The Quiet Day Curve

Dave Typinski, October, 2017

This article explains what a quiet day curve (QDC) is and how to make one.

The Problem

When analyzing radio telescope data, the difference between the amplitudes of the signal of interest and the prevailing radio background noise must be quantified. This is easier at frequencies above 100 MHz, for the background noise level is relatively constant and well defined, most of it coming from the radio telescope itself.

At frequencies below 30 MHz, however, the background noise is dominated by the galactic background (GB) radiation (18 kK at 30 MHz) which is over 10 dB hotter than a modern telescope receiver's self-generated noise.[1]

It is, therefore, desirable to have a reference GB curve against which observational data can be compared.

There are several factors associated with making such a reference curve:

◊ the GB changes on a diurnal cycle due to waxing and waning of the ionosphere,

◊ the GB changes on an almost-diurnal cycle due to the galactic plane passing through the antenna beam,

◊ the timing of the presence of the galactic plane within the antenna beam relative to UTC changes on a yearly cycle due to Earth's orbit around the Sun, and

◊ the GB is polluted with band noise and RFI during the daytime.

We can use the fact that the GB varies uniformly in sidereal time to overcome the difficulty of looking at the GB using the UTC time scale. Use of histograms helps overcome the problems associated with ionospheric attenuation, daytime band noise, and sporadic radio frequency interference (RFI).

The Solution

By analyzing many days of data taken under quiet observing conditions and looking at the data on a sidereal time scale, one can generate a usefully smooth curve for the GB. This curve is called a quiet day curve or QDC because it removes as much as possible the effects of solar activity and terrestrial noise from ground-based observations of the GB.

Originally, the method was to make a continuous piece of strip chart paper that looped through a chart recorder once every sidereal day, the pen marking multiple curves one atop the other on the same piece of chart paper.[2] This method requires quiet observing conditions lest RFI contaminate the GB.

We can achieve the same result by making periodic histograms of digital strip chart data in sidereal time for many days, then summing the histograms of each daily identical sidereal time period. The use of histograms provides some degree of RFI rejection.[3] Sporadic RFI will not interfere with this type of analysis; but, relatively constant RFI still presents a problem. The histogram technique is not a panacea for RFI.

How it Works

Let us use some single frequency (strip chart) data as an example. Radio Sky Pipe (RSP) measures and records amplitudes, usually in terms of antenna temperature, once every 100 milliseconds. Most observers have RSP con-
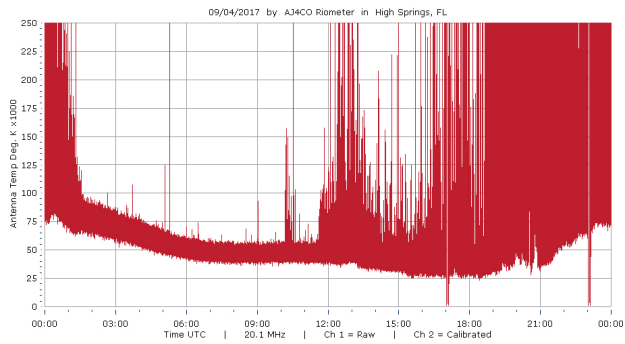
Figure 1 – A typical 24-hour daily dataset from the AJ4CO riometer, this one on 04 Sep 2017. The two negative-going spikes are the twice-daily step calibrations that run 3 hours before and after each Jupiter transit (2000 UTC on this day). Sunrise occurred at 1109 UTC and the galactic plane transited at ~0030 UTC. The normal increase in daytime band noise is plainly visible. The channel 1 data is raw, not-yet-calibrated data and is not shown for clarity.

Data Point #	Time (UTC)	T_{ant} (kK)
1	07:59:59.903	42.5
2	08:00:00.184	49.8
3	08:00:00.200	43.6
4	08:00:00.309	46.8
5	08:00:00.387	45.0
6	08:00:00.497	51.8
7	08:00:00.606	44.4
8	08:00:00.716	43.9
9	08:00:00.794	47.6
10	08:00:00.903	56.0
11	08:00:01.012	48.2
12	08:00:01.091	44.3
13	08:00:01.200	49.4
14	08:00:01.309	54.4
15	08:00:01.387	44.2
16	08:00:01.497	47.8
17	08:00:01.606	48.6
18	08:00:01.684	45.7
19	08:00:01.794	45.9
20	08:00:01.903	46.1

Table 1 – Data point time stamps and antenna temperature values from Fig 2B. Note the missing data point near 08:00:00.000 UTC and the varying period between each sample. This is a good example of how real data is often imperfect.
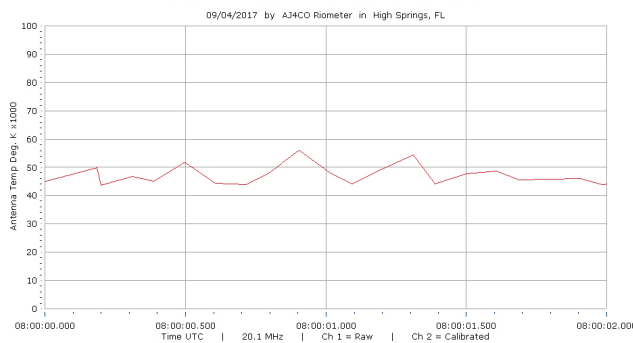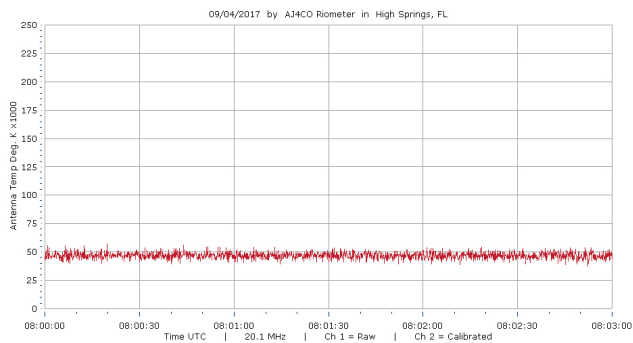

Figure 2A (upper) – A 3 minute section (0800 to 0803 UTC) of Figure 1.
Figure 2B (lower) – A 2 seconds section (20 data points from 08:00:00 to 08:00:02 UTC) from Fig 1. The vertical scale in Fig 2B is altered compared to Figs 1 and 2A for clarity.
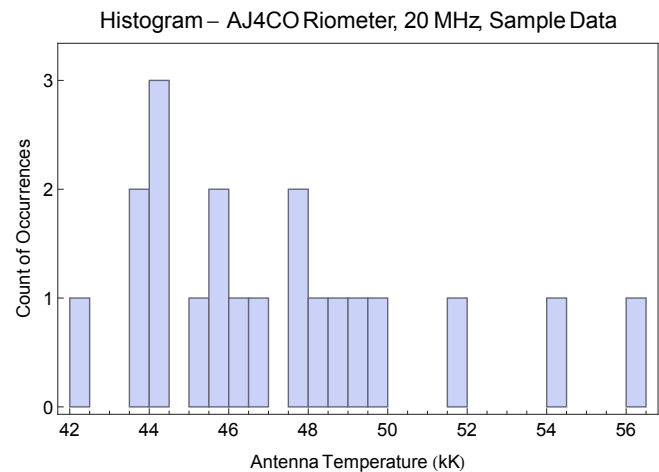

Figure 3 – A histogram made from the data shown in Fig 2B and Table 1. The histogram shows the number of occurrences within each 500 K antenna temperature bin.
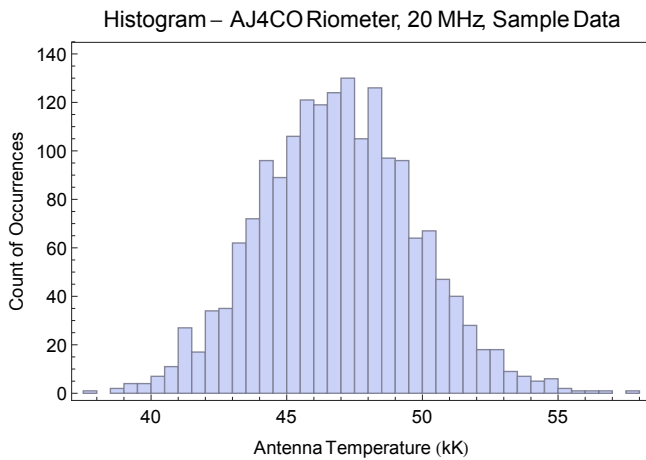
Figure 4 – A histogram made from the 3 minutes of data (1,800 data points from 0800 to 0803 UTC) shown in Fig 2A. The bell-shaped curve of the histogram occurs because bandwidth-limited noise (what the GB is to a tuned receiver) has a Gaussian distribution.

figured to record 24 hours of data in each file. Each data file (the .spd file created by RSP) record contains one or more data points and an associated time stamp.

A 24-hour data set is plotted in Fig 1 and zoomed-in views are shown in Figs 2A and 2B.

Table 1 shows two seconds, or 20 samples, from the data. These 20 data points are used to make a histogram as shown in Fig 3.

To make a histogram, one must decide how to "bin" the data. In this case, each bin is 500 K wide in temperature. We use the convention of labeling the bins with their lower bound.

For example, a T_{ant} of 49.4 kK would fall into the 49.0 kK bin and a T_{ant} of 49.6 kK would lie in the 49.5 kK bin.

One must also decide how many data points to put into a single histogram. Since the data represents a time series, the number of data points per histogram also determines the amount of time represented by a histogram.
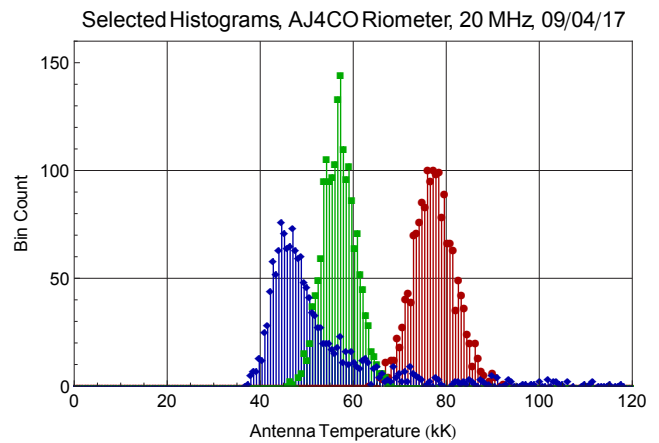

Figure 5 – Three histograms from the 04 Sep 2017 riometer data file. Red (right) is 0400 to 0403 UTC, green (center) is 1000 to 1003 UTC, blue (left) is 1900 to 1903 UTC. Each histogram represents three minutes of data, or 1,800 data points.
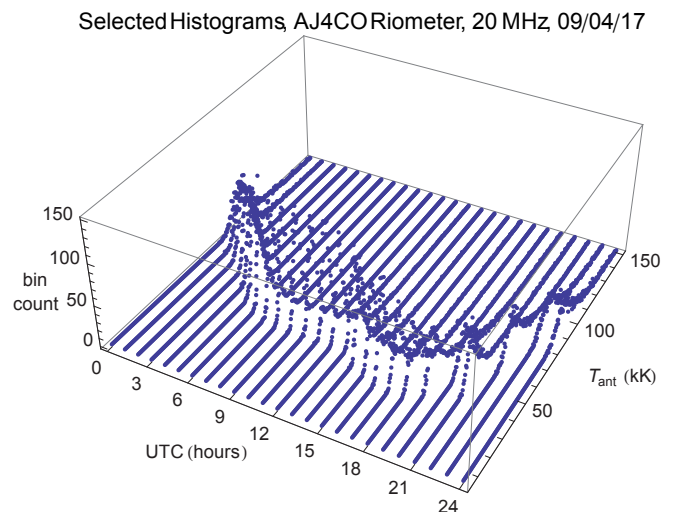

Figure 6 – 24 histograms from the 04 Sep 2017 riometer data file, each three minutes of data beginning at the times noted on the x axis.

In Fig 3, there are 20 data points, or two seconds of data. In Fig 4, there are three minutes, or 1,800 data points, in the histogram. It is a balancing act between too few data points per histogram, which leaves the result noisier than desired, and too many data points per histogram, which may smooth away the signal of interest. For analysis of the
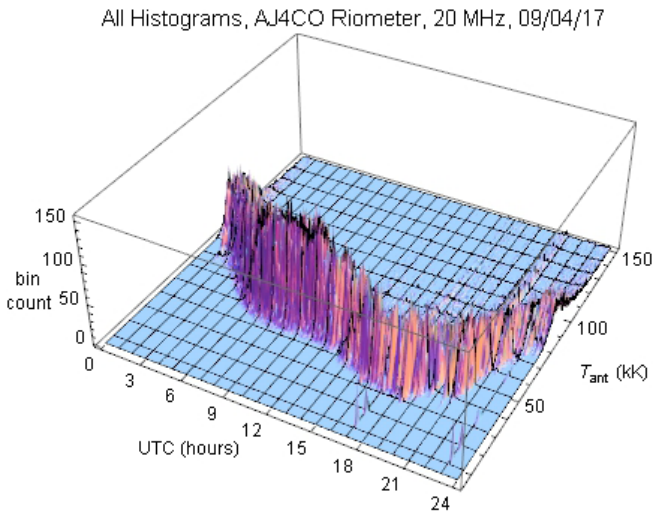
Figure 7 – All 480 histograms from the 04 Sep 2017 riometer data. Each histogram covers three minutes of data. Same representation is Fig 6, but with all the interstitial histograms shown.
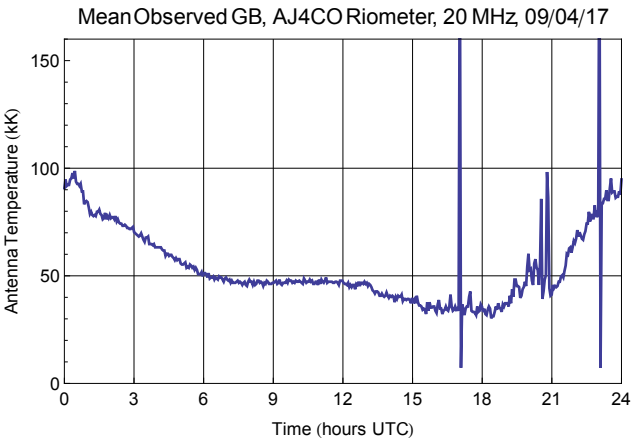

Figure 8 – This curve represents the peaks of all 480 histograms shown in Fig 7. These peaks are the values that were recorded most often in each 3-minute segment of riometer data from 04 Sep 2017. The spikes at 1700 and 2300 are from calibrations. The noise between 1900 and 2100 is daytime band noise and RFI.

GB, we chose to divide the data into 3-minute chunks. This means that in a 24-hour period, there are 480 histograms.

Fig 5 shows three such histograms, each representing 3 minutes of data, stacked together in the same plot. Fig 6 shows 24 histograms from the same data file with 1 hour separation between the histograms. Figure 7 again shows histograms from the same data file, this time with all 480 histograms stacked together side-by-side.

Looking at Fig 7, we trace the path along the peaks from one histogram to the next, we arrive at the curve shown in Fig 8. This represents the antenna temperature most often recorded during each 3-minute block of data.

We could make such a curve once a day. However, we would find that the GB peak occurs roughly 4 minutes earlier each day. If we tried to combine such daily curves in universal time (UTC), the result would be useless.

The solution is to plot each daily curve not against UTC, but against sidereal time. One

method for making a set of histograms in local mean sidereal time (LMST) from data recorded in universal time is shown in Fig 9.

Note that we use LMST instead of local apparent sidereal time (LAST). LAST includes a small correction for the equation of the equinoxes which amounts to no more than 1.1 seconds.[4] Since we are binning our data in chunks of 3 minutes, this small adjustment is not required and LMST is good enough.

RSP records timestamps in its SPD files in the Microsoft Visual Basic (VB) Date data type, an 8-byte number. To find LMST, we must first convert the VB date value to Julian date (JD), then convert from JD to LMST.

The VB Date data type is a 64-bit (i.e., double precision) floating point number representing the number of days and fraction thereof since 30 Dec 1899 at 0000 UTC. To convert this number to JD, we use the following equation:

$$JD = VBDate + 2415018.5$$

To convert JD to LMST, we use the following

algorithm:

LMST = GMST + Lon

where

GMST = 280.46061837
    + 360.98564736629(JD-2451545)
    + 0.000387933T^2
    - T^3/38710000

and

T = (JD-2451545)/36525

and

Lon = observer's longitude.

First, Greenwich mean sidereal time (GMST is calculated, then converted to LMST by addition of the offset for the observer's longitude. In the algorithm above, GMST, LMST, and Lon are in terms of degrees and western longitudes are negative.[5]

The conversion from UTC to LMST can be made as the data is recorded. RSP version 2.0.0 and higher can write timestamps in LMST when taking data. The conversion can also be made after data is written with UTC time stamps. RSP version 2.7.20 and higher can open a UTC time-stamped SPD file, display it in LMST, and save a new file with timestamps converted to LMST. Or, the equations above can be used to manually convert UTC time stamps when binning the data into histograms.

The data from 04 Sep 2017 was processed via the method shown in Fig 9 and plotted against sidereal time as shown in Fig 10.

Since the curve in Fig 10 is produced from the peaks of histograms, not directly from the amplitude data itself, small excursions in amplitude are washed away when many days are brought together in the same histogram.

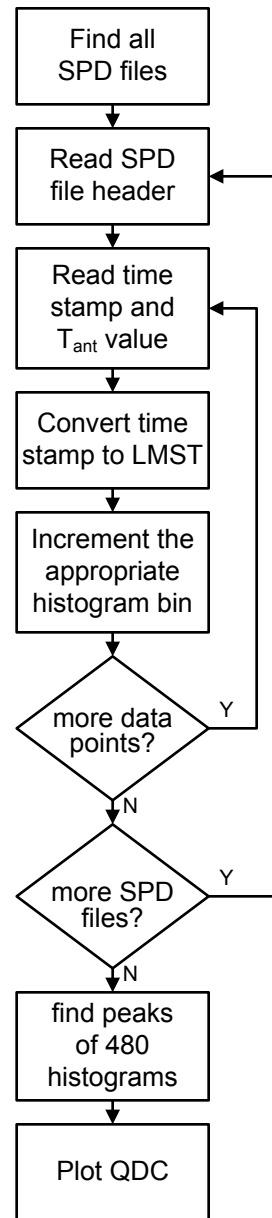Figures 11, 12, and 13 show plots from 70 days of riometer data.

Figure 9 – Data processing flow chart to create a QDC from UTC-time-stamped SPD files.

We have used the peaks of each histogram to represent the prevailing GB. For the relatively small amount of data, two months, this is useful. The result is the quiet day curve, or QDC, for these two months.

For data sets spanning many months to years, there is another method that is useful. Instead of picking the peak of each histogram, we
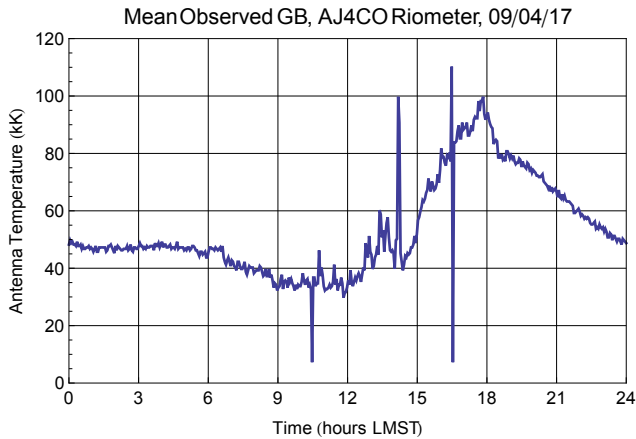
Figure 10 – Riometer data from 04 Sep 2017, histogram peaks plotted in sidereal time. Same as Fig 8, but plotted against LMST instead of UTC.
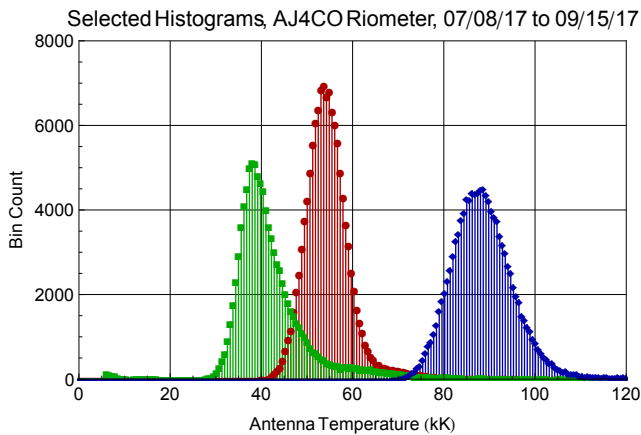

Figure 11 – Three histograms from 70 days of riometer data (note that the bin counts are more or less a factor of 70 greater than the bin counts shown in Figs 4 through 7). Red (center) is 0400 to 0403 LMST, green (left) is 1000 to 1003 LMST, blue (right) is 1900 to 1903 LMST. Each histogram represents three minutes of data, or 1,795 data points, or 3 minutes of sidereal time.
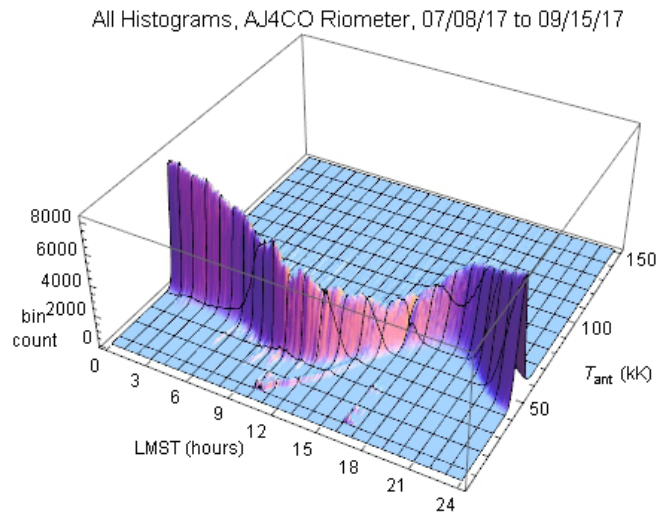

Figure 12 – All 480 histograms from 70 days of riometer data. Each histogram covers three minutes of data. Same representation as Fig 7, but covering 70 days and plotted against sidereal time instead of universal time. The deviation below the norm from 1000 to 1500 LMST is from a magnitude X8.2 (among the strongest ever recorded) X-ray flare on 10 Jul 2017 that produced a very deep radio blackout.
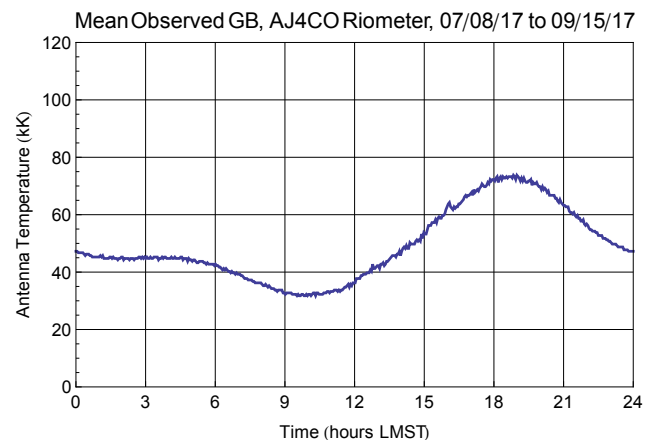

Figure 13 – This curve represents the peaks of all 480 histograms shown in Fig 10. These peaks are the values that were recorded most often in each 3-minute segment of 70 days of riometer data. Same representation as in Fig 8, but with many more days of data and plotted against sidereal time instead of universal time.

may choose the inflection point on the upper side of the bell curve above the peak.[6] For small data sets, we feel that the peak of the histogram is more useful as the upper side is often contaminated with RFI during the daytime hours to the point where the upper inflection point is indeterminate. However, for large data sets, the changes in ionospheric absorption will smear the data lower in temperature. For example, we see the GB peak at 1800 LMST is about 75 kK for the months of July and August plus the first half of September. During these months, the inner galactic plane was transiting between 0100 and 0500 UTC, which at AJ4CO Observatory in the Summer is at or right after sunset.

This means the ionosphere was more opaque than it will be when the inner galactic plane transits several hour later into the very early morning. That won't happen until the Spring of 2018. At that time, we expect that the observed GB peak at 1900 LMST will be somewhat higher than 75 kK.

Therefore, to account for this broadening of the histograms' bell-shaped curves due to the yearly variation in ionospheric attenuation at each fixed time in LMST, we plan to pick points on the upper side of the curves once we have a full year of riometer data available for analysis. This, then, will create an overall QDC comprising all UTC-to-LMST offsets.

However, one may also create a set of QDCs, one QDC for each month of data. In some instances, a set of monthly QDCs may be a more useful reference than a single one-size-fits-all QDC. Future work will compare and evaluate monthly versus yearly QDCs.

Comparison with Other Methods

Using histogram analysis has the advantage of being tolerant of sporadic RFI where other analytic methods fail. Fig 14 compares the results of analyzing one month of riometer data
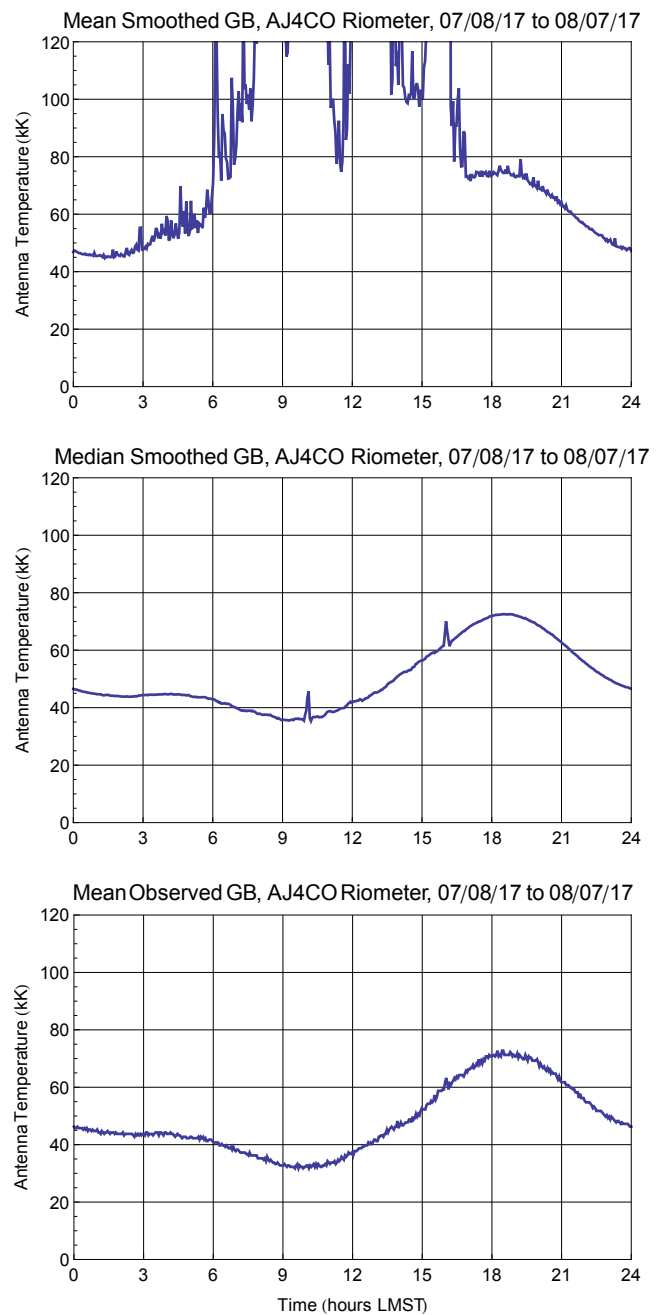

Figure 14 – A (top), 08 Jul 2017 to 07 Aug 2017 riometer data analyzed using mean smoothing (averaging) across 1,800-sample blocks of data (3 minutes). B (center), same data analyzed using median smoothing across 1,800-sample blocks of data. C (bottom), same data analyzed using histograms of 1,800 samples each. In all three cases, the data was accumulated in 3-minute LMST blocks, combined (stacked), then each stacked-block analyzed for mean, median, and histogram peak to create the plots shown here.

using mean smoothing (averaging), median smoothing, and histogram analysis.

Performing mean or median smoothing across multiple data files faces one with the difficulty of how to best combine daily data files that have slightly different numbers of samples for an identical time span. That is, one cannot simple stack multiple daily data files point for point, since each file will have a different number of data points.

For the mean and median smoothing analyses, we are folding the data (sometimes known as chart stacking) in sidereal time with data time-stamped in terms of UTC. We thus used an approach similar to that described for the histogram analysis. We converted each data point's time stamp to LMST to determine into which 3-minute LMST segment to accumulate that data point.

The difference in procedure with the mean and median smoothing analyses is that the appropriate LMST bin is not simply a counter that is incremented (as in the histogram method), but instead becomes a list of all antenna temperature values observed within the associated 3 minutes of LMST during, in this case, one month of observation. We binned all data points of all 31 SPD files in this manner, then calculated the mean and median for each 3-minute segment to produce the plots shown in Figs 14A and B.

Mean smoothing is not tolerant of sporadic RFI as shown in Fig 14A. As expected, the results of median smoothing shown in Fig 14B produce a much smoother curve than mean smoothing.[7] However, median smoothing is not tolerant of the daily step calibrations that could be considered RFI in this situation.

The curves produced using mean and median smoothing indicate antenna temperatures slightly higher than the histogram method; for example, 1200 LMST in Figs 14B and C. We attribute this difference to the mean smoothing method having a lower tolerance of RFI contamination. While the median smoothing method is far more tolerant of RFI than mean smoothing is, the histogram method is the most accurate of these three methods under sporadic RFI conditions.

It will be interesting to see how median smoothing compares to the histogram method after a complete year of riometer data is available.

Summary

Making a QDC by using histogram analysis lets one arrive at an accurate value for the galactic background. This is useful when analyzing the temperature of other signals. Histogram analysis produces a more accurate GB curve than mean smoothing and median smoothing when sporadic RFI is present.

References

[1] Typinski, *The Galactic Background in the Upper HF Band*, SARAJ (2013).

[2] Little and Leinbach, *The Riometer—A Device for the Continuous Measurement of Ionospheric Absorption*, Proc IRE (1959).

[3] Armstrong, et al., *The Day to Night Absorption Ratio in Auroral Zone Riometer Measurements*, Planet Space Sci (1977).

[4] *Approximate Sidereal Time*, US Naval Observatory Astronomical Applications Department (2017).

[5] Meeus, *Astronomical Algorithms*, Willmann-Bell (1991).

[6] Krishnaswamy, et al., *The Inflection Point Method of Determining Riometer Quiet Day Curves*, Radio Science (1985).

[7] Typinski, *Implementation of Median Smoothing in Radio-SkyPipe II*, SARAJ (2012).