A Descriptive Analysis of GLOBE Data Collection Prior to and During the COVID-19 Pandemic.

Ksenia Lepikhina Denise Yost

July 15, 2021

The GLOBE Implementation Office
The GLOBE Program

Sponsored by: NASA Supported by: NSF NOAA Implemented by: UCAR

1 Introduction

The Global Learning and Observation to Benefit the Environment (GLOBE) Program is an international citizen science and education program. GLOBE provides students, scientists, and citizens worldwide with the "opportunity to participate in data collection and the scientific process, and contribute meaningfully to our understanding of the Earth system and global environment"[1]. The program collaborates with the community to promote science education, enhance environmental literacy and stewardship, and promote scientific discovery.

The GLOBE Program puts a heavy emphasis on data collection and distribution. The data collected includes earth observations such as tree height, cloud type, precipitation, etc. Training is required for all individuals who enter data into the GLOBE database and GLOBE partners and Country Coordinators typically conduct trainings for GLOBE protocols. Teachers are trained through in-person workshops and eTraining modules while citizen scientists receive training by following prompts in the GLOBE Program's App, GLOBE Observer. Training is a strong part of the Program and serves to ensure high quality data collection. Without training, entering data is not permitted. More information about the GLOBE protocols can be found on globe.gov. Scientists benefit greatly from the contributions of citizen scientists and students because they contribute complementary, usable data that enhances their projects. For example, these data are used to ground-truth matching satellite measurements, expanding the capacity for data collection and data quality that is otherwise limited by time and resource constraints.[2]

1.1 COVID-19 Pandemic

On March 11, 2020, the World Health Organization (W.H.O.) declared "SARS-CoV-2" (COVID-19) a pandemic – a globally disruptive event. As of December 31, 2020, 83.5 million cases and 1.85 million deaths globally and 20.06 million cases and 353.6 thousand deaths in the United States had been reported [3]. In the United States (US) at the end of May 2020, 36.5% of COVID-19 patients were admitted to the ICU (compared to 17.6% admissions for influenza patients in the same time frame) and 21% of COVID-19 patients died while hospitalized (relative to 3.8% influenza patients)[4]. Additionally the median hospitalization duration for COVID-19 patients by the end of May was 8.6 days relative to 3 days for the median duration of influenza patients[4].

In response to the fluctuating increase of COVID-19 cases since January 2020, public health officials[5] encouraged frequent handwashing, public mask wearing, social distancing by maintaining 6 feet of space (or approximately two meters) between people, and avoiding enclosed and crowded spaces – efforts that are known to reduce the spread of the virus[6]. The W.H.O. provided a list of quarantine suggestions for individuals who felt sick[7]. Nearly all countries instituted some form of lockdown, restrictions, or curfew either on a city, county, state, or country level[8].

1.2 Structure

This paper summarizes and provides a broad, descriptive analysis of GLOBE data submissions based on data entry types, GLOBE regions, and selected protocols over time. We discuss potential pandemic impacts on GLOBE data collection and submission and note that data collection and eTraining continued throughout 2020 despite the pandemic. In addition, the analyses of GLOBE data collection trends show that pandemic-related restrictions (e.g. mandated stay-at-home orders in 42 U.S. states and territories between March and May [9]) potentially decreased the quantity of citizen science data submissions received.

Although many citizen science organizations have reported success (in terms of increased community data submission) in 2020 despite the pandemic[9][10][11][12], few have tracked the change in data submissions over time through the end of 2020. This report aims to provide an overview of GLOBE data submissions throughout the pandemic. In addition to tracking submissions during the pandemic, GLOBE is uniquely positioned as an international citizen science and educational organization which allows for worldwide

comparisons of data entered during the pandemic. GLOBE is particularly notable because it collaborates directly with schools as well as citizens. This paper aims to provide information about GLOBE data in 2020 contextualized by COVID-19 and relative to previous years and across regions.

1.3 GLOBE Countries

As of December 31, 2020, the GLOBE Program had established formal partnerships, with the support of NASA and the Department of State, with 124 countries. Specifically, GLOBE regions include Africa (27 countries), Asia and Pacific (A&P) (18 countries), Europe and Eurasia (EE) (44 countries), Latin America and Caribbean (LAC) (20 countries), Near East and North Africa (NENA) (13 countries) and North America (NA) (2 countries). A map of the GLOBE regions and participating countries can be seen in Figure 1. A full list of the countries can be found on the GLOBE website and in the Appendix, Figure 1. A larger version of the map in Figure 1 can be found on the GLOBE website.

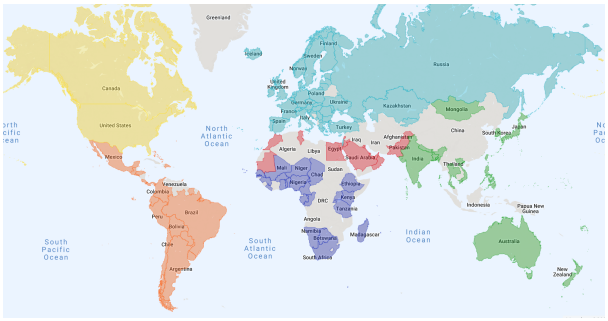

Figure 1: GLOBE Regions

GLOBE countries in the LAC region are represented by the color orange, GLOBE countries in the Africa region are represented by purple, GLOBE countries in the A&P region by green, GLOBE countries in the EE by blue and GLOBE countries in NA by yellow. These colors will be used to represent these regions throughout this paper. A color key for these regions can be seen in Figure 2.

The gradient colors within the color key represent individual countries in the GLOBE regions. Country specific data are not investigated in this paper and the country color gradient is used as a representation of countries within the regions.

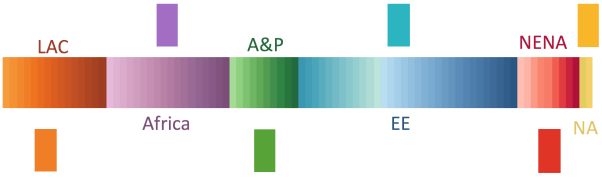

Figure 2: Region Color Key

2 Data

To conduct our analyses we used data from six different sources. The primary data sources were Our World in Data, The New York Times, Worldometers, and GLOBE. The first two data sources were selected because they track and update COVID-19 case information daily. Worldometers was selected as a source because it is "run by an international team of developers, researchers, and volunteers with the goal of making world statistics available in a thought-provoking and time relevant format to a wide audience around the world."[13]

The first COVID-19 dataset, compiled by Our World in Data[14], contains daily summary data for countries and territories such as the number of new COVID-19 cases, the number of new COVID-19 deaths, the total number of COVID-19 cases, the total number of COVID-19 deaths, weekly COVID-19 cases, weekly COVID-19 deaths, biweekly COVID-19 cases and biweekly COVID-19 deaths for each day in 2020. The second dataset contains similar COVID-19 information for each state in the United States, compiled by The New York Times[15]. The third dataset, also compiled by Our World in Data[14], contains each country or territory and it's current population estimate. The fourth dataset, found on Worldometers[13], containing country/territory area data was used to calculate population density.

The first GLOBE dataset is retrieved from the GLOBE database. This dataset consists of data submissions organized by protocol, submission type (the GLOBE Program's app (GLOBE Observer App), email data entry, mobile data entry, web forms, and automated weather stations), and by location (country and GLOBE re-

gion) for each day in 2020 as well as data on the number of e-trained or workshop trained individuals. The second GLOBE dataset consists of GLOBE ADAT (Advanced Data Access Tool) and Visualization Tool access statistics. GLOBE protocol, submission count, and location data are publicly available.

In the figures that follow in section 3, GLOBE data were analyzed from 2016 or 2017 through 2020. The primary reason for these selected years is that the GLOBE Observer App was released for Android on April 20, 2016 and on iOS on April 22, 2016. Some data submissions start in 2017 to include a full year of data and the resulting analysis includes data captured from 2017-2020. For all other analyses, the GLOBE data begins with January 2016.

Using these six data sources, we analyzed GLOBE data through the lens of the COVID-19 pandemic.

3 Results

Entry of data into GLOBE's databases requires training of both teachers and citizen scientists and data may be submitted in a number of ways. Data submission types include automated weather station measurements, web forms, email submissions, and App entries. Below we present trends in data entry over time as a whole and for each of these means of entry as there are differences in accessibility and utility by region and country. We examine data entry by region and by week in order to view yearly trends related to school year cycles and global events. Throughout the figures that follow we use a color scheme consistent to that presented in section 1.3 for region and country, as well as a color scheme for years that is consistent across all graphics.

3.1 Participating Countries

To begin our investigation into the impacts of COVID-19 on GLOBE data collection, we start by looking at the total number of "participating" countries by year. Here, we define "participation" as having submitted data at least once. By the end of 2020, NA had 2 participating countries, NENA had 7 participating countries, EE had 26, A&P had 9, Africa had 11, and LAC had 12.

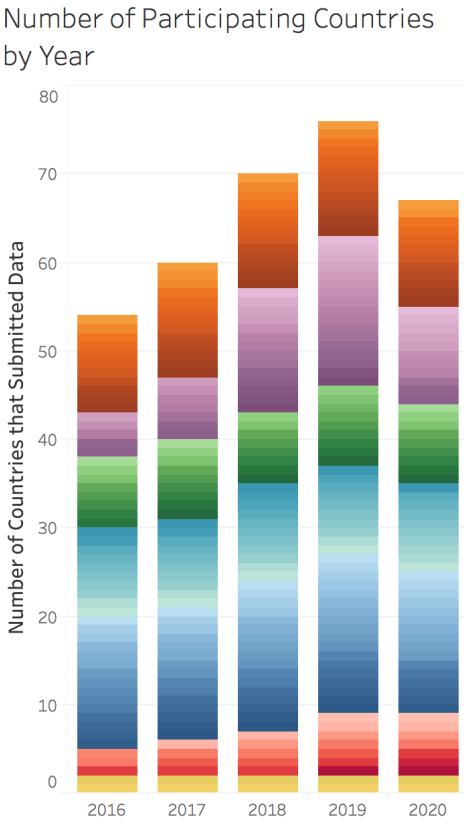

Figure 3: GLOBE countries that submitted data

In Figure 3 we note that A&P, EE, LAC, and NA have seen relatively consistent participation. The Africa region had a slight decrease in countries submitting data, however the total number of countries participating is still greater than the number observed in 2017. The NENA region had an increase in participation over time and the total participating countries remained steady between 2019 and 2020.

In addition to observing which countries submitted data from 2016 to 2020, this paper also investigates the total number of COVID-19 cases as of December 31, 2020 by population density in GLOBE countries to assess whether particular GLOBE regions were more heavily affected by COVID-19 in terms of population density.

The total number of COVID-19 cases by population density for GLOBE countries (all values rounded to the nearest integer) are presented in Figure 4. The equation used to calculate the total COVID-19 cases by population density can

be seen below.

$$\frac{\sum_{i=1}^{365} \text{COVID-19 cases}}{\text{population density}} = \frac{\text{total COVID-19 cases in 2020}}{\text{population in 2020/country area}}$$

where i is the day of the year in 2020.

As of December 31, 2020, there are a number of countries that had cases by population density less than 1, such as Bermuda (0.199), Fiji (0.386), Marshall Islands (0.005), Mauritius (0.327), Seychelles (0.460), and Taiwan (0.469) that are not seen on the map. Of all GLOBE countries with values greater than 1, Vietnam had the lowest number of total COVID-19 cases by population density and the United States had the highest. The NA GLOBE region had the highest number of COVID-19 cases by population density, LAC had the second highest and EE had the third highest.

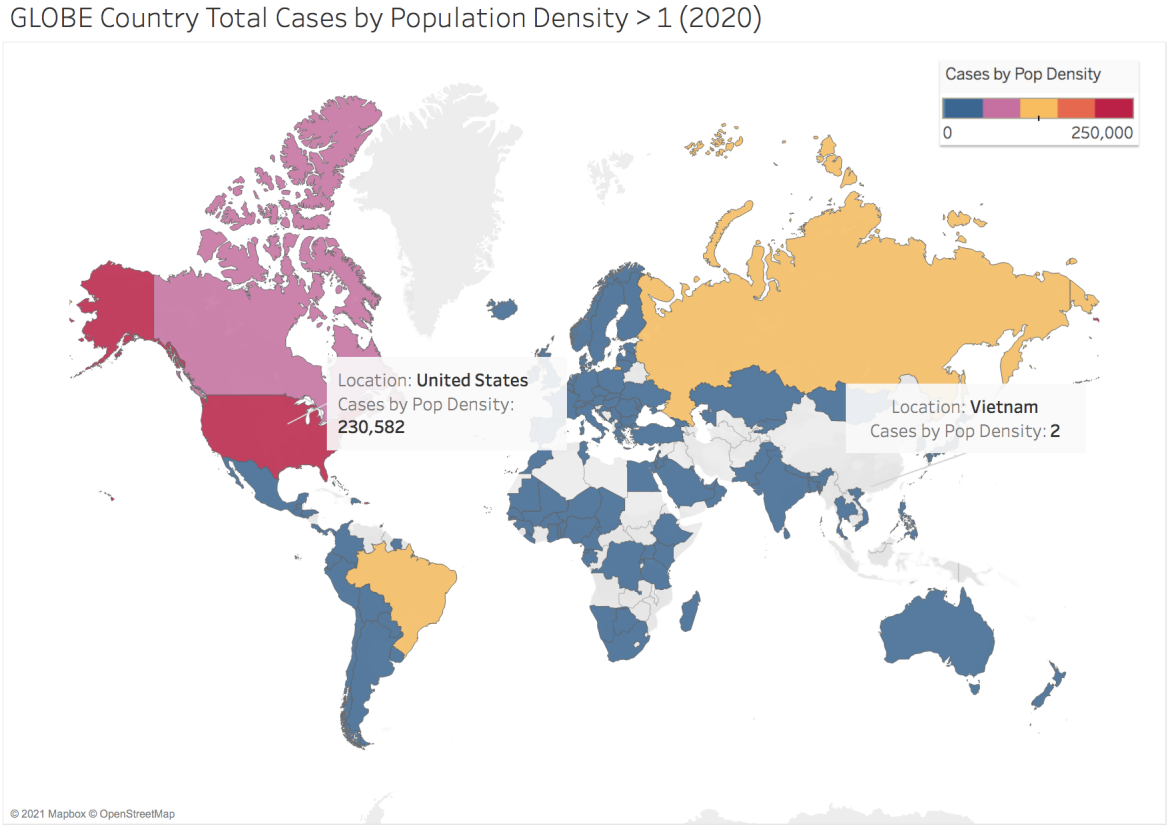

Figure 4: GLOBE Country Total COVID-19 Cases by Population Density.

3.2 Weekly Submissions

We analyzed the total number of GLOBE data submissions over the past 5 years (2016-2020). In the following figures, submissions are reported by year and by submission type. Each figure shows the data trend over the course of a year grouped by week. In other words, each data point represents the total number of submissions recorded during a given week. Week 1 and week 53 are excluded from the figures because they contain different numbers of days from year to year.

Note that for the following figures, COVID-19 was declared a pandemic on March 11, 2020 (week 11) by the W.H.O.; many GLOBE countries (particularly in A&P) were dealing with the crisis much earlier.

Figure 5 shows the total number of GLOBE submissions over the past five years. The data show an increase in the total number of weekly submissions over time. From 2016 to 2020 we observe progressively more submissions and in 2020 an average of approximately 160,000 sub-

missions per week compared to approximately 60,000 per week in 2016.

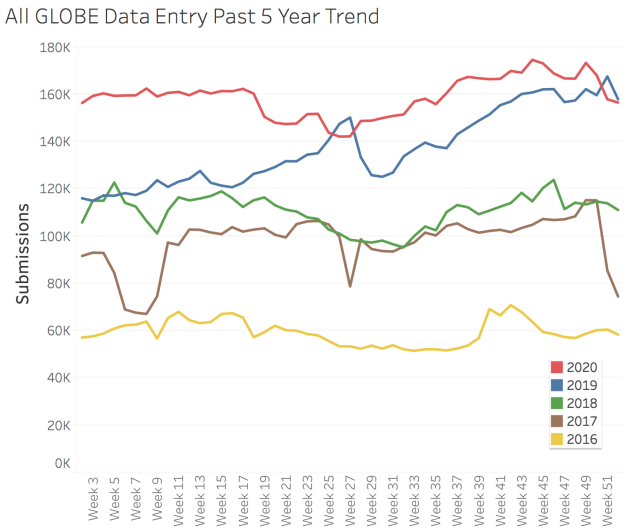

Figure 5: GLOBE Total Weekly Submissions.

Figure 6 shows the total number of weekly GLOBE submissions over the course of a year for the past 5 years *excluding* automated weather stations. Automated weather stations can submit up to 96 measurements per day because they can be taken at as little as 15 minutes intervals. The majority of these automated weather stations are in the United States. Compared to Figure 5, we see an immediate drop in the total number of submissions annually, and we see the largest difference between the two datasets in 2020. For 2016-2019, we see an average weekly difference between non-automated (citizen science) submissions and all submissions of approximately 20,000 per week. In 2020, that difference increases to nearly 90,000 per week.

Notably, the number of weekly citizen science submissions were already decreasing in the latter half of 2019.

3.2.1 Submission Types

GLOBE has five primary methods for data including email data entry ("Email DE"), the mobile data entry app ("Mobile DE"), the GLOBE Observer mobile app, web forms, and automated weather stations. The mobile data entry app is in the process of being merged with the GLOBE Observer App.

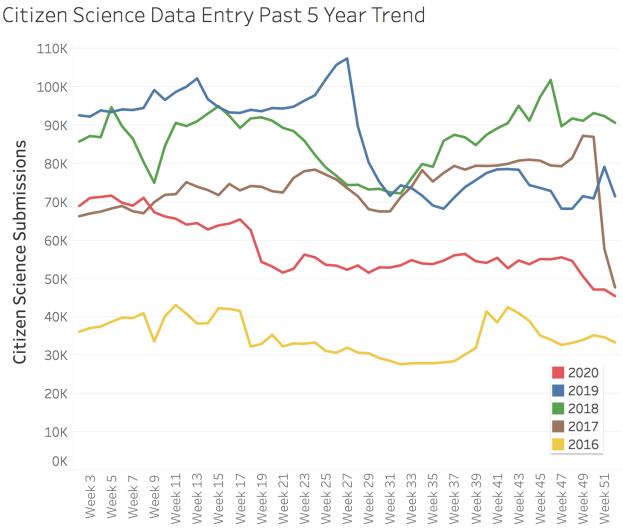

Figure 6: GLOBE Weekly Citizen Science Submissions

Starting with GLOBE's email data entry in Figure 7, we can see that use of the data entry app has decreased throughout 2020 on average by approximately 400 submissions per week.

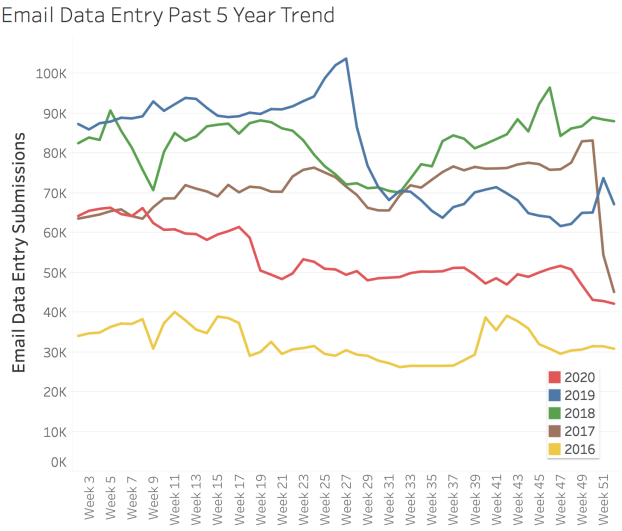

Figure 7: GLOBE Weekly Email Data Entry Submissions

Email data entry from 2017-2019 ranged from 65,000 measurements per week on average in 2017 up to 98,5000 measurements per week on average in 2019. In 2020, we see a steady decrease from 65,000 measurements to about 42,000 measurements per week. While the total number of submissions using email data

entry was still greater in 2020 than it was in 2016 (35,500 per week on average), the effects of COVID-19 may have impacted the quantity of submissions submitted through email data entry.

Note that Figure 6 and Figure 7 share striking similarities. This indicates that email data entry accounts for the majority of citizen science observations.

Looking at web form submissions in Figure 8, we see that the total weekly submissions appear to follow a general pattern for a school year in the northern hemisphere (more submissions during the spring, week 2 - week 19, and during the fall, week 33 - week 51). Entries during the fall school year period appear to include a peak during weeks 39-45, with submissions dropping off into the end of the year, coinciding with holiday breaks. On average, COVID-19 did not appear to have an impact on web form weekly submissions.

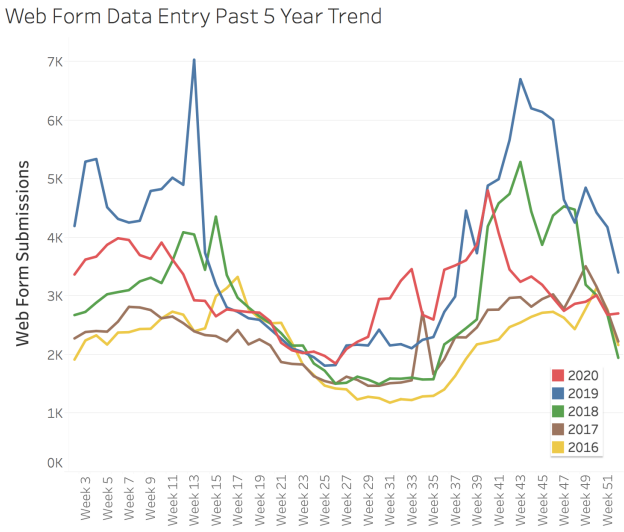

Figure 8: GLOBE Weekly Web Form Data Entry Submissions

Third, for the GLOBE Observer App data submissions in Figure 9, we note that there has not been a substantial decrease in the average weekly submissions in 2020 compared to the past 4 years. Although previous years saw higher unique spikes which are likely the result of data collection campaigns, 2020 was largely consistent across all weeks with an average of 960 submissions per week.

Next, we can see that the total number of weekly submissions through the data entry app (Figure 10) have remained stable throughout 2020 and the trend did not vary substantially from past years. In other words, the GLOBE Observer App, which is used by students and citizen scientists, has not seen any major difference in data submission frequency in 2020 compared to previous years.

Figure 11 shows the total number of weekly automated weather station submissions over time. These automated submissions contribute greatly to GLOBE's database. Between the end of 2019 and the end of 2020 we have seen an increase of about 40 thousand weekly submissions.

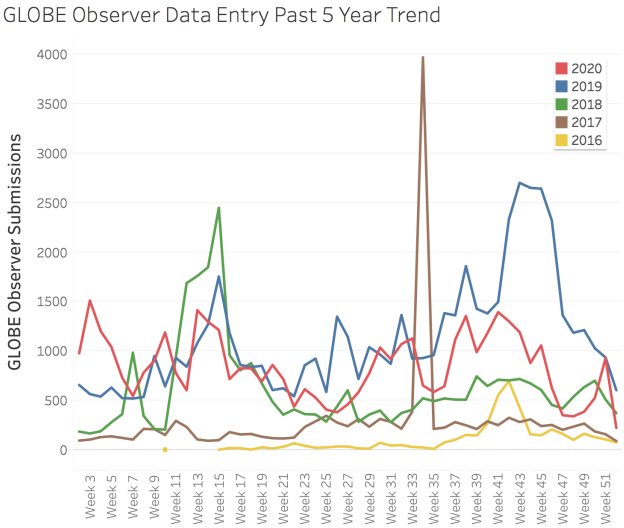

Figure 9: GLOBE Weekly GLOBE Observer App Submissions

GLOBE received approximately 300 submissions from mobile data entry, 65,000 from email data entry, 970 from the GLOBE Observer App, 3,200 from web forms, and 85,000 from automated weather stations per week in 2020. In 2019, approximately 700 submissions from mobile data entry, 98,000 from email data entry, 530 from the GLOBE Observer, 3,500 from web forms, and 60,000 from automated weather stations were recorded per week.

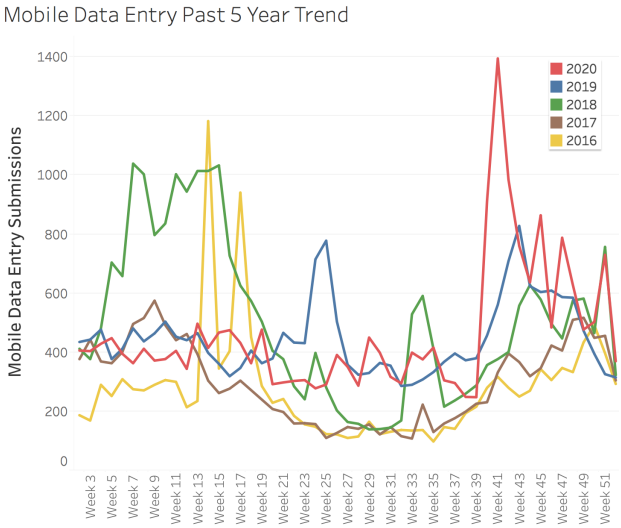

Figure 10: GLOBE Weekly Mobile Data Entry Submissions

Although GLOBE is receiving fewer citizen science submissions through mobile and email data entry, web forms, the GLOBE Observer App, and automated measurements continue to contribute to the increase in GLOBE data entered into the database.

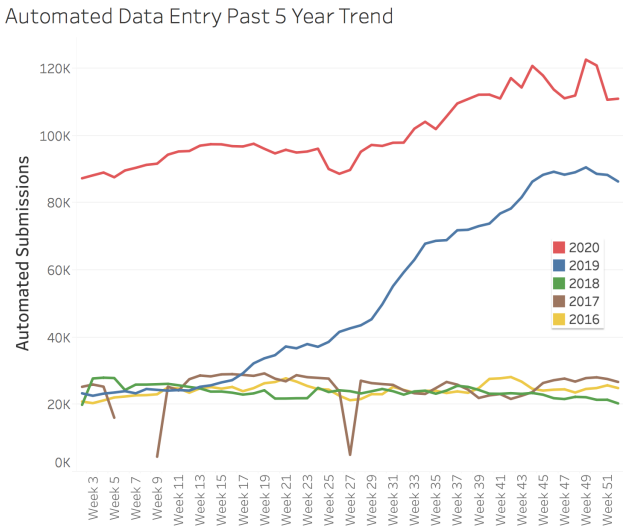

Figure 11: GLOBE Weekly Automated Submissions

3.3 Submissions by Region

The following figures show the total GLOBE measurements by year, divided by region. The colors in the plots correspond to Figure 2.

The total GLOBE measurements shown in Figure 12, show a clear increase in total yearly measurements over time. In 2020, GLOBE members in NA submitted data approximately 7.5 million times, NENA members submitted data approximately 87,000 times, LAC members submitted data approximately 22,000 times, EE members submitted data approximately 460,000 times, A&P members submitted approximately 180,000 times and Africa members submitted approximately 18,000 times.

North America automated data entry submissions were at nearly 5.5 million in 2020 and the remaining, approximately 2.5 million measurements, were citizen science related (Figure 13). Figure 13 also shows that even although a number of regions saw a decrease in citizen science data submissions from 2019 to 2020, NA and A&P regions did not see a substantial decrease.

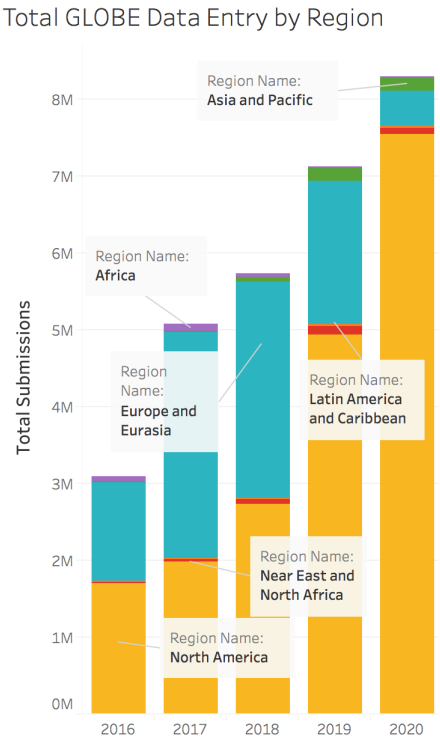

Figure 12: GLOBE Submissions by Region

In the NA GLOBE region there was a clear increase in total data submissions from 2016-2020; as we'll see in further figures, this increase can be attributed to the increase in automated measurements. In Africa, there was a sharp de-

crease in total data measurements since 2017 (when GLOBE received nearly 100,000 submissions from the Africa region). It is possible that the decrease in measurements from 2019 to 2020 for EE and NENA may be attributable to COVID-19 stay-at-home restrictions but could be investigated further given that similar decreases were not observed in other regions in which restrictions were also in place.

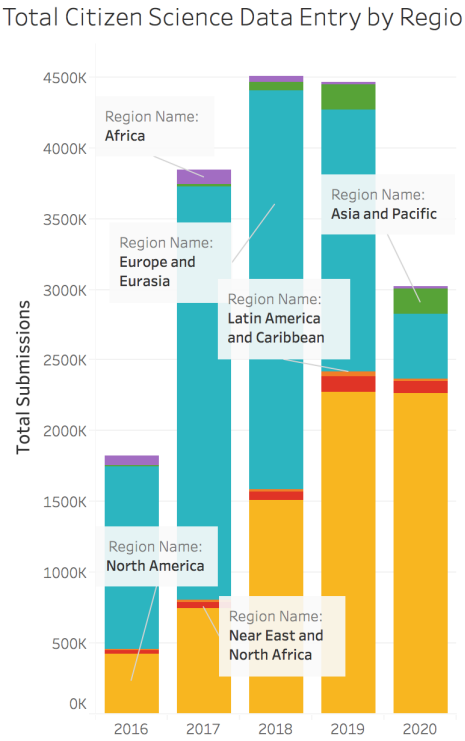

Figure 13: GLOBE Weekly Citizen Science Submissions by Region

3.3.1 Submission Types

Next we observe the regional distribution of web form based submissions over the past five years (Figure 14). Note that there is an order of magnitude difference in total submissions between web forms and email data entry. Web form submissions are most popular in the A&P, EE, and NENA regions.

The A&P region submitted more data via web form in 2020 compared to previous years, whereas, EE, LAC, and NA submitted less data than in past years. There was a steady increase in total web form submissions from 2016 to 2019, followed by a decrease in 2020.

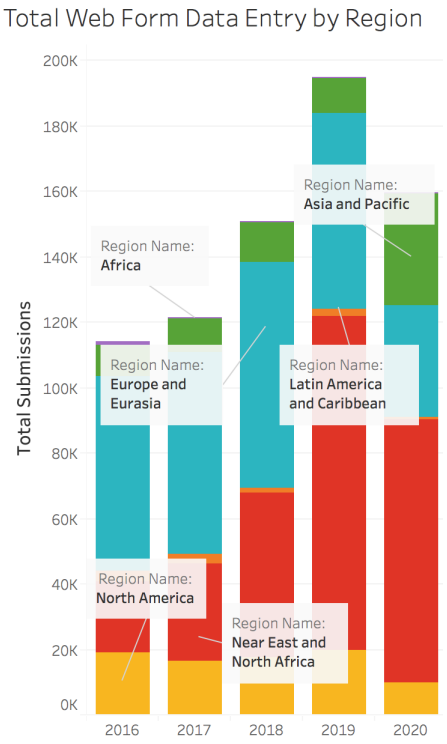

Figure 14: GLOBE Weekly Web Form Data Entry Submissions by Region

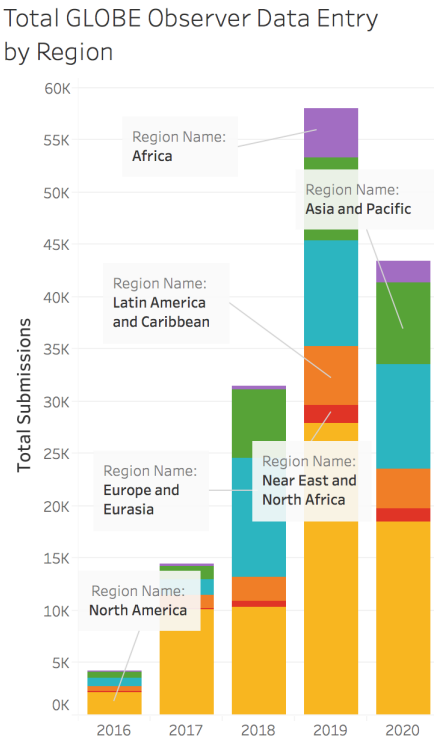

Figure 15: GLOBE Weekly GLOBE Observer App Submissions by Region

Figure 15 shows that the GLOBE Observer

App attracts submissions uniformly across all regions. In this figure, we see that A&P and EE have continued to submit approximately the same amount of data from 2018-2020. All other regions have seen a slight decrease in total submissions in 2020 compared to 2019.

Automated submissions were excluded from the submission type regional analysis because the total number of automated submissions has steadily increased from 2016 through 2020 and all automated submissions came from the NA region. Mobile data entry submissions were excluded because the app is in the process of being merged into the GLOBE Observer App. Email data entry submissions were excluded because the data is nearly identical to the weekly citizen science data entry figure.

3.4 Selected Protocols

In this section, we present data for two specific protocols: the cloud protocol and the mosquito protocol.

These data include 2017-2020 submissions as the GLOBE Observer App (most clouds and mosquitoes submissions are entered via this app) became available in 2016 and submissions were sparse in the 1st year the app was available.

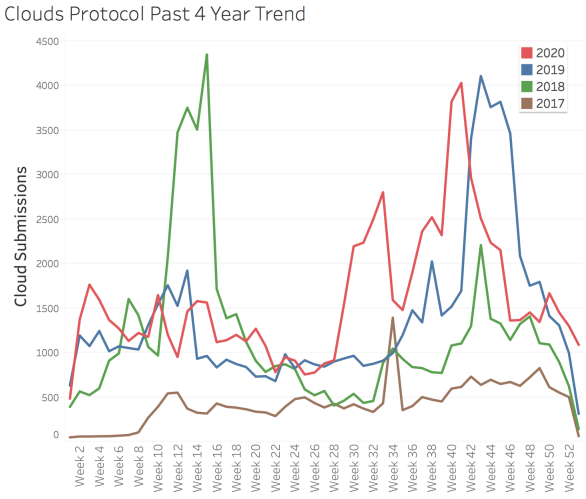

Figure 16: GLOBE Weekly Cloud Submissions Over Time

First, looking at the GLOBE clouds protocol data submissions in Figure 16 we notice that there was a steady increase in the average number of weekly cloud submissions from 2017 to 2019. In 2018, there is a spike in the spring which can be attributed to the GLOBE Observer Spring Clouds Challenge. There is also a spike around fall 2019 which can be attributed to the GLOBE Observer Fall Cloud Observation Data Challenge. In 2020 we see a similar spike in the fall due to GLOBE Observer's 2020 Community Cloud Challenge: Science is Better Together. Overall, the cloud submissions have continued to steadily increase, annually, on average. The school year seasonality can be seen here as well.

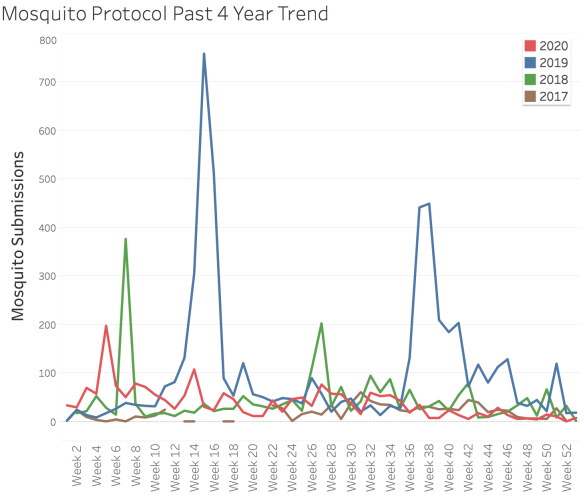

Figure 17: GLOBE Weekly Mosquito Submissions Over Time

Next we focus on the four-year mosquito protocol data submission trend (Figure 17). Between 2017 and 2019, there was a noticeable increase in weekly submissions. In 2020, the total number of weekly submissions remains at an average of approximately 50 submissions per week. There were no large spikes in mosquito data submission in 2020, and submissions remained relatively stable as in past years. The spikes in 2019 may be attributed to the GLOBE Observer Mosquito Blitz in the spring and the GLOBE Mission Mosquito (GMM) webinars in the fall.

3.5 Mosquitoes

The mosquito protocol was selected for this paper to study whether there was a heightened interest or awareness in public health issues related to COVID-19.

At GLOBE the mosquito protocol was integral to the success of a recently completed collaborative project with the Department of State called the "GLOBE Zika Education and Prevention project." The project pushed the edge of citizen science impact and enlisted thousands of students, teachers, and communities to collect data on mosquitoes for a global mapping project. The project exceeded it's goals, increasing mosquito studies and creating connections between public health officials to better control mosquitoes and reduce mosquito borne infectious disease. The GLOBE mosquito protocol continues to be used in ongoing studies as a powerful tool for mitigating local mosquito populations.

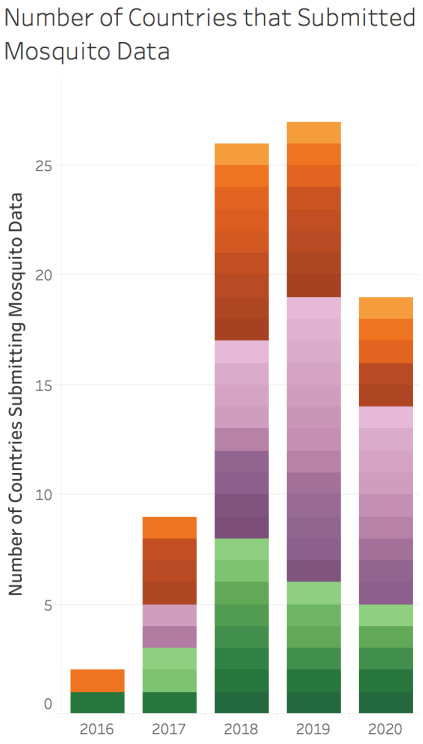

Figure 18: GLOBE Countries Submitting Mosquito Data

Figure 18 is a representation of the countries that have submitted mosquito data. The Zika Project started in 2018 and targets the Africa, A&P and NENA regions. As mentioned in section 1.3, purple colors represent countries in the Africa region, green colors represent countries in the A&P region, and orange colors represent countries in the LAC region.

When the project started in 2018, GLOBE saw a sharp increase in the number of distinct participating countries (from nine countries to 26). In 2019, GLOBE had one additional country submit mosquito data, with fewer countries from the A&P region and more countries from the Africa region participating. In 2020, many countries within these GLOBE regions were heavily impacted by COVID-19 which may have influenced the total number of participating countries[16].

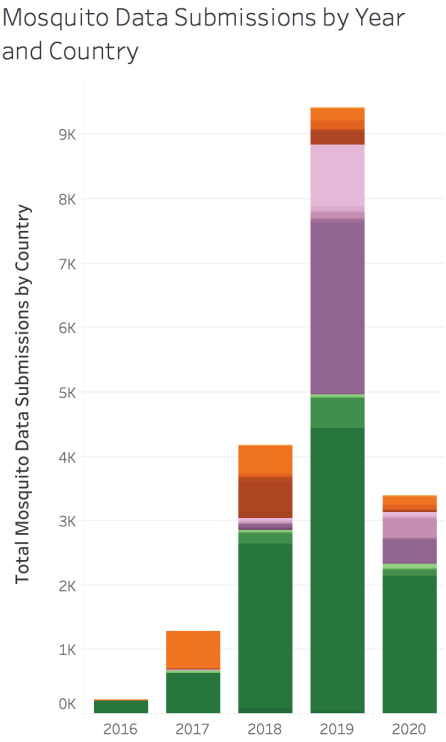

Figure 19: Yearly GLOBE Country Mosquito Submissions

Finally, the total data submissions by region (Figure 19) show that after the Zika project started in 2018, there was a substantial increase from 2017 to 2018 and then again between 2018 and 2019. The Zika project was heavily impacted by COVID-19 as travel restrictions were implemented, which prevented further protocol and community training, and the project began wrapping up in late 2020. In 2020, there was a sharp decrease in submissions across each of the three Zika Project GLOBE regions but the total

data submitted was still greater than 2017.

3.6 Training

GLOBE provides protocol e-Training and in-person protocol workshop training to teach educators, students and citizen scientists how to correctly follow GLOBE protocols to collect data. In this section, we briefly look at the total number of individuals that are trained weekly.

First, Figure 20 shows the number of individuals that are e-trained weekly from April 2016 - December 2020. The vertical axis shows the total number of people that are e-trained with a maximum of approximately 155 people.

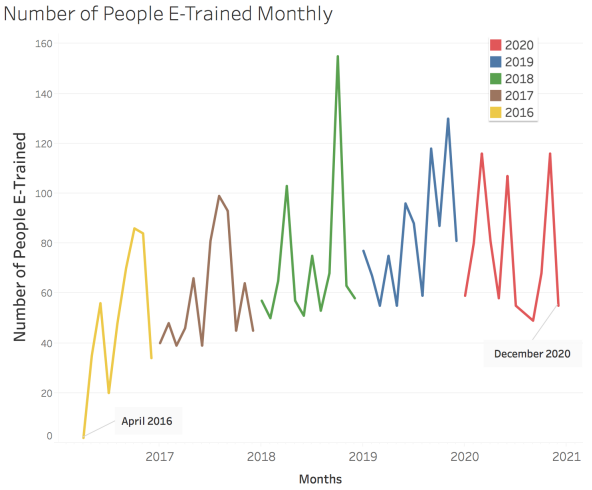

Figure 20: Number of People E-Trained Weekly

It's notable that the number of people taking the GLOBE e-Training have remained approximately the same over the past five years. In fact, in early 2020, prior to the declaration of the pandemic in week 11, GLOBE saw a greater number of individuals being trained compared to previous years. For the remainder of the year, the number of trained individuals remained about the same as those trained in 2016-2019.

Second, the number of individuals that were in-person workshop trained weekly from April 2016 - December 2020 are shown in Figure 21. The horizontal axis shows the week of the year and the vertical axis shows the total number of people that were workshop trained with a maximum of approximately 1,350 people. There were substantially fewer individuals trained in in-person workshops in 2020 than in previous years, likely related to COVID-19 related restriction for in-person activities.

Figure 21: Number of People Workshop Trained Weekly

4 Related Work

Citizen science programs had a great opportunity to encourage participation in science and data collection during the pandemic as some individuals had more flexibility, time at home, and opportunities to be engaged outdoors during this period. Many citizen science organizations, including GLOBE, began encouraging more citizen science participation and student engagement (e.g. via online working activities and data entry)[17].

Although GLOBE did not see an increase in citizen science participation throughout the pandemic, many[11][18] citizen projects reported an uptick of citizen science submissions around the beginning of the pandemic (March 11 – early April). Undark Magazine reported that Zooniverse saw "200,000 participants contributed over five million classifications across all projects, the equivalent of 48 years of full-time research"[11]. Projects such as Stall Catchers (supporting research for Alzheimer's disease)[11], COVID-19 Citizen Science (tracking COVID-19 Symptoms)[19], and Globe at Night (a citizen science organization collecting data of the night sky and light pollution)[20], are just a few examples of citizen science programs that saw increased participation throughout the pandemic. InsideCli-

mate News wrote that "Citizen science [is seeing] a boom in participation... motivated by cuts [to] environmental protections and COVID-19 impacts on data gathering and justice."[21]

Other science, education and student engagement programs – such as the GLOBE Program – did not see an increase in data submissions in 2020. For example, SABAP2 (Southern African Bird Atlas Project) stated that "most citizen scientists in South Africa were not able to submit full protocol lists during the Alert Level 5 lockdown"[22] due to "mobility [being] severely constrained"[22]. Haeftan, et al. (2020) identified factors that could potentially be problematic for citizen science organizations in the midst of COVID-19. These factors include: (a) Limited time and resources to develop and implement project, (b) variable student access to resources, (c) challenges engaging students, (d) health and safety concerns, etc.[23] Many of these, in particular, access to resources and health and safety concerns, are likely sources of impact to the GLOBE citizen science effort.

5 Conclusions and Future Work

This paper provides a descriptive analysis of GLOBE data over time with a focus on the potential impacts of COVID-19. The goal of this paper was to provide broad insight into GLOBE data collection trends, the potential impact of COVID-19 restrictions on data collection worldwide and on the continuation of GLOBE protocol training, in-person and online.

The GLOBE Program is a successful science, education, and citizen science program that has reached a global audience since 1995. In this paper, we presented a number of summaries detailing GLOBE data collection and entry over time. These data show that GLOBE did not experience an increase in citizen science data submission during the time period investigated and that the resulting restrictions due to the COVID-19 pandemic may have had a slight negative impact on some data collections. GLOBE saw a decrease in email data entry submissions in 2020 compared to previous years. Web form, GLOBE Observer App, and mobile data entry were not impacted by the pandemic.

It is challenging for a causal link between COVID-19 restrictions and GLOBE citizen science participation by data submission to be established as there are a number of potential confounders such as lack of access to protocol materials and educators. In addition to teachers not being able to teach in-person and a lack of resources, social distancing and stay-at-home orders may have also played a role in data submission trends as students couldn't get outside to take Earth science measurements. Despite the overall increase in data submissions, we posit that COVID-19 restrictions and additional factors have resulted in an overall decrease in citizen science weekly data submissions (particularly with email data entry) and decrease in in-person workshop trained users. GLOBE paused in-person workshops in early 2020 to mitigate the spread of the COVID-19 pandemic. From Figure 19 we postulate that COVID-19 was such an overwhelming public health crisis that people may have had little personal capacity or interest to investigate other public health areas like vector-borne disease, although we expect an increased interest in this area as restrictions are lifted.

An area to investigate further is whether the observed decreases in GLOBE data submission during the investigated time period were due to COVID-19 or because of other external factors. For example, many of the GLOBE protocols require additional materials (e.g. thermometer) that many people may not have at home. Teachers are typically trained in the GLOBE protocols and guide their students in protocol measurements. With COVID-19 restrictions in place and schools being closed, it's possible that students did not have access to the materials they needed to complete some of the protocols. Most of the GLOBE protocols included in the GLOBE Observer App do not require any additional material and these protocols did not appear to see a substantial decrease compared to other data submission methods over time. It would be interesting to analyze whether GLOBE schools that experienced different COVID-19 related restrictions in different countries submitted different amounts of data.

Another area to investigate is whether or not schools located in cities versus rural areas were able to submit more, less, or the same number of submissions as previous years. Figure 13 shows a substantial decrease in citizen science submissions particularly in the EE region. A question that arises is whether or not this is due to many schools being in high density cities as opposed to the US, for example, where many of the schools are in rural and suburban areas. A potential dataset that could help with such an analysis is the CDC's Urbanization Classification dataset[24].

Data quality could be an additional topic for further investigation. Data quality is one of GLOBE's highest priorities. This paper researched data quantity – how much data GLOBE received over a given time period by region. It's possible that data submitted throughout the pandemic may be more precise due to people having more time to dedicate to submission quality. Alternatively, it's possible that submission quality was negatively impacted as a result of fewer in-person training sessions and less hands-on learning facilitated by teachers for students.

Citizen science opportunities such as those provided by the GLOBE Program provide the public with a means to be active participants in collecting meaningful scientific data. Scientists benefit from increased availability of data that are used directly to enhance research data sets or to validate research. Citizen scientists' data also benefit educators who use GLOBE data to educate students and answer local Earth science questions at local scales globally. Students and teachers alike also use the GLOBE app to enter data. Citizen science enhances data collection because it engages larger audiences and provides multiple opportunities for people to play an active role in science.

We conclude this paper by providing a number of suggestions. First, GLOBE is in the process of translating eTraining modules to be more inclusive of all citizen scientists, which participate from locations worldwide and would benefit from training materials language adaptations. Second, GLOBE may consider which protocols can be simplified further so that citizen scientists that do not have access to some of the more sophisticated tools needed for some protocols can still participate and submit measurements outside of a school or educational environment that has access to such tools. Third, GLOBE may be interested in migrating the e-Training modules to a Learning Management System built with schools in mind, such as Google Classroom (or a similar platform) and updating the current slides to a PDF format in order to increase visibility and ease of use (while also retaining the availability of the current PowerPoint format). Fourth, GLOBE might consider optimizing the Advanced Data Access Tool to reduce the time it takes to download a data file summary and work on updating the Visualization system to simplify data visualization. The latter can be accomplished by asking for the community and members outside of the GLOBE community to provide feedback or by making the code bases for those systems open source.

After 26 years, GLOBE continues to grow with more countries joining each year, giving the opportunity for increased data coverage which will lead to more measurements and perhaps the closure of data gaps for certain protocols. In addition, it is likely that the composition of the suite of protocols that are in the program will change, altering the nature of the data available. Finally, as technology changes, different and more advanced data collection instruments may appear and be integrated into the program. However, GLOBE's commitment to maintaining data as well as data quality, its engagement of the community in the collection and use of the data, will remain steadfast.

6 Acknowledgements

The authors of this paper would like to thank Mindi DePaola (GLOBE Program, Program Specialist) and Dave Salisbury (GLOBE Program, Data Engineer) for providing guidance on the structure of the paper, suggesting the goals of the paper and offering feedback. The GLOBE Implementation Office would like to thank NASA for the continued sponsorship and the NSF, NOAA and the Department of State for the continued support.

References

[1] *About The GLOBE Program*. URL: https://www.globe.gov/about/overview.

[2] *Citizen Science, Important Tool for Researchers*. Mar. 2021. URL: https://magazine.scienceconnected.org/2018/09/citizen-science-important-tool/#:~:text=Citizen%20Science%20Informs%20Research%20and,the%20contributions%20of%20citizen%20scientists.&text=This%20is%20great%20help%20for,to%20time%20and%20resource%20constraints..

[3] *COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)*. Mar. 2021. URL: https://coronavirus.jhu.edu/map.html.

[4] Jordan Cates, Cynthia Lucero-Obusan, Rebecca M Dahl, et al. *Risk for In-Hospital Complications Associated with COVID-19 and Influenza - Veterans Health Administration, United States, October 1, 2018–May 31, 2020*. Oct. 2020. URL: https://www.cdc.gov/mmwr/volumes/69/wr/mm6942e3.htm.

[5] *Advice for the public on COVID-19*. 2020. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.

[6] *Guidance for Cleaning and Disinfecting Public Spaces, Workplaces, Businesses, Schools, and Homes*. URL: https://www.cdc.gov/coronavirus/2019-ncov/community/reopen-guidance.html.

[7] *Considerations for quarantine of contacts of COVID-19 cases*. URL: https://www.who.int/publications/i/item/considerations-for-quarantine-of-individuals-in-the-context-of-containment-for-coronavirus-disease-(covid-19).

[8] *COVID-19 Government Response Tracker*. URL: https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker.

[9] CoCoRaHS. *Countdown to 50 million! Sometime this weekend we will receive our 50 millionth daily report!! What an accomplishment, and thank you to each and every volunteer. The "winning" station will receive a certificate but may wish to stay anonymous. Good luck!* Apr. 2020. URL: https://twitter.com/CoCoRaHS/status/1250546102601805824?s=20.

[10] *Citizen scientists across the globe unite for the City Nature Challenge April 24 – 27, 2020*. 2020. URL: https://www.calacademy.org/press/releases/citizen-scientists-across-the-globe-unite-for-the-city-nature-challenge-april-24-%E2%80%93-27.

[11] Andy Withers James Dinneen. *Covid-19 Can't Stop Citizen Science*. Apr. 2020. URL: https://undark.org/2020/04/17/covid-19-citizen-science/.

[12] The Zooniverse. *We've received over 5 million classifications across all projects in the last week!That's three times the amount for the same week in 2019.We're not sure what to say, other then "Thank you, you are all unbelievable!"*. Apr. 2020. URL: https://twitter.com/the_zooniverse/status/1246111810752786433?s=20.

[13] *Worldometers: Real Time World Statistics*. 2021. URL: https://www.worldometers.info/.

[14] Hannah Ritchie. *Coronavirus Source Data*. 2021. URL: https://ourworldindata.org/coronavirus-source-data.

[15] Nytimes. *nytimes/covid-19-data*. 2021. URL: https://github.com/nytimes/covid-19-data.

[16] Munyaradzi Makoni. "COVID-19 in Africa: half a year later". In: *The Lancet Infectious Diseases* 20.10 (2020), p. 1127.

[17] Andrew Moore. *How You Can Contribute to Scientific Research During the COVID-19 Pandemic*. Apr. 2020. URL: https://cnr.ncsu.edu/news/2020/04/citizen-science-coronavirus/.

[18] *Citizen Science and COVID-19: The Power of the (Distanced) Crowd*. June 2020. URL: https://www.wilsoncenter.org/blog-post/citizen-science-and-covid-19-power-distanced-crowd.

[19] *Covid19 Citizen Science - Participant Counter and World Map of Participation*. 2021. URL: https://covid19.eurekaplatform.org/.

[20] *Maps and Results*. 2021. URL: https://www.globeatnight.org/maps.php.

[21] Maddie Kornfeld. *As Scientists Struggle with Rollbacks, Stay At Home Orders and Funding Cuts, Citizens Fill the Gap*. Nov. 2020. URL: https://insideclimatenews.org/news/23062020/citizen-science-coronavirus/.

[22] Sanjo Rose, Jessleena Suri, Michael Brooks, et al. "COVID-19 and citizen science: lessons learned from southern Africa". In: *Ostrich* 91.2 (2020), pp. 188–191.

[23] Shanice Van Haeften, Andelija Milic, Beth Addison-Smith, et al. "Grass Gazers: Using citizen science as a tool to facilitate practical and online science learning for secondary school students during the COVID-19 lockdown". In: *Ecology and evolution* (2020).

[24] CDC. *NCHS Urban-Rural Classification Scheme for Counties*. June 2017. URL: https://www.cdc.gov/nchs/data_access/urban_rural.htm#2013_Urban-Rural_Classification_Scheme_for_Counties.

Appendix

Current list of GLOBE countries by GLOBE region (see also: GLOBE Community Map):

Region	Country
Africa	Benin
Africa	Botswana
Africa	Burkina Faso
Africa	Cameroon
Africa	Cape Verde
Africa	Chad
Africa	Congo
Africa	Ethiopia
Africa	Gabon
Africa	Gambia
Africa	Ghana
Africa	Guinea
Africa	Kenya
Africa	Liberia
Africa	Madagascar
Africa	Mali
Africa	Mauritius
Africa	Namibia
Africa	Niger
Africa	Nigeria
Africa	Rwanda
Africa	Senegal
Africa	Seychelles
Africa	South Africa
Africa	Tanzania
Africa	Togo
Africa	Uganda

Region	Country
North America	Canada
North America	United States of America

Region	Country
Asia and Pacific	Australia
Asia and Pacific	Bangladesh
Asia and Pacific	Fiji
Asia and Pacific	India
Asia and Pacific	Japan
Asia and Pacific	Maldives
Asia and Pacific	Marshall Islands
Asia and Pacific	Micronesia
Asia and Pacific	Mongolia
Asia and Pacific	Nepal
Asia and Pacific	New Zealand
Asia and Pacific	Palau
Asia and Pacific	Philippines
Asia and Pacific	Republic of Korea
Asia and Pacific	Sri Lanka
Asia and Pacific	Taiwan Partnership
Asia and Pacific	Thailand
Asia and Pacific	Vietnam

Region	Country
Near East and North Africa	Bahrain
Near East and North Africa	Egypt
Near East and North Africa	Jordan
Near East and North Africa	Kuwait
Near East and North Africa	Lebanon
Near East and North Africa	Mauritania
Near East and North Africa	Morocco
Near East and North Africa	Oman
Near East and North Africa	Pakistan
Near East and North Africa	Qatar
Near East and North Africa	Saudi Arabia
Near East and North Africa	Tunisia
Near East and North Africa	United Arab Emirates

Region	Country
Europe and Eurasia	Austria
Europe and Eurasia	Belgium
Europe and Eurasia	Bulgaria
Europe and Eurasia	Croatia
Europe and Eurasia	Cyprus
Europe and Eurasia	Czech Republic
Europe and Eurasia	Denmark
Europe and Eurasia	Estonia
Europe and Eurasia	Finland
Europe and Eurasia	France
Europe and Eurasia	Georgia
Europe and Eurasia	Germany
Europe and Eurasia	Greece
Europe and Eurasia	Hungary
Europe and Eurasia	Iceland
Europe and Eurasia	Ireland
Europe and Eurasia	Israel
Europe and Eurasia	Italy
Europe and Eurasia	Kazakhstan
Europe and Eurasia	Kyrgyz Republic
Europe and Eurasia	Latvia
Europe and Eurasia	Liechtenstein
Europe and Eurasia	Lithuania
Europe and Eurasia	Luxembourg
Europe and Eurasia	Malta
Europe and Eurasia	Moldova
Europe and Eurasia	Monaco
Europe and Eurasia	Montenegro
Europe and Eurasia	Netherlands
Europe and Eurasia	North Macedonia
Europe and Eurasia	Norway
Europe and Eurasia	Poland
Europe and Eurasia	Portugal
Europe and Eurasia	Romania
Europe and Eurasia	Russia
Europe and Eurasia	Serbia
Europe and Eurasia	Slovak Republic
Europe and Eurasia	Slovenia
Europe and Eurasia	Spain
Europe and Eurasia	Sweden
Europe and Eurasia	Switzerland
Europe and Eurasia	Turkey
Europe and Eurasia	Ukraine
Europe and Eurasia	United Kingdom

Region	Country
Latin America and Caribbean	Argentina
Latin America and Caribbean	Bahamas
Latin America and Caribbean	Bermuda
Latin America and Caribbean	Bolivia
Latin America and Caribbean	Brazil
Latin America and Caribbean	Chile
Latin America and Caribbean	Colombia
Latin America and Caribbean	Costa Rica
Latin America and Caribbean	Dominican Republic
Latin America and Caribbean	Ecuador
Latin America and Caribbean	El Salvador
Latin America and Caribbean	Guatemala
Latin America and Caribbean	Honduras
Latin America and Caribbean	Mexico
Latin America and Caribbean	Panama
Latin America and Caribbean	Paraguay
Latin America and Caribbean	Peru
Latin America and Caribbean	Suriname
Latin America and Caribbean	Trinidad and Tobago
Latin America and Caribbean	Uruguay

Figure 1: GLOBE Countries

This appendix includes a number of figures that provide additional information and context for data presented in the main body of this paper.

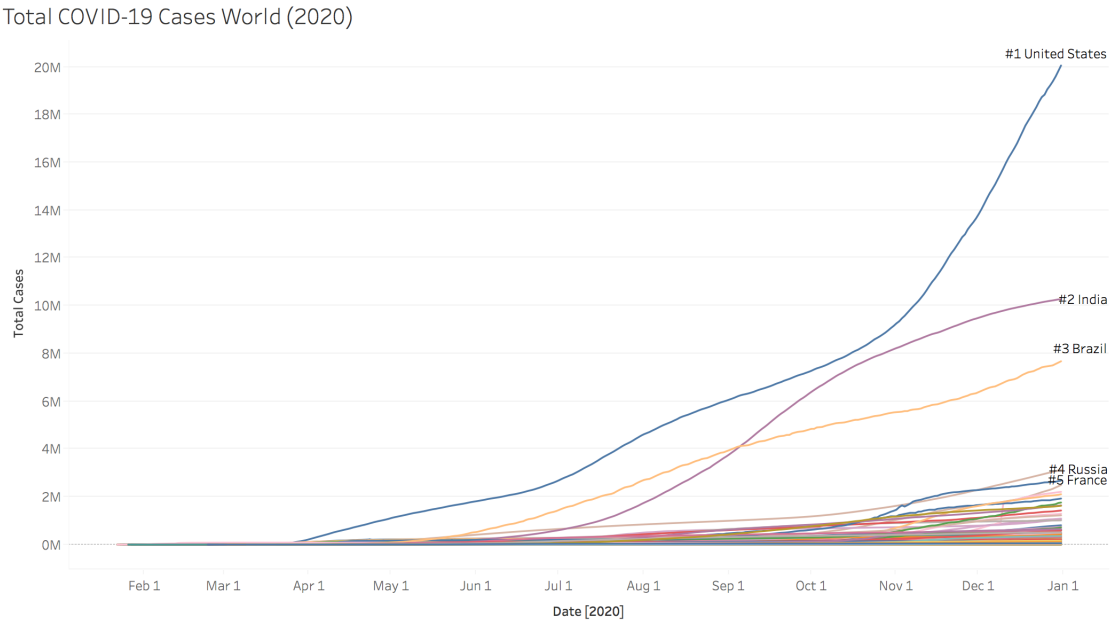

Figure 2: Total COVID-19 Cases World

Figure 2 shows the growth of total cases in all countries around the world in 2020 (Jan 1 - Dec 31). It also shows the five countries that had the highest number of total COVID-19 cases as of December 31.

The following figures (3, 4, 5, 6, 7, 8) show the total density of GLOBE data submissions over time by region.

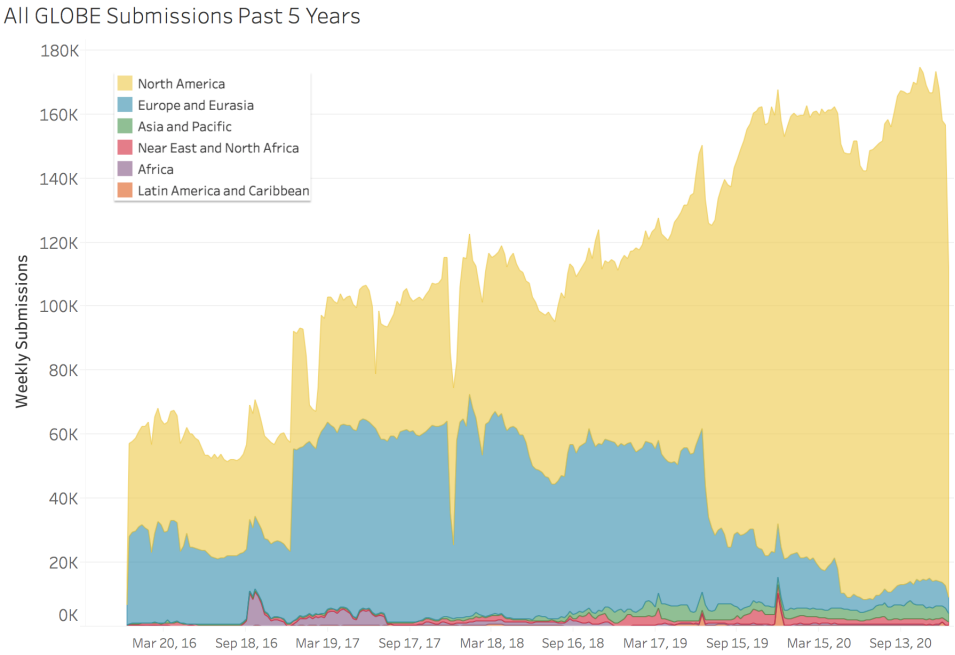

Figure 3: Total GLOBE Submissions by Region Over Time

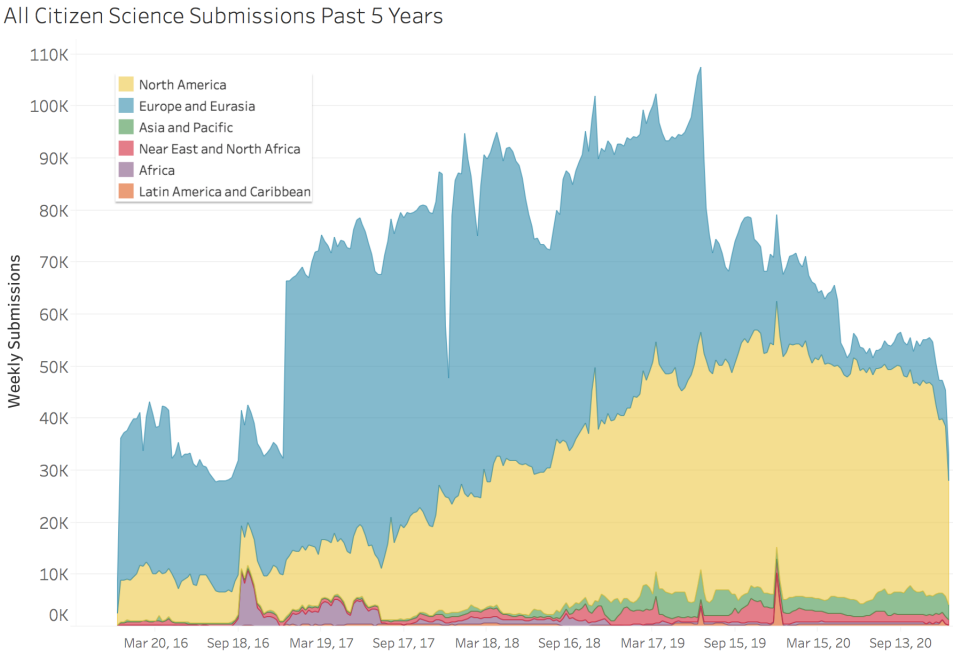

Figure 4: Total GLOBE Citizen Science Submissions by Region Over Time

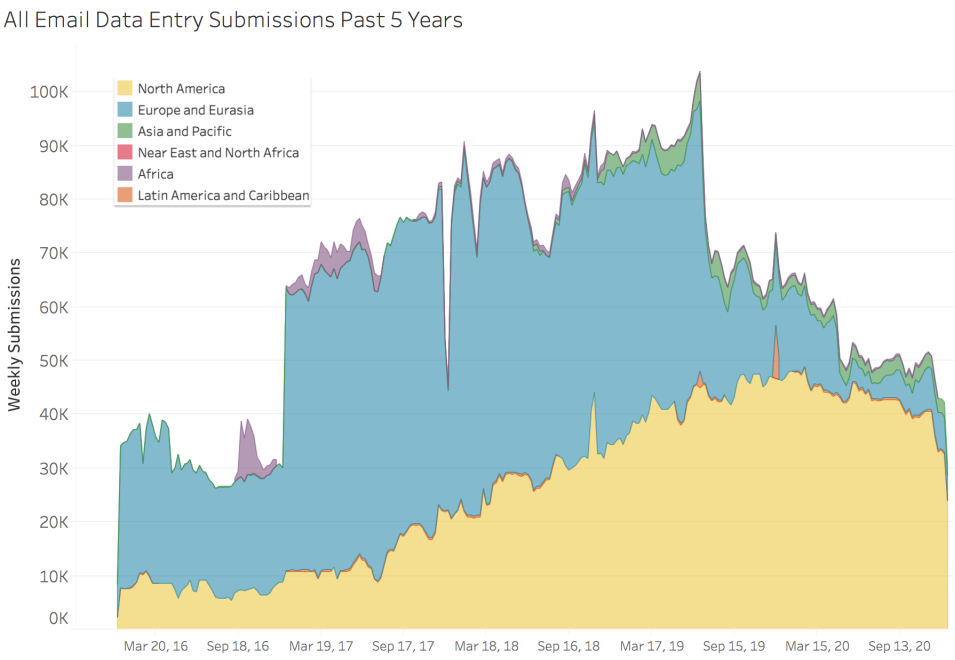

Figure 5: Total Email Data Entry GLOBE Submissions by Region Over Time

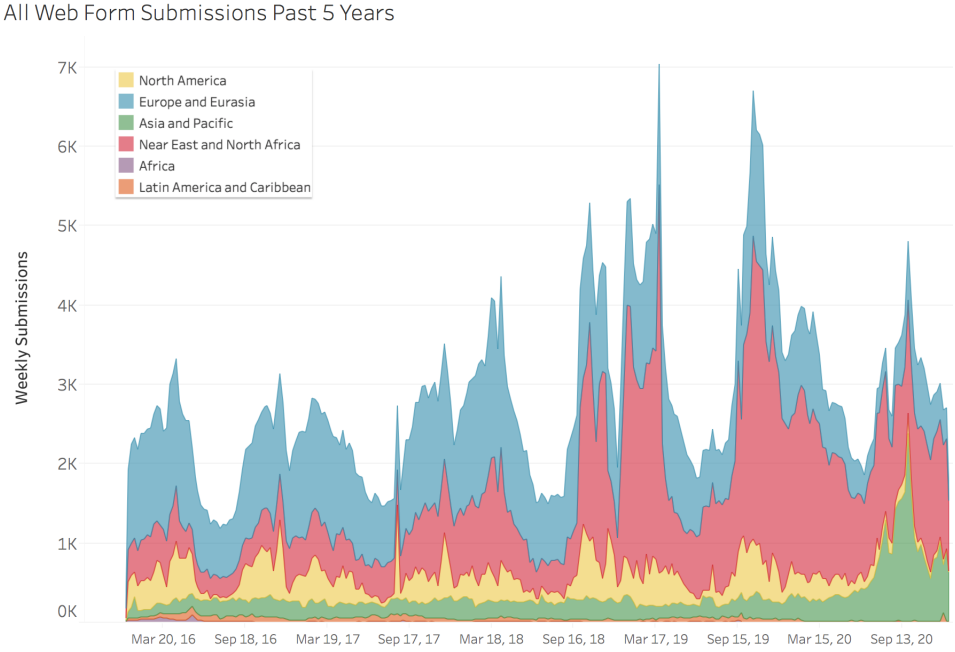

Figure 6: Total Web Form Submissions by Region Over Time

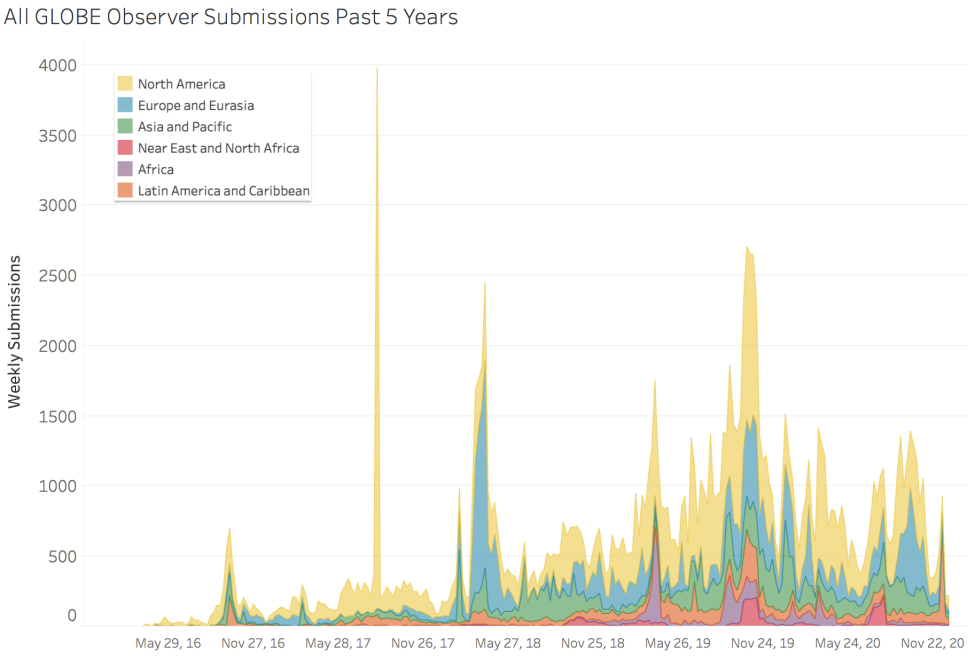

Figure 7: Total GLOBE Observer Submissions by Region Over Time

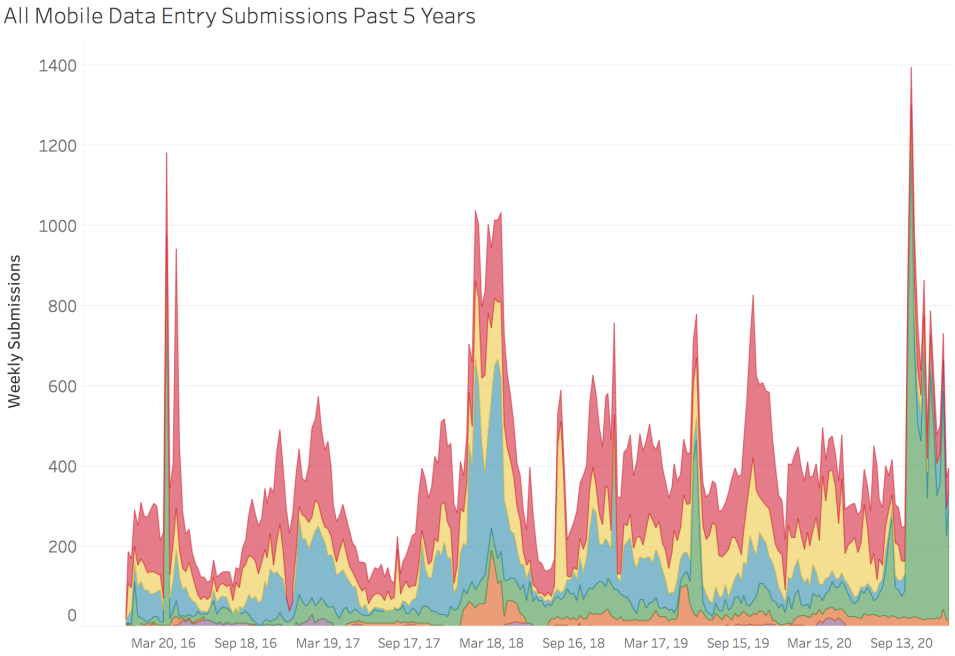

Figure 8: Total Mobile Data Entry Submissions by Region Over Time

The next four figures show the log-log relationship between data submissions (by type) and total cases. The vertical axis shows the log number of data submissions and the horizontal axis shows the total number of COVID-19 cases for GLOBE countries.

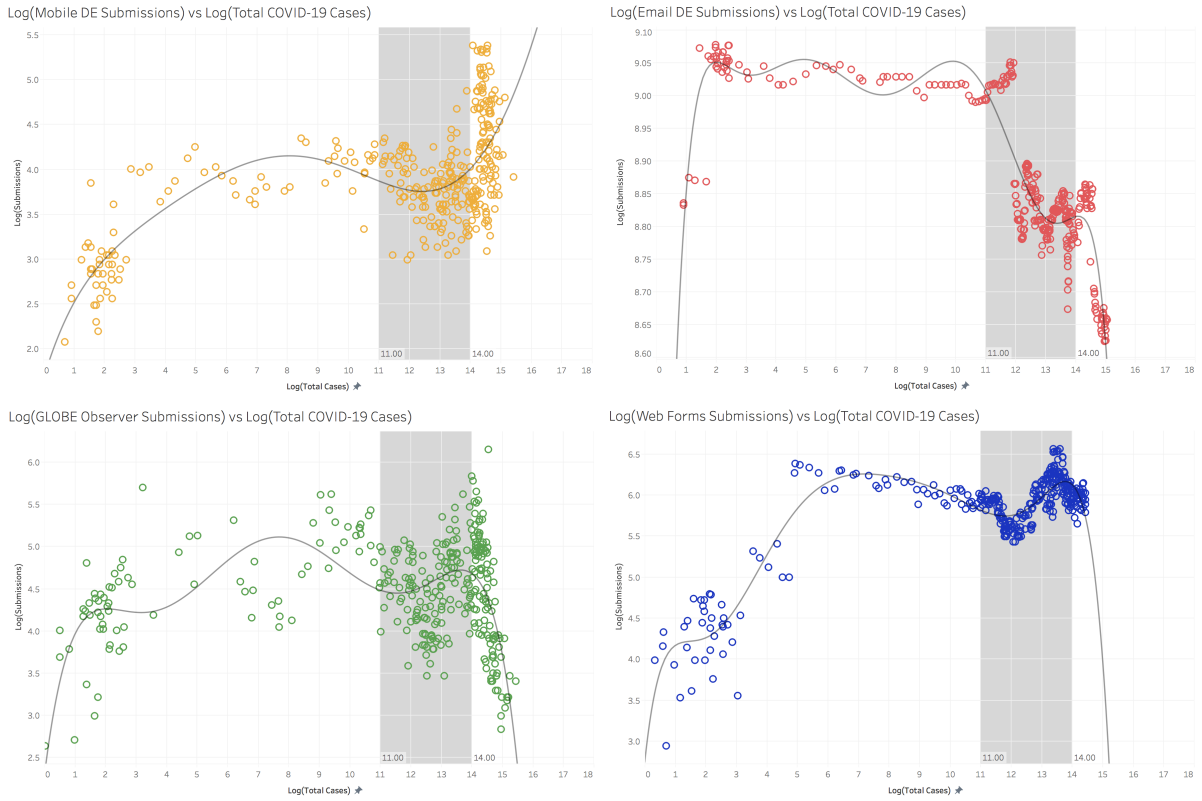

Figure 9: Log-Log View of Submissions vs COVID-19 Cases

The upper-left hand figure with yellow data points shows mobile data entry submissions, the upper-right hand figure with red data points shows email data entry submissions, the lower-right hand figure with green data points shows GLOBE Observer submissions, and the lower-right hand figure with blue data points shows web form submissions.

In each of these figures, the vertical gray bar represents where the log total cases are between 11 and 14 (or in other words, where the cases are between $e^{11} \approx 59,874$ and $e^{14} \approx 1,202,604$). The total COVID-19 cases and submissions represent only GLOBE countries. The log total cases at 11 corresponds with the date March 22, 2020 and the log total cases at 14 correspond with the date May 4, 2020. These figures show a clear dip in log submissions within that time frame.

Next, we look at the total number of data requests from the GLOBE Visualization System by sphere in 2019 and 2020 (Figure 10). The left side charts show 2020 data and the right side charts show 2019 data. The key for the sphere colors can be seen on the far left and far right sides of the figure. The pie charts show the ratio of sphere data requested and the bar charts show the total data submitted by sphere. Shades of orange represent atmosphere protocols, brown shades represent pedosphere protocols, blue shades represent hydrosphere protocols, and green shades represent biosphere protocols.

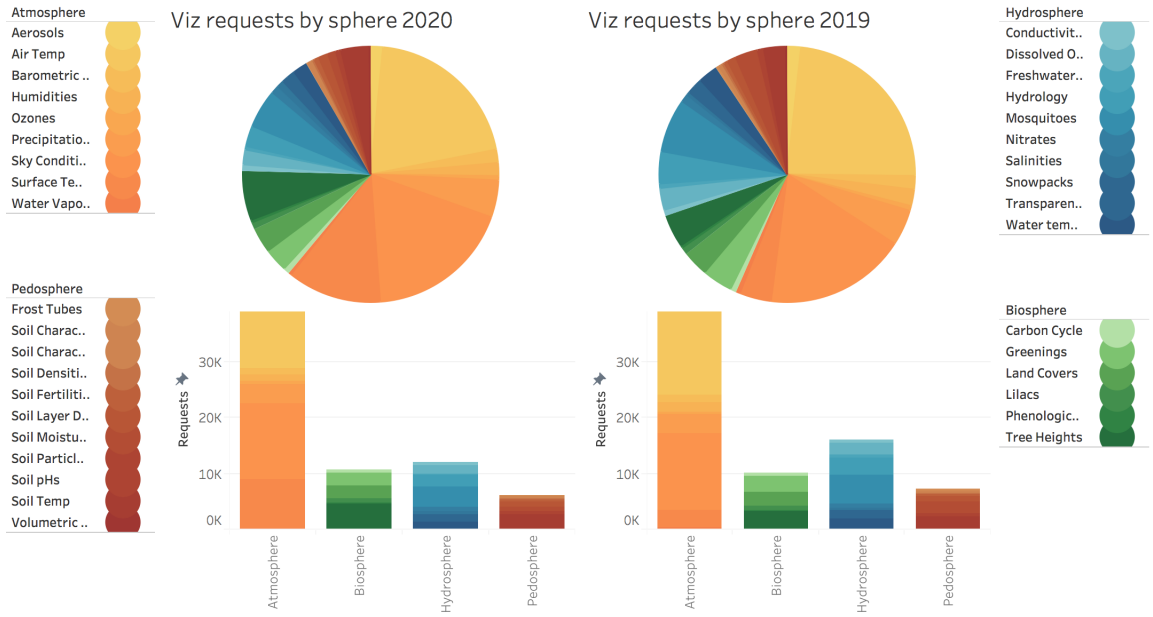

Figure 10: GLOBE Visualization Requests by GLOBE Sphere

Comparing 2019 and 2020, we see that percentage-wise, atmosphere protocols have grown in popularity between 2019 and 2020. Numerically, the total number of requests for atmosphere protocols have remained approximately the same between the two years. Surface temperature was the most requested atmosphere protocol in both 2019 and 2020. Biosphere received slightly more request interest in 2020 compared to 2019 and tree height was the most popular protocol in both years. Hydrosphere and pedosphere show a slight decrease in visualization requests. The most popular protocol requested from hydrosphere in 2019 and 2020 was the mosquito protocol. From pedosphere, the most popular protocol was soil temperature.

Figure 11 shows similar bar charts but for the ADAT system. Here, we are looking at data downloaded from 2017 to 2020 by protocol/sphere. In 2020, although we see a more uniform distribution of data downloaded by sphere, we have nearly a quarter of the interest in atmosphere in 2020 compared to 2018, an equivalent amount of interest in biosphere in 2020 as in 2019, half the interest in hydrosphere in 2020 compared to 2018 and half the interest in pedosphere in 2020 compared to 2018.

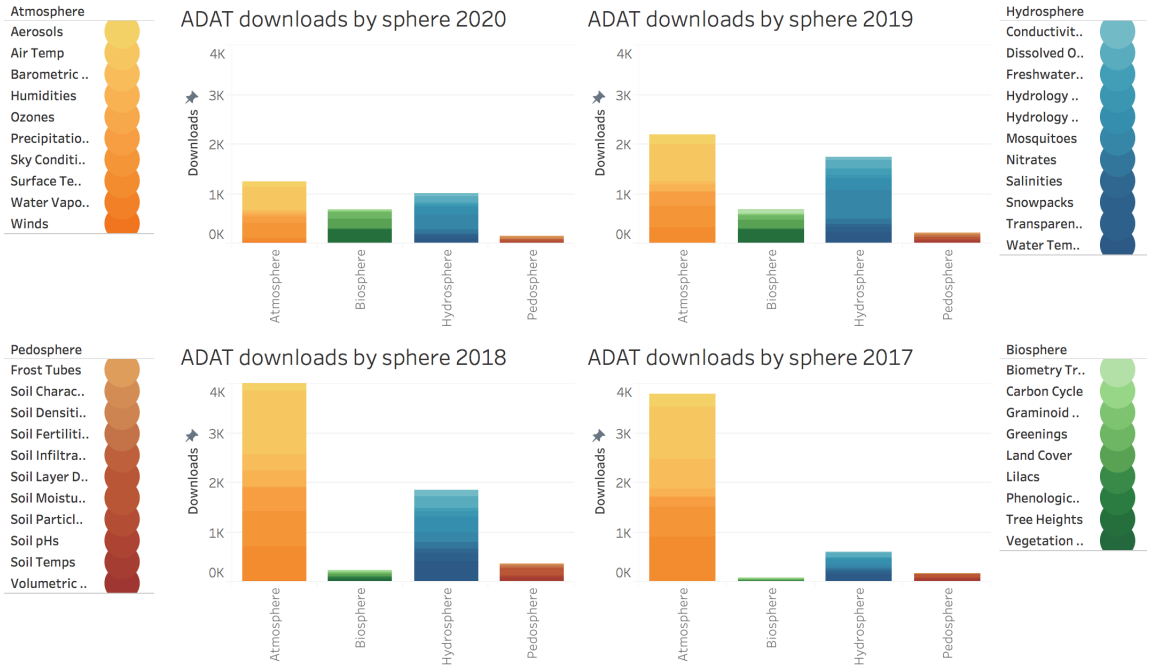

Figure 11: Downloads from ADAT by GLOBE Sphere Totals