How well do we know the sunspot number?

Leif Svalgaard[1]

[1]Hansen Experimental Physics Laboratory, Stanford University
650 Via Ortega, Stanford, CA 94304, USA
email: `leif@leif.org`

Abstract. We show that only two adjustments are necessary to harmonize the Group Sunspot Number with the Zürich Sunspot Number. The latter being inflated from the 1940s on to the present by 20% due to weighting of sunspot counts according to size of the spots, and corrected by inflating the earlier values as well. The Group Sunspot Number before ∼1885 is too low by ∼50%. With these adjustments a single sunspot number series results. Of note is that there is no longer a distinct Modern Grand Maximum.

Keywords. Sun: activity, sunspots

1. Motivation

A hundred years after Rudolf Wolf's death, Hoyt et al. (1994) asked "Do we have the correct reconstruction of solar activity?" After a heroic effort to find and tabulate many more early sunspot reports than were available to Wolf, Hoyt et al. thought to answer that question in the negative and to provide a revised measure of solar activity, the Group Sunspot Number (GSN) based solely on the number of sunspot *groups*, normalized by a factor of 12 to match the Wolf numbers 1874-1991. Implicit in that normalization is the assumption or stipulation that the 'Wolf' number is 'correct' over that period. In this paper we shall show that that assumption is likely false and that the Wolf number (WSN) must be corrected. With this correction, the difference between the GSN and WSN (Fig. 1) becomes disturbing: The GSN shows either a 'plateau' until the 1940s followed by a Modern Grand Maximum, or alternatively a steady rise over the past three hundred years, while the (corrected) WSN shows no significant secular trend. As the sunspot number is often used as the basic input to models of the future evolution of the Earth's environment (e.g. Emmert & Beig (2011)) and of the climate (e.g. Lean & Rind (2009)), having the correct reconstruction becomes of utmost importance, and the difference illustrated in Fig. 1 becomes unacceptable.
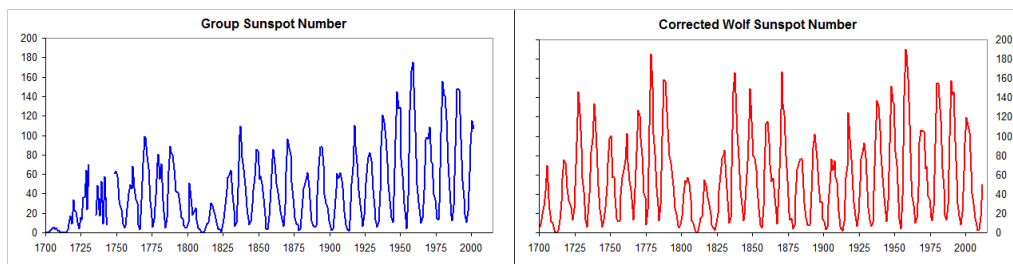

Figure 1. Contrasting views of solar activity. Left (blue curve): yearly average Group Sunspot Number. Right (red curve): The yearly average Wolf Sunspot Number corrected as described in the text. We use the designation 'Wolf' number to include the values derived by Wolf's successors in Zürich and after 1981 in Brussels.

2. The Wolf Number

Sunspots occur in groups with a coherent magnetic configuration (although this detail was not known to Wolf) and a definite evolution over several days – emergence and coalescence of smaller spots, spreading in longitude, and decay. Today we call such a group an active region and give each a number for tracking purposes. Recognizing that the birth of a new group was a much more significant event than the addition of a single new spot to the several spots already there, Wolf fashioned his Relative Sunspot Number, R, as the sum of a weighted number of groups, G, and the total number, S, of individual spots regardless of their size: $R = 10G + S$. The weight factor of 10 was chosen based both on experience (average number of spots per group) and on convenience for calculation. As telescopes of different apertures allow a different population of sunspots to be seen, Wolf introduced a scale factor, k, in the definition: $R = k(10G + S)$ to be able to reduce the observations to a common standard, which was with $k = 1$, Wolf himself when using the standard (80 mm aperture) Fraunhofer refractor at magnification 64. The k-factor depends on many things: telescope, seeing, observer visual acuity and experience, counting method, etc. and although these factors do not impact the number of groups and of spots the same way, the practice is to stick with the simplest scheme with a single k-factor for each observer at a given time. It is now clear that the homogeneity of the series and a correct measure of long-term trends depend critically on the k-factor, determined or adopted, for each observer.

Wolf published several versions and (importantly) revisions of this sunspot series. For observations before Wolf's own began in 1849, he relied primarily on Staudach's drawings 1749-1799 (Arlt (2009)) and on Schwabe's observations 1825-1848 (Arlt & Abdolvand (2011)). Up to the mid 1860s Wolf used two superb refractors built by renowned optician Joseph von Fraunhofer. One of these still exists and is used to this day to continue the Wolf number. But increasing travel (he became president of the Swiss Geodetic Commission) forced Wolf to use, eventually exclusively, smaller portable refractors (aperture 37 mm at magnification 40), which also still exist in use to this day (Fig. 2). With the smaller telescopes fewer spots (and a third fewer groups) could be seen, and Wolf determined a k-factor of 1.5 for those to match the count from the larger telescope. Observing with the larger telescope, Wolf decided (following Schwabe) not to count the smallest spots and (as he called them – "grayish") pores, visible only during exceptional seeing. With the smaller, portable telescopes these spots could not be seen anyway.

Wolf hired several assistants from the 1870s on. One of those, Alfred Wolfer, eventually became Wolf's successor. Wolfer found that Wolf's practice of omitting the smallest spots was not reproducible and advocated counting all spots that could be seen. During a 16-

Figure 2. Left: the 80 mm Fraunhofer refractor used since 1855 by Wolf and successors until 1996. At the eye piece can be seen the Merz polarizing helioscope employed to reduce light intensity to comfortable levels. Center: the same telescope in use today by Thomas Friedli (person at right). Right: the 37 mm portable telescope used by Wolf since the mid 1860s.

year period of simultaneous observations, Wolf determined that the sunspot number using Wolfer's counting method should be multiplied by a k-factor of 0.60 to match the Wolf scale for the 80 mm telescope, although Wolf compared Wolfer's count not to own counts using the 80 mm telescope, but to the count with the 37 mm portable, multiplied by 1.5. Subsequent observers at the Zürich and Brussels centers to this day adopt that same k-factor, which actually cannot be measured as Wolf is not around anymore. For the years 1876-1893, the published sunspot numbers are an average of Wolf's [using the 37 mm] and the assistants' [using the 80 mm], with k-factors applied. In 1980 the responsibility for production of the sunspot number was transferred to the SIDC in Brussels. There, rather than relying on a principal observer supplemented by secondary observers, observations are averaged over a network of ∼65 observers, all scaled to the reference station, Locarno in Southern Switzerland for which a k-factor of 0.60 is simply adopted.

3. The Weighting Scheme

In 1945 Max Waldmeier became director of the Zürich Observatory and in charge of production of the Wolf Number. He noted (Waldmeier 1948) Wolfer's different counting procedure, namely also counting the smallest spots that Wolf was omitting. But also claimed that Wolfer started (around 1882) weighting the sunspot count by the size and structure of each spot. In Waldmeier (1968) the weighting scheme is described as follows: "A spot appearing as a fine point is counted as one spot; a larger spot, but still without penumbra, gets the statistical weight 2, a smallish spot with penumbra gets 3, and a larger one gets 5." Presumably there would also be spots with weight 4. However, we have found *no* mention of this scheme in any of the, otherwise meticulous, reports or papers by Wolf, Wolfer, and Brunner (director 1926-1945), as was also noted by Kopecký et al. (1980). Nor do other sunspot observers (professional or amateur) not affiliated with the Zürich observers employ such a weighting scheme (or any), which is, indeed, generally unknown.
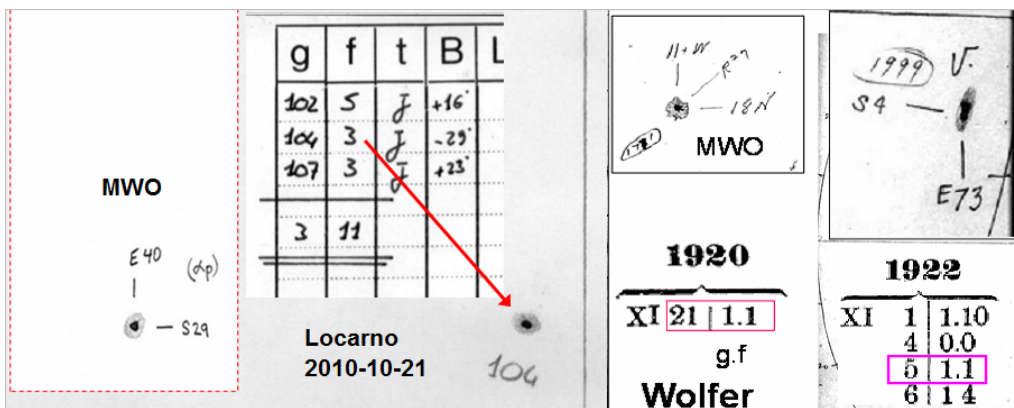

Figure 3. Left: Part of a drawing made at Locarno, showing that the spot with penumbra designated '104' was counted three times (weight 3). The insert shows the spot as observed at Mount Wilson Observatory (MWO). Right: Two spots with the same area (∼220 μhemispheres) on drawings from MWO and counted with weight 1 by Wolfer (in each case as one group with one spot (Wolf's notation 'groups.spots' = '1.1'), as the group was the only group on the disk).

It is easy to show (Fig. 3) that Wolfer did not employ any weighting, contrary to the claim by Waldmeier. There are many days (43, in fact) when there was exactly one group on the disk having exactly one spot with an area of $\geqslant$∼200 μhemispheres that should have been counted with weight 3, for a total sunspot number of at least $0.6 \cdot (10 \cdot 1 + 3) = 7.8 \equiv 8$, yet were all reported as $0.6 \cdot (10 \cdot 1 + 1) = 6.6 \equiv 7$.

4. The Effect of the Weighting

To determine how much the weighting scheme inflates the sunspot numbers we for each day 'undo' the weighting by counting on Locarno's drawings each spot only once for comparison with the sum of the weighted counts (Fig. 4). Locarno was founded as 'backup' station for Zürich. The station, located on the other side of the Alps has often complementary weather.
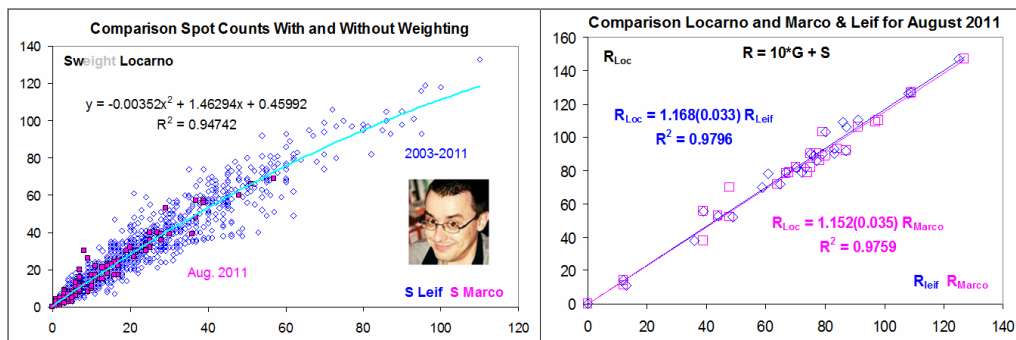

Figure 4. Left: (blue open diamonds) The weighted sunspot count compared to the simple count for 2557 days, 2003-2011. The double-blind control count is shown by pink squares. Right: Comparison of calculated relative sunspot numbers for August, 2011; blue open diamonds: Leif Svalgaard, pink open squares: Marco Cagnotti, Locarno (pictured). The standard errors of the slopes of the regression lines are given in parentheses.

As the drawings from the projected images are at times only indicative (and thus not always a true reflection of the actual, reported observations done visually at the eyepiece) it was felt necessary to perform a double-blind control test. The observer at Locarno would also record the unweighted count for each day in addition to the reported 'official' visual count. This was done for the month of August, 2011 (and is continuing). As is clear from Fig. 4, there is no significant difference between the resident observer and the probing solar physicist, thus preliminarily validating the technique. The result is that the weighting for a typical number of spots increases the spot count by 30-50%. But we are interested in the effect on the relative sunspot number where the group count will dilute the effect by about a factor of two, as shown in the right panel of Fig. 4 where an inflation of 15-17% is evident. The inflation weight factor, W, can be determined for each of the years 2003-2011 and we find that there is a weak dependence on the sunspot number: $W = 1.13 + R/2500$. For a typical sunspot number of $R = 100$, W becomes 1.17.

Waldmeier also introduced a new classification of groups (the Zürich classification) based on understanding of the evolution of the group rather than mere proximity of the spots. This tends to increase the number of groups over that that proximity would dictate. We find that, on average, on a fifth of all days an additional group is reported which means that the relative sunspot number increases by about 3% due to this inflation of the group count brought about by the better understanding of what constitutes a group. Kopecký et al. (1980) quote the observer Zelenka suggesting a possible influence of the new Zürich classification of groups. The combined effect of the weighting and the classification might thus be of the order of 20%, which can be corrected for by increasing the pre-1945 values by a factor 1.20, as done in Fig. 1. Can we see this effect by comparing with other solar activity measures?

5. The Waldmeier Discontinuity

The Group Sunspot Number during most of the 20th century is based on the Greenwich Helio Photographs and might be suspected of having a constant calibration over that interval. We can form the ratio WSN/GSN when neither is too small to examine its stability. As shown in upper left of Fig. 5 there is a clear discontinuity corresponding to a jump of a factor 1.18 between 1945 and 1946 as we would expect if the weighting was introduced at that time. The sunspot area, S_A, (Balmaceda et al. (2009)) has a strong (albeit slightly non-linear) relationship with the sunspot number: $R = A \, S_A^{0.732}$. The ratio $A = R / S_A^{0.732}$ plotted in the right-hand panel of Fig. 5 shows the same discontinuity around 1945, with an increase by a factor of 1.21 (as established by the histograms of A in the lower left of the Figure).
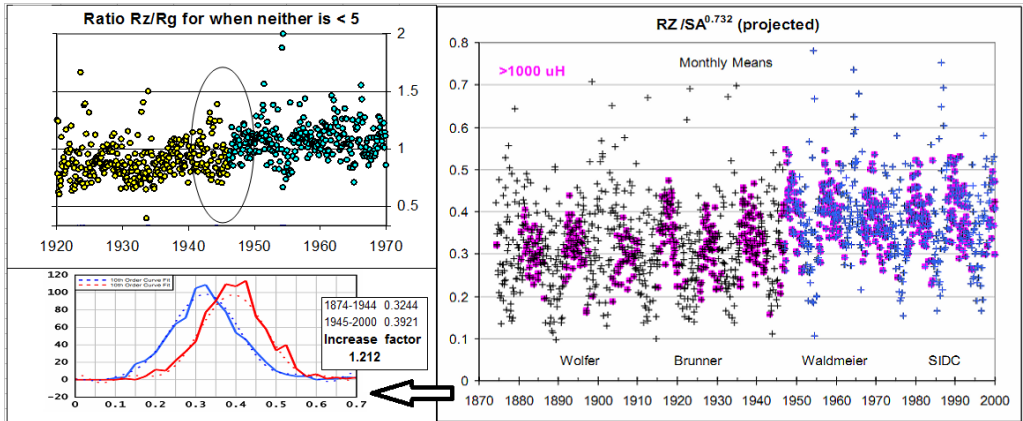

Figure 5. Upper left: Ratio between WSN and GSN, monthly values 1920-1969. Right: Ratio between WSN and linearized sunspot area. Monthly averages greater than 1000 μhemispheres are marked by pink squares. Lower left: Histograms of that ratio before and after 1945.

From ∼40,000 Ca II K-line spectroheliograms from the 60-foot tower at Mount Wilson between 1915 and 1985 a daily index of the fractional area of the3 visible solar disk occupied by plages and active network has been constructed (Bertello et al. (2008)). Monthly averages of this index is strongly correlated with the sunspot number, with the expression $R = 27235 \, CaK - 67.14$ before 1946. After that year, the sunspot number calculated from this expression is too high by 19% (Fig. 6). An equivalent conclusion was reached by Foukal (1998).
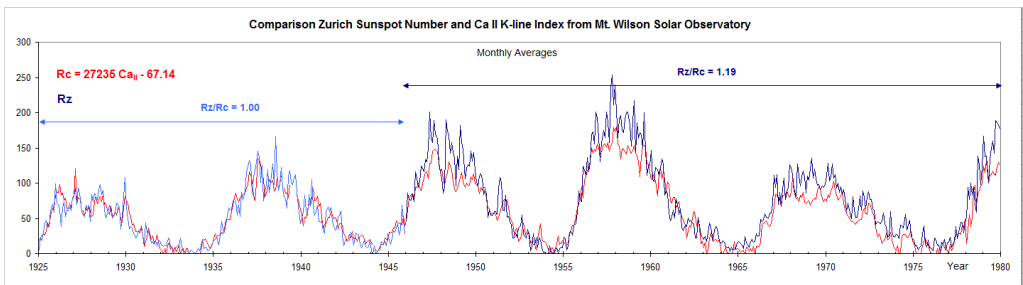

Figure 6. Sunspot number (red curve) calculated from MWO CaII K-line spectroheliograms compared to observed Wolf Numbers (blue curve).

6. Terrestrial Evidence

The F2 layer critical frequency is the maximum radio frequency that can be reflected by the F2 region of of the ionosphere at vertical incidence and has been found to have a profound solar cycle dependence which showed a difference between cycle 17 and cycle 18 (Ostrow & PoKempner (1952). It was necessary to shift the sunspot number 21% between the two cycles in 1945 to obtain the same relationship. Again we see the same discontinuity.

Wolf (1859) discovered the linear relationship between the amplitude, v, of the diurnal variation of the Declination of the 'magnetic needle': $v = a + b R$, and used it to calibrate the Wolf Number. The variation of the Declination (Fig. 7) was discovered in 1722 by George Graham and reliable measurements stretch back in to the 1780s.
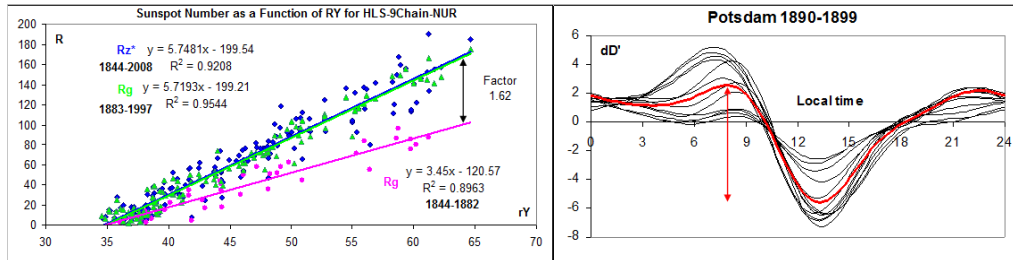

Figure 7. Right: Twelve Monthly curves showing the variation during the day of the magnetic Declination. The red curve and arrow show the yearly average variation which is a strong function of the sunspot number (controlled by Far Ultraviolet (FUV) radiation). Left: The relationship between a many-station series of diurnal variation of the East-West component of the geomagnetic field – which is calculated from the variation of the Declination and which is actual physical parameter controlled by solar FUV – and the Wolf Number (blue diamonds) back to 1844. The Group Sunspot Number has the same relationship (green triangles) after 1882, but is seriously too small before (pink dots).

It is clear that there is a consistent relationship between the diurnal range and the Wolf sunspot number, establishing an objective calibration, but that the Group Sunspot Number is seriously too low before the 1880s.

7. Implications

The possible implications from this reassessment of the sunspot series are so serious that a series of workshops are planned to resolve these questions. The first was held at Sunspot, NM in September, 2011, (`http://ssnworkshop.wikia.com/wiki/Home`).

References

Arlt, R. 2009 *Sol. Phys.*, 255, 143
Arlt, R., & Abdolvand, A. 2011 *Proc. IAU Symposium 273*, 6, 286
Balmaceda, L.A., Solanki, S.K., Krivova, N.A., & Foster, S. 2009, *J. Geophys. Res.*, 114, A07104
Emmert, J.T., & Beig, G. 2011, *J. Geophys. Res.*, 116, A00H01
Foukal, P. 1998, *Geophys. Res. Lett.*, 25, 2909
Hoyt, D.V., Schatten, K.H., & Nesmes-Ribes, E. 1994, *Geophys. Res. Lett.*, 21, 2067
Kopecký, M., Ružičková-Topolová, B., & Kuklin, G.V. 1980, *Bull. Astron. Inst. Czech*, 31, 267
Lean, J.L., & Rind, D.H. 2009, *Geophys. Res. Lett.*, 36, L15708
Ostrow, S.M., & PoKempner, M. 1952, *J. Geophys. Res.*, 57, 473
Waldmeier, M. 1948, *Astron. Mitteil.*, Nr. 152
Waldmeier, M. 1968, *Astron. Mitteil.*, Nr. 285
Wolf, J.R.. 1859, *Mittheil. Sonnenfl.*, Nr. 9