

Data Challenges and More: Using Events to Support Citizen Science with NASA and the GLOBE Program

Authors: Kristen Weaver, Brian Campbell, Marile Colon Robles

App-based Earth science observations during a total solar eclipse. A month-long clouds data collection challenge promoted via social media. A full-year student research campaign with regular webinars featuring scientists. All of these are examples of using special events to recruit and support volunteer observers, and for collecting scientifically usable data. But how do you judge the effectiveness of these techniques? Does an event produce more usable data than during non-event periods? Do citizen scientists who join for an event stick around and continue participating in the project? We will give examples of how the GLOBE Program and the NASA GLOBE Observer app team have used various events to draw in new citizen scientists and promote concentrated data collection, and share lessons learned regarding events as recruitment tools and in terms of retention of participants after the event concludes, as well as adaptations to event processes necessary during the coronavirus pandemic.



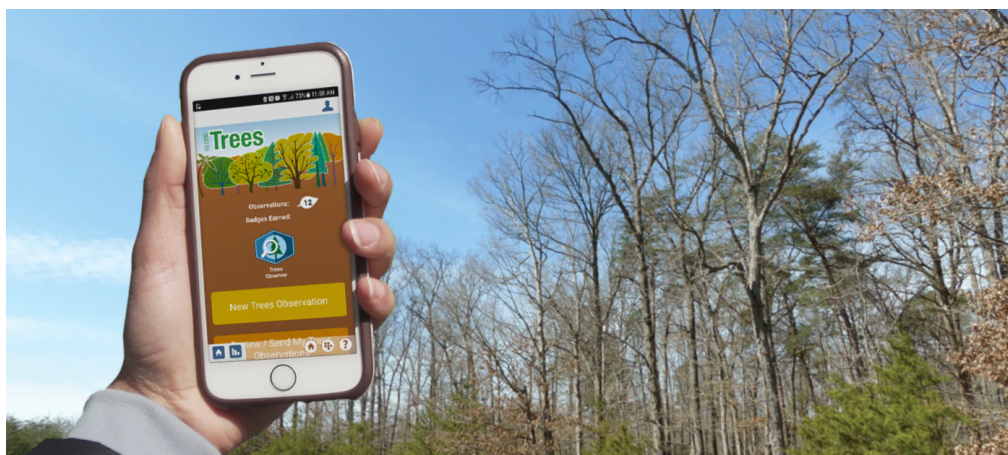
Introduction

The Global Learning and Observations to Benefit the Environment (GLOBE) Program has a long history of involving students in Earth science data collection and analysis as an education and citizen science initiative in existence since 1995. In 2016, GLOBE expanded to include citizen scientists in GLOBE countries in data collection through GLOBE Observer (GO), a smart phone app-based program coordinated by the NASA Earth Science Education Collaborative (NESEC) and supported by the GLOBE Implementation Office. The app is free to download, although an email account is needed to register. All of the needed training is contained with the app, allowing observations of clouds, land cover, tree height and mosquito habitats. Basic observations require nothing beyond the smartphone itself, although some optional data requires additional equipment (e.g. a measuring tape for tree circumference, or a clip-on magnifier to identify mosquito larvae). A

special GO tool that allows reports of air temperature using an external thermometer is also released on a periodic basis around total solar eclipse events.

The goals for GLOBE Observer include.

- Increase GLOBE data density both spatially and temporally through audience expansion (beyond formal education)
- Increase student and scientific research. This means that scientists and students can access data and assess data quality, and that the data collected are useful to the scientific community
- Participants feel they are part of a community or a bigger purpose



Using Events to Encourage Data Collection

The GLOBE Observer team has used multiple events to drive increased data collection, in two main categories:

- Data collection challenges
 - These tend to be more broad-based, using mostly social media to get the message across, and generally require short-term participation (a month seems to be the sweet spot, although some have been longer or shorter)
 - Examples of this type of challenge include data collection during total solar eclipses (starting with August 2017 in North America, but also smaller events surrounding the July 2019 and December 2020 eclipses in South America) seasonal Cloud Challenges, GO on a Trail, and the current Community Trees Challenge, running through May 15th, 2021.



- Student research campaigns
 - These are more focused on school-based GLOBE participants, connecting with teachers and students through webinars, and entail more sustained participation.
 - Examples of this type of event are [GLOBE Mission Mosquito](#) and [Trees Around the GLOBE](#), as well as campaigns such as the U.S. Air Quality campaign, which is supported by NESEC team members, but uses protocols outside the GLOBE Observer app.



Judging the Effectiveness of Events

Do challenges increase the quantity of data collected?

The criteria for judging the effectiveness of data challenges versus other types of campaigns are not necessarily the same, although there is some overlap. One of the primary goals of both types of events is to increase data collection of a particular type, or in a particular place, so let's start with pure numbers.

In figure 1 from Amos, et al. (2020), the spikes in observations from the GO app during events is quite visible. It's worth noting that the y-axis is cut off at 3,000 observations per day, but on the day of the 2017 North American Total Solar Eclipse more than 18,000 observations were submitted. While all four GO tools are included in this chart, Clouds is the main driver, as the most popular tool, and the focus for each of the challenges listed in this time frame.

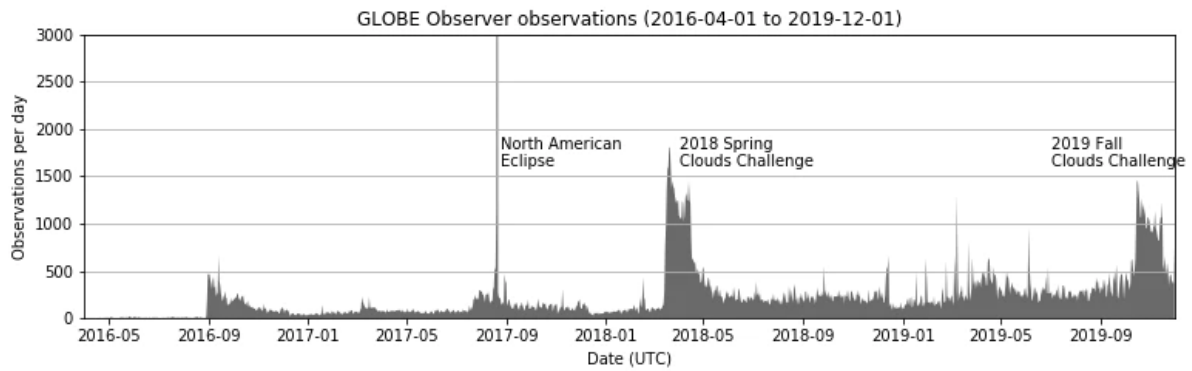


Figure 1. Time series of observations made with the NASA GLOBE Observer mobile app from 1 April 2016 to 1 December 2019 (Fig. 4 from Amos et al., 2020).

Challenges drive observations by existing users, but also the creating of new accounts. Figure 2 from Colón Robles, et al. (2020) noted that over 27,000 new accounts were created during the Spring Clouds Challenge in 2018. Sub-spikes in that time frame may be correlated with specific social media and other news articles.

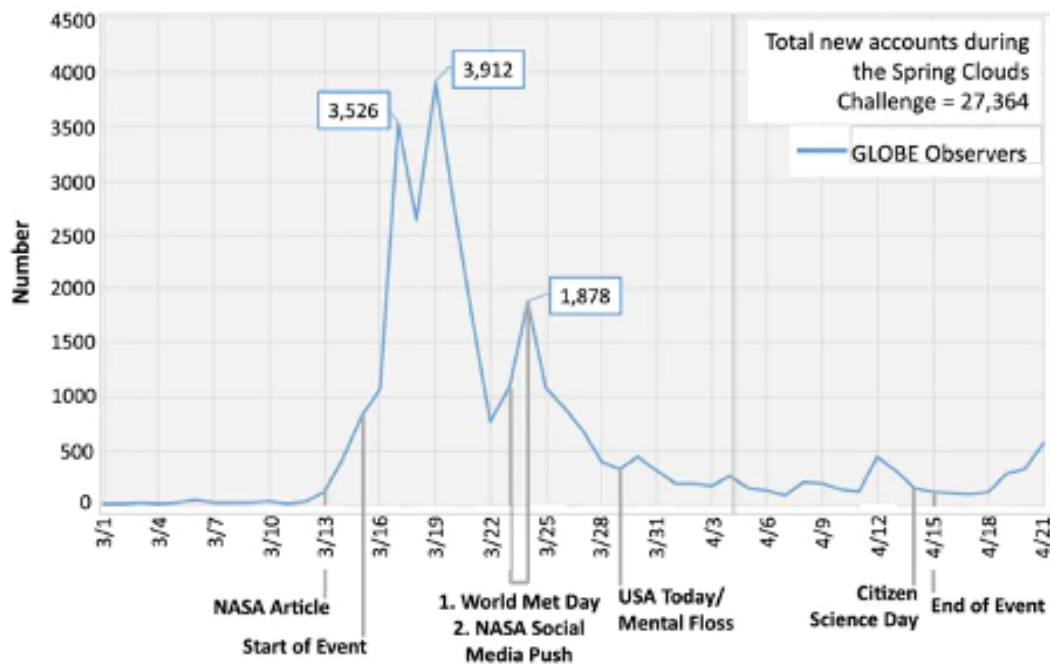


Figure 2. Number of new app users. The spikes correlate with the aftermath of major product releases (NASA article, 13 March, and World Met Day, 23/24 Mar). Also, the NASA article (released on 13 Mar 2018)

continued to gain momentum through sharing by other media outlets [i.e., Forbes (28 Mar 2019), Mashable (28 Mar 2018), ABC (29 Mar 2019), USA Today (29 Mar 2019), Mental Floss (29 Mar 2019), and more]. The total number of GLOBE Observer accounts increased by over 30,000 from 74,954 accounts on 1 Mar 2018 to 106,128 total accounts by 30 Apr 2018 (Fig. 8 from Colón Robles, et al., 2020).

Is there an observable impact due to COVID-19 pandemic?

More recently, the COVID-19 pandemic changed the way the GO team thinks about challenges. Previously, we emphasized observations heavily, and recognized top observers in each GLOBE region. However, since not everyone is able to get outside to take measurements during these times, starting with the Community Clouds Challenge in 2020, we changed the way we promoted the challenges to emphasize other ways to participate beyond just data collection. More about that can be found on the poster "[Beyond Data Collection: How COVID-19 Changed How We Do Data Challenges](#)" here in CitSciVirtual.

In figure 3 (data pulled by month from the [GLOBE API](#)), the months that contained part of a challenge are noted in green, and it is clear that the most recent clouds challenge in 2020 has lower numbers than the previous two challenges, and in some cases lower numbers than other non-challenge months.

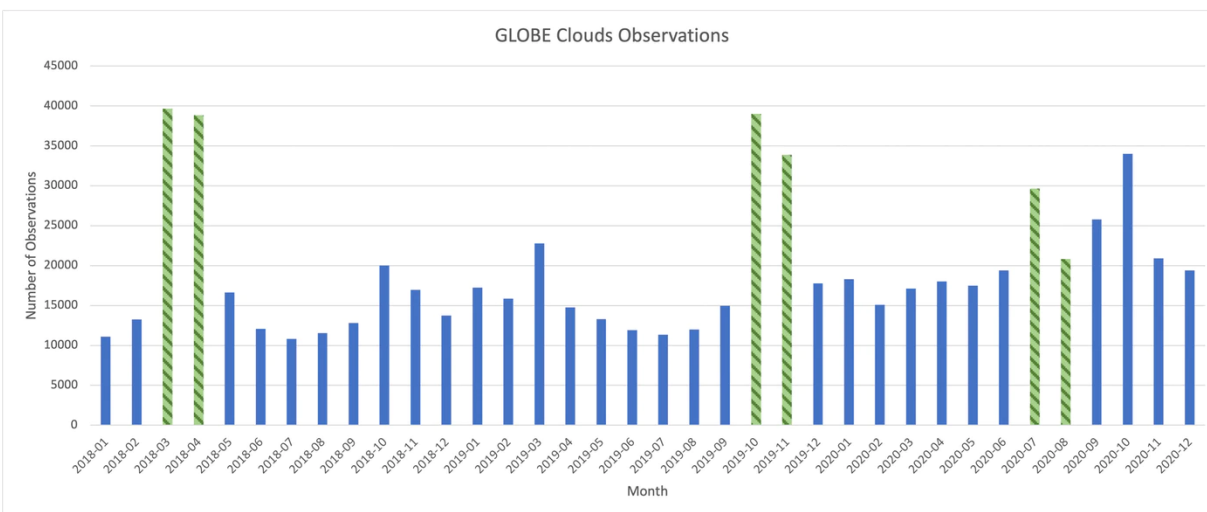


Figure 3. GLOBE Cloud observations per month from January 2018 to December 2020.

However, by separating out the years and looking at the average number of clouds observations by month in challenge versus non-challenge months (figure 4), it's possible to see that even in these unusual times, challenges increase the number of observations submitted compared to non-challenge periods.

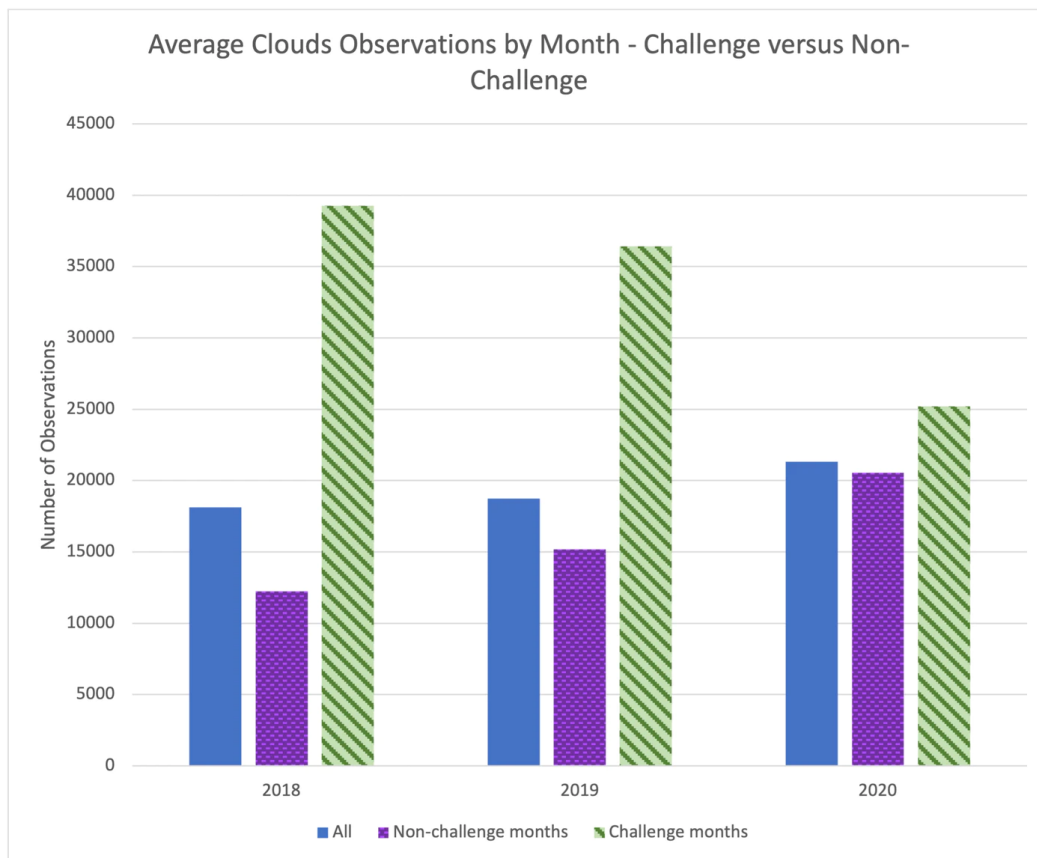


Figure 4. GLOBE Cloud observations averaged per month, highlighting challenge months.

How student research campaigns compare to data challenges?

The impact on data collection of a student research campaign can be a bit harder to ascertain, since they tend to have smaller numbers of participants, but more in-depth involvement. For example, the Trees Around the GLOBE Student Research Campaign that began in September 2018, has held 32 regular webinars (plus 8 special ones) with overall attendance of over 2,000 participants (primarily students and teachers). In addition, 40 blogs have been published receiving more than 110,000 views on the website. In the campaign time frame thus far (September 2018 to April 2021), there have been nearly 33,000 tree height measurements and over 18,000 each of land cover and green-

up/green-down observations, the protocols covered by the campaign. However, while it is clear that there is a great deal of engagement, it's harder to tell how much of the data collection itself can be attributed specifically to campaign activities.

We do know that there are project reports that have been submitted the GLOBE International Virtual Science Symposium (IVSS) that are supported by campaign efforts. There is not currently a checkbox upon submission of such reports to indicate involvement in the Trees Around the GLOBE campaign, but there is for GLOBE Mission Mosquito. In the 2021 IVSS cycle, 51 student reports were submitted as connected with that campaign, many through the [STEM Enhancement in Earth Science Mosquito Mappers](#) virtual internship program that NESEC supports.

Does an event produce more usable data than during non-event periods?

This is a question that still need more investigation. In addition to Colón Robles, et al. (2020), two other papers have examined GLOBE challenge data sets in an in-depth way: Dodson, et al. (2019) on the 2017 North American Solar Eclipse and Kohl, et al. (2021) on GO on a Trail challenge. Dodson, et al. found that data from greater numbers of participants did reduce the uncertainty of the measurements, in that case temperature depressions during the 2017 total solar eclipse reported using the special GO Eclipse tool. Dodson et al. state that “group sizes larger than 50 are desirable for giving results robust enough for the amplitude of the cloud-temperature relationship to be reliably detected” referring to their Monte Carlo test results (p. 2383). Kohl, et al. discussed the characteristics of the GO on a Trail challenge data set of land cover observations collected from June to October 2019, especially examining the location accuracy of the geolocated

photos and the quality of volunteer-assigned land cover classification labels, but did not compare the data to a non-challenge time frame. More investigation in this area is needed.

Do citizen scientists who join for an event stick around and continue participating in the project?

The GO team is very interested in pursuing this question further, but does not yet have definitive answers. A survey conducted with GLOBE Observer users in 2019 and reported on as part of the recent paper “Going Beyond Hooked Participants: The Nibble-and-Drop Framework for Classifying Citizen Science Participation” by Fischer, et al. 2021, found that the top reason cited for ceasing contributions was that the user had signed up for a specific event and did not feel the need to continue. The authors speculate that these users were never actually fully recruited and “hooked” by the project, so could perhaps be re-recruited to help participant retention. While not directly related to this event discussion, the same article comments that another common reason cited for dropping participation in GO was not feeling like they were making a meaningful contribution. Specifically, that there was not enough feedback given about observations. The GO team is working on ways to improve our communication with participants, both in general and as a follow-up to challenges and other events, and can hopefully increase retention numbers in future.

Summary

In conclusion, data challenges bring a large number of observations compared to normal observation periods, even during COVID-19

pandemic. Although large number of new users join during a data challenge, a drop in observers is noted once the event is over. Student research campaigns bring in observations, but their most noticeable contribution is student projects and investigations.

The work done by the GO team shows that multiple venues for groups to be involved with GLOBE contribute to number of observations and student research projects, more work needs to be done to best keep new observers within the projects and to analyze data from data challenges beyond clouds.

References:

Amos, H. M., Starke, M. J., Rogerson, T., M., Colón Robles, M., Andersen, T., Boger, R., et al. (2020). GLOBE Observer data: 2016–2019. *Earth and Space Science*, 7(8), e2020EA001175.

<https://doi.org/10.1029/2020EA001175>

Colón Robles, M., Amos, H. M., Dodson, J. B., Bouwman, J., Rogerson, T. M., Bombosch, A., Farmer, L., Burdick, A., Taylor, J. E. & Chambers, L.H. (2020). Clouds around the world: How a simple data challenge became a worldwide success. *Bulletin of the American Meteorological Society*, 101(7), E1201–E1213. <https://doi.org/10.1175/BAMS-D-19-0295.1>

Dodson, J.B., Colón Robles, M., Taylor J.E., DeFontes, C.C., Weaver K.L., (2019). Eclipse Across America: Citizen Science Observations of the 21 August 2017 Total Solar Eclipse. *Journal of Applied Meteorology & Climatology*, 58(11). <https://doi.org/10.1175/JAMC-D-18-0297.1>.

Fischer, H., Cho, H. and Storksdieck, M., 2021. Going Beyond Hooked Participants: The Nibble-and-Drop Framework for Classifying Citizen Science Participation. *Citizen Science: Theory and Practice*, 6(1), p.10. DOI: <http://doi.org/10.5334/cstp.350>

Kohl H.A., Nelson P.V., Pring J., Weaver K.L., et al. (2021). GLOBE Observer and the GO on a Trail Data Challenge: A Citizen Science Approach to Generating a Global Land Cover Land Use Reference Dataset. *Frontiers in Climate*, 3:620497. <https://doi.org/10.3389/fclim.2021.620497>

Greenbelt, MD · Updated Wed, April 28



Kristen Weaver

I'm the deputy coordinator for the GLOBE Observer project, based at NASA's Goddard Space Flight Center in Greenbelt, Maryland.