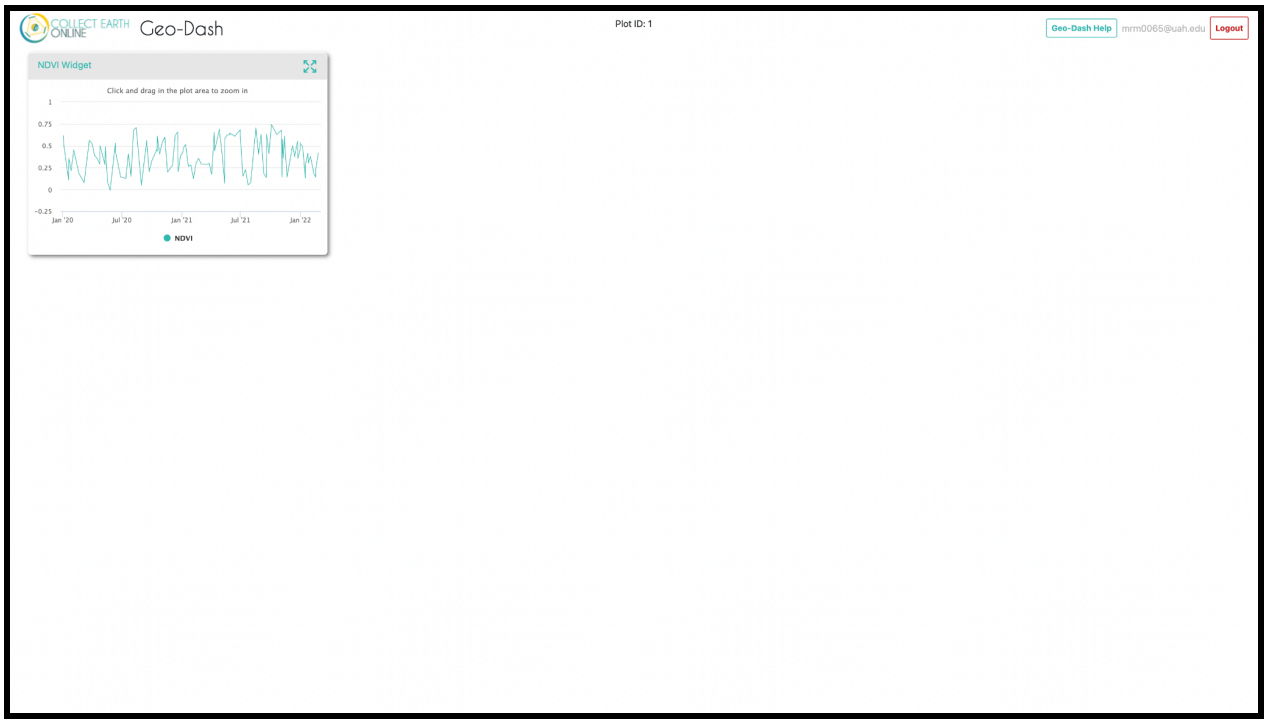
Flagging Plots

You can also flag plots within Collect Earth Online. When you flag a plot, no data will be collected for this plot. Flagging plots is useful if there is unavailable imagery for a specific plot due to cloud cover, haze, etc. To flag a plot, click the red **"Flag Plot"** button below the confidence slider. This button is indicated by the arrow in the image below.

After clicking this button, you will see a popup saying **"This plot has been flagged"** on the right side of the screen. In the textbox under the text that says **"Flagged Reason"**, you may type in the reason you flagged this plot. This textbox is indicated by the arrow in the image below. ***Then, click the "Save" button towards the bottom right of the page.*** The **"Save"** button is located directly below the **"Flag Plot"** button, shown in the image above. Clicking the **"Save"** button will automatically bring you to the next plot.

Geo-Dash Widgets

If you have enabled Geo-Dash Widgets for your project, the Geo-Dash screen will appear in a separate tab in your browser, to the right of the tab hosting the Collect Earth Online Project. This tab will appear when you click the **"Go to first plot"** button, which we covered on page 9 in the Navigating to the Data Collection page in your Project section. GeoDash will also automatically display after you collect data for each point and travel to a new plot. This GeoDash screen will look similar to the screenshot shown below.



At the top of the screen we can see the Plot ID of the data we are viewing, and in the top left you can see your widget(s). Clicking and dragging along the graph will zoom in, allowing you to see the values in greater detail, as shown in the image below. Mousing over the line on the graph will show you the specific values at different points in time.

Importing Data from Collect Earth Online

In order to download the data you collected for your project, visit the "ITC: NRM2" Institution homepage in Collect Earth Online. [Click here to visit our Institution homepage.] After visiting this link, make sure you are logged in by clicking on the **"Login/Register"** button in the upper right corner. Your screen will now appear similar to the screenshot shown below (except you will see more projects underneath the "Maganini_Test_Project" Project). Click on your project.

You will see several icons to the right of your image interpretation project. Clicking the **"P"** icon will download the plot data for your project, whereas clicking the **"S"** icon will download the sample data for your project. The difference between the plot data and sampling data is that the plot data will download the data for your assessment unit as a whole, whereas the sampling data will capture all of the points classified in your project. Downloaded data from CEO will be in the WGS84 EPSG:4326 projection.

Plot Data

Downloaded Plot Data" (which can be obtained by clicking the **"P"** button to the right of your image interpretation project) will have the following columns in a .csv (Comma-separated value) file.

I. **PLOT_ID**: the CEO-assigned unique sample plot number or the user-provided Plot ID (if you uploaded a .csv or .shp file in the sampling design step)
II. **CENTER_LON** and **CENTER_LAT:** geographic coordinates of the center of your sample plots.
III. **SIZE_M** and **SHAPE**: size in meters and the shape (circle or square) of the sample plot.
IV. **SAMPLE_POINTS**: the number of samples in each plot.
V. **EMAIL**: the email address of the person that classified this plot.
VI. **FLAGGED**: will be "FALSE" for plots that were not flagged, will be "TRUE" for plots where a user has flagged due to lack of available usable imagery (due to cloud cover, haze, etc.)
VII. **FLAGGED_REASON**: A user-provided reason for flagging the plot, if flagged.
VIII. **CONFIDENCE**: User confidence percentage.
IX. **COLLECTION_TIME**: date and time (in UTC) the user collected the plot data
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_DATES**: Number of SecureWatch dates used
XIII. If you used a .csv or .shp 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 of the following columns will have information about each of the survey questions, broken down by answer. They are labeled as **"QUESTION TEXT:ANSWER TEXT**." For example, a column titled "LULC: Built Surface" indicates that "LULC" was the question and "Built Surface'' was the answer. These are quantified as a percent of the sample points in the plot that were assigned that answer.
    A. For example, suppose you have four sample points within each plot (i.e. assessment unit) and two answers (e.g. land cover class) to choose from. If one sample point is assigned to one answer and the other three points to the second answer, the downloaded data will say '25' for the first answer and '75' for the second answer.

Sample Data

Downloaded Sample Data (which can be obtained by clicking the **"S"** button to the right of your project) will have the following columns in a .csv (Comma-separated values) file:

I. **PLOT_ID**: the CEO-assigned unique sample plot number or the user-provided Plot ID (if you uploaded a .csv or .shp file in the sampling design step)
II. **SAMPLE_ID**: the CEO-assigned unique sample point number.
III. **LON** and **LAT**: are the geographic coordinates of the center of your sample points.
IV. **EMAIL**: the email address of the person that classified this plot.
V. **FLAGGED**: will be "FALSE" for plots that were not flagged, will be "TRUE" for plots where a user has flagged due to lack of available usable imagery (due to cloud cover, haze, etc.)
VI. **COLLECTION_TIME**: date and time (in UTC) the user collected the plot data
VII. **ANALYSIS_DURATION**: Amount of time, in seconds, the user spent analyzing the plot
VIII. **IMAGERY_TITLE**: Name of the imagery layer that the user had selected while analyzing the plot
IX. **IMAGERY ATTRIBUTIONS:** Any attributions for the imagery 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. All of 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.

Now that you have your downloaded data, you can transform the data into your final product. If you are doing an inventory style land cover survey, this will include analyzing the statistical patterns in the data, which can be done by importing the .csv file into Python or another coding environment. If you collected training data for a machine learning model, your next step would be to upload the data into a geospatial computing environment such as Google Earth Engine. Whatever your end goal may be, Collect Earth Online allows you to obtain the reference data needed to accomplish your project's objectives. Happy data collecting!

External Links

External Link 1: The Collect Earth Online Data Collection Manual provides a more in-depth look at collecting data for an image interpretation project. Click here to view the Collect Earth Online Data Collection Manual.

Acknowledgements

Collect Earth Online has received financial support from NASA, The U.S. Agency for International Development (USAID), SERVIR, the Food and Agriculture Organization (FAO), the U.S. Forest Service, SilvaCarbon, Google, and Spatial Informatics Group. It was co-developed as an online tool housed within the OpenForis Initiative of FAO.

Collect Earth Online was initially developed by SERVIR, and is now supported by a broad base of partners. CEO was inspired by Collect Earth, a desktop software developed by FAO. The development team includes Arthur Luz, Jordan Combs, Matt Spencer, Richard Shepherd, Oliver Baldwin Edwards, Sif Biri, Roberto Fontanarosa, Francisco Delgado, Githika Tondapu, Billy Ashmall, Nishanta Khanal, John Dilger, Karen Deyson, Karis Tenneson, Kel Markert, Africa Flores, Emil Cherrington, and Eric Anderson.

The Collect Earth Online curriculum was organized by SERVIR's Science Coordination Office with individual modules created by NASA's Earth Observatory, the Spatial Informatics Group, and SERVIR SCO. Individual modules were developed by Crystal Wespestad (Spatial Informatics Group), Holli Riebeek (NASA Earth Observatory), Robert Simmon (NASA Earth Observatory), Billy Ashmall (SERVIR Science Coordination Office) Micky Maganini (SERVIR Science Coordination Office), NASA Earth Observatory, NASA, and the US Agency for International Development. Review of the material was conducted by SERVIR's Science Coordination Office, specifically Kelsey Herndon, Emil Cherrington, Billy Ashmall, Diana West, Katie Walker, Lauren Carey, Jacob Abramowitz, Jake Ramthun, Natalia Bermudez, Stefanie Mehlich, Emily Adams, Stephanie Jimenez, Vanesa Martin, Alex Goberna, Francisco Delgado, Biplov Bhandari, and Amanda Markert. Crucial insight regarding the development of the curriculum materials was provided by Claudia Paris and Andrea Puzzi Nicolau.

Review of the material was also conducted by Bart Krol and Laura Cray of ITC (The Faculty of Geo-information Science and Earth Observation at the University of Twente). The course and unit banner images were created by Gianluca Ambrosi of ITC.

Sources

- Development Team: https://sams.servirglobal.net/detail/7
- All other info: https://www.collect.earth/about/