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

Leif Svalgaard, Ed W. Cliver, and Philippe Le Sager
December 2002

The Lockwood et al. Conjecture

Lockwood *et al*. (1999) have conjectured that the mean coronal source field over the interval 1901-1995 has increased 131% or more than doubled. Figure 1shows their basic result:

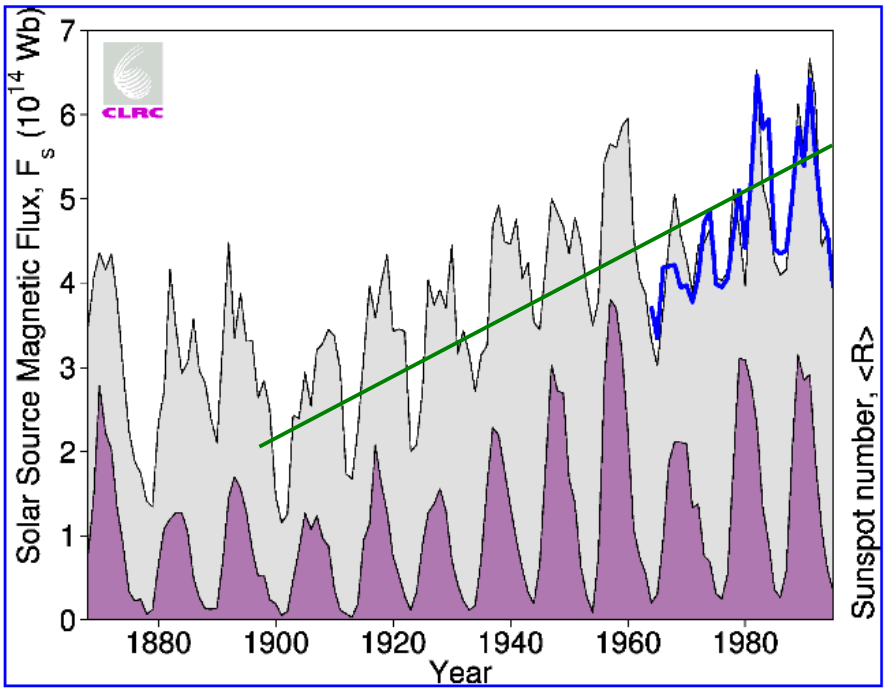

Figure 1**.** Annual means of the coronal source magnetic flux. Those derived from the *aa* index, F_s, are shown by the area shaded grey, whereas those from near-Earth measurements of the IMF during solar cycles 20-22, F_{so}, are shown by the thick line. The darker shaded area shows the variation of the smoothed sunspot number (from Lockwood et al (1999)). The green line shows the purported increase in the magnetic flux.

Lockwood *et al*. base their study on the *aa* geomagnetic activity index (Mayaud 1972) that is compared with spacecraft measurements of the interplanetary magnetic field since 1963. Under the assumption that the *aa*-index has a constant calibration over time, their conjecture that the interplanetary, and by (not at all certain) implication, the Sun's magnetic flux has more than doubled in the last 100 years would seem to follow. Extraordinary claims require extraordinary evidence. In the present paper we produce evidence that the interplanetary magnetic field has *not* changed systematically over the last 100 years, but that instead, the *aa*-index does *not* have a constant calibration over time.

Cheltenham/Fredericksburg, 1901-2000

The USGS Geomagnetic observatory Fredericksburg (Geog. Lat. 38.20N Long. 282.63E; Geomagn. Lat. 49.13N, Long. 352.23E, Elev. 69 m) has been in operation since 1956 (with a brief gap at 1981-83). Fredericksburg replaced the Cheltenham observatory (Geog. Lat. 38.70N, Long. 283.20; Geomagn. Lat. 49.13N, Long. 352.23E, Elev. 72 m) that had operated uninterrupted from 1901 through 1956, overlapping Fredericksburg during most of 1956. Hourly values of the geomagnetic field components H, D, and Z can be downloaded from WDC-C2 in Kyoto. This 100-year long series of field measurements have constant calibration (serving as the U.S. standard station).

Solar quiet-time Daily Variation (*Sq*)

The horizontal component shows a regular variation though the day as shown in Figure 2. This variation (the so-called *Sq*-variation) is caused by winds and tides in the ionosphere and does not constitute what we today call "geomagnetic activity".

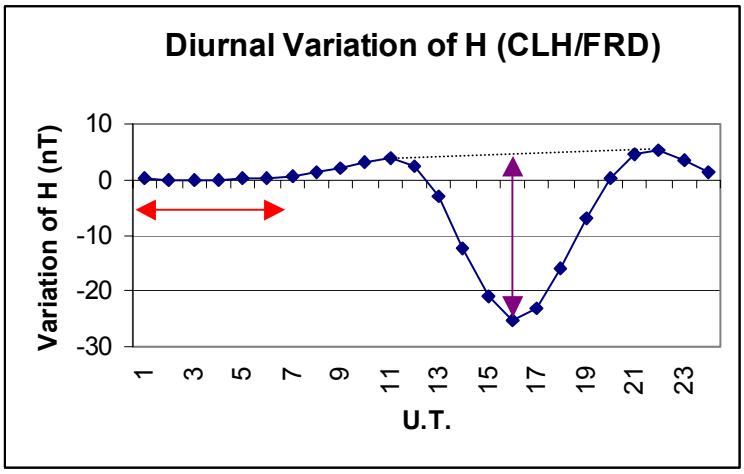

Figure 2. The daily variation of the H-component at Cheltenham/Fredericksburg for the years 1901-2000. The zero-level has been taken as the value at 3:30 UT. The definition of the local night-time is shown by the horizontal arrow. The definition of the amplitude of the *Sq*-variation is shown by the vertical arrow.

The conductivity of the mid-latitude ionosphere is controlled mainly by ultraviolet radiation from the Sun. The amplitude of the daily *Sq* variation is then controlled by the disk-integrated solar *UV*-flux, which in turn varies significantly with the solar cycle, specifically, with the sunspot number. The *Sq*-variation is not noticeable during local night hours (marked with a horizontal arrow on Figure 2). We can therefore divide the daily variation into two parts, one part (during the day) that is the *Sq*-variation plus whatever geomagnetic disturbances might be present, and one part (during the night) where only disturbances (geomagnetic "activity") are seen.

Sunspot Number Variation of the *Sq*-Amplitude

That the *Sq*-amplitude varies with the sunspot cycle has been known for a long time (e.g. Ellis 1880; Wolf 1859). Many other solar, interplanetary, and geomagnetic parameters vary with the solar cycle as well. It is important to note, that the *Sq*-amplitude varies with the (photospheric) *UV*-flux (or its proxy, the sunspot number) and *not* with solar wind parameters such as the solar wind speed (Figure 3).

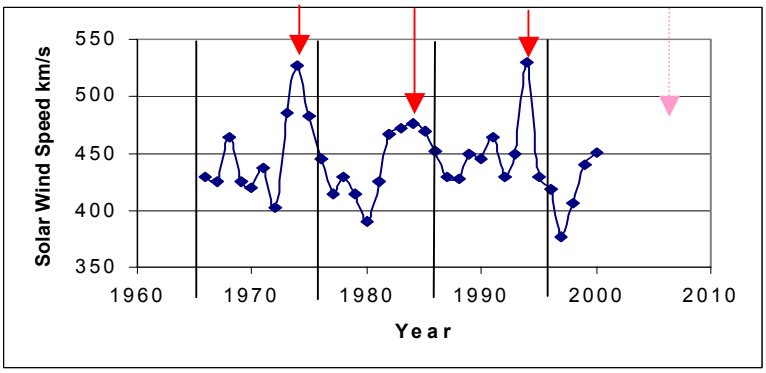

Figure 3. Solar wind speed since 1966. As is well known, the solar wind speed often peaks sharply in the year(s) just before the end of the solar cycle when strong long-lived solar wind streams from coronal holes are present. This is clearly seen in the figure, e.g. for the years 1974 and 1994. Other years where geomagnetic records indicate strong solar streams include 1930 and 1943.

As can be seen in Figure 4, the *Sq*-variation does not exhibit such peaks near the end of each cycle:

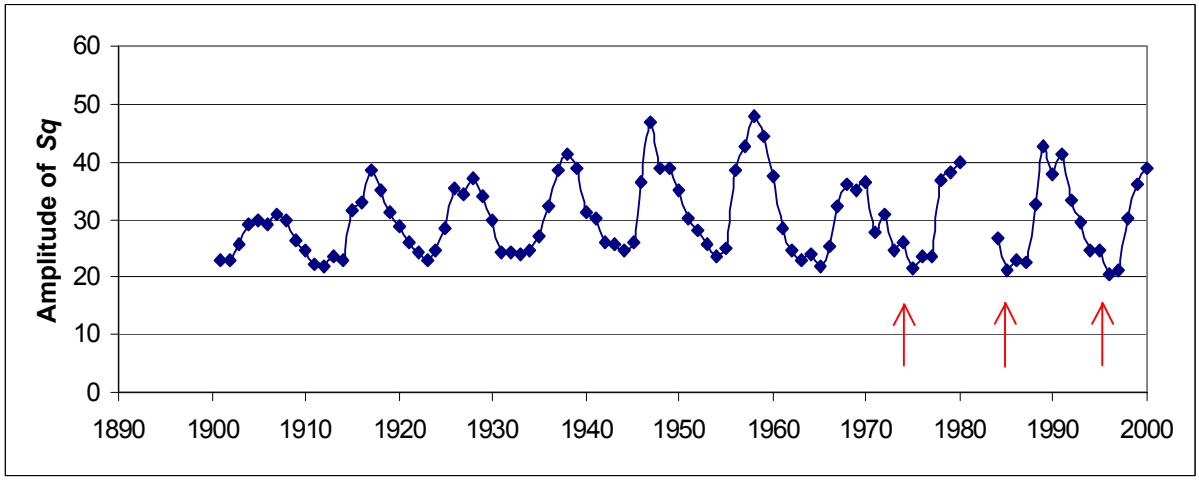

Figure 4a. The amplitude (in nT) of the *Sq*-variation at Cheltenham/Fredericksburg during the 20th century. Note the *absence* of peaks at the times when the solar wind velocity was peaking near the end of most cycles.

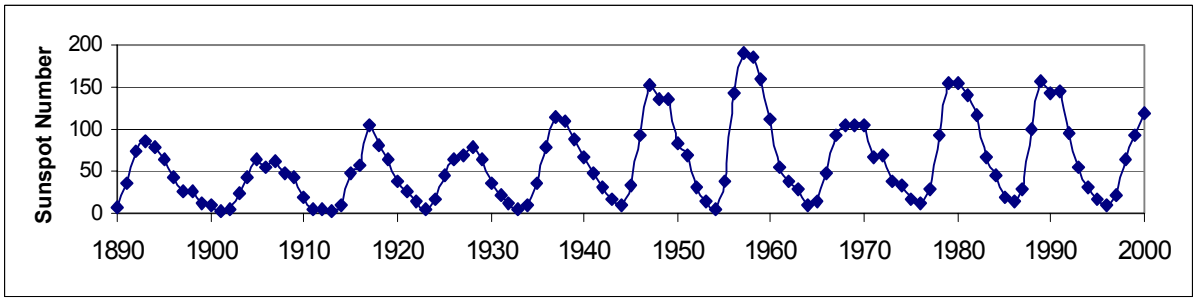

Figure 4b.The-sunspot number during the interval 1890-2000. Note how well it tracks the amplitude of the *Sq*-variation. Yearly averages are shown in Figures 4a and 4b.

The *Sq*-amplitude has a very strong, and well-known, correlation ($R^2 = 0.901$) with the sunspot number. It is rare to find a (non-solar) quantity with such a strong correlation with the sunspot spot number (Figure 5):

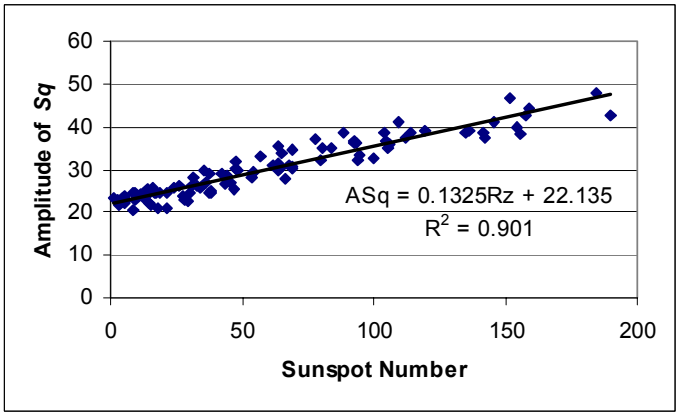

Figure 5. Amplitude of the average yearly *Sq*-variation at the combined station-pair Cheltenham/Fredericksburg for the interval 1901-2000 as a function of yearly average sunspot number (R_Z). Already Wolf (1859) suggested the following linear relationship: $ASq = a(1+mR_Z)$. Wolf (and others) found $10^4 m$ to be typically between 50 and 100 depending on latitude. The relationship in Figure 5 corresponds to $m = 60$.

Sunspot Number Variation of the Interplanetary Magnetic Field Strength

We noted the strong solar wind streams just before the end of each solar cycle see (Figure 3). There are no corresponding peaks or enhancements of the magnitude of the solar wind total magnetic field:

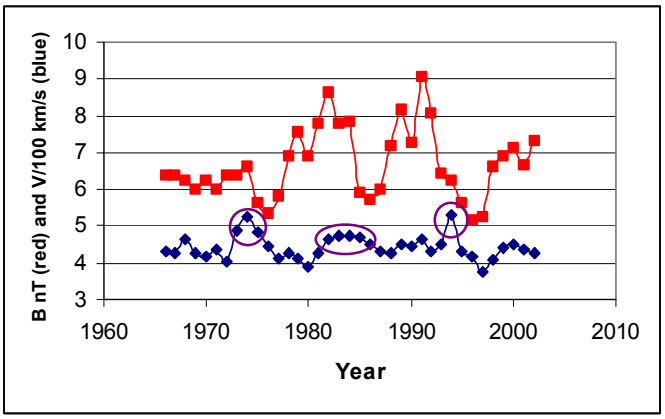

Figure 6. Yearly averages (red) of the interplanetary magnetic field magnitude (nT) and of the solar wind speed (blue, in units of 100 km/s) for the total time where we have reliable spacecraft measurements near the earth. Note that the strong solar wind streams near the end of each solar cycle have no corresponding peaks of enhanced magnetic field. If anything, the magnetic field is low within the high-speed streams. Also note, that there is *no significant long-term trend* in neither of the two quantities nor any significant correlation between them.

It seems well established that there is a sunspot *number* dependence of the interplanetary magnetic field strength (Figure 7). The correlation is only moderately strong (R^2 = 0.5298), and is not unexpected if one makes the (reasonable, but not compelling) argument that more magnetic flux on the Sun eventually means more magnetic flux in interplanetary space at the Earth:

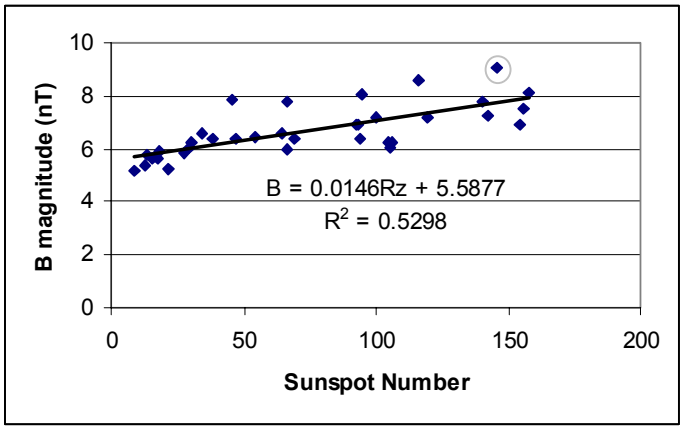

Figure 7. Correlation between yearly means of the sunspot number and of the field magnitude (in nT) of the near-earth inter-planetary magnetic field. The five data points that lie clearly above the best-fit line are from the interval 1982-1992, but so are the lowest data points, so we do not think there is a trend here either. For a sunspot number of zero, the best-fit linear relation $B = 5.5877 + 0.0146R_z$ predicts a field magnitude of 5.6 nT.

Yet Another Index: The Interhour Variation

The long time-series of hourly values of the magnetic components from several stations form a rationale for investigating if an index can be constructed from them with the following properties:

- Completely objective (no subjective determination of baselines, quiet background, or the like)
- Easy to duplicate by other researches for verification purposes
- Long-term (potentially going back to about 1835 or beyond)
- Physically Quantitative (expressed directly in or proportional to magnetic variations measured in nT)
- Easy to understand (existing indices are derived via a process that is convoluted, intricate, and in its details largely unknown - to most researchers)
- Suitable for studies of long-term variations (on time scales of weeks and longer)
- Having the property that its limitations are clearly stated and understood
- Constant calibration over time for any given station (not requiring conversion tables or daily or seasonal adjustment tables)

We take our starting point in the observation that *Sq* is almost absent in the middle of the (local) night and concentrate on the variation of the H-component during these night hours. Figure 8 shows a typical variation during several consecutive days.

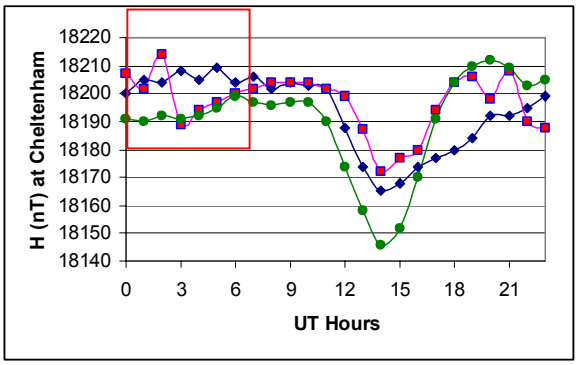

Figure 8. The run of H during several days (in May 1942) at Cheltenham. The first day (blue diamonds) starts out relatively quiet, then becomes more active at the end of the day. Activity continues into the second day (pink squares) before settling down during the third day (green circles), that however exhibits a lingering depression of the H-component due to the equatorial ring current. By only considering the local night hours from 0 through 6 UT, we eliminate variation of Sq. During the night hours (red box) considerable variation of the value of the H-component is seen from hour to hour.

As a measurement of the variation from hour to hour, we compute the differences between the measured values of the H-component without regard to their sign from one hour to the next through the night. The average value of these absolute valued differences for the night (assigned to the UT day) we call the InterHourly Variation (the *IHV*). Formally:

$$IHV^{H}(\text{day}) = \sum_{h=h1}^{h=h2} [\text{abs}\,(H_h - H_{h+1})] \,/\, (h2\text{-}h1+1) \qquad (1)$$

Where the hourly index *h* runs from *h1* through *h2* (from hour 0 UT through 5 UT for Cheltenham/Fredericksburg). If *h1* and *h2* are on opposite sides of 0 UT, the hourly index is taken modulo 24 (this is a potential quibble point). We could define similar indices for the D, Z, X, or Y components. If we omit the superscript (the ^{H} in IHV^{H}), the H-component is assumed.

Interstation Calibration

The requirement of "constant calibration" can be relaxed when simultaneous observations exist for several geomagnetic stations or observatories allowing a simple "cross calibration" provided it is clear that the index derived from one station differs but by a constant factor from that derived from another station. An example of this is shown in Figure 9. Cheltenham and Fredericksburg are located were close to each other (only a few miles apart) and operated simultaneously for 273 days during 1956. The very strong correlation shown in Figure 9 justifies calibrating the IHV-index values derived from (the older station) Cheltenham to corresponding values as were they derived from Fredericksburg by multiplying them by 0.970.

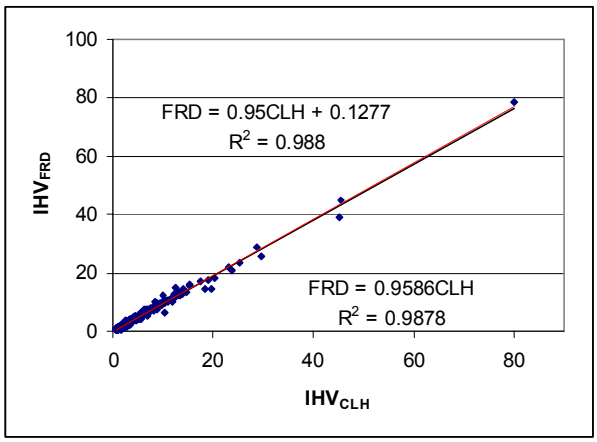

Figure 9. 273 pairs of simultaneous IHV values derived from the stations Cheltenham (CLH) and Fredericksburg (FRD) during most of 1956 are very strongly correlated. The red line is the best-fit line forced through the origin and the blue line (if you can distinguish them) is the best-fit line with an offset. The correlation coefficients are very high and not really significantly different so we first adopt the fit without an offset, leading to $IHV_{FRD} = 0.959\ IHV_{CLH}$. As most of the values are concentrated near the origin one could also simply calculate the calibration constant as the ratio between the mean values $\langle IHV_{FRD}\rangle / \langle IHV_{CLH}\rangle = 5.98 / 6.16 =$ **0.970**, and this is what we finally adopted.

For days with no measurements (chiefly the years 1981-83) at either Cheltenham or Fredericksburg, the IHV values were computed from Niegmek (NGK) which has a calibration factor to Fredericksburg of 0.8615, determined the

same way as in Figure 9. The end result is an unbroken time series of IHV for each day covering the one hundred-year interval 1901-2000.

Calibration Against the *am*-Index

The *am* geomagnetic activity index (Mayaud, 1980) is generally considered the most "accurate" geomagnetic index derived from a reasonably dense network of midlatitude stations in both hemispheres. In this section we'll investigate how our *IHV*-index correlates with the *am*-index on several time scales. We first note that because we only use a few night hours to define IHV, the *IHV*-index is not meant to replace the *am*-index. In Figure 10 we compare the daily values of IHV_{FRD} to the average *am*-index during the 6 hours for which we computed the *IHV*-index for each day during the year 1968.

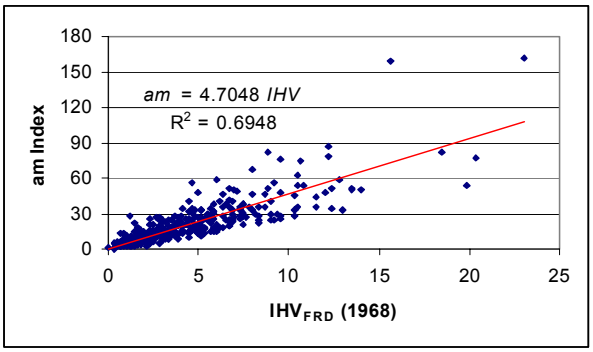

Figure 10. Relationship between all daily values of the *IHV*-index for the station Fredericksburg and the values of the *am*-index for the same 6-hour intervals of the day during the year 1968. A decent linear relation with correlation coefficient R^2=0.6948 is indicated. The correlation is very similar for other years throughout the interval from 1959 to 2000 for which the *am*-index is available.

The *am*-index is a *range* index over a time interval while the *IHV*-index is a *variability* index from one interval to the next. Statistically, a correlation between ranges and variability is to be expected, growing stronger with increasing number of intervals. While it was comforting to verify that even on a one-day basis, *am* and *IHV* were nicely correlated, what we really want is to verify that longer time averages are more strongly correlated. It is especially important to verify that the correlation (and calibration) is constant over time as we wish to investigate long-term trends.

The *am*-index is available since the beginning of 1959. We now compute 27-day averages of *am* and of IHV_{FRD} (synchronized with 27-day Bartels rotations and for the same 6-hour UT-intervals). To easily compare the two quantities, we have multiplied IHV by 4.641 (the ratio between their averages) so that the two time series have the same mean values. We can then plot the 27-day averages as a function of time (the rotation number) and visually compare them. Figure 11 is the rather tedious result, but it is important to convince the reader that the correlation is good and strong, and that there are no systematic differences between the two time series:

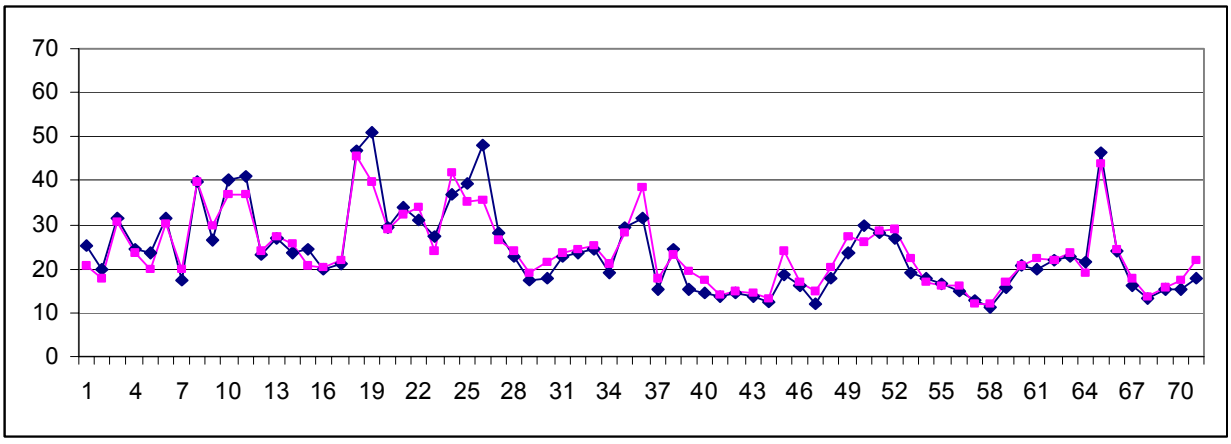

Figure 11a. Bartels rotation averages of *am* (blue diamonds) and of IHV_{FRD} (pink squares) for 1959-64 (the first 71 rotations since 1959). The interval from 1969 to 2000 was divided into 8 segments each 71 rotations of length. Figure 11a through Figure 11h show each segment. *IHV* values multiplied by 4.641.

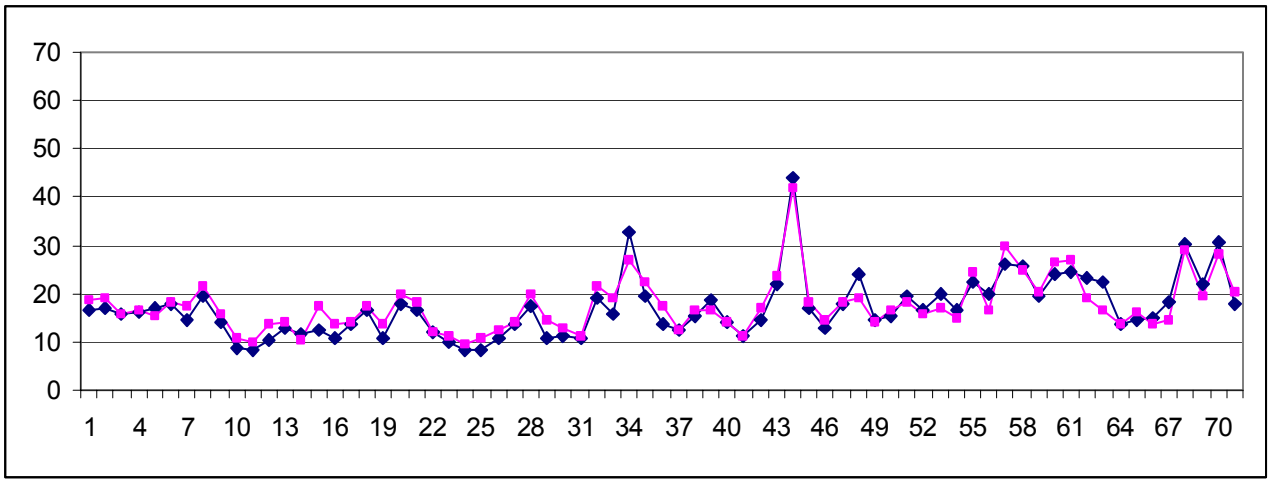

Figure 11b. Like Figure 11a, but for the next 71 rotations (1964-69)

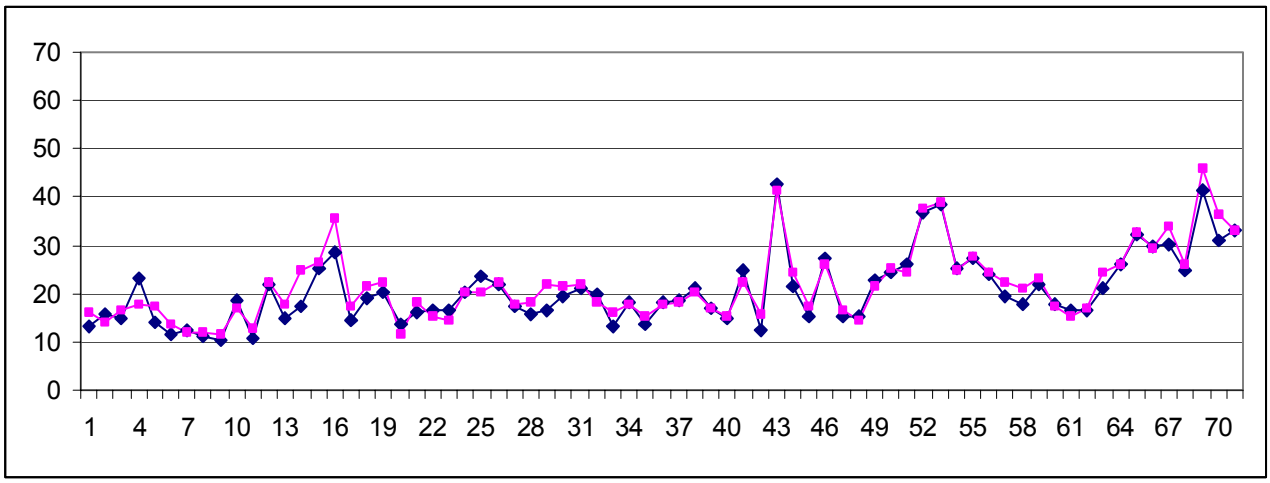

Figure 11c. Like Figure 11a, but for the next 71 rotations (1969-74)

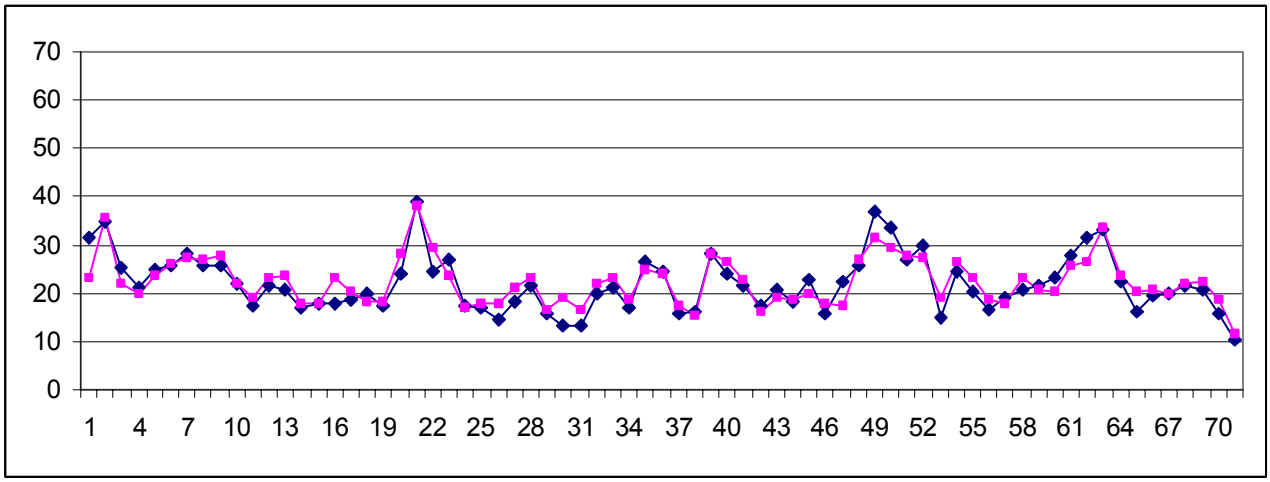

Figure 11d. Like figure 11a, but for the next 71 rotations (1974-79)

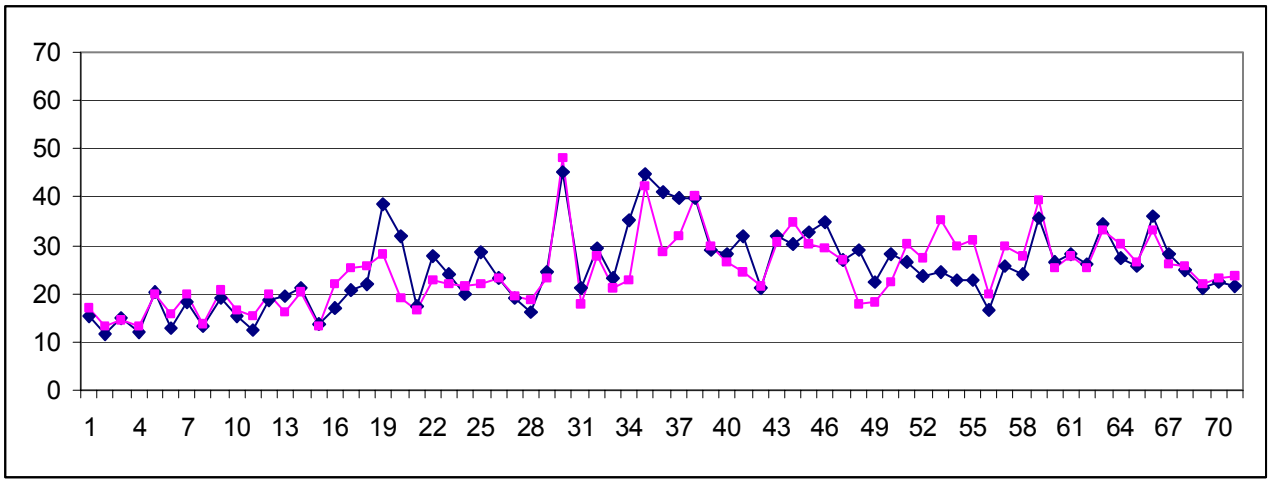

Figure 11e. Like Figure 11a, but for the next 71 rotations (1980-84)

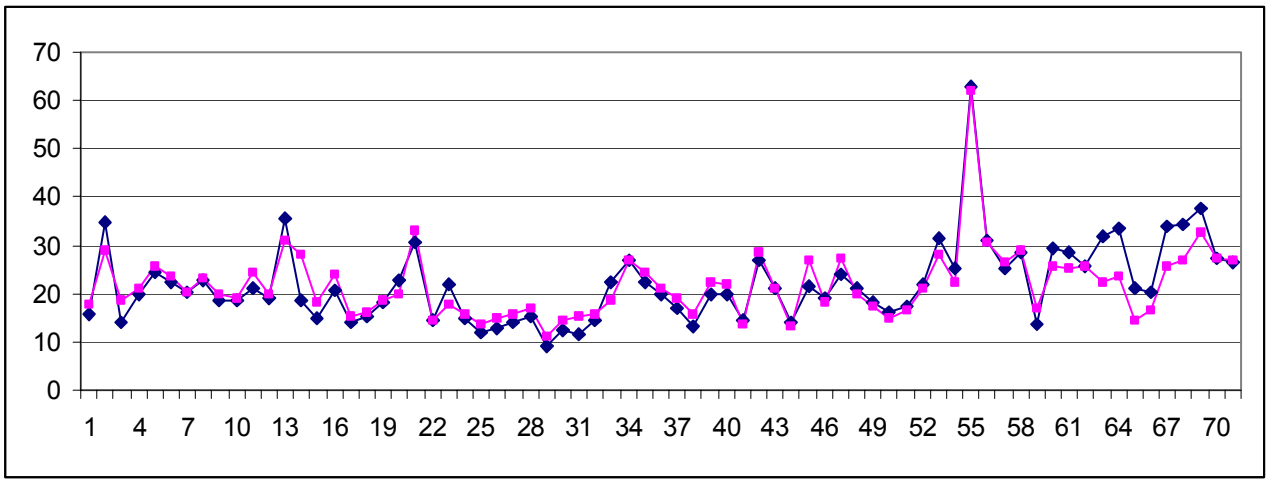

Figure 11f. Like Figure 11a, but for the next 71 rotations (1985-90)

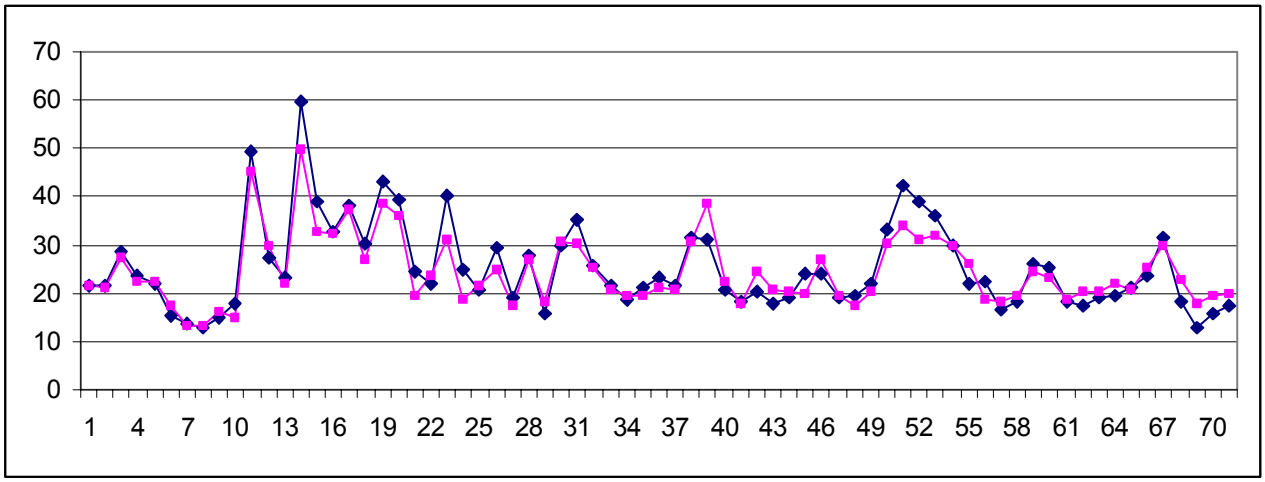

Figure 11g. Like Figure 11a, but for the next 71 rotations (1990-95)

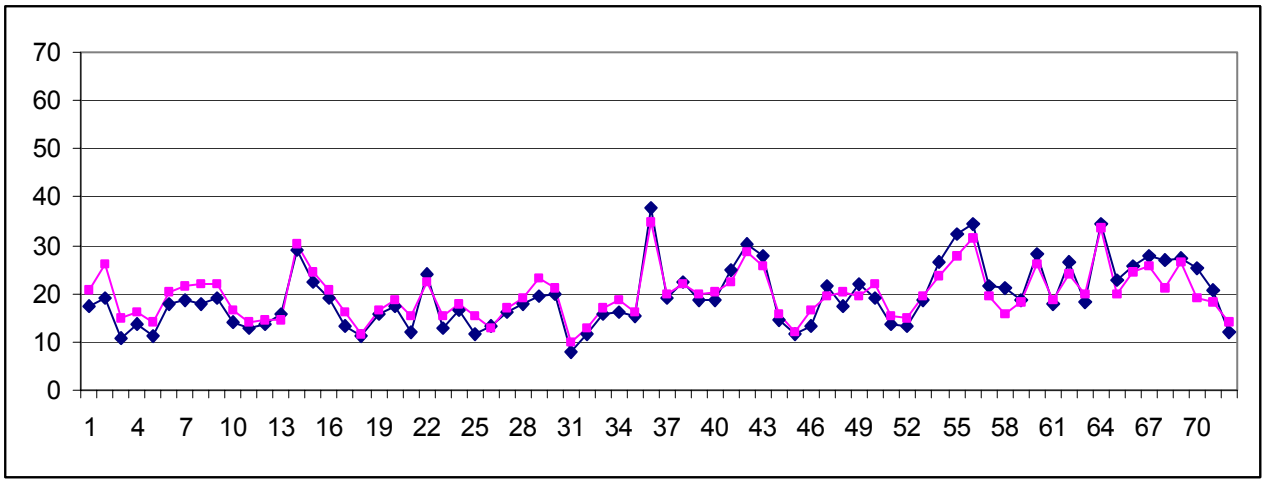

Figure 11h. Like Figure 11a, but for the last 72 rotations (1995-2000)

We urge the reader to study these comparisons carefully and come away convinced that the correlation is strong throughout (in fact $R^2 = 0.7991$) and that, although occasional differences exist, no *systematic* differences between the two data sets are evident. This is equivalent to saying that we can reconstruct *am* from IHV_{FRD} with fair accuracy even on a time scale as short as one solar rotation.

Reconstructing the *am*-Index

Figure 12 shows the relationship between yearly average values of the *IHV*-index for Fredericksburg and yearly average values of the *am*-index. Working with yearly averages increases the correlation to better than 0.9.

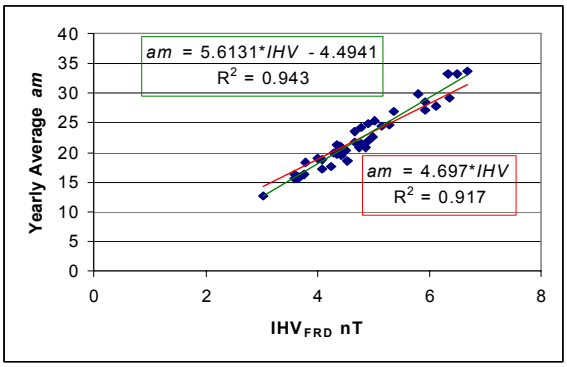

Figure 12. Yearly average of the *am*-index for each year during 1959-2000 compared to the yearly average *IHV*-index for Fredericksburg. As usual, the values used for calculation of *am* refer to the same UT-intervals used in the calculation of *IHV*. Lines with and without an offset are least-squared fitted as shown. The correlation is highest (R^2=0.943) for the line with an offset. We can use the result to reconstruct *am*:

$$am = 5.6131 * IHV_{FRD} - 4.4941 \quad (2)$$

Using eq.(2) we can now reconstruct *am*. The result is shown in Figure 13:

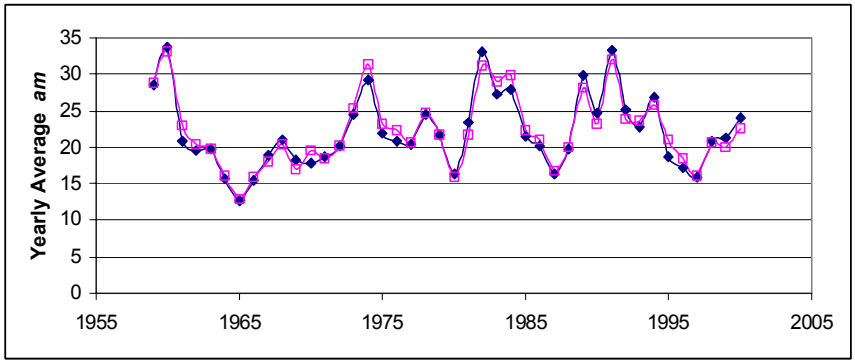

Figure 13. Reconstructed *am*-index (pink squares) for the interval 1959-2000 compared to the observed values (blue diamonds). Yearly averages are shown. No significant differences are apparent and no significant long-term trend either. Using the expression without the offset gives very closely the same result.

Reconstructing the *aa*-Index

Having successfully reconstructed the *am*-index for the entire interval where it is known, we can apply the same method to reconstructing the *aa*-index for the same interval. Figure 14 shows the relationship between yearly average values of the *IHV*-index for Fredericksburg and yearly average values of the *aa*-index. Working with yearly averages increases the correlation to better than 0.9.

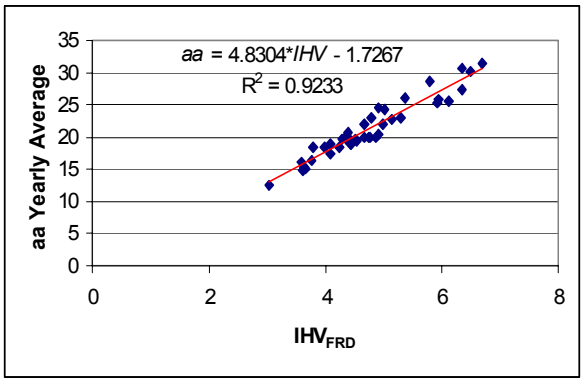

Figure 14. Yearly average of the *aa*-index for each year during 1959-2000 compared to the yearly average *IHV*-index for Fredericksburg. As usual, the values used for calculation of *aa* refer to the same UT-intervals used in the calculation of *IHV*. The least-squared fit best line is shown. The correlation (R^2=0.9233) is almost as high as for *am*. We can use the result to reconstruct *aa*:

$$aa = 4.8304 * IHV_{FRD} - 1.7267 \quad (3)$$

Using eq.(3) we can now reconstruct *aa*. The result is shown in Figure 15:

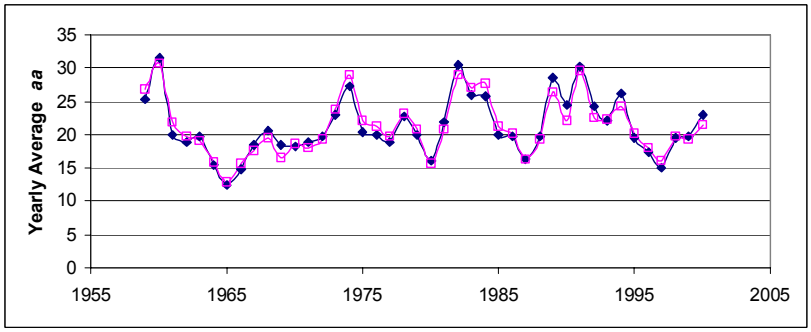

Figure 15. Reconstructed *aa*-index (pink squares) for the interval 1959-2000 compared to the observed values (blue diamonds). Yearly averages are shown. No significant differences are apparent and no significant long-term trend either.

Reconstructing the *ap*-Index

Having successfully reconstructed the *am*-index for the entire interval where it is known, we can apply the same method to reconstructing the *ap*-index for the same interval. Figure 16 shows the relationship between yearly average values of the *IHV*-index for Fredericksburg and yearly average values of the *ap*-index. Working with yearly averages increases the correlation to better than 0.9.

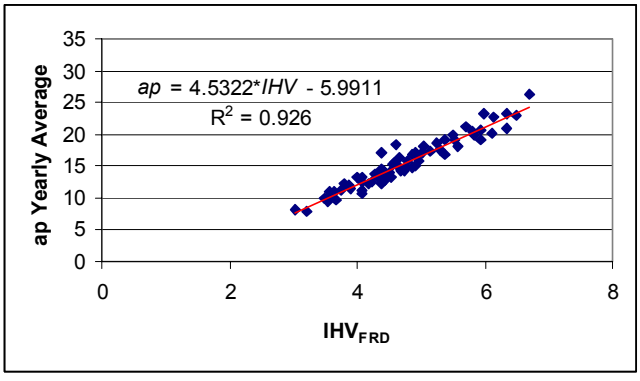

Figure 16. Yearly average of the *ap*-index for each year during 1959-2000 compared to the yearly average *IHV*-index for Fredericksburg. As usual, the values used for calculation of *ap* refer to the same UT-intervals used in the calculation of *IHV*. The least-squared fit best line is shown. The correlation (R^2=0.926) is almost as high as for *am*. We can use the result to reconstruct *ap*:

$$ap = 4.5322 * IHV_{FRD} - 5.9911 \quad (4)$$

Using eq.(4) we can now reconstruct *ap*. The result is shown in Figure 17:

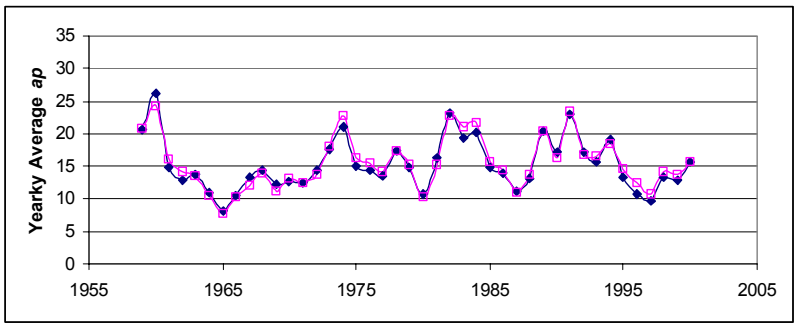

Figure 17. Reconstructed *ap*-index (pink squares) for the interval 1959-2000 compared to the observed values (blue diamonds). Yearly averages are shown. No significant differences are apparent and no significant long-term trend either.

Judging from Figures 13, 15, and 17 we can with fair success reconstruct all the range indices *am*, *aa*, and *ap* for the interval 1959-2000, at least their yearly average values. We deliberately chose the *same* interval for all three indices in order to be able to compare the efficacy of the reconstructions. So far, no significant differences are apparent. It does indeed seem that the *IHV*-index can be used as a basis for reconstruction of the standard geomagnetic indices. The very direct and straightforward method of determining the IHV values and thereby reconstructed standard indices seems to carry promise of extending the reconstructions as far back in time as we have hourly values available for mid-latitude stations, some 150 years or more.

Reconstructing the *aa*-Index for the Last One Hundred Years.

Bolstered by our success in reconstructing all three standard mid-latitude indices for the last forty years, we employ the exact same method and reconstruct *aa* for the past one hundred years. The result is shown in Figure 18:

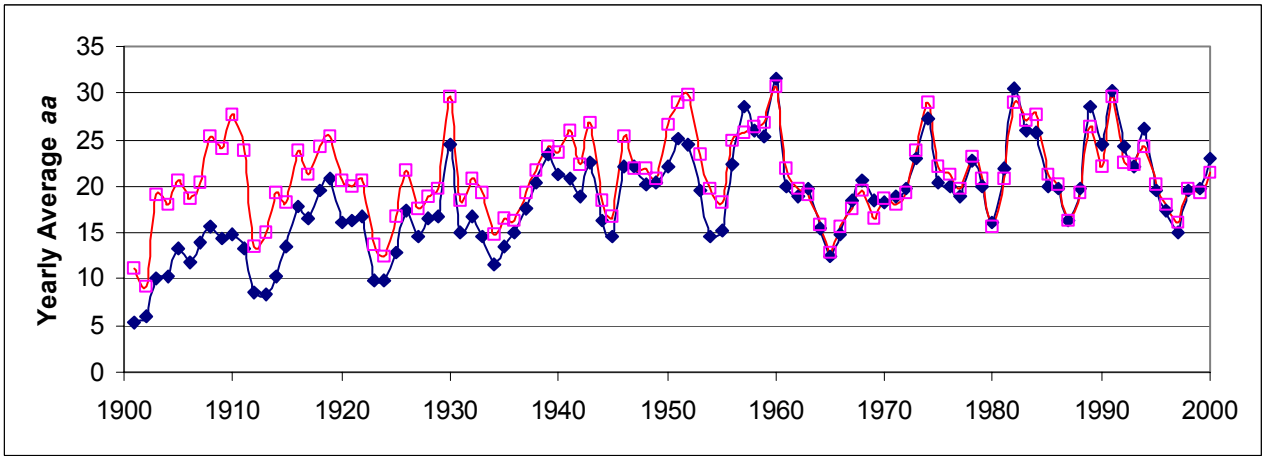

Figure 18. Observed (blue diamonds) and reconstructed (pink squares) yearly average values of the *aa*-index (during the two three-hour periods from 0 to 6 UT). The Reconstruction was based on *IHV* values from Cheltenham and Fredericksburg (CHL reduced to FRD levels) and on eq.(3).

The surprising result is that for the last forty years, the observed and the reconstructed values track each other rather closely, but that before, the two values diverge progressively as we go back in time to the point of differing by almost a factor of two in the early part of the 20th century. Because the rise of *aa* (as observed and reported by ISGI (Service International des Indices Geomagnetique)) since 1900 is the crucial observation on which Lockwood *et al.* base their claim that the Sun's coronal magnetic field has doubled since 1900, we find ourselves making an extraordinary claim too: that there is no evidence for a doubling of the Sun's coronal magnetic field in the last one hundred years. The remaining part of this paper provides further support for our claim.

The Differences Between Reconstructed and Observed Values

Figure 19 shows the differences between reconstructed and observed values of the *aa* and *ap* indices. Since the scale of the Figure is much larger than that of Figures 15 and 17, what small differences there were in the 1959-2000 interval (of the order of only ±2 nT) now show up as being clearly controlled by the sunspot number. Before about 1956, the differences grow large for both indices:

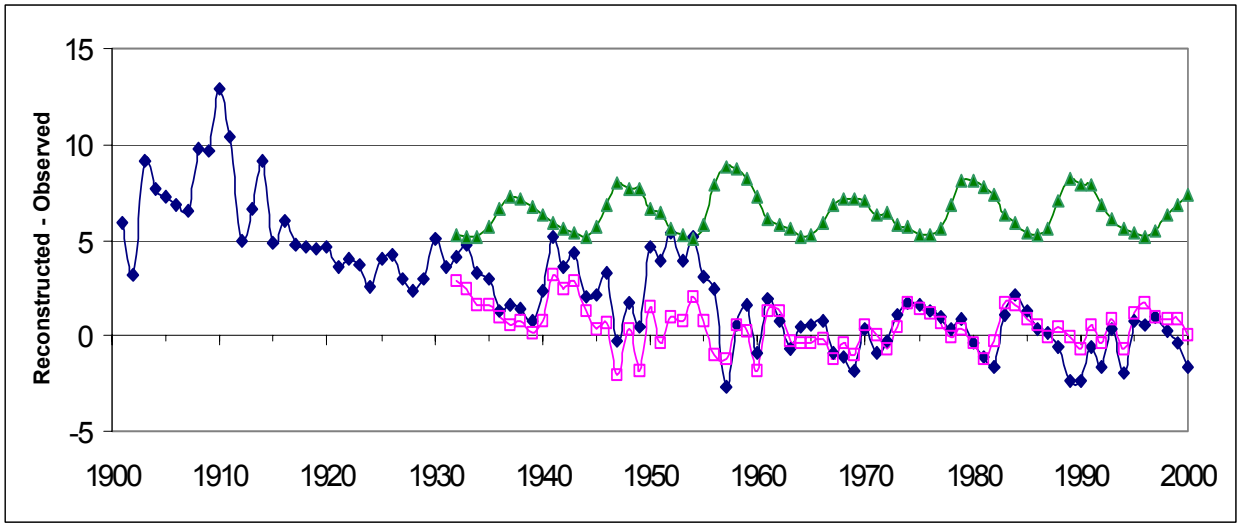

Figure 19. The differences between the reconstructed and observed yearly average values for the *aa*-index (blue diamonds) and for the *ap*-index (pink squares). The sunspot number cycle is indicated for cycles since 1930 (green triangles).

Because the *aa*-differences and the *ap*-differences generally track each other and are inversely correlated with the sunspot number, we infer that the cause of the discrepancies lying in the reconstruction process and that a small sunspot number dependent correction would improve the fit. Such a correction would violate the directness and simplicity of the reconstruction, but seems in order if agreement within 2 nT or less is sought. In any case, even without the correction, the result remains that serious doubts have been cast on the calibration or the homogeneity of the *aa*-index before about 1956.

Figure 20 shows the differences as a function of the sunspot number:

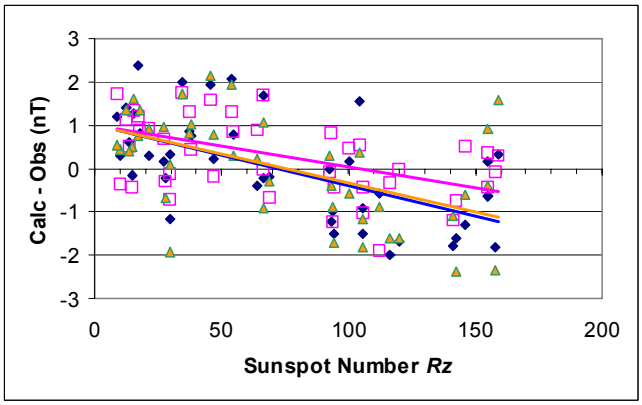

Figure 20. The differences between yearly mean values of reconstructed and observed standard indices *am* (blue diamonds), *ap* (pink squares), and *aa* (orange triangles) for each year of the interval 1959-2000. While the scatter is appreciable (R^2 of the order of 0.3) the trend is nevertheless clear and the same for all three indices.

The following approximate relationship for the difference, *D*, holds equally well for all three indices:

$$D_{year} = 1 - R_{zYear}/75 \qquad (5)$$

When the sunspot number is low, the difference is positive, and when the sunspot number is high, the difference is negative. For the average sunspot number over the entire interval of about $\langle R_z \rangle = 75$, the difference is zero. All this suggests that we can rewrite the empirical relationship as

$$D_{year} = k\,(\langle R_z \rangle - R_{zYear}) / \langle R_z \rangle \qquad (6)$$

Where k is a calibration quantity approximately equal to 1.0 for *IHV* values calculated for Fredericksburg. Eq.(6) applies, strictly speaking, to yearly averages, not to daily values.

With our *D*-correction, we can now reconstruct once again *am* for the 1959-2000 interval and also plot the result and the "reconstruction" error in Figure 21:

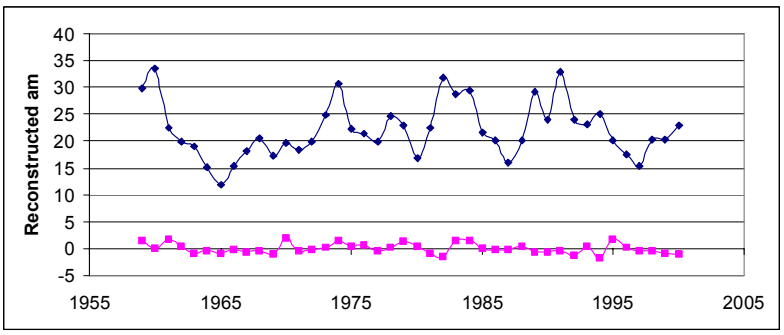

Figure 21. Reconstructed *am*-index for the interval 1959-2000 with the correction given by eq.(6) (blue diamonds) and at the bottom (pink squares), the discrepancy between the reconstructed and the observed *am*-values. All values are yearly means.

The sunspot number dependent (small) *D*-correction can be understood as a (slight) "spill-over" of the enhanced diurnal *Sq*-variation into the early and late night hours interfering with our assumption that there be no *Sq*-variation during the local night and adding to the variability of the field components and hence increasing the *IHV*-values.

Reconstructing the *ap*- and *aa*-Indices (Again)

Applying the *D*-correction to reconstructed values for *ap* and *aa* we get Figures 22 and 23. There is now no hint of a sunspot cycle modulation of the deviations from observed values:

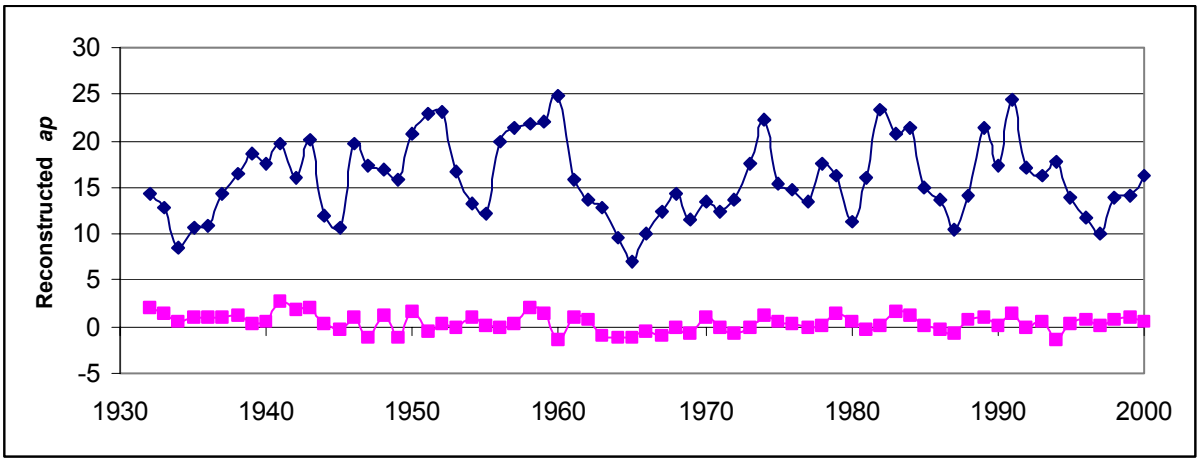

Figure 22. Reconstructed *ap*-index (blue diamonds) for the interval 1932-2000 using IHV_{FRD} with the sunspot number *D*-correction. There is no significant long-term trend in the reconstructed *ap*-values. At the bottom is shown the deviation from the observed values. Yearly averages are shown.

Although there are significant inter-cycle variations, no long-term trend is apparent in the reconstructed yearly averages of the *ap*-index going back to 1932. The deviations from observed values are overall small and insignificant. The assertion that the Sun's coronal magnetic field has doubled in the last one hundred years now hinges critically on what happened during the first 31 years of the 20th century.

The use of relationships discovered for the 1959-2000 internal to reconstruct indices for times prior to 1959 can be questioned. The above reconstruction of ap back to 1932 seemed to fare reasonable well, justifying the use method back to 1932. We'll now make the leap and assume that we can build on the success and use the method with *aa* even further back, to 1901.

The result is shown in Figure 23:

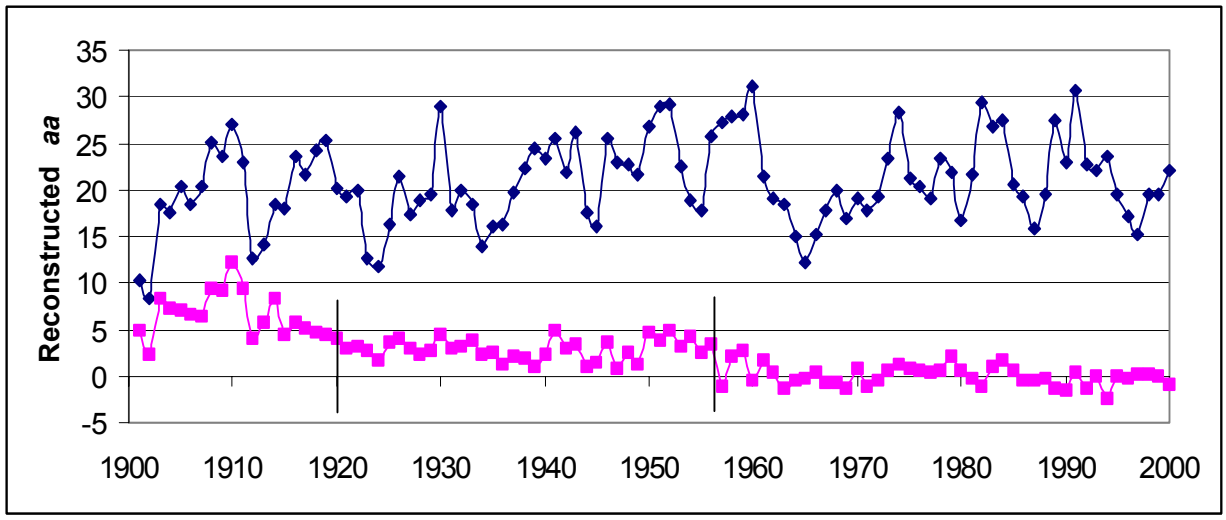

Figure 23. Reconstructed yearly averages (blue diamonds) of the *aa*-index (applying the *D*-correction) for the interval 1901-2000. At the bottom are shown the deviations from the "observed" values of *aa* (pink squares).

There is no significant long-term trend of the reconstructed *aa*-index. There *are*, however, disturbing structures in the deviations. One might divide the interval into three sub-intervals, from 1901 through 1919 (mean deviation +6.6), from 1920 through 1956 (mean deviation +2.6), and from 1957 to the present (mean deviation -0.1 or effectively nil). The dividing years, 1920 and 1957, are, of course, subject to debate and could be different by a few years. Whether or not there are trends *within* the sub-intervals could also be debated, but the data so far does not seem to support much speculation here.

Combining the reconstructed *ap* and *aa*-indices in one Figure gives us Figure 24:

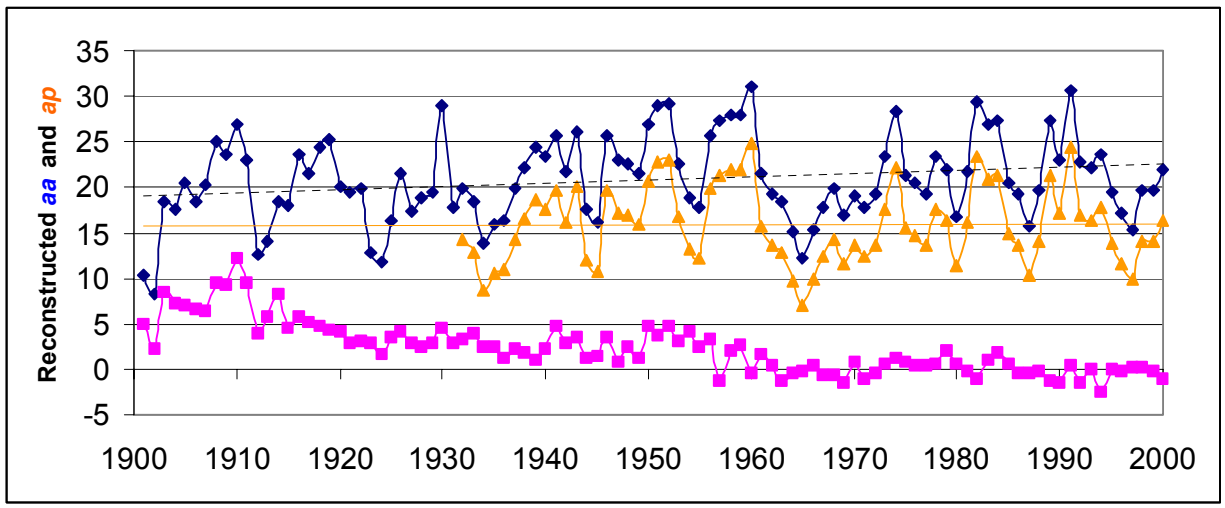

Figure 24. Reconstructed yearly averages of *aa* (blue diamonds) and *ap* (orange triangles). Also Shown are the deviations of reconstructed *aa* from "observed" *aa* (pink squares).

The possible (weak) trend in *aa* is not significant ($R^2 = 0.0457$). One could cling to the fact that 1901-02 values *were* very low, but that hardly justifies the claim that the Sun's field has doubled in the last 100 years (if based on *aa*).

What About Other Stations?

Basing a sweeping statement on data from a single station is always risky business. Several other geomagnetic observatories have been in operation for a hundred years or more. Reverting to the simplest possible situation we compare the (almost) raw *IHV*-index for several such stations in Figure 25:

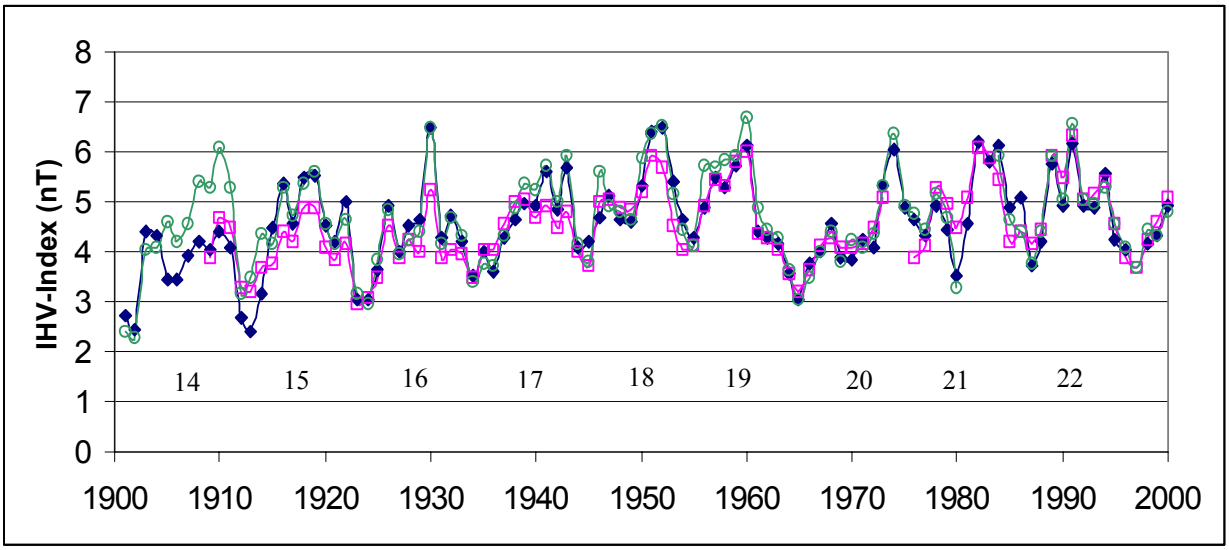

Figure 25. The *IHV*-index for Cheltenham/Fredericksburg (Geomagn. Lat. 51.8° N, green circles), Niemegk (Geomagn. Lat. 48.8° N, blue diamonds), and Tucson (Geomagn, Lat. 39.7° N, pink squares). To compare the indices, they have been normalized to have the same mean value during the 1984-96 interval. Other than that, no corrections or other dubious calibrations have been applied.

The agreement is fairly good with the exception that Cheltenham is significantly elevated during the years 1908-11. The reason for this discrepancy has not been investigated yet. It was good we checked and that we did not attach particular significance to this anomaly. There is little, if any, effect of the latitude of the stations. Fredericksburg and Tucson are on opposite sides of the *Sq*-vortex. It is clear that cycle 14 was a tiny cycle and that cycles 18 and 19 were great cycles, but apart from that, no secular trend seems to be called for. In particular, cycles 15 and 16 were not any less active than cycle 20.

Reconstructing the Solar Wind Speed

In this section we investigate the dependence of the solar wind speed, *V*, with the *IHV*-index and with the sunspot number (*Rz*). Figure 26 shows yearly averages of *V* as a function of *IHV* and of *Rz*:

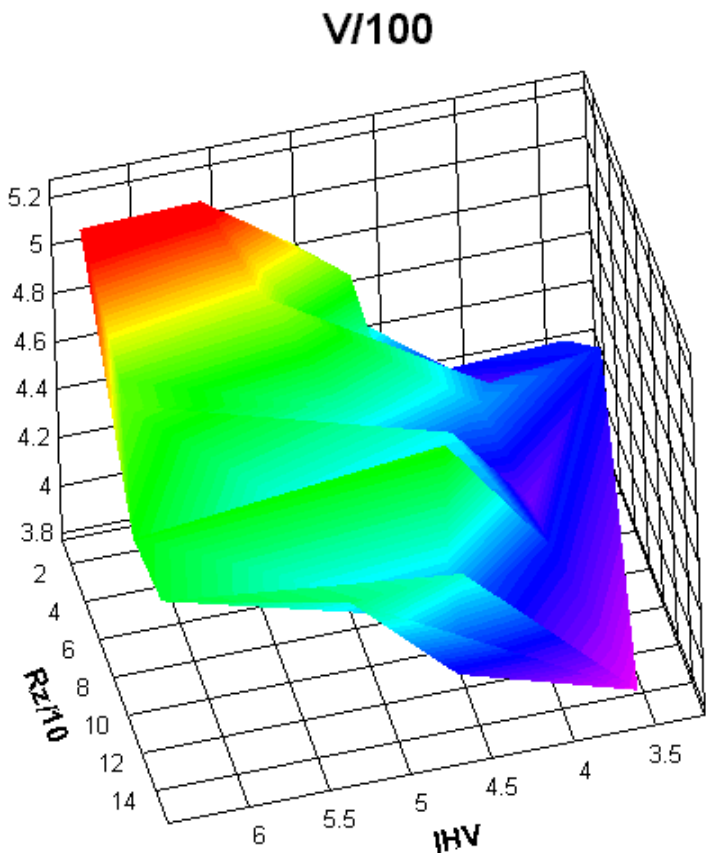

Figure 26. Yearly average values of *V* (plotted as *V*/100 km/sec) as a function of the *IHV*-index (nT) and of the sunspot number *Rz* (plotted as *Rz*/10) using data from 1964-1998.

A respectable relationship is apparent, with a summit in the upper left corner at high *IHV*-values and low sunspot number, with a valley in the lower right corner at low *IHV* and high *Rz*, and with foothills at intermediate values. Least-square fitting a sloping plane to the landscape yields:

$$V \text{ km/sec} = 308 + 31 * IHV - 0.20 * Rz \qquad (7)$$

The assumptions here are that

- The interplanetary field strength, *B*, is regulated by *Rz*
- The interplanetary field strength (or *Rz*) is not correlated with the wind speed *V*

As these assumptions are only approximately true, the systematic errors (caused by weak assumptions) overwhelm the random errors, so no error ranges or formal uncertainties make sense in eq.(7).We can gain an appreciation of the accuracy with which eq.(7) actually can reproduce the observed solar wind speed by comparing the calculated values with observations. This is done in Figure 27.

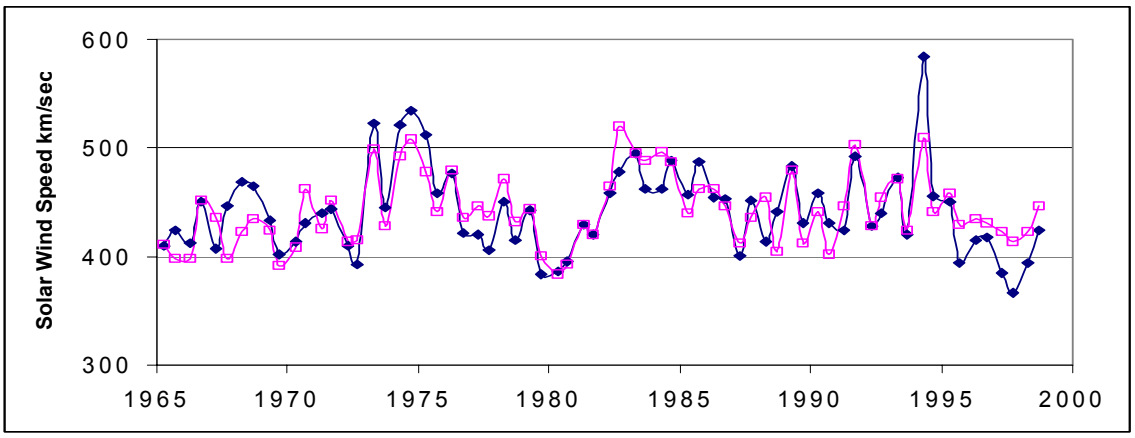

Figure 27. Reconstructed values of solar wind speed (pink squares) versus observed six-month averages (blue diamonds) for the interval 1965-1998.

Even on the shorter time scale of six months, eq.(7) does a pretty good job. Its greatest failings happen for some of the extreme values like in 1994, but the overall agreement is good. How good, can be seen in Figure 28:

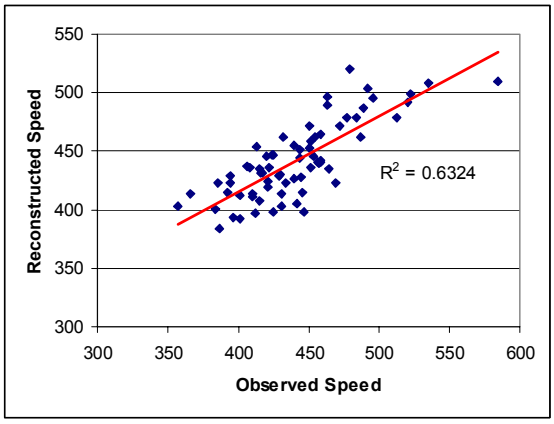

Figure 28. The Correlation between solar wind speed observed by spacecraft and the reconstructed solar wind speed for six-month averages during 1964-1998 is $R^2 = 0.6324$. Input to equation (1) was the raw *IHV*-index for Fredericksburg and the Zürich sunspot number and nothing else. By averaging over longer time-intervals (one-year, two-years, and on) the correlation coefficient increases, but the statistical significance decreases as we have a lesser number of degrees of freedom. (The ultimate is the complete loss of significance with the correlation going to one when we have only two points left…)

Using both Cheltenham/Fredericksburg and Niegmek (the latter multiplied by 0.8615) we can compare the reconstructed solar wind speed from two stations (and with the observed solar wind speed):

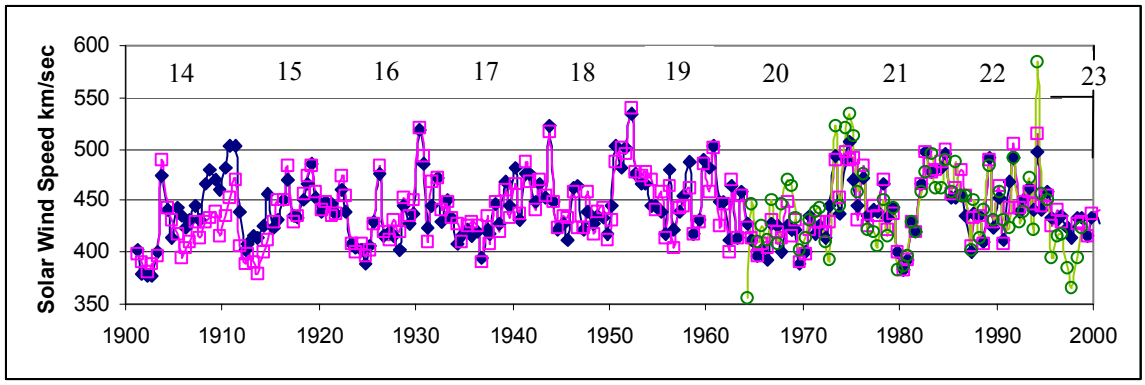

Figure 29. Reconstructed solar wind speed using CLH/FRD (blue diamonds), NGK (pink squares) compared with observed values (green circles) for six-month intervals during the 20th century.

Apart from the tendency to slightly higher values in the middle of the century, there does not seem to be any significant long-term trend. The extreme values are probably underestimated, but we shall make no attempt at guessing by how much. The high-speed streams during the declining phase of every solar cycle are clearly visible.

Stamper's Reconstruction

Lockwood *et al*. (1999) recognize and confirm the importance of the solar wind dynamic pressure as an input parameter to geomagnetic activity range indices of the *aa*-type as first pointed out by Svalgaard (1977). They include the dynamic pressure (the "ram" pressure) in a factor, *f*, isolating and expressing the solar wind speed dependence on geomagnetic activity as:

$$f = k(\theta)\,(nV^2)^{0.281}\,V = k(\theta)\,(nV^2)^{\omega}\,V \qquad (8)$$

Here n is the solar wind (mass) density, V is the speed, and $k(\theta)$ incorporates the IMF clock-angle which does not vary on time scales longer than one year. Lockwood *et al*. found the value of the exponent ω to be 0.281, while Svalgaard found 0.333. They go on to say: "the variation of [f] is dominated by that of the solar wind speed V. the annual mean of V rises in the declining phase of solar cycles because the Earth repeatedly intersects fast solar wind streams from low-latitude extensions of coronal holes. These occur every 27 days and so also raise the geomagnetic recurrence index, I. Hence we expect f and I to increase together in the declining phase of the sunspot cycle. **However, I can remain high at sunspot minimum (whereas V is lower) because *aa* values are low and relatively constant**. Hence we adopted a relationship for a predicted f of the form:

$$f_p = s\,I^{\beta}\,aa^{\lambda} + c \qquad (9)$$

where the exponents and constants are found from a regression fit". With the values ($\beta = 0.263$, $\lambda = 1.303 \approx 5\beta$) found by Lockwood *et al*., we can rewrite eq.(9) as:

$$f_p = s\,((I{\cdot}aa^5)^{\beta} + 72.6) \qquad (10)$$

The constant term ($c/s = 72.6$) is needed because while I does go to zero at times, V does not. The quantity I is Sargent's Recurrence Index (Sargent 1986). It would have been nice if the recurrence index would be a good proxy for the solar wind speed. It is however not, as acknowledged by Lockwood *et al*. in the emphasized statement quoted above. To compensate for this, they now introduce the quantity "$(I{\cdot}aa)^5$" with what we would consider to be a sleight of hand. The argument is important and subtle, so we shall be somewhat pedantic. I can be high even if *aa* is not high. This is taken to mean that there is no fast stream after all, even if activity is recurrent. We can make the interpretation of all this explicit and clearer with a table:

Recurrence	**Activity**	**Interpretation**
Recurrence index I is low	Activity *aa* is low	Nothing going on. Speed V is low
Recurrence index I is low	Activity *aa* is **high**	Transient blast. Speed V is **high**
Recurrence index I is **high**	Activity *aa* is low	Sector structure. Speed V is low
Recurrence index I is **high**	Activity *aa* is **high**	In high-speed stream. Speed V is **high**

The *aa*-index enters into the expression for the predictor f_p with such a large exponent (to the 5th power) that, in the end, the *aa*-index effectively controls whether or not we think we have a high-speed solar wind. This is not new, of course, authors too numerous to mention (*e.g.* Svalgaard 1977, $am \sim V^{2.25}$) have shown that *aa* (or *ap* and *am*) vary with at least the square of the solar wind speed, V, but only linearly with the interplanetary magnetic field, B, so the wind speed usually wins. But not always; high *aa*-values can also be caused by high interplanetary magnetic fields (which can be recurrent too), so we have not achieved a clear-cut separation of B and V, however plausible the above interpretation is. Although there is little doubt that the very strong peaks in *aa* (e.g. in 1973 and 1930) were caused by strong high-speed solar streams, it is somewhat more difficult to a put solid number to what is due to V and what is due to B, and even more so if we consider times with moderate activity. The suggestion by Lockwood *et al*. has merit as a step on the way, but we have to watch out for subtle quantitative effects of taking our assumptions too far.

The Recurrence "Background"

Figure 29 shows the variation over time of the Sargent Recurrence Index during the interval 1868-1997. The maximum value of *I* is fairly constant with time being around 0.4, but there is a significant "background" level covering most of the solar cycles at times away from the "stream spike" during the declining phase of the cycle. Moreover, the background level *does* show a pronounced difference between the first half of the interval (up through about 1922 - being on the average 0.18) and the second half (from 1923 on - being on the average 0.07):

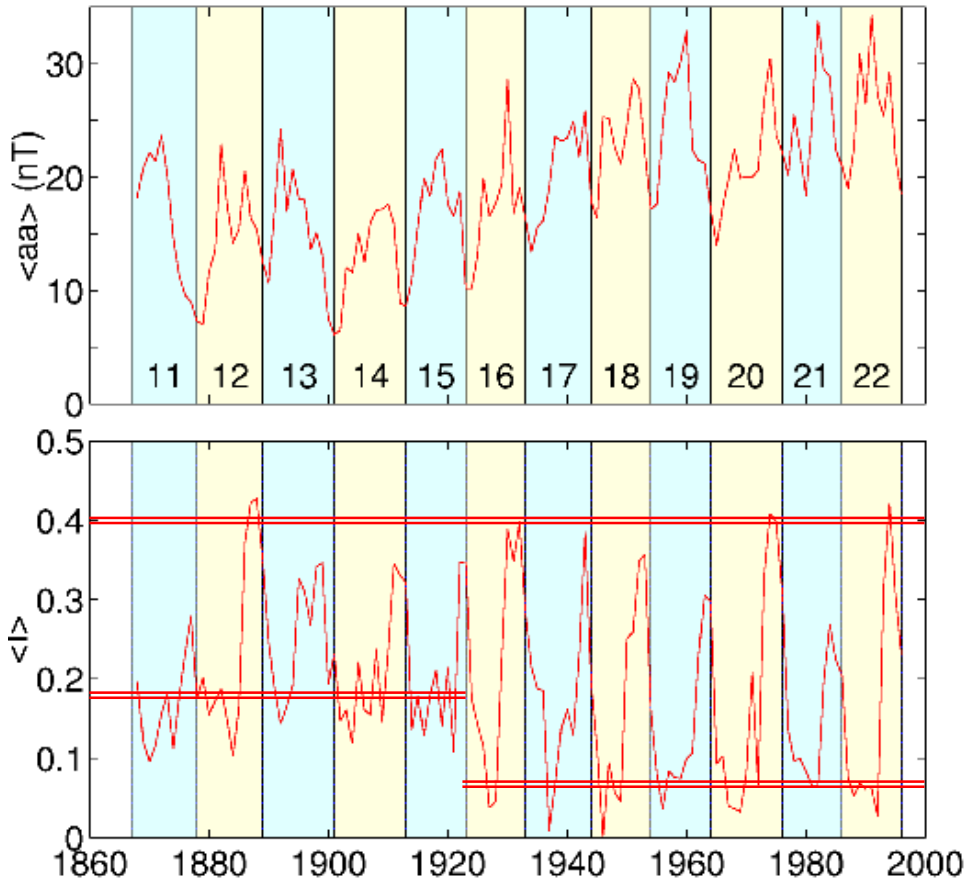

Figure 30. Yearly means of the *aa*-index (top) and of Sargent's recurrence index (bottom). (After Lockwood *et al*., 1999). The sunspot cycles are indicated by numbers and by alternating different shading. The maximum recurrence level is indicated as well as the different "background" levels before 1920 and after about 1935.

In view of the (up to now) unexplained difference in background level it may not be correct to assume that the relationship between solar wind speed, recurrence index, and the *aa*-index that was constructed based on data after 1964 would be valid for years before 1920. But let's for now stay with the assumption that eq.(10) is valid and see where that leads us. We look at the years away from the clear high-speed stream for cycles 14, 15, and 22, 23. We can apply eq.(10) for each interval as shown in the following table:

Cycle and years	**Average *aa***	**Background *I***	***V* calculated**	***V* change**
Min 1901	6.11	0.18	359.5 km/sec	-82.8 km/sec
14 (1901-09)	12.70	0.18	391.2	-51.1
15 (1913-20)	16.91	0.18	413.5	-28.8
22 (1987-93)	26.54	0.07	442.3	***V* observed**
23 (1996-2002)	20.58	0.07	414.2	422.2

Assuming constant density n, we derive the following ratio from eqs (9) and (10):

$$W = f / f_p \sim V^{1.561} / ((I \cdot aa^5)^{0.263} + 72.6)$$

If f_p is an accurate predictor of f, W should be constant and close to unity (with statistical significance 99.999999957% according to Lockwood *et al.* ☺), so that

$$V_{calc} = K ((I \cdot aa^5)^{0.263} + 72.6)^{0.641} \qquad (11)$$

Where K (= 21.76) is calculated such that V_{calc} matches the observed solar wind speed V_{obs} = 442.3 km/sec for the non-stream years in cycle 22 where we have spacecraft data. The solar wind speed according to the above considerations would seem to have been 30-50 km/sec slower during the first two cycles of the 20th century. This rather large difference might be observable in comet tail aberration data or in the magnitude of *sudden commencements*. Our reconstructed solar wind speeds for these early cycles were not significantly lower than for recent cycles.

Another way to investigate long-term trends that largely avoids the solar wind speed problem is maybe to *exclude* years with high recurrence index.

The Bare-Bones Lockwood-Stamper Formula

If we remove the overgrowth of Physics that obscures the essential features of the formula that Lockwood *et al.* give for the magnitude of the interplanetary field, we are left with:

$$B \text{ nT} = B_0 \, [aa / ((I \cdot aa^5)^{0.263} + 72.6)]^{1.295} \qquad (12)$$

Where the single constant B_0 (= 46.85 nT) subsumes the various other constants and calibration factors. This lays bare the numerology all such reconstructions really are. We should remember that all values are to be taken as *annual* means and that eq.(12) is not applicable to shorter intervals, such as 3-hourly values or daily averages.

Sargent's Recurrence Index

Lockwood *et al.* used Sargent's recurrence index, which they say is defined for the jth 27-day Carrington (We believe they mean "Bartels") rotation as

$$I_j = (1/13) \sum_{k=-6}^{k=+6} c(j+k, j+k+1) \qquad (13)$$

where c is the correlation coefficient between two consecutive 27-day intervals of *aa* values. Defined in this way, the index is really a running mean over 13 rotations (about one year). Now, on top of this, Lockwood *et al.* also compute annual means of I, which they denote as $\langle I \rangle$. It is not quite clear to me how to compute $\langle I \rangle$, as a year is not an integral number of 27-day Bartels rotations. We assume that $\langle I \rangle$ and I are not significantly different since I is already a running (almost) yearly mean. Figure 31 shows our calculated values of I.

As is clear from eq.(13) and also from our Figure 31, the I-index can be negative at times (and it is), yet Figure 29 (from Lockwood *et al.* 1999) does not show any negative values. Indeed eq.(9) does not *allow* negative values for I (what is $(-1)^{0.263}$?), so we are missing a definitional subtlety somewhere. Although our values show relatively good general agreement with Figure 29, there are considerable differences in the details (including the negative values). Sargent's original paper also shows negative values.

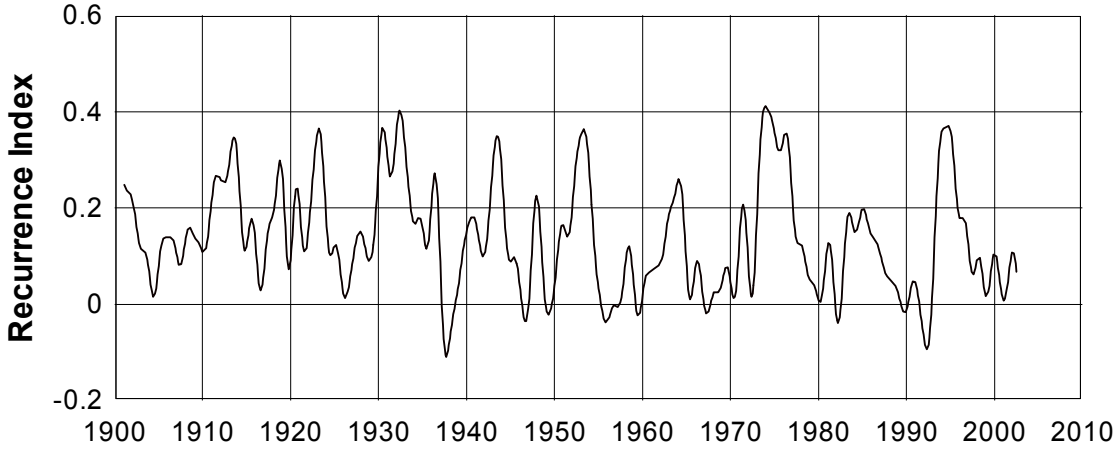

Figure 31. The Sargent *I*-index computed from eq.(13) for the interval 1901-2002. This is essentially a yearly running mean of the cross-correlation between adjacent 27-day rotations.

The raw correlation coefficients for each rotation are also interesting:

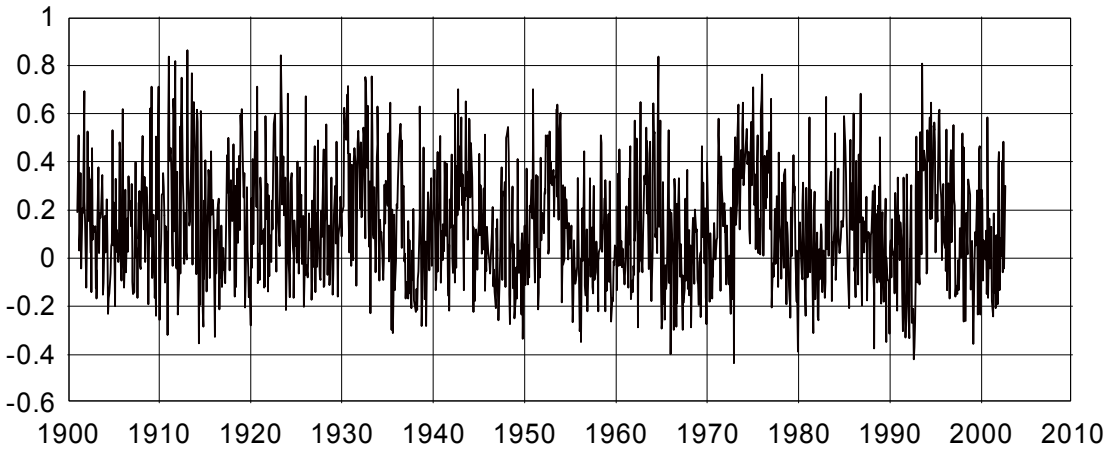

Figure 32. The individual 27-day rotation correlation coefficient values for the interval 1901-2002.

Correcting for the Semiannual Variation

The semiannual variation of geomagnetic activity (Sabine 1856, Russell and McPherron 1973, Svalgaard *et al*. 2002) is a modulation of activity controlled by the *geometry* of the interaction between the solar wind and the Earth's magnetic field. As such, the Sun's magnetic field should not show any semiannual variation.

It is instructive to consider what difference the semiannual variation of geomagnetic activity indices makes for the long-term trends under investigation. First we show (Figure 33) that for the observed solar wind speed and interplanetary magnetic field strength, there is no significant semiannual variation apart from a small axial effect that may be present. The axial effect is what would result from a variation of solar wind properties with the Earth's latitude with respect to the Sun's equatorial plane. Within the 7.25º latitude excursions during the year no firm observational evidence has been found for latitudinal dependencies on intrinsic solar wind properties, although for much larger excursions it is clear that the fashionable slow/fast solar wind dichotomy exists.

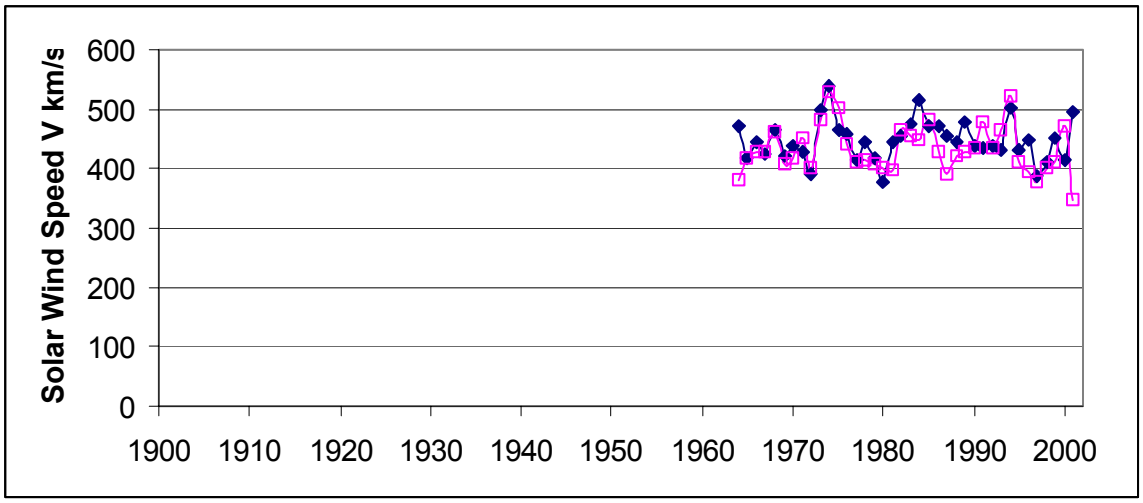

Figure 33a. Observed solar wind speed for equinoctial (blue diamonds) and solstitial (pink squares) months. Equinoctial months are March, April, September, and October. Solstitial months are June, July, December, and January. No long-term trend is evident.

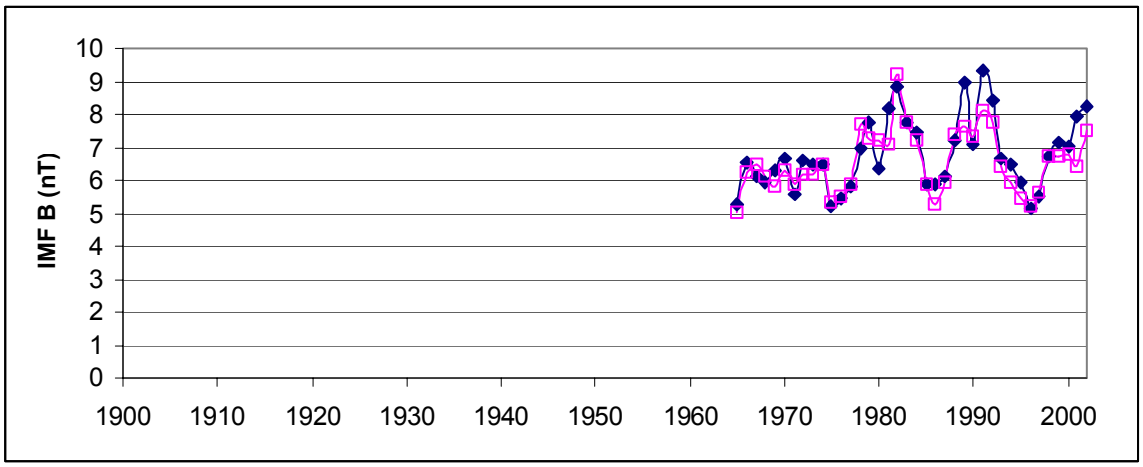

Figure 33b. Observed interplanetary magnetic field magnitude for equinoctial (blue diamonds) and solstitial (pink squares) months. As usual, we shall remark that there is no long-term trend not explainable mainly by higher sunspot numbers away from the (constant) minimum values.

In Figure 33b we *may* see an indication of axial enhancement of the magnetic field strength around 1990, but no systematic difference between the interplanetary magnetic field strengths during equinoctial months (MAR, APR, SEP, OCT) and solstitial months (JUN, JUL, DEC, JAN) is evident. Values of V and B for intermediate months (MAY, AUG, NOV, FEB) are also not systematically different from those during equinoctial and solstitial months. All this is not surprising and should need no further elaboration.

On the other hand, plotting the *IHV*-index, Figure 34 reveals (as we must find) a significant and interesting difference between equinoctial and solstitial months. The semiannual variation shows itself as solstitial values being lower than equinoctial values. There is significant information in the details of this comparison, as we shall see shortly. We here remind the reader that the semiannual variation has three components: a weak "axial" component, a somewhat stronger "excitation" component caused by the Russell-McPherron effect (Russell and McPherron, 1973), and a dominant "suppression" component related to the angle between the solar wind and the Earth's magnetic dipole axis (Svalgaard *et al*., 2002).

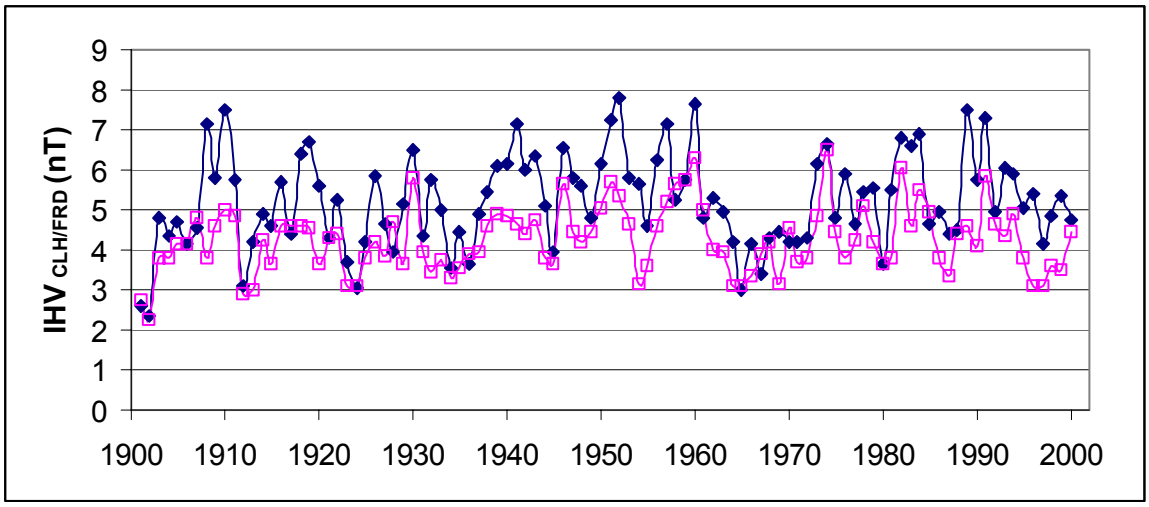

Figure 34. *IHV*-index (Cheltenham/Fredericksburg) for equinoctial (blue diamonds) and for solstitial months (pink squares).

Note how the solstitial values at sunspot minima show little variation with time. In particular, we note how the minimum years 1954 and 1996, which show a strong Russell-McPherron excitation (Cliver et al. 2003), relax towards the overall base level for other sunspot minima. The large solar wind streams in 1974 affected activity regardless of season.

Difference between the "Official" Mayaud *aa*-Index and "Our" *aa*-Index

Because the *IHV*-index is defined on only a few hours each night, we have compared it with *aa* defined on those *same* hours (the first two 3-hour periods from 00 UT to 06 UT). This creates a small (but in some respects - as we shall see - important) difference between the "official" *aa*-index defined on all eight 3-hour intervals and our subset defined on only the first two intervals. Figure 35 shows their close kinship:

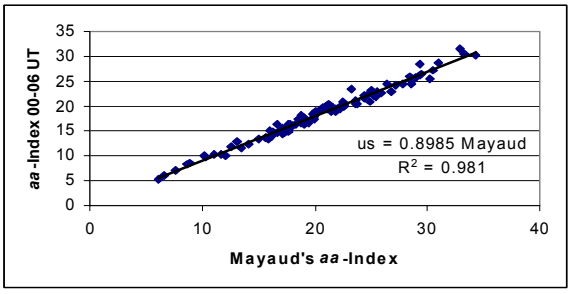

Figure 35. Relationship between the subset of the *aa*-index that is defined using only the first two 3-hour intervals and the full set using all eight 3-hour intervals. The correlation is very tight with $R^2 = 0.981$. Our values are slightly lower:

Our ***aa*** = 0.8985 * Mayaud's
Mayaud's = 1.113 * Our ***aa***

It is also of interest to see the long-term relationship between the official *aa*-index and our subset:

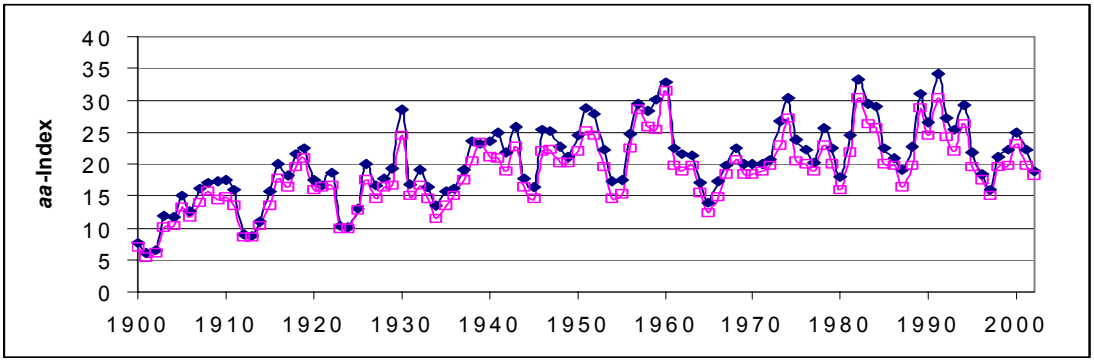

Figure 36. Mayaud's *aa*-index (blue diamonds) compared to our *aa* values (pink squares) for yearly means through the interval 1901-2002.

Figure 37 shows monthly averages of the original *aa*-index plotted such that for each month we show all eight 3-hour intervals. This allows us to see the systematic differences caused by *aa* not being derived from several stations equally spaced in longitude:

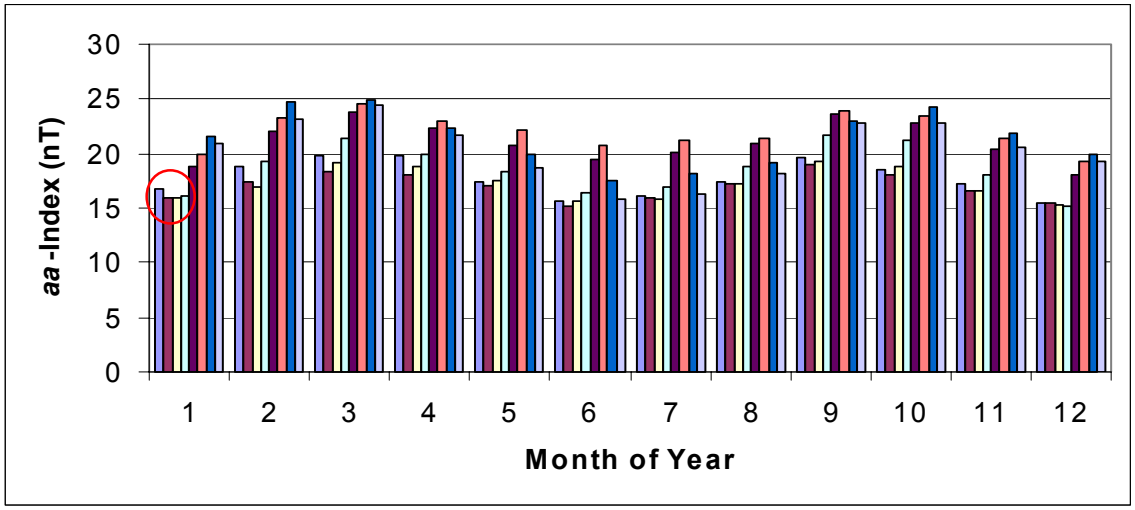

Figure 37. Monthly averages of the official *aa*-index for each of the eight 3-hour intervals within each month, (all years 1868-2001).

We emphasize that the UT-variation seen here is largely an artifact. The *aa*-index should not be used in its original form to study UT-variations of geomagnetic activity. Although Svalgaard *et al.* (2002) have used the *am*-index to compute calibration factors to remove the artificial UT-variation, we shall here use the uncorrected time series. In Figure 37 we see the semiannual variation very clearly and also note how the first two 3-hour intervals are systematically too low. This explains the difference seen in Figure 36. The (artificial) UT-dependence is rather large, 25-30%, and is the rationale behind our decision only to work with *aa* values for the actual hours over which the *IHV*-index is defined. The artificial UT-variation is such that the second half of the UT day is "more" active (have higher *aa*-values) than the first half. If the *aa*-index were homogeneous we would expect this imbalance to be constant over time. Figure 38 shows that this is not so:

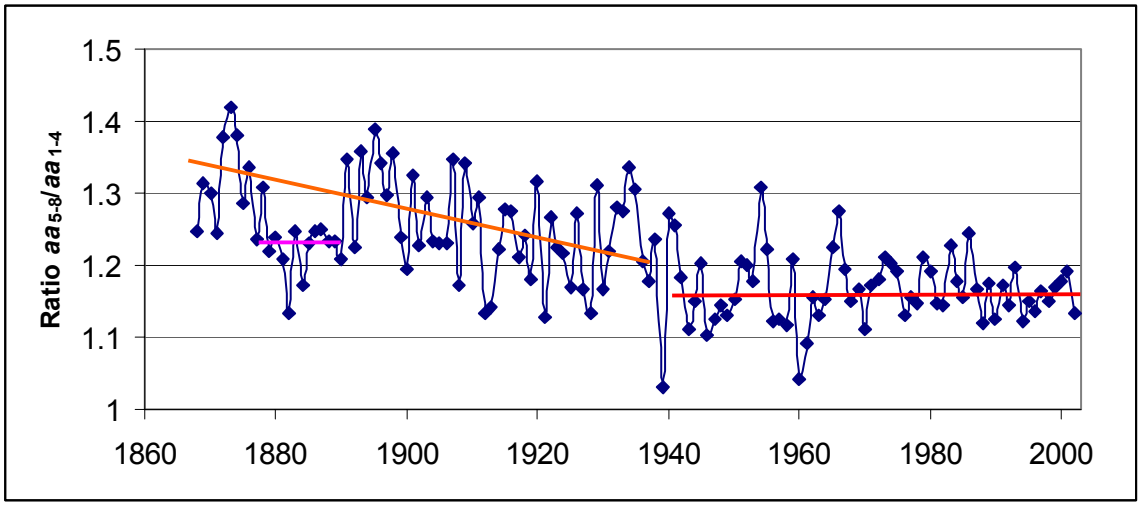

Figure 38. Yearly ratios between *aa* for the last four intervals (12-24 UT) of the day and for the first four intervals (00-12 UT).

After about 1969 the ratio fluctuates less, but in any case seems to stay fairly constant (average 1.163) back to about 1940 preceded by a general rise towards the beginning of the series (1.35, an increase of 16%). If we interpret the rise to be due to a decrease of the denominator (and hence also the mean value), we see a first hint of a crack in the armor of *aa*-homogeneity.

The Rosenberg-Coleman Effect

The Rosenberg-Coleman effect (Rosenberg and Coleman, 1969) is the tendency for the polarity of the interplanetary magnetic field observed at the Earth to be unbalanced on either side of the heliospheric current sheet. For six months centered on March/September the Earth tends to be north/south of the current sheet more often and the polarity of the Sun's northern/southern magnetic pole tends to be observed more often. This effect is only observed when the Sun's polar fields are strong (at or just after sunspot minimum) and disappears as the polar fields disappear. Figure 39 shows the effect nicely in the variation of 27-day Bartels rotation averages of the radial component, B_x, of the IMF:

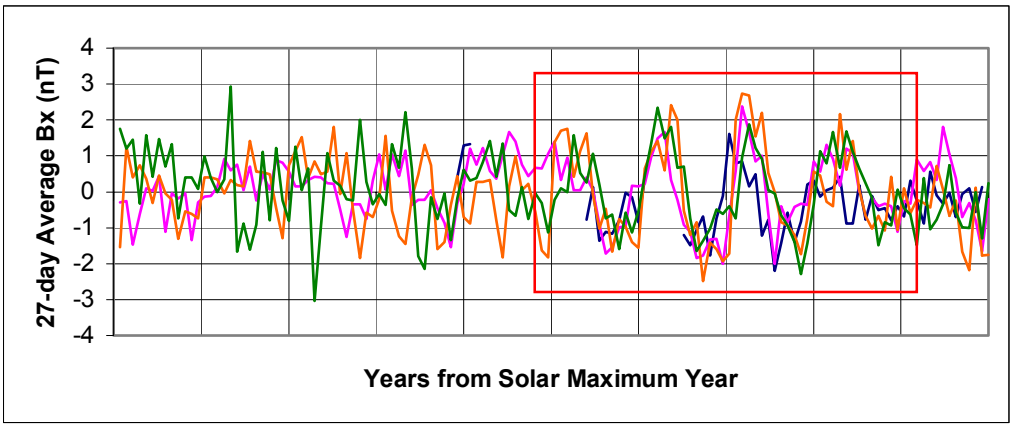

Figure 39. 27-days averages of the IMF radial component, B_x, plotted as a function of the time since sunspot maximum (taken formally as January 1st in 1960, 1970, 1980, and 1990). For the intervals 1960-69 and 1980-89, values are plotted with the opposite sign (because the Sun's magnetic field switches polarity at or just after sunspot maximum). The yearly variation of the dominant polarity manifests itself as a similar variation of the rotation averages. The effect is pronounced for the years shown in the red box (the four years falling 5 to 8 years after cycle maximum). The color code is: 1960-69 (blue), 1970-79 (pink), 1980-89 (orange), and 1990-99 (green).

The Russell-McPherron effect (Russell and McPherron 1973) gives rise to a very small semiannual variation of geomagnetic activity during times when there are equal amounts of both IMF polarities. A strong Rosenberg-Coleman effect can, however, create such an imbalance between the two polarities, that the Russell-McPherron effect results in a significant enhancement of the semiannual variation, leading to a 22-year cycle of geomagnetic activity (Russell and Mulligan, 1996; Chernosky 1966). This effect causes some minima years (*e.g.* 1954 - see Figure 40 - and 1996) to have higher geomagnetic activity during the equinoxes than others without requiring a higher solar magnetic flux. These purely geometrical factors must be eliminated when trying to infer properties of the Sun. By using only solstice data (see figure 34), we remove the combined effect of the Rosenberg-Coleman and Russell-McPherron effects.

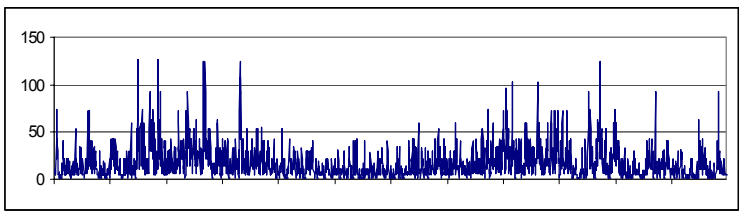

Figure 40. The *aa*-index in the sunspot minimum year 1954 showing the strong semiannual variation that results from an extreme Rosenberg-Coleman effect.

Including Stations De Bilt, Witteveen, Wingst

The Dutch stations De Bilt, Witteveen, and Wingst form a series going back to 1903. The station coverage is detailed in the following table:

Station	IAGA Abbrev.	Geog. Lat.	Geog. Long.	Geom. Lat.	Geom. Long	Operating Interval	Calibration factor
De Bilt	DBN	52.1 N	5.2 E	53.4 N	90.6 E	1903-1938	0.748
Witteveen	WIT	52.8	6.7	53.8	92.4	1938-1984	0.867
Wingst	WNG	53.8	9.1	54.2	95.2	1943-present	1.000

The calibration factor is what is needed to reduce the station's *IHV*-value to that of Wingst, *i.e.* $IHV_{Wingst} = 0.867\ IHV_{Witteveen}$. The factors are derived from intervals of simultaneous operation, either directly or via an intermediate station. Figure 41 show the result of including these stations. It is becoming clear that the values from Cheltenham during 1908-11 are simply too high as the other stations all agree on lower values:

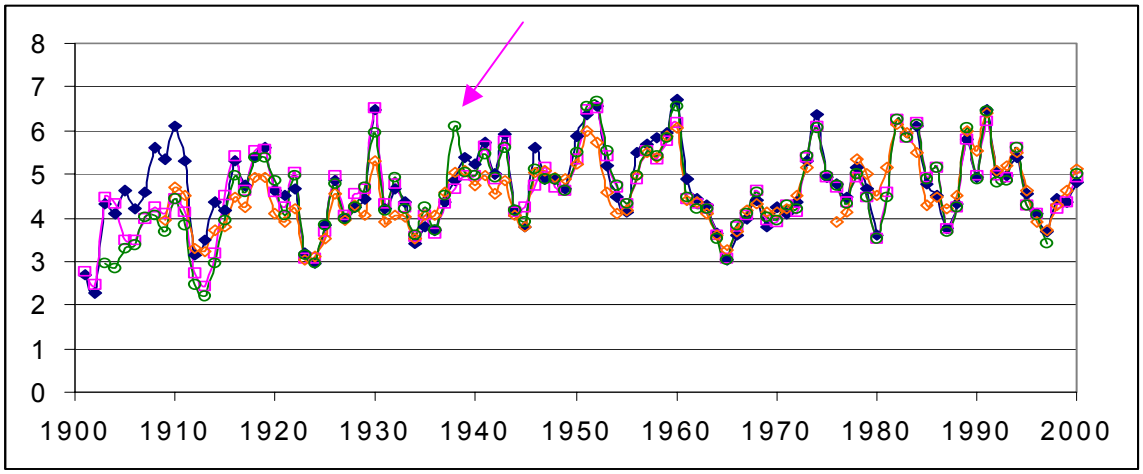

Figure 41. The *IHV*-index derived from Cheltenham/Fredericksburg (blue diamonds), Niemegk (pink squares), Tucson (orange diamonds), and De Bilt/Witteveen/Wingst (green circles). All stations have been normalized to have the same average value during the intervals 1950-73 and 1984-97. Yearly averages are shown.

The value for De Bilt for 1938 is anomalously high (arrow in Figure 41). This is caused by a slightly different UT-interval (so it is included for some stations not for others) for the nighttime in combination with the famous geomagnetic storm on January 25th, 1938, where the disturbance field exceeded 1000 nT (see Figure 1 in section 6.13 of Chapman and Bartels (1940)). Over the whole 100-year interval there has only been a handful such storms. These extremes are in a sense filtered out of the *aa* (and related) indices by the fact that they are all accorded a *K*-figure of 9, which is then translated back into a range of 400 nT. We could impose a similar cut-off, but I have chosen not to, in keeping with the simplicity of the method.

A separate investigation should be pursued to understand the situation in 1908-11, but for the present purpose this is hardly necessary as the other stations show reasonable agreement through the century including the interval 1908-11. We shall thus conclude that the *IHV*-index from many stations (including the *aa*-stations) does not support the claim of a doubling of the Sun's coronal magnetic flux, while not denying that there are significant changes from cycle to cycle, some solar cycles (*e.g.* #19) having more flux than others (*e.g.* #14).

High-Resolution Examples of the Daily Variation

As an aid to interpretation of the Inter-Hour Variability index, we offer Figure 42:

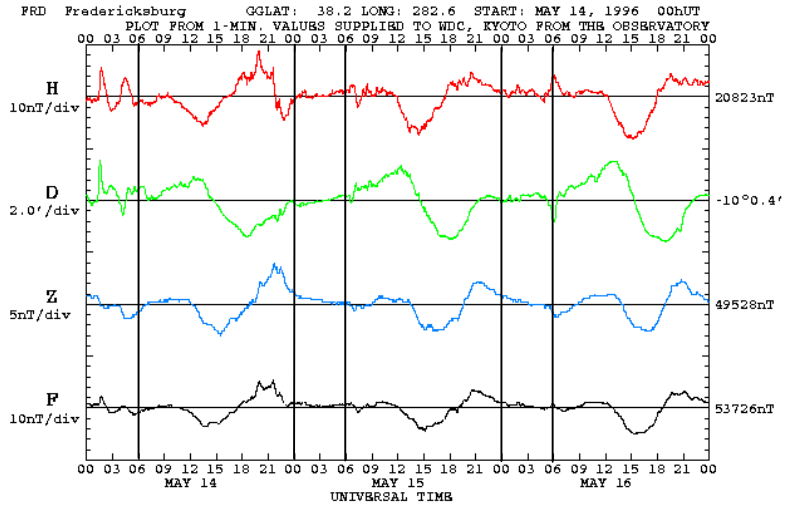

Figure 42. Traces of the magnetic elements on several consecutive days at Fredericksburg during the minimum year 1996. The daily *Sq*-variation is clearly seen. During local late night, the Sq-variation vanishes. We consider variations at such times to be pure geomagnetic activity controlled by the solar wind.

Three six-hour intervals are shown on three consecutive days. The first interval is clearly disturbed, while the second interval is much quieter, and the third interval is moderately disturbed. These example days are very typical and thousands of similar examples could be given. It is also apparent that an even "better" activity index might be obtained if calculating the nighttime variability from finer resolution data, *e.g.* from 1-minute values rather than 1-hour values.

Derivation of the *aa*-Index

The *aa*-index is derived from hand-scaled magnetograms from two almost antipodal observatories (at invariant magnetic latitudes of approximately 50˚); one in the United Kingdom and the other in Australia (Mayaud, 1980; Menvielle and Berthelier, 1991). The subauroral observatories that have contributed to the derivation of the *aa* index are shown in the following table:

Name	IAGA Abbreviation	Geog. Lat.	Geog. Long.	Geom. Lat.	Geom. Long	K=9 Limit	Weight	Used
Greenwich	GRW	51.5	360.0	53.8	85.1	?	1.007	1868-1925
Abinger	ABN	51.2	359.6	53.6	84.6	520	0.934	1926-1956
Hartland	HAD	51.0	355.5	54.2	80.3	530	1.059	1957-…
Melbourne	MEL	-37.8	145.0	46.6	222.3	?	0.967	1868-1919
Toolangi	TOO	-37.5	145.5	46.1	222.7	510	1.033	1920-1978
Canberra	CNB	-35.3	149.0	43.4	226.1	420	1.084	1979-…

For each 3-h interval (00:00-03:00, 03:00-06:00, etc.), *K* indices (Bartels et al. 1939) are derived for the two antipodal observatories. The K index is a quasi-logarithmic number between 0 and 9 that is assigned to the middle end of these specified 3-h intervals. It is derived by measuring the maximum deviation [in nanotesla (nT)] of the observed field from the expected quiet-time level, for each of the three magnetic-field components. The largest of the three maxima at each observatory is converted to a K index by using a look-up table appropriate to that observatory. The K indices measured at the two antipodal observatories are then converted back into amplitudes and an individual aa index is the average of the two amplitudes, weighted to allow for the small difference in latitude of the northern and southern observatories, or for the slight changes in the locations of the two antipodal observatories. Since an individual three-hourly value of the aa index is derived from just two K indices, it provides only an approximate indication of the actual level of planetary geomagnetic activity. However, half-daily, or daily, averages of the aa index give an acceptably accurate indication of geomagnetic activity on a global scale and over a significantly longer time-interval than any other index of geomagnetic activity. Mayaud (1973) published a unique 100-year (1868-1968) series of three-hourly aa indices, which is based on K indices that were all measured by the author himself to ensure the homogeneity of the time-series. The aa indices have continued to be published

subsequently, in order to provide a rapidly available world-wide index of geomagnetic activity that is physically meaningful on a half-daily or daily time-scale.

The *aa* index comes from a request made by the Royal Society of London at the Madrid IAGA in 1969 by S. Chapman for C_i before 1884. Early records from Greenwich and Melbourne provide a long series of data. Where early individual data records are missing, other stations provide approximate data instead.

The *IHV*-Index for the "*aa*" Stations

Hourly values are not (readily) available for Greenwich and Melville. For the replacement stations hourly values are available from about 1925 onwards. We can construct *IHV* indices for all stations used in the derivation of *aa*. The result is shown in Figure 43.

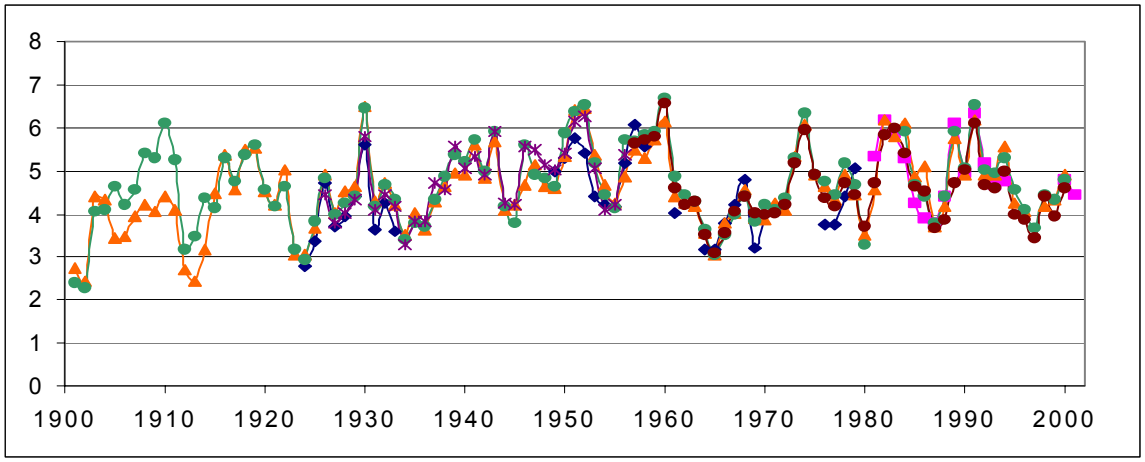

Figure 43. The IHV-index for several stations: Toolangi (blue diamonds), Canberra (pink squares), Hartland (brown circles), and Abinger (purple crosses). In addition, Cheltenham-Fredericksburg (green circles) and Niemegk (orange triangles, adjusted as in Figure 25).

The four "*aa*"-stations (Greenwich and Melbourne were not plotted due to lack of data) show substantial agreement. No adjustments of the data of any kind have been performed. The same lack of a long-term trend as we found before (e.g. Figure 21) is apparent. Figure 44 is a less cluttered version of Figure 43:

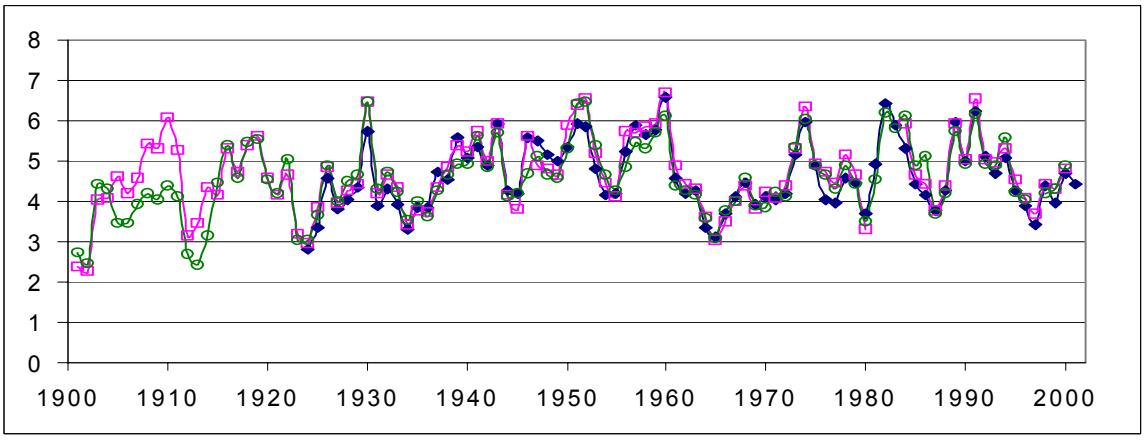

Figure 44. The *IHV*-index for the *aa*-stations (blue diamonds), CLH/FRD (pink squares), and Niemegk (green circles) since 1901. *IHV* is calculated for the H (or X) component.

In constructing the single *IHV* value for the "*aa*"-stations, we simply averaged the individual *IHV* values for each day over all *aa*-stations that had data for that day (most of the time we would have two such values, one from the

northern hemisphere and one from the southern hemisphere). No weighting or other calibration was performed. No systematic differences between simultaneous *IHV* values are apparent (apart from the interval 1908-11 during solar cycle 14). This result eases the concern caused by possible effects from slow changes in auroral zone location or geomagnetic pole drifts (*e.g.* Clilverd *et al.* 1998).

Using the D-component

It is generally accepted (reference?) that disturbances of the nighttime D-component (or the Y-component) at midlatitude stations are caused by field-aligned currents during substorms. Calculating the IHV^{Y}-index from the Y-component could then give us a measure of substorm activity. Let us first remind the reader of the basic relationships between the components

$$Y = H \sin D \qquad \text{differentiating we get:}$$
$$\Delta Y = \Delta H \sin D + \Delta D\, H \cos D \qquad (13)$$
$$\Delta Y \approx \Delta D\, H \cos D \; = X \cdot \Delta D \qquad (14)$$

Equation (14) follows from (13) when H is much larger than ΔH and D is small For Niemegk, where recent values of the components are about $D = 1°\ 20'$ and $H = 18700$ nT, we have

$$\Delta Y_{NGK} = \Delta D_{NGK}(\text{in 0.1 arc minutes}) / 34377 * 18700 * 0.9997$$
$$\Delta Y_{NGK} = \Delta D_{NGK}(\text{in 0.1 arc minutes}) \cdot 0.5437 \text{ nT} \qquad (15)$$

Figure 45 shows the result for Niemegk:

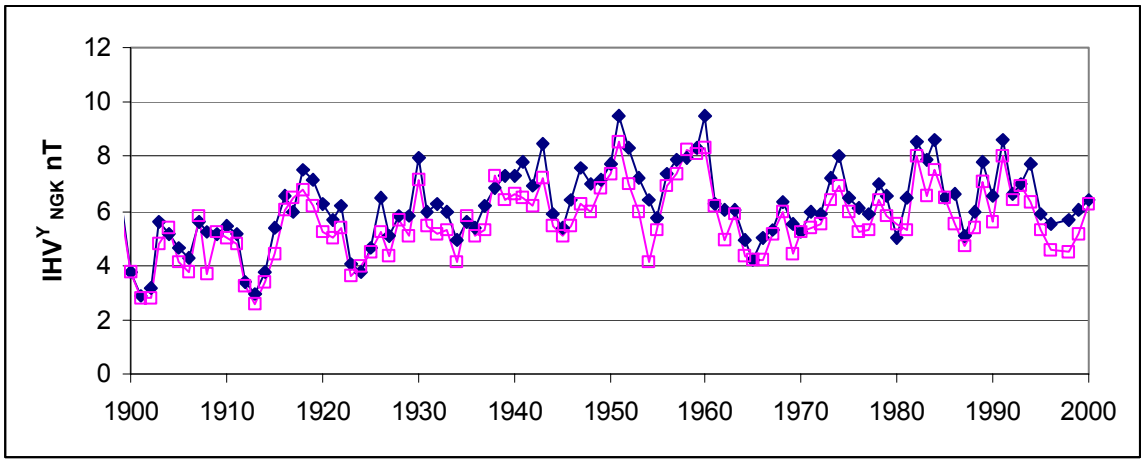

Figure 45. The *IHV*-index for the *Y*-component at Niemegk (all year: blue diamonds; solstices: pink squares).

We see the familiar picture**:** a very low cycle 14, but thereafter no long-term trend. Note, how closely the minima for the solstices line up with 1945 being a notable exception.

Direct Measurements of the Sun's Magnetic flux Show No Increase

Hildner *et al.* (2000) point out that the claim that the Sun's coronal magnetic field has doubled in the past 100 years depends crucially upon the inferences that 1) interplanetary measurements at the Earth can be used to characterize the entire heliosphere, and 2) stronger interplanetary fields imply stronger solar fields.

Line-of-sight photospheric field measurements from both Mt. Wilson (MWO) and Wilcox Solar (WSO) Observatories are available for approximately the last three and two solar cycles, respectively. (Data from MWO are available from ~1967 to present and WSO data from ~1976 to present). The data from both observatories agree well with one another during the ~24 year interval for which there are mutual observations. Their analysis of these data shows no increase in the solar magnetic flux over the last three solar cycles in contradiction to that deduced by Lockwood *et al.* Nor does the Sun's open flux, the magnetic field escaping into interplanetary space, show a secular

increase over the last three cycles, according to calculations of the flux reaching a spherical coronal ``source surface" at a height of 2.5 R⊙. They believe that the Sun's average coronal magnetic flux is not rapidly increasing now. They further believe that near-Earth interplanetary measurements can only poorly, if at all, be generalized to characterize the entire heliosphere, and that the strength of the interplanetary magnetic field is not simply related to the overall solar magnetic field strength. It would seem that our analysis makes it unnecessary to make such qualifications.

Cosmic Ray Muons Show No Increase of the Sun's Magnetic Flux

Stozhkov *et al*. (2001) derive a tight relationship between the muon flux measured with ionization chambers and the IMF field strength and reconstruct *B* back to 1937 (Figure 46). No increase since then is apparent in concordance with the present study:

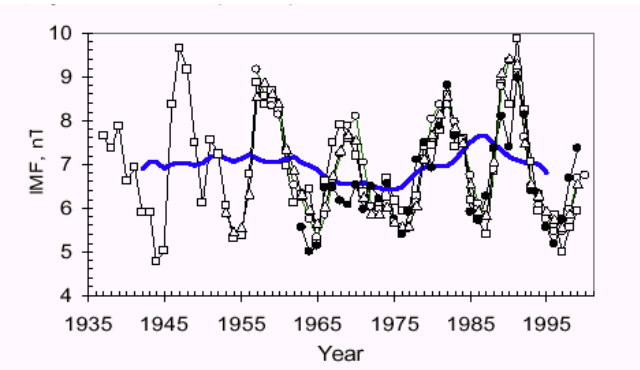

Figure 46. The experimental (●) and reconstructed values of the IMF. The reconstruction was made using stratospheric data (○), neutron monitor Climax data (△), and muon data (□). The thick blue line is the average value of the IMF smoothed with a period of 11 years. (After Stozhkov *et al*. (2001)). The agreement is not so good for solar cycle 20.

Because the modulation of cosmic rays also depends on the large-scale geometry of the heliosphere and not just on the "classical" (and somewhat discredited?) theory of diffusion off scattering centers (presumably high *B*-concentrations), the reconstruction is somewhat less firm than one might desire, especially as it is poor for solar cycle 20.

Beryllium in Greenland Ice Cores

The long-term variations in the cosmic ray flux can be determined by examining light **radionuclides**, such as ^{14}C (Carbon 14) and ^{10}Be (Beryllium 10). Both of these radioactive elements are produced exclusively by cosmic rays, and a record of both is available in ice cores and ocean sediments. The record of Carbon 14 tends to be smeared out over several decades due to its long residence time in the atmosphere. Beryllium 10 has a much lower concentration than Carbon 14, but has a temporal resolution of about a year in ice cores. Beryllium 10 (half-life $1.5 \cdot 10^6$ years) produced in the atmosphere by cosmic rays readily attaches to aerosols and gets snowed out onto ice caps, leaving a clear signal in the ice core of variations in the cosmic ray flux.

Figure 47 (after Yiou et al. (1997)) shows the recent changes in the Beryllium 10 concentration in the Greenland ice core. The decrease in Beryllium 10 since 1900 is often taken as a confirmation of the claim that the solar magnetic flux has increased by almost a factor of 2 since 1900.

You cannot have it both ways. The maximum ^{10}Be-concentration (in the red oval) was found in 1892 (±5 years) and is so large that it demands a very small solar magnetic flux, yet according to Lockwood *et al*.'s own figure the flux at that time was fairly large, comparable to the flux in 1920-35 (green rectangle). Also, the strong peak around 1951 (purple rectangle) should correspond to a low solar magnetic flux, yet it occurs in the middle of the very active cycle 18. Plotting the ^{10}Be-concentration and the inferred solar magnetic flux on the same Figure would show these disagreements in embarrassing detail. It would seem that other factors beside the solar magnetic flux might be important in determining the ^{10}Be-concentration, and thus that it is doubtful that the ^{10}Be-concentration can be taken as a confirmation of the inferred solar magnetic flux values. As noted by Beer et al. (1991) "the use of ^{10}Be to reconstruct the history of solar activity is still in the development and testing stage".

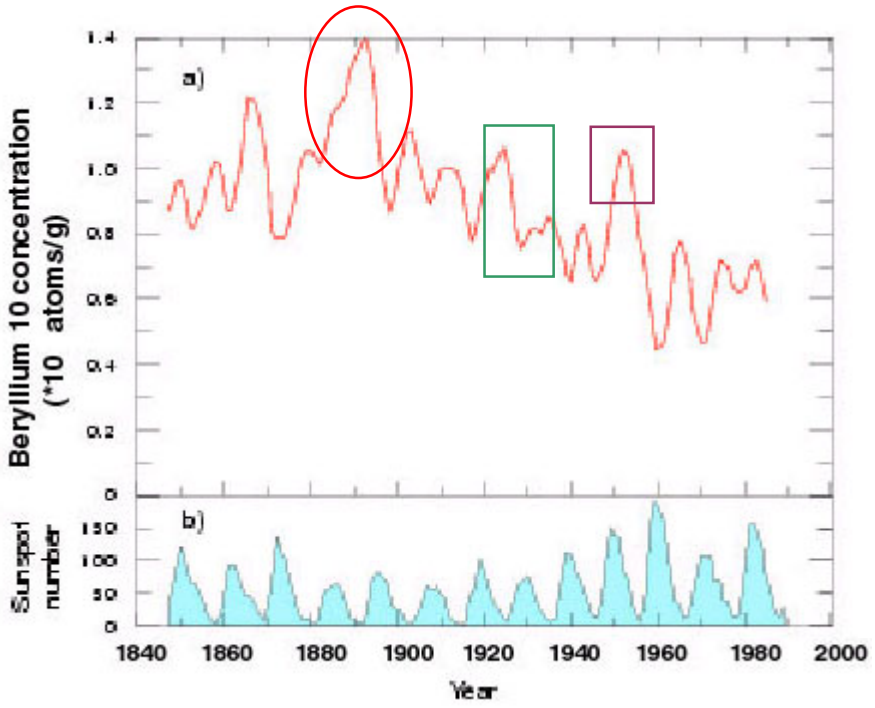

Figure 47. Concentration of ^{10}Be in an ice core from Dye3 (Greenland).

The top curve is a nine-point running mean, after Yiou *et al*. (1997).

The bottom curve shows the familiar sunspot number.

The oval shows the ^{10}Be-concentration for the years 1880-1900.

The Measured Solar Magnetic Flux

Arge et al. (2002) present and discuss all the direct measurements of the Sun's magnetic flux made at three solar observatories since 1974. If anything a light decrease is observed, Figure 48:

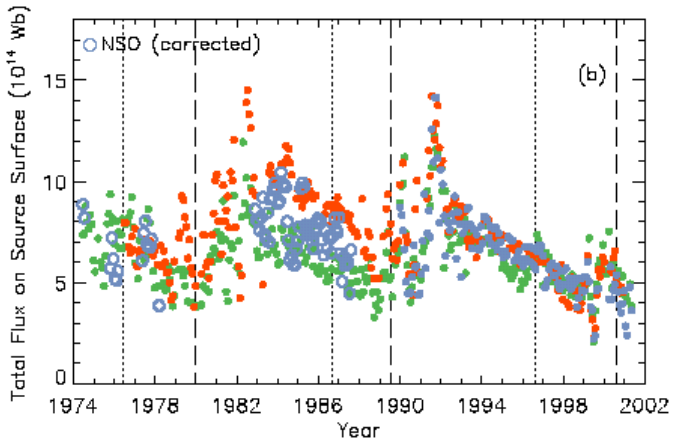

Figure 48: Total open (unsigned) flux at the source surface for each Carrington rotation since 1974. The fluxes were calculated using a potential field source surface model. The color code is: green dots: Mount Wilson Observatory; orange dots: Wilcox Solar Observatory; blue dots: the National Solar Observatory. No significant trend is apparent.

Solar Polar Crown, 1878-1994

On the other hand, Makarov (1994) found a gradual decrease of the mean latitude of polar zonal boundaries of the global magnetic field of the Sun at the minimum of activity (from 53 degree in 1878 up to 38 degree in 1994). He infers that the area of the polar zone of unipolar magnetic field at sunspot minima has risen by a factor of 2 during 1878-1996 (*sic*). It means that the behavior of the geomagnetic index, *aa*, may be explained by the increase of an area of polar zone in this period according to speculation by Makarov.

Analyzing the Quantization of the *aa*-Index

Because the *aa*-index is first scaled as *K*-values and then converted back into a range in nT, a quantization results. The standard conversion table between the quasi-logarithmic *K*-values and linear ranges in nT is:

K	0	1	2	3	4	5	6	7	8	9
Range	2.5	7.5	14	30	54	96	160	264	414	800

A complication is that two stations are involved in the derivation of the *aa*-index. The average of the two ranges is computed. If one station has a range of 7.5 ($K = 1$) and the other a range of 14 ($K = 2$), an average (with rounding) of 11 is used. A further complication is that the lowest range-value is reported as 2 (for averaging reasons actually the value 2.3 is used in the calculation, to avoid that (2.5 + 2.5)/2 =2.5 rounded to 3 results in a final value of 3 if both values were 2.5). Bubenik and Fraser-Smith (1977) discuss the quantization problem and also looked at the original, separate *K*-indices for each hemisphere. Their conclusion was that "the distribution of the N and S *K*-indices are regular, having well-defined maximums at low values of each index. Thus the irregularities in the *aa* distribution must arise in the process by which N and S are averaged to give *aa*. The process of standardization, conversion, correction, and averaging that is employed (Mayaud, 1973) is complex, and **a complete analysis of it is difficult**." Figure 49 shows how the *aa* distribution varies with time (from year to year):

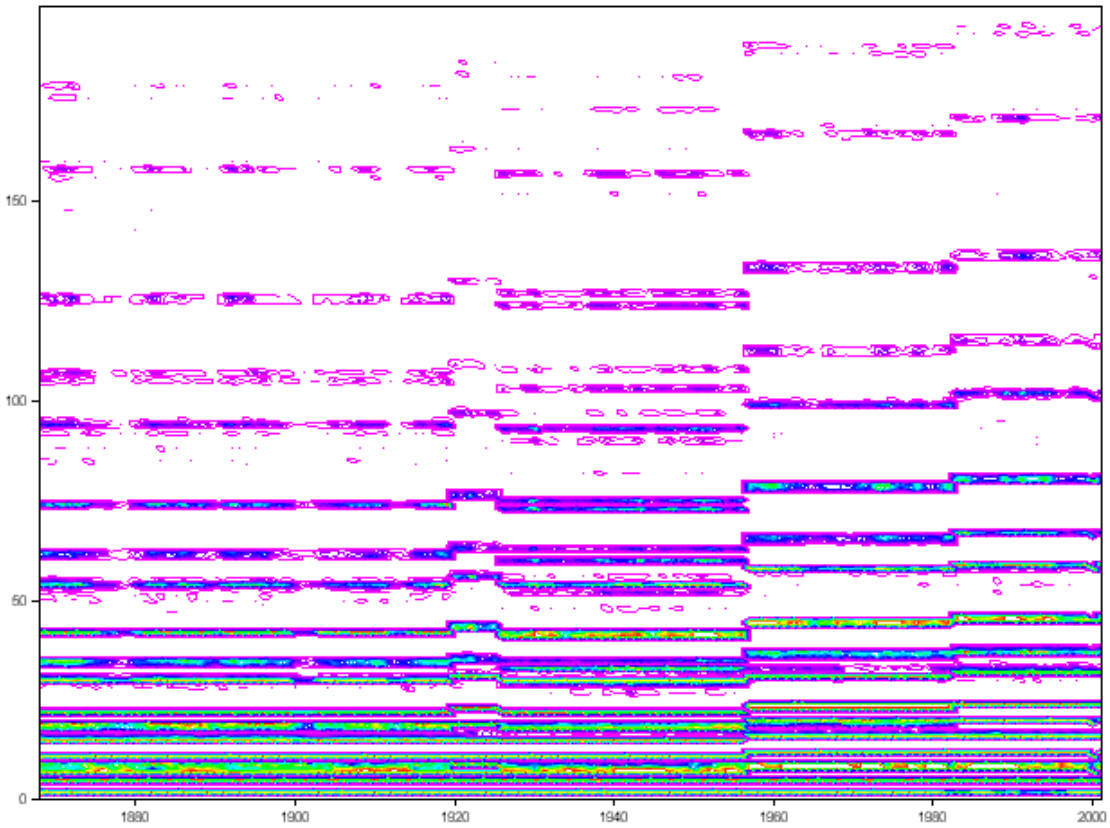

Figure 49. The distribution of *aa*-values in bins of 1 nT over the entire interval 1868-2001. The sparsely populated bins above *aa* = 200 are not shown, but discontinuities are present up there as well.

Several discontinuities stand out. As Bubenik and Fraser-Smith (1977) point out, "the probability density distributions of the geomagnetic indices can usefully indicate the existence of certain undesirable artificial components in the indices". This certainly seems to be the case here. The following table shows the precise date of each discontinuity:

Discontinuity Date	Remark	Change	Accum. change	Corr. factor
1868/01/01	Baseline	0.0%	1.0000	1.0820
1920/01/01	Change of station Melbourne to Toolangi	+3.2%	1.0315	1.0490
1926/01/01	Change of station Greenwich to Abinger	-4.2%	0.9887	1.0944
1957/01/01	Change of station Abinger to Hartland	+6.5%	1.0534	1.0272
1983/01/01	?	+2.7%	1.0820	1.0000
2000/01/01	?	-1.3%	1.0681	1.0130
2001/01/01	Back to level before 2000/01/01	+1.3%	1.0820	1.0000

Each discontinuity occurs at a change of year, i.e. at the change from December 31st to January 1st. The first three are clearly related to a change of station. It is not known what happened in 1983. The change from Toolangi to Canberra took place at the beginning of 1979. Maybe the conversion tables were only updated several years later. The situation during the year 2000 is curious. Apparently a new conversion table was used and then abandoned again in 2001.

Inspection of the nature of the discontinuities reveals that they are all of the same type, that is, they all amounts to a largely proportional change across all values of *aa*. The accumulated change column contains the factor by which *aa* has (artificially?) increased between the original baseline (1868-1919) and the interval in question. Conversely, one has to divide by that number to reduce the values to the baseline interval. If we elect to keep the modern day values (except for the year 2000) the same, we reduce the reported, older *aa* values to the 1983-2001 interval by applying the factor in the "Corr." column. If one does that, the discontinuities mostly disappear, or at least become small and random.

It is, of course, of interest to note that the discontinuities (except for the short intervals 1920-26 and 2000) all have an *upward* trend making the *aa*-index larger than it should be. We may also note that this does not explain "the factor of two" in the *aa* values. Figure 50 shows the effect of correcting the *aa*-index (reducing it to its modern level):

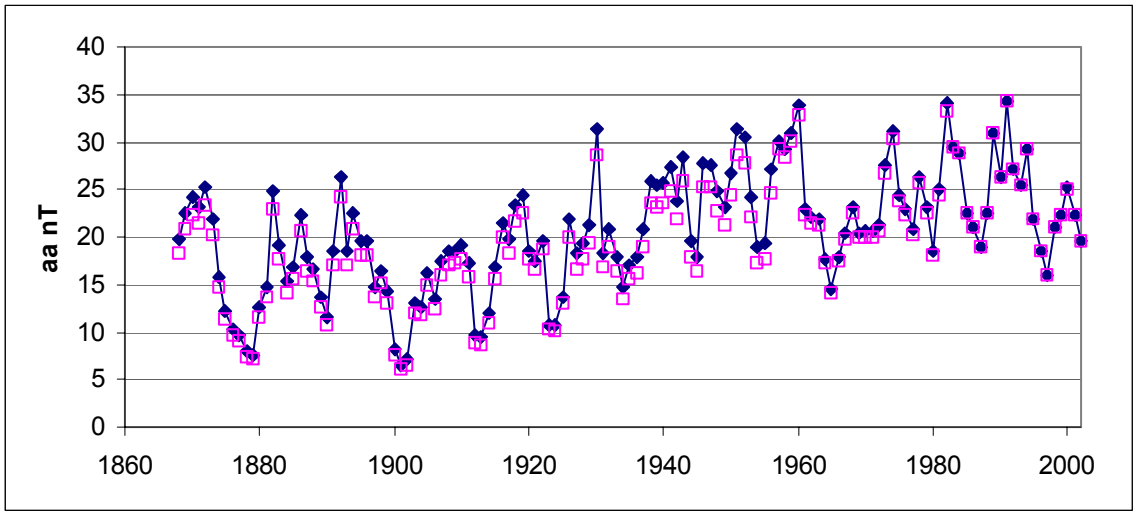

Figure 50. The *aa*-index corrected to match the present day value distribution (blue diamonds and line). The uncorrected *aa*-index values are shown by pink squares. Yearly averages shown.

Before 1957 the difference between the corrected *aa*-index and the reported values is typically up to about 10%. While this difference helps explain some of the differences between the *aa*-index reconstructed from *IHV* values and the observed values, we still have a long way to go to explain why *aa* shows such a large change through most of the 20th century. The suspicion is beginning to form that the scaling of the *K*-indices that form the basis for the *aa*-index

may be systematically in error for the earlier years. This is a serious claim and we shall look closer at this in a later section.

Reconstructing *aa*, Again

With our corrected *aa*-index we can now calibrate *IHV* again against the corrected *aa*-index. As before, we only use data from the first two 3-hour intervals (00-06 UT). We also only use data when *am* is available (1959-2000), so we can compare the reconstructed *aa* versus *am*. Figure 52 shows the relationship between $aa_{\text{corrected}}$ and IHV_{FRD}:

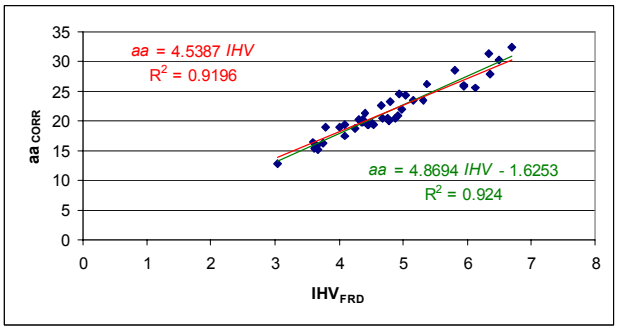

Figure 52. The relationship between yearly means of IHV_{FRD} and $aa_{\text{corrected}}$ for the interval 1959-2000. Fitting two lines to the data, both with and without an offset, we conclude that there is no significant difference and choose the fit without an offset, as its physical interpretation is more straightforward. The relationship is:

$$aa = 4.5387 \cdot IHV_{\text{FRD}} \qquad (16)$$

Armed with eq.(16) we now reconstruct *aa* from IHV_{FRD}. The differences between the calculated and the observed (but corrected) *aa*-values are shown in Figure 53:

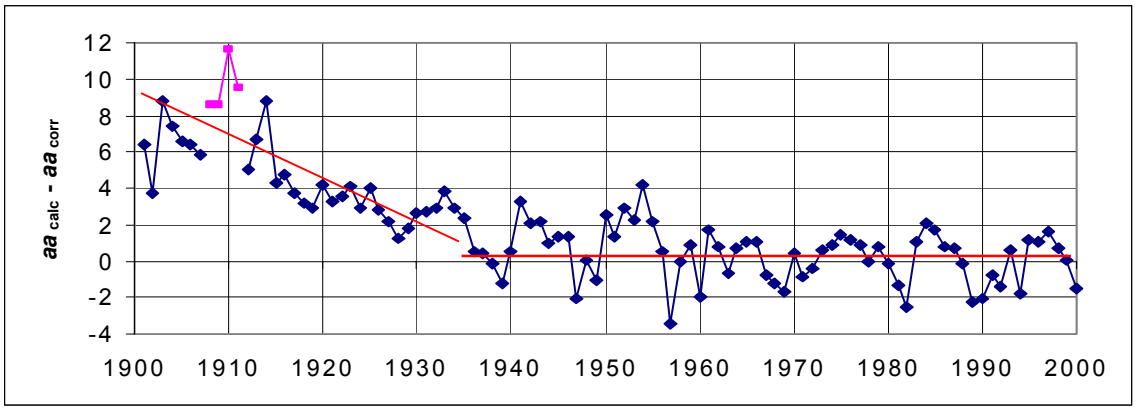

Figure 53. Yearly means of the difference between the reconstructed *aa* index (using eq.(16)) and the observed (but corrected) *aa*-index values. The anomalous years 1908-11 are shown with pink squares.

From the late 1930s until the present, the mean difference is small (0.3 nT), but before that the difference gets progressively larger. There is still (as in Figure 19) a clear sunspot cycle variation of the difference, in anti-phase with the sunspot number. To this we turn next.

Solar Cycle Variation of the Reconstruction Error

We have assumed that the solar cycle variation of the error in the reconstructed values are due to a problem with *IHV* and not with *aa*. We now explore that assumption. Our assumption was that the solar variation of the *Sq*-variation (see Figure 4) "spills" over into the *IHV* values. *IHV* for Cheltenham/Fredericksburg was constructed using only data between 00 and 06 UT during local nighttime where the normal *Sq* is at a minimum, or even absent (Figures 2 and 40).

A nice way to bring out the solar cycle variation is to use a "polar" plot:

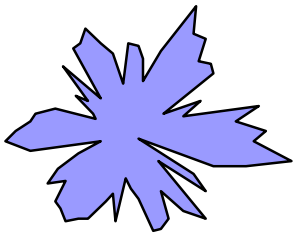

Figure 54. The deviation of the reconstructed *aa*-index from the observed (but corrected) *aa* values for the interval 1936-2000 covering six solar cycle. The six "petals" show the six solar cycles clearly. This polar plot is constructed by using the year as the angular coordinate and the deviation as the magnitude. Zero deviation is a circle about halfway out from the center to the outer boundary.

Figure 55 shows the FFT power spectrum of the differences:

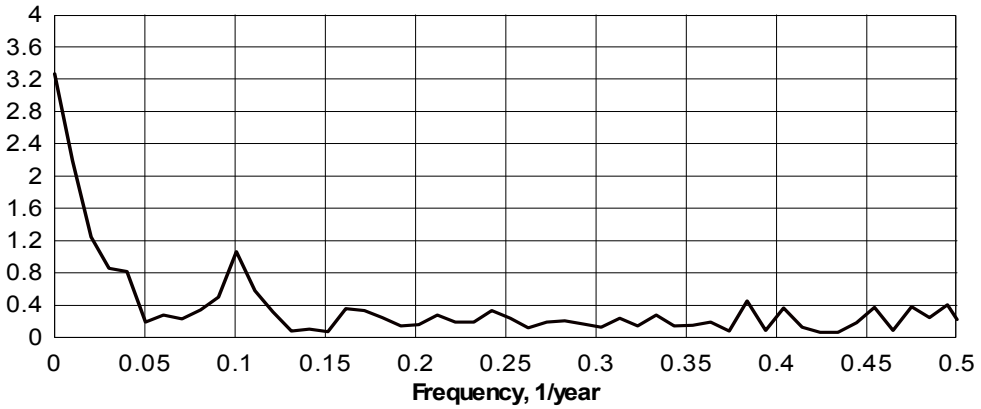

Figure 55. FFT power spectrum of the difference between the reconstructed *aa* values and the observed (but corrected) *aa* values. Yearly averages were used for the interval 1901-2000. The years 1908-11 were omitted (or rather: interpolated from 1907 and 1912).

There is power at the lowest frequencies (corresponding to the decrease from 1901 to about 1935) and a clear peak at frequency 0.10/year, corresponding to a cycle period of 10.0 years. The rest of the power variations are probably just noise (unless you want to make something of the "knee" corresponding to 22 years).

We can also plot the deviation *D* directly against the sunspot number:

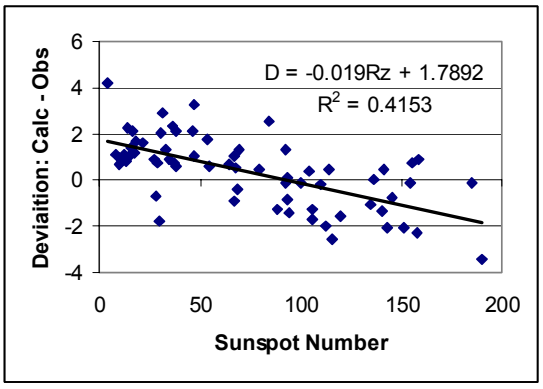

Figure 56. The deviation or difference, *D*, between the yearly averages of the reconstructed *aa* values and observed (but corrected) *aa* values as a function of the sunspot number for the interval 1935-2000. The correlation is better (R^2=0.4153) than in Figure 20 that were using uncorrected values. We have:

$$D = 1.79 - 0.019 \cdot R_z \qquad (17)$$

Combining eqs. (16) and (17), we can now write

$$aa_{\text{reconstructed}} = -1.79 + 4.5387 \cdot IHV_{\text{CLH/FRD}} + 0.019 \cdot R_z \qquad (18)$$

When both *IHV* and R_z are zero, eq.(18) leads to a negative value (-1.79) of *aa*. This is not so good. It is also not clear why the coefficient for R_z is positive. I would have expected it to be negative. If we fit an equation of the form: $aa = a + b\ IHV + c\ R_z$, a multi-variable regression gives:

$$aa_{\text{reconstructed}} = 0.6815 + 4.0241 \cdot IHV_{\text{CLH/FRD}} + 0.0198 \cdot R_z \qquad (19)$$

At least the constant term is now positive, but the coefficient for R_z is still positive. A sunspot number of 100 adds about 2 to *aa* over and above what we would expect from *IHV*. It is, of course, also possible that this sunspot-excess is a defect in the *aa*-index, resulting possible from insufficient elimination of the *Sq* variation (which *is* larger at sunspot maximum). We see the same solar-cycle variation in the *ap*-deviations (Figure 19), so it is not clear what is going on.

Distribution of *am*-Index Values

Values of the *am*-index are not quantized, as they are averages of many stations. Figure 57 shows the variation with time of the distribution of *am* values:

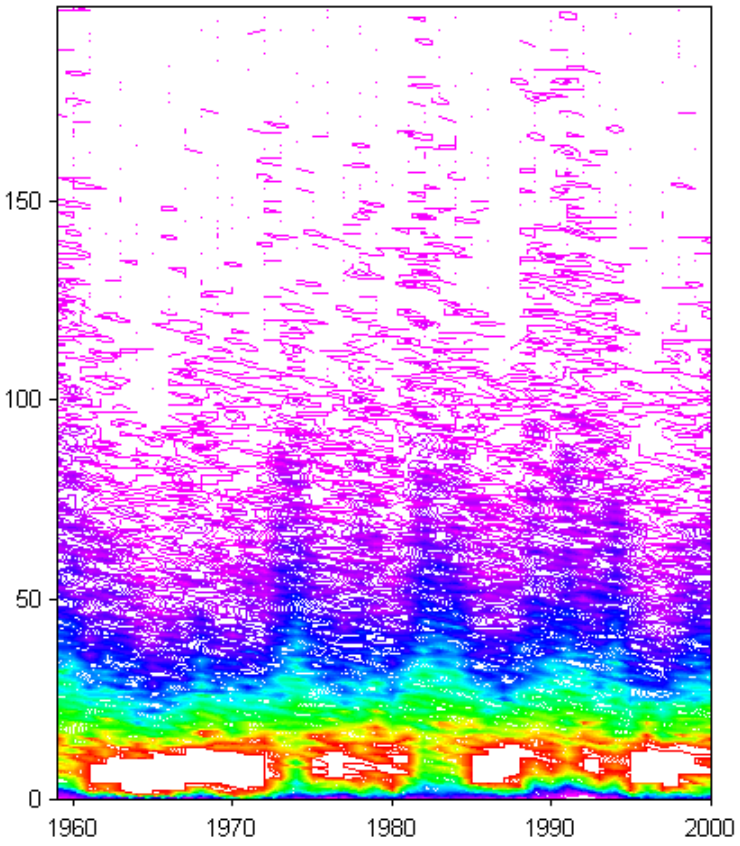

Figure 57. Distribution of *am* values for the interval 1959-2000 in 1-nT bins. Bins for values above 200 are not shown. The "saturation" effect at the lower bins (white areas) is the result of artificially "capping" the counts to get better resolution for the lower-count bins. The number of cases in each bin is color-coded (violet - low; to red - high).

It is instructive to compare Figures 57 and 47. The year 1991 stands out as one of the most active years on record. The (northern) summer months of 1991 were extraordinary in the number and intensity of geomagnetic disturbances. Because *V* and *B* are usually not correlated, by random chance we might have times where both *V* and *B* are significantly higher than normal. Such a time was the summer of 1991. Richardson *et al*. (2002) study the sources of geomagnetic activity during nearly three solar cycles, alas without drawing special attention to the activity in 1991.

The Polar Cap Potential (Godhavn 1926-2002)

It is generally believed that a sunward, transpolar, two-cell ionospheric Hall current system flows over the polar caps, or that at least an equivalent current system can be so defined. This current in thought to be a measure of convection electric fields generated by coupling of the solar wind and interplanetary magnetic field with the Earth's magnetosphere. In addition, field-aligned current may be partly responsible for some of the magnetic disturbances observed, especially during local winter. Whatever the exact mechanisms are, the magnetic effects from these currents are observed facts. The corresponding equivalent current is approximately constant across the greater part of the polar cap, meaning that all stations observe about the same magnetic effect as that of a current sheet rotating overhead with the direction to the Sun.

To the extent that the magnetic observatory at Godhavn (now properly named Qeqertarsuaq) is (sometimes barely) within the polar cap most of the time, the magnetic effect of the polar cap current is readily observed. The size of the magnetic variations varies with season and sunspot number (and is the subject of a separate paper). We minimize the influence of seasonal and sunspot number influences by computing 11-year averages. The Godhavn observatory started operation in 1926, so our averaging intervals will be taken as 1926-36, 1936-46, …, to 1986-96. The idealized effect of a rotating overhead current sheet is two equal-amplitude sine waves in the *H* and *D* components out of phase by 90˚ or six hours. Because Godhavn is not deep within the polar cap, the variation is somewhat distorted and the phase difference is almost nine hours rather than six. Figure 58 shows the variations throughout the UT-day:

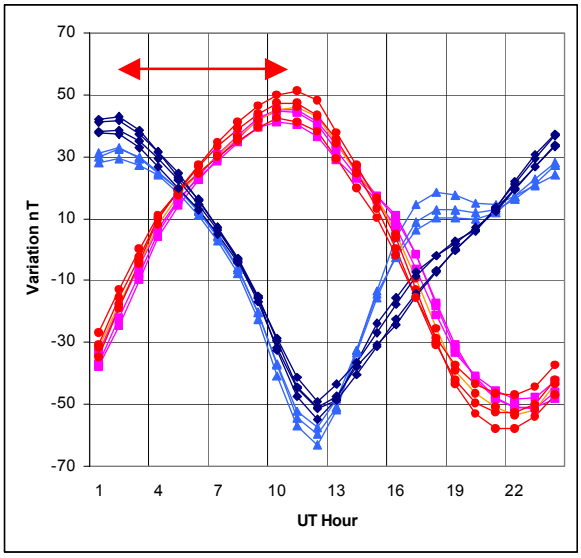

Figure 58. Variation through the UT-day of the *H* (bluish) and *D* (reddish) components at Godhavn for the seven 11-year intervals from 1926 through 1996 (1926-36, 1936-46, …). The values are plotted as deviations from the daily mean values. An interesting fine-structure of the H-variation is seen in the three most recent cycles (light-blue triangles). This feature has little impact on the overall character of the variation, but does caution against the assumption that conditions during recent times, where solar wind data is available, are representative of the entire time interval. An arrow shows the phase-shift between Δ*H* and Δ*D* component variations. Both component variations are expressed in nT.

It is remarkable how small the variation in amplitude of the magnetic effect of the overhead equivalent current is across the seven sunspot cycles. This is brought out explicitly in Figure 59:

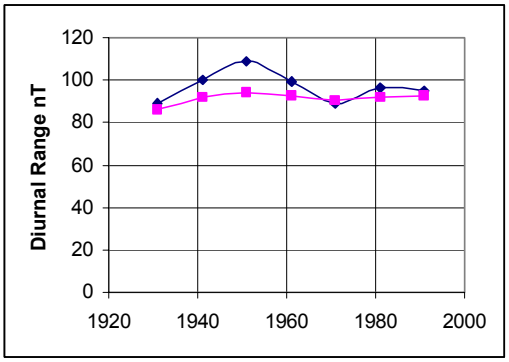

Figure 59. Variation of the average diurnal range at Godhavn for *H* (blue diamonds) and *D* (pink squares) [both expressed in nT] for the seven 11-year intervals plotted at the midpoint of each interval (1931, 1941, 1951, …, 1991).

To convert Δ*D* from tenth of arc minutes to nT, use:

$$\Delta D\,(\mathrm{nT}) = H\,(\mathrm{nT}) \cdot \Delta D\,(0.1') \,/\, 34377$$

The "fine-structure" in *H* causes additional noise and may to due to the fact that the Earth's northern magnetic pole is on the move:

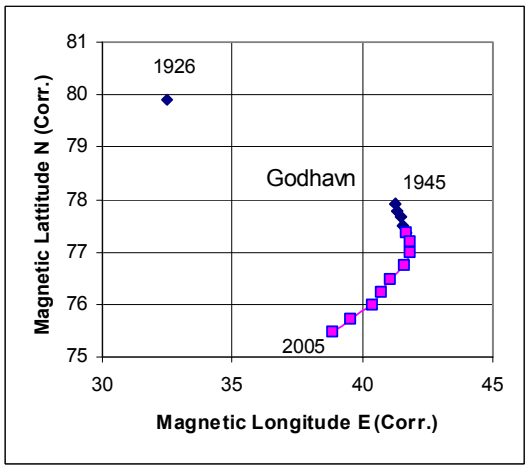

Figure 60. Corrected geomagnetic coordinates for Godhavn from 1945 to 2005. The point labeled '1926' is the location according to the geomagnetic coordinates given in the 1926 "Yearbook" (pole at 78.5 N and 291 E).

Part of the large displacement from 1926 to 1945 is undoubtedly artificial and due to the use of different reference models of the main field.

The locations for which spacecraft data exist are plotted as pink squares; the rest as blue

Godhavn is changing from being a "polar-cap"-station to becoming a "near-auroral-oval"-station. This means, among other things, that one must be careful in extrapolating conclusions drawn from comparisons between spacecraft data and magnetic observations at Godhavn back to before the space age (circa 1963). This is especially true for numerical "calibrations" and correlations derived from the recent data (*e.g.* Kjærgaard Andreasen, 1997).

The *D*-component seems considerable less sensitive to the changing geomagnetic location of the station and we'll concentrate on that component in what follows. The main conclusion is that based on the *D*-component there does not seem to be any significant long-term trend in the strength of the polar cap convection.

Observations at Gjøahavn 1903-1905

When the noted Norwegian explorer Roald Amundsen "wintered" over on King William's Land with his ship "Gjøa", he named the place "Gjøahavn". On September 12th, 1903, "Gjøa" was anchored at Gjøahavn. A magnetic observatory was constructed and continuous photographic recordings of all three components began on November 1st, 1903 and continued until the end of May 1905, thus (bracketing and) covering the entire year of 1904. The year 1904 had an average sunspot number of 42.1 and was in the rising phase of sunspot cycle 14.

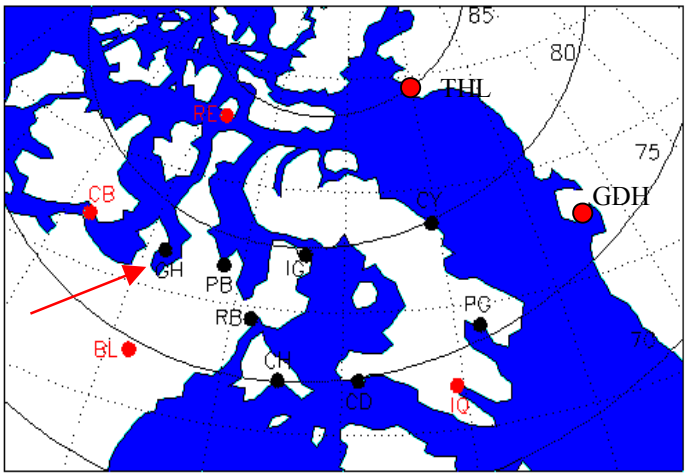

Figure 61. Location of Godhavn (GDH) and of Gjøahavn (GH - red arrow). Note how Gjøahavn lies about halfway between the Baker Lake (BL) and Resolute Bay (RE) observatories.

Circles of corrected geomagnetic latitude are shown. Note that Gjøahavn (corr. geomagnetic coordinates: 78.2ºN and 323.4ºE) is a polar cap station on par with Godhavn. Modern data (but spotty) exists for Gjøahavn.

We now select years with a similar level of sunspot activity in the rising phase of a solar cycle (to stay away from the recurrent activity during the declining phase). There are six such intervals:

Year and station	Average sunspot number
1904 Gjøahavn	42.1
1935 Godhavn	36.1
1945 Godhavn	33.2
1955 Godhavn	38.0
1966 Godhavn	47.0
1997-98 Godhavn	42.9 (average of the two years)

These intervals have similar dependence on the sunspot number (as the sunspot number is almost the same). The UV-flux should then also be almost the same and the ensuing ionization should be the same.

Figure 62 shows the *D*-variation through the UT-day for these years:

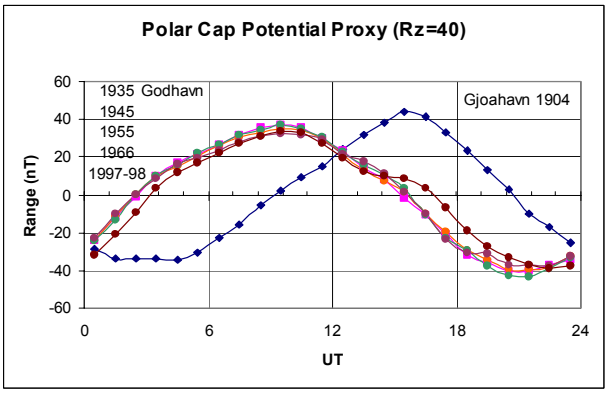

Figure 62: Variation of *D* (in nT) from the daily mean for the six yearly intervals where the average sunspot number was rather constant (about 40) for Gjøahavn (blue diamonds) and Godhavn (various other colors). The five hour phase shift between the two stations reflect the 77º difference in corrected geomagnetic longitude. No significant difference in amplitude or range is apparent.

It would seem that the diurnal range is rather constant with time (for similar values of the sunspot number). Figure 63 shows this explicitly:

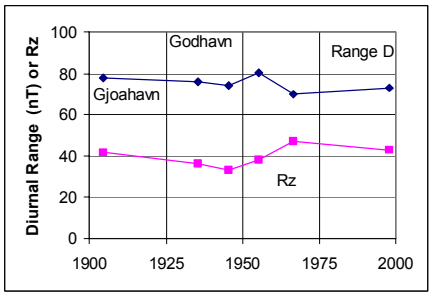

Figure 63. Variation with time of the diurnal range of the *D*-component for (near) polar cap stations. Yearly means are shown for years where the sunspot number was around 40 (indicated by pink squares in the bottom part of the Figure). The years selected are also all in the rising phase of the sunspot cycle, so they should have had very similar solar and solar wind properties. Indeed, this is borne out by the lack of significant variation with time.

For the same years from 1904 to 1966, Lockwood *et al*. calculate that the solar source surface magnetic flux changed from $2.6 \cdot 10^{14}$ Wb to $4.0 \cdot 10^{14}$ Wb (1997-98 is not on their figure, but extrapolating their curve would suggest a value of $5 \cdot 10^{14}$ Wb). We draw the conclusion that the general level of geomagnetic activity has not changed over the same time interval, but we need to demonstrate two things:

1) The diurnal range is substantially constant in space within the polar cap
2) The diurnal range does increase when geomagnetic activity increases

To this we turn next.

Polar Cap Potential (Several Stations, 1997-98)

Figure 61 showed the location of several polar cap observatories (marked with red circles). We now compute the diurnal variation of the *D*-component (actually the *Y*-component, as it does not make much difference). The stations range from deep polar cap stations (like Thule) to near-polar-cap-boundary stations (like Baker Lake or Iqaluit). Even over this range, the range of the variation of the Y-component through the day is remarkably stable as shown

in Figure 64. This fact is, of course, the reason people are talking about a "sun-following overhead polar-cap-wide return current sheet".

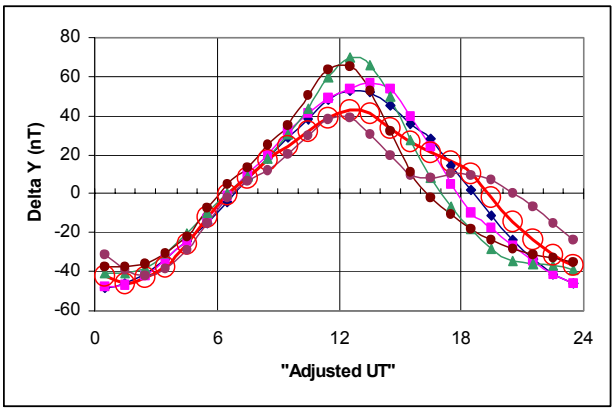

Figure 64. The diurnal variation of the deviation of the *Y*-component from its mean value for all the permanent geomagnetic observatories shown in Figure 61 for the two full years 1997-98. The curves have been shifted left or right by one or two hours to compensate for the different longitudes of the stations. The station Cambridge Bay (CB on Figure 61) is shown with a thick red line and large red circles. Cambridge Bay has almost the same corrected geomagnetic latitude as Gjøahavn.

Gjøahavn had almost the same corrected geomagnetic latitude as Cambridge Bay, so the two stations should have very similar *D*- (or *Y*-) variation. We can thus use one as a proxy for the other and the state that Cambridge Bay (and by extension Gjøahavn) has a diurnal variation comparable to any of the other observatories (in particular Godhavn) demonstrating the first point.

The Polar Cap Potential, Thule 1947-2001

The Danish Meteorological Institute has operated a geomagnetic observation for more than half a century at Thule in northern Greenland. Due to interference from the nearby Thule Air Force Base, the station was moved to the village of Qaanaaq in the mid-1950's. The geomagnetic field is sufficiently similar at both locations that no distinctions between the two stations are normally made. Regular observations of the geomagnetic field at Thule started in 1947 (although there is also data for the Second Polar Year 1932-33). Since Thule is so close to the center of the polar cap, the diurnal variations of the geomagnetic field components (X, Y, and Z) are particular regular and clear. Because the declination changed significantly during the more than fifty years, we define a "local" coordinate system for each year with the X-axis (denoted X') along the yearly mean direction of the horizontal force. This allows us to compare data from several years without worrying about secular changes. Figure 65 shows the average diurnal variation of the three components (in our local coordinate system) at Thule:

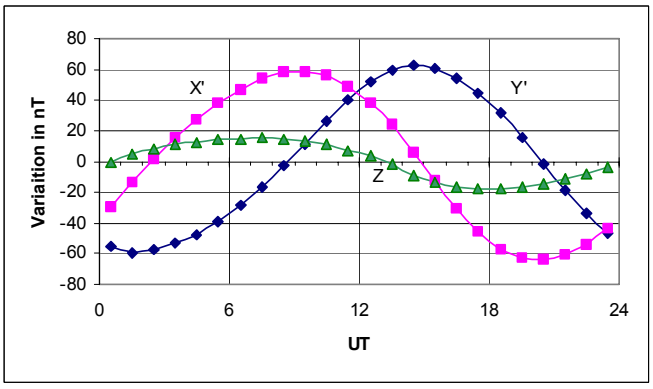

Figure 65. Diurnal variation of the geomagnetic field components (in our "local" system) at Thule for the interval 1947-2001. The X' and Y' components show nearly perfect sinusoidal curves with the six-hour phase shift so characteristic of an equivalent current sheet rotating overhead. The auroral electrojets have a clear (but weak) effect on the variation of the vertical component. All values are departures from the daily means.

A striking way of depicting the regularity of the variation is shown in Figure 66. Here is plotted the path of the end point of the horizontal force vector as it varies throughout the UT-day. For an ideal current sheet rotating overhead, this path would be a circle. The radius of the circle (in nT) is then a measure of the magnitude of the polar cap potential.

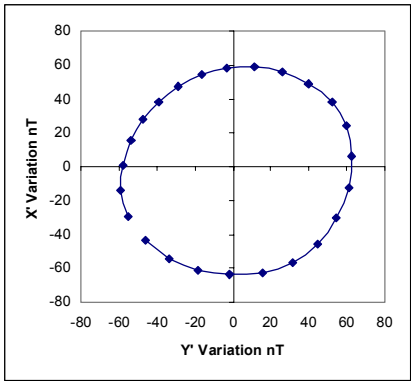

Figure 66. Vector diagram of the diurnal variation of the horizontal force of the geomagnetic field at Thule for the interval 1947-2001. During the day the (almost constant) magnetic effect of the overhead equivalent current rotates clockwise as the Earth (and Thule) rotates underneath the current sheet. The cross-polar cap current maintains its direction towards the Sun. This behavior has been known since the observations during the First Polar Year in 1882-83. On disturbed days the same rotation is seen, except that the radius is greater. The radius of the circle also undergoes seasonal (greater in local summer) and sunspot cycle variations (increases with sunspot number).

The yearly averages of the magnitude of the polar cap potential is strongly correlated with geomagnetic activity, *e.g.* as measured by the *aa*-index. Because the end-point of the rotating vector does not trace out a perfect circle we analyze the total range separately for the horizontal two components (X', Y' or H, D) as a function of the yearly average values of *aa* in Figure 67:

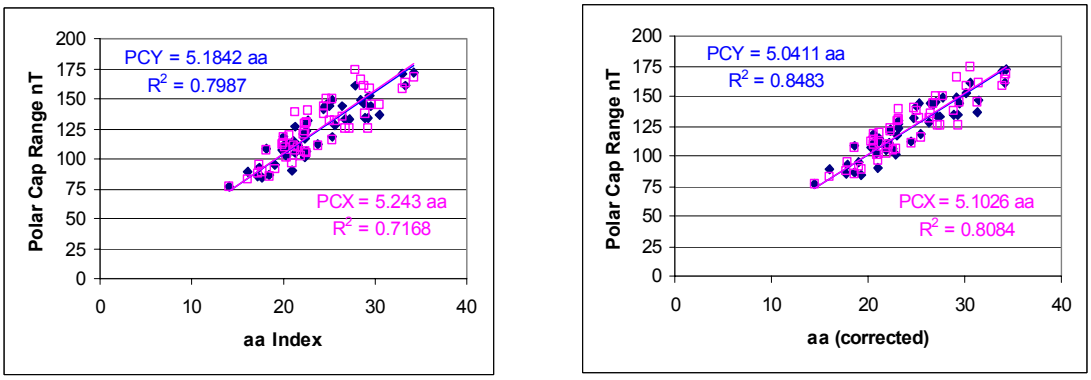

Figure 67. The relationship between the range (from minimum to maximum) of the polar cap potential as measured by the diurnal variation of the horizontal components (Y' = blue diamonds; X' = pink squares) and yearly averages of the *aa*-index (left panel). The right-hand panel shows the correlation using the corrected *aa*-index. Data from Thule for 1947-2001.

Note that the correlation coefficient is higher for the Y'-component and also higher when using the **corrected** *aa*-index. The latter observation tends to bolster the claim that *aa* needs to be corrected for the discontinuities shown in Figure 49. Figure 68 shows the evolution with time of the polar cap range:

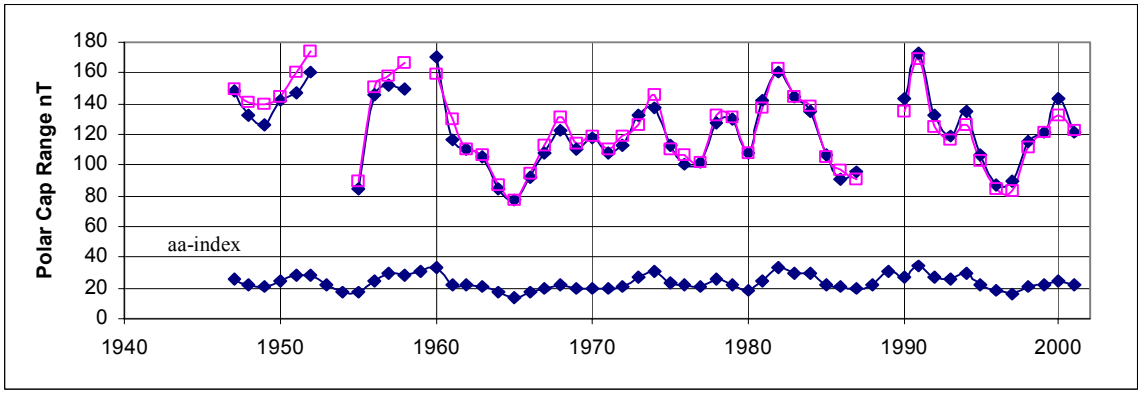

Figure 68. Polar cap potential range for Thule (Y' = blue diamonds (top); X' = pink squares).

As before, no long-term trend is apparent. We see the usual fact that some sunspot cycles are more active than others.

Spatial Variation of Polar Cap Current

In this section we investigate the latitude variation of the polar cap current. In other words, as we move from the very center of the polar cap towards the auroral oval, the image of a simple current sheet rotating with the direction to the Sun begins to break down. The Canadian stations are ideally located for this.

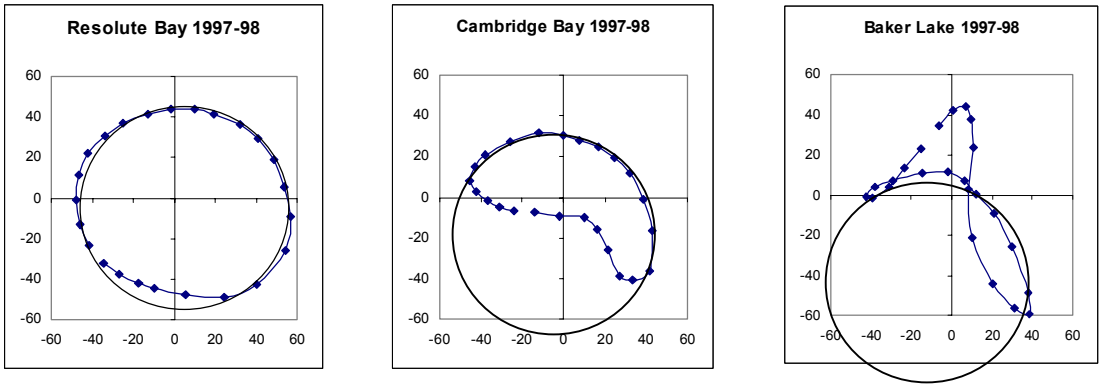

Figure 69. Horizontal force vector diagrams for three observatories covering a range of latitude from well north of to well south of Gjøahavn. This covers the range from near the center of the polar cap to near the edge. Note that Cambridge Bay has nearly the same latitude as (in fact is very near) Gjøahavn (see Figure 61). We have chosen the years 1997-98 that have the same average sunspot number as 1904 when Gjøahavn was operating. Circles are fitted to the polar cap current effects on each diagram. Note that these circles have the same size.

Near the center of the polar cap (Resolute Bay) the vector diagram is nearly circular. The range is 100 nT (for Thule it was 102 nT). There is a hint of a deformation on the dayside (lower left portion). This deformation grows progressively larger as we move away from the center of the polar cap, but at all stations, the nightside variation fits a circle of the same size as the Resolute Bay circular variation. This has its cause in the eccentricity of the auroral oval ensuring that during the night all stations are well within the polar cap.

The Polar Cap Current at Gjøahavn 1904

We can compare the polar cap current vector diagrams for Cambridge Bay, Godhavn, and Gjøahavn for years where the sunspot number was about the same (about 40). For Godhavn, we have chosen the earliest year possible, because the auroral oval in modern times has moved closer to the station:

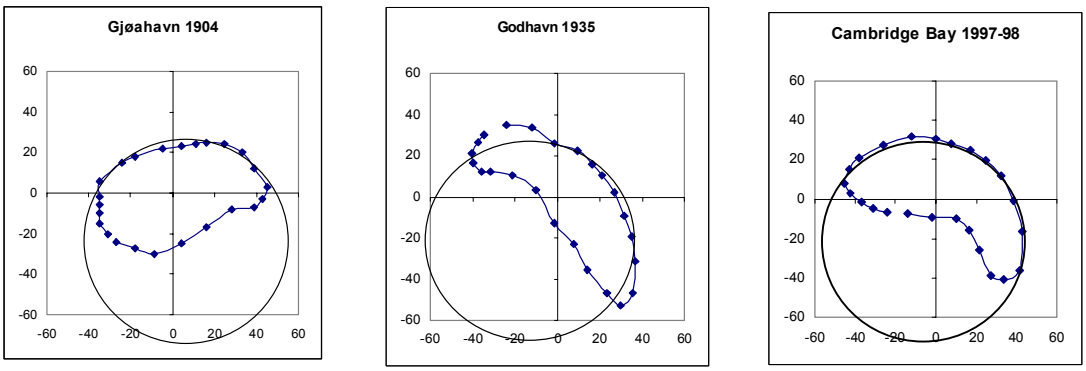

Figure 70. Horizontal force vector diagrams for three stations at closely the same position within the polar cap for years with approximately the same sunspot number (≈40). A circle (of diameter 100 nT) is fitted to the nighttime variation of each station.

It is clear that no long-term trend is present. In particular, the two stations that are physically closest (Gjøahavn and Cambridge Bay) show very similar polar cap current effects. On all the vector diagrams shown above and in the previous sections, it is apparent that the influence of being close to the day-time auroral manifest itself the *least* in the Y' (or D-) component, and that it here usually is felt as a slightly smaller range of the variation of the Y' (or D-) component through the UT-day.

Long-term Variation of the Polar Cap Current Strength.

Using the observed insensitivity of the Y'-component to closeness to the auroral oval we can compile all the data for Thule, Godhavn, and Gjøahavn into a single Figure:

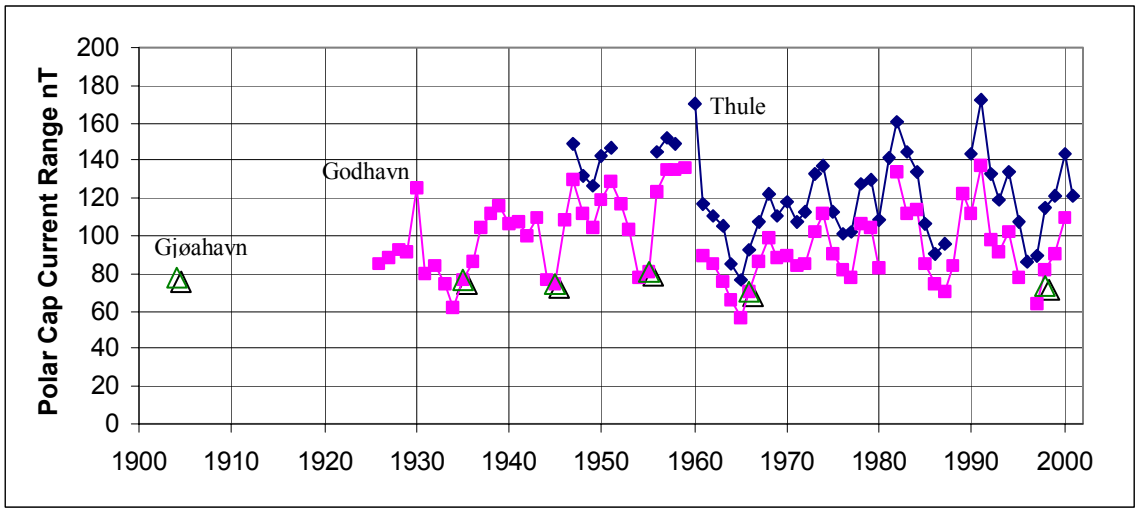

Figure 71. The polar cap current range in the Y'-component for Thule (blue diamonds), Godhavn (pink squares and triangles after 1926), and Gjøahavn (triangle 1904)

Godhavn Y'-component alone shows a range that is about 20% less than Thule's polar cap current range, but the correlation between the two stations is good (R^2 = 0.94; as good as the "unprecendentedly high" case quoted by Lockwood *et al.*). While a solar cycle variation is evident, no long-term trend is apparent. In particular, the 1904 and the 1997-98 polar cap potentials are of very similar magnitude as are all the other year's for which the sunspot number was the same as in 1904 (≈40, triangles).

The Polar Cap Electric Field

Troshichev and Andrezen (1985) have shown that ground geomagnetic disturbances measured at a single near-pole station correlate highly ($r > 0.8$) with the interplanetary "merging electric field". Their Polar Cap Magnetic index (essentially a normalized and seasonally adjusted version of our Polar Cap Current range) correlates with the product of the magnitude of the interplanetary magnetic field (apart from a geometric factor), B, and the solar wind speed, V: $PC \sim V \cdot B$. We also find such a relationship:

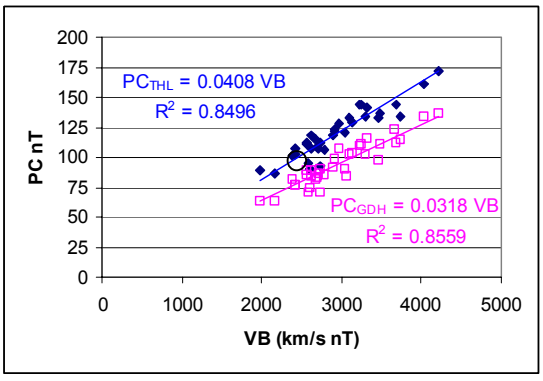

Figure 72. The dependence of the Polar Cap current on the product $V \cdot B$ (V in km/s and B in nT). Shown are the size of the vector "circle" at Thule (blue diamonds) and the range in the Y'-component at Godhavn (pink squares) for the interval 1966-2001. Yearly average values were used. The solar wind data has random data gaps, but the polar cap current ranges were calculated using the full year in order to average out the seasonal variation. Note that our values are easy to understand as no normalizations or calibrations were performed.

Estimation of *V* and *B* in 1904

From the relationships given in Figure 72, we can express a measure, *VB*, of the electric field in two different ways:

$$VB \text{ (km/s nT)} = 24.5\ PC_{THL} \qquad (20)$$

$$VB \text{ (km/s nT)} = 31.4\ PC_{GDH} \qquad (21)$$

For Gjøahavn (for the year 1904) we found a fit to the effect of the nighttime polar cap current to be 100 nT and for the range of the Y'-component we found 78.1 nT. Identifying these with PC_{THL} and PC_{GDH}, respectively, we estimate VB_{1904} = 2450 and VB_{1904} = 2452 (km/s nT), respectively. We get then a consistent estimate of VB_{1904} of 2451 (km/s nT). What is the uncertainty of this? We have not conducted a formal error analysis because the major errors are in the assumptions and thus are systematic rather than random. It seems that an error greater than ±250 (km/s nT) would make this datapoint (if plotted in Figure 72) an "outlier", so we shall make that a guesstimate of a conservative uncertainty (10%).

Using the relationship between *B* and the sunspot number, R_z, (42.1 for 1904) from Figure 7, we get

$$B_{1904} = 5.59 + 0.0149\ R_z = 6.22 \text{ nT, and thus } V_{1904} = 2451 / 6.22 = 394 \text{ km/s}$$

Using eq.(7) to estimate the solar wind speed from *IHV* and R_z, we obtain for 1904 (using $IHV_{CLH/FRD}$):

$$V_{1904} = 432 \text{ km/s, and thus } B_{1904} = 2451 / 432 = 5.67 \text{ nT.}$$

The mean values from these independent estimates are:

$$R_{z1904} = 42.1 \qquad V_{1904} = 413 \text{ km/s} \qquad B_{1904} = 5.95 \text{ nT} \qquad (22)$$

None of these are particularly out of the ordinary. They are very typical for a year with a sunspot number of about 40 in the *rising* part of the sunspot cycle. Other such years include:

Year	Sunspot Number	Solar Wind Speed	IMF Magnitude
1966	47.0	430 km/s	6.37 nT
1977	27.5	414	5.84
1987	29.4	430	6.03
1997	21.5	376	5.25
1998	64.3	406	6.61
Mean ➔	37.9	411	6.02

In a sense it is comforting that 1904 was so ordinary, but on the other hand it is also disappointing that nothing unusual was going on. It would have been exciting had the Sun's coronal magnetic field indeed doubled since then as part of a long-term change.

Applying the Lockwood-Stamper-Wild formula (eq.(12)) to the relevant quantities for 1904 yields:

Year	aa-Index	Recurrence Index	IMF Magnitude
1904	11.685	0.01885	3.79

If the interplanetary magnetic field were indeed that low, we would need a solar wind speed of 646 km/s to bring the value of *VB* up to 2451 km/s nT. Such a high value does not seem credible. We could ask: "what value of *aa* would in the Lockwood *et al*. formula give *B* = 5.95 nT when the recurrence index I_{1904} is 0.01885?". The answer is aa_{1904} = 18.04. The reconstructed *aa* for 1904 (Figure 18) is 18.05. This close agreement is pure luck, of course, but is consistent with the view that values of *aa* in the beginning of the 20th century are much too low.

Summary of the Problem

We have covered a lot of ground and a certain numbness is beginning to set in. To counteract that, we shall here summarize our findings and set the stage for further development. Figures 73a through 73e show the variation of the several indices and quantities that we have been working with since 1868 (as far as they are known) but only plotted for years where the average sunspot number was about 40. In a few cases two consecutive years have been combined to encompass $R_z \approx 40$. All the years selected were in the *rising* phase of the solar cycle where the recurrence index was low and thus no strong, recurrent high-speed streams were present:

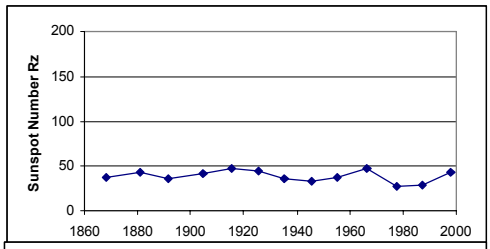

Figure 73a. The sunspot number, $R_z \approx 40$, for the years in the rising phase of the sunspot cycle that were selected. The years are 1868, 1880-81, 1891, 1904, 1915, 1935, 1945, 1955, 1966, 1977, 1987, and 1997-98. The values shown are assigned to either the midyear (e.g. 1915.5) or the mid-interval (e.g. 1881.0).

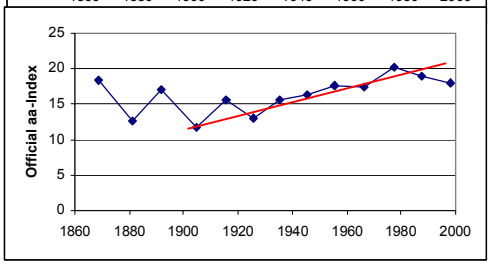

Figure 73b. The *aa*-index for the years selected in Figure 72a. This is the uncorrected *aa*-index as reported by the World Data Centers. The red line marks the "linear change" that has attracted so much attention. Note also the large inter-cycle changes before 1930 and the downward trend during the last two solar cycles.

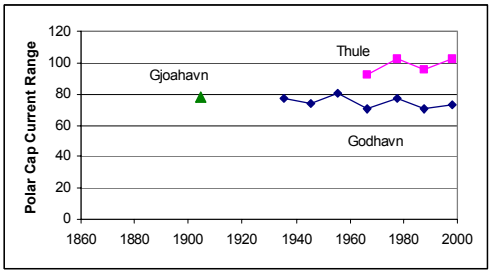

Figure 73c. The diurnal range (in nT) of the magnetic effect of the cross polar cap current as measured using the local Y'-component of the geomagnetic field. Thule near the center of polar cap sees a slightly larger and much more regular variation than do Godhavn and Gjøahavn, but there is no long-term trend in the polar cap potential.

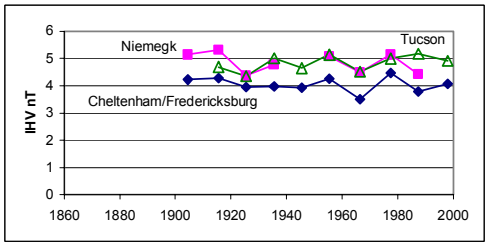

Figure 73d. The Inter-Hour Variability (*IHV*) index for the three midlatitude stations shown. The *IHV*-index correlates very well with the *am*-index and can be used as a close proxy for that index, leading to the assertion that no long-term trend in the *am*-index is apparent.

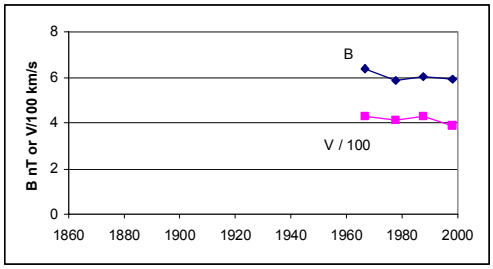

Figure 73e. Observed values of the magnitude, *B*, of the interplanetary magnetic field (nT) and of the solar wind speed, *V*, (in 100 km/s). Again, no trend is present. Note that the years in all of the Figures in this panel are the same.

The following table summarizes the data:

Year	Sunspot number	Recurr. index	*aa*-index	$IHV_{CLH/FRD}$	IHV_{NGK}	IHV_{TUC}	IMF *B*	Solar wind speed *V*
1868.5	37.6		18.3					
1881.0	43.3		12.7					
1891.5	35.6	0.078	17.1					
1904.5	42.1	0.019	11.7	4.22	5.13			
1915.5	47.4	0.199	15.7	4.28	5.32	4.68		
1925.5	44.4	0.094	13.1	3.95	4.36	4.35		
1935.5	36.1	0.104	15.7	3.98	4.78	5.01		
1945.5	33.2	0.151	16.4	3.92		4.65		
1955.5	38.0	0.000	17.6	4.26	5.09	5.13		
1966.5	47.0	0.020	17.4	3.50	4.49	4.52	6.37	430
1977.5	27.5	0.095	20.3	4.46	5.14	5.11	5.84	414
1987.5	29.4	0.069	19.0	3.78	4.42	5.18	6.03	430
1998.0	42.9	0.053	18.6	4.06		4.92	5.93	391
Mean	**38.8**	**0.080**	**16.4**	**4.04**	**4.84**	**4.83**	**6.04**	**416**

Going from *aa* to Solar Magnetic Flux

We have demonstrated that the rise in the *aa*-index over the last 100 years probably is an artifact of some sort (again, nobody is saying that solar cycle average of *aa* did not rise during the first half of the 20th century), but the actual claim by Lockwood *et al.* is that the solar magnetic flux in the corona has doubled, *i.e.* not the *aa*-index per se.

The Influence of the Earth's Decreasing Dipole Moment

Our eq.(12) omitted (or rather subsumed into B_0) the variation of the Earth's magnetic dipole moment M_E. Taking the Earth's magnetic dipole moment to be variable, introduces a factor $h = (M_E/M_0)^{-0.8636}$ where M_0 is M_E at a suitable reference epoch. Taking the year 1900 as reference epoch, the value of M_E/M_0 in 1990 was 0.948 (*e.g.* Russell and Luhmann, 1997), thus $h_{1990} = 1.047$. So, even if geomagnetic activity (as expressed by *aa*) were constant, an increase of 4.7% of the calculated solar magnetic field would be called for, solely due to the decrease of the Earth's magnetic moment according to the Lockwood *et al.* formula. Their original formula (see their paper for nomenclature) is reproduced here:

$$P\alpha = \{k\pi/2\mu_o^{(1/3-\alpha)}\}\, \boldsymbol{M_E^{2/3}}\, m_{sw}^{(2/3-\alpha)}\, N_{sw}^{(2/3-\alpha)}\, v_{sw}^{(7/3-2\alpha)}\, B_{sw}^{2\alpha} \sin^4(\theta/2) = \boldsymbol{aa_p} / s_a$$

From this we see that in their formulation, the *aa*-index should *increase* with M_E. Their argument was that the solar wind kinetic energy is incident on the geomagnetic field, which presents a roughly circular cross section to the flow. A fraction (α) of the incident energy (P) is extracted and is eventually measured as *aa* (with scale-factor s_a). Since a weaker terrestrial dipole means a smaller cross section, it follows that *aa* should decrease with decreasing M_E.

Let us look at the physics here. As the Earth's magnetic moment gets smaller, a given solar wind dynamic pressure would compress the magnetosphere just like a solar storm would (albeit on a different scale!), and we would register more geomagnetic activity because the gradients are higher and we are also simply closer to the magnetopause and ring currents. It is this that I think is behind the equinoctial mechanism for the semiannual variation. This viewpoint is quite the opposite of the Lockwood *et al.* argument. They claim that a *larger* magnetosphere results in more geomagnetic activity. My claim is that a *smaller* magnetosphere means more geomagnetic activity, not because we intercept more kinetic energy, but because the gradients are higher and distances to the currents are smaller. This view is admittedly a gross simplification, but I think that the semiannual variation validates the view: when the stronger geomagnetic field away from the equator is facing the sun and pushes the magnetopause out, activity goes down (valley digging).

Maybe we can take the case of the "day the solar wind disappeared" on May 11th, 1999. At 15h UT, the proton density dropped to 0.2 cm^{-3}, the speed to 330 km/s, while the magnetic field was steady at 6 nT. Similar conditions prevailed during most of day. The magnetosphere swelled to five or six times its normal size (30-fold increase of cross section) almost offsetting the decrease of incident kinetic energy density. The *am*-index was very low throughout the day (daily mean = 2.1) as was the *aa*-index (daily mean = 3.9). Intuitively, the magnetosphere shields us from the effects of the solar wind and the stronger the Earth's magnetic field is, the stronger the shielding ought to be. A counter-argument might be that the energy incident on and captured by the magnetosphere is focused on the edges of the terrestrial polar caps and that therefore a bigger magnetosphere has more energy to focus on where we measure its effect. This whole question is actually vital, as there are signs that the dipole is rapidly weakening and may disappear in a few hundred years or so.

In any case, what *is* important is what Lockwood *et al.* thought, because that goes into their claim of doubling of the Sun's coronal magnetic flux.

The Influence of Changing Background Recurrence Index

In Figure 29 we noted that there seemed to be a change in the "background" Recurrence Index, I_0, around 1920 from about 0.18 to 0.07. According to eq.(12) that would *increase* the magnitude of the interplanetary magnetic field by 9.5% for an average *aa*-value of 20. The change is quite sensitive to the value of the background recurrence index. Lowering the background value to half of 0.07 would entail an increase of B of 16% compared to $I_0 = 0.018$. Setting the recurrence index to zero would result in an increase of 62%. Some of the I-values read off Lockwood *et al.*'s Figure are indeed near zero. (I even calculate some that are negative…)

The following table shows the effects on the calculated solar coronal magnetic flux of the various influences we have identified:

Reason for change from 1900 to 1990	Factor	Cumulative factor to 1990
Calibration error of *aa*	1.082	1.082
Decrease of Earth's dipole moment M_E	1.047	1.133
Background I_0 from 0.18 to 0.07	1.095	1.240

We can account for about a 25% increase, but not a factor of two (or more - as claimed by Lockwood *et al.*)
So we need to look at the possibility that either some serious problem besets the *aa*-index or that we don't have a good understanding (as we think we have) of how all these things hang together.

The influence of the value of the recurrence index is particularly troublesome near zero. Figure 74 shows the calculated solar magnetic field (using eq.(12)) for a yearly average value of *aa* of 20:

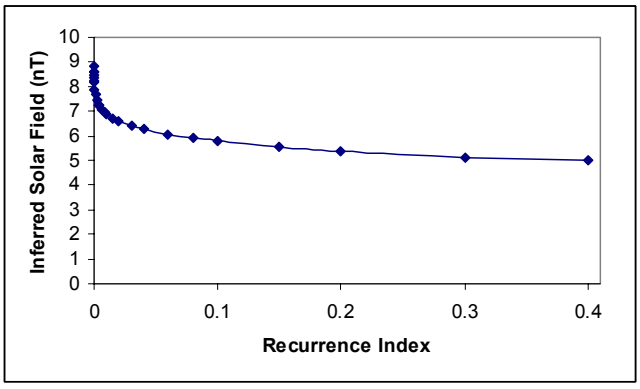

Figure 74. Calculated value of the solar coronal magnetic field using the Lockwood *et al.* formula with a mean value of *aa* = 20 as a function of the Sargent recurrence index. As the recurrence index goes from 0.4 to zero, the calculated magnetic field rises from 5.0 nT to 8.8 nT. Most of the increase takes place for very small values of the recurrence index. This is not physically reasonable. According to eq.(11), the solar wind speed for a recurrence index of zero is 339 km/s regardless of the value of *aa*.

Analysis of *aa*-Values

The lowest value that *aa* can have is 2. This is in marked contrast with *ap* and *am* that can go down to zero. A low yearly average value of *aa* could be caused by many cases of *aa* = 2. Figure 75 shows the number of 2s per year:

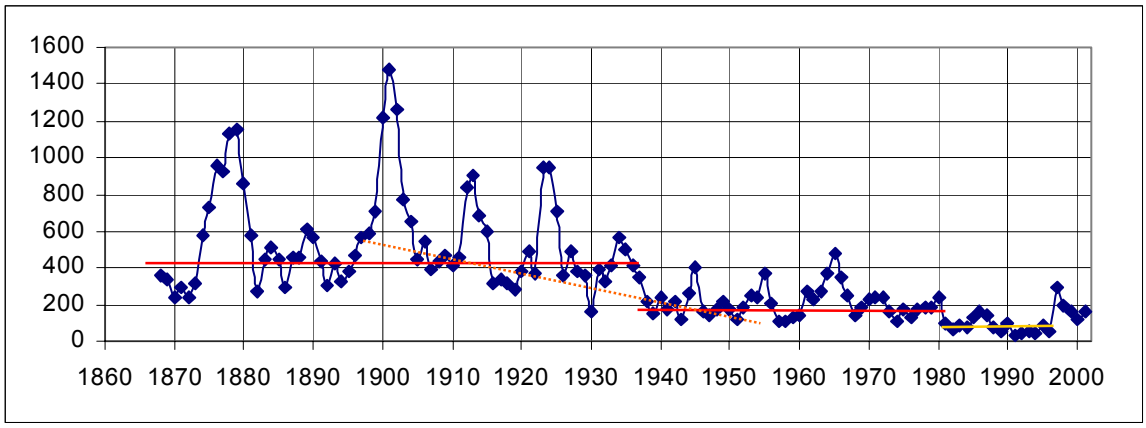

Figure 75. Number of *aa* values equal to 2 per year for the interval 1868-2001.

There is a clear "inverse" sunspot cycle variation (as we would expect). Relatively narrow peaks at sunspot minima stand out (with notable exceptions during the eight-year interval 1874-81 and the six-year interval 1899-1904). Away from these peaks the distribution seems to be, at least, bimodal with a count of about 400 before 1938 dropping to a count of about 200 thereafter. There is a possibility that the count is too low from 1981 through 1997. If so, that would help elevate the *aa*-index. People with the "linear change" syndrome might argue that we have a linear drop from a count of 500 around 1900 to a low of 100 in the 1950s. Linear change argues for physical cause while discontinuity argues for instrumental problem. In Figure 38 we saw another example of a discontinuity in the same year, 1938.

The year 1901 has a remarkable number of cases where *aa* = 2, namely 1484 of them out of a total number of 3-hour intervals for the year of 2920, *i.e.* 50.8%. More than half of the time there was *no measurable activity at all*.

Clilverd *et al*. (2002) notes that "Three instrumental effects have been identified where significant changes in quiet-time conditions can be seen *i.e.* 1938, 1980, and 1997" They find that three instrumental changes occurred coincident with these critical dates. These changes occurred at Northern Hemisphere sites, while the Australian sites remained close to their usual values.

At the beginning of 1938, the horizontal and declination variometers that were installed at Abinger were replaced by a single LaCour magnetograph system. The old system had scale values of 2.62 nT/mm in H and 3.29 nT/mm in D, whereas the new system was less sensitive with scale values of 4.40 nT/mm and 4.85 nT/mm in H and D, that is 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, which is 1.9 mm on the old variometers, but only 1.1 mm on the new, so it is likely that fewer occurrences of *K* = 0 were scaled with the new system (in the opinion of Clilverd *et al.,* although I would think it would work the other way around - the less sensitive system can't see small variations and thus mislead the observer to think that there is no variation and scale K = 0). In 1980 the presence of locally generated noise (*sic*) was noted, and in 1997, a new low-noise magnetometer was introduced at Hartland. The actual number of *K* = 0 scalings at Hartland in 1997 was 500, more typical of the years before 1938.

Comparison of data from the *aa* observatory sites with Sodankylä (Finland) shows that Sodankylä scales the same number of *K* = 0 values as the Australian sites from the beginning of observations at Sodankylä in 1914 until the present, but that in 1938 there is a significant change of the Northern Hemisphere values, concurrent with the change of scale values at Abinger. The difference continues to grow until 1997 when the new instrument was installed at Hartland.

Analysis of *IHV* Distribution

The distribution of values is a powerful tool (as pointed out by Bubenik and Fraser-Smith (1977)) to check for discontinuities, changes of calibration, instruments, observers, or methods. We shall now employ this tool to the *IHV* values (from Cheltenham/Fredericksburg). Figure 76 shows the result:

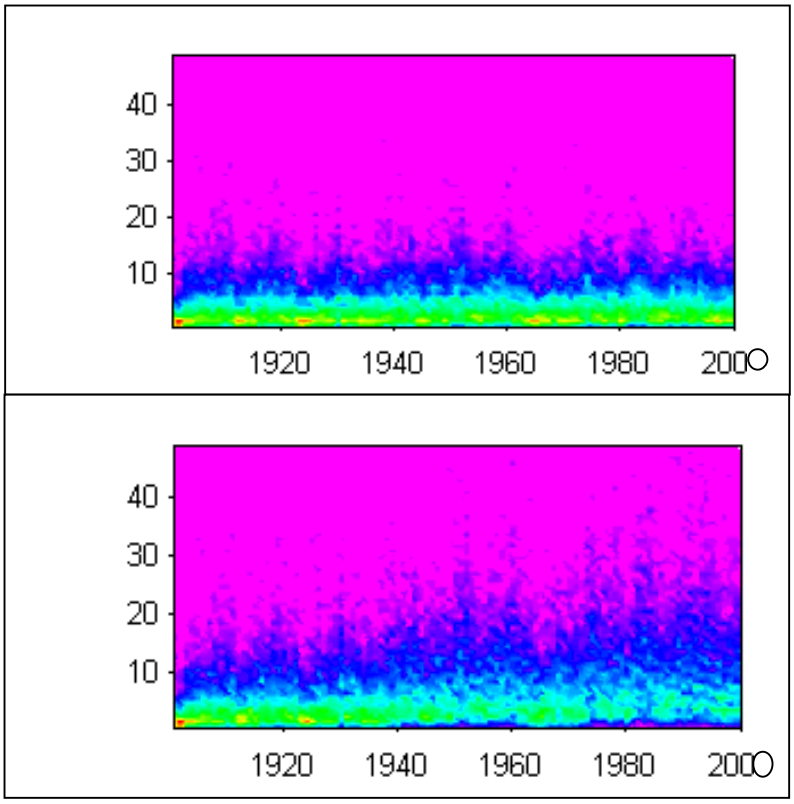

Figure 76a. Distribution of IHV values from the stations Cheltenham / Fredericksburg in 1 nT-bins for the entire interval 1901-2000. The very sparsely populated part of the distribution above 50 nT is not shown. The "rainbow" color-coding is such that from red through violet goes from higher towards lower counts.

Figure 76b. Same format as in Figure 76a, except that an artificial increase of a factor of 2.31 over 91 years (as claimed for the solar flux by Lockwood *et al.*) has been applied to the *IHV* values. This is what the *IHV* plot should look like if such an increase had in fact taken place. (A bug in the plotting software mangled the labels).

The distribution of *IHV* values does not seem to show any secular trend over the years. The solar cycle is (of course) clearly visible. To settle the question if this type of plot would show a secular trend clearly enough had there been a trend, we introduced an artificial trend equal to that claimed by Lockwood *et al*. The result is shown in Figure 76b and demonstrates that such a trend would have been clearly visible. The statistical argument here is that the set of days that in 1901 would have had an *IHV* index value of 10 (due to the then prevalent value of the interplanetary magnetic field, *B*) would in 1991 have an *IHV* value of 23.1 (due to *B* being that much higher and to the empirical relation: $IHV_{CLH/FRD} = 0.001589\ VB$ (km/sec nT, using data since 1966), neglecting the small effects of the changing dipole moment of the Earth, after all, Figure 76b is for illustration only).

The Atmospheric Potential Gradient

Harrison (2002) reports that the current flowing in the global atmospheric electrical circuit substantially decreased during the twentieth century. Fair-weather potential gradient observations in Scotland and Shetland show a previously unreported annual decline from 1920 to 1980, when the measurements ceased. A 25% reduction in PG occurred in Scotland 1920–50, with the maximum decline during the winter months. This is quantitatively explained by a decrease in cosmic rays increasing the thunderstorm-electrosphere coupling resistance, reducing the ionospheric potential. The data after about 1960 is suspect because of nuclear weapons testing, but the decrease 1920-50 seems well established. This is consistent with the cycles up to #19 being more active.

The Godhavn Observatory

At the General Assembly of the International Union of Geodesy and Geophysics (IUGG) at Madrid on October 24th, 1924, the section on Terrestrial Magnetism and Electricity adopted the following resolution:

> "Because of the geographical position of Greenland and of the importance of continuous measurements of magnetism and electricity in the auroral regions for the study of terrestrial magnetism and electricity, it is highly desirable that a permanent observatory dedicated to these subjects be established at the most suitable location on the West coast of Greenland."

The then director of the Danish Meteorological Institute (DMI) *Dan la Cour* persuaded the Minister of the Interior, *C.N. Hauge*, to propose to the Danish Parliament that a permanent observatory be established at Godhavn, on the island of Disco on the West coast of Greenland. The project was approved in the spring of 1925 and on February 1st, 1926, the new observatory started operations. Early observers include *G. Ljungdahl*, *Johannes Olsen*, and *Knud Lassen*, assisted by the Greenlander *Ole Mølgård*.

We owe to the foresight of the founders of the observatory and to the wisdom of the Danish Government in funding the observatory continuously to this day (and hopefully beyond) the existence of a virtually unbroken series of high-quality magnetic measurements going back now more than three-quarters of a century. The original equipment has, of course, been upgraded to modern instruments, but the quality and the care in data reduction have stayed true to the high standards set by the early observers.

Godhavn is situated at 69.25º N geographic latitude, 306.47º E geographic longitude. The geomagnetic coordinates of the observatory have changed somewhat over the years due to movement of the magnetic poles and also depend on the reference model of the geomagnetic field that is used. In the 1926 yearbook, the geomagnetic coordinates are quoted as 79.9º N and 32.5º E. Using the GEO-CGM code (v.9.9) provided by NSSDC, the "corrected" geomagnetic coordinates come out as 75.8º N, 40.4º E. Through the 75 years the observatory has operated, its geomagnetic latitude has decreased about 5º. This means that Godhavn has become less of a polar cap station over time. We may note that although the observatory retains its international station designation *GDH*, it should today properly be called *Qeqertarsuaq*, the local Eskimo name for Godhavn.

Getting and Scrubbing the Data

Hourly values of the three components (X, Y, Z or H, D, Z) of the geomagnetic field at Godhavn were obtained from the World Data Center for Geomagnetism in Kyoto, Japan. As always, when converting data to machine-readable form, clerical errors creep in. A typical example is the Z-component for the Second Polar Year, 1932-33, where the tables in the printed yearbook contain deviations from a base value Z_0. The tabulated values are defined as the ΔZs that give the value of the Z-component from $Z = Z_0 - \Delta Z$. All the other published yearbooks for Godhavn use the formula $Z = Z_0 + \Delta Z$. The difference in sign seems to have escaped the data-entry personnel. In 1976, small shifts in the base values for H (+50 nT) and Z (+285 nT) occurred due to modernization of the observatory (moving to new nearby buildings). Data prior to 1976 have been adjusted to eliminate the base line jumps. Correction of the clerical errors and base line adjustments result in a very homogenous data set extending from 1926 through 2001 consisting of more than 650,000 hourly values for each component. Figure 1 shows the run of monthly means of the three components.

Secular Variation

The secular variation of the components is clearly shown in Fig. 76. Most of the secular variation is caused by changes in the general geomagnetic field. The initial fear that the field would be contaminated by variable sedimentation of magnetic basaltic sand into the sea by the nearby *Røde Elv* (*elv* means river) seems to be unfounded. It is interesting to note that the local sand is so magnetic that it was deemed unsuitable for use in the construction of the observatory. It was considered to bring the 100 tons of sand needed all the way from Denmark. Luckily, a source of non-magnetic sand was found on Kronprinsens Eiland about 100 km away. The local anomaly caused by the sand and the generally rather magnetic basaltic bedrock amounts to an almost constant vector difference of about (+240, -270, +780) nT from the values computed from the IGRF model, that apart from the constant difference reproduces the secular variation very accurately.

Our interest in the secular variation in the present paper is mainly to be able to *remove* it, as we are interested in variations of the geomagnetic field that have external causes.

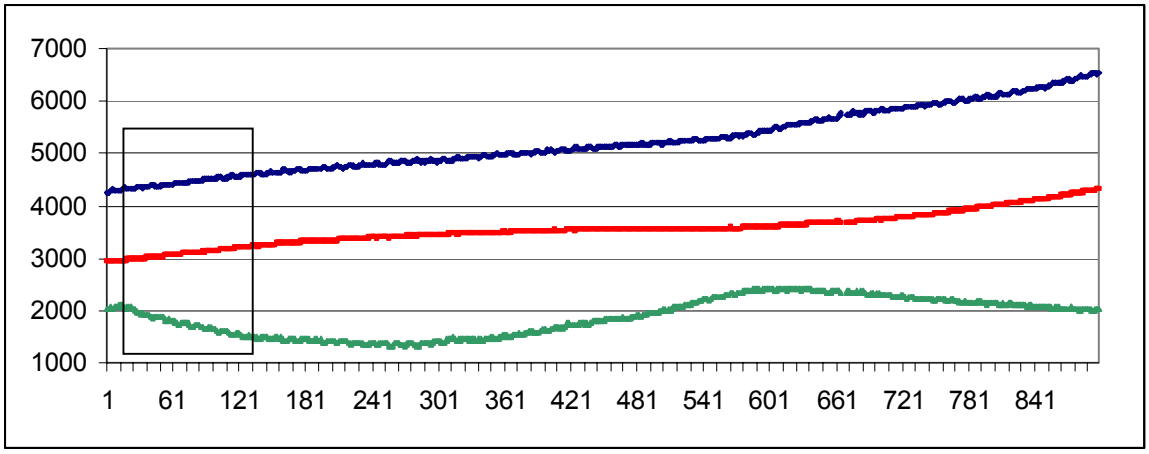

Figure. 77. Variation of the X (blue), Y (red), and Z (green) components at Godhavn 1926-2000. The curves have the same scale (in nT), but have been shifted vertically to fit within the same ordinate range. Unsmoothed monthly means are shown. The "box" outlines a portion of the graph that is shown magnified in Fig. 78.

We first note that there is no sign of a sunspot cycle variation in any of the components. There does seem to be persistent "wiggles" suggesting a *seasonal* variation. To this we turn next.

Seasonal variation

Figure 78 shows a ten-year, magnified slice (1927-36) of the variations of the three components.

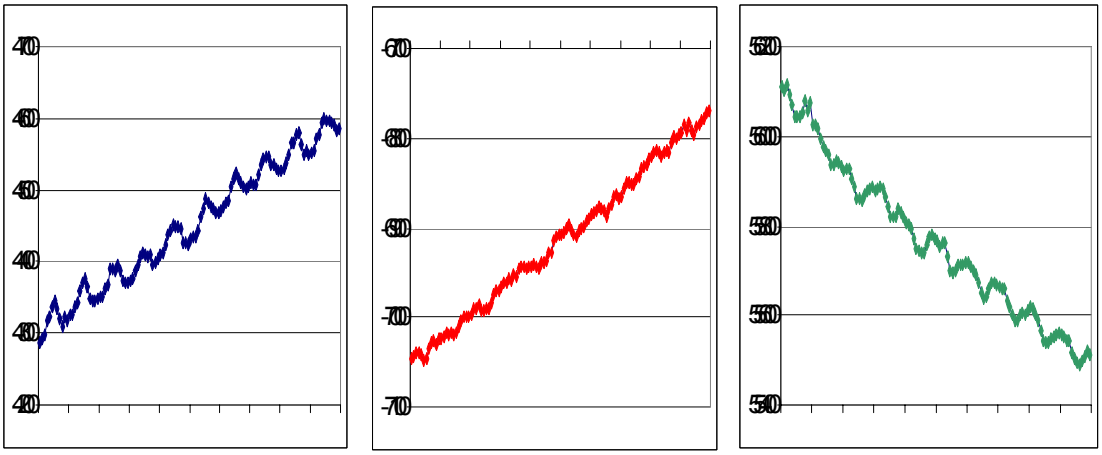

Figure 78. Variations of the X (left), Y (middle), and Z (right) components during 1927-36 showing a clear seasonal variation superimposed on the general secular variation of the X and Z-components, with a much smaller variation of the Y-component. Monthly means are plotted.

We could bandpass-filter the data to remove the secular variation (e.g. subtract a one-year running mean). This might also remove some physics and make it harder to decide if the seasonal variation is a wave with equal positive and negative excursions, a summertime maximum, a wintertime minimum, or has some other morphology. Several authors [e.g. Troshichev *et al*., 2000] have argued that the values observed on very quiet winter days should be used as a reference level, so that the secular variation can be removed by subtracting this very quiet reference level

interpolated to the day in question. Little justification has been given for this (seemingly very reasonable) assumption.

Definition of Background Field Level

The geomagnetic field as observed is a superposition of a very slowly varying background field of (almost entirely) internal origin and a rapidly varying additional component of (almost entirely) external origin (including associated induced contributions from the earth and the sea). Removing these rapid time variations is a problem in magnetic surveying. We are having the opposite problem here. The rapidly varying component is at times of very small amplitude. The frequency distribution of the measured field components (hourly values in our case) over a given period (e.g. a month) is observed at times to be sharply peaked, especially during winter months. Figure 79 shows the frequency of hourly mean values of the X-component during December 1997 compared with distribution during June of the same year. The sharp winter time peak at 6358 nT is evident compared to the much broader summer time distribution.

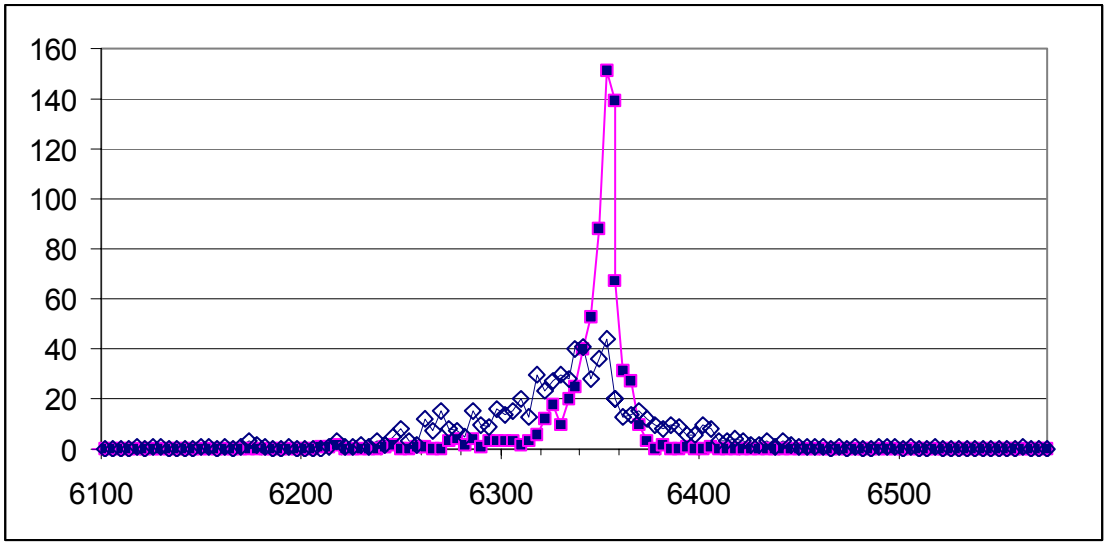

Figure 79. Distribution of hourly values (in 4 nT bins) of the Godhavn X-component during December, 1997 (filled squares) and during June, 1997 (open diamonds).

We can proceed to construct histograms like Figure 79 for every month of the entire dataset and to locate the peak on each histogram. The time series of values of X (or Y or Z) where the peak for each month is located could be used as a definition of a "background" field in the sense of "*minimum variability*". One could have hoped that this time series would not show any seasonal variation thus immediately making our definition useful. Of course, this turns out not to be so. Nature is subtle.

The interest in removing the background level comes about because there are gaps in the data. For example, if we want to compare the response of a field component to the polarity of the interplanetary magnetic field we could compute the average hourly value for a given hour separately for each polarity and compare the results. If now there were by chance more observations of a given polarity during time intervals where the secular values of the component were high and more observations of the opposite polarity during intervals where the secular values of the component were low, the difference between the hourly means for the two polarities would be severely contaminated by the combination of secular variation and uneven data coverage and might be completely swamped by the artificial difference due to uneven data availability. Similar considerations apply for the effect of missing data on superposed epoch analyses.

Figure 80 shows a contour plot of the distribution of values of the Godhavn X-component for all the years where we have data. The plot was generated by counting, for each month, how many times an hourly value fell within 4-nT wide bins and graphing the result. There is a very large amount of information in this plot,

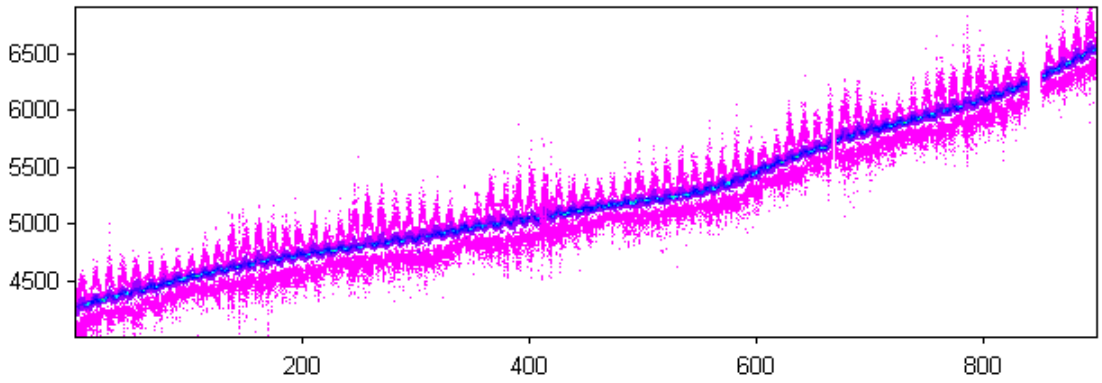

Figure 80. Distribution of values of the Godhavn X-component (in 4-nT bins) for each month during the entire 75-year interval from 1926 through 2000. In addition to the gap in 1996, two smaller gaps are visible. The abscissa shows the number of the months since the beginning of the observatory.

so it merits careful study. A small piece of the plot is shown in Figure 81: The wintertime peaks are very narrow and well defined. For any interval of a few years one can confidently fit an almost straight line to their positions. This line defines our "minimum variation" background reference level. The other field components show very similar behavior and reference levels for them can be defined in a like manner. The yearly variation is very marked as well, and there is a clear sunspot cycle variation, with the yearly variation being larger during sunspot maximum years.

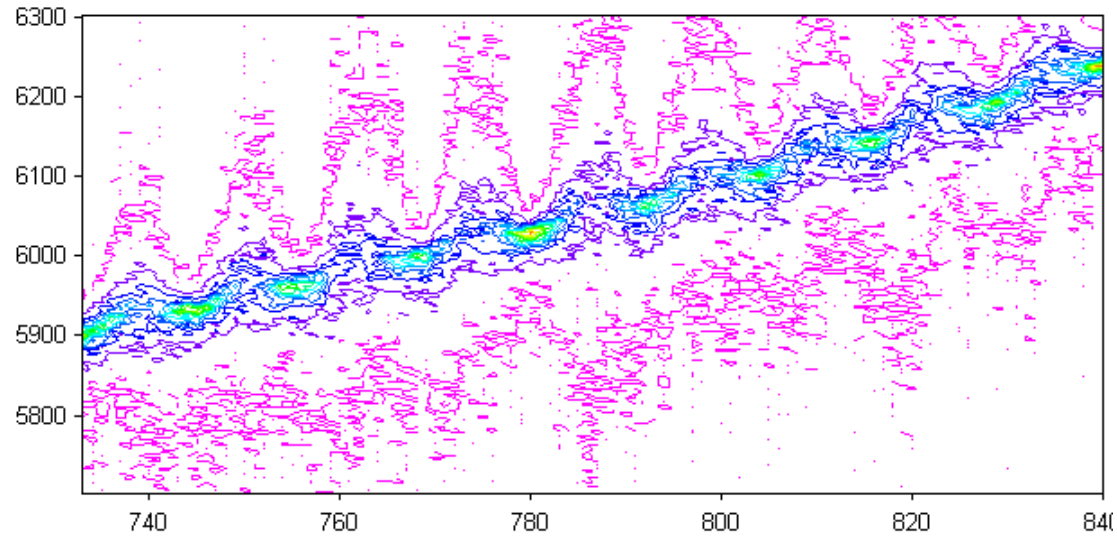

Figure 81. Detail of Figure 80 for the years 1987 through 1995. Any other interval looks pretty much the same as this. One cannot from inspection of the Figure tell which interval is shown.

A 3D-view of the same data is shown in Fig. 82. The sharpness of the wintertime peaks is particular striking in this representation.

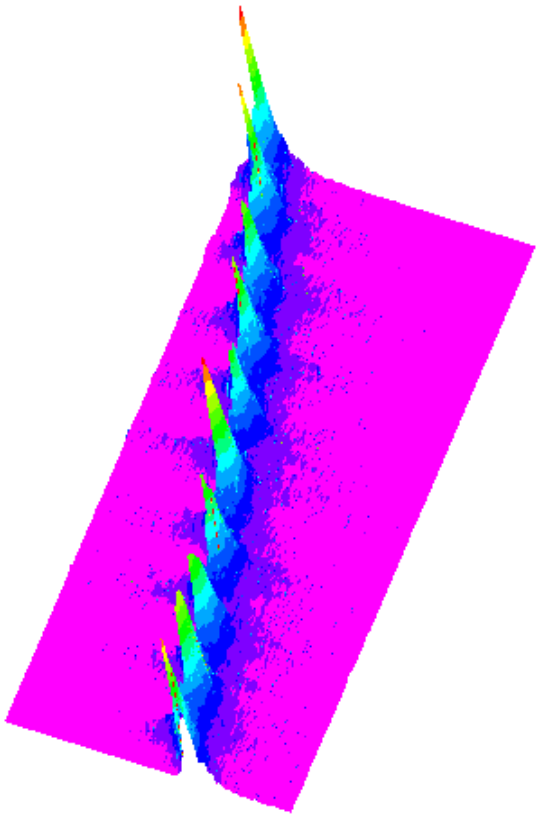

Figure 82. 3D-view of the frequency of values of the X-component at Godhavn from 1987 through 1995.

We choose the center of the bins where the frequency distribution peaks during the months of December and January as representative of the base reference level for the field component in question and assign the mean of these to values to the start of the year. Linear interpolation between successive such values give us the base reference level for each day.

The high quality of the Godhavn data is crucial in securing a background level that is almost free of instrumental bias and that is remarkably stable given the logistical difficulties in the Arctic. The present author, having been an observer in Greenland himself, appreciates the effort and care that have gone into the Godhavn observations.

Fig. 83 shows the frequency distribution with the reference level removed. A sunspot cycle variation is clearly visible. It is also clear that there is a pronounced and well-organized yearly variation in positive values (above the reference level) while the yearly variation in the negative values (while present) is much less distinct.

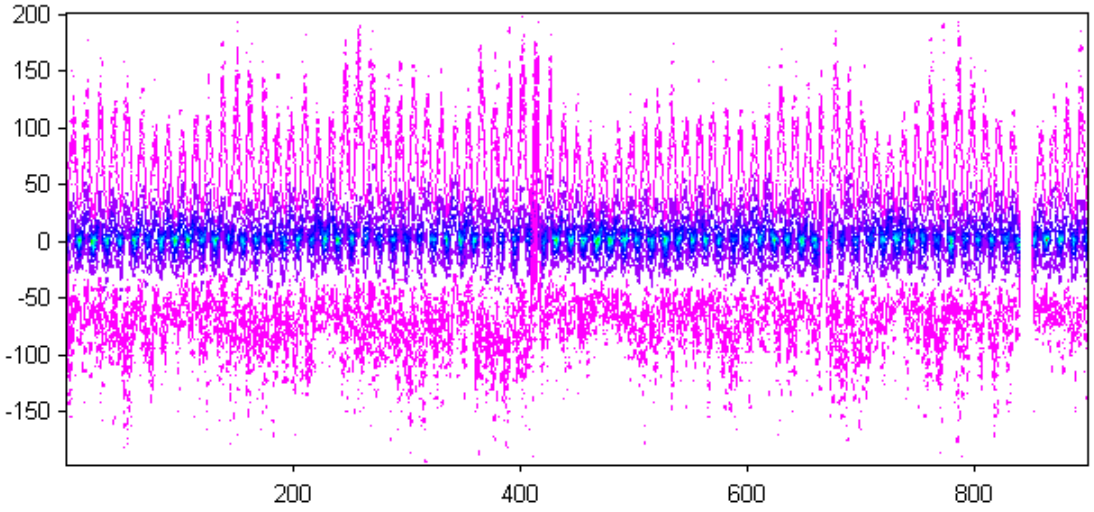

Figure 83. Distribution of values of the Godhavn X-component (in 4-nT bins) for each month during the entire 75-year interval from 1926 through 2000. The “minimum variability” reference level has been subtracted from the component values. Three data gaps introduce artifacts in the contour levels. No Long-term trend is apparent.

Fig. 83 showed the distribution for each year separately. Computing the *average* frequency distribution over the year produces Fig. 84:

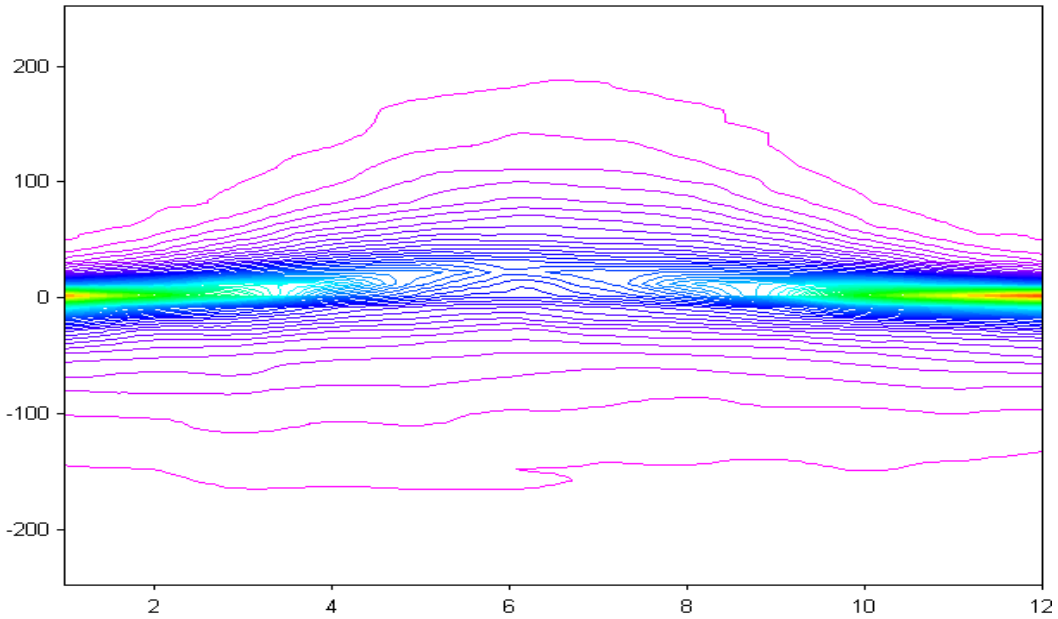

Figure 84. Average distribution of Godhavn residual X-component values (after removing the base reference level) through a year computed from the entire interval 1926-2000. The abscissa is in months and the ordinate is the bin in nT units. The bins were 4 nT wide and the count is plotted at the center of each bin. The vertical axis is the average count for each bin over a period of a month. This count ranges from zero to 105.

The distribution has an interesting fine structure. This is shown in Fig.85. The distribution is bimodal with a horizontal branch that extends throughout the year centered on zero (the base level). The existence of this branch shows that the "underlying" base level does not change with the seasons. We had determined the reference level from the wintertime peaks and then interpolating linearly through the rest of the year. It is reassuring to observe that the level so determined actually is visible as the horizontal branch. There is another branch of the distribution corresponding to the annual position values variation.

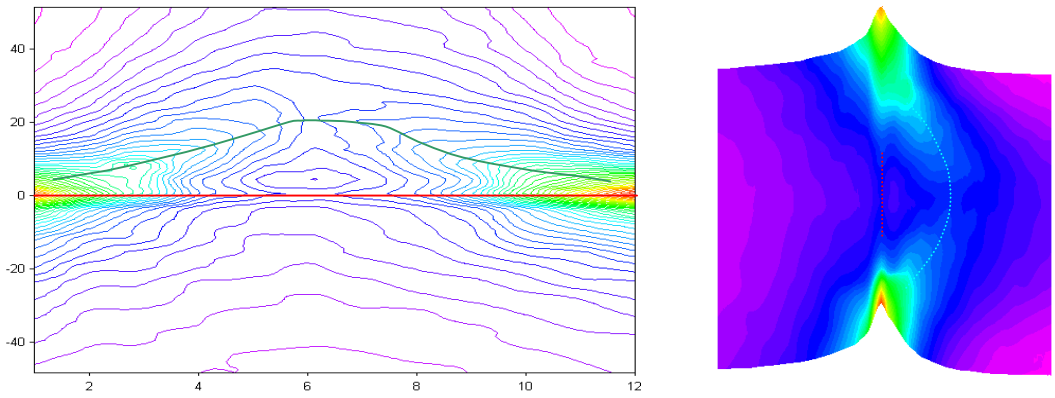

Figure 85. Enlarged view of the middle part of Figure 84 shown both as a contour plot and in a 3-D view (rotated 90º). The horizontal branch of the distribution is highlighted with a red line and the other branch is highlighted with a green line. The abscissa is labeled in months and the ordinate in nT.

To investigate if the annual variation is related to general geomagnetic activity, we repeat the above analysis separately for the five International Quiet Days and for the five International Disturbed Days of each month. Since 1932, *ten* International Quiet Days per month are defined and ranked from the most quiet to the least quiet. We pick the five most quiet of the ten. Before 1932, only five Q-days per month were defined. The results are shown in Figures. 86a and 86b:

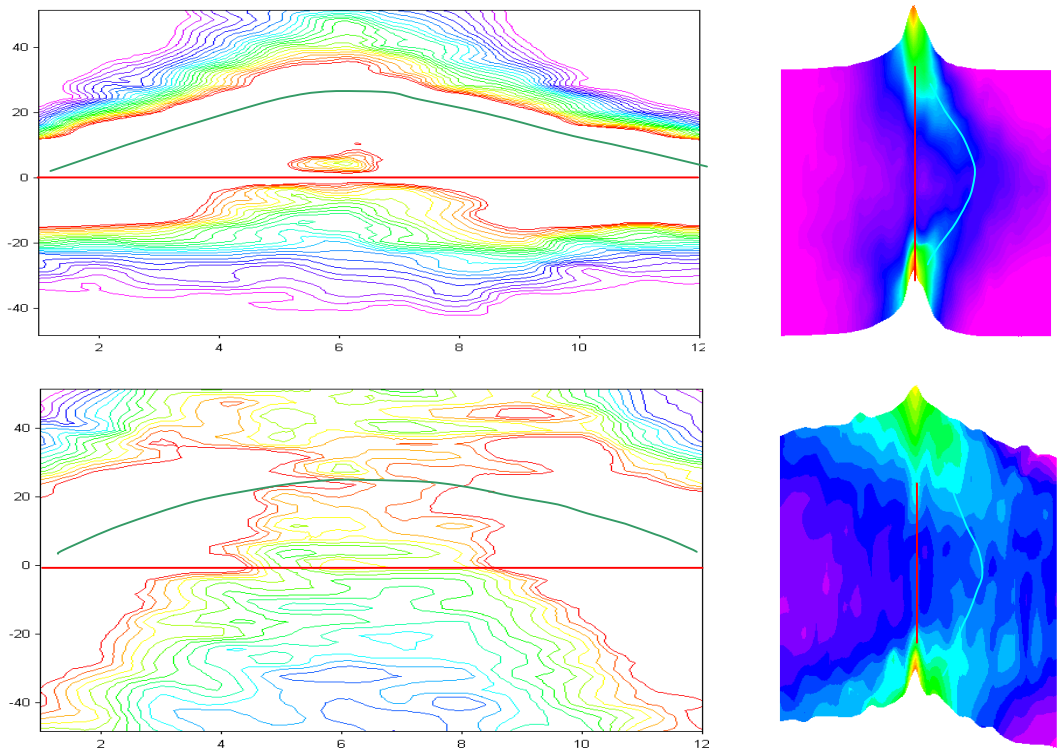

Figure 86. (a; upper panel) Frequency distribution of Godhavn X for International Quiet Days and (b; lower panel) for International Disturbed Days in the same format and for the same interval as in Figure 84. The contour plots were deliberately saturated to bring out the structure at low count values.

As before, it is reassuring to see that the reference level is visible as a horizontal branch for both Quiet days (Figure 86a) and for Disturbed days (Figure 86b). Remarkably, the summertime maximum in the annual variation (defined by the trace of higher counts in the frequency distribution) seems to be of the same magnitude regardless of the disturbance level.

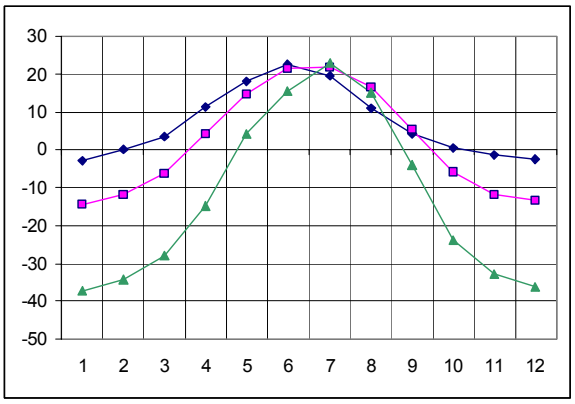

Figure 87. Annual variation of Godhavn X-component residual values (after removing the base reference level) for quiet days (blue diamonds), "ordinary" days (pink squares), and disturbed days (green triangles). [1926-2000]

The summertime peak is the same for all levels of disturbance. Note the slight shift of phase for disturbed days.

The analysis of the range of variations at Godhavn (especially as show in Figure 83 strongly supports the statement that no long-term trend in activity is present. At usual, some solar cycles are stronger than others are and as expected cause activity to be duly higher for these cycles.

Sudden Increases in Solar Activity?

Duhau and Chen (2002) report a "sudden increase of solar and geomagnetic activity after 1923". Although their paper is not very clear to me, there is evidence from the number of occurrences of geomagnetic "sudden commencements" that sudden increases might be observed. Figure 88 shows the number of sudden commencements (SC) since 1868:

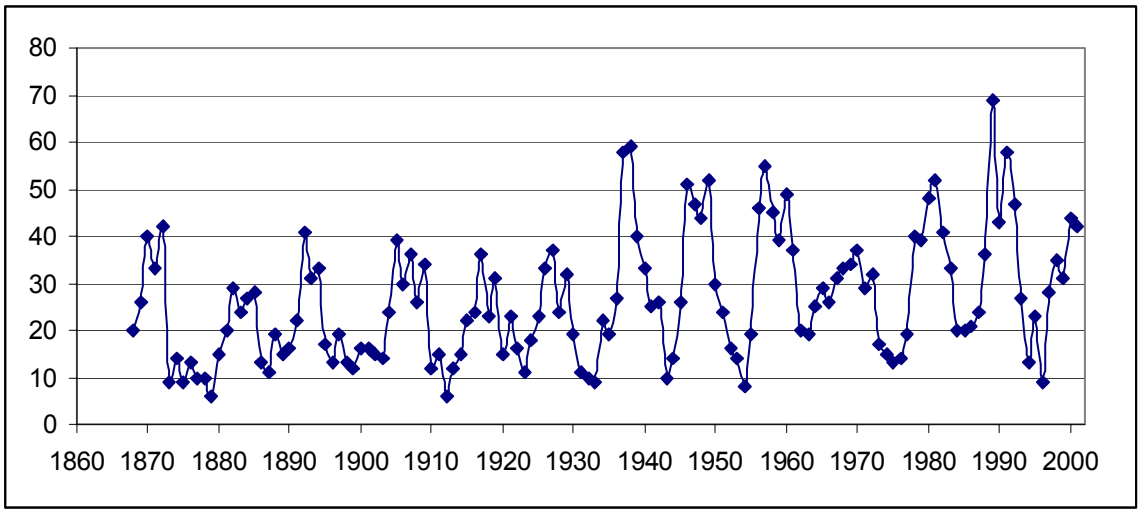

Figure 88. Number of "Sudden Commencements" per year during the interval 1868-2001.

At solar minimum the number of SCs is rather constant (around 10) with the exception of the minima in 1965 and 1985. At solar maximum the count seems to follow the sunspot number. The change around 1935 does seem very abrupt and the possibility of a selection effect exists (higher sensitivity or such). It is interesting to note that since 1993 the count is very similar to the pre-1935 counts.

The human brain (and that of most other animals) has evolved to be alert to abrupt changes and we may make too much of the sharp increase around 1935. If we simply plot the yearly count of SCs against the yearly mean sunspot number we get a much less alarming relationship:

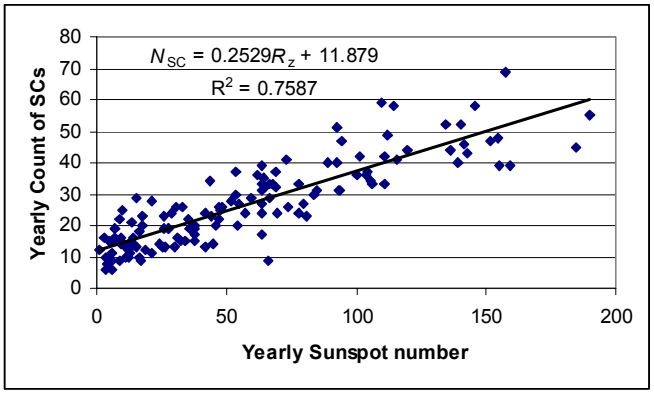

Figure 88. The number of sudden commencements per year as a function of the yearly mean sunspot number for the interval 1868-2001.

This plot seems simply to show that more sunspots means more sudden commencements, which is not unexpected and with which scientists of half a century ago would certainly agree without much wonderment

The cause of the higher counts during the minima 1965 and 1985 is not known and some people might hang the whole argument of a doubling of the Sun's coronal magnetic flux on those few years.

To eliminate the psychological effect of dramatic changes, we show in Figure 89 the result of removing the solar cycles with highest sunspot numbers. This should reveal if any true, underlying secular change exists:

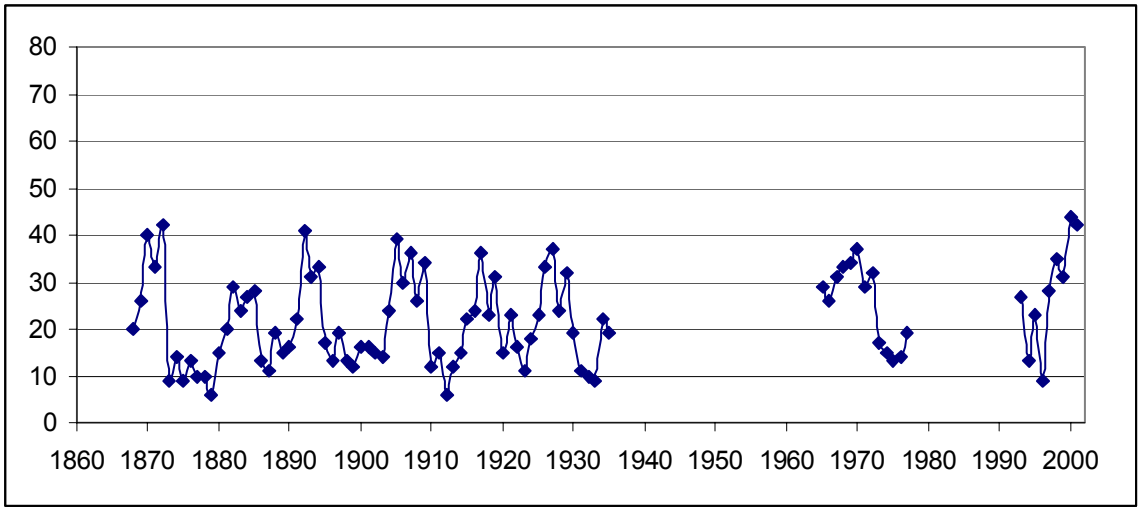

Figure 89. Yearly count of sudden commencements for solar cycles with a maximum sunspot number less than about 120

There are features about the counts of sudden commencements that are puzzling and not understood. Figure 90 shows the average number of SCs per month:

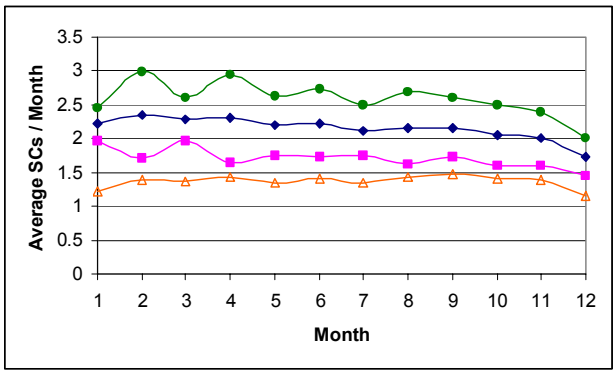

Figure 90. Average number of sudden commencements (SCs) per month (normalized to 30 days of length) for several intervals: 1868-1935: pink squares, 1936-2001: green circles, and entire interval 1868-2001: blue diamonds.

Orange triangles show the lack of a clear semiannual variation. Detrended values are plotted offset by a count of one for the orange triangles.

There is a significant decrease through the year from January to December. This decrease shows up in both the data before 1936 and after that year. If we remove this seasonal trend there is no clear semiannual variation of the counts although one might find a weak hint. The outstanding feature is the strange decrease of 22% in the count from January to December.

Sudden Commencements

Sudden commencements (for a recent treatment see Tsunomura, 1998 and references therein) are caused by a sudden compression of the magnetosphere as the result of a sudden increase in the dynamic pressure of the solar wind and are therefore not directly related to the interplanetary magnetic field. The sudden increases of the dynamic pressure are, however, indirectly correlated with the magnetic field as well as with all the other solar wind properties because they are related to shock waves that compress the solar wind and increase the magnetic field in the compression region. There is a well-known *signature* of a SC in solar wind properties (*e.g.* Gosling *et al.*, 1991). The solar wind speed jumps at the time of the SC and stays high for several days. The solar wind density peaks sharply at the SC, but then decreases quickly and after a day becomes lower than before the SC (in the so-called rarefaction region). The interplanetary magnetic field also peaks sharply at the SC, then decreases, but less rapidly

than the density, to reach the pre-SC value about 2.5 days after the SC. Figure 91 shows superposed epoch analyses of solar wind parameters and of *aa* from one day before the SC to four days after the SC:

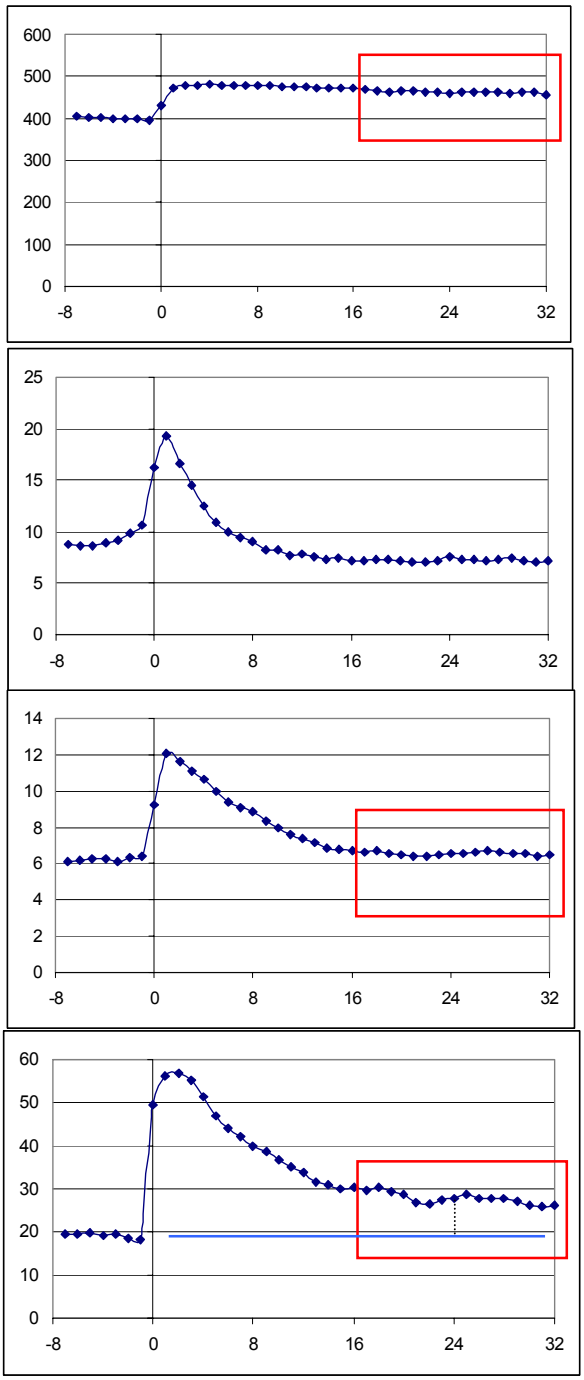

Figure 91. Top: Solar wind speed (km/s) from one day before a SC to four days after the day using 750 SCs. Next: Solar wind density (protons/cm^3) for 700 SCs. Next: Interplanetary magnetic field strength (nT) for 800 SCs. Bottom: *aa*-index for 1197 SCs. The analyses cover all SCs in the interval 1964 through 2001 whenever data was available. Various gaps explain the different number of SCs used for each Figure.

We note that on days three and four after the SC, the magnetic field is down to almost its pre-SC level, while the solar wind speed remains elevated. This is duly reflected in the *aa*-index that although having reached a stable plateau is still elevated over its pre-SC value. This observation gives us the **Holy Grail** of proxy-research, namely a means of *separating* the influence of the magnetic field and of the solar wind speed. This was the main problem for Lockwood *et al.* as well that they tried to solve invoking the recurrence index.

Because B has returned to its pre-SC level, the elevation of the *aa*-index during days three and four is due to the solar wind speed V alone. Small, second order corrections can be made for the influence of the density (which is weak anyway) and for the residual enhancement of B, but are really not needed.

An expression of the form

$$aa = 0.114\, B\, V^{2.362} \qquad (23)$$

is an excellent fit (with $R^2 = 0.99$) to the observed *aa*-index response to the passage of an SC depending on the values of the interplanetary magnetic field B in nT and of the solar wind speed V in 100 km/s units. The formula pertains to values after 1963.

If the superposed epoch analysis is done for a longer period after the SC, we simply see "more of the same". The solar wind speed continues to drop, but slowly. The magnetic field reaches the pre-SC level and stays there, and the density continues to recover slowly to its pre-SC value.

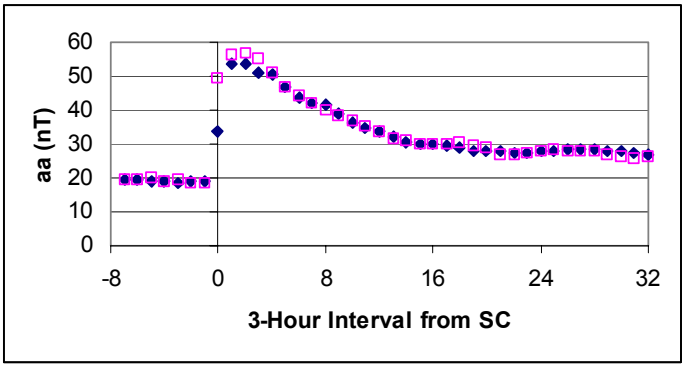

Figure 92. Superposed epoch analysis of the *aa*-index in response to a sudden commencement. 1197 SCs in the interval 1964-2001 were used. Observed values of *aa* are shown as blue diamonds, while the open pink squares show the response calculated from eq.(23) using the observed response of *B* and *V*.

Because the density was not incorporated in eq.(23), the calculated values for *aa* are slightly lower just after the SC where the density has a sharp peak, but otherwise the fit is very good away from the density peak.

We can perform a similar superposed epoch analysis separately for every sunspot cycle since number 11 where the *aa*-index begins. Figure 93 is a very compact representation of the result:

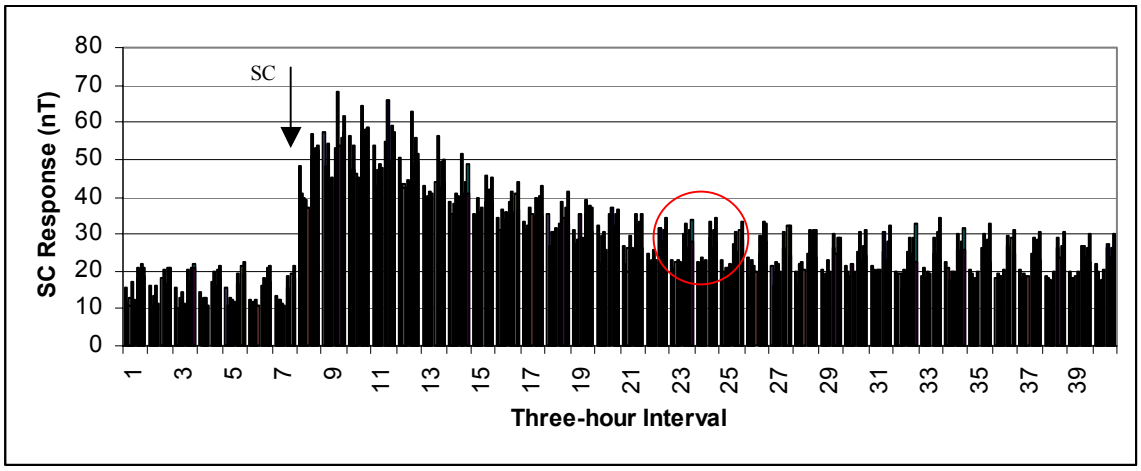

Figure 93. Superposed epochs of *aa* keyed to the time of sudden commencements. For each three-hour interval 13 touching vertical bars are plotted, in sequence for solar cycles 11 through 23. The height of each bar is the average value of *aa* for that three-hour interval at that epoch.

One can see the evolution of the response to an SC with time within each set of vertical bars for each three-hour interval. It is apparent from Figure 93 (examine the variation within the red circle) that the first six solar cycles show a very similar response, but then the whole response seems to be shifted uniformly upwards by 5-10 nT for the last seven cycles. It is probably not a coincidence that closer examination shows that the "shift" takes place between 1937 and 1938, coincident with the instrumental change at Abinger identified by Clilverd *et al*.

We next perform the superposed epoch analysis for the two intervals, before 1938 and after the start of 1938. The result is shown in Figure 94. The responses are very similar except for the "shift" between the two curves. To the extent that the interplanetary fast stream that ploughs into the slower moving ambient solar wind compresses the material in front of it to cause the characteristic velocity-density response one would expect that the "scooped up" material would reflect properties of the ambient solar wind. If, for instance, the interplanetary magnetic field was higher during a certain epoch (as possibly indicated by the level of activity before the SC) the compressed material would also have a higher magnetic field and its geomagnetic effect would be correspondingly greater.

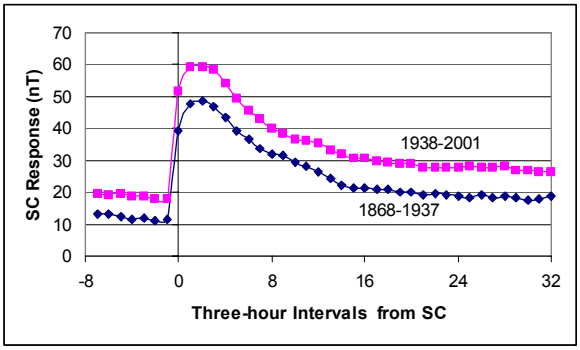

Figure 94.Superposed epoch analysis of the response of the *aa*-index to a sudden commencement. Blue diamonds show the average response to the 1503 SCs that occurred in the years 1868-1937. Pink squares show the average response to the 2054 SCs that occurred in the years 1938-2001. The analysis extends from 1 day before the SC through 4 days after the SC.

The pre-SC value of *aa* in Figure 94 is 12.1 nT before 1938 and 18.9 nT thereafter. If eq.(23) is valid (Lockwood *et al*. also assume that things behaved in the past basically as they do now), then the response should scale in the same ratio (we assume first that the solar wind speed V behaves similarly). For the pre-1938 data, the response peaks at 48.8. This should scale with a factor of 18.9/12.1 = 1.56 to predict a response of 48.8 * 1.56 = 76.2 nT. This is not what is observed; the two curves are very nearly homologous with a much smaller scale factor (only 1.09), but with a significant *offset* (the "shift"). This is illustrated in Figure 95:

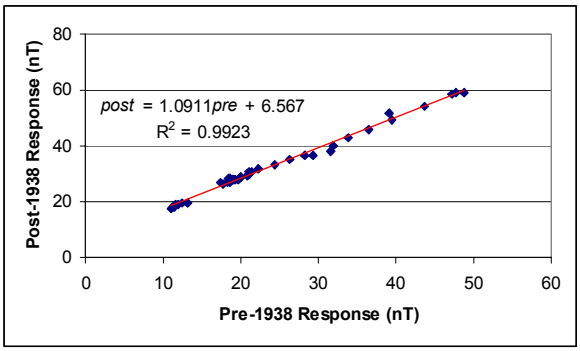

Figure 95. The average response to an SC for the interval 1868-1937 scales to the average response for the interval 1938-2001 with a scale factor of 1.09 and an offset of 6.6 nT. The similarity between the two responses is extremely high ($R^2 > 0.99$).

If we remove the shift from the post-1938 response, we get the result in Figure 96:

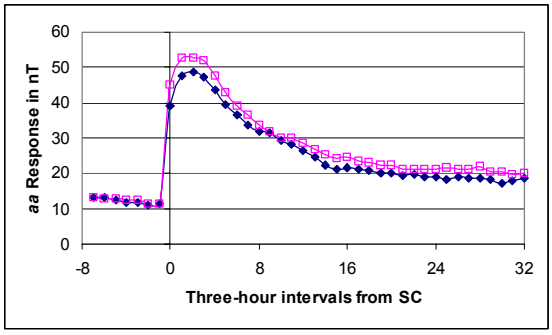

Figure 96. Superposed epoch analysis of the response of the *aa*-index to a sudden commencement. Blue diamonds show the average response to the 1503 SCs that occurred in the years 1868-1937. Open pink squares show the average response to the 2054 SCs that occurred in the years 1938-2001. An offset of 6.6 nT has been subtracted from the post-1938 response.

The 9% difference between the amplitudes of the responses for the two time intervals is of approximately the same size as the change in calibration of *aa* between the two intervals, consistent with the statement that the response is identical for both intervals if we attribute the shift in base level to an artifact in *aa*. This is actually a puzzling result, as I would expect some clear difference considering that the average sunspot number for the pre-1938 interval was 41.87, while is was almost double that, 75.08, for the post-1938 interval (there is your doubling of the Sun's magnetic field!). It seems that each SC has about the same effect, but that there are more of them when the sunspot number is higher, so their cumulative effect is higher when the sunspot number is higher.

Response of the Horizontal force to SSCs

Studying the geomagnetic field in Bombay, Nanabhoy Moos (1910), the original director of the Indian Institute of Geomagnetism, discovered the existence of a unique pattern in geomagnetic disturbances confirming an observation already made by Broun (1861). Arranging the data in time from the start of the disturbance he discovered the 'classic' geomagnetic storm signature: the positive H-component onset of disturbance, followed by a rapid depression of H, and finally a slow recovery to the previous quiet level. This pattern was further investigated by Sidney Chapman (1918, 1935) who used the name '*Dst*' for this average storm-time presentation of field disturbances. Strictly speaking, the regular daily variations (mostly *Sq*) and main field levels should be removed first. Chapman (1951) introduced the now familiar terms, 'sudden commencement', 'initial phase', 'main phase', and 'recovery phase' to describe the typical *Dst* storm characteristics. Campbell (1996) reviews (and criticizes) the conventional view of ascribing *Dst* to the magnetic effects of first a compression of the magnetosphere by a shock wave from the Sun, followed by the development and subsequent decay of a massive equatorial ring current at about 4-8 Earth radii distance. Possibly, the main and recovery phases are just the combined and overlapping effects of many smaller injection events that each decay on their own (much shorter) time-scale. Whatever the exact mechanism, the unique signature remains.

We have performed such a 'storm-time' superposed epoch analysis on the horizontal component, H, at Cheltenham/Fredericksburg using the sudden commencement list as key times. The initial impulse, main and recovery phases are clearly recognizable. Because the SSCs occur at all times during the day, *Sq* averages out. We subtract the main field by subtracting the value 13 hours before the SSC (because the approaching shock wave begins to have effect about 12 hours before the SSC (see later)). The result is a close approximation to *Dst*, except that we do not divide by the cosine of the dipole latitude - in keeping with our policy of introducing as few calibration factors as possible. In years where geomagnetic activity as measured by the *aa*-index is high, the magnitude of the main phase as measured by *Dst* is also found to be large (no surprise). This is clearly seen in Figure 97, where the analysis was done for the interval 1959-2000, but separately for years when the yearly average of *aa* was less than and greater than the overall average for 1959-2000, respectively:

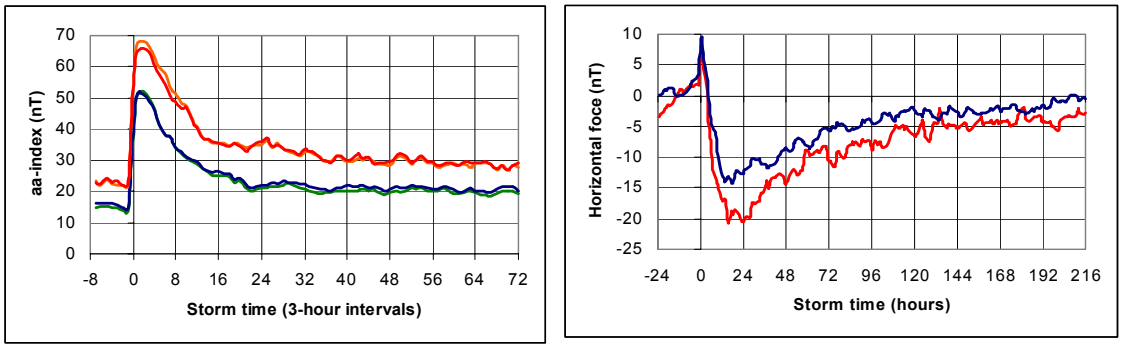

Figure 97. (Left) Superposed epoch analysis of *aa* and *am* for the years when both indices are available (1959-2000). The upper curves (*aa* red, *am* orange) are for years when the yearly average of *aa* was greater than the average value (22.6 nT) for the interval. The lower curves (*aa* blue, *am* green) are for years when the average value of *aa* was smaller than the overall average. The difference between the *aa* and *am* responses is very small, attesting to the stability of the *aa*-index since 1959.

(Right) Superposed epoch analysis of H at Fredericksburg for the same interval as above for *aa*. The upper curve (blue) is for years with higher than average *aa*. The lower curve (red) is for years with lower than average *aa*. The typical number of SSCs used for each plot is between 400 and 700 depending on the grouping of years and on (the small amount of) missing data.

The main purpose of Figure 97 was to demonstrate that the *Dst*-like response is capable of distinguishing between years that have genuinely different average values of the *aa*-index. The years of the two groups are distributed throughout the interval of analysis, and the close similarity of the *aa* and *am* responses validates the *aa* values for the interval. We infer that a difference in the *aa*-response corresponds to a similar difference in the *Dst*-response in a

straightforward manner. Because the yearly average of *aa* itself was used as the selection criterion we would expect that the pre-SSC values of *aa* would be substantially different for the two groups of years, which, in fact, they are.

We now look at the opposite question. Figure 98 shows the result of repeating the analysis, except this time we divide the years (since 1959) in two groups depending on the average size of the *Dst*-like response for each year. We define the size of the response as the difference between the mean value of *Dst* (actually H).for storm-time hours 12 through 216 and the mean value of Dst for storm-time hours -24 through -2 (avoiding the rapid rise near storm-time 0). Because in some years there are rather few SSCs, the noise and scatter are at times considerable, so we only use the calculated size to place the year in one of the two groups: with less than average size and with more than average size:

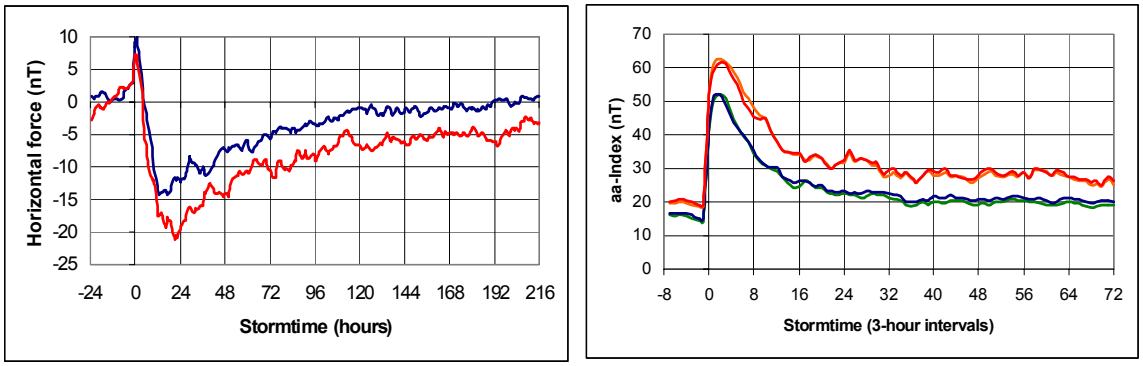

Figure 98. (Left) Superposed epoch analysis of *Dst* at Fredericksburg for the same interval as in Figure 97 for *aa*. The upper curve (blue) is for years with lower than average size of the H response. The lower curve (red) is for years with higher than average size of the H response.

(Right) Superposed epoch analysis of *aa* (blue and red curves) and of *am* (green and orange curves) for the same groups of years as above. In this panel, blue and red curves correspond to like colors in the left-side panel.

We infer that a difference in the *Dst*-response corresponds to a similar difference in the *aa*-response in a straightforward manner. Because the *Dst*-response (and not *aa*) was used a selection criterion, we would expect the pre-SSC *aa* values to be less different than was the case when we used *aa* as the selection criterion, as, in fact, we find (ideally, there should be *no* difference in pre-SSC *aa* values, but Nature is seldom this kind).

The interval 1959-2001 was chosen because we have the *am*-index derived from many stations available.

We now divide the years into three groups: prior to 1938, from 1938 through 1966, and later than 1966. Originally I divided the data into groups of equal size (33 years) using 1933 as the divider between group 1 and 2, but I changed that to 1938 because of the change of magnetograph at Abinger at that time. The division at 1966 was just left. Using the SSCs as key epochs we perform the superposed epoch analysis of *Dst* and *aa*. The result is shown in Figure 99:

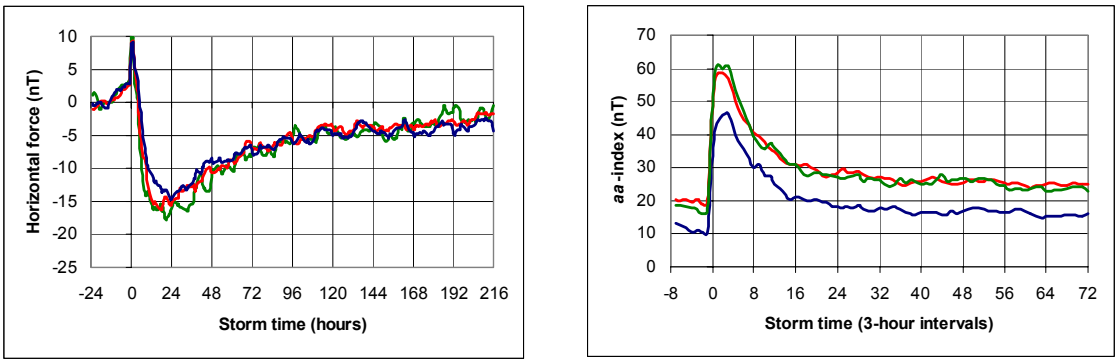

Figure 99. (Left) Superposed epoch analysis of *Dst* at Cheltenham/Fredericksburg. The blue curve is for 1901-1937, the green is for 1938-1966, and the red curve is for 1967-2000.

(Right) Superposed epoch analysis of *aa* for the same groups of years as above. In this panel, blue, green, and red curves correspond to like colors in the left-side panel.

There is little difference between the *Dst* responses for the three groups, but the *aa* response for the early years is clearly lower by about 10 nT.

“Winter Quarters”, Antarctica 1902-03

The National Antarctic Expedition of 1901-04 under Robert F. Scott had a set of Eschenhagen magnetographs, which were in operation at the Winter Quarters (on Ross Island) of the Expedition for nearly two full years (1902-03), under the immediate supervision of the physicist, Mr. L.C. Bernacchi. The magnetic curves were measured by Bernacchi and the results were discussed by Charles Chree (1912). Winter Quarters (at 77.8° S, 166.6° E) was very close to the present-day magnetic observatory Scott Base (corrected geomagnetic coordinates 79.96° S, 327.6 ° W). Figure 100 shows the location of the station:

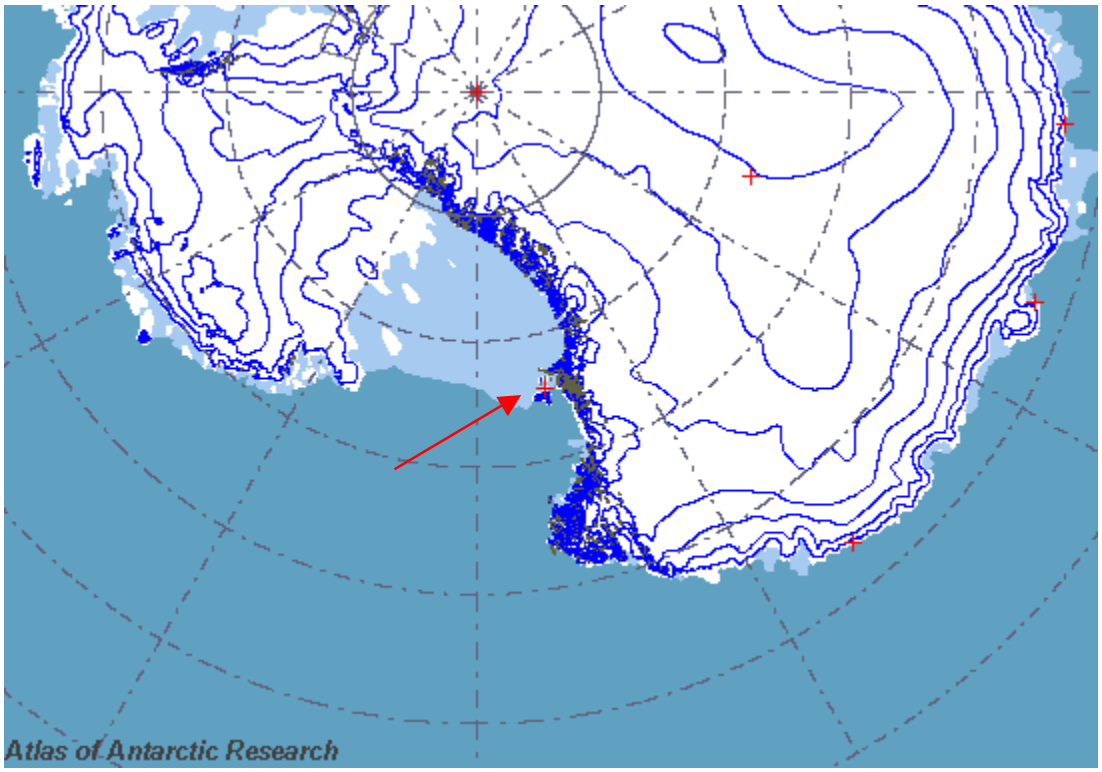

Figure 100. Location of Winter Quarters (and present-day Scott Base) on Ross Island

The magnetograph set up at Winter Quarters was unfortunately too sensitive for so highly disturbed a station and on the most disturbed days some loss of trace occurred, especially of the H-trace. All days of incomplete trace were omitted when deriving the diurnal variation, and (in the words of Charles Chree) “it is thus not improbably that the range given for H in Table XXXIV, is smaller than it would have been if all days utilized for the D or V (now we call it Z) inequalities had been available”. Here is a lesson of using the *same* days when averages are calculated. As we shall use the range of the calculated Y-component in our analysis, our results are not seriously affected by missing H trace.

Winter Quarters was a polar cap station in the same sense as present-day Scott Base and Godhavn and Gjøahavn in the Northern Hemisphere. All stations are at around 80° geomagnetic latitude. At quiet times near noon, the auroral ovals approach these stations and deform the diurnal variation, but during the nighttime the stations are well within the ovals. The following Figures show the locations of the stations and of the auroral ovals to put the observations in perspective and to emphasize how similar the stations are located within the polar caps.

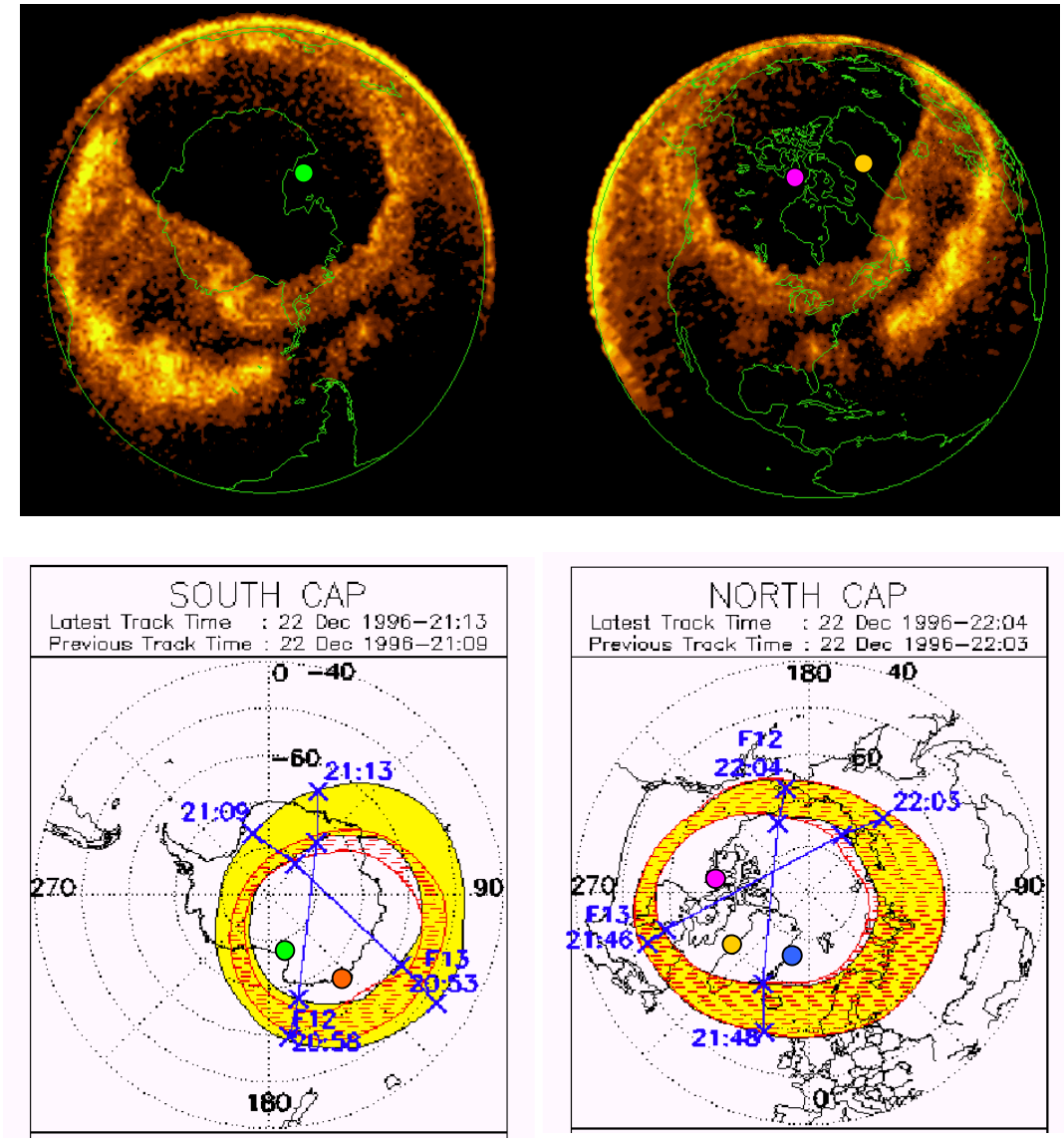

Figure 101. Top: Survey images from the *Dynamics Explorer-1* Spin-scan Auroral Imager's (DESAI) Vacuum Ultra-Violet (VUV) photometer (PI: L. Frank) from March 14, 1989. The picture of a large aurora was taken over the Southern Hemisphere only, but has been remapped along field lines to the Northern Hemisphere giving a close approximation to what the instrument would have seen there at this time. This is an example of a large polar cap.

Bottom: The auroral ovals as observed by the POLAR satellite (?). The location of Godhavn is shown by a yellow dot (on all plots of this Figure), that of Gjøahavn by a pink dot, that of Winter Quarters/Scott Base by a green dot, that on Adelie Land (see later) by an orange dot and that of Danmarkshavn (see later) by a blue dot. The activity levels for the bottom panels were moderate (K_p between 3 and 4), so the size of the polar cap is more representative of "normal" conditions. For very quiet times, the polar cap shrinks further. The plots were produced as a part of the UPOS project.

As in previous sections we can estimate the polar cap range (proportional the polar cap potential) from the diurnal range of the Y-component (or even better: the force component perpendicular to the local horizontal field direction, Y' - simply the D-component expressed in force units). Figure 102 shows the results.

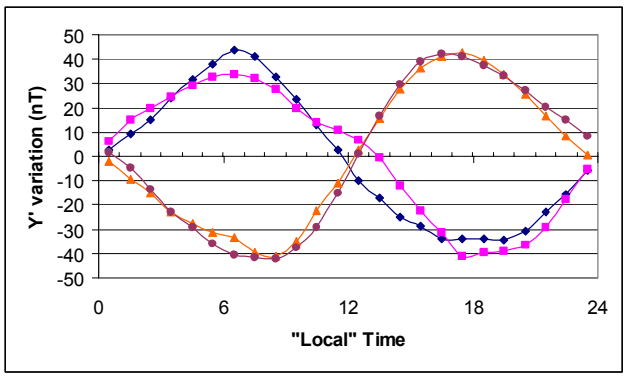

Figure 102. The diurnal variation of the Y'-component at two stations in the northern polar cap (Godhavn pink squares, Gjøahavn blue diamonds) and two stations in the southern polar cap (Winter Quarters orange triangles, Scott Base purple dots). The curves have been shifted so that the maxima and minima line up. This compensates for the very different time zones of the stations. At the nominal noon, the perturbations are zero as the magnetic effect is directed towards the Sun.

The average sunspot number for the two years 1902 and 1903 was 14.1. Selecting years with as close as possible the same number of sunspots and for which modern data from Scott Base is available we get 1964 (R_z=10.2), 1965 (15.1), 1975 (15.5), 1976 (12.6), 1985 (17.9), 1986 (13.4), 1995 (17.5), 1996 (8.6), and 1997 (21.5). These years were used to construct the curve for Scott Base in Figure 102. Similarly, the curve for Godhavn was constructed from years *after* 1964 for which the sunspot number was close the 42.1, that was the sunspot number for 1904, when data from Gjøahavn was available. We are thus comparing the polar cap potential for 1902-04 with the average since 1964. As far as we can tell, there is no difference between the two epochs, again not supportive of the claim that the Sun's magnetic field has doubled in the last 100 years. For this to have happened, the solar wind speed must have halved in the last 100 hundred years which does not seem credible.

The following table summarizes the results:

Average Epoch (year)	Average *Rz*	Station	Corr. Magn. Lat.	Polar Cap Range
1903.0	14.1	Winter Quarters	-80°	84.4 nT
1982.1 (During 1964-99)	16.0	Scott Base	-80°	83.9
1904.5	42.1	Gjøahavn	78°	78.1
1985.0 (During 1964-99)	37.9	Godhavn	76°	75.0

The polar cap range is generally 15% larger in the southern polar cap owing to its smaller size (the same potential drop has to take place over a smaller distance, so the gradient is higher). It is then no surprise that the ranges quoted above for the southern stations are slightly larger than those for the northern stations in spite of the smaller sunspot number for their epoch.

Cape Evans, Antarctica 1911-12

Magnetographs were in operations at Cape Evans, the base station of the British (Terra Nova) Antarctic Expedition during 1911 and 1912. Cape Evans (at 77.63° S, 166.40° E) is very close to Winter Quarters and Scott Base. For our intent and purpose, they can be considered one station. A second set of magnetographs was in operation at the base station of the Australasian Antarctic Expedition in Adelie Land (at 67.00° S, 140.67° E, very close to the present-day station Dumont d'Urville) during 1912 and 1913. There were synchronous records for seven months, April through October, 1912.

Charles Chree was charged with the reduction and discussion of both sets of magnetic registrations. Unfortunately in Chree (1923) he only discusses the seven months of simultaneous data (covering one winter and most of the equinoctial months on either side). Using modern data from Scott Base and Dumont d'Urville we can compensate for this and get a figure representative of the whole year. It turns out that you have to multiply the range derived from the seven winter and equinoctial months by 1.17 to get a yearly value. The diurnal range in Y was 50 nT at both Cape Evans and Adelie Land for the seven simultaneous months. This corresponds to a range of 50 * 1.17 = 60 nT for the year. However, this figure represents only a *lower limit* to the true range because days where the trace was

not complete were excluded. These days were usually the most disturbed (that's why the trace is missing, confused or unreadable), so the result suffers from a systematic lack of disturbed days, thus leading to a too low value for the diurnal variations. It might be possible by a careful reanalysis of the original data to get a better estimate of the true range, but already the lower limit we have obtained is larger than what would be expected if the Sun's magnetic field were lower by a factor of two or more in 1912 than today. A full discussion of the data from Cape Evans is given by Chree (1921) and might be worthwhile to consult. Here we summarize the results derived from his 1923 paper:

Average Epoch (year)	Average *Rz*	Station	Corr. Magn. Lat.	Polar Cap Range
1912.6	3.6	Cape Evans	-80°	>60 nT
1912.6	3.6	Adelie Land	-76°	>60

Danmarkshavn, Greenland 1906-07

The Danish expedition 1906-08 to map the northeastern coast of Greenland erected a magnetic station at Danmarkshavn (geographic coordinates 76.77° N, 341.26° E) and magnetographic observations of the variations of the Declination were carried out over a full year from October 1906 to October 1907 by the meteorologist Alfred Wegener (Brückmann, 1914). The Danish Meteorological Institute operates a modern automatic station at the same location, so we have the opportunity to compare observations taken more than 90 years apart.

References

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

Bartels, J., N. H. Heck, and H. F. Johnston, The three-hour-range index measuring geomagnetic activity, *J. Geophys. Res.,* **44**, 411 (1939).

Beer, J., G.M. Raisbeck and F. Yiou, The variations of ^{10}Be and solar activity, in '*The Sun in Time*', eds.. C.P. sonnett *et al*., (University of Tucson), pp 343-359 (1991).

Broun, J. A., On the horizontal force of the earth's magnetism, *Proc. Roy. Soc. Edinburgh*, **22**, 511 (1861).

Brückmann, W., Magnetische Beobactungen der Danmark-Expedition 1906-08, Meddelelser om Grønland, **47**, part II/8, Copenhagen (1914).

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

Ellis, W., On the relation between the diurnal range of magnetic declination and horizontal force at Greenwich, 1841 to 1877, and the period of solar spot frequency, *Phil. Trans. London (A)*, **171**, 541-60 (1880).

Campbell, W. H., Geomagnetic storms, the Dst ring-current myth and lognormal distributions, *J. Atmos. Terr. Phys.*, **58**, no. 10, 1171-1187 (1996).

Chapman, S., An outline of a theory of magnetic storms, *Proc. Roy. Soc. London (A)*, **95**, 61-83 (1918).

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

Chapman, S., *The Earth's Magnetism*, (John Wiley: New York), 127 pp (1951).

Chernosky, E. J., Double sunspot-cycle variation in terrestrial magnetic activity 1884-1963, *J. Geophys. Res.*, **71** (3), 965-974 (1966).

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

Chree, C., British (Terra Nova) Antarctic Expedition 1910-1913. *Terrestrial Magnetism*. 548 pp, London, Harrison & Sons (1921).

Chree, C., Magnetic Phenomena in the Region of the South Magnetic Pole, *Proc. Roy. Soc. London (A)*, **104**, 165-91 (1923).

Clilverd, M.A., T.D.G. Clark, E. Clarke, and H. Rishbeth, Increased magnetic storm activity from 1868 to 1995, *J. Atmos. Sol. Terr. Phys.*, **60**, 1047-1056 (1998).

Clilverd, M.A., T.D.G. Clark, E. Clarke, H. Rishbeth, and T. Ulrich, The causes of long-term changes in the *aa* index, *J. Geophys. Res.*, **107**, submitted (2002).

Cliver, E.W. *et al.* (2003) [1954 and 1996 semiannual varition] ◄

Duhau, S and Ch. Y. Chen, The sudden increase of solar and geomagnetic activity after 1923 as a manifestation of a non-linear solar dynamo, *Geophys Res. Lett.*, **29**, No. 13, 10.1029/2001GL013953, 6 (2002).

Gosling, J.T., D.J. McComas, J.L. Phillips, and S.J. Bame, Geomagnetic activity associated with earth passage of interplanetary shock disturbances and coronal mass ejections, *J. Geophys. Res.*, **96**, 731 (1991).

Gustafsson, G., N.E. Papitashvili, and V.O. Papitashvili, A Revised Corrected Geomagnetic Coordinate System for Epochs 1985 and 1990, *J. Atmos. Terr. Phys.*, **54**, 1609-1631 (1992).

Harrison, G., Twentieth century secular decrease in the atmospheric potential gradient, *Geophys. Res. Lett.,* **29**, No. 14, *5* ,10.1029/2002GL014878 (2002).

Kjærgaard Andreasen, G., Reconstruction of past solar wind variations: Inversion of the geomagnetic response at Godhavn, *J. Geophys. Res.*, **102**, 7025-7036 (1997).

Lockwood M., R. Stamper, and M.N. Wild, A Doubling of the Sun's Coronal Magnetic Field during the Last 100 Years, *Nature* **399**, 3 June 1999. Pages 437-439 (1999).

Makarov, V.I., *Solar Phys*., **150**, 359 (1994).

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

Mayaud, P.N., A Hundred Years Series of Geomagnetic Data 1868-1968, *IAGA Bulletin, No.***33**, 255 pp (1973)

Mayaud, P.N., Derivation, Meaning, and Use of Geomagnetic Indices, AGU (Washington D.C.) (1980).

Menvielle, M., and A. Berthelier, The *K*-derived planetary indices: description and availability, *Rev. Geophys.,* **29**, 415-432 (1991). Correction: Ibid., **30**, 91 (1992).

Moos, N. A. F, *Colaba Magnetic Data, 1846 to 1905*, Bombay, 781 pp (1910). See also Chapman (1918).

King, J. H., *et al.*, OMNI tape, NASA/NSSDC, http://nssdc.gsfc.nasa.gov/omniweb/ow.html. (2002).

Richardson, I.G., H.V. Cane, and E.W. Cliver, Sources of geomagnetic activity during nearly three solar cycles (1972-2000), *J. Geophys. Res.*, **107**, SSH 8-1, 10.1029/2001/JA000504 (2002).

Rosenberg, R. L. and P. J. Coleman, Jr., Heliographic latitude dependence of the dominant polarity of the interplanetary magnetic field, *J. Geophys. Res.*, **74** (24), 5611-5622 (1969).

Russell, C.T. and R.L. McPherron, Semiannual variation of geomagnetic activity, *J. Geophys. Res.*, **78**, 92-108 (1973).

Russell, C.T. and 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).

Russell, C.T. and J.G. Luhmann, Encyclopedia of Planetary Sciences, edited by J. H. Shirley and R. W. Fainbridge, 208-211, Chapman and Hall, New York, (1997).

Sabine, E., On periodical laws discoverable in the mean effects of the larger magnetic disturbances - No III, *Phil. Trans. Roy. Soc. London*, **146**, 357 (1856).

Sargent, H.H. III, The 27-day recurrence index, in "Solar Wind-Magnetosphere Coupling" edited by Y.Kamide and J.A. Slavin, pp143-148, Terra Scientific, Tokyo (1986).

Stozhkov, Y.I, N.S. Svirzhevsky, V.S. Makhmutov, and A.K. Svirzhevskaya, Long-term cosmic ray observations in the atmosphere, Proceedings of ICRC 2001, 3883 (2001).

Svalgaard, L., E.W. Cliver, and A. Ling, The semiannual variation of great geomagnetic storms, *Geophys. Res. Lett.*, **29**, X-1 (2002).

Troshichev, O.A., R. Yu. Lukianova, V.O. Papitashvili, F.J. Rich, and O. Rasmussen, Polar cap index (PC) as a proxy for ionospheric electric field in the near-pole region, *Geophys. Res. Lett.*, **27**, 23, pp. 3809-3812 (2000).

Tsunomura S., Characteristics of geomagnetic sudden commencement observed in middle and low latitudes, *Earth Planets Space*, **50**, 755–772 (1998).

Wolf, R., Astronomische Mitteilungen, *Vierteljahrsschrift d. Naturforsh. Ges. Zürich*, **9**, 221 (1859).

Yiou, F., G.M. Raisbeck, S. Baumgartner, J. Beer, C. Hammer, S.J. Johnsen, J. Jouzel, P.W. Kubik, J. Lestringuez, M. Stievenard, M. Suter, and P. Yiou, Beryllium 10 in Greenland Ice Core Project ice core at Summit, Greenland, *J. Geophys. Res.*, **102**, 26783-26794 (1997).