Documentation of GLOBE Eclipse Data For 2023 and 2024

Before major solar eclipses, GLOBE Observer adds a special limited time tool to the app to allow citizen scientists observers to submit air temperature data in addition to the usual data types available. For eclipses after the initial 2017 event, the appearance of the GLOBE Eclipse tool has been constrained geographically to roughly the area experiencing the eclipse. More details about the previous eclipses are available on the GLOBE Eclipse Data page. The description below provides details specifically about the data for the 2023 Annular Solar Eclipse and 2024 Total Solar Eclipse.

Eclipse Path and Tool Visibility

The 14 October 2023 annular solar eclipse began in the United States, traveling from the coast of Oregon to the Texas Gulf Coast. Weather permitting, the annular eclipse was visible in Oregon, Nevada, Utah, New Mexico, and Texas, as well as some parts of California, Idaho, Colorado, and Arizona. The annular eclipse continued to Central America, passing over Mexico, Belize, Honduras, and Panama. In South America, the eclipse traveled through Colombia before ending off the coast of Natal, Brazil, in the Atlantic Ocean. The GLOBE Eclipse tool was made accessible in the GLOBE Observer app for a polygon roughly covering the area where any portion of the annular eclipse would be visible, represented by this GeoJSON file for 2023 (also available on the GLOBE Eclipse data page and as text in the appendix below.) Data collection of air temperature via the tool was only open on 14 October itself.

On 08 April 2024, the total solar eclipse began on land in Mazatlán, Mexico, with the path of totality extending from there through parts of the states of Texas, Oklahoma, Arkansas, Missouri, Indiana, Ohio, Pennsylvania, New York, New Hampshire, Vermont and Maine, as well as parts of the provinces of Ontario, Quebec and New Brunswick. At least a partial eclipse was visible (weather permitting) in nearly all of the United States (including Hawaii, although only the very southeastern portion of Alaska), Canada, Central America and the Caribbean. In addition, the end of the eclipse was visible further east in the Atlantic, including a partial eclipse in southern Greenland, Iceland, Ireland and the northern United Kingdom, right before sunset. For the total solar eclipse, the GLOBE Eclipse tool was made accessible in the GLOBE Observer app for a polygon roughly covering the area where a partial eclipse would be visible, represented by this GeoJSON file for 2024 (also available on the GLOBE Eclipse data page and in the appendix.) Data collection of air temperature via the tool was open from 01 April to 08 April.

NASA has several sources of information about the eclipse paths in 2023 and 2024, such as the NASA Eclipse Explorer, and visualizations and shapefiles from NASA's Scientific Visualization Studio for the 2023 annular eclipse and 2024 total solar eclipse.

General Information About the Data

For the curated datasets on the GLOBE Eclipse data page, the GLOBE Application Programming Interface (API) was used to retrieve the data, specifically the option to get data constrained by a

GeoJSON shape. The shapes where the app was visible for the eclipse were used for the data retrieval (see details in the Eclipse Path and Tool Visibility section above.) Data are made available in JSON and GeoJSON file formats directly from the API, and also converted into comma separated value (CSV) and Excel (XLSX) file formats. The date the data were retrieved from the API is also noted in the filename.

For the CSV and XLSX files, to promote ease of use by those less familiar with GLOBE data the following adjustments were made to the raw data pulled from the API. The XLSX format preserves some formatting that may make the files easier to work with.

1. The columns were renamed to remove the duplicative “air_temps,” “sky_conditions” (for clouds), “land_covers,” and “winds” prefixes before the more descriptive column headers.
2. The latitude, longitude, and elevation columns were renamed to siteLatitude, siteLongitude, and siteElevation, to distinguish between those site values based on the Military Grid Reference System for automatically created sites in the app and the measured values from the device. More details about how those site locations are calculated is available in the GLOBE Data User Guide in Section 1.2.4 Spatial Characteristics. Measured values are only available for clouds and land cover data, not air temperature or wind, but the headers were changed for all three data types for consistency within the eclipse dataset.
3. A dataId column was added, providing a unique identifier for each data point.
    a. The format for the dataId is as follows, using the columns with the given identifiers, plus the measuredAt date reformatted for year, month, day, hour and minute format: protocol-siteId-userId-YYYYMMDDhhmm
    b. As an example, for a clouds observation (protocol=sky_conditions) taken at site ID 262095, by user ID 85840211, on August 9, 2023 at 1938 UTC the dataId will be: sky_conditions-262095-85840211-202308091938
4. For the globeTeams column, the data from the API contains a list of GLOBE Teams an individual user was a member of at the time of data collection, surrounded by square brackets []. For the eclipse data, a few manual tags were also added to this column for users who weren’t members of any GLOBE Teams, noting membership in significant groups the GLOBE Observer team was working with during the eclipses, as identified by the organization name. Currently this extra classification has only been done for clouds and air temperature, not land cover or wind. The extra notations are:
    a. *CAP Org - No GLOBE Team*: For Civil Air Patrol squadrons participating in data collection but without their own GLOBE Team or who didn’t join an overarching CAP team
    b. *Library Org - No GLOBE Team*: For libraries participating in data collection but without their own GLOBE Team, or who didn’t join the overarching GLOBE Eclipse Libraries Team
    c. *TMN - No GLOBE Team (Eclipse Educators):* For Texas Master Naturalists who didn’t join the overarching Eclipse Educators team for the organization.

Compared to the order when pulled from the API, columns in the CSV/XLSX eclipse area datasets were reorganized to be in more logical groupings. Protocol and observation identifiers (dataId/observationId) were placed at the beginning, followed by date and time of measurement, and date and time of solar noon if included (clouds and air temperature). Following that come the

columns with country, organization, and GLOBE Teams information, and then site information (ID, name, latitude, longitude and elevation.) For the protocols with measurement latitude and longitude (land cover and clouds) and location source and accuracy (land cover only), those columns are next. Clouds and land cover data also have a data source field, indicating data collection via the GLOBE Observer app, online data entry forms, or email data entry.

Clouds data also indicates whether the user has been through additional GLOBE training beyond the in-app training, and if they are considered a “citizen science” user versus a GLOBE Educator. Following the userId column are the protocol-specific columns. In general photo URL fields, comments or field notes fields, and similar types of data values were grouped together (such as cloud types observed at different altitude levels and any land cover classification information.) More details about the GLOBE data more broadly, and specific descriptions of each of the protocol data values, can be found in the GLOBE Data User Guide.

Additional specific details related to each type of data in the eclipse dataset are described below.

Air Temperature Data

GLOBE users who are trained can always enter air temperature data, either through Atmosphere Data Entry within the GLOBE app or via online data entry forms. The GLOBE website has more detail about the Air Temperature Protocol.

For the special Eclipse tool, training was not required to submit air temperature data. The instructions emphasized that a separate thermometer was needed, and to not use a weather app on the phone. As part of the initial settings when first opening the tool, users were asked to confirm the type of thermometer and given the options of liquid filled, digital, weather station and other (although the responses were not stored with the data.) Participants were given reminders (including a countdown in the tool and a pop-up message) to collect air temperature data every ten minutes until 30 minutes before maximum eclipse, at which point the reminders increased to every five minutes until 30 minutes after maximum eclipse, when the frequency returned to every ten minutes.

Air temperature values reported through the Eclipse tool (by trained or untrained users) can be distinguished in the GLOBE Application Programming Interface (API) by the field labeled as “airtempsComments” and in the GLOBE Advanced Data Access Tool (ADAT) by the field “air temps:comments.” Data points with comments in the format {"temp":"36.2","units":"C"} were entered via the Eclipse tool. The unit value will be C or F depending on whether the app user reported measurements in Celsius or Fahrenheit although temperatures are converted as needed and stored in Celsius only in the “airtempsCurrentTemp” or “air temps:current temp (deg C)” fields for the API/ADAT respectively. Other values in this comments field may indicate automated weather stations, such as (Phase 4 Email Data Entry), Earth Networks, or Weatherbug. Trained GLOBE users entering data via other means (online web forms or the Data Entry app, which was incorporated into GLOBE Observer for the atmosphere protocols in July 2021) also have the opportunity to enter custom comments associated with the data, including observations about current conditions.

Several columns were added specifically for air temperature with additional information, based on looking up the site information in the GLOBE database. This is primarily relevant for automated

weather stations or other data collection sites set up for trained users. Less information is available for those using the GLOBE Eclipse tool. These columns are.

1. sensorType: providing general information about the type of sensor used to collect the air temperature data. The options are based on the instrument choices available when a site is set up in the GLOBE system, plus a separate determination of weather stations using the site names or other details provided.
    a. AWS: Automated Weather Station
    b. Data Logger: specified by trained user during site setup
    c. Digital Multi-Day: specified by trained user during site setup
    d. Digital Single-Day: specified by trained user during site setup
    e. Liquid Current Temp: specified by trained user during site setup
    f. Other/Unknown: not enough information was provided during site setup to specify
    g. Other (via Eclipse Tool): a known user of the Eclipse tool, but no other details available
2. sensorDetails: if further description beyond the general sensor type categories was provided in the site details, it's included here (ex: Davis weather station, Kestrel data logger, etc.)
3. siteDetails: a summary of any information provided by the user about the location, including surface cover, building or obstacles, height of thermometer, etc.

Clouds Data

During the eclipse, users were directed to use the Clouds tool and use the regular process of taking a cloud observation. In the Eclipse tool, reminders were given to take cloud observations every 30 minutes except for the period a half hour before and after maximum eclipse, during which the reminders were every 15 minutes. More detail about the Clouds Protocol can be found on the GLOBE website.

In addition to the dataId column added to the CSV and XLSX file formats, clouds data (called sky_conditions in the API), has an identifier called observationID which is used for the satellite comparison process done by the GLOBE Clouds team. The format of that observationID was used as the basis for the dataId added for the other data types, so they are very similar values for each observation. Additional information about the satellite comparisons for clouds, including curated clouds datasets, can be found on the Clouds Data page. Both values are included in the eclipse curated datasets for clouds.

Land Cover Data

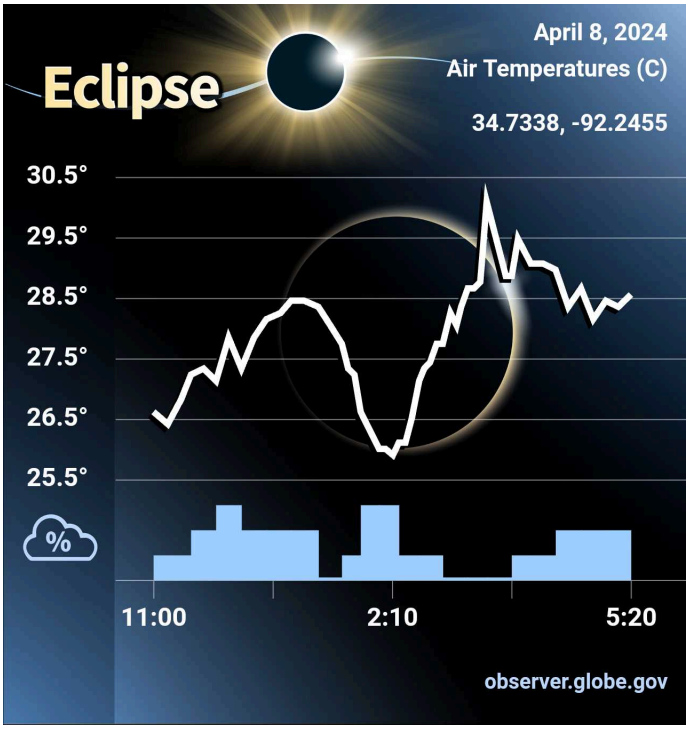
New for the eclipses in 2023 and 2024 compared to previous eclipse events, users were asked to take a single observation with the Land Cover tool at their data collection site, to set the scene for the other types of data collected. A request for this data was included as part of the initial setup when the Eclipse tool was opened for the first time, as well as in a single pop-up message. Other than those prompts, the data was collected as described on the Taking Land Cover Observations page.

Wind Data

The Wind protocol as part of the GLOBE Program is not as fully developed as most other protocols, and is primarily used for automated weather station data reported to the database. Wind data included as part of the eclipse curated data sets therefore comes from automated weather stations and a few trained users with handheld anemometers who entered measurements through Data Entry, either in GLOBE Observer or on the website, or through email data entry. It is included in the eclipse datasets because it may prove useful for researchers. Aside from the renaming/reorganization of columns as described in the general information section above, the only other change in the CSV/XLSX wind eclipse dataset compared to the original API pull was to remove the "protocolId" which had a numeric code designation, which was duplicative of the "protocol" column with the label *winds* and which was not present in the other data types included in this dataset.

"Eclipses" Data in the API

Users accessing data through the GLOBE API will notice an "eclipses" protocol available to download. While this does include the air temperature data collected via the GLOBE Eclipse tool in the form of an array with measured_at and current_temp values, it was primarily used to generate the graph for display to users in the app on the day of eclipse data collection. It may be useful if one wishes to access URLs to the images of those graphs (example below), but all the reported temperature values are also available as separate data points via the air temperature protocol as noted above, and are easier to use for analysis in that form.



Documentation prepared by Kristen Weaver, reviewed by Marilé Colón Robles and Holli Kohl, last updated 15 August 2025.

Appendix

GeoJSON format for shape where the app appeared during the 2023 eclipse:


```
{ "type": "FeatureCollection",
   "features": [
     {  "type":"Feature",
        "properties": { },
        "geometry": {
          "coordinates": [
            [-64.14121403490442,81.57392677293225],
            [-84.63884094874881,82.51883449192306],
            [-152.23502186243678,70.82051084577648],
            [-161.35033362166016,58.83619431724563],
            [-72.75885928657947,-36.06805167698425],
            [-60.05864923152744,-39.140261070928936],
            [-8.100793138130598,-33.91910211762262],
            [-55.681153561432666,61.73863797892909],
            [-42.60229438951197,80.22276202782226],
            [-64.14121403490442,81.57392677293225]
          ],
          "type": "LineString"
        }
     }
   ]
}
```


GeoJSON format for shape where the app appeared during the 2024 eclipse:


```
{  "type": "FeatureCollection",
   "features": [
     {  "type": "Feature",
        "properties": { },
        "geometry": {
          "coordinates": [
            [-171.68853321276669,32.10750494843059],
            [-133.93881481315074,-30.0595299457743],
            [-79.39313459872736,3.641910904536644],
            [-74.7528367528288,12.816715403289834],
            [-26.13058541102049,13.602364151972267],
            [-6.762385706400067,48.423706586099996],
            [0.5006891828324456,62.52619222856205],
            [30.35999706078897,75.57615624940522],
            [32.17576578309729,81.25307984618925],
            [-28.95511453461131,84.4202875927736],
            [-75.10735037753747,84.2635655128432],
            [-113.4402456262652,81.46885436632431],
            [-136.0049394627668,68.91368721769882],
            [-141.27837696276643,59.804384124844745],
            [-171.68853321276669,32.10750494843059]
          ],
          "type" : "Polygon"
        }
     }
   ]
}
```