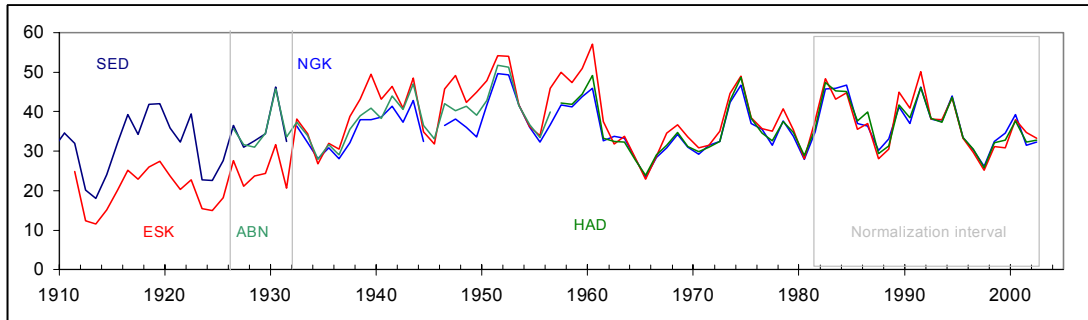
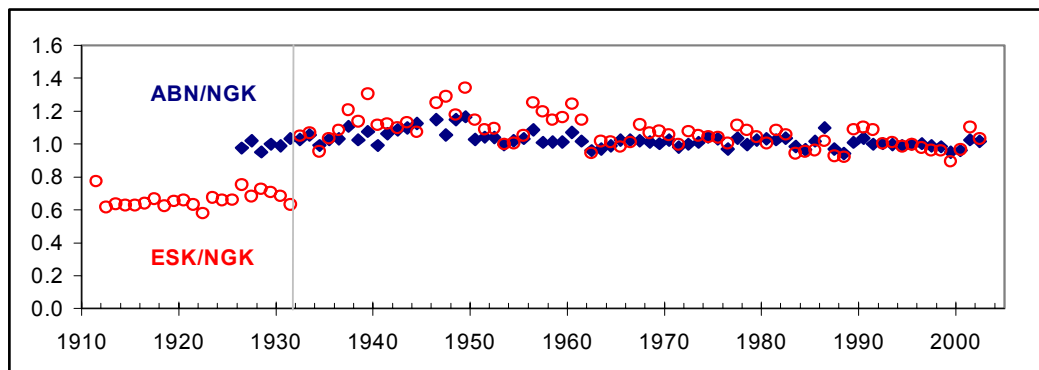


Comment on Clilverd *et al.* “Reconstructing the long-term aa index”
***Journal of Geophysical Research*, vol 110(A7), CiteID A07205**
Doi:10.1029/2004JA10762.

Clilverd *et al.* compute IHV (H) from the British stations ESK and LER under the assumption that these long-running stations have changed so little over time that long-term trends are meaningful. ESK was recording X and Y before 1932.00, and H and D thereafter. LER has been recording H and D since 1926. The change of instruments and elements observed at ESK in 1932 unfortunately resulted in a more than 50% difference in IHV (H), with IHV being too low before 1932. To show this, we need to compare IHV from ESK with IHV calculated using other stations. The authors also used NGK since 1890. It may have escaped their attention (although Linthe is from NGK) that NGK only began observing in 1932 and that data before 1932 is really from SED (Seddin) back to 1908 and POT (Potsdam) before 1908. Under the assumption that SED and NGK form a homogeneous series, we can compare ESK with SED before 1932 and with NGK thereafter. In addition, the British stations ABN (Abinger) is available from 1926-1956 and HAD (Hartland) thereafter. We now form two long-running series: SED-NGK and ABN-HAD and first compare these to check if they agree. As activity is slightly different for the two series (different latitude and Earth conductivity) we normalize IHV (H) for both series to have the same mean over the interval 1982-2002. We also normalize ESK the same way. The result is shown in Figure 1 below:



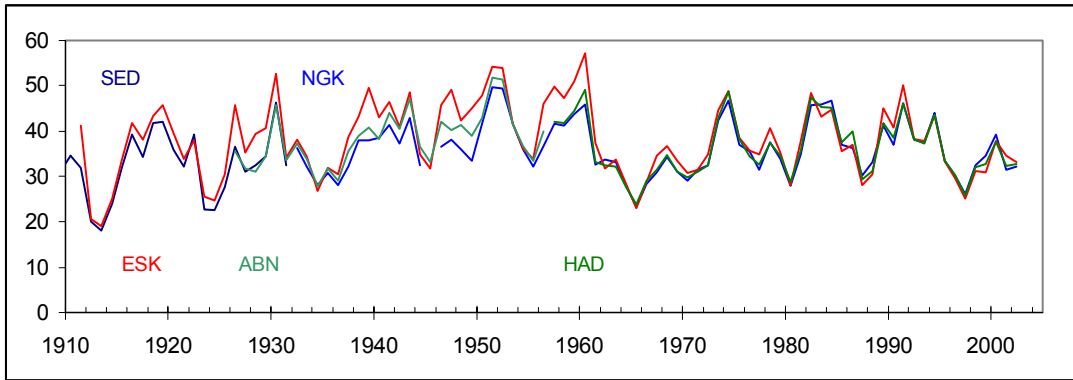
It is evident that ABN and SED agree 1926-1931, that ABN and NGK agree 1932-1956 and that HAD and NGK agree 1957-2002, with only a slight problem 1946-1951 (unknown why). It is also evident that ESK agrees well with both ABN and NGK for the years 1932-1938 and with HAD after 1960. There are some disagreement at sunspot maxima 1939-1960 (again not known why), but the good agreement 1932-1938 and the clear disagreement 1910-1931 allow us to establish the fact IHV from ESK was much too low compared to ABN and SED before 1932. This discontinuity is perhaps illustrated better in Figure 2 below, where we plot the ratio ABN-HAD/SED-NGK and ESK/SED-NGK:



We discount the possibility that SED, NGK, and ABN *conspire* to make it *look* that it is ESK that jumps (this is based on comparisons with several other stations, including LER). We have to accept that ESK IHV

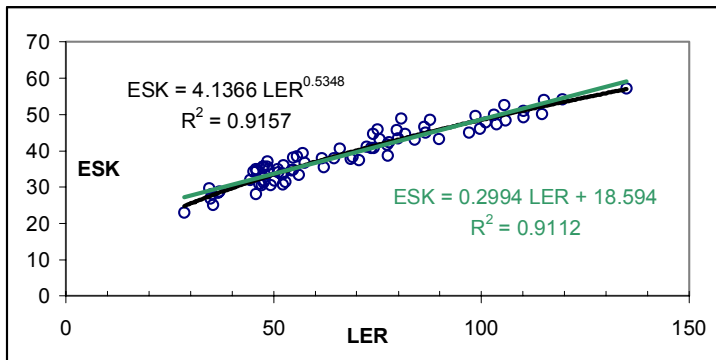
is too small by a factor 1.66 before 1932. The factor 1.66 is the ratio of the average ratios for 20 years on either side of the discontinuity.

We can now correct the ESK values before 1932 by multiplying by 1.66. The result is shown in Figure 3:



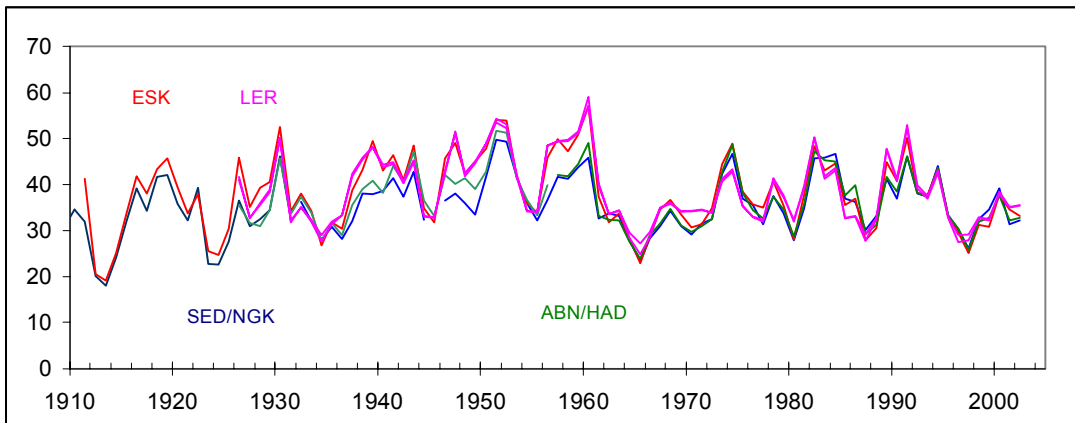
As before, the ESK values are higher near sunspot maxima (this was also noted by Clilverd *et al.*) before about 1960. At this point we do not know the reason for this. Alternatively, the lower latitude stations ABN, SED, and NGK may be too low, but at least the behavior of ESK is consistent over the interval 1910-1960.

It would be of interest to include LER in the analysis. The LER IHV values are much larger than ESK IHV values due to the higher latitude of LER. To be able to compare the LER and ESK series we first plot ESK against LER since 1932 in Figure 4 below:

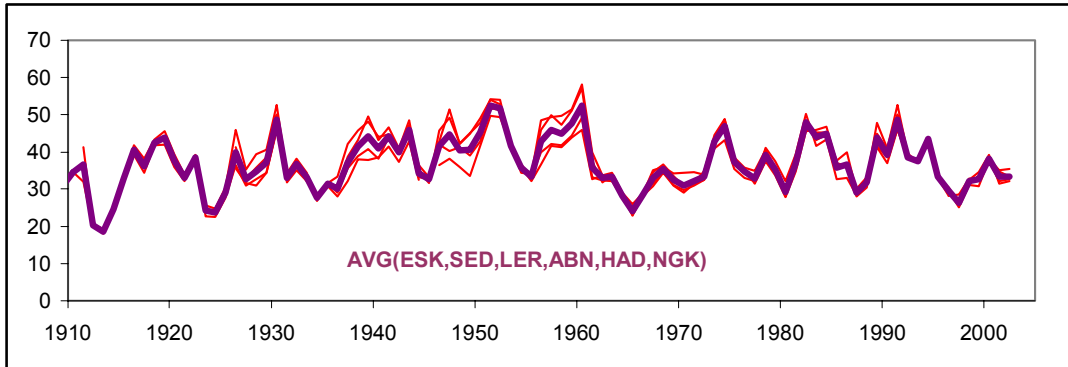


The relationship is not quite linear and there is a large offset. We fit both a power law and a linear dependence as shown. With these fits we can reduce LER to the level of ESK (where we are using the normalized value for ESK, but not for LER).

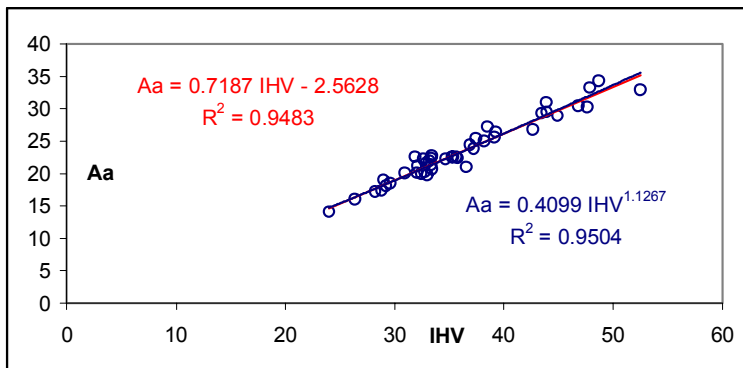
We can now add LER (reduced to ESK) to the plot comparing all stations (Figure 5 below):



Note that LER shows the same behavior as ESK regarding sunspot maxima before 1960. Also note that the corrected ESK agrees very well with LER. It is possible that the “sunspot maxima” effect is latitude-related, but it could also be due to a difference in data reduction. Without further study, we cannot get much closer. If we pass no judgement and simply take the average of all stations we get Figure 6 below:

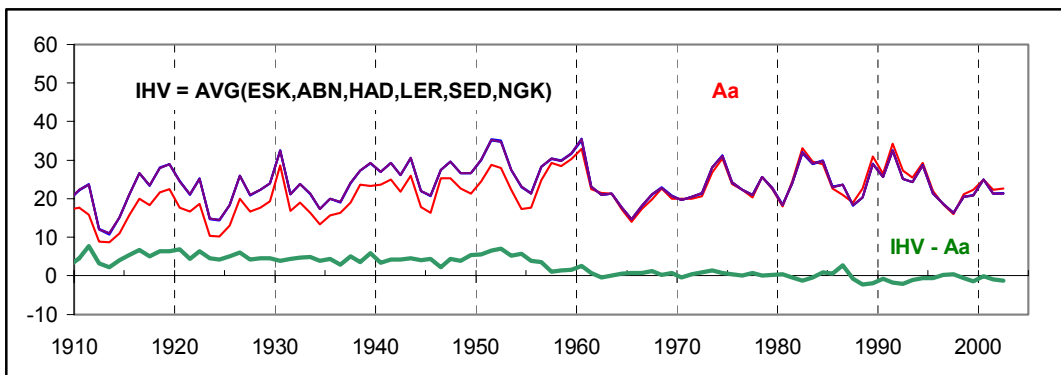


To compare the resulting average IHV with Aa it is convenient to reduce both to a common mean over some interval. We use the interval 1959-2002 as Aa is well correlated with the much better index Am over that interval. For the reduction we first plot Aa versus IHV for 1959-2002 (Figure 7 below):



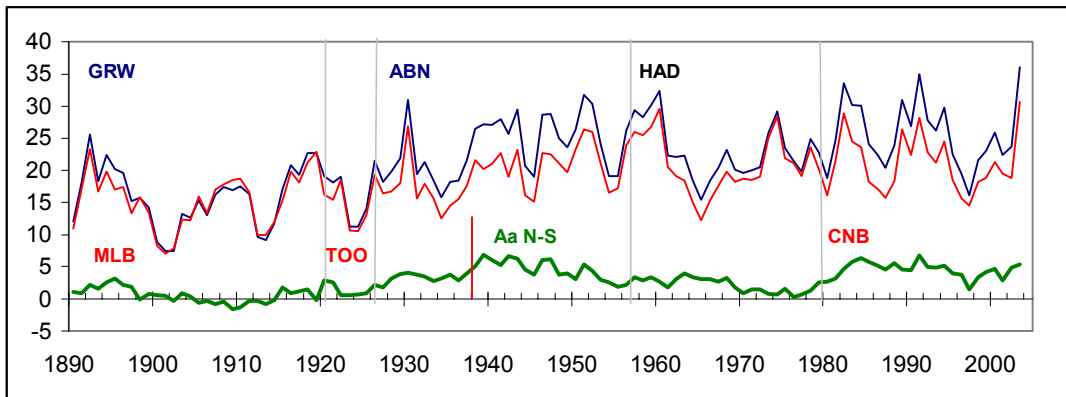
As in Figure 5, we fit both a linear relationship and a power law. There is not a significant difference between the two fits. In reducing IHV to Aa-units we calculate IHV reduced to Aa from the observed average IHV over all stations using both regression formulae and take the average.

The result compared to Aa is shown in Figure 8 below:

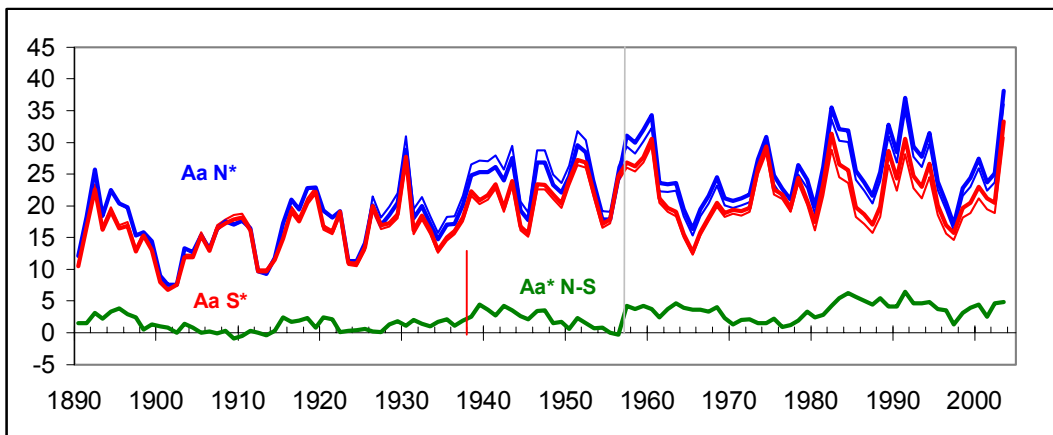


Also plotted is the difference IHV-Aa. It is clear that this difference is close to zero (average = 0.04) after 1957, but significantly larger than zero (average = 4.68) before that, as claimed by Svalgaard et al.

There are four long-running K-series: Aa North, Aa South, NGK, and SOD. We convert the K-indices to equivalent amplitudes using the standard table. We then plot these four series. First we discuss the Aa series. Figure 9 below shows the Aa series and their difference:

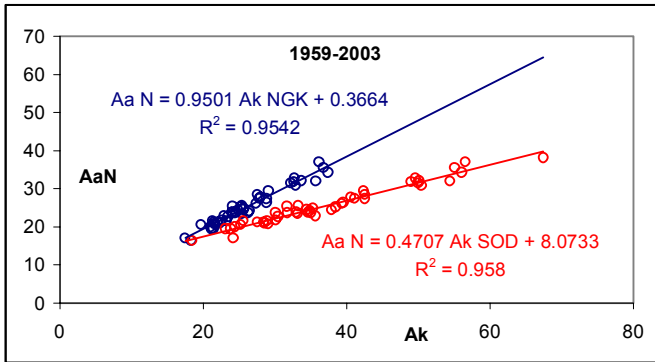


Vertical lines show the times of change of observatories. Note that Table 2 in Clilverd et al. has the years wrong. A short red vertical line shows the time of change of instrument at ABN. The Aa values have *not* yet been corrected using the official weight factors. Note that the introduction of HAD in 1957 did not change the difference between North and South. This is an indication that HAD shows the same activity as ABN; something we already saw in Figures 1 through 3. Let's see what happens when we introduce the official weights. The result is shown below in Figure 10:

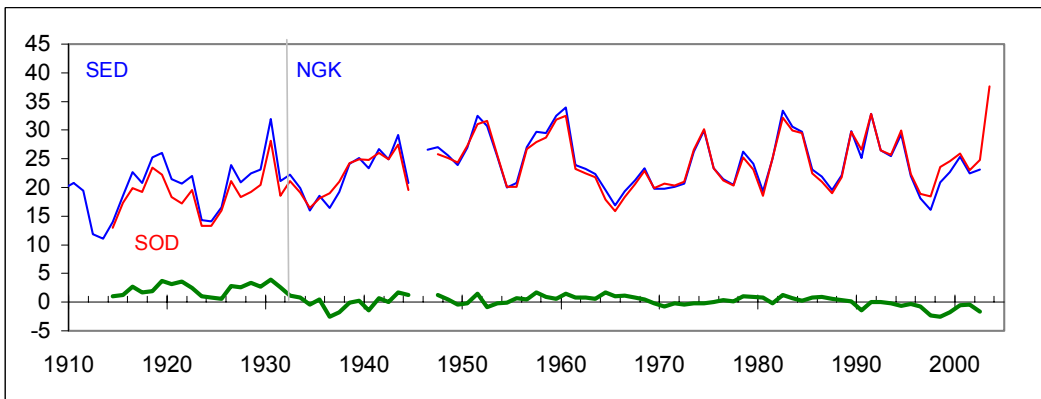


The weighted values are connected with heavy curves, the original values with light curves. Note that the official weights introduce a jump in 1957, presumably partly responsible for the change in the difference between IHV and Aa at the same time. In general, the official (weighted) values of Aa N-S after 1957 are higher than before that time, the difference progressing from near zero around 1910 to near 5 nT at the present time. Unless there is some reason that the Northern Hemisphere should have become more active than the Southern, we may see here hints of calibration problems with Aa that IHV was also indicating (see Figure 8).

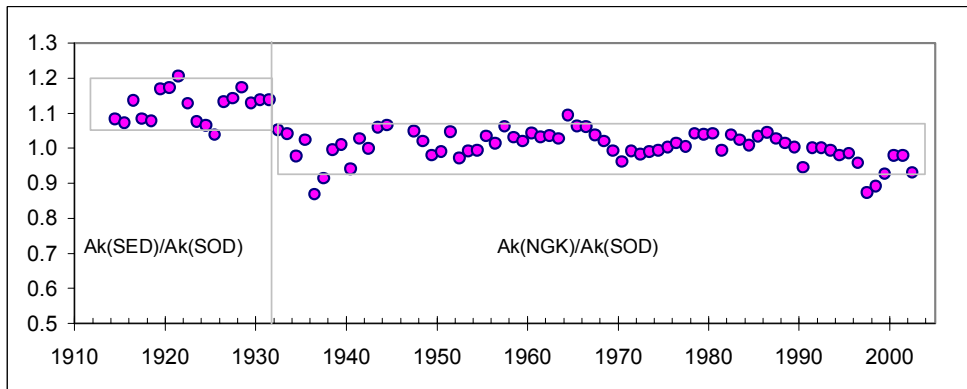
We now wish to compare Ak NGK and Ak SOD. As usual, it is convenient to reduce the data to a common mean over an interval where the calibration is good. We again use 1959-2003 and compare NGK and SOD to AaN for that interval (this will reduce NGK and SOD to the level of AaN). The result is shown in Figure 11 below:



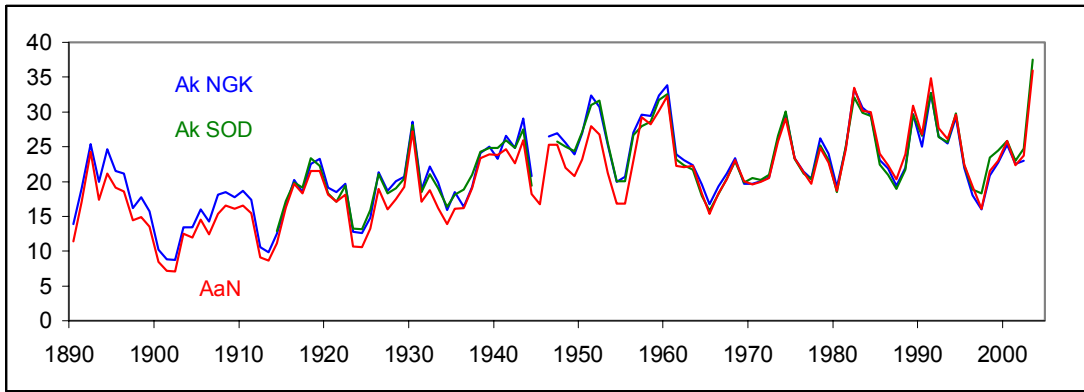
Using the regression formulae we now reduce Ak NGK to equivalent Aa for the whole period 1890-2002, and Ak SOD for the period 1914-2003. If the data can be relied upon to show the long-term trend, the two series should substantially agree. Figure 12 below shows the result (with the difference (SED,NGK - SOD) shown in green):



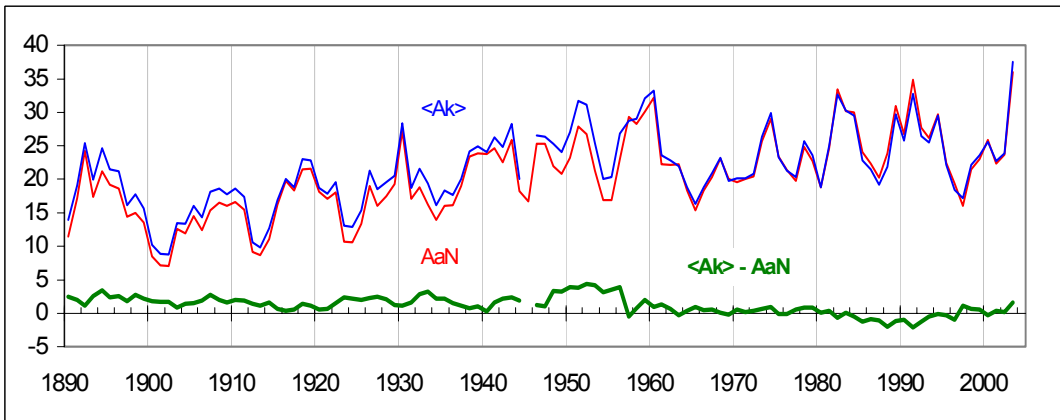
We note that NGK and SOD agree after 1932, but that SED and SOD do not agree before 1932. The assumption that both SOD and SED-NGK show the same trend is thus not justified. As with ESK there is a discontinuity in 1932. Forming ratios as in Figure 2, we see this more clearly in Figure 13 below:



The ratio is 1.119 before 1932 and 1.003 thereafter. Either SED is too high by a factor of $1.119/1.003 = 1.116$ or SOD is too low by the same factor. At this point it is hard to know which is which. As there was a station change between SED and NGK in 1932 while there was no change to SOD might indicate that it is SED that is too high. If we assume that this is the case and correct Ak for SED (but not for POT) we get the result in Figure 14 below. We also show the weighted AaN:



The two Ak series (SOD and POT-SED-NGK) now agree quite well. They still disagree with AaN before 1957, but this we would expect based on Figure 8. To show this, we form the average of the Aks and compute the difference from AaN as shown in Figure 15 below:



Again we see the jump in 1957 consistent with the IHV-result (Figure 8).

Conclusion 1: The IHV-series from ESK and LER do not have the same calibration at all times. There is a jump in 1932.

Conclusion 2: The K-indices from SOD and POT-SED-NGK do not have the same calibration at all times. There is a jump in 1932.

These two conclusions invalidate the analysis of Clilverd *et al.* that assumes that there are no such jumps. Calibrating on the basis of the *earliest* solar cycle is not a good idea in view of our conclusions. Why not use the *latest* solar cycle and base the calibration on modern data where we know that aa is good (validated by am)?