No Doubling of the Sun's Coronal Magnetic Field during the Last 100 years

L. Svalgaard *, E. W. Cliver †, and P. Le Sager ‡.

* *Easy Toolkit, Inc., Houston, TX 77055, USA.* † *Space Vehicles Directorate, Air Force Research Laboratory, Hanscom AFB, MA 01731, USA.* ‡ *Prairie View A&M University, Prairie View, TX 77446, USA.*

Lockwood et al. and others [1-4] have suggested that the total magnetic field leaving the sun has more than doubled since 1900. The claim is based on a similar increase of the *aa* index [5] that measures the irregular geomagnetic variations. An explicit assumption was that the *aa* index has had constant calibration over time. We report here several independent lines of evidence that strongly suggest that the assumption of constant calibration of the *aa* index is not valid and that therefore the *aa* index cannot be used as a basis for inferring a secular trend in geomagnetic activity. There does not seem to be a long-term trend in geomagnetic activity and thus no similar trend in the magnetic field of the sun.

The *aa* index seeks to measure solar wind induced activity. The observer attempts to distinguish, for each three-hour Universal Time interval, between fluctuations caused by solar electromagnetic radiation and fluctuations caused by the solar wind and to assign a code (*K* index) to the interval characterising the magnitude of the latter [6]. This involves the ability of the observer to correctly distinguish between the two classes of fluctuations and the availability of appropriate conversion tables to reduce the index to a standard station. As stations change from time to time, new conversion tables have to be drawn up and intercalibrated with the previous tables. The entire process cannot easily be duplicated and the quality of the index values is difficult to gauge [7].

For studies of the long-term behaviour of geomagnetic activity we do not need a time resolution as fine as three hours. The fluctuations caused by solar EUV radiation are mainly confined to local day hours and are effectively absent deep into the night. A new index [8] of solar wind type activity uses only data from six hours centred on local midnight. For a given station, the average difference (taken without regard to sign) between each hourly value and the next of the horizontal component of the geomagnetic field for the six hours around local midnight is a measure (the *IHV* index) of the average ***i****nter-****h****our* ***v****ariability*. The average *IHV* index for each day can be calculated using several stations going back to the beginning of the twentieth century [9]. Mean values of this measure correlate very well with yearly means of the *am* index (correlation coefficient 0.98). The *am* index [10] is derived (since 1959) from a network of stations well distributed in longitude in both hemispheres and is found to correlate closely with solar wind properties [2]. Using data since 1959, we find [8] that we can reconstruct with high confidence (R^2 = 0.96) yearly average *aa* values from yearly average *IHV* values using the relationship *aa* = 5 *IHV*. The relevance of this reconstruction is that the *IHV* index is completely objective, does not require complicated conversion procedures and tables, and can be readily calculated by anybody.

Purely geometrical effects [11] create a 22-year cycle in geomagnetic activity. We largely eliminate their influence by only considering data from the four solstitial months. Figure 1 shows reconstructed (using Cheltenham/Fredericksburg) and observed *aa* indices for the interval 1901-2000 for the solstitial months of each year. It is clear that the reconstructed and observed *aa* indices match very closely back until 1957. Before 1957, the observed *aa* index is lower than the reconstructed value, and the difference becomes even greater before 1938, where it reaches 5-10 nT. The years 1957 and 1938 correspond to times of significant instrumental changes as explained below. The reconstructed *aa* is free from observer bias and does not rely on conversion tables. A similar reconstruction can be made of the standard planetary measure of geomagnetic activity, the *ap* index [10] going back to 1932. Reconstructing the *ap* index, we find that there is no discrepancy between reconstructed and observed values before 1957, and no long-term trend during the last two-thirds of the last 100 years.

We have performed 'storm-time' superposed epoch analysis of the horizontal component, H, at the Cheltenham/Fredericksburg magnetic observatory using 'sudden commencements' [12] (SSCs) as key times for three groups of years: 1901-37, 1938-66, and 1967-2000. No significant difference in the response to an SSC between the three groups was found (Figure 2a). The response of the *aa* index (Figure 2b) is also similar for the three groups, except that the *aa* values for the 1901-37 group are ~7 nT lower. This is the second piece of evidence that *aa* is systematically too low before 1938.

The polar cap (PC) index [13] was developed to quantify magnetic disturbances in the polar cap caused by electric fields generated by the "merging" electric field determined by the product of the solar wind speed (V) and the interplanetary magnetic field (B) modulated by a function of the angle between the interplanetary magnetic field and the Earth's magnetic field at the sunward front of the magnetosphere. We average over a full year and take the diameter (*Range*) of the nearly circular diurnal variation of the end-point of the horizontal component vector as a proxy for the merging electric field. For the interval 1965-2000 where good coverage of interplanetary properties is available, our proxy predicts the observed yearly averages of BV from $BV_{\text{pred.}}$ (mV/m) = 0.0248 $Range_{\text{THL}}$ (nT) with correlation coefficient 0.924 from the yearly average ranges measured at the (near-pole) Thule geomagnetic observatory. Figure 3 shows the polar cap variation ranges measured at Thule from 1947, at Scott Base from 1957, and at Godhavn from 1926. There is a clear solar cycle dependence related to a similar dependence of the interplanetary magnetic field and to a somewhat enhanced conductivity in the polar cap ionosphere at sunspot maximum, but apart from that, no long-term trend is apparent of the ranges and thus also not of the product BV. This conclusion is bolstered by the very similar variations measured in 1904 at Gjøahavn [14] and in 1902-03 at Winter Quarters [15] close to the present-day magnetic observatory Scott Base (these stations are approximately geomagnetic conjugates to Godhavn).

The lowest value the *aa*-index can attain is 2 (nT). This is an admission of the fact that it is very hard to unambiguously identify variations as being caused by the solar wind when activity is very low (less than 5 nT). A low yearly average value of *aa* could be caused by many cases of *aa* = 2. Figure 4 shows the number of 2s per year. There is the clear expected "inverse" sunspot cycle variation. Relatively narrow peaks at sunspot minima stand out. Away from these peaks, the distribution seems to be bimodal with a count of about 400 before 1938 dropping to a count of about 200 thereafter with a further drop from 1981 through 1997. Clilverd et al.[16] investigated the possibilities of instrumental changes to the *aa*-index. They note that "Three instrumental effects have been identified where significant changes in quiet-time conditions can be seen *i.e.* 1938, 1980, and 1997". These changes occurred at Northern Hemisphere sites, while the Australian sites remained close to their usual values. At the beginning of 1938, the variometers that were installed at Abinger were replaced by a single magnetograph system. The new system had only about half the sensitivity of the old system. To scale a *K*-value of 0 (*aa* = 2) would require identifying variations less than 5 nT (only 1.1 mm on the new recordings) as *distinct* from the regular *Sq* variation. This is likely to be more difficult on the less-sensitive newer system, meaning that possibly fewer occurrences of *K* = 0 were scaled with the new system. In 1980, the presence of locally generated “noise” was noted, further diluting the number of low *K*-values, and in 1997, a new low-noise magnetometer was introduced at Hartland. The actual number [16] of *K* = 0 scalings at Hartland in 1997 was back up to 500, more typical of the years before 1938. The systematic removal of low *K*-values beginning in 1938 causes *aa* to be higher from 1938 on.

The distribution of values is a powerful tool[7] to check for discontinuities, changes of calibration, instruments, observers, or methods. The distribution of *aa*-values has several discontinuities occurring at a change of year (*e.g.* between 1956 and 1957) and (except for the short interval 1920-26 and the year 2000) all having an *upward* trend making the *aa*-index larger than it should be. As the total effect of the changes in calibration is only about 10%, the discontinuities by themselves do not explain the apparent doubling, but do point to problems with the conversion tables.

We have presented several lines of evidence that each casts doubt on the validity of the assumption of constant calibration of the *aa* index. For each, in isolation, there might be other explanations. Collectively, they offer strong evidence that there is no doubling of the sun's open magnetic flux during the last 100 years in agreement with direct measurements of the sun's magnetic flux[17, 18].

1. Lockwood, M., R. Stamper, & M. N. Wild, A Doubling of the Sun's Coronal Magnetic Field during the Last 100 Years, *Nature*, **399**, 437-439 (1999).

2. Svalgaard, L., Geomagnetic activity: Dependence on solar wind parameters, in *Coronal Holes and High Speed Wind Streams, ed. J. B. Zirker*, Colorado Associated University Press: Boulder, p 371 (1977).

3. Solanki, S. K., M. Schlüssler, & S. M. Fligge, Evolution of the Sun's large-scale magnetic field since the Maunder minimum, *Nature*, **408**, 445-447 (2000).

4. Cliver, E. W. & A. G. Ling, Secular change in geomagnetic indices and the solar open magnetic flux during the first half of the twentieth century, *J. Geophys. Res.*, **107**, no A10, 1303, SSH-11, doi:10.1029/2001JA000505 (2002).

5. Mayaud, P. N., The *aa* indices: a 100-year series characterising the magnetic activity, *J. Geophys. Res.*, **72**, 6870-6874 (1972).

6. Mayaud, P. N., *Derivation, Meaning, and Use of Geomagnetic Indices*, American Geophysical Union: Washington D.C. (1980).

7. Bubenik, D. M. & A. C. Fraser-Smith, Evidence for strong artificial components in the equivalent linear amplitude geomagnetic indices, *J. Geophys. Res.*, **82**, 2875-2878 (1977).

8. Svalgaard, L., E. W. Cliver, & P. Le Sager, IHV: A new geomagnetic index, *Adv. Space Res.* (in press) (2003).

9. World Data Centre for Geomagnetism, Kyoto, Japan, http://swdcwww.kugi.kyoto-u.ac.jp/index.html.

10. Menvielle, M., and A. Berthelier, The K-derived planetary indices: description and availability, *Rev. Geophys.*, **29**, 415-432 (1991).

11. Russell, C. T. & T. Mulligan, The 22-year Variation of Geomagnetic Activity: Implications for the Polar Magnetic Field of the Sun, *Geophys. Res. Lett.*, **22**, 3287-3288 (1995).

12. Chapman, S., Electric current systems of magnetic storms, *Terr. Magn. Atmos. Electr.*, **40**, 349-370 (1935).

13. Troshichev, O. A., R. Yu. Lukianova, V.O. Papitashvili, F. J. Rich, & O. Rasmussen, Polar cap index (PC) as a proxy for ionospheric electric field in the near-pole region, *Geophys. Res. Lett.*, *27*, 3809 (2000).

14. Steen, A. S., N. Russeltvedt, & K. F. Wasserfall, The Scientific Results of the Norwegian Arctic Expedition in the Gjøa, 1903-1906, Part II, Terrestrial Magnetism, in *Geofysiske Publikasjoner*, **7**, Oslo: Grøndahl & Søns Boktrykkeri (1933).

15. Chree, C., *Studies in Terrestrial Magnetism*, Macmillan: London, 206 pp (1912).

16. Clilverd, M. A., T. D. G. Clark, E. Clarke, H. Rishbeth, & T. Ulrich, The causes of long-term changes in the *aa* index, *J. Geophys. Res.*, **107**, no. A12, 1441, SSH-4, doi:10.1029/2001JA000501 (2002).

17. Arge, C.N., E. Hildner, V. J. Pizzo, & J. W. Harvey, Two solar cycles of nonincreasing magnetic flux, *J. Geophys. Res.*, **107**, no A10, 1319, doi:10.1029/2001JA000503 (2002).

18. Kotov, V. A. & I. V. Kotova, Does the solar magnetic field increase?, *Astron. Lett.*, **27**, 260-266 (2001).

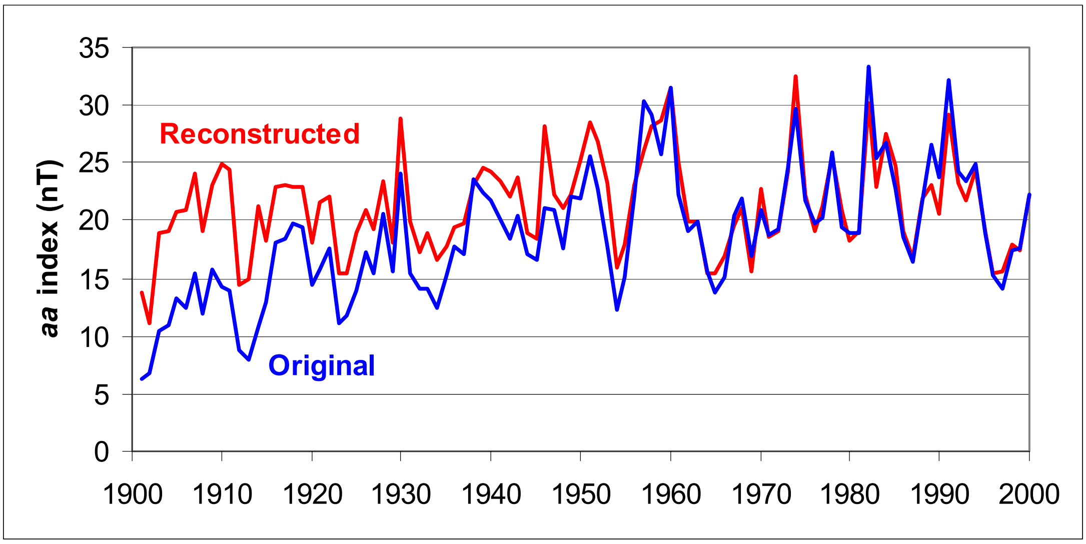


Figure 1. Observed (blue) and reconstructed (red) *aa* index for solstitial months for each year since 1901. The reconstruction is based on the IHV-index from Cheltenham/Fredericksburg.

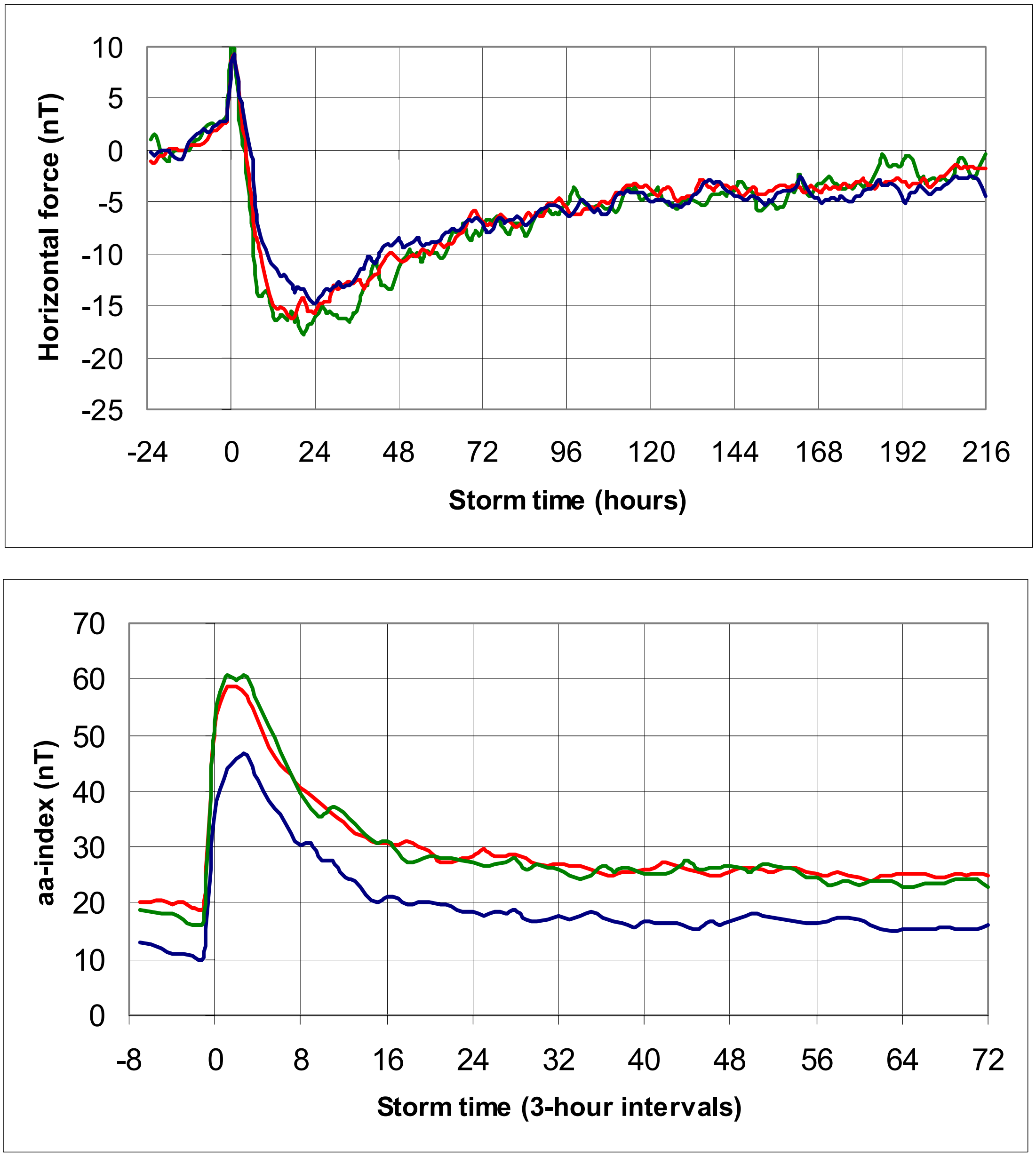


Figure 2 (a). Superposed epoch analysis of the horizontal component of the geomagnetic field at Cheltenham/Fredericksburg to the passage of SSCs. (b) Superposed epoch analysis of the response of the *aa* index. Storm time" starts at the time of a sudden commencement. The number of SSCs used for each interval are 824 for 1901-1937 (blue), 925 (green) for 1938-1966, and 879 (red) for 1967-2000.

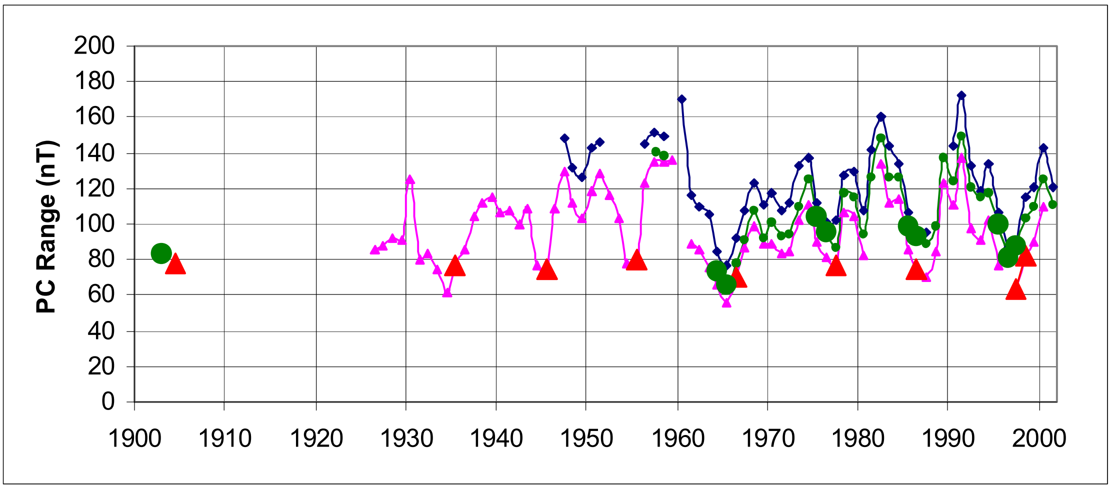


Figure 3. The range (nT) of the horizontal magnetic perturbation caused by the convection currents in the polar cap. Yearly means are shown for Thule (small blue diamonds), Godhavn (small pink triangles), and Scott Base (small green circles). The large red triangle for the year 1904 represents the range measured at Gjøahavn. Large red triangles from 1926 on show the range observed at Godhavn for years with average sunspot number approximately the same (42) as in 1904. The large green circle for 1902-03 represents the range measured at Winter Quarters. Large green circles from 1957 on show the range observed at Scott Base for years with average sunspot number approximately the same (14) as in 1902-03.

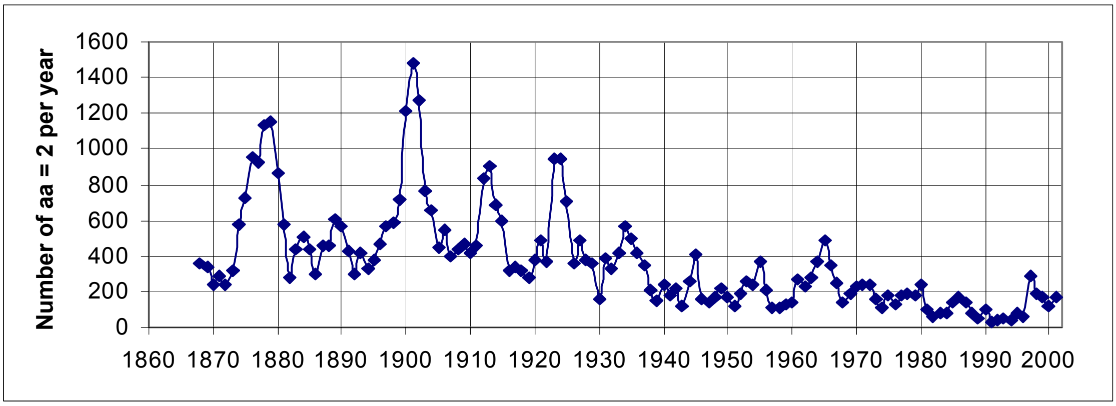


Figure 4. The number of three-hour intervals per year where the *aa* index had its lowest value (2 nT).

No Doubling of the Sun's Coronal Magnetic Field during the Last 100 years

Leif Svalgaard*, Edward W. Cliver†, and Philippe LeSager‡

* *Easy Toolkit, Inc., Houston, TX 77055, USA.*

† *Space Vehicles Directorate, Air Force Research Laboratory, Hanscom AFB, MA 01731, USA.*

‡ *Prairie View A&M University, Prairie View, TX 77446, USA.*

Lockwood *et al.* and others [1-4] have suggested that the total magnetic field leaving the sun has more than doubled since 1900. Observations[5-6] by polar explorers Amundsen and Scott early in the 20th century of the geomagnetic response to the solar wind do not support such a claim. This finding is consistent with a recent recalibration[7] of the geomagnetic *aa*-index[8] (on which the claim was based) that shows that the index is systematically too low before 1957 leading to the spurious secular trend.

The sunspot number, waxing and waning in the 11-year solar cycle, is a measure of the sun's magnetic field. The paper by Lockwood *et al.*[1] shows how the solar cycle variation of the inferred magnetic flux is superposed on a 100-year increasing trend in a "general" background flux[9]. The increase, if real, is significant, in fact, greater than the solar cycle variations. Strictly speaking, the "open solar magnetic flux" referred to by Lockwood *et al.* is not based on measurements of solar magnetic fields (direct measurements since 1974 show no increasing trend[10]), but is inferred from solar wind properties at the earth (inferred from geomagnetic activity), and then extrapolated to the entire sun. Lockwood *et al.* essentially infer the magnitude, B, of the interplanetary magnetic field and the speed, V, of the solar wind plasma. The inferred values of B and of V have risen since 1900 by factors of 1.96 and 1.14, respectively. Their product, V

times B, a measure of the interplanetary electric field, has thus risen by a factor of 2.23. The implication of Lockwood *et al.*[1] is that VB at the beginning of the 20th century should have been less than half of what it is now.

The horizontal component of the geomagnetic field measured within the polar regions has a particularly simple average diurnal variation: the end point of the component vector describes a circle with a diameter (the "range", E) of typically 100 nT. E is controlled by season (ionospheric conductivity) and by the interplanetary electric field VB. Averaging over a full year eliminates the dependence on conductivity and yearly average ranges have a strong linear dependence, $VB = k\ E$, (correlation coefficient 0.95) on VB as measured by spacecraft. This relationship holds for any station within the polar caps with only a slight variation of k. We have determined k for three polar cap stations: Thule (data back to 1932, $k = 24.4$ for V in km/s, B and E in nT) and Godhavn (to 1926, $k = 31.3$) in the northern polar cap and Scott Base (to 1957, $k = 27.4$) in the southern polar cap. Figure 1 shows VB determined from these stations as well as observed by spacecraft. The agreement between these values is substantial.

The noted Norwegian explorer Roald Amundsen wintered over with his ship "Gjøa" at "Gjøahavn". Magnetic recordings began on November 1st, 1903 and continued through May 1905[5]. The National (British) Antarctic Expedition of 1901-04 under Robert F. Scott operated magnetographs at the Winter Quarters of the Expedition for nearly two full years (1902-03)[6]. The range of the diurnal variation has been determined for these two sets of observations. Magnetographs were in operation at Cape Evans, the base station of the British (Terra Nova) Antarctic Expedition during 1911 and 1912. Cape Evans and Winter Quarters are co-located with Scott Base. The range was also determined for this station. However, this value represents only a *lower limit* to the true range because disturbed days, where the trace was not complete, were excluded. These early determinations are also shown in Figure 1 using k-values derived from the modern data (using the *k*-value for Godhavn for Gjøahavn as well, as these two stations have nearly the same geomagnetic latitude). It seems clear from Figure 1 that *VB* does not show the predicted decrease by a factor of 2.23 back to 1900. *VB does* depend on the sunspot number, R_z, (because *B* does): $VB \approx 2000 + 10.6\ R_z$ and its envelope shows similar changes, but relaxes to about the same non-zero value for zero R_z.

Figure 2 shows several independent determinations of *VB*. The values calculated from *aa* show the change in the calibration of the *aa*-index.

1. Lockwood, M., R. Stamper, & M. N. Wild, *Nature*, **399**, 437-439 (1999).

2. Svalgaard, L., in *Coronal Holes and High Speed Wind Streams, ed. J. B. Zirker*, Colorado Associated University Press: Boulder, p 371 (1977).

3. Solanki, S. K., M. Schlüssler, & S. M. Fligge, *Nature*, **408**, 445-447 (2000).

4. Cliver, E. W. & A. G. Ling, *J. Geophys. Res.*, **107**, no A10, 1303, SSH-11, doi:10.1029/2001JA000505 (2002).

5. Steen, A. S., N. Russeltvedt, & K. F. Wasserfall, The Scientific Results of the Norwegian Arctic Expedition in the Gjøa, 1903-1906, Part II, Terrestrial Magnetism, in *Geofysiske Publikasjoner*, **7**, Oslo: Grøndahl & Søns Boktrykkeri (1933).

6. Chree, C., *Studies in Terrestrial Magnetism*, Macmillan: London, 206 pp (1912).

7. Svalgaard, L., E. W. Cliver, & P. Le Sager, IHV: *Adv. Space Res.* (in press) (2003).

8. Mayaud, P. N., *J. Geophys. Res.*, **72**, 6870-6874 (1972).

9. Parker, E. N., *Nature*, **399**, 416-417 (1999).

10. Arge, C.N., E. Hildner, V. J. Pizzo, & J. W. Harvey, *J. Geophys. Res.*, **107**, no A10, 1319, doi:10.1029/2001JA000503 (2002).

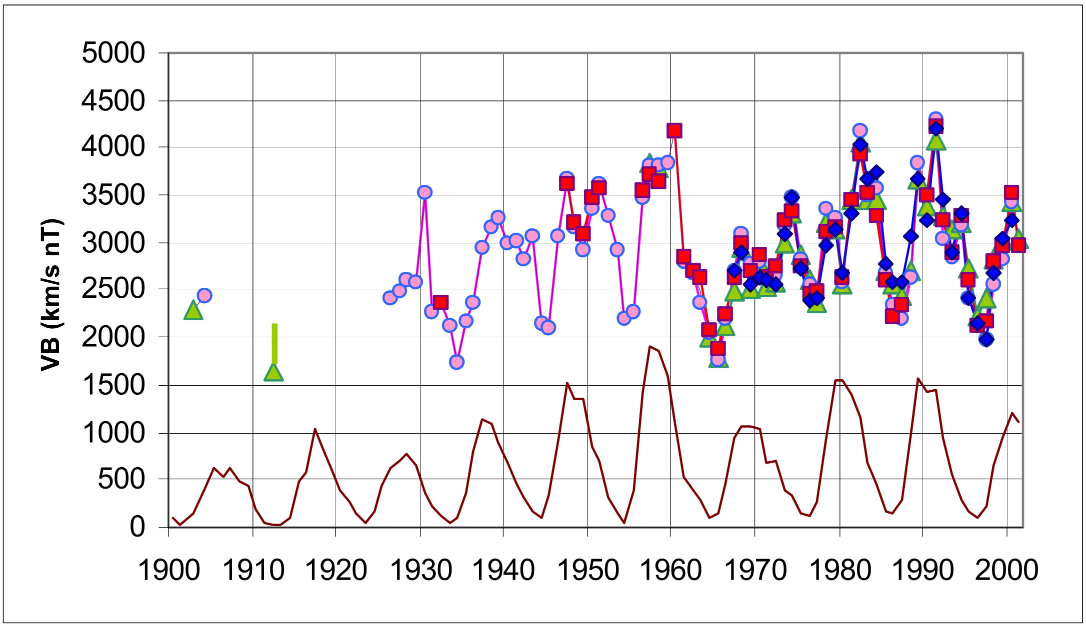


Figure 1. Yearly means of the product of solar wind speed *V* (km/s) and interplanetary magnetic field magnitude *B* (nT) measured by spacecraft (blue diamonds), inferred from the polar cap current range using Thule (red squares), Godhavn (pink circles), and Scott Base (green triangles). The pink circle at 1904 is for Gjøahavn. The green triangle at 1902-03 is for Winter Quarters, and the lower limit green triangle for 1911-12 is for Cape Evans. When only one symbol is shown for a given year it often simply means that it literally falls on top of the other symbols for that year. The lower curve shows the variation with time of the yearly smoothed sunspot number.

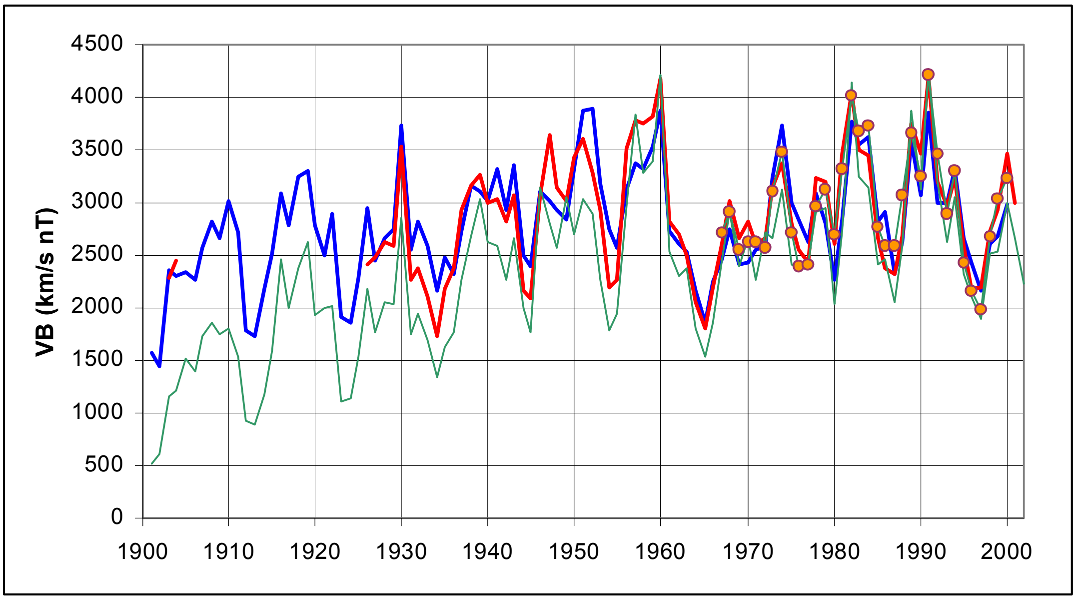


Figure 2. Yearly means of the product *VB* derived from several sources. Light green line is computed from the *aa*-index using the formulae in Lockwood *et al.*[1]. Blue line is 600 km/s times the *IHV*-index computed (and intercalibrated) from several stations following the prescription in Svalgaard *et al.*[7]. Multiplying by 600 km/s normalises the data to have the same mean (since 1966) as *VB* measured by spacecraft. Red line is computed as the mean of *VB* calculated for all stations used in the present paper, and orange circles show *VB* calculated from yearly means of *V* and *B* measured by spacecraft. All determinations essentially agree, except the one calculated from *aa*, which going back in time begins to deviate around 1957 and then deviates progressively more for earlier and earlier years. We interpret this discrepancy as a sign of progressive calibration errors in the *aa*-index.

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to L.S. (e-mail: nature@leif.org.).

Review of paper: No doubling of the Sun's Coronal Magnetic Field during the last 100 years

By: Svalgaard et al.

Submitted to Nature

January 2003

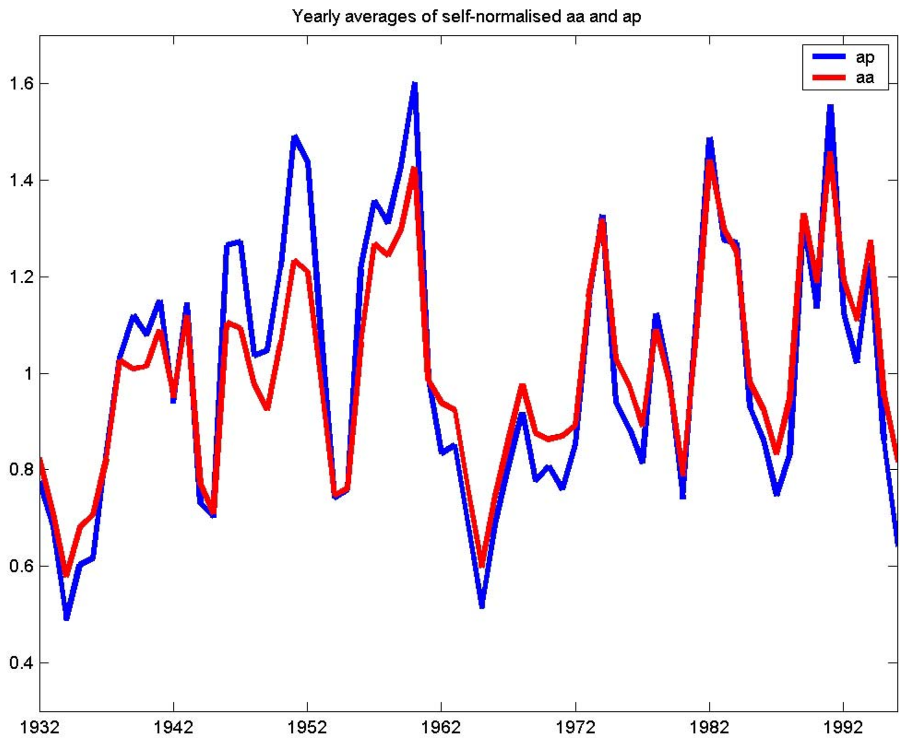
Because it comes from authors for whom I have a great deal of respect, I read this paper with great interest and in great detail. I am afraid (and I am also rather surprised) to say that I found the paper poor and extraordinarily one-sided in its discussion. I am also surprised because one of the authors has recently published a paper which comes to precisely the opposite conclusion to the present paper (Cliver and Ling [2002] find that the changes in both the aa index drift and the open solar flux are real).

Both of us (L.S. and E.C) have reached the opposite conclusion in the past. In fact, L.S. said in his 1977 Coronal Hole chapter: "Since the year 1900 the general level of geomagnetic activity has increased at least two-fold […] It seems more likely to the present author that it is mainly the magnetic field magnitude that has changed […]". It is only after a thorough and painstaking investigation that we very recently have come to our present standpoint, which we do not take lightly. The space limitation of Nature precludes incorporating the 100 pages and 100 figures on which our result is based. A document detailing that investigation is enclosed. We had instructed the editor from Nature to send it to Lockwood and the Referees, but it is evident that that did not happen.

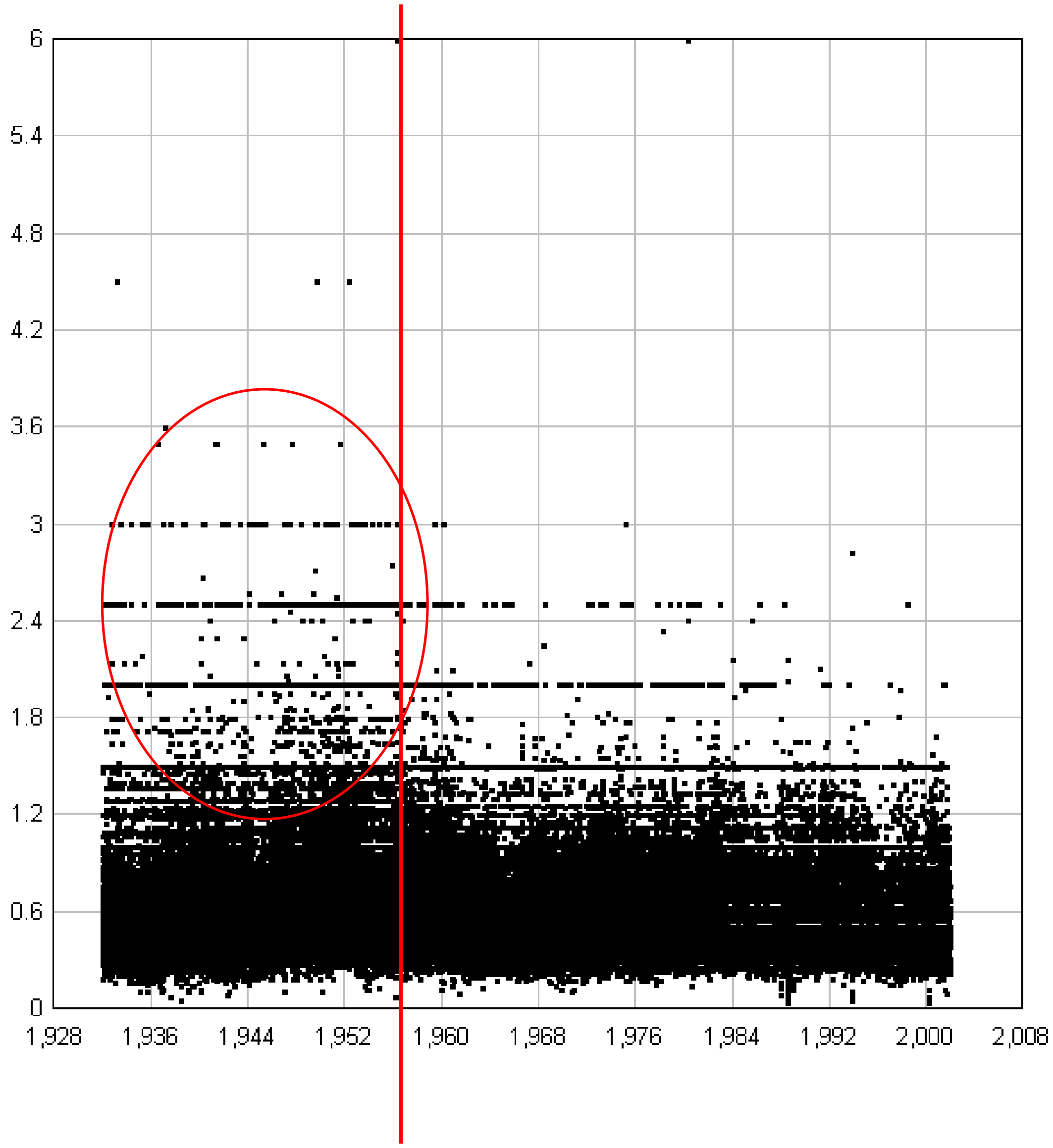
1. The problems start with the title. It is definitive statement that the drift of the open solar flux did not happen. This claim is, I will show below, simply not substantiated by the analysis presented and is contradicted by fully independent evidence that the authors ignore. However, even at face value, this title is not an accurate reflection of what this paper is about: were the authors' analysis to be valid (which I do not believe it to be), a title like "no long-term rise in geomagnetic activity" would be appropriate, not the one given. The paper is an attack on the validity of the aa geomagnetic index and says nothing at all about the calculation of open flux from aa. The authors' reconstructed curve shows some rise and the authors have not used it to compute a new variation of the open solar flux.

We are sure that if aa had not shown an apparent doubling, no paper with the title "A doubling of..." would have appeared in Nature. The claim hinges completely on aa, so if no doubling of aa happened, it seems clear that no new recalculation of the open solar flux would be needed to make the statement that there is "No doubling of ….". Strictly speaking both Lockwood et al's and our paper do not *prove* anything about the sun's flux, but are only supporting evidence for the supposition that there was/was not a doubling. In that light, the title of our paper has the same relation to the substance of the paper as the title to the Lockwood et al paper had to theirs. A compromise would be to place a question mark at the end of our title.

2. The aa index was published, clearly showing the rise in the first half of the last century, in 1972 [Mayaud, 1972]. To put the present, belated, criticism of the aa index into context, the paper attacks the basis of the way that the aa index is compiled – via the use of the range of the horizontal component variation in 3-hour windows and the generation of "k indices". Thus it also attacks the widely-used Ap and Kp geomagnetic indices which are compiled in a similar manner [Mayaud, 1980]. In relation to the long-term change in geomagnetic activity, it should be noted that Ap extends back to 1932 and it agrees very well with the aa index at all times - and not at all well with the authors' reconstruction between 1932 and about 1960 (when the latter diverges from aa). The first attached figure (in file aafigure1.jpg and aafigure1.eps) shows annual means of aa and Ap since the start of the latter in 1932. Both have been normalised by dividing by the mean for the entire interval.

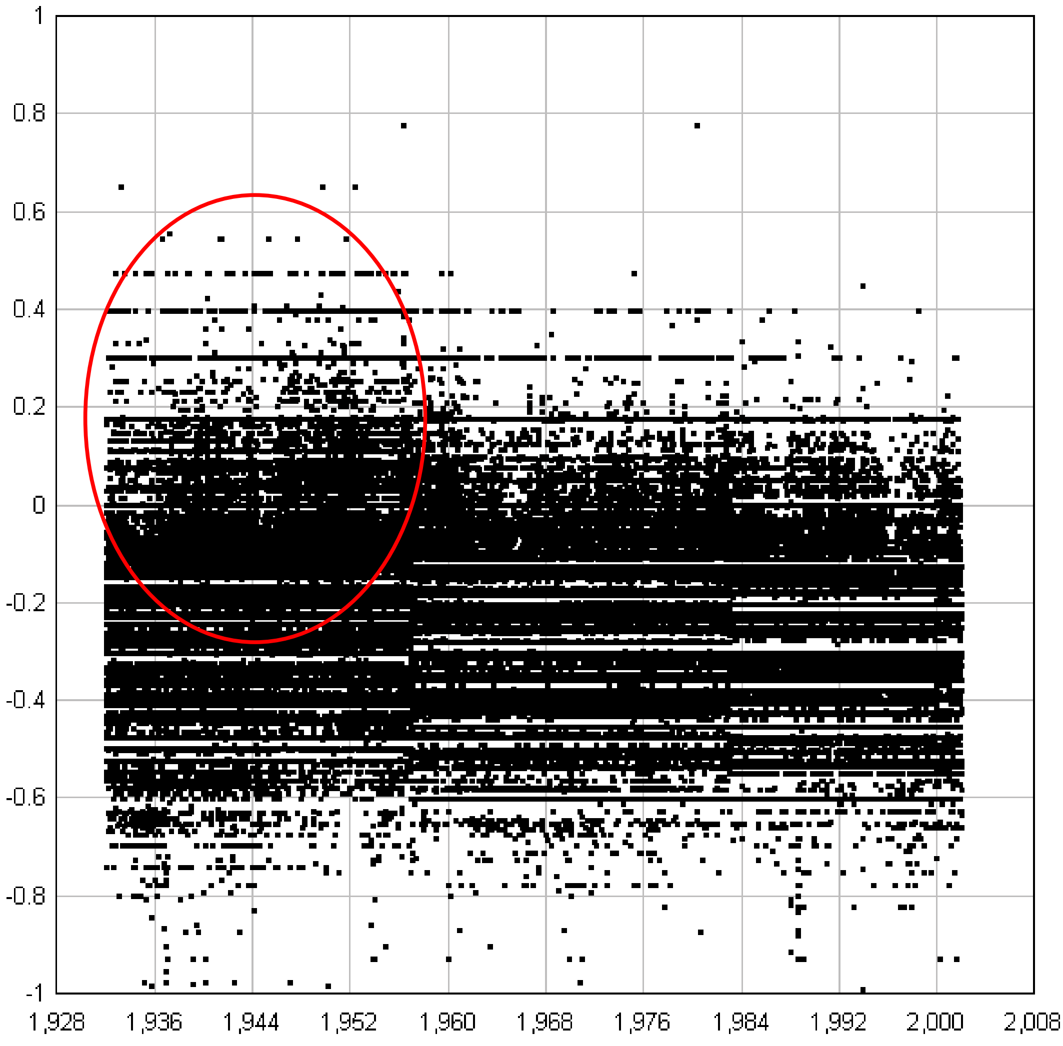


I guess that the eye cannot see what the mind doesn't want. Just looking at Figure 1 it is evident that before 1957, the blue (ap) line is generally **above** the red (aa), while after 1956 the blue line is **below** the red, simply meaning that either aa or ap has drifted: either the pre-1957 ap are too high or the pre-1957 aa are too low. Our analysis showed that it is aa that is too low. We did not compare aa with ap as we wanted to have *independent* verification, but if we play the comparison game we get again a break after 1956. Here is a plot of the ratio between every single ap and aa value since 1932:



The red line is drawn between 1956 and 1957. It separates regions with different ap/aa distributions. To the left of the line there are considerably more values above ap/aa =1.2 than to the right of the line, especially above 2.0.

Since the distribution is very skewed it is often a good idea to plot the distribution of the logarithm instead (discarding the very few zero values). This is our next Figure:



Here there is even no need for a red line: the discontinuity between 1956 and 1957 is obvious. Before the discontinuity, the ratio ap/aa is too high; that is: aa is too low. The horizontal “white lines” are due to the quantization of aa.

It can be seen that the drift in Ap during 1932-1960 is present and very similar to that in aa. In my view, the claim that Mayaud’s procedures for generating aa and Ap are flawed is much, much more contentious than what is the authors’ main rationale for this paper – namely the change, or lack of it, in the open flux in the Sun. Those procedures are laid down and justified in Mayaud’s seminal book [Mayaud, 1980] in much more detail than the lightweight criticism made of them in these papers. Mayaud’s work has been accepted by the community for over 30 years: the number of papers in solar-terrestrial physics based on Mayaud’s work is truly vast – it seems very strange to me to write this criticism after 30 years and then aim it at just one of those papers.

Well, we didn’t know about this problem 30 years ago. In a sense, our poor paper *is* aimed at all the other papers, especially the ones using aa values as precursors to predict sunspot cycle maxima. The Lockwood et al. 1999 Nature paper is a proper and specific aim because their claim of a doubling of the sun’s field has been used by other scientists in the heated debate about climate change and therefore reaches far outside of the solar/geo physics field.

3. The reporting of other papers in this paper is extraordinarily one-sided. In particular, no mention is made whatsoever of the fact that cosmogenic isotopes and the earliest cosmic ray data DO support the doubling of the open flux – despite the fact that this is

in the literature. The 10Be isotope abundance data from ice sheets shows a downward drift that matches the upward drift in inferred open flux and which has an amplitude which is comparable to the present amplitude of the solar cycle variation [e.g. Beer, 2000] – consistent with the changes in both aa and the inferred open flux [Lockwood, 2001]. McCracken and McDonald [2001] also show that this drift is also seen in the earliest cosmic ray observations (mainly ionisation chamber data [Neher, 1953, Forbush, 1958]). The 14C isotope production rate curve is only usable up to the first atomic bomb explosions – but it too fits the open flux variation up to this time [Stuiver and Quay, 1980; Lockwood, 2002]. The 44Ti isotope found in meteorites also supports the long term increase on open flux over the past 150 years [Bonnino et al., 1998]. Note that the cosmogenic isotopes also tell us that the open flux changes of the past 150 years are typical – this is the normal behaviour of the Sun [Beer, 2000; O'Brien, 1979, Bard 1997].

The cosmogenic isotope connection is an important one and we are fully aware of it. It is however thrice indirect. The first indirection is from solar magnetic fields to cosmic rays. The second indirection is from cosmic rays to isotopes. The third indirection is from isotopes to deposition processes. As far as the first link is concerned, there are some studies (e.g. Stoshkov, Geomagnetism and Aeronomy, vol 42, No 6, pg 711, 2002) that purport to show that "The CR flux near the Earth's orbit [… and hence B …] from 1937 to the present has been constant to an accuracy of 5%". The extreme space limitation on a Letter to Nature precludes any discussion of pros and cons of the various viewpoints found in recent literature.

Another example of the one-sided citation is the use of the Arge et al [2002] paper which reports no change in open flux over the last 2 solar cycles – but the failure to mention the Wang and Sheeley [2002] paper in the very same issue of JGR that finds no contradiction with the open flux variation derived from aa [see also Lockwood, 2002; 2003].

Our inclusion of the Arge and Kotov papers is relevant because they refer to *direct measurements* of the sun's magnetic flux.

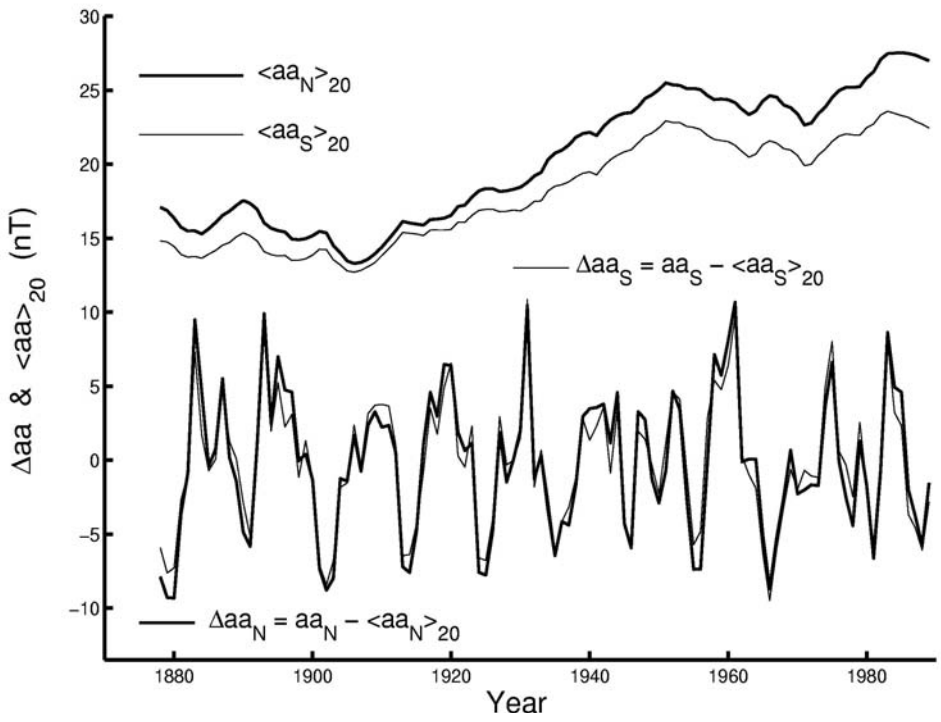
Other papers have pointed to changes in the Sun which follow those in the inferred open solar flux very closely. For example there are long-term changes in the peak and 11-year averaged sunspot number [Lockwood et al., 1999a; b; Cliver and Ling, 2002], the latitude of sunspots [Pulkkinen et al., 1997], the spread of sunspot latitudes (the width of the butterfly wings) [Foster and Lockwood, 2000], and the inferred solar irradiance [Lean, 2000]. Papers on the occurrence of low-latitude terrestrial aurorae [Pulkkinen et al., 2001] also show changes that follow the aa index variation very closely indeed and so very strongly support the long term change in both aa and the open flux. None of these supporting studies is even mentioned by the authors.

Since aa is a "**primary**" measurement its quality should not be gauged by "supporting" studies.

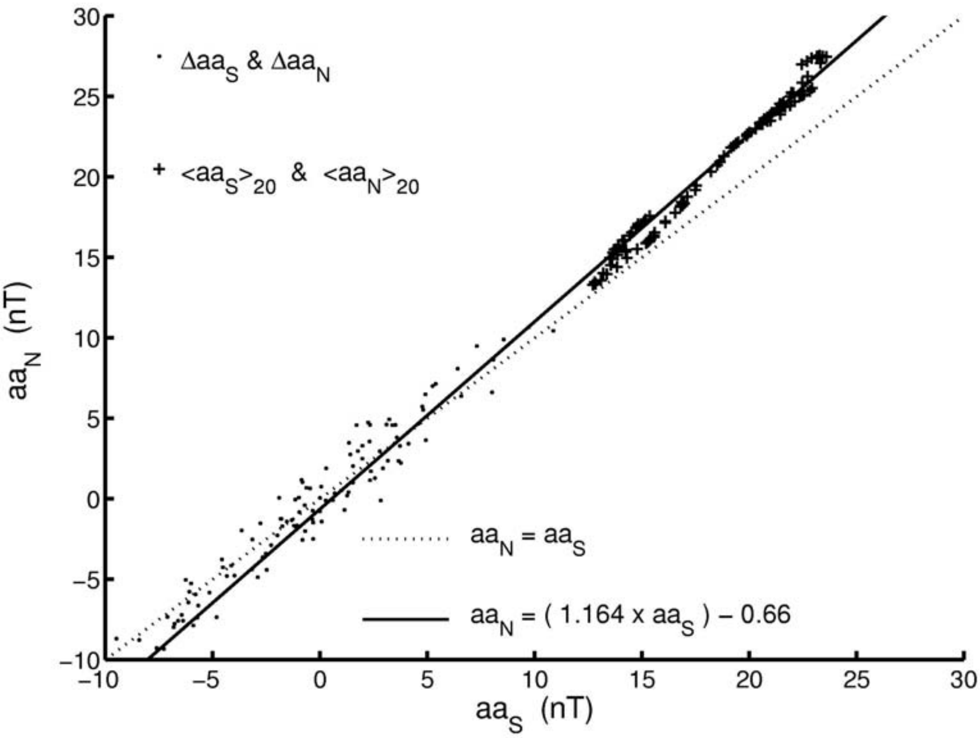
4. Let me now deal with the substance of the paper's claim. Nowhere does the paper mention the fact that aa is the average of a northern and southern aa indices (aa_N and aa_S), both of which show almost exactly the same long-term trend.

A re-scaling of a few years, e.g. 1902 and 1996 might be instructive. As our Figure 4 shows, more than half of all 3-hour intervals of 1902 was shown by the aa index to have no measurable activity at all. This should be easy to recheck. Such a check is in progress, but our paper was not really about *what* is wrong with the aa index, but simply that *something* is wrong. The detailed analysis belongs in a speciality journal, not in Nature

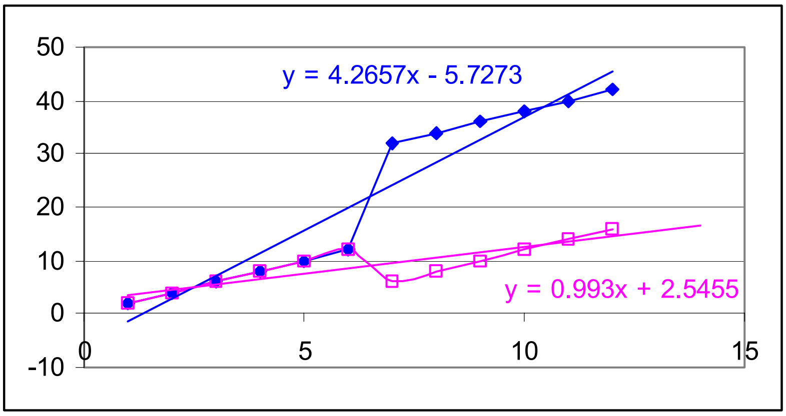
This is demonstrated by the plot given in the files aafigure2.eps and aafigure2.jpg [published by Lockwood, 2002 and 2003]. This compares the aa data from the northern and southern hemisphere stations. The bottom curves give the variations of annual means, de-trended by subtracting 20-year means. It can be seen that the two hemispheres give almost identical solar cycle variations (but there is a consistent scaling difference that makes the northern hemisphere data persistently greater by a factor of 1.164). The upper curves show the long-term drifts by giving the 20-year means. Both hemispheres show the same waveform of long-term drift (the scaling difference factor of 1.164 is again apparent).



To emphasise how consistent these two data sets are, and how constant the factor 1.164 is, the figure shows the data in the last figure as a scatter plot.



The plot is flawed. One cannot calculate a meaningful trend line when mixing <aa> and Δaa. The following plot illustrates the problem with two artificial data sets. The first six points are the deltas (identical for the two sets) and the last six points are the mean values (different for the two sets). Even if the two data sets have the same trend (slope) [which is the case for our artificial cases, but not a given for the aa data sets], the mean value can be anything:



It is mixing apples and oranges to fit a common line to both deltas and means.

In both hemispheres, 3 stations have been used to generate the data series, but the potential discontinuities in the data sequences took place at different times in the two date series. Yet aa_N and aa_S both show a gradual change - and the same gradual change in the two hemispheres [Lockwood, 2002]. The conclusion has to be that Mayaud's intercalibrations and methods are extremely good – except that his northern hemisphere scaling factor to convert the k-indices into aa are 16.4% too high for the northern hemisphere (but this factor is constant and has not drifted).

Simply measuring off their Figure 2 it is evident that the factor is *not* constant. It varies from 1.03 in (say) 1915 to 1.20 in 1990. In his formula on Figure 3 there is an offset (-0.66). Such an offset (like on my Figure above) can be used to make a common fit with a single "constant". The fact is that the ratio between aa_N and aa_S has drifted.

There is a lengthy discussion of possible discontinuities in the data from the northern hemisphere. It fails to mention that as well as overlap in the data series, these data have been intercalibrated by re-installing contemporary magnetometers, of to the original sites. Against this, it is astonishing to me that the authors fail to mention, let alone discuss, that there is a discontinuity in the data they use between the Fredricksberg and Cheltenham sites and that this has its main effect at exactly the time that their reconstructed indeed starts to diverge from aa!

Our long "justification paper" goes into details about how during 1956 the two stations operated simultaneously and that an intercalibration was indeed made. The correlation coefficient between the two stations was 0.994 (R^2 = 0.988) and the calibration factor was 0.97 or so close to unity as to make no difference (even though it was applied anyway).

In effect, the authors put an untested intercalibration between their two stations ahead of the exceptionally well-tested intercalibrations of two sets of data (aa_N and aa_S). The authors have effectively correlated their IHV from Fredericksburg/Cheltenham with aa from later years and done a simple extrapolation.

Several other stations were also used in our analysis (Niemegk, Tucson, De Bilt, Witteveen, Wingst, Toolangi, Canberra, Hartland, and Abinger). The results were the same as for Cheltenham and Fredericksburg.

Because aa (and open solar flux) peaked around 1986 this is of strictly limited validity. The open flux computation is theory-based and not a simple correlation such as the authors use.

Apart from the variation of the Earth's magnetic moment, the calculation boils down (see our long paper) to the following formula:

$$\boldsymbol{B\ \mathrm{nT} = B_0\ [aa\ /\ ((I \cdot aa^{5})^{0.263} + 72.6)]^{1.295}}$$

Where the single constant B_0 (= 46.85) subsumes the various other constants and calibration factors. This lays bare the numerology all such reconstructions really are. If aa does not increase by a factor of two or more, *B* will not either whether using a simple correlation or a complicated formula. The main problem (as Lockwood et al are keenly aware of) is the separation of the influence of the interplanetary magnetic field strength and solar wind speed. The above formula (derived from a possibly flawed aa-index) is not the last word and it seems premature to apply it to our reconstructed index.

5. The authors believe and state that their use of hourly means is more objective and reliable than the K-indices. This is a fallacy. Hourly means would have been scaled by

hand and would have been more subject to error than the K indices (maximum and minima, and thus range, being much more accurately defined by a manual scaler than a mean). Mayaud's extensive and thorough work laying down the rules for generating K indices means that they are very much more reliably homogeneous than hourly means. Uncertainties will then have been amplified when the authors take the difference between two successive means. Substorm expansion phases can last less than 10 min – so the authors' use of the difference between hourly means is therefore highly problematical. Range is much more meaningful as it picks out the maximum disturbance levels in each 3-hour interval.

As our long paper shows (and the upcoming - and cited - paper in Adv. Space. Res.) the IHV index correlates extremely well with the am-index (R^2 = 0.943 for yearly means). The usefulness of IHV hinges on this high correlation. We stated that IHV was objective (we never said "more reliable") and meant by that simply that the calculation we did can be checked by anybody using published data. The same cannot be said about the aa index.

Lockwood is mistaken on the accuracy of ranges versus means. To measure a maximum requires first to locate where the maximum is, then to measure the distance of that point from the baseline (which is a separate line on the magnetogram), the same for the minimum. To measure the hourly mean you repeat the above for many points within the hour, say for every five minutes. There is no difference in accuracy between measuring two such points (max and min) and 36 of them (every five minutes). The point he makes about amplifying uncertainties by subtracting successive means can be equally made when calculating the range = maximum - minimum. In, fact, it is the other way: by subtracting values you decrease the uncertainty because you remove any errors related to the baseline varying with time. Furthermore a mean over 12 values per hour and then over six such hours is a better determined quantity than just two range values over those hours. All of this is really so obvious that it does not deserve much discussion. But it misses the point completely, namely that the K index is NOT just the range. The K-index is the deviation from a time-varying quiet background (Sq+Lq+sfe), or more precisely from anything caused by solar EUV and not by the solar wind. This is the crucial point (as Bartels and Mayaud have tried to drive home - seemingly with little success). Deciding if a wiggle on the magnetogram is caused by EUV or by the solar wind is non-trivial and become harder the smaller the variation is and thus very dependent on the sensitivity of the magnetograph. I have myself scaled thousands (nay, tens of thousands) of hours both in calculating hourly means and scaling K-indices so I know in my bones what the issues are.

6. In the super-posed epoch analysis, the authors confuse ring current “storm” effects and auroral electroject “substorm” effects. The mid-latitude indices Ap, aa and Kp all have a major influence from substorm effects taking place at higher latitudes, as well as ring current effects and it is the former which introduces the IMF control that gives the computation of open flux. On the other hand, the authors isolate ring-current storms using the superposed epoch analysis based on the timing of storm sudden commencement. The relationship of sudden-commencement storms and substorms is complex and substorms will have contributed to aa at all times relative to the storm sudden commencements. Thus the authors superposed epoch analysis in no way shows

there is an instrumental effect lowering sensitivity in the earlier years, as they claim – it is equally consistent with the idea that substorm activity levels were genuinely lower in earlier years (consistent with the lower aa levels and caused by the open solar flux being lower).

Figure 97 of our long paper shows that (since 1959)

1) dividing the years into two groups with lower than average aa and higher than average aa, produces a similar difference in the Dst (proxy) response, and

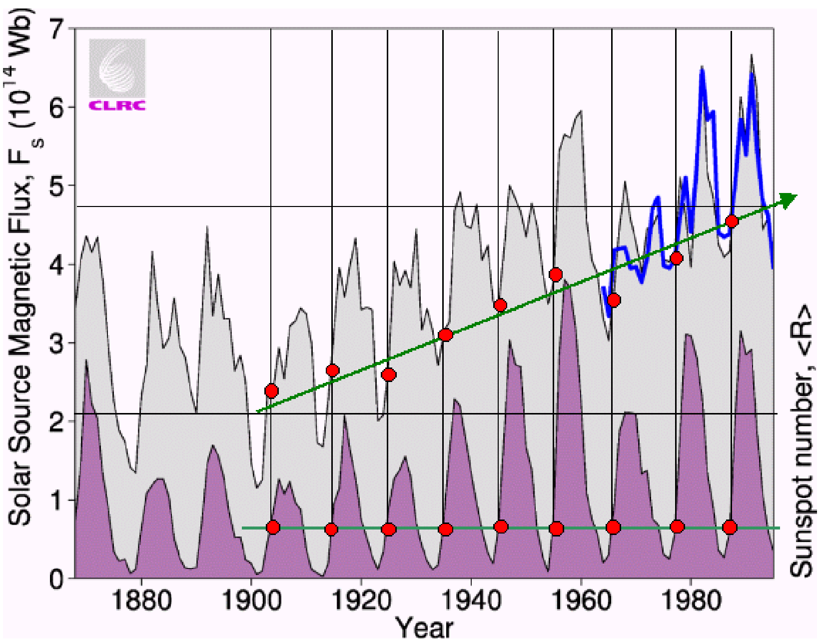
2) dividing the years into two groups with lower than average Dst-response and higher than average Dst-response, produces a similar difference in aa. If the quality of aa before 1959 were as good as afterwards, we would expect a similar behaviour for the early years. Instead we find that before 1938, aa is way too low. There is no confusion on our part.

7. The use of polar cap indices introduces a similar set of physics-based inter-relation and interpretation problems. As the authors state, these are indices of convection, or the DP2 current system, which is only part (and a smaller one at that) of the auroral electrojet and substorm currents, which are dominated by the DP1 current system. Furthermore the authors make no attempt to intercalibrate the various data sequences in their figure 3.

This is on purpose to keep things straightforward. It is enough that the various stations all show a “flat” response not compatible with a doubling of B. There is a strong relation between B times V and the polar cap potential. Our own analysis finds a correlation with $R^2 = 0.86$.

Even after all these important caveats, I cannot even agree with the authors' interpretation of their figure 3. I see more similarities with the aa variation (in particular there is an upward drift between 1928 and 1950 in the Godhavn data) than with their reconstruction in figure 1.

We had emphasised that the variation with time for constant sunspot number was "flat". There is, of course, a simple correlation with the sunspot number in the sense that higher sunspot number means higher B and thus higher polar cap potential. There seems to be general confusion on this point, with people talking about "the rise" from 1920 to 1960. We do not disagree with the notion that more sunspots mean more magnetic flux. What we saw in Lockwood's justifiably famous Figure:



was (as Parker also saw in that same number of Nature) that what was claimed was an increase of the "general" or background flux independent of the sunspot number. We have modified the Figure (above) as follows: In each cycle from #14 onwards we identified the time (during the rising phase to stay away from recurrent streams) when the sunspot number was 40, and marked it on the graph by a red dot. Then followed a vertical line upward to intersect the calculated flux curve and marked that point by a matching red dot. What we interpret as the meaning of "A doubling of the …" is the fact that the upper red dots so produced lie on an upwards rising line (the green arrow) that indeed corresponds to more than a doubling of the flux. In addition to that general rise (which would be of fundamental importance if true) there are several other "rises" over and above the general rise (three such have been drawn on the Figure). These rises (one of them being the "upward drift between 1928 and 1950" and the "drift in Ap during 1932-1960" mentioned by Lockwood above) are just variations of the size of sunspot cycles and are not of fundamental importance as the "general rise" would be.

In summary, were the authors' criticisms to be valid, this would be an important paper.

We agree with this assessment and also with the impact that it would have on many other papers as noted by Lockwood:

But its importance would lie in the fact that all studies of global geomagnetism from the last 30 years which used either Ap or aa (which would be most of them) would need re-assessment – this is more significant than any conclusion about the change in open solar flux. The change in the open flux is well supported by completely independent evidence from cosmogenic isotopes. The change in aa is well supported by fully independent observations the aurora and is reflected in many know changes in the Sun. The analysis of geomagnetic indices is not nearly as detailed nor as careful as the original work by Mayaud that it criticises. Arguments about inter-calibration problems in the northern hemisphere are manifestly invalid as the southern hemisphere aa data shows almost exactly the same behaviour.

Simple inspection of Lockwood's own Figure 3 shows that the ratio between aa_N and aa_S rises from 1.03 to 1.20 between 1915 and 1990.

In making their arguments, the authors confuse the various causes of geomagnetic activity (storms, substorms, DP1, DP2).

I do not regard this paper as worthy of publication.

References

Arge, C.N., E. Hildner, V.J. Pizzo, and J.W. Harvey, Two cycles of non-increasing magnetic flux, *J. Geophys. Res., 107(A10), paper* SSH16, 2002.

Bard, E., G.M. Raisbeck, F. Yiou, J. Jouzel, Solar modulation of cosmogenic nuclide production over the last millenium: comparison between ^{14}C and ^{10}Be records, *Earth and Planet. Sci. Lett., 150,* 453-462, 1997.

Beer, J., Long-term indirect indices of solar variability, *Space Sci. Rev., 94 (1/2),* 53-66, 2000.

Bonino, G., G. Cini Castognoli, N. Bhabdari and C, Taricco, Behavior of the heliosphere over prolonged solar quiet periods by ^{44}Ti measurements in meteorites, *Science, 270,* 1648-1650, 1998.

Cliver, E.W., and A.G. Ling, Secular change in geomagnetic indices and the open solar flux in the first half on the twentieth century, *J. Geophys. Res., 107(A10), paper* SSH11, 2002.

Forbush, S.E., Cosmic ray intensity variations during two solar cycles, *Geophys. Res.*, 63, 651-669, 1958.

Foster, S., and M. Lockwood , Long-Term Changes in the Solar Photosphere Associated with Changes in the Coronal Source Flux, *Geophys. Res. Lett., 28*, 1443-1446, 2001.

Lean, J., Evolution of the Sun's spectral irradiance since the Maunder minimum, *Geophys. Res. Lett., 27 , 2425-2428,* 2000.

Lockwood, M., Long-Term Variations in the Magnetic Fields of the Sun and the Heliosphere: their origin, effects and implications, *J. Geophys. Res., 106*, 16021-16038*,* 2001

Lockwood, M., Long-term variations in the open solar flux and links to variations in earth's climate, *In "From Solar Min to Max: Half a solar cycle with SoHO"*, Proc. SoHO 11 Symposium, Davos, Switzerland, March 2002, ESA-SP-508, pp 507- 522, ESA Publications, Noordvijk, The Netherlands, 2002.

Lockwood, M., Twenty-three cycles of changing open solar flux, *J. Geophys. Res.* In press, 2003

Lockwood, M., R. Stamper, and M.N. Wild, A doubling of the sun's coronal magnetic field during the last 100 years, *Nature*, 399, 437-439, 1999a.

Lockwood, M., R. Stamper, M.N. Wild, A. Balogh, and G. Jones, Our changing Sun, *Astron. Geophys*, *40*, 4.10-4.16, 1999b.

Mayaud, P.N., The *aa* indices: A 100-year series characterising the magnetic activity*, J. Geophys. Res., 77*, 6870-6874, 1972.

Mayaud, P.N. , Derivation, Meaning and Use of Geomagnetic Indices*, Geophysical Monograph **22***, Am. Geophys. Union, Washington D.C., 1980

McCracken, K.G., and F.B. McDonald, The long-term modulation of the galactic cosmic radiation, 1500-2000, *in Proc. 27th. Int. Cosmic Ray Conference*, Hamburg, 2001

Neher, H.V., V.Z. Peterson, and E.A. Stern, Fluctuations and latitude effect of cosmic rays at high altitudes and latitudes. *Phys. Rev., 90*, 655-674, 1953.

O'Brien, K., Secular variations in the production of cosmogenic isotopes in the Earth's atmosphere, *J. Geophys. Res., 84*, 423-431, 1979.

Pulkkinen, P.J., J. Brooke, J. Pelt, and I. Tuominen, Long-term variation of sunspot latitudes, *Astron. Astrophys., **341***, L43-L46, 1999.

Pulkkinen, T.I., H. Nevanlinna, P.J. Pulkkinen, and M. Lockwood, The Earth-Sun connection in time scales from years to decades to centuries, *Space Sci. Rev., Space Sci. Rev., 95(1/2),* 625-637, 2001.

Stuiver, M., and P.D. Quay, Changes in atmospheric carbon-14 attributed to a variable Sun, *Science, 207*, 11, 1980.

Wang , Y.-M., and N.R. Sheeley Jr., Sunspot activity and the long-term variation of the Sun's open magnetic flux, *J. Geophys. Res., 107(A10), paper* SSH10, 2002.

==

Referee 1

This paper attempts to show that inference, reported by Lockwood et al., that the solar coronal magnetic field has doubled in strength in the last hundred years, derived from correlation studies with the aa geomagnetic activity index, is false. I am not convinced by the arguments put forward for the following reasons:

1. The authors base their argument in the observation that the reconstructed aa time series, derived from the IHV index, does not show an increasing trend. What is clear in Figure 1 is that, when going back in time the reconstructed aa time series diverges from the original time series in 1957. I believe this happens to coincide with the year of changeover between Cheltenham and Fredericksburg, the two observatories used to construct the IHV index. The authors pass no comment on this, but it is clearly highly relevant to their argument. Unless the authors address this point the paper cannot be taken seriously. Ideally the study should be made using a number of observatories which have operated continuously for as long as possible. This would not be difficult to do and would remove this very obvious objection to their argument.

See our response to Lockwood on this point. We *did* intercalibrate and also used several other stations all showing the same result.

2. The authors state that IHV is "completely objective". In fact the hourly mean values on which it is based would have been scaled by hand. The process of scaling hourly

mean values by hand is more difficult than scaling range values, and just as open to observer bias. There should at least be some recognition of the problems associated with handscaling hourly mean values, and a discussion of how it may affect the results.

See our response to Lockwood on this point. Scaling mean values is just a simple measurement problem with no bias but as for any other direct measurement with random measurement errors. The error of the mean is much smaller than the error of individual values. Systematic errors in baseline values are avoided by taken the difference between successive hourly mean values.

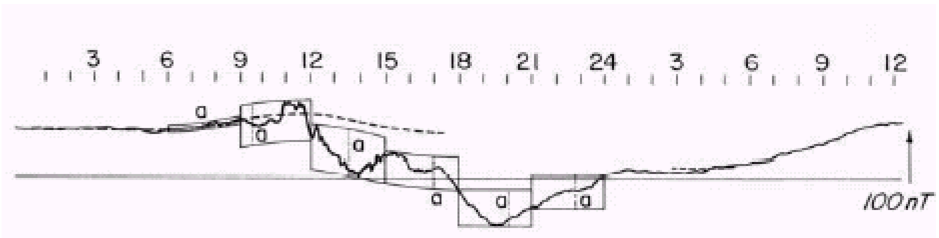
Scaling K-values are qualitatively different, because the observer has to distinguish between wiggles caused by solar EUV radiation and wiggles caused by the solar wind.

3. The quote from a paper by Clilverd et al. that instrumental effects were detected in the aa time series in 1938, 1980 and 1997 is misleading. These effects could only be detected in studies of the quietest geomagnetic periods (K=0 or 1) and had a negligible effect on the annual average of aa, which is what the authors are using.

Not so. As our Figure 4 shows, more than half of all values in 1901 had a K-index of 0, i.e. no measurable activity at all. Other early years also had a large number of quiet intervals. The quiet intervals during these early years are a large part of the argument that aa was much lower around 1900 than now, so the issue of the quietest values is critical.

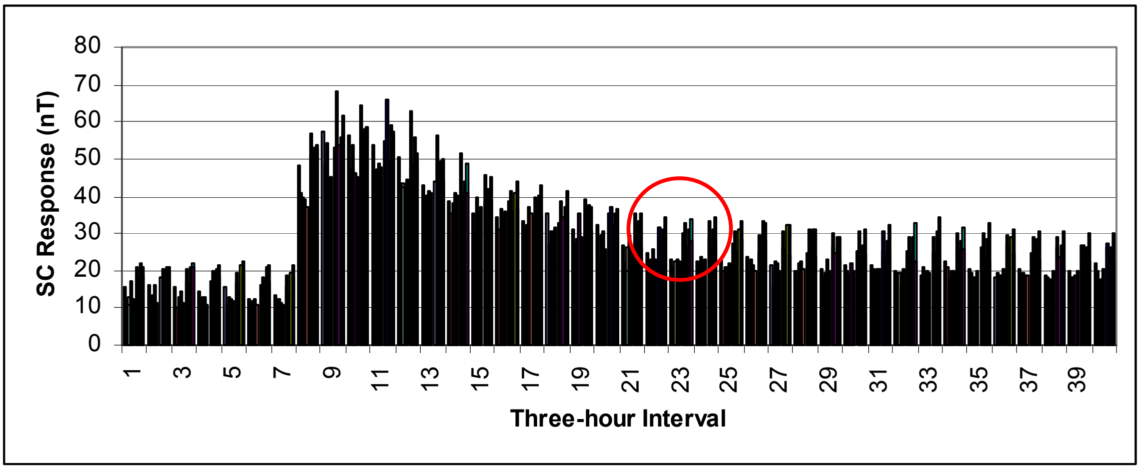
4. In their superposed epoch analysis the authors seem to misunderstand the difference between the plots of H and the plots of aa. The aa index responds more to the general level of substorm activity (being derived from 3-hourly ranges, and substorm activity typically having timescales of tens of minutes to one hour) and the hourly mean H is responding more to the global ring-current distribution. The superposed epoch analysis will remove much of the variation in H caused by the substorm processes which would be of a random nature, leaving the well-known profile of the ring-current effects as shown in the figure. A comparison of these plots cannot be used to support their argument.

The following famous Figure (originally due to Mayaud himself) illustrates the issue:



During this magnetic storm (with a healthy Dst) the amplitudes of the 3-hour K-variations vary roughly with the depth of the curve. Large Dst = large K-variation. By the way, the Figure also illustrates how the K-amplitude is measured as the range between two curves that follow the "quiet" variation and is thus no simple range from minimum to maximum. Our own Figure 2b also shows that there is a typical aa-response to a SSC, so that large Dst = large aa. It is true that the superposed epoch analysis averages out the substorm influence, but that only means that our Figure 2 is a good proxy for the ring current strength. What our Figure 2a shows is that there is no (or only a small, at best) difference between ring current strength over the three divisions of the century. Now, Dst also depends on the strength of the interplanetary magnetic field, so one would expect a factor of two change to have a noticeable effect on the Dst response.

We performed the superposed epoch analysis of aa with SSCs as key times separately for each solar cycle from #11 to the present. The following figure is the result (discussed in our long paper):



For each three-hour interval 13 touching vertical bars are plotted, in sequence from cycle 11 through 23. The height of each bar is the average value of aa at that epoch. It is apparent from this Figure that the first six solar cycles show a very similar response, but then the whole response seems to be abruptly shifted uniformly upwards by 5-10 nT for the last seven cycles. Closer examination (using a 10-year moving window) shows that the abrupt shift takes place between 1937 and 1938. It is this bimodal nature of the response that makes us conclude that a discontinuity is present between 1937 and 1938.

The above were my original opinions formed before reading Lockwood's reply. I endorse all of the points in Lockwood's reply, and so I recommend that this paper should not be published in its present form.

See our response to Lockwood's points.

** Referee 2

General Comments

The title of the paper by Svalgaard clearly states the motivation for writing the paper but the title is not proven by the paper. The topic of the paper is whether the aa index has maintained its calibration over the years and whether it is a reliable proxy for the activity level of the Earth's magnetosphere. The paper should either devote more space to the topic outlined in the title or the title should be changed to the equally important issue of the accuracy of the aa index.

The accuracy of the aa-index is important, but such a title would not alert the wider (e.g. climate) community to the possibility that the aa-index has anything to do with the sun's magnetic field. What they have seen and remember is the claim of a doubling of the sun's field. Most of them have no idea what the aa-index is. One could argue that Lockwood's paper should on the same grounds not have had the title "A doubling …" or as their JGR paper had: "Solar cause of a rise in ….". The "A doubling …" title was sensationalism as the paper did not show that sun's field has been measured to double, but only that one could speculate that that is what might have caused the apparent rise of aa. Anyway, the title issue is not a showstopper for us. As suggested above, placing a question mark at the end of our title might be a softening touch.

It is difficult for magnetometers to be incorrectly calibrated as their data are used in global internal magnetic field models that change slowly. Any errant station with perhaps more than 0.1% error would be noted quickly. It is possible that resolution would be of concern but in general that would not affect disturbed conditions only quiet times. The lower panel of Figure 2 clearly shows that all values of aa are lower during the blue period. [The upper panel shows that the magnetometers have the same calibration for three epochs but we knew that had to be true. See above].

We concur with this

Returning to Figure 2 there are two measures of geomagnetic disturbance shown. In the top panel there is basically a measure of the energy pumped into the magnetosphere during storms. There is very little difference in the three epochs. In the bottom panel is the aa-index that effectively measures the currents in the auroral ionosphere as sensed at mid-latitudes. These values can be affected by ionospheric conductivity and by the size of the auroral region. The ionospheric conductivity can be affected by solar EUV a quantity that is solar cycle dependent and could vary if the sun's magnetic field changed.

We examined this carefully. Figure 4 of our “long” paper (that none of the participants were given) shows that the amplitude of Sq (which is a proxy of the EUV) does not have any long-term trend. More over, aa was designed to be independent of the EUV flux. Using antipodal stations is part of that scheme.

The authors' new IHV index will not because it is created from nighttime values. Thus, it is quite possible for IHV and aa to have a long-term variation in their ratio without either being wrong. Moreover, the aa index because it measures substorms and the auroral currents could easily have long term variations that are different than those seen in the ring current. In short, all the variations seen could be real and there is no error in any of the indices.

Except that the amplitude of the aa-response to an SSC in the early years is not substantially different from the later years. What seems to have changed is the base level of aa. We end our paper with the statements: “We have presented several lines of evidence that each casts doubt on the validity of the assumption on constant calibration of the aa index. For each, in isolation, there might be other explanations. *Collectively*, they offer strong evidence that there is no doubling […]”. Although “all the variations seen could be real”, they would conflict with the other evidence.

What does this say about the long-term variation of the Sun? It is very, very clear that the Sun does vary on the long term. There is much evidence for that from sunspot numbers to carbon-14 data. Did the Sun's magnetic field double? It is difficult to say. However, a much different paper would have to be written to prove or disprove that statement. This paper sheds no light on that topic.

Since the original claim that the field doubled was based on a doubling of aa, if aa did not double, we infer that that argument for a doubling is not valid. Did it really double? Or halve? Or whatever? As the referee points out, “it is difficult to say” with real measurements. To the extent that Lockwood’s paper sheds light on the topic, so does ours.

I strongly recommend that the authors refocus their paper on the aa index versus the IHV index and whether the differences point to a problem with aa or whether they provide new insight into the variation of the Sun and its interaction with the Earth's magnetic and plasma environment. With this refocus, a reply by Lockwood would seem unnecessary.

Our long paper examines the issue of aa and IHV derived from many stations and using both the H and D components in great detail and the conclusion (one we didn’t like ourselves) was hard to escape, namely that it is aa that is at fault. Our paper does not provide new insight into the variation of the sun, but if aa is at fault, so did not Lockwood et al’s paper either. Our provocative title was simply mirrored after their title. We could soften our title by adding a question mark.