



SOLAR ACTIVITY – PAST, PRESENT, FUTURE

A Personal Journey

Leif Svalgaard

Stanford University, California, USA¹

Keywords

Solar Activity, Cycle Prediction, Time Variation, Historical Data, Solar Wind

Abstract

As our civilization depends increasingly on space-borne assets and on a delicate and vulnerable earth-bound infrastructure, solar activity and its potential impact becomes of increasing importance and relevance. In his famous paper on the Maunder Minimum, Eddy (1976) introduced the notion that the Sun is a variable star on long time scales. After the recent decade of vigorous research based on cosmic ray and sunspot data as well as on geomagnetic activity, an emerging consensus reconstruction of solar wind magnetic field strength has been forged for the last century. The consensus reconstruction shows reasonable agreement among the various reconstructions of solar wind magnetic field the past ~170 years. New magnetic indices open further possibilities for the exploitation of historic data. The solar wind is a direct result of solar magnetic activity providing an important link to the effects on the Earth's environment. Reassessment of the sunspot series (no Modern Grand Maximum) and new reconstructions of Total Solar Irradiance also contribute to our improved knowledge (or at least best guess) of the environment of the Earth System, with obvious implications for management of space-based technological assets or, perhaps, even climate. Several lines of evidence suggest that the Sun is entering a period of low activity, perhaps even a Grand Minimum. Average space weather might be "milder" with decreased solar activity, but the extreme events that dominate technological effects are not expected to disappear. Prediction of solar activity has a poor track record, but the progression of the current Cycle 24 is in accordance with its behaviour predicted from the evolution of the solar polar fields, so perhaps there is hope.

Introduction

Solar activity is the result of solar magnetic fields. If our Sun had no magnetic field it would be as dull as models of stellar constitution proscribe and we would not have this conference. The magnetic field makes the Sun *interesting*, which before the development of our technological civilization was of little consequence, but that the Sun is a variable magnetic star is today of immense practical importance; in fact, a potential danger to our modern way of life. Increasingly, our civilization depends on environmental conditions, communication devices, and infrastructures that are vulnerable to solar magnetic variability [NRC, 2008]. The famous Carrington Event in 1859 [e.g. Cliver & Svalgaard, 2004] can be said to mark the birth of this concern, although the technological effects, some even damaging with attendant economics loss, of geomagnetic disturbances [accompanied by brilliant aurorae] on the nascent telegraph communication capability were already noted more than a decade before.

¹ Leif Svalgaard, W.W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305, USA; email: leif@leif.org

To assess the impact of solar activity and the chances of effective mitigation of its effect we need to monitor and understand not only current space weather, but also space *climatology*: what is the equivalent of a ‘hundred-year flood’? Direct telescopic observation of solar activity, of course, begins with the discovery 402 years ago of sunspots. Our understanding of that historical record forms the basis for interpreting the indirect evidence both from natural archives (e.g. ^{10}Be from ice cores) and human naked-eye observations (aurorae, 日誌 [ri-zhi] blemishes on the sun) stretching much farther back in time.

The Sunspot Record(s)

The historical sunspot record was first assembled by Rudolf Wolf in 1850s and has been continued by Wolf and his successors followed by a more ‘institutionalized’ approach later in the 20th century until today and hopefully beyond. Wolf’s original definition of the *Relative Sunspot Number* for a given day as $R = 10 \text{ Number of Groups} + \text{Number of Spots}$ visible on the solar disk has stood the test of time and recognizes [in Wolf’s own words] that the emergence of a coherent group of spots (an active region) is much more important than the addition of yet a few spots to an existing group. The factor of 10 has also turned out to be a good choice as historically a group contained on average ten spots.

A fundamental problem is the homogeneity of the series, that is: does a relative sunspot number of 100 mean the same level of solar activity today as it did in 1938, in 1872, or in 1739? And what is a useful definition of ‘solar activity’ anyway? From the viewpoint of solar effects on our technological infrastructure, the *solar wind* – the ever present expanding outer atmosphere of the Sun – is perhaps the most relevant element, although bursts of highly energetic particles and radiation also degrade devices and spacecraft and threaten humans in space. Almost all solar indices and solar wind quantities show a relationship with the Relative Sunspot Number [SSN], so homogeneity and proper calibration of the SSN become of utmost importance.

Hoyt et al. [1994] in a series of papers at the centenary of Wolf’s death asked “Do we have the correct reconstruction of solar activity?” and proposed to answer the question in the negative. A heroic effort from an extensive search of archives and primary sources yielded ~350,000 observations, many not available to Wolf, covering the interval 1610-1993. An earlier study showed that the ratio of individual spots to groups is nearly a constant. Theoretical arguments [Schaefer, 1993] showed that Wolf’s Relative Sunspot Number may be set equal to a constant times the number of sunspot groups, so an index based solely on the number of sunspot groups can simulate the Wolf SSN. Hoyt & Schatten called this index the Group Sunspot Number [GSN] and found that it appears that solar activity has steadily increased since 1700 to what might be called a Modern Grand Maximum in the latter part of the 20th century. Before ~1885 the GSN is significantly smaller than the Wolf SSN that does not support the idea of a Modern Grand Maximum. This discrepancy is not satisfactory and must be resolved so solar-terrestrial researchers have a stable and unique dataset to work with.

Recognizing the need to resolve this issue, a number of workshops on the calibration of the sunspot number have been sponsored by the National Solar Observatory (NSO), the Royal Observatory of Belgium (ROB), and the Air Force Research Laboratory (AFRL) as an effort to provide the solar community with a vetted long-term (single) sunspot number series and the tools to keep it on track. The first workshop was held at Sunspot, New Mexico [yes, there is such a place] in September 2011, followed by a second workshop in Brussels in May 2012. Further meetings will take place in 2013 in Tucson, AZ and then in Switzerland. We are considering a special Topical Issue of Solar Physics for the eventual joint publication of the SSN series and the accompanying historical, procedural, and scientific papers. In this paper, I’ll report on the progress made so far.

An efficient way of comparing the Wolf Sunspot Number and the Group Sunspot Number is to plot the ratio between them as shown in Figure 1. That removes most of the solar cycle

variation and will show obvious discontinuities caused by non-solar related changes in the calibration. Figure 1 shows two clear discontinuities, one near 1945 and one near 1885.

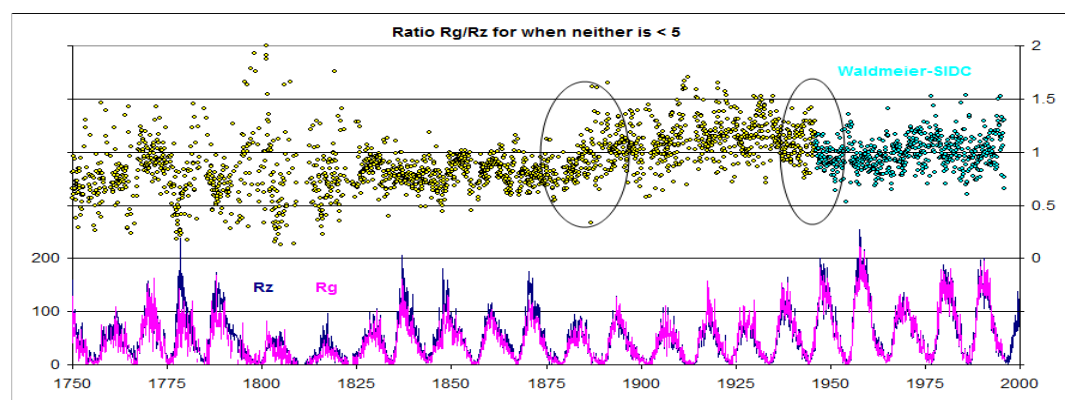


Figure 1: Ratio between monthly values of the Group Sunspot Number (R_g) and the Wolf (Zürich, International) Sunspot Number.

Explaining those will go a long way to resolve the differences between the two series. We shall show that the 1945 discontinuity is a problem with the Wolf Number while the 1885 discontinuity is a problem with the Group Number. When both problems are corrected, there is no longer a significant difference between the two series.

The Waldmeier Discontinuity

Some time in the 1940s the observers in Zürich began to *weight* sunspots when counting them. The director of the Zürich Observatory, Max Waldmeier described [Waldmeier, 1968] the procedure thus “A spot like 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.” This weighting increases the spot count by 45% on average and, since the spot count is half of the Relative Number, the SSN by approximately half that. The Locarno Station in Southern Switzerland has since 1957 served as an auxiliary [site](#) for Zürich [as the weather on opposite sides of the Alps often is complimentary] and is still today the reference station for the modern sunspot number maintained by SIDC in Brussels, as all other observers are normalized to Locarno’s count. Locarno is still weighting the spots according to Waldmeier’s prescription, so the weighting carries over fully into the current sunspot number. Normalization is done by applying a factor, k , in the formula for the Relative Sunspot Number $R = k(10G + S)$ such that different observers of the number of groups, G , and the number of spots, S , arrive at the same relative number, R . The k -factor depends on several things: telescope aperture and magnification, observer acuity, atmospheric seeing, and the precise way spots are recognized and grouped. The Zürich observers after Wolf chose to count all spots that were visible, while Wolf did not include the smallest spots near the limit of detectability, in order to be compatible with Schwabe’s observations. Consequently, a k -factor of 0.6 was, at first, empirically determined and later simply adopted to reduce the sunspot number to the original Wolf scale.

Figure 2 shows the effect of weighting using a typical drawing from Locarno. There is also a small contribution from an improved classification of sunspot groups introduced at about the same time. The combined net effect is to increase the sunspot number since ~1945 by ~20%. This explains the discontinuity in the ratio GSN/SSN at that time. A strong check and confirmation of the effect of the weighting have been carried out the past year by the Locarno observers, counting both with and without weighting, so the magnitude of the effect of the weighting is now established and is no longer an open issue. The effect of the weighting turns out to be almost independent of solar activity, so a simple corrective action would be to multiply all numbers before 1945 by 1.20. Figure 3 shows the result of the correction. Such ‘wholesale’ correction is not without precedent. In 1861 Wolf published his first long list of the SSN covering the years 1749-1860. Around 1875, Wolf increased all the numbers on the

1861 list before 1848 by 25%, based on measurements of ‘the magnetic needle’, which we’ll hear more about later in this paper.

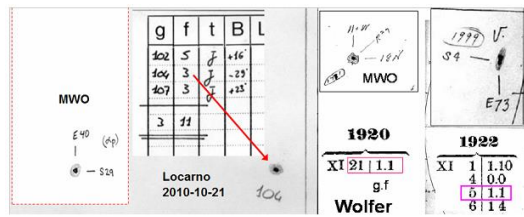


Figure 2: (Left) Part of a drawing made at Locarno, showing that the spot with penumbra designated 104 was counted three times (weight 3); the insert shows the same spot from Mt. Wilson. (Right) Two spots with the same area on drawings from Mt. Wilson and counted with weight 1 by Wolf, in each case as one group with one spot (Wolf’s notation: *groups.spots = 1.1*), as the group was the only group on the disk

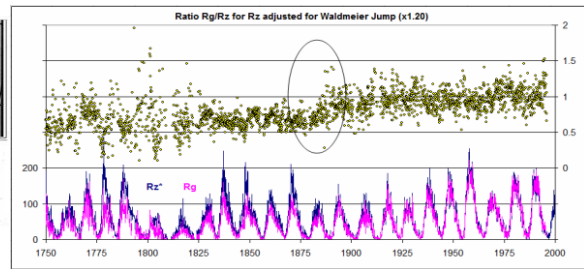


Figure 3: Removing the effect of weighting after 1945 removes the discontinuity in 1945. There remains the discrepancy around 1885.

The lower panel shows the corrected SSN (R_z) in blue and the GSN in pink. Note that the corrected SSN reaches about the same level in each century

Group Sunspot Number Calibration

Hoyt & Schatten [H&S] quickly found that the constant that converted a group count to a simulated Wolf Relative Number was not a constant at all, but varied with the observer, and was, in effect, a k -factor to be applied to the number of groups. A decision must be made as to which observer should have a k -factor of unity. H&S chose the ‘Helio-Photographic Results’ tabulations from Royal Greenwich Observatory [RGO] covering the years 1874-1975 as their ‘standard’ observer. Because Wolf and successor Alfred Wolfer observed from 1848 through 1928 with overlap 1876-1893, their combined observations can serve as a ‘backbone’ for the determination of k -factors for other observers overlapping with them before the RGO series begins. H&S determined that the ratio between k -factors for Wolf and Wolfer was 1.021, i.e. that both observers saw very nearly the same number of groups. And herein lies the problem.

Because of extensive travel and other duties, Wolf from the 1860s until his death in 1893 exclusively used a small, handheld telescope, while Wolfer used a larger 80mm Fraunhofer telescope similar to what Wolf used earlier. Both telescopes still exist (Figure 4) and are even in use today by Thomas Friedli in Belp, near Bern, continuing the Wolf tradition.



Figure 4. (Left) the 80 mm Fraunhofer refractor used since 1855 by Wolf and successors. (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.

Figure 5 shows that for the time of overlap between Wolf and Wolfer, Wolfer counted [as appropriate for the larger telescope] 1.653 times as many groups as Wolf, and not only 1.021 times as determined by H&S. This discrepancy is the main cause of the 1885 discontinuity in the GSN/SSN ratio. We can understand the reason why Wolf saw so few groups by considering that the small spots making up groups with Zürich classifications a and b (groups containing spots without penumbra) are usually not visible in Wolf’s small telescope. Such groups make up about a half of all groups.

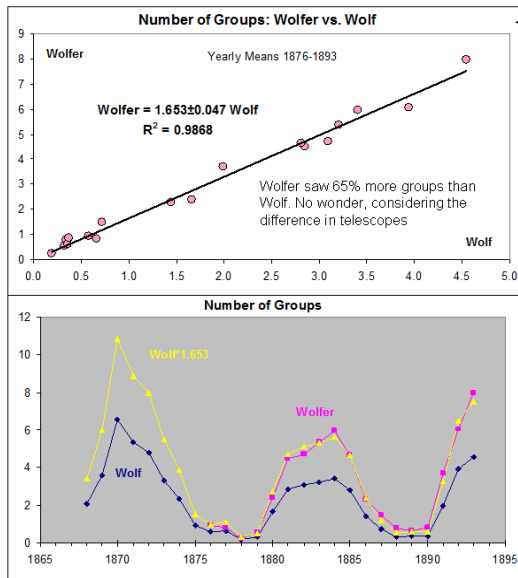


Figure 5: (Top) Number of groups observed by Wolfer compared to the number observed by Wolf. (Bottom) Applying the *k*-factor of 1.653 makes Wolf's count (blue) match (yellow) Wolfer's count (pink)

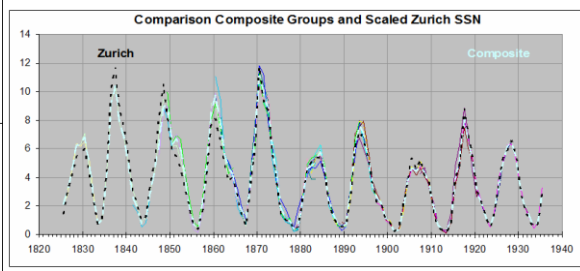


Figure 6: Different colors (hard to see because they fall on top of each other) show the individual contributions to the composite (bright cyan). The black dashed curve is the Zürich SSN / 12 (to convert it to number of groups).

Using the Wolf-Wolfer composite record as a backbone, we can now confidently determine *k*-factors for 22 mutually overlapping observers stretching back to Schwabe and forward to Brunner and construct a composite series. As Figure 6 shows there is now no significant difference between the GSN and the SSN. So, with only two adjustments: ~20% for SSN before 1945 and ~50% for GSN before 1885, the discrepancy is resolved. The issue of the very early data, say before 1825, is still open and is the target for the next SSN Workshop, but if people can accept the current series without adjustments, then they might also accept that we, for now, assume that there are no further adjustments warranted for the early data.

Geomagnetic Calibration of Sunspot Numbers

Wolf [1852] discovered a beautiful connection between sunspots and the diurnal variation of the Earth's magnetic field. He marvelled "Who would have thought just a few years ago about the possibility of computing a terrestrial phenomenon from observations of sunspots", and he at once realized that such a relation could be used as an independent check on the calibration of the sunspot number. Today we understand the physics of that connection and can fully validate Wolf's assertion. Solar Far Ultraviolet [FUV] radiation creates and maintains the conducting *E*-layer in the ionosphere. Thermal winds driven by solar heating move the charges across the Earth's magnetic field setting up an ionospheric dynamo with currents generating magnetic effects observable on the ground. As the Earth rotates under the currents (which are fixed in direction to the Sun) a characteristic diurnal perturbation of the geomagnetic field is readily observed (discovered by Graham in 1722). The diurnal variation is best seen in the East component of the geomagnetic field. The 10.7 cm microwave flux from the Sun is a good proxy for the FUV flux and Figure 7 shows how well the amplitude of the diurnal variation tracks 10.7, thus validating Wolf's procedure using modern data.

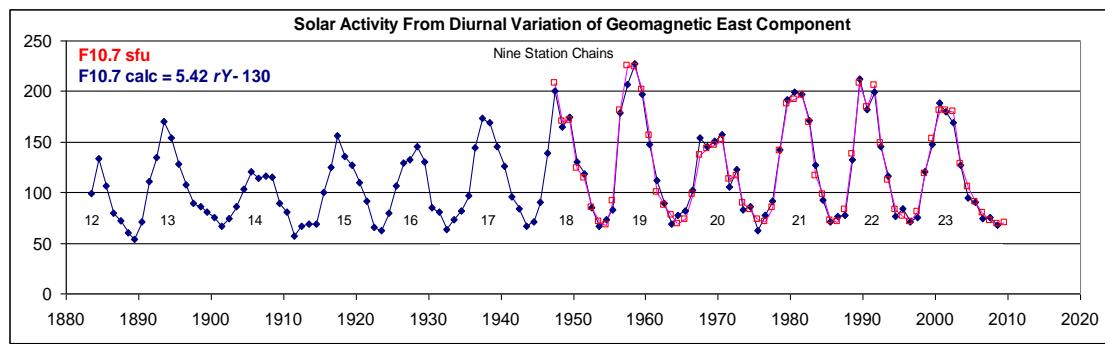


Figure 7: 10.7 flux correlation with the range, *rY*, of the diurnal variation of East Component

The Figure also shows the ‘equivalent’ 10.7 flux calculated from the observed range (average of nine stations) of the diurnal variation for many solar cycles before the advent of the flux radiometers. It is clear that the well-understood physics of causes of the diurnal variation of the geomagnetic field provides a reliable way of assessing the past variation of solar flux, and hence the magnetic activity responsible for it; a variation for which the sunspot number is a proxy. Careful application of this method fully supports the two adjustments of the sunspot series described above based on the diurnal variation as observed since the 1780s.

Solar Wind Properties in the Past

Direct in-situ observation of the solar wind goes back 50 years and it was clear from even the earliest data that geomagnetic activity (separate from and superposed on the regular diurnal variation just discussed) was directly controlled by the expanding solar atmosphere – by the expansion speed and by the strength and direction of the magnetic field dragged out from the Sun. Recent research [Svalgaard & Cliver, 2005, 2007, 2010; Lockwood & Owens, 2011] has shown that it is possible to infer the solar wind speed and the magnetic field strength from suitable, newly defined, indices of geomagnetic activity that have been found to respond to different combinations of these solar wind parameters, allowing the influence of each to be separately extracted and calibrated by comparison with the space-based data measured near the Earth, effectively inverting the ‘response function’ of the Earth to the solar wind. Figure 8 shows one result of this inversion.

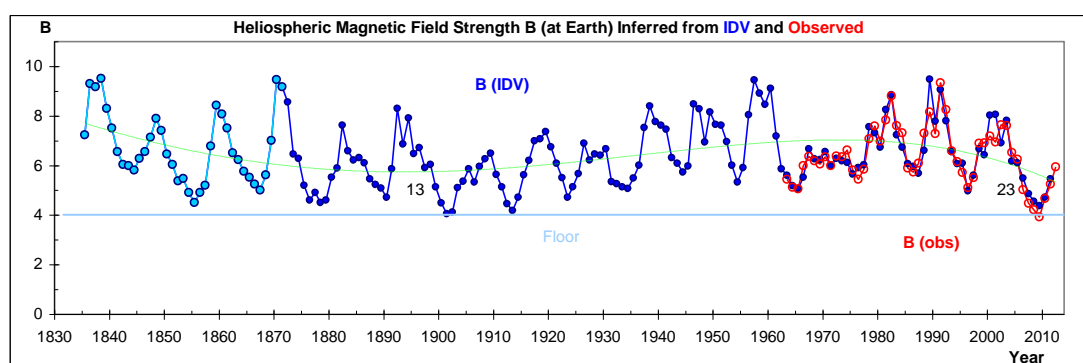


Figure 8: Heliospheric magnetic field strength B at Earth inferred from the *IDV* geomagnetic index (blue) and observed by spacecraft (red). Data before 1872 (light blue) are preliminary and can be improved by adding more 19th observatory data.

Data exist to carry this inversion back to the earliest systematic observations of the geomagnetic field in the 1830s. We find that the reconstruction of the solar wind is consistent with the re-assessment of the sunspot number series described above, in particular that there also does not seem to have been a Modern Grand Maximum in solar wind parameters. Another finding of interest is that even during periods of extremely low solar activity [e.g. the years 1901-1902, 2008-2009] the solar wind is still present having a respectable magnetic field of about 4 nT. An important research issue at present is whether this minimum state of solar magnetic activity, a ‘floor’, is a general feature, at all times [Schrijver et al., 2010]. Recent work by Owens et al. [2012] suggests that “even a steady decline in sunspot number may result in a plateau in the Open Solar (magnetic) Flux”.

The Cosmic Ray Record

Cosmic ray particles reaching the Earth are mostly produced outside the solar system during supernova explosions and subsequent acceleration in the Galaxy. Two time-varying magnetic ‘shields’, the solar magnetic field and the geomagnetic field, modulate the cosmic ray flux. The weaker these fields, the higher is the cosmic ray intensity near the surface of the Earth. Ionization chambers and neutron monitors have directly monitored the intensity of cosmic radiation since the 1930s [Steinhilber et al., 2012]. Before that, no direct measurements exist, and cosmogenic radionuclides, are used as a proxy for cosmic radiation, especially ¹⁰Be and

^{14}C , produced by cosmic rays colliding with atmospheric nitrogen and oxygen. Thus, the production rates of these nuclides are related to the flux of the incoming cosmic rays. As with the sunspot number, there are issues with the proper calibration of the cosmic ray proxies. What are measured are not variations of the production rate, but of the deposition rate, which in addition to the incoming flux also depend on atmospheric circulation and climate in general. Inversions of the ionization chamber data to extract the strength of the solar magnetic field are discordant with inversions of the neutron monitor data and with the result of the geomagnetic constructions. This issue will eventually be resolved and a special ISSI workshop towards this goal is ongoing [Svalgaard et al., 2011].

Of special interest is the cosmic ray record of so-called ‘Grand Minima’, like the Maunder Minimum. The solar magnetic field (expressed as the near-Earth solar wind magnetic field) extracted from the cosmic ray record falls to zero or at times is un-physically negative during Grand Minima (Figure 8) while at the same time a vigorous solar cycle modulation of the cosmic ray flux is observed [e.g. Berggren et al., 2009, Figure 9] indicating to this author that significant solar magnetic field was present.

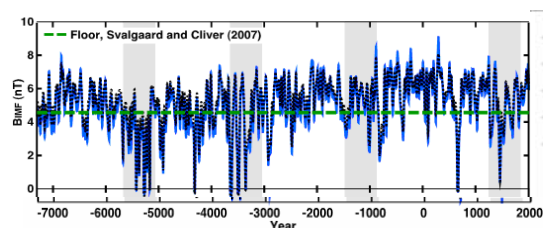


Figure 8: Heliospheric magnetic field strength at Earth inferred from the cosmic ray record [Steinhilber et al., 2010]. The deep excursions to zero or even un-physical negative values are not understood and may be artefacts caused by too aggressive extrapolation from modulation potential to field strength. This is presently an open and controversial issue.

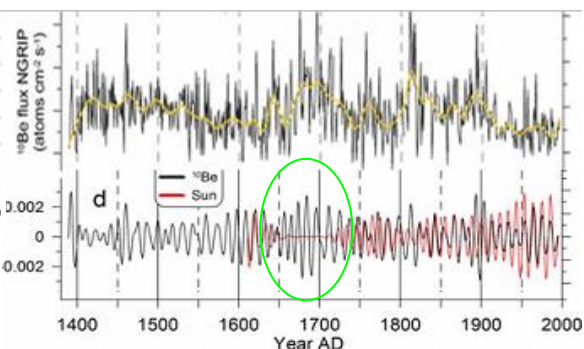


Figure 9: (Top) ^{10}Be flux from the NGRIP ice core (Greenland). (Bottom) The ^{10}Be data band pass filtered at 8–16 years [Berggren et al., 2009]. The red curve is the filtered GSN series.

Similarly, observations of the spicule forest (the ‘red flash’) during the total solar eclipses in 1706 and 1715 seem to require the presence of bright network structures, and thus of substantial solar photospheric and chromospheric magnetism during at least the last decades of the Maunder Minimum [Foukal & Eddy, 2007].

The very long cosmic ray record (when calibrated and understood correctly – including the effect of solar cosmic rays) provides the necessary material for statistical studies of the frequency and features of extremes of solar activity (the ‘500-year floods’). The first order of business is to understand why the variations are discordant compared to other solar indicators the past 400 years. This effort is ongoing and the results are not yet in sight.

Predicting Solar Activity

At this point in time it has become of great practical and societal importance to predict solar activity and space climate, rather than just recording them, e.g. for planning mitigation of the effects of extreme solar events. The NASA/NOAA international Panel for Predicting the Solar Cycle examined 75 ‘predictions’ of the current Cycle 24, basically covering the full spectrum of possible outcomes centered on the climatological mean, from extremely small to extremely large [Pesnell, 2012]. The Panel ended up (barely) endorsing the ‘precursor’ methods as the most promising where some property of the Sun near minimum is used as a predictor of the following cycle. On physical grounds, the magnitude of the polar fields of the Sun seems to be a good candidate as a precursor as it is thought that the polar fields act as a ‘seed’ for the dynamo producing the next cycle [e.g. Jiang et al., 2007]. Leading up to the minimum in 2008, the polar fields were the weakest ever observed (since the invention of

reliable solar magnetographs in the 1950s) prompting Svalgaard et al. [2005] and Schatten [2005] to predict that cycle 24 would be ‘the smallest cycle in a hundred years’. This prediction has, so far, held up well, providing (together with predictions using other precursors, such as the geomagnetic *aa*-index) a successful test of the precursor notion. Figure 10 compares the polar fields and the size of the next cycle for the past several cycles.

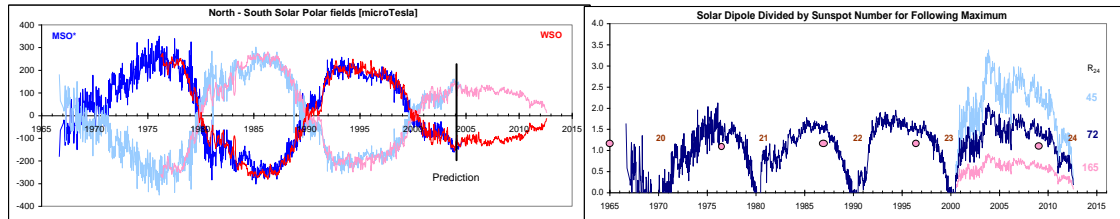


Figure 10: (Left) The axial magnetic ‘dipole moment’ of the Sun, defined as the difference between the (signed) magnetic field strength near the North Pole and near the South Pole, for the past four solar cycles. To facilitate comparison of cycles, a ‘ghost’ mirror image is also plotted. Data from Mount Wilson Observatory (blue) has been scaled to match observations from Wilcox Solar Observatory (red). (Right) The polar field variation at the left scaled by the maximum smoothed sunspot number for the next cycle suggesting the same variation for each cycle, i.e. that the cycle maximum is controlled by the polar fields at the preceding minimum. For Cycle 24, the maximum was not known, but a very high value (165) or a very low value (45) does not fit the pattern. A maximum size of 72 for Cycle 24 seems to be the just right ‘Goldilocks’ value.

Figure 11 shows how the prediction is doing. The quantity plotted is the total number of active regions per month on the disk within 70° of Central Meridian (which is on average 2.25 times the sunspot number).

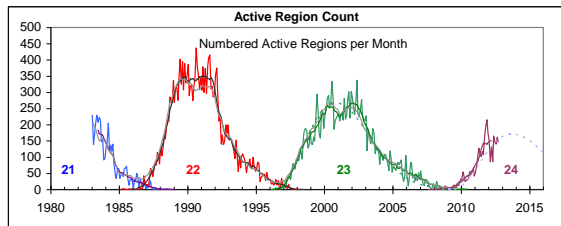


Figure 11: Numbered active regions within 70° of CM per month. Different cycles are plotted with different colors. The predicted Cycle 24 is shown as the dashed purple curve. We are very close to a drawn out solar maximum at this time of writing.

Recent Changes in the Sun?

Historically all solar indices have been closely correlated as they all derive from the same source: the variable magnetic field. In fact, the various reconstructions of past (and predictions of future) activity all rely on the implicit assumption that the correlations stay the same over time. This is likely to be true for indices that have a close physical connection, like the 10.7 and UV fluxes, but is not given a priori for correlations that are more indirect, e.g involving the sunspot number: the processes creating visible sunspots are varied and not fully understood. And indeed, while there has long been a stable relationship between the 10.7 flux and the sunspot number, allowing one to calculate or map one from the other, that relationship has steadily deteriorated in the past decade to the point where the sunspot number for a given flux has decreased by about a third (Figure 12).

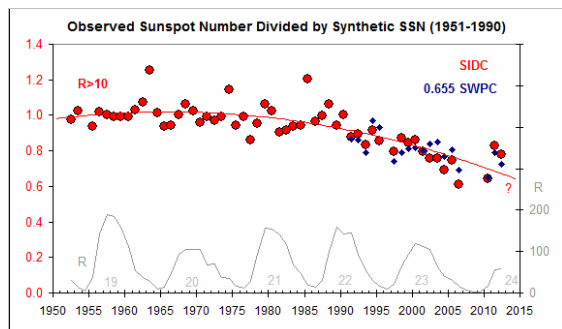


Figure 12: The observed SSN divided by a synthetic SSN computed from a polynomial fit to the 10.7 flux over the interval 1951-1990. Red points are SIDC sunspot numbers while blue points are SWPC (NOAA) sunspot numbers scaled down by the their average *k*-factor of 0.655. Only years when the SSN was above 10 are included. Since ~1990 the observed SSN is progressively lower for a given 10.7 flux.

A similar decrease of the sunspot number for a given amount of magnetic flux in the plages surrounding active regions has occurred as well as a decrease of the number of spots per active region. This is unprecedented in the observational record. We interpret this decrease as a loss of primarily the small spots [Clette & Lefèvre, 2012].

Observations by Livingston & Penn since 1998 until the present show that the average magnetic field in sunspots has steadily decreased by 25% [Livingston et al., 2012], regardless of the fact that we are now again at the maximum of a solar cycle, so there has not been a solar-cycle-related reversal of the trend. Since their magnetic fields cool sunspots, a decreasing field means that sunspots are getting warmer and that their contrast with the surrounding photosphere is getting smaller, making the spots harder to see. There is a minimum field strength in visible spots of about 1500 Gauss [0.15 T] and as that 1500 G threshold is approached, magnetic fields appear at the solar surface which do not seem to form dark sunspots or pores. Owens et al. [2012] suggest that the photospheric flux emergence in such cases may take place in flux tubes with field too weak, or of too small a diameter, to form sunspots, citing Spruit [1977]. The observed distribution of number of spots vs. field strength has been shifting steadily towards that limit. *If*, and that is a big *IF*, this trend continues, the number of visible spots in the next cycle [and perhaps beyond] may fall to values not seen since the Maunder Minimum, but without dramatic changes in the emerging magnetic flux. Without the dark spots, Total Solar Irradiance might even be a bit higher. It is not clear what this will mean for the impact of solar activity on the Earth's environment, if any, but it portends exciting times for solar physicists.

Discussion

Our technological civilization has reached a point where solar activity and its prediction on all times scales have become significant factors in maintaining and safekeeping of the technological infrastructure, both on the ground and in space. New capabilities, in instrumentation, deployment, computer storage and power, and – last, but not least – increased awareness promise progress towards predictive improvements. On the other hand, if the Sun is moving into a new regime of lower activity, a period of uncertainty may make life hard for the forecaster. The possible, recent re-assessment of past solar activity should provide a better benchmark for theoretical modelling to meet. If the discrepancy between the sunspot number and other solar indices continues, it may be that the SSN, for a while, is no longer a good measure of solar activity and forecasters (and users) may be forced to rely on other indices for operational use. This does not mean that we should stop deriving the sunspot number from the usual visual observations; on the contrary, the evolution of the SSN must be followed closely in order to provide a continuing basis for assessing the historical record.

References

- Berggren, A-M., J. Beer, G. Possnert, *et al.* (2009), A 600-year annual ^{10}Be record from the NGRIP ice core, Greenland, *Geophys. Res. Lett.*, 36, L11801, doi:10.1029/2009GL038004
- Clette, F. and L. Lefèvre (2012), Are the sunspots really vanishing? Anomalies in solar cycle 23 and implications for long-term models and proxies *J. Space Weather Space Clim.*, 2, A06
- Foukal, P. and J. Eddy (2007), Did the Sun's Prairie Ever Stop Burning? *Sol. Phys.*, 245, 247, doi:10.1007/s11207-007-9057-8
- Hoyt, D. V., K. H. Schatten, and E. Nesme-Ribes (1994), The one hundredth year of Rudolf Wolf's death: Do we have the correct reconstruction of solar activity? *Geophys. Res. Lett.*, 21 (18), 2067, doi:10.1029/94GL01698
- Jiang, J., P. Chatterjee, and A. R. Choudhuri (2007), Solar activity forecast with a dynamo model, *Mon. Not. Roy. Astron. Soc.*, 381, 1527, doi:10.1111/j.1365-2966.2007.12267.x
- Livingston, W., M. J. Penn, and L. Svalgaard (2012), Decreasing Sunspot Magnetic Fields Explain Unique 10.7 cm Radio Flux, *Ap. J.*, 757(1), L8, doi:10.1088/2041-8205/757/1/L8

Lockwood, M. and M. J. Owens (2011), Centennial changes in the heliospheric magnetic field and open solar flux: The consensus view from geomagnetic data and cosmogenic isotopes and its implications, *J. Geophys. Res.*, 116, A04109, doi:10.1029/2010JA016220

NRC: National Research Council (2008), *Severe Space Weather Events--Understanding Societal and Economic Impacts: A Workshop Report*. Washington, DC, The National Academies Press

Owens, M. J., I. Usoskin, and M. Lockwood (2012), Heliospheric modulation of galactic cosmic rays during grand solar minima: Past and future variations, *Geophys. Res. Lett.*, 39, L19102, doi:10.1029/2012GL053151

Pesnell, W. D. (2012), Solar Cycle Predictions (Invited Review), *Sol. Phys.*, The Sun 360, doi:10.1007/s11207-012-9997-5

Schaefer, B. (1993), Visibility of Sunspots, *Ap. J.*, 411, 909

Schatten, K. H. (2005), Fair space weather for solar cycle 24, *Geophys. Res. Lett.*, 32, L21106, doi:10.1029/2005GL024363

Schrijver, C. J., W. C. Livingston, T. N. Woods, and R. A. Mewaldt (2011), The minimal solar activity in 2008–2009 and its implications for long-term climate modeling, *Geophys. Res. Lett.*, 38, L06701, doi:10.1029/2011GL046658

Spruit, H. C. (1977), Heat flow near obstacles in the solar convection zone, *Sol. Phys.*, 55, 3, doi:10.1007/BF00150871

Steinhilber, F., J. A. Abreu, J. Beer, and K. G. McCracken (2010), Interplanetary magnetic field during the past 9300 years inferred from cosmogenic radionuclides, *J. Geophys. Res.*, 115(A1), A01104, doi:10.1029/2009JA014193

Steinhilber, F., J. A. Abreu, J. Beer, *et al.* (2012), 9,400 years of cosmic radiation and solar activity from ice cores and tree rings, *Proc. Nat. Acad. Sci.*, 109(16), 5967

Svalgaard, L., M. Lockwood, and J. Beer (2011), Long-term reconstruction of Solar and Solar Wind Parameters, http://www.leif.org/research/Svalgaard_ISSI_Proposal_Base.pdf

Svalgaard, L. and E. W. Cliver (2005), The IDV index: Its derivation and use in inferring long-term variations of the interplanetary magnetic field strength, *J. Geophys. Res.*, 110, A12103, doi:10.1029/2005JA011203

Svalgaard, L., E. W. Cliver, and Y. Kamide (2005), Sunspot cycle 24: Smallest in a hundred years? *Geophys. Res. Lett.*, 32, L01104, doi:10.1029/2004GL021664, 2005

Svalgaard, L. and E. W. Cliver (2007), Interhourly variability index of geomagnetic activity and its use in deriving the long-term variation of solar wind speed, *J. Geophys. Res.*, 112, A10111, doi:10.1029/2007JA012437

Svalgaard, L. and E. W. Cliver (2010), Heliospheric magnetic field 1835–2009, *J. Geophys. Res.*, 115, A09111, doi:10.1029/2009JA015069

Waldmeier, M. (1968), Die Beziehung zwischen der Sonnenflecken-relativzahl und der Gruppenzahl, *Astr. Mitteil. Eidgen. Sternw Zürich*, Nr. 285

Wolf, J. R. (1852), Neue Untersuchungen über der Periode der Sonnenflecken und ihre Bedeutung, *Mittheil. der naturforsch. Gesellschaft in Bern*, Nr. 245

Acknowledgements

The author thanks Ed Cliver for serving as a long-suffering sounding board for my wilder ideas. I also thank Phil Scherrer at Stanford for support and I have benefited from participation in the International Teams in Space Science [ISSI] Workshop on ‘Long-term reconstruction of Solar and Solar Wind Parameters’ and from the Sunspot Number Workshops <http://ssnworkshop.wikia.com/wiki/Home>

SWSC130003 – Response to Referee reports

My responses are in red
The revised manuscript has changes highlighted in yellow

Detailed scientific reports from Referee 1:

Report on L. Svalgaard paper:

This is an interesting paper that I can recommend for publication after the following revisions:

p 2; para 2, line 1: ...first assembled by Rudolf Wolf...

done

p 2; para 4: the statement that the ratio of spots to groups is constant and the later finding that the GSN/ SSN ratio was smaller prior to 1885 is confusing. This needs some clarification.

The constancy was mentioned by H&S [in the Data Collection section of their 1994 paper cited] as motivation for creating the GSN. The constant ratio is not between GSN and SSN, but expresses the number of spots per group [about 10]. I changed 'which' to 'that' to clarify that it is the Wolf SSN that does not support the idea of a Modern Grand Minimum

Fig 1: I can (sort of) see the changes reported in R_g/R_z in Fig 1, but they would probably be clearer if a smooth were run through those rather noisy data.

I don't know how to do this in a non-objectionably way. I have at times in the past drawn three lines through the points to show the discrepancy, but it invariably attracts criticisms, so I decide to let the reader draw his own conclusions.

p 3; para 2, line 7: ...auxiliary site....

done

p 4; last para: Is there some explanation why the H&S result was so different ?

Given in the text: Wolf used the small telescope since the mid 1860s and could not see the smallest groups [which make up half of all groups]. H&S assumed a constant k-factor for Wolf throughout, thus mixing in the observations with the larger telescope. An additional factor is that the RGO data is not homogeneous. Space constraints prevented me from elaborating on this. A fuller explanation can be found here: slides 38-48 of

<http://www.leif.org/research/Reconstruction%20of%20Sunspot%20Number.pdf>

I wonder if the Journal would permit as Supplementary Material the various links that I cite in the Response.

p 5; para 1, last sentence: I don't understand this statement. Please explain better or remove.

I assume that you mean this statement: "if people can accept the current series without adjustments, then they might also accept that we, for now, assume that there are no further adjustments warranted for the early data."

I'll try to explain: there are people who defend the GSN (often on the grounds that it provides a convenient secular change which may support the notion that a secular increase of solar activity is the cause of global warming since the 17th century). So they accept that GSN 1610-1885 is correct as it stands (and also of course since

1885). Apart from the discontinuity in 1885 we find that the GSN is quite homogeneous since 1800. As we think we have an explanation for the discontinuity that is unique to its time, we assume for now that there are no further discontinuities going back in time. This assumption is the topic for the next SSN workshop. But at this point we go with the smallest number of discontinuities [one]. If one believes that the GSN is good as it stands then one cannot reasonably argue that there might be more discontinuities.

Fig 7: The agreement in this plot after 1947 seems almost too good to be true. Is it? Also, it would be interesting to say something about how the reconstructed F10.7 prior to 1947 compares with the Ca K variation now available back to 1915 from analyses of the MWO, SPO and Kodaikanal data.

The agreement is real and is that good. Slides 6-12 of <http://www.leif.org/research/Rudolf%20Wolf%20Was%20Right.pdf> explains why. Again due to space constraints it was not possible to go into details. Same with the Ca K variation, see slide 24 of <http://www.leif.org/research/HAO-Seminar,%20How%20Well%20Do%20We%20Know%20the%20SSN.pdf> It is actually possible to see the Waldmeier discontinuity ~1945 comparing with the Ca K data [and with several other indices, e.g. sunspot areas], but again space limitations...

p 6; para 1, last sentence: This statement seems to require that plage area and SSN track accurately in all cycles. Do they? Also, the data on diurnal variation shown in Fig 7 only extend back to 1885, so can they really be used to test the circa 1885 change in GSN/SSN?

Actually they do not. As shown in Figure 12 the tracking has broken down since ~1990. There is no evidence that this happened in the last 300 years before that [I personally believe that it also broke down during the Maunder Minimum explaining why there were few spots <http://www.leif.org/research/SSN/Svalgaard12.pdf> Figure 7 was to show that the diurnal variation is a very accurate sensor (correlation coeff. 0.99) of solar activity [at least expressed by F10.7]. Data on the diurnal variation actually exists back to 1722 [with some gaps before 1781] and fully support the adjustment of GSN in 1885 [see some of the links I gave, and also slide 15 of <http://www.leif.org/research/Solar-Petaluma--How%20Well%20Do%20We%20Know%20the%20SSN.pdf>].

p 6; Fig. 8: What is the IDV?

IDV is a geomagnetic index developed in Svalgaard & Cliver [2005, 2010] as mentioned. It has the interesting property only to be sensitive to the solar wind magnetic field and not to its speed [allowing us to determine the magnetic field].

p 6; para 3: Should clarify to what extent different aspects of the solar field reflected in e.g. number of spots, in configuration and strength of background fields in polar and separately, in lower latitudes, are related and also to a large extent separate.

I would love to go into this subject, but space limitations do not permit me. Should the editor lift the limits and grant me the necessary extra space I'll include a section on the solar dynamo and the spatio-temporal evolution of the field.

p 6; last para, first sentence: ... solar system by mechanisms that include supernova... (the source mechanism is still somewhat controversial).

I have added a phrase alluding to the necessary acceleration of the cosmic rays but since the precise mechanism is not central to the paper I cannot delve any further into this fascinating subject

p 7: it is worth pointing out that the radionuclide record, although useful for studying periods of moderate and low solar activity, is less well suited to setting upper limits on such activity in the past. The reason is that a sun more active than anything we have experienced recently would produce an SEP flux that would begin to compensate the decrease in GCR caused by heliomagnetic shielding. So at some point the relation between cosmic ray flux and activity would change sign from negative (now), to positive. Calculations by Usoskin et al a few years ago confirmed that such a SEP effect can be significant for Be 10, even in the largest flares presently recorded. I added a [too] short phrase on this “including the effect of solar cosmic rays”.

p 9, para 1; line 3: should read: Clette & Lefevre, 2012
done

General comment: The figures (except for 1, 4,7, 8, 11, 12) are too small. Perhaps there is space to enlarge them; they are all interesting so it would be a pity to remove any.

As I understand it, the PDF version is not the version that will appear in the Journal. I have larger versions of the Figures and can supply these for the final stage of production.

I thank the referee for insightful comments.

Detailed scientific reports from Referee 2:

The paper is an excellent summary of the latest efforts to homogenize the SSN record, and touches on the important applications of this effort with regard to spaceweather and other terrestrial effects. The paper will have a broad audience of readers and will be of much interest.

I have two small comments which could help clarify two parts of the paper:

(1) Figure 2 is confusing in that the left-most panel seems to show a MWO drawing which is not mentioned in the caption. Please remove this part of the figure, or revise the caption to assign the Locarno drawing to the (center) of the figure.

I added ‘the insert shows the same spot from Mt. Wilson’. This is important because the plot on the right show MWO drawings, so one must be sure that they are compatible with Locarno drawings.

(2) Can the author mention a physical reason as to why the heliospheric magnetic field would have a floor at 4 nT, as discussed around Figure 8?

The floor is an observed quantity. Its cause is a research issue and no simple explanation has been put forward. Perhaps a hand-waving argument will do [although there is no space to put it in the paper and it is not critical for the message of the paper: The magnetic field on the Sun drives the solar wind, and as there is a minimum value [250 km/s escape velocity at the source surface] of the solar wind speed, otherwise the particles could not escape, it may not unreasonable to postulate that if there is a solar wind at all there will also be a minimum magnetic field.