The CWOP Solar Radiation Data Archive

Lucy Hancock and Russ Chadwick

18th Symposium on Meteorological Observation and Instrumentation - Innovative Measurements, Annual Meeting of the American Meteorological Society 2016

Introduction

This presentation concerns the Solar Radiation Data Archive, which is an offshoot of the Citizen Weather Observers Program. The program is to present the Archive, some tools that we have developed to make it easier to use, and what one author is doing with it.

Part 1. The Archive

The Citizen Weather Observers Program is a public-private partnership which gathers weather data from individuals and makes it available in support of weather forecasting.

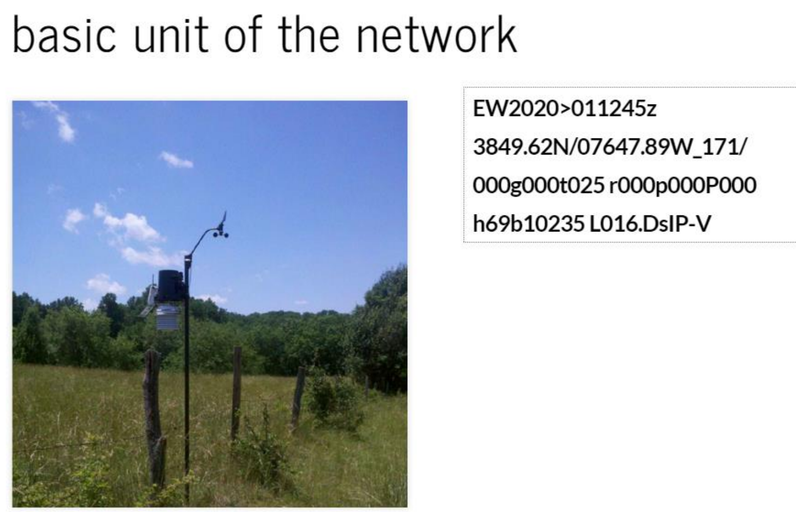


Figure 1

Figure 1 shows one station in that network and one report.

Altogether the program has registered at one time or another about 17,000 stations, of which about 9,000 report on any given day. It is a torrent of data. In real time, it is readily accessible, but getting access to the archives is a little bit more of a challenge and it used to be even more of a challenge than it is now. In 2009, Russ Chadwick decided to initiate an archive specifically of all of the observations in the data stream which recorded nonzero values for solar irradiance. His goal was to accumulate data support for the development of solar energy.

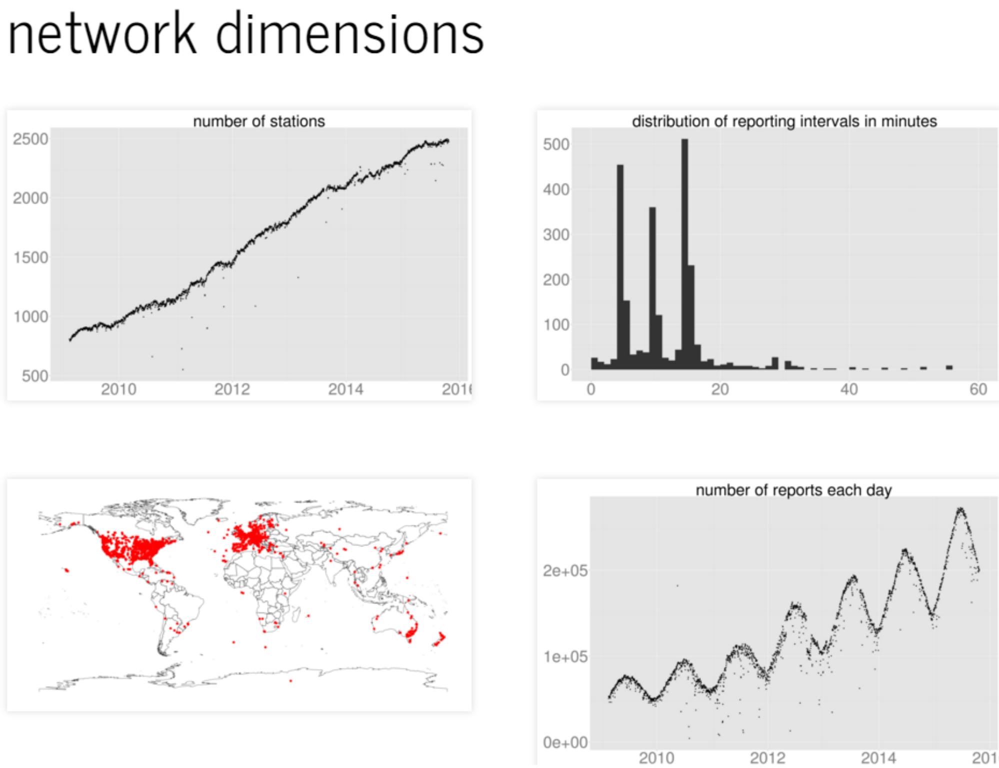


Figure 2

Figure 2 presents the development of that archive from then until now. As Figure 2 (top left) shows, when the archive was initiated it was collecting data from about 700 stations a day and as of October 2015 it was something like 2500.

The geographic distribution of the data as of October 2015 is presented in Figure 2 (bottom left). The distribution has always been something like this, even when it was sparser.

Stations tend to record once per five, ten, or fifteen minutes (see top right); some record less often and some more often.

The number of daily reports collected by the archive is presented at lower right. It has a seasonal signal which reflects Northern Hemisphere summertime.

Altogether as of October 2015 in this archive of solar radiance measurements there were more than three hundred million observations; the archive had heard at one time or another from more than 5500 stations; and it had recorded observations from four million station-days.

That data has now been parsed, and is available in an Amazon Redshift database which everyone is warmly invited to clone, from which it is easy to extract tidy data according to whatever criteria a user can dream up. (See Annex A and links at end of presentation.)

Part 2. Tools for QA/QC

The next user-support issue is that this is crowd-sourced data; there are typos. Although the respective volunteers correct them quickly, still a user would want to be able to identify them.

To simplify the task of quality assurance we have added to the archive a modeled value of solar irradiance to correspond to each measurement. Figure 3 is an example of this addition.

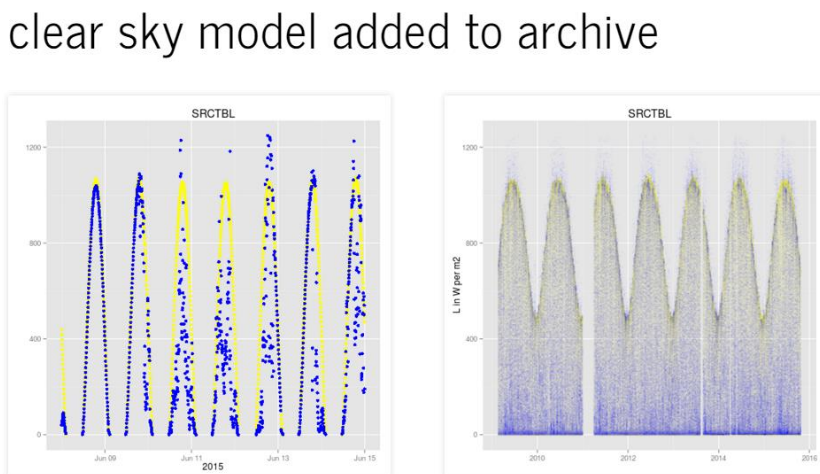


Figure 3

Figure 3 presents NOAA data (Table Mountain). The observations are in blue and the modeled values are in yellow. On the right is the archive's full record from that station.

Figure 4 presents four more stations: top left is a station from Finland; top right is a station from Australia, 17S; below left is from Bulgaria; below right is from Florida.

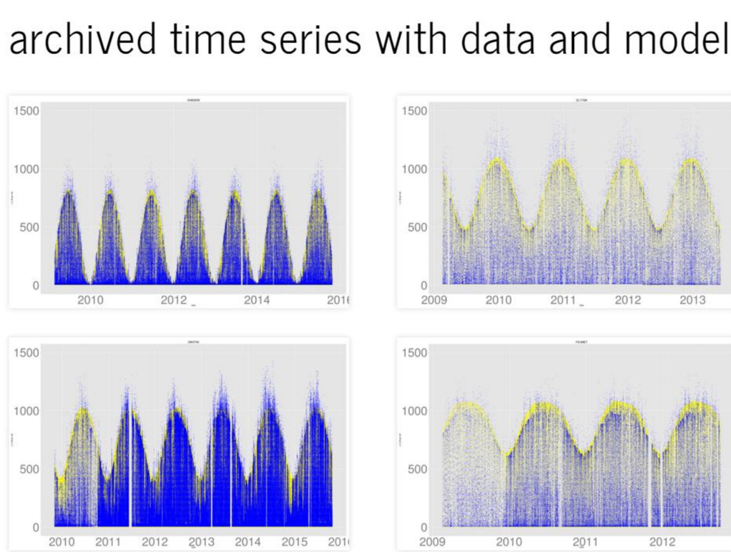
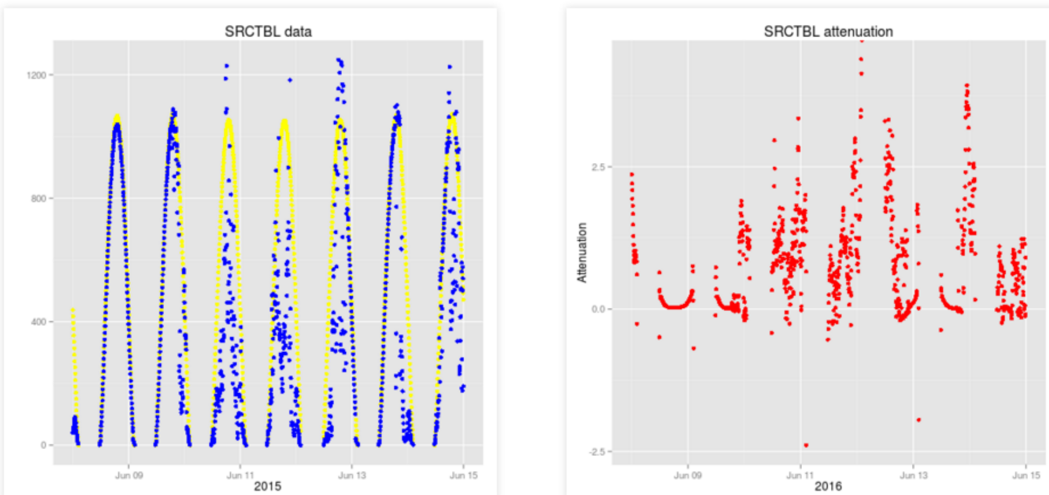


Figure 4

The purpose of Figure 4 is to emphasize to you the global grasp of this dataset and how useful it would be to you for your applications.

Still, visual comparison of data to model isn't a very convenient approach to quality assurance and quality control. Therefore, to the database has also been added at each value a comparison between observed and modeled values, using the usual definition of attenuation.

attenuation added to archive



$$\tau \equiv -\ln\left(\frac{L_{\text{obs}}}{L_{\text{modeled}}}\right)$$

Figure 5

In Figure 5, to the left is that NOAA data again (observations in blue, model in yellow); to the right is the corresponding time series of attenuation of solar radiation. As to that, we are reassured to see these “saucer” shapes on clear days which tell us that attenuation was just about zero, or near zero, at solar transit, and was much more than that at dawn and at dusk.

We are now ready to use the data for the first quality assurance exercise: looking for site shading. Usually we know nothing about the sites, and we don't know whether someone is just doing their best, putting up a station within an apartment air shaft, or what are the circumstances for each station. Here we are going to show that it is actually possible to work out site shading using the tools just illustrated, working directly from the archive.

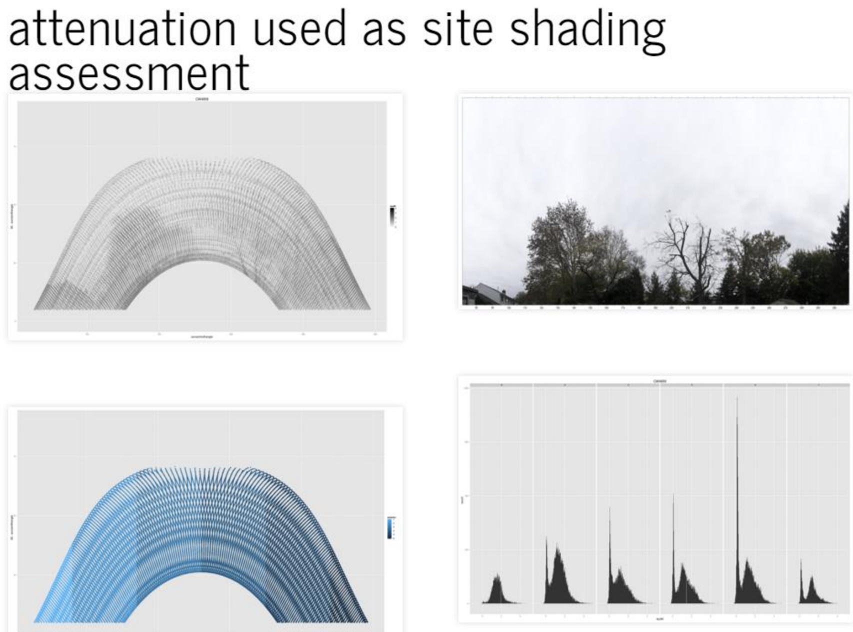


Figure 6

In Figure 6, we are analyzing data from one station: every value ever collected of attenuation. In the top right figure, all of the values of attenuation have been plotted on a spatial grid where solar azimuth angle is on the x axis and solar zenith angle is on the y axis. In effect it's just ordinary space.

That can be compared with a photo verification at top right, taken from the very same station. We see that when you plot attenuation, you can actually identify where the persistent shading is. In both figures (top right and to left) we see a big sycamore tree which is located near the station. This slanting branch, this is a walnut tree in this person's yard. One sees also a pin oak and the neighbor's house. We have been able to assess site shading; this reassures us that we will be able to use the solar radiation archive data in spite of not having that particular very important aspect of metadata.

In order to understand what these anomalous shading circumstances look like in the data, here (lower left) that sky is divided into six sectors, which are - vaguely, perhaps - visible here as shades of blue. On the lower right is a histogram of attenuation from each of these sectors. Thus farthest left is a histogram of attenuation from the sector with the neighbor's house, and so on over to the right, the histogram from the sector you see directly above it. Each of these histograms of attenuation is some realization or another of what appears to be fundamentally a two-humped distribution, where the clear sky is over on the left and the site shading is systematically over on the right.

four skies from the archive

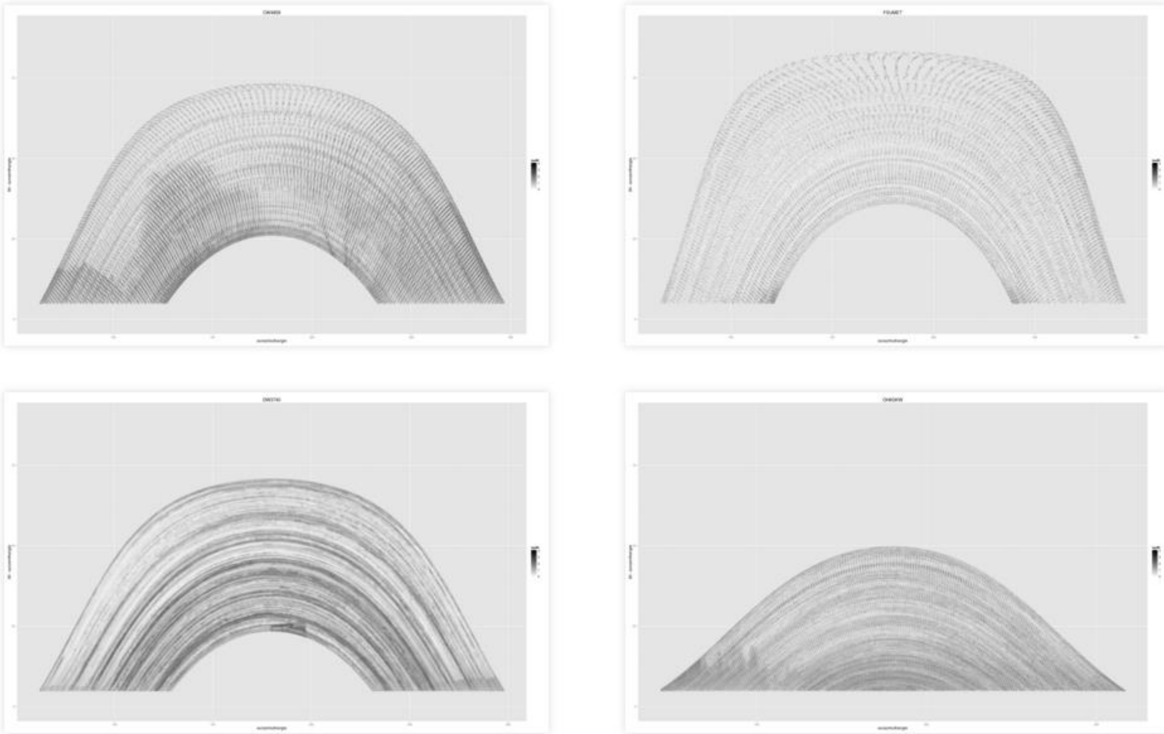


Figure 7

Figure 7 is the site shading analysis at four sites – the one we just did and three more places in order to make sure we know what shading looks like in other places and that the site shading assessment didn't happen to work well in Figure 6 by chance.

- Figure 7 top left is the site shading assessment from the site shown in Figure 6.
- Figure 7 top right is a site which is called FSUMET. I don't know who they are but it sounds like FSU Met. It appears that this station has a very nice clear sky with the possible exception of what might be some bushes which are causing shade at dawn and at dusk at midwinter.
- At lower left in Figure 7 is a station which is located in Bulgaria, and from its latitude and its longitude it's located in a city center. It looks like a very good sky. He's caught the corner, apparently, of some building in the city center; everything else looks really good.
- At lower right in Figure 7 is a station in Finland. It's a very nice clear sky, except that over in the east, this is a stand of fir trees; it is beautiful if you look at it closely.

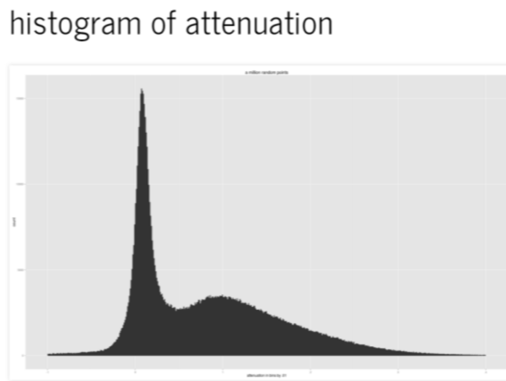


Figure 8

Figure 8 is based on a draw of a random million observations from our immense archive. It presents a histogram of the attenuation values from those observations. The goal here is to measure the range of attenuation values. Values fall mainly between minus two and five. Under the assumptions under which model irradiance was calculated (see Annex B), this is a characteristic histogram.

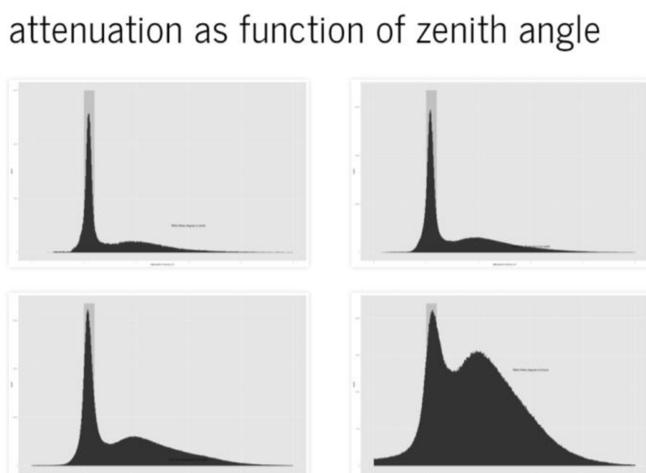


Figure 9

Figure 9 presents data from a random draw of ten million points, histograms of attenuation sorted by zenith angle of sun at the time of the observation. Top left is attenuation only for observations which were made within 22.5 degrees of zenith; top right is observations from 22.5 to 45 degrees from zenith; lower left is observations from 45 to 67.5 degrees from zenith; lower right is observations within 22.5 degrees of the horizon

The value of Figure 9 is once again so that we can understand what physical values should look like. The mode of the clear sky distribution does not change with zenith angle. That's surprising.

The parsed data has a number of limitations and I have presented a few of them below (Figure 10 together with (blue) an indication of the context in which each may contribute systematic error. Do please get in contact with the authors to understand more about these if there is anything puzzling here.

known limitations of the parsed data

- - Metadata sometimes has typos. **Errors most affect tropics.**
- Two different sampling strategies are in use. **Effect on stats.**
- Archive intervals are not reported. **Would affect glinting stats.**
- Sensors may tilt. **Overbright nearer meridian - noon and winter?**
- Protocol for reporting L>999 is used inconsistently. **Midsummer.**
- L values above 2000 are outside the range of most stations. **Midsummer glinting.**

Figure 10

Part 3. Application of the Archive: the question of an Earth ring

The next and last topic presents one application of the database, namely to answer the question of whether or not the Earth has a planetary ring of its own.

(A visualization of a notional ringed Earth from the point of view of the Sun, over a one year cycle, is at https://www.youtube.com/watch?v=TQisD91tGo4.)

project: can there be a planetary ring

WHAT TO LOOK FOR:
equatorial orientation is usual
therefore annual cycle
usually densest around 1.5-3 R(planet)
fuzzy structure for a rocky planet
composed of typical aeolian sediments

REASONS TO NOT DO THIS:
a ring would decay very rapidly
only the moon could make it persistent
but moon was long thought dead

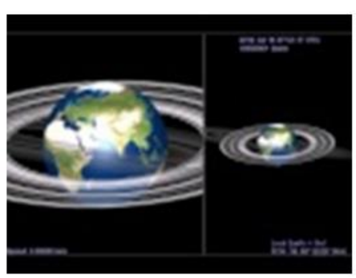


Figure 11

If Earth did have a ring, it would be meteorologically incredibly consequential. As for why everybody doesn't already study this important question, there are two big areas of discouragement and concern:

(a) "Why hasn't it been seen, and especially by spacecraft?"
(b) "It is certain to decay rapidly. And then what would replenish it? Is the Moon geologically active enough to replenish it? Isn't the Moon dead?"

We can't spend too long on this; that would really be a different presentation. But a few thoughts:

There are four known planetary rings in the Solar System: all at the gas giants in the outer Solar System and with rings that spacecraft can see. But an Earth ring could be ***more*** difficult to observe than these distant entities for several reasons.

- First, even if there were a ring in the Inner Solar System, it would tend to be blown away by the solar wind with greater pressure than in the Outer Solar System; thus an Earth ring would be sparse especially when solar wind is strong even if there is a tendency for a ring to collect.
- Secondly, and perhaps more importantly, there is this subtle issue that Earth is a rocky planet, and as such it is more nearly round than the gas giants. It is the oblateness of a planet, specifically, that encourages flattening of a cloud of particles in orbit into a disk presenting a high density contrast against the background. So because our planet is rocky and therefore more-nearly round we should probably expect a fuzzy ring and not the beautiful sharp disk that the Saturn observers consider.
- And then the third issue is that, unique in the Solar System, we have this huge and important Moon and it is orbiting us not in our equatorial plane but thirty degrees away. So once we're envisioning some kind of mass exchange (adequate to replenish a ring), we're actually talking about two rings, and they would be in different planes, and what it comes down to is, you can't really expect, at a random draw in space, to see an Earth ring in an edge-on aspect at good contrast against a dark sky. It would be actually a very rare thing, from any point of view, to be able to do that.

As for whether the Moon could replenish it, that's extremely important. But the consensus that the Moon is dead has weakened lately.

We're not going to spend any more time on that here - our actual topic here today is what solar radiation data has to say about what structure, if any, is consistent with solar irradiance data.

Our approach is going to be this:

Any wintertime observation of solar radiation is a test of the ring hypothesis. Consider this observer here in Figure 12, measuring sunlight at solar transit (approximately noon). This green person is making the observation on some day of winter that is given (up to ambiguity between a day in SOND, or a day in DJFM) by the declination of the Sun. This green person is an observer at some latitude, given here by the zenith angle ZA at solar transit. In this case the extent to which solar radiation is dimmed for this observer is a test of the optical depth of a ring at radius R shown here as one side of a triangle where we have two angles and a side. Thus this radius R, the ring radius which would attenuate solar radiation for this observer here, is something which is easy to calculate (and the derivation is Annex C). It has been calculated; it has been added the database. Thus for every wintertime observation in the database, there is associated a value R for what is the ring radius that was being tested by that particular value of attenuation of solar radiation. (Note that the derivation assumes transit, but the expression is actually more widely applicable than that. This is why it is included at every point, though you have seen only why it works at transit.)

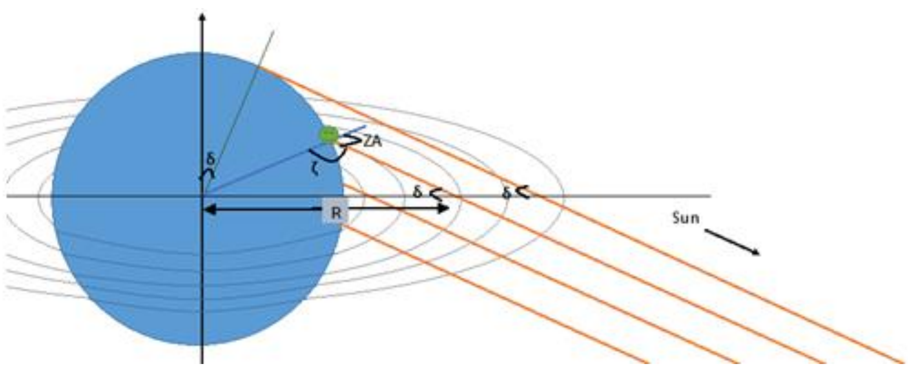


Figure 12

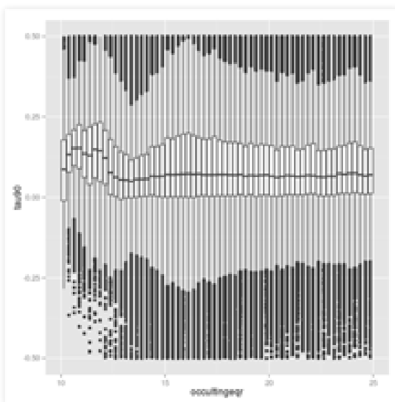
What we're going to do is simply organize all the values of wintertime attenuation in the database by their corresponding R, to see whether we find some kind of a consistent structure, some association of attenuation and R which lends itself to a consistent physical interpretation of a shady ring which is out there.

The analysis is limited to values of attenuation less than 0.5, so we're sticking with that very clear sky distribution over on the right as in Figures 6, 8 and 9, and shedding by this means most attenuation caused by buildings, trees and the leaves and so on, which were mostly though not entirely comprised among values of attenuation greater than 0.5.

The analysis it takes in only the Northern Hemisphere, where most data is; thus it better averages out factors we are not interested in such as weather.

And the analysis considers only observations made with the Sun within five degrees of meridian transit because the expression for R has been derived here (Figure 12 and Annex C) for an observation is made at solar transit.

attenuation as a function of attributed ring



- densest at about 16K km from center of Earth, about 2.5R(Earth),
- $\tau_{ring_{edge-on}} \approx 0.1$ (taken as the anomaly above average value of τ),
- $\tau_{ring} \approx 0.02$ if the above is scaled by sin of ten degrees (i.e., half of 23.5).

Figure 13

Figure 13 shows that when you line up all these attenuation values by preparing a boxplot to show the distribution of shadiness for each incremental value of radius 10K to 20K km from the center of the Earth, what you find is that yes, the data does seem to line itself up and have a feature. Figure 13 is consistent with the idea that something shadier happens on days where the Sun would be shaded by a ring of radius of about 16 thousand km from the center of the Earth. The anomaly is not always presen: the median value of attenuation does not change much from about 13K to 20K. But around 16K km, a larger number of high values of attenuation occur.

As far as it goes what we have found is that the solar radiation data are consistent with the interpretation that Earth has a climate-important ring system in its equatorial plane. Clearly, more work needs to be done.

Closing

The one thing one may very firmly say in closing is that if we do have a ring system, then even if it is sparse, even if it is fuzzy, even if it is erratic, even if it is discontinuous and even if there are two of them, this database will be able to find it; it can give us the answer, “Yes” or “No.”

acknowledgments

CWOP Volunteers

Steve Dimse, Philip Gladstone,Ted Lum, Davis Instruments, Weather Display

Postgres community, Aginity, R Community, RStudio Community,
Javier Corripio (insol), Hadley Wickham (ggplot, dplyr, stringr, inter alia), Coursera and JHU, Celestia Community, Amazon Web Services, Lewis Aslett for AMI, Stack Overflow community

links

- CWOP generally: http://wxqa.com
- CWOP solar radiation archive: http://wxqa.com/lum_search.htm
- Parsing routine: www.github.com/lohancock/solardataparser
- Markdown for this presentation: https://github.com/lohancock/solar-data-parser/blob/master/ams2016.Rpres
- To clone the database of parsed solar radiation data, contact either author (russ4cwop@gmail.com,lohancock@aol.com)

Annex A

annex: variables in the archive

each observation comprises 34 variables including: **parsed data: 23 variables** original report, station name, archive date, date/time as extracted from report, dateflag to highlight nominal dates out of range, calculated date and time z, latitude, N/S, longitude, E/W, wind direction, wind speed, wind gust, temperature (F), rain this hour, rain last 24 hours, rain today, relative humidity, barometric pressure, solar radiation measurement as given in report, solar radiation as interpreted, flag for number of characters in report, tech suffix providing some description of hardware and software **from model: 8 variables** julian day,sun azimuth angle, sun zenith angle, modeled solar insolation, diffuse component of modeled insolation, day length, solar declination, equation of time **from lookup:** station height **computed** attenuation **computed** occultingeqr

Annex B – Details of Clear Sky Model

Used insol routine by Javier Corripio
https://cran.r-project.org/web/packages/insol/insol.pdf

Used latitude, longitude, z and relative humidity supplied in observations
Used temp(F) supplied in reports, converted to tempK
Used height looked up at gpsvisualizer.com
Applied to all calculations: visibility = 90 km, albedo 0.5, ozone=.02

Annex C – Geometry of occulting ring in the equatorial plane

Q: On a winter day, solar declination = δ, the sun for some observer has zenith angle ZA at transit (ZA is the angular distance from the zenith to the sun's position).

What is the radius R of the orbit of material in an equatorial ring that shades the sun at transit for this observer?

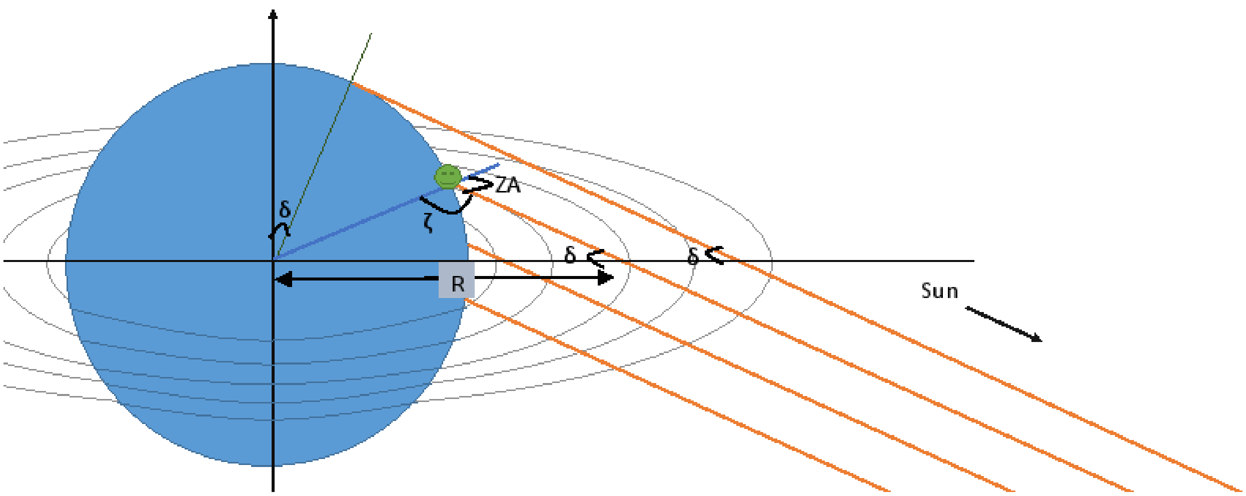
A: In the figure below, the blue line extends from the center of the earth to the observer in question. Orange lines depict incoming solar radiation. The angle ζ is the supplement of ZA; that is, ζ=180-ZA. Observer position is denoted thus:

By the Law of Sines,

$$\frac{\sin\delta}{r_{earth}} = \frac{\sin\zeta}{R}$$

Furthermore, sin ζ = sin ZA

$$\boxed{R = \sin ZA * \frac{r_{earth}}{\sin\delta}.}$$



$$R_{\varnothing} = sin(ZA) * \frac{R_{\oplus}}{sin(\delta)}$$

where

$R_{\varnothing}$ ≡ radius of ring;
$R_{\oplus}$ ≡ radius of Earth;
ZA ≡ zenith angle of sun;
δ ≡ solar declination angle.