# IMPROVEMENTS FOR THE AAVSO SUNSPOT NUMBER

**Bradley E. Schaefer**
Physics Department, JWG 463
Yale University
New Haven, CT 06520-8121

*Presented at the 85th AAVSO Annual Meeting, November 2, 1996*

## Abstract

The AAVSO sunspot number has served as a measure of solar activity on a time scale of decades. Nevertheless, this paper presents four ways in which the current AAVSO handling of sunspot data could be improved: (1) publishing error bars; (2) requiring new sunspot observers to use small apertures and high magnification; (3) using a logarithmic, rather than linear, average of the scaled sunspot counts; and (4) publishing all raw data and all procedures.

## 1. Introduction

Along with the International sunspot number, the AAVSO sunspot number is the primary measure of the spot activity on our Sun. With the high importance of sunspots for many modern science questions, both numbers must be scrutinized for improvements and for reliability. Such questioning is a normal part of science, yet until now, the sunspot numbers have never been critically examined by outsiders. I have performed an analysis of the propagation of uncertainties in both numbers. My results for the International (and Zurich) numbers will appear elsewhere. For the AAVSO number, I found a subtle mathematical flaw which inevitably leads to an artificial increase in the published numbers (Schaefer 1997). I identified several other features of the AAVSO number that can be substantially improved. Altogether, there are four easy improvements that the AAVSO should make.

## 2. Report error bars

The error bar is at the heart of the scientific method. Any observation is useless without a knowledge of the measurement uncertainty. For example, an isolated report that the Hubble Constant is 80 km $s^{-1}$ $Mpc^{-1}$ is of no utility, since a one-sigma error bar of 40 km $s^{-1}$ $Mpc^{-1}$ implies that the measure is useless, while a one-sigma error bar of 4 km $s^{-1}$ $Mpc^{-1}$ implies a decisive new result. Similarly for sunspots, an isolated number without error bar is of poor utility. Thus, a number of 80 might be $\pm$40 (for an individual observer including scaling uncertainties), or $\pm$4 (for 100 scaled observers), with radical differences in subsequent interpretation. An extreme but valid position is to say that any measure without a known uncertainty is not science.

The uncertainties in the various sunspot numbers change drastically with time. For example, the average number of daily observations changes from day to day and decade to decade, so that the random errors will be beaten down to varying extents. Another example is that the uncertainty near the peak of cycles is much larger than near the minimum, but it does not vary as a simple proportion, since quantization errors dominate for low numbers. These uncertainties are vital to any correct studies of the sunspot time series. That is, without uncertainties, analysts can only equally weight the input, so that their results will be unfairly driven by the poorest points. Without quoted error bars, a sunspot number should not be analyzed.

The error bars are easy to calculate. For the daily numbers, it will be simply the standard deviation of the scaled counts from individual observers divided by the square root of the number of observers. There will be additional terms added in quadrature to the random error component, for example, arising from the uncertainty in the K-coefficients. The monthly and yearly averages have uncertainties that can be found by the normal propagation of errors, and are likely to be dominated by the day-to-day variation.

All sunspot numbers (except the Hoyt Number) *never* report any uncertainties. The AAVSO can take the lead by calculating and publishing sunspot numbers with error bars. This is an easy and obvious improvement, and may possibly be done retroactively.

### 3. Small aperture and high magnification

A major problem of all the numbers that work as K(10G+F) is that the K term varies by large amounts on all time scales. In my study of sunspot visibility (Schaefer 1993), I found that the cause of much of this variability is the changing seeing and observer acuity. I also found a means to minimize this variability. First, by going to a small-aperture telescope (diameter <~2"), the visibility of the spots will be dominated by the diffraction disk as compared to the seeing disk. In a diffraction-limited case, the effects of seeing become negligible. Second, by going to high magnification, the visibility of the spots will be dominated by the diffraction disk as compared to the eye's blur circle. Again, the variation from the eye's resolution become negligible. With small aperture and high magnification, the diffraction is the primary determinant of visibility, and since the telescope aperture remains precisely constant, the variation of the K-coefficient will be minimized. The whole idea here is to make the atmosphere-telescope-observer system diffraction-limited, because aperture does not vary.

Thus another major improvement to the various sunspot numbers would be to ensure that everyone observing does so at small apertures at high magnification. Fewer than 5% of the Wolf, International, or AAVSO observers use sufficiently small apertures.

Any sudden, drastic shift to small apertures would be unwise because of logistics problems for established observers; in addition, there would be a problem of continuity. So the best way to implement this improvement is to *require* all new observers to use small apertures and high magnifications. This would not effect established observers. Yet over a decade or so in time, this would result in most observers having a stable K-coefficient. In any case, there are long-term shifts in observer equipment (such as the shift to larger aperture over the last few decades), so it is better to mandate a desirable situation than to let a poor situation evolve by chance.

### 4. Correct inflation

The AAVSO procedure for calculating its sunspot number has a subtle mathematical flaw which inflates the reported number by ~0.3% every time the K-coefficients are recalculated (Schaefer 1997). There is no doubt about the existence of this inflation, as it is a forced mathematical consequence of the AAVSO procedure. I estimate that the AAVSO number has inflated by ~6% since the founding of the AAVSO Sunspot Program in 1944. Fortunately, this inflation is not yet prominent, although even a 6% spurious increase in apparent solar activity can have profound long-term effects on public policy and health concerning the greenhouse gases and global warming.

One obvious improvement is to correct the AAVSO procedure so as to remove this flaw. Schaefer (1997) suggests an easy and minimal change, wherein the individual scaled observer's counts are logarithmically averaged instead of linearly averaged. This correction is trivial to implement.

A second obvious improvement is to deflate the historical AAVSO numbers so that the entire record is on a uniform basis. There are two methods for this deflation. One possibility is to pull out all old raw data and completely re-analyze them with the above inflation-free procedure. Unfortunately, the raw data are not available. The second possible deflation method is to change the published values by some calculated deflation factor. The deflation factor can be calculated from the average rms scatter over time of each individual observer's K-coefficient. These data may not be available, but the Zurich sunspot number publishes appropriate numbers and AAVSO observers likely have similar statistics. Then this deflation should be applied for all data after every recalculation of the K-coefficients. This second method is easy to implement reliably.

## 5. Publish raw data and procedures

Ideally, all input data and procedures should be published. For the AAVSO sunspot number, the raw data might appear in the *AAVSO Solar Bulletin*, the analyzed data might appear yearly in the *Journal of the AAVSO*, and the procedures might appear in one article in the *Journal of the AAVSO*.

Why publish the raw data? It is because later improvements (such as Hoyt's group number), later corrections (such as the deflation of the AAVSO numbers), or later additions (such as error bars) may require the original raw data to work from. Already in this paper, I have indicated that the old raw data are needed; unfortunately, the AAVSO historical sunspot raw data are apparently widely scattered and/or destroyed, so it is unclear how much utility can be derived from five decades of observations. Let us not make the same mistake again. A publication of the raw data will allow future researchers to make improvements, corrections, and additions, so that all the work by AAVSO members can retain its usefulness.

Why publish the procedures? Because there are many subjective and irregular procedures. For the AAVSO, the precise criterion (and its exceptions) for times to recalculate the K-coefficients is unpublished, and this has a drastic effect on the deduced inflation. Also unstated is under what conditions new observers are included, and when observers with discrepant K-coefficients are ignored. For the Zurich number, Wolf used geomagnetic data, and Wolfer weighted spots by their area, so the Zurich number is a time-variable measure of various activity indicators with an unknown mixture (Hoyt 1995). Also, the free choice of the Zurich preliminary sites might influence the final number. If a sunspot number cannot stand up to detailed scrutiny, then it should never appear in print.

Ideally, the procedure should involve no subjective choices by the analyzer. This subjectivity could be precluded by using a computer program that makes automated choices based on published criteria. Moreover, the analyzer's method for handling special or unusual cases could be codified into the program and published; procedural changes are fine so long as they are published. Sunspot numbers are subjective enough without the analyzer's particular choices or biases entering in.

## 6. Summary

In my analysis of the propagation of uncertainties in the AAVSO number, I have identified four specific improvements, all of which are easy to implement. With these improvements, the AAVSO can take the lead in sunspot number reliability.

## References

Hoyt, D. V. 1995, private communication.
Schaefer, B. E. 1993, *Astrophys. J.*, **411**, 909.
Schaefer, B. E. 1997, *J. Amer. Assoc. Var. Star Obs.*, **26**, 40.