Eclipse across America: Citizen Science Observations of the 21 August 2017 Total Solar Eclipse

J. BRANT DODSON AND MARILÉ COLÓN ROBLES

Science Systems and Applications, Inc., Hampton, Virginia

JESSICA E. TAYLOR

NASA Langley Research Center, Hampton, Virginia

CAYLEY C. DEFONTES

*NASA Internships, Fellowships and Scholarships, NASA Langley Research Center,
Hampton, Virginia*

KRISTEN L. WEAVER

Science Systems and Applications, Inc., Hampton, Virginia

(Manuscript received 5 November 2018, in final form 21 June 2019)

ABSTRACT

On 21 August 2017, North America witnessed a total solar eclipse, with the path of totality passing across the United States from coast to coast. The major public interest in the event inspired the Global Learning and Observations to Benefit the Environment (GLOBE) Observer to organize a citizen science observing campaign to record the meteorological effects of the eclipse. Participants at 17 585 observing sites collected 68 620 temperature observations and 15 978 cloud observations. With 7194 sites positioned in the path of totality, participants provide a nearly unbroken record of the cloud and temperature effects of the eclipse across the contiguous United States. The collection of both temperature and cloud observations provides an opportunity to quantify the cloud–temperature relationship. The unique character of citizen science, which provides data from a large number of observations with limited quality control, requires a method that leverages the large number of observations. By grouping observing sites along the path of totality by 1° longitude bins, the errors from individual sites are averaged out and the meteorological effects of the eclipse can be determined robustly. The data reveal a distinct relationship between prevailing cloud cover and the eclipse-induced temperature depression, in which overcast conditions reduces the temperature depression by about one-half of the value from clear conditions. A comparison of the GLOBE results with mesonet data allows a test of the robustness of the citizen science results. The results also show the great benefit that research using citizen science data receives from increased numbers of participants and observations.

1. Introduction

On 21 August 2017, North America experienced a total solar eclipse, with the umbral shadow passing

across the contiguous United States from the West to East Coasts (Fig. 1). It was the first total solar eclipse since 1979 for which totality was visible from the contiguous United States. The path of totality crossed the west coast of Oregon at 1716 UTC, moved northwest to southeast across 14 states, and exited the eastern coast of South Carolina at 1849 UTC. Because the eclipse was relatively easily accessible by the majority of U.S. citizens, and because of the heightened public awareness of the event through both traditional and social media, a large majority of the U.S. population witnessed the eclipse in some form. A preliminary estimate suggests that about 215 000 000 adults viewed the eclipse in

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JAMC-D-18-0297.s1>.

Corresponding author: J. Brant Dodson, jason.b.dodson@nasa.gov

DOI: 10.1175/JAMC-D-18-0297.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

FIG. 1. Path of totality and maximum obscuration values across the CONUS during the eclipse (the image is from NASA's Scientific Visualization Studio).

various manners (the majority of them directly), about 2 times the number that viewed the 2018 Super Bowl event (Miller 2018).

The large public interest in the 2017 eclipse provided an opportunity for a study of the eclipse using citizen science. Scientists have called on the public to provide meteorological observations in the past, often from amateur weather stations (e.g., Hanna 2000). More recently, citizen scientists without well-equipped pre-existing weather stations have provided valuable data for the National Eclipse Weather Experiment, in which citizen scientists in the United Kingdom gathered meteorological observations of the 20 March 2015 partial solar eclipse (Barnard et al. 2016; Hanna et al. 2016; Hanna 2018). Inspired by these past efforts, the Global Learning and Observations to Benefit the Environment (GLOBE) Program (Finarelli 1998; Muller et al. 2015), used the recently released GLOBE Observer mobile application (hereinafter referred to as the GO app) to collect observations from the general public during the 2017 eclipse. A citizen science campaign for the 2017 eclipse was organized and was promoted to the public as “How Cool is the Eclipse? Explore the Sun–Earth Connection with GLOBE Observer.” The GO app was designed as a tool for the general public to collect and submit data the gathering of which does not require equipment or extensive training (GLOBE Observer 2017a).

The project, hereinafter referred to as Eclipse Across America (EAA), capitalized on the public enthusiasm

to gather scientific data about the meteorological effects of the eclipse and utilized the mobile app with basic tutorials supplied within to make cloud observations and air temperature measurements (GLOBE Observer 2017b). Participants collected temperature and cloud data from 17 585 observing sites across North America, about one-half within the path of totality, solely by using the GO app. The large number of participants combined with their ability to travel to meteorologically important areas provided nearly unbroken coverage of the path of totality, with a density that was often much greater than preexisting meteorological observing networks. The volume of observations is critical for conducting robust studies with citizen science observations (Kelling et al. 2015; Aceves-Bueno et al. 2017) and complements automated network data of the eclipse.

The meteorological effects of solar eclipses have been of interest in the atmospheric science community for over a century (Aplin et al. 2016). The relatively large and abrupt externally forced perturbation to top-of-the-atmosphere insolation (as compared with the diurnal and annual cycles) functions as a type of “natural experiment” that displays the atmospheric response to such a perturbation. There are very few kinds of these natural or unplanned experiments—large volcanic eruptions or the study of contrails during the grounding of commercial flights on 11 September 2001 (Travis et al. 2002). This provides a relatively rare set of cases from which to test theoretical understanding and numerical simulations of atmospheric processes.

There have been a number of meteorological effects of solar eclipses identified in literature. While there are several studies about perturbations in wind, pressure, and chemistry (e.g., Clayton 1901; Bojkov 1968; Chimonas and Hines 1970; Anderson and Keefer 1975; Aplin and Harrison 2003; Tzanis et al. 2008), this paper will focus on the temperature and cloud fields. First, the most obvious meteorological effect is the reduction in temperature associated with the reduction in insolation. This cooling typically begins within half an hour of first contact (i.e., when the moon begins obscuring the sun), and maximizes within 20 min of totality (or maximum obscuration for partial and annular eclipses). There have been a wide range of temperature depressions associated with total solar eclipses, ranging from less than 1° to more than 10°C (e.g., Anderson 1999; Eugster et al. 2017). Several factors likely control the magnitude and timing of temperature depression—the time of day or year, the duration of the eclipse and totality, the surface characteristics, the prevailing cloud cover, and so on. Detangling these multiple factors remains an important topic of studying eclipse-induced meteorological conditions.

In this study we will use the GO data to test the previously reported hypothesis that larger cloud cover reduces the eclipse-induced temperature depression substantially enough to be a significant factor in the observed variability in temperature depression. This reduction should theoretically occur through two mechanisms. First, because the temperature depression is caused by the reduction in insolation, and because the strongest cooling occurs at the surface, the presence of clouds should reduce the local temperature depression simply because there is less surface insolation to reduce. Second, clouds reduce the rate of surface cooling by reducing longwave cooling of the surface. This combination of effects should result in reduced temperature depressions in regions with larger cloud cover. Past studies have reported the impact of cloudy conditions on the maximum temperature anomalies observed (e.g., Hanna 2000; Barnard et al. 2016; Hanna et al. 2016). Hanna et al. (2016) found that variations in cloud cover had a greater impact in surface cooling than variations in the eclipse maximum obscuration. There is likely some modulation of the cloud effect from the optical depth and altitude of the clouds; however, previous studies mostly focus on cloud cover.

The primary scientific question this paper will address is, What is the observed relationship between the prevailing cloud cover and the local eclipse-induced temperature depression as seen by EAA? Quantifying the effect of cloud cover is desirable because it potentially modulates the other meteorological anomalies induced by the eclipse, such as perturbations to mesoscale and

synoptic wind fields related to the temperature depression (Aplin and Harrison 2003). We will answer this question both on the national and state spatial scales, and the results will test the hypothesized cloud–temperature relationship. Also, the presence of multiple meteorological observing networks along the path of totality allows us to answer a second question: How do the citizen science observations compare with high density meteorological station observations? Also, because of the large public participation in the eclipse, we have an opportunity to quantify the influence of increased participation on the GO-based results. Not only is this necessary for identifying the limits of certainty of the results, but it also may inform the broader citizen science community of the benefits of encouraging more participation. Thus, we address a third question: What is the effect of public participation numbers on the robustness of the results?

Section 2 describes GO, EAA, and the resulting data, as well as the solar obscuration and meteorological station data. Section 3 explains the method used to calculate the eclipse-induced temperature depression from GLOBE data. Section 4 summarizes the meteorological conditions on the day of the 21 August 2017 eclipse. Section 5 examines the eclipse-induced temperature depression and the influence of cloud cover on a continental spatial scale and quantifies the uncertainty related to the number of GO observations, showing the benefit of greater numbers of participants. Section 6 narrows the focus to the states of Oklahoma and Nebraska, which have mesonets that allow comparison with meteorological station data. Section 7 answers the scientific questions that were posed above and discusses lessons learned from the event.

2. Data

a. GLOBE Observer and eclipse across America

GO is an extension of the international, NASA-funded GLOBE Program created to involve the general public as citizen scientists through data-collection tools facilitated through the app. The GO app was modified to support the eclipse by using the standard cloud data-collection tool already available in the app, with a special feature for air temperature that was temporarily added (GLOBE Observer 2017b). For cloud observations, observers locate, photograph, and classify the cloud type and percent cover of the overhead sky (GLOBE Program 2017). For temperature measurements, participants were asked to acquire a meteorological thermometer, either a traditional alcohol-filled glass model or a digital version, and collect observations in a shaded area (GLOBE Observer 2017c). The goal was to have

TABLE 1. Numbers of observing sites and observations during the 2017 eclipse.

Temperature obs (global)	80 079
Cloud obs (global)	20 172
Obs sites (CONUS)	17 585
Temperature obs (CONUS)	68 620
Cloud obs (CONUS)	15 978
Obs sites (totality)	7194
Temperature obs (totality)	27 312
Cloud obs (totality)	3819
Sites collecting both temperature and cloud obs (totality)	1322

participants measure the eclipse-induced temperature depression directly in the field, to complement the temperature measurements from automated weather stations.

EAA attracted substantial public interest and participation. Table 1 summarizes the number of observing sites, temperature and cloud observations for the eclipse. An “observing site” is defined as a location with unique latitude and longitude information obtained through the GO app for which temperature and/or cloud observations were submitted (GLOBE Observer 2017c). The participants were requested via automated reminders to collect air temperature data at 10-min intervals beginning 2 h before the time of maximum solar obscuration and ending 2 h after. For the period 0.5 h before and after the maximum eclipse, the reminder interval was reduced to every 5 min. In addition, after every third air temperature measurement, the participant was also prompted to report a cloud observation. Cloud cover was recorded as one of six qualitative values, which are listed in Table 2, based on GLOBE’s cloud protocol (GLOBE Program 2017). Further details about GO and the eclipse are provided by GLOBE Observer (2018).

Figure 2 shows the time series of the number of GO observations taken in the contiguous United States (CONUS) during the eclipse, for all observing sites and those in the path of totality. The citizen scientists in the United States provide an unbroken time series of observations from the beginning to end of the 6-h time domain, including in the path of totality. There is a noticeable dip in the number of observations in the 10-min time span before and after maximum obscuration, occurring mainly in the path of totality. However, this does not compromise the ability to estimate the temperature minimum occurring within 0.5 h after the maximum obscuration. The number of observations decreases rapidly within the first 45 min after maximum obscuration and then steadily afterward as many participants likely did not continue viewing the eclipse until the end of partiality.

TABLE 2. Cloud-cover category names and associated estimated coverage values.

Name and abbreviation	Cover range (%)
No clouds (NON)	0
Mostly clear (CLR)	0–10
Isolated (ISO)	10–25
Scattered (SCT)	25–50
Broken (BKN)	50–90
Overcast (OVC)	90–100

EAA was designed to encourage participants to collect and report multiple temperature observations from observing sites during the eclipse. While many did so, overall the number of sites drops roughly exponentially with increasing number of observations. The vast majority of sites reported substantially few data points than requested (about 16 for sites experiencing the mean eclipse duration of 160 min). About one-half of the reporting sites recorded only one temperature observation, which cannot be used to calculate a temperature depression (which, by definition, requires at least two data points). This result both highlights the need for large numbers of participants to collect the data necessary for citizen science–based eclipse studies and provides motivation for future citizen science eclipse campaigns to design methods to encourage citizen scientists to submit multiple observations.

To test the robustness of GO cloud data, the total cloud-cover reports were collocated to geostationary satellites (*GOES-13* and *GOES-15*) and compared with satellite measurements of total cloud cover (Colón Robles et al. 2019). Figure 3 shows the mean collocated satellite total cloud cover, with 1 standard deviation from the mean, in comparison with ground observations based on GLOBE’s total cloud-cover categories (GLOBE Program 2017). Satellite total cloud-cover reports are noticeably less than ground observations in partly cloudy to cloudy conditions, with a mean satellite cloud cover of 66% when GO reports overcast (OVC; 90%–100%). Figure 4 shows the probability density functions (PDFs) for satellite measurement of total cloud cover in comparison with ground observations. Figures 3 and 4 suggest that GO reports and satellite reports have the greatest disagreement when GO reports isolated (ISO) and scattered (SCT), because satellites most often report 0%–10% clear (CLR) conditions.

b. Solar obscuration data

To relate the meteorological observations of the eclipse with the passing of the lunar shadow, we use solar obscuration to characterize the astronomical component of the eclipse. The dataset providing solar obscuration values is provided by NASA’s Scientific

FIG. 2. (a) Time series of the number of observations per minute before, during, and after the time of maximum obscuration. Black lines and red lines represent all observations and the observations in the path of totality, respectively. (b) The number of observing sites stratified by the number of observations collected.

Visualization Studio (Wright 2016). Solar obscuration is the fraction of the solar disc occulted (i.e., covered) by the moon as seen from a given location at a given time. Obscuration values for all locations on Earth are provided at 10-s intervals for nearly the full duration of the eclipse, on an equidistant cylindrical map projection with 0.044° horizontal resolution. This equates to a spatial resolution of about 3.75 km by 4.9 km at 40°N latitude. Errors regarding the position and timing of this dataset are small when compared with the 1-min time resolution of GO data, and 5 min of mesonet data, because the lunar shadow moves roughly an order of magnitude farther than obscuration data resolution (i.e., roughly 50 km min^{-1} at 40°N). Figure 5 shows an example of the reduction of insolation during the eclipse as the shadow passes across the eastern United States, which is related to the solar obscuration and derived from the same data used in this study.

c. Oklahoma and Nebraska Mesonets

A key advantage of the 2017 eclipse is that, in addition to automated measurements from National Weather Service stations, observations are available from several state mesonetworks situated in and near the path of totality. These “mesonets” have spatial densities large enough to sample meteorological variability on spatial scales much smaller than a U.S. state, which is useful for comparison with GO data.

The Oklahoma Mesonet (Brock et al. 1995; McPherson et al. 2007) is a network of 120 environmental monitoring stations covering every county of Oklahoma. Data are collected, quality controlled, and distributed by the

Oklahoma Climatological Survey. The instrumentation collects a wide variety of meteorological information; the fields relevant for this analysis are surface air (2 m) temperature, 10-m air temperature, and total insolation. Observations are provided at 5-min intervals for the 6-h period before, during, and after the eclipse (1500–2100 UTC). At the time of the eclipse, all stations but

FIG. 3. Comparison of cloud-cover (CLDC) category reported by GO and the collocated satellite observations. For each GO category on the x axis, the y axis displays the mean collocated satellite cloud-cover percentage (the open squares) and 1 standard deviation above/below the mean (the asterisks and whiskers). Horizontal dashed lines indicate the percentage threshold cloud-cover values for each GO category according to Table 2; e.g., ISO is 0%–10%.

FIG. 4. PDFs of the geostationary satellite (GEO)-derived CLDC categories per collocating GO observation category of (a) NON, (b) CLR, (c) ISO, (d) SCT, (e) BKN, and (f) OVC. First, each GEO observation is converted into a cover category on the basis of percentage (e.g., a satellite report of 12% is labeled “ISO”). Then the GEO observations are separated by collocating GO category. Last, the PDF of GEO cloud-cover categories for each GO category is calculated. For example, for all GO observations of NON, the collocating GEO reported cloud-cover values consistent with NON 76% of the time, with CLR 15% of the time, with ISO 4% of the time, etc.

one provided continuous coverage of the event. The path of totality passed to the north of Oklahoma; however, all stations observed a maximum solar obscuration value greater than 75%. Hanna et al. (2016) showed such small offsets from totality to have a smaller effect on temperature depression than cloud cover.

The Nebraska Mesonet (Mahmood et al. 2017; Shulski et al. 2018) consists of 67 monitoring stations, which cover the majority of the 93 Nebraskan counties. The Nebraska State Climate Office, part of the University of Nebraska–Lincoln, operates the network. Only surface air temperature and insolation are available, but with the same 5-min resolution and 6-h time range as the other mesonets. At the time of the eclipse, 62 of the stations provided meteorological observations. The eclipse centerline passed across nearly the entire state from west to east, exiting the southern Nebraska–Kansas border in the extreme southeastern corner. Twenty-one mesonet sites experienced 100% maximum solar obscuration. This provides an ideal opportunity to

compare GO observations within the umbral path with mesonet data across the entire state in a wide variety of cloud-cover conditions.

3. Compositing method

The GO data from the 2017 eclipse provide unique challenges to researchers that are different from data provided by automated observation networks and stations staffed by well-trained observers. Data from the latter sources are usually stringently quality controlled and possess quantifiable sources of bias and error, but they are relatively fixed in number and cannot be relocated for special events like eclipses. In contrast, citizen science campaigns such as the EAA provide much greater numbers of temporary observing sites from participants who can move into meteorologically interesting areas such as the path of totality, but are also more difficult to quality control and provide fewer observations per site. Because of these differences, the

FIG. 5. Insolation values at 1825 UTC during the eclipse, derived from the solar obscuration values used in this study (the image is from NASA's Scientific Visualization Studio).

method for analyzing the 2017 eclipse data from GO must be designed differently from one that uses conventional station data, specifically to take advantage of the large volume of observations while minimizing the relatively larger sources of error.

The primary method we use is to combine multiple temperature observations from groups of observing sites located near each other into single composite time series across the 6-h time span of the eclipse. This method is comparable to other studies that use high spatial density networks with low quality controlled datasets (e.g., [Heimann et al. 2015](#)) First, the mean temperature from each observing site is removed from the corresponding time series, to yield a temperature anomaly time series. To reduce noise from insufficient site sampling, only observing sites that contributed at least three temperature observations, and at least one observation both before and after the time of maximum obscuration, are included in the composite time series. Second, the temperature observations for each site are binned into minutes before and after the local time of maximum solar obscuration. For each minute, if there is more than one observation, the mean temperature anomaly is calculated, and is used as the composite temperature anomaly for that minute. A linear interpolation is used to fill minutes without any observations. Finally, a smoothing algorithm using a boxcar average is applied

to reduce the minute-to-minute noise in the composite time series. See Fig. S1 in the online supplemental material for more details of the process of creating the composite.

From the composite time series, the eclipse-induced temperature depression is calculated as the difference in maximum temperature before totality (or maximum partiality) and the temperature minimum near totality. The timing of the depression is the difference between the time of totality and the time of the temperature minimum.

To examine the effect of cloud cover on temperature variability, the observing sites along the path of totality are grouped by 1° longitude bins, and a composite temperature time series is calculated for each longitude bin. In addition, the prevailing cloud cover is calculated for each longitude bin and is categorized based on GO's total cloud-cover categories. The prevailing cloud cover is calculated as the statistical mode of all cloud observations collected in the longitudinal bin over the duration of the eclipse. The longitudinal bins are then grouped by cloud cover, and a single composite time series is calculated from the longitudinal bin composites (i.e., a composite of the composites). When calculating this single composite, the longitudinal bin composites are weighted by the number of sites in each longitudinal bin, so to not bias the single composite toward bins with

FIG. 6. Weather maps depicting the synoptic weather conditions during the 21 Aug 2017 eclipse: (a) analysis of surface conditions (12-h QPF) from the Weather Prediction Center (WPC) for 1800 UTC, (b) 12-h quantitative prediction of precipitation (in.; 1 in. = 2.54 cm) from the WPC beginning 1200 UTC, and (c) composite *Aqua* MODIS true-color imagery of CONUS cloud cover on the day of the event. Overpasses were at 1330 local standard time. National and U.S. state boundaries are outlined in gray. The image gives the general location of cloudy and clear regions of the CONUS without major interference from the umbral shadow, as might occur with geostationary imagery.

fewer sites. In effect, this approach allows sites that reported temperature but not cloud cover to “borrow” cloud-cover observations from sites nearby that did.

When examining the results, it is necessary to use statistical testing to determine when the temperature depressions calculated from different composite time series are in agreement, or when there is significant disagreement. The testing method is discussed in the [appendix](#), and statistical results described in [sections 5 and 6](#) are calculated from this method.

4. Overview of event and meteorological conditions

a. Synoptic meteorological maps

The 2017 eclipse path crossed a range of meteorological conditions in North America. [Figure 6a](#) displays a surface analysis from the Weather Prediction Center at 1800 UTC 21 August, when the eclipse was ongoing. The temperature spread across the CONUS was about 18°C and was about 14°C along the path of totality, from 19°C (66°F) in western

Oregon to 33°C (91°F) in North Carolina. Note that the temperatures are instantaneous, not averages, and are likely reduced a few degrees by the eclipse. There was a low pressure center and associated weak frontal boundary in the Midwestern United States. This concurred with a region of precipitation ([Fig. 6b](#)) and large cloud cover ([Fig. 6c](#)) across several Midwestern states, including in the path of totality. Other regions with cloud cover and precipitation include New Mexico and the Texas Panhandle, and the southern East Coast concurrent with another frontal boundary.

b. Cloud observations from GLOBE Observer

[Figure 7](#) summarizes the GO observations for the eclipse, displaying the mean temperature, temperature depression, and prevailing cloud cover recorded by every observing site across the 6-h time range of 1500–2100 UTC. The prevailing cloud-cover map ([Fig. 7a](#)) shows a wide range of cloud conditions across the United States, including in the path of totality. In general, the west had clearer skies than the east, which is

FIG. 7. (a) The GO observations of prevailing CLDC across the 6-h time span (1500–2100 UTC) during the 21 Aug 2017 eclipse. The definitions of cloud category are given in Table 2. Triangles with black outlines represent observing sites that reported at least two temperature observations and thus can be included in the composite time series. The black line spanning the CONUS is the path of totality (specifically the centerline of the eclipse), and the red lines are contours of maximum obscuration. (b) The mean temperature (TAIR) across the 6-h time span. (c) The temperature depression (TDEP) calculated from temperature reports from each individual site. Gray triangles represent sites that reported only one temperature observation or did not report at least one temperature observation both before and shortly after the local time of maximum obscuration.

consistent with the summer climate of the United States (Warren et al. 1986). The most notable overcast conditions occurred in the Midwest, extending north from Kansas and Missouri into the Great Lakes region and the Dakotas. The overcast condition consisted largely of optically thick clouds and is consistent with the satellite and precipitation maps. Because GO participants took many cloud and temperature observations in this region, despite the eclipse itself being obscured, this provides an opportunity to examine the effect of the regional cloud cover on the regional temperature depression. Other cloudy regions, such as the southern East Coast and New Mexico are also seen in the GO data.

c. Temperature observations from GLOBE Observer

The observations of mean temperature (Fig. 7b) provide a quick check of the robustness of the data; however,

they do not necessarily correspond exactly with the instantaneous temperatures shown in Fig. 7a. As the eclipse occurred during boreal summer, the GO-observed temperature spread from $\sim 15^{\circ}\text{C}$ in the northwest to $\sim 35^{\circ}\text{C}$ in the south is appropriate for the season. Nevertheless, disagreement between adjacent sites of 3°C or more are not uncommon (e.g., around southeast Texas, central Oklahoma, and southern Wisconsin). These discrepancies may arise from temperature gradients in the real atmosphere, or from a variety of measurement artifacts. The temperature discrepancies in Oklahoma are further examined using mesonet data in section 6.

d. Temperature depression observations from GLOBE Observer

Figure 7c shows an estimate of the eclipse-induced temperature depression across the CONUS. The temperature

FIG. 8. (a) The GO-reported CLDC by longitude. Dots indicate the CLDC mode for individual stations; solid lines indicate the CLDC mode for each 1° longitude bin. Red and black colors indicate observations in the path of totality and all observations, respectively. (b) As in (a), but for TDEP.

depression is calculated for each observing site individually by subtracting the lowest temperature recorded within an hour after maximum obscuration from the highest temperature reported within 2 h before maximum obscuration. The temperature depression could not be calculated for all sites, as not all reported temperature frequently enough to allow a calculation. The results show a range of reported temperature depressions from 0° to 10°C, and even a few negative depressions (i.e., a temperature rise during the eclipse). Furthermore, there are frequent disagreements between neighbors of 3°C or more. The magnitude of discrepancies is much more important for temperature depression than mean temperature, because the associated uncertainty is the same order of magnitude as the temperature depressions themselves.

To more precisely illustrate the distribution of meteorological variables across the CONUS, Fig. 8 displays the cloud cover and temperature depressions sorted by longitude. Cloud cover is overall greater east of 105°W than west (the approximate longitude of the Rocky Mountains) and is noticeably apparent along the path of totality. The highest values of temperature depression are found to the west of 105°W, and the spread of temperature depression values is larger in the west than the east. This suggests an association between cloud cover and temperature depression controlled by location relative the Rockies; this is addressed in section 5d and the Nebraska results.

Figure 9 shows the histograms of temperature depression magnitude and timing, for both all data and for sites in the path of totality. The histogram mean, median, and standard deviation values are shown in Table 3. The standard deviation for sites in totality is slightly larger than that for all sites (0.08°C); however,

sites in totality reported depressions more heavily skewed toward large values and away from negative values. The mean temperature depression for sites in totality is 0.86°C larger than the mean for all sites (p value from two-tailed t test < 0.01), which is consistent with the greater reduction in insolation for sites in totality.

The mean temperature minimum for all sites in the CONUS occurred 91 s later than that for only sites in totality. However, the histogram of temperature minimum times has a long tail stretching almost an hour and a half after the time of maximum obscuration. It is unclear whether these long delays in the temperature minima are physical or caused by sites reporting temperature observations long after the time of maximum obscuration, in effect delaying the apparent timing beyond the actual meteorological time of the temperature minimum. This possibility is supported by a slight but statistically significant negative correlation (-0.20 ; p value < 0.01) between the temperature depression magnitude and timing. This issue suggests that the limitations of the GO data prevent a meaningful examination of the timing of the temperature depression, so determining the amplitude is the main focus of the results.

5. Composite time series of temperature for the contiguous United States

a. Mean composite time series

Figure 10 displays the temperature composite time series for all observing sites in the CONUS, and the composites for only sites inside the path of totality and outside the path. The composites show a clear single temperature depression beginning about an hour before

FIG. 9. Histograms of (a) TDEP and (b) time of temperature minimum (TTMN) calculated from observing sites for all of the CONUS (black curve) and from those along the path of totality (red curve). The thick black or red vertical line depicts the mean temperature depression amplitude [for (a)] and timing [for (b)] for all sites or for sites along the path of totality, respectively. Histogram bin sizes are 1°C and 1 min, respectively.

maximum obscuration, and reaching a maximum (i.e., temperature minimum) within 15 min after. The compositing method depicts the mean temperature depression caused by the eclipse across the entire United States (after smoothing) as 2.38°C (Table 4). This is about 0.9°C less than the mean temperature calculated from the individual temperature depressions, and likely reflects the lower sensitivity of the compositing method to outlier values. Observations in the path of totality have larger temperature depression than those outside because of the larger reduction in insolation. The temperature depression of the smoothed composite time series is 3.12°C , significantly different (p value < 0.01) from the temperature depression outside the path (1.90°C).

How do these results compare with past observations of eclipses? Anderson (1999) lists the temperature depression amplitude and timing for eight events (including three annular eclipses, for which maximum solar obscuration values are typically 90% or greater). The 2017 temperature depressions for all sites and for those in the path of totality fall on the low end of most past observations, with the exceptions of the 1979 total eclipse (2.0°C) and the 1994 annular eclipse for one of two observations (0.4°C).

More recent eclipses have temperature effects similar to the 2017 event. For example, a total eclipse observed in the United Kingdom and parts of mainland Europe on 11 August 1999 produced temperature depressions of $2.0^{\circ}\text{--}3.0^{\circ}\text{C}$ in clear regions near the path of totality (6°C locally), and $0.5^{\circ}\text{--}2.0^{\circ}\text{C}$ in overcast regions in the path of totality (Hanna 2000). For the 29 March 2006 total eclipse across Africa and Eurasia, observations taken in and near Greece (which has obscuration of at least 75%) showed temperature depressions spanning $1.6^{\circ}\text{--}3.9^{\circ}\text{C}$ in clear to partly cloudy conditions (Founda et al. 2007). The NEWex citizen science campaign for the 20 March 2015 eclipse in the United Kingdom (Barnard et al. 2016) found a mean temperature drop of 2.2°C . The eclipse was partial, with maximum obscuration ranging from 85%–95% across the United Kingdom, and was observed in a variety of sky conditions from clear to overcast.

b. Sensitivity of composite time series to prevailing cloud cover

To test for the hypothesized cloud cover–temperature depression relationship using the 2017 eclipse data, we subdivide the data by GO-reported cloud cover.

TABLE 3. Temperature depression amplitude (TDEP) and timing (TTMN) for the histogram of CONUS GO individual observing sites—for all sites and for those in the path of totality.

	TDEP mean ($^{\circ}\text{C}$)	TDEP median ($^{\circ}\text{C}$)	TDEP σ ($^{\circ}\text{C}$)	TTMN mean (min)	TTMN median (min)	TTMN σ (min)
All sites	3.25	2.78	2.94	17.88	15.00	14.03
Path of totality only	4.11	3.89	3.02	16.36	13.00	12.68

FIG. 10. The composite time series of temperature change from GO data, for (a) all observing sites, (b) only observations from the path of totality, and (c) only those outside totality, from 3 h before to 3 h after the time of maximum obscuration. The black line is the unsmoothed composite time series, the red line is the smoothed time series, and the gray area is the 95% confidence interval of the unsmoothed time series. The numbers of observations are shown in Table 1.

Our initial attempts revealed that the GO data becomes sensitive to sample size when the number of observing sites included in a composite time series is reduced well below 100. This necessitates an approach that maximizes the number of sites included in each composite. Because the United States spans 45° longitude along the path of totality, grouping the 7194 GO sites in the path of totality into 1° bins gives about 160 sites per bin on average. This helps to mitigate the problem of low sample size obscuring the relationship.

Figures 11a–f shows the composites, and the cloud–temperature relationship becomes very clear. The temperature depression decreases nearly monotonically as the prevailing cloud cover increases, from 4.08°C for no clouds (NON) to 1.74°C for OVC (Table 5). This result is statistically significant (p value < 0.01) and is consistent with the hypothesized relationship and correlate with findings by Hanna et al. (2016) of the importance of cloud cover in controlling the surface temperature response during an eclipse. Broken (BKN) and OVC have similar temperature depressions. Because OVC does not necessarily mean 100% cloud cover, this similarity may result from GO participants not clearly distinguishing between BKN and OVC. This may have been difficult in the convective weather in the Midwest, as convective systems often produce multiple cloud types resulting in a complex-appearing sky that changes rapidly with time. Figures 3 and 4 show the difficulty of cleanly separating OVC and BKN.

The systematic overestimation of cloud cover by GO relative to satellites suggests that the GO results may underestimate the amplitude of the cloud effect on the temperature depression. Figure 3 shows that the OVC

category has a satellite-derived mean cloud cover of 66%. If this apparent overestimate is true, then the actual effect of 100% cloud cover may be approximately 20% larger than indicated by GO, assuming that the cloud–temperature relationship is approximately linear as the data suggest.

c. Composite results for the eastern contiguous United States (ECONUS) only

The results in Fig. 11 and Table 5 seem to confirm the hypothesized cloud–temperature relationship. However, there may be other explanations for this apparent relationship. Cloud-cover variability was not distributed evenly across the United States during the eclipse. Rather, clearer or cloudier skies respectively occurred mostly in the western or eastern United States (Fig. 8a). The western United States has an overall drier climate than the eastern United States, with drier soil and less vegetation, and a resulting larger Bowen ratio (i.e., the ratio of surface sensible heat flux to latent heat flux) (Anderson et al. 2007). This means that the west may have greater surface cooling than the east because of climatological surface conditions and not solely cloud cover (though cloud cover and surface moisture are closely linked).

TABLE 4. Temperature depression amplitude and standard deviation for the CONUS GO time series, for all observing sites, those in the path of totality, and those outside the path (i.e., partiality only). In addition to raw values, smoothed values are given in parentheses.

	TDEP ($^\circ\text{C}$)	TDEP σ ($^\circ\text{C}$)
All sites	2.71 (2.38)	2.19 (2.21)
Path of totality only	3.71 (3.12)	2.14 (2.34)
Partiality only	2.31 (1.90)	2.27 (2.06)

FIG. 11. The composite time series of temperature change from GO observations for the CONUS data subdivided into 1° longitudinal bins along the path of totality and then grouped by prevailing cloud-cover amount: (a) none, (b) clear, (c) isolated, (d) scattered, (e) broken, and (f) overcast. Black or red lines are the unsmoothed or smoothed composite time series of the composites, respectively, and the gray area is the 95% confidence interval of the unsmoothed time series. (g)–(l) As in (a)–(f), but limited to the eastern CONUS (east of 102°W).

TABLE 5. Number of longitude bins, number of observing sites (No. obr), number of observations (No. obs), and temperature depression amplitude and standard deviation for the CONUS and ECONUS composite time series in the path of totality, by cloud-cover type. In addition to raw values of amplitude and timing, smoothed values are given in parentheses.

	No. lon bins	No. obr	No. obs	TDEP (°C)	TDEP σ (°C)
CONUS					
NON	17	1802	7280	5.16 (4.08)	2.76 (2.41)
CLR	8	1303	5412	4.96 (3.47)	1.94 (2.03)
ISO	4	244	1093	4.54 (3.33)	2.28 (2.20)
SCT	9	2956	10 771	3.31 (2.38)	1.92 (1.91)
BKN	4	446	1467	3.72 (1.73)	1.08 (1.49)
OVC	3	439	1289	3.71 (1.74)	1.90 (1.87)
ECONUS					
NON	0	0	0	— (—)	— (—)
CLR	4	1109	4557	3.73 (2.64)	1.32 (1.18)
ISO	3	244	1093	4.54 (3.33)	2.15 (2.11)
SCT	9	2956	10 771	3.31 (2.38)	1.92 (1.91)
BKN	4	446	1467	3.72 (1.73)	1.10 (1.49)
OVC	3	439	1289	3.71 (1.74)	1.96 (1.80)

To test the possible influence of surface climatology, the data spatial domain is restricted to the eastern United States, east of 102°W, and the composites are recomputed (Figs. 11g–i). The dividing longitude is based on Fig. 8b, which suggests a break in the statistical properties of the eclipse temperature effect near this longitude. The reduced spatial domain excludes the western Great Plains, the Rocky Mountains, and the U.S. Northwest. The remaining longitude bins cover portions of the United States with thicker vegetation and wetter soil, so the effect of cloud cover on temperature depression is less obscured by surface conditions. Because the clearer conditions occurred primarily in the western United States, the restriction of the spatial domain only affects the CLR category. The temperature depression generally decreases as cloud cover increases, although not monotonically. CLR has a temperature depression about 0.6°C lower than ISO (p value < 0.01) but still 0.3°C larger than SCT (p value < 0.01). There is 0.9°C separation between CLR and OVC (p value < 0.01), supporting the hypothesized cloud–temperature relationship.

The unexpected small value of the CLR temperature depression is not easily explained. It is likely not merely an error in the GO-reported cloud cover; testing for this possibility using the collocated satellite data yields no evidence of this. Figure 11h shows that CLR in the eastern CONUS has a larger confidence interval earlier than 1 h before the eclipse than the other categories in Fig. 11. It is possible that some source of uncertainty in the GO data is affecting the composite pretotality temperature maximum, yielding the unexpectedly low temperature depression and limiting the ability to determine more precise quantitative values. Whether this arises from unknown issues with the data, or from natural variations in meteorology not related to cloud

cover, cannot be determined with the compositing methodology. Creating methods to determine an exact source of uncertainty in the data—or whether one exists—is the topic of ongoing work in the GO community.

d. Quantitative robustness of composite time series

In addition to the scientific value of these results, they also inform us of the robustness of the citizen science data and useful methodologies for deriving reliable results. Because the citizen science data are not as rigorously quality controlled as conventional ground station and satellite data, individual observing sites must be grouped to give sample sizes large enough for the individual errors to average out. The results presented in this section clearly demonstrate the ability to derive scientifically useful estimates of eclipse-induced meteorological variability, given a sufficiently large number of participants. It should be noted that the specific number of necessary participants may vary between different citizen science efforts, requiring researchers to check their specific data to determine that number. Thus, we must develop a process of determining the required number for this dataset using this methodology.

The GO results show an obvious quantitative distinction in temperature depression between clear and cloud scenes. However, it is not clear whether the differences in temperature depression are large enough to be significant, given the lower quality control of GO data relative to conventional meteorological station data. To test the robustness of the data, we need to estimate the uncertainty of the temperature depression values given by the composite time series. The nature of the GO data makes simple statistical tests difficult, so instead we use Monte Carlo–style tests to determine the spread of temperature depression values, and in particular composites constructed from low numbers of sites.

FIG. 12. Temperature depressions from the Monte Carlo tests from 1° longitudinal bins along the path of totality at (a) 87° , (b) 84° , and (c) 123° W. Gray pluses mark a single temperature depression for one trial (of 100) for a given sample size (from 10 to 100 samples). The solid black line is the mean temperature depression for each sample size, and the dashed lines are the $+1$ and -1 standard deviation bounds at each sample size.

Because the test relies on random samples of sites, large populations of sites are required. First, we select three 1° longitude bins along the path of totality that contain large numbers of sites. The bins at 84° , 87° , and 123° W all contain 200+ sites with at least two temperature observations, and therefore they are chosen for the experiments. Second, each bin has 10 sites chosen randomly, and from the 10 random sites the (smoothed) composite time series, temperature depression, and time of temperature minimum are calculated and recorded. Third, the sampling and compositing are repeated 100 times for each bin, yielding PDFs of temperature depression for site sample sizes of 10. Fourth, the 100 trials are repeated for sample sizes of 15, 20, 25, and so on, up to 100.

The results (Fig. 12) show that the PDFs for temperature depression have much greater spread for low sample sizes than high. Above a sample size of 50, the standard deviation of the temperature depression is $\sim 0.4^\circ\text{C}$; for a sample size of 100, it is $\sim 0.33^\circ\text{C}$. Because two standard deviations contain about 95% of the data in a normal distribution, a difference between temperature depressions greater than 2 standard deviations, or about 0.83°C , is likely to be a statistically significant difference when the number of sites contributing to the composite is greater than 50. For site numbers below 50, the standard deviation is 0.55°C or larger, which makes comparisons of temperature depressions problematic.

6. State composite results

The results for the United States provide strong evidence of the hypothesized cloud–temperature relationship.

However, it is also useful to check the GO temperature observations against automated station data to test the robustness, and determine the degree of bias and artificial spread, in the GO data. The relative high station density of mesonets is very useful for this task, as it provides information about the spatial variability of temperature at a much finer scale than National Weather Service stations.

a. Oklahoma

Oklahoma provides a useful opportunity to test the robustness of the data in cloud-free conditions. Most of Oklahoma did not have large variability in cloud cover during the eclipse (Fig. 13a), so discrepancies between GO and station temperature data cannot (Figs. 13b,c) be attributed to cloud variations.

Figure 14 shows the composite time series for Oklahoma Mesonet and GO temperature data for three spatial domains. The first is the entirety of Oklahoma minus the Panhandle, which was excluded because it is drier and at a higher altitude than the rest of the state (Arndt 2012). The latter two are 1° by 1° boxes centered around Oklahoma City and Tulsa, Oklahoma, that contained clusters of GO sites on the day of the eclipse (Fig. 13c). Note that we compare the mesonet composite time series with the smoothed GO time series (i.e., the red line), rather than the unsmoothed data, to stay consistent with the analysis of U.S. composite time series.

GO depicts a statewide temperature depression of 1.88°C , which is 0.14°C lower than the mesonet-observed depression (Table 6). The difference is not statistically significant by the standard threshold of 0.05 (p value of 0.33), so GO and the mesonet appear to be in approximate agreement on the statewide spatial scale. Looking at

FIG. 13. As in Fig. 7, but focused on Oklahoma. In addition, in (b) and (c) the circles represent data from the Oklahoma Mesonet.

smaller regions, varying the mesonet stations included in the mesonet composite time series through spatial subsetting appears to create a variance in the mesonet temperature depression of about 0.25°C (p value 0.98), spanning from 1.87°C when focused on Oklahoma City to 2.12°C when including all of central and northeast Oklahoma. This “real” temperature variability can also be used to explain the 0.17°C discrepancy between the GO and mesonet observations for Tulsa (although with a p value of 0.24, it is not clear that explanation is needed). However, the discrepancy in Oklahoma City between GO and the mesonet of 0.72°C is $\sim 0.45^{\circ}\text{C}$ larger than any observed spatial variation in the mesonet and is statistically significant (p value of 0.02). A likely contributor to the large discrepancy is the relative low number of GO sites in Oklahoma City included in the composite time series (33; see section 5d). The similar number of GO sites in Tulsa (33) also suggests that the relatively small discrepancy between GO and mesonet may be coincidental. However, the larger number of sites across the state (107) yields a more robust estimate of the temperature reduction. Similarly, the temperature depression difference of 0.21°C between GO and

mesonet for central and northeastern Oklahoma falls within the range of mesonet-observed spatial variability, which can be expected with the larger number of GO sites (99).

b. Nebraska results

Unlike Oklahoma, Nebraska had a wide variety of cloud conditions across the state during the eclipse (Fig. 15). Cloudiness generally increased from west to east, with clear skies in the Panhandle and overcast skies associated with convection south of Omaha. This large variability in cloud cover allows a useful test of the hypothesized relationship between temperature depression and cloud cover as depicted by both GO and the Nebraska Mesonet.

Figure 16a shows the composite time series of GO data for all Nebraska, and Fig. 16b shows the composite time series in the path of totality. The temperature depression of 2.53°C in the path of totality is smaller than the national mean (Table 7). The zonal cloud-cover gradient across Nebraska makes longitude a useful proxy for cloud cover, which can be used for direct comparison with mesonet data (which does not report

FIG. 14. Composite time series of temperature change from GO observations in Oklahoma: (a) results from all observing sites in Oklahoma, (b) the data from a 1° by 1° bin centered on Oklahoma City (OKC; 35°–36°N, 98°–97°W), (c) the data from a 1° by 1° bin centered on Tulsa (TUL; 35.5°–36.5°N, 96.5°–95.5°W), (d) data from the combination of sites from both Oklahoma City and Tulsa, and (e) data from Oklahoma north of 35°N and east of 98°W. Black or red lines are the unsmoothed or smoothed composite time series, respectively, and the gray area is the 95% confidence interval of the unsmoothed time series. (f)–(j) As in (a)–(e), but using Oklahoma Mesonet data. Here, the smoothed composite time series are omitted because they are almost identical to the unsmoothed data.

TABLE 6. Number of observing sites/stations (No. obr/sta), number of observations, and temperature depression amplitude and standard deviation for GO and Oklahoma Mesonet composite time series by region of Oklahoma. In addition to raw values of amplitude and timing, smoothed values for GO data are given in parentheses.

	No. obr/sta	No. obs	TDEP (°C)	TDEP σ (°C)
GO				
All of Oklahoma except the Panhandle	107	674	4.48 (1.88)	2.35 (1.47)
Central and northeastern Oklahoma	99	638	4.48 (1.91)	2.35 (1.46)
Oklahoma City	31	195	7.32 (2.59)	1.86 (1.79)
Tulsa	33	213	4.86 (1.79)	0.69 (0.92)
Oklahoma City + Tulsa	64	408	5.87 (1.87)	1.47 (1.46)
Oklahoma Mesonet				
All of Oklahoma except the Panhandle	121	8712	2.02	0.57
Central and northeastern Oklahoma	66	4752	2.12	0.39
Oklahoma City	9	648	1.87	0.40
Tulsa	9	648	1.97	0.47
Oklahoma City + Tulsa	18	1296	1.92	0.44

cloud cover). To test the effect of cloud cover on the temperature depression, the GO observations in the path of totality are binned by 2° longitude bins, and the composite time series calculated for each longitudinal bin. The temperature depression decreases by approximately 1.1°C from west to east (p value < 0.01), with the largest drop occurring from center-west to center-east (p value < 0.01 , Figs. 16d,e). This is consistent with the hypothesized cloud–temperature relationship, although the large amount of noise in the data associated with the limited number of observations precludes a more precise calculation of the contrast in temperature depression between west and east.

The mesonet data are also binned by longitude, and the composite time series are shown in Figs. 16g–i. The temperature depression decreases from far west to far east, similar to the GO results. This strongly supports that the GO result is indeed picking up on a real signal, and not merely a result of noise. One notable discrepancy is the depression for central west (1.96°C) is lower than that of central east (2.71°C), a statistically significant difference (p value < 0.01) that is not the case for GO. There are a few possibilities that may explain this discrepancy. First, the number of GO sites for the center west longitude bin is 34, which may be too small of a sample for a robust result according to the Monte Carlo tests (see section 5d). This raises questions about the trustworthiness of the GO results in this longitude box. Second, examining the individual station time series shows that for central west the times of the temperature minima were spread across the hour after totality, which results in a smaller (less negative) temperature minimum in the composite time series. The other longitude bins do not show this degree of spread in the timing of the temperature minimum. This is an unfortunate limitation of the mesonet data, having only a few sites in the

path of totality for each longitude bin, and so it is not possible to determine if the discrepancy is real or a sampling artifact. Third, this longitude bin exists in the gradient zone between clear skies to the west and cloudy skies to the east (see Fig. 15a), and the varying cloud cover between stations affected the observations. Finally, the warm front shown in Fig. 6a may have influenced the temperature time series. Regardless of the cause, the results show that such discrepancies between the GO and mesonet data are not sufficiently large to mask the agreement in the cloud–temperature relationship.

7. Summary and conclusions

The “Great American” total solar eclipse on 21 August 2017 provided a significant opportunity for a citizen science campaign to observe cloud and temperature properties associated with the eclipse. The large number of observations from GO allowed for the testing of a hypothesized relationship between prevailing cloud cover and eclipse-induced temperature depression on both national and regional scales. In addition, the path of totality passed over or near multiple mesonet stations, allowing cross comparisons between GO observations and meteorological stations data with high spatial density.

The challenge of using the citizen science observations to produce robust results is developing a methodology that takes advantage of the large volume of observations to compensate for the lack of rigorous quality control for individual observations. We use the GO observations to calculate prevailing cloud cover and composite time series of temperature before, during, and after the eclipse in 1° and 2° longitudinal bins along the path of totality. This approach maximizes the number of observing sites contributing to the prevailing cloud cover and temperature observations

FIG. 15. As in Fig. 13, but focused on Nebraska and its mesonet.

on regional spatial scales, and it allows the examination of the cloud–temperature relationship. In addition, it provides a method of relating the mesonet station observation with the GO data on regional and statewide spatial scales.

These results provide strong evidence of the modulation of the eclipse-induced temperature depression by cloud cover and yield a quantitative estimate of the relationship when other modulating factors such as differences in regional climatology and surface conditions are reduced. Specifically, the results yield answers to three primary scientific questions as follows.

- a. *What is the observed relationship between the prevailing cloud cover and the local eclipse-induced temperature depression as seen by citizen scientists?*

While the effect of cloud cover on the timing of the temperature minimum shortly after maximum obscuration

cannot be determined because of the magnitude of uncertainty in the data, the amplitude of the effect can be robustly detected. Local regions with larger cloud cover tend to have smaller temperature depressions than those with lower cloud cover. The regions of large cloud cover during this eclipse were associated with deep convective systems and frontal boundaries, and the cloud cover was often optically thick. The reduction in temperature depression in overcast conditions relative to clear conditions is at least on the order of 50%, similarly observed in past studies (Hanna 2000; Hanna et al. 2016). This relationship is apparent in the GO data on both a continental scale (specifically the contiguous United States east of 102°W) and on a U.S. state scale (e.g., Nebraska). The overall conclusion is likely not merely a coincidence of cloud cover being influenced by other local climatological factors (e.g., surface type), as the relationship holds even in regions that are climatologically similar. Nor is the conclusion

FIG. 16. As in Fig. 14, but for composite time series of temperature change from GO observations in Nebraska: (a) results from all observing sites in Nebraska, (b) data limited to sites in the path of totality, and data from the path of totality subdivided into 2° longitudinal bins as (c) “far west” (104° – 102° W), (d) “center west” (102° – 100° W), (e) “center east” (100° – 98° W), and (f) “far east” (98° – 96° W). (g)–(l) As in (a)–(f), but using the Nebraska Mesonet data.

TABLE 7. Number of observing sites/stations, number of observations, and temperature depression amplitude and standard deviation for GO and Nebraska Mesonet composite time series by region of Nebraska. In addition to raw values of amplitude and timing, smoothed values for GO data are given in parentheses.

	No. obr/sta	No. obs	TDEP (°C)	TDEP σ (°C)
GO				
All of Nebraska	476	1984	3.98 (2.30)	2.45 (1.91)
Nebraska in path of totality	388	1656	5.23 (2.53)	2.65 (1.85)
Nebraska—far west	90	317	8.82 (3.41)	2.16 (1.57)
Nebraska—central west	34	245	6.11 (3.54)	1.22 (1.95)
Nebraska—central east	112	448	6.38 (2.61)	3.22 (1.96)
Nebraska—far east	152	646	5.04 (2.29)	3.01 (1.65)
Nebraska Mesonet				
All of Nebraska	62	4464	2.03	1.61
Nebraska in path of totality	21	1512	1.95	1.05
Nebraska—far west	4	288	3.2	1.04
Nebraska—central west	4	288	1.96	1.06
Nebraska—central east	8	576	2.71	0.57
Nebraska—far east	5	360	1.24	1.63

just a chance of noise in the GO data, as shown by the statistical tests and Monte Carlo results. However, the details of the relationship in the GO data may be affected by these other factors, limiting the ability of GO data to provide an exact quantitative value of the cloud–temperature relationship.

b. How do the citizen science observations compare with high density meteorological station observations?

In Nebraska, which had large variability in local cloud cover along the path of totality, both the GO and mesonet observations depict a sensitivity of the temperature depression to cloud cover. As with the CONUS results, overcast conditions reduce the temperature depression by about 50% relative to clear conditions. Despite some discrepancies in the details, both GO and Nebraska Mesonet data show comparable statistically significant relationships. In clear conditions, the spread in GO temperature depressions across different regions in Oklahoma is about 2 times that reported by the Oklahoma Mesonet but is still small enough to robustly detect the hypothesized cloud–temperature relationship.

c. What is the effect of public participation numbers on the robustness of the results?

A Monte Carlo test demonstrates that for the data collected for the 2017 eclipse, greater numbers of participants appreciably reduces the uncertainty of the composite-calculated temperature depressions. In particular, group sizes larger than 50 are desirable for giving results robust enough for the amplitude of the cloud–temperature relationship to be reliably detected. As interest in citizen science continues to grow in the coming years and receives increasing acceptance as a

part of “traditional” science, methodologies such as the Monte Carlo–based test will gain importance in evaluating the robustness of the data and associated conclusions.

The robustness test demonstrates the value of encouraging the public to participate in citizen science efforts. The nature of citizen science requires large numbers of participants to overcome the limitations of the observations. Fortunately, the August 2017 eclipse demonstrated that there is a sizable public interest in astronomy and meteorology that can be encouraged into action, and thus there were a sufficiently large number of participants in EAA to conduct an investigation of the eclipse-induced changes in weather. The success of EAA is encouraging for expanded citizen science campaigns for other eclipses, such as the 2019 and 2020 total solar eclipses in southern South America. Furthermore, the public interest is not limited to rare events such as eclipses but extends to general observations of the Earth system. GO collected over 140 000 observations of cloud conditions worldwide in 2017 via the GLOBE Clouds protocol, and only 20 172 of them were associated with a special event (i.e., the eclipse). In 2018, GLOBE Clouds collected more than 50 000 cloud observations in one month alone, without any associated rare astronomical event, just by soliciting participation from the public. Cloud observations are just one element of GO, which covers a wide range of Earth system properties, from meteorological conditions to mosquito populations. Citizen science is rapidly becoming a critical component of the scientific endeavor, and the full potential and benefit to science are likely yet to be discovered.

Acknowledgments. This project is based upon work supported by NASA under Award NNX16AE28A.

The authors sincerely thank all of the volunteers who participated in the GLOBE Observer effort to collect temperature and clouds data during the 2017 eclipse. We are grateful to A. Laney (alaney@mesonet.org) and G. McManus (gmcmamus@mesonet.org) for access to Oklahoma Mesonet data and to S. Cooper (scooper6@unl.edu) and M. Shulski (mshulski3@unl.edu) for access to Nebraska Mesonet data. We also thank Dr. Helen Amos and reviewers for their helpful comments. GO data are available on the Internet (<https://observer.globe.gov/eclipse-data-analysis>), as are data and images from the NASA Scientific Visualization Studio (<https://svs.gsfc.nasa.gov/4466> and <https://svs.gsfc.nasa.gov/4518>).

APPENDIX

Statistical Testing of Temperature Depression Differences from Composite Time Series

Determining the statistical significance of the difference between two temperature depressions from different composite time series requires some particular consideration. There is not a straightforward statistical test for the methodology used here. There are two complicating factors for which one must account in devising a statistical testing method. First, the temperature depression (TDEP) is not a mean but is rather a difference of means [the mean maximum temperature (TMAX) minus the mean minimum temperature (TMIN)]. So, comparing two TDEPs is not a simple “difference of means” test but rather is a “difference of difference of means,” which did not have an immediately obvious solution. Second, the results use a smoothing window, and therefore it is not immediately obvious what sample sizes and standard deviations should be used in calculating the test value.

We use the following procedure, which adapts Welch’s *t* test (Welch 1947) to account for these complicating factors:

- 1) For a given composite time series, the standard deviation of TMAX is the standard deviation of all temperature observations taken in the 20-min window before/after the time of TMAX (because the smoothing window is 20 min).
- 2) The sample size of TMAX is the number of temperature observations.
- 3) Steps 1 and 2 are repeated for TMIN.
- 4) The standard deviation of TDEP is the square root of the sum of the variances (squares of the standard deviations) for TMAX and TMIN, using the additive properties of variances.
- 5) The sample size of TDEP is the sum of sample sizes of TMAX and TMIN.

- 6) When testing the differences of TDEP between two composite series, we use the two-tailed *t* test for samples of unequal size and variance, as described by Welch (1947), to estimate the standard errors and degrees of freedom of the combined TDEP sample sizes and standard deviations. The *p* significance value is calculated using Welch’s *t* test.

Because of the sensitivity of the *p* value to sample size, increasing statistical significance of the results can be accomplished both by greater public participation and by grouping larger numbers of participants into the composites. There is an additional benefit of the smoothing process—because of the much larger number of samples collected during 20 min versus a single minute, the smoothed TDEP differences tend to have much greater significance than the unsmoothed values.

One possible concern associated with this method is the assumption that the smoothed TMAX/TMIN values are reasonable substitutions for the mean temperature values of the observations from the 20-min windows used to calculate standard deviation and sample size. They are similar but not exactly the same. Using mean temperature estimates of TMAX/TMIN instead of the smoothed TMAX/TMIN value has almost no effect on the statistical test results. Therefore, the smoothed TMAX/TMIN values are sufficient for this method.

REFERENCES

- Aceves-Bueno, E., A. S. Adeleye, M. Feraud, Y. Huang, M. Tao, Y. Yang, and S. E. Anderson, 2017: The accuracy of citizen science data: A quantitative review. *Bull. Ecol. Soc. Amer.*, **98**, 278–290, <https://doi.org/10.1002/bes2.1336>.
- Anderson, J., 1999: Meteorological changes during a solar eclipse. *Weather*, **54**, 207–215, <https://doi.org/10.1002/j.1477-8696.1999.tb06465.x>.
- Anderson, M. C., J. M. Norman, J. R. Mecikalski, J. A. Otkin, and W. P. Kustas, 2007: A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing; 2. Surface moisture climatology. *J. Geophys. Res.*, **112**, D11112, <https://doi.org/10.1029/2006JD007507>.
- Anderson, R. C., and D. R. Keefer, 1975: Observation of the temperature and pressure changes during the 30 June 1973 solar eclipse. *J. Atmos. Sci.*, **32**, 228–231, [https://doi.org/10.1175/1520-0469\(1975\)032<0228:OOTAP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1975)032<0228:OOTAP>2.0.CO;2).
- Aplin, K. L., and R. G. Harrison, 2003: Meteorological effects of the eclipse of 11 August, 1999 in cloudy and clear conditions. *Proc. Roy. Soc. London*, **459A**, 353–371, <https://doi.org/10.1098/rspa.2002.1042>.
- , C. J. Scott, and S. L. Gray, 2016: Atmospheric changes from solar eclipses. *Philos. Trans. Roy. Soc.*, **374A**, 20150217, <https://doi.org/10.1098/rsta.2015.0217>.
- Arndt, D., 2012: Oklahoma’s climate: An overview. Oklahoma Climatological Survey Doc., 8 pp., http://climate.mesonet.org/county_climate/Products/oklahoma_climate_overview.pdf.

- Barnard, L., A. M. Portas, S. L. Gray, and R. G. Harrison, 2016: The National Eclipse Weather Experiment: An assessment of citizen scientist weather observations. *Philos. Trans. Roy. Soc.*, **374A**, 20150220, <https://doi.org/10.1098/rsta.2015.0220>.
- Bojkov, R. D., 1968: The ozone variations during the solar eclipse of 20 May 1966. *Tellus*, **20**, 417–421, <https://doi.org/10.3402/tellusa.v20i3.10020>.
- Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.*, **12**, 5–19, [https://doi.org/10.1175/1520-0426\(1995\)012<0005:TOMATO>2.0.CO;2](https://doi.org/10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2).
- Chimonas, G., and C. O. Hines, 1970: Atmospheric gravity waves induced by a solar eclipse. *J. Geophys. Res.*, **75**, 875, <https://doi.org/10.1029/JA075i004p00875>.
- Clayton, H. H., 1901: The eclipse cyclone and the diurnal cyclones. *Ann. Astron. Obs. Harv. Coll.*, **43**, 5–33.
- Colón Robles, M., T. Rogerson, and J. B. Dodson, 2019: NASA GLOBE Clouds: Documentation on how satellite data is collocated to ground cloud observations (v1.0), NASA Doc., 13 pp., <https://www.globe.gov/documents/16792331/0/NASA+GLOBE+Clouds+Satellite+Matches+to+Ground+Observations/c15f4fed-9739-46aa-98e5-3b1b7aad83e8>.
- Eugster, W., C. Emmel, S. Wolf, N. Buchmann, J. P. McFadden, and C. D. Whiteman, 2017: Effects of vernal equinox solar eclipse on temperature and wind direction in Switzerland. *Atmos. Chem. Phys.*, **17**, 14 887–14 904, <https://doi.org/10.5194/acp-17-14887-2017>.
- Finarelli, M. G., 1998: GLOBE: A worldwide environmental science and education partnership. *J. Sci. Educ. Technol.*, **7**, 77–84, <https://doi.org/10.1023/A:1022588216843>.
- Founda, D., D. Melas, S. Lykoudis, I. Lissaridis, E. Gerasopoulos, G. Kouvarakis, M. Petrakis, and C. Zerefos, 2007: The effect of the total solar eclipse of 29 March 2006 on meteorological variables in Greece. *Atmos. Chem. Phys.*, **7**, 5543–5553, <https://doi.org/10.5194/acp-7-5543-2007>.
- GLOBE Observer, 2017a: The GLOBE Program. Accessed 29 March 2019, <https://observer.globe.gov/>.
- , 2017b: How cool was the eclipse? The GLOBE Program, accessed 29 March 2019, <https://observer.globe.gov/eclipse2017>.
- , 2017c: Total eclipse: Temperature and clouds. The GLOBE Program, accessed 29 March 2019, <https://observer.globe.gov/documents/19589576/20846667/GLOBE+Observer+Eclipse+Info+Sheet+v2.pdf/2e3cf210-b48b-43fa-82ac-09a0c37a989e>.
- , 2018: Eclipse data analysis. The GLOBE Program, accessed 31 October 2018, <https://observer.globe.gov/eclipse-data-analysis>.
- GLOBE Program, 2017: GLOBE cloud observations featuring NASA Satellite Comparison. NASA, accessed 29 March 2019, <https://www.globe.gov/web/s-cool>.
- Hanna, E., 2000: Meteorological effects of the solar eclipse of 11 August 1999. *Weather*, **55**, 430–446, <https://doi.org/10.1002/j.1477-8696.2000.tb06481.x>.
- , 2018: Meteorological effects of the 20 March 2015 solar eclipse over the United Kingdom. *Weather*, **73**, 71–80, <https://doi.org/10.1002/wea.2820>.
- , and Coauthors, 2016: Meteorological effects of the solar eclipse of 20 March 2015: Analysis of UK Met Office automatic weather station data and comparison with automatic weather station data from the Faroes and Iceland. *Philos. Trans. Roy. Soc.*, **374A**, 20150212, <https://doi.org/10.1098/rsta.2015.0212>.
- Heimann, I., V. B. Bright, M. W. McLeod, M. I. Mead, O. A. M. Popoola, G. B. Stewart, and R. L. Jones, 2015: Source attribution of air pollution by spatial scale separation using high spatial density networks of low cost air quality sensors. *Atmos. Environ.*, **113**, 10–19, <https://doi.org/10.1016/j.atmosenv.2015.04.057>.
- Kelling, S., D. Fink, F. A. La Sorte, A. Johnston, N. E. Bruns, and W. M. Hochachcka, 2015: Taking a ‘Big Data’ approach to data quality in a citizen science project. *Ambio*, **44** (Suppl. 4), 601–611, <https://doi.org/10.1007/s13280-015-0710-4>.
- Mahmood, R., and Coauthors, 2017: Mesonets: Mesoscale weather and climate observations for the United States. *Bull. Amer. Meteor. Soc.*, **98**, 1349–1361, <https://doi.org/10.1175/BAMS-D-15-00258.1>.
- McPherson, R., and Coauthors, 2007: Statewide monitoring of the mesoscale environment: a technical update on the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.*, **24**, 301–321, <https://doi.org/10.1175/JTECH1976.1>.
- Miller, J. D., 2018: Americans and the 2017 eclipse: A final report on public viewing of the August total solar eclipse. University of Michigan Center for Political Studies Rep., 22 pp., <https://isr.umich.edu/wp-content/uploads/2018/08/Final-Eclipse-Viewing-Report.pdf>.
- Muller, C. L., L. Chapman, S. Johnston, C. Kidd, S. Illingworth, G. Foody, A. Overeem, and R. R. Leigh, 2015: Crowdsourcing for climate and atmospheric sciences: Current status and future potential. *Int. J. Climatol.*, **35**, 3185–3203, <https://doi.org/10.1002/joc.4210>.
- Shulski, M., S. Cooper, G. Roebke, and A. Dutcher, 2018: The Nebraska Mesonet: Technical overview of an automated state weather network. *J. Atmos. Oceanic Technol.*, **35**, 2189–2200, <https://doi.org/10.1175/JTECH-D-17-0181.1>.
- Travis, D. J., A. M. Carleton, and R. G. Lauritsen, 2002: Contrails reduce daily temperature range. *Nature*, **418**, 601, <https://doi.org/10.1038/418601a>.
- Tzani, C., C. Varotsos, and L. Viras, 2008: Impacts of the solar eclipse of 29 March 2006 on the surface ozone concentration, the solar ultraviolet radiation and the meteorological parameters at Athens, Greece. *Atmos. Chem. Phys.*, **8**, 425–430, <https://doi.org/10.5194/acp-8-425-2008>.
- Warren, S. G., C. J. Hahn, J. London, R. M. Chervin, and R. L. Jenne, 1986: Global distribution of total cloud cover and cloud type amounts over land. NCAR Tech. Note NCAR/TN-273+STR, 230 pp., <https://doi.org/10.5065/D6GH9FXB>.
- Welch, B. L., 1947: The generalization of ‘Student’s’ problem when several different population variances are involved. *Biometrika*, **34**, 28–35, <https://doi.org/10.1093/biomet/34.1-2.28>.
- Wright, E., 2016: Insolation during the 2017 Eclipse. NASA Scientific Visualization Studio, accessed 2 February 2018, <https://svs.gsfc.nasa.gov/4466#79133>.