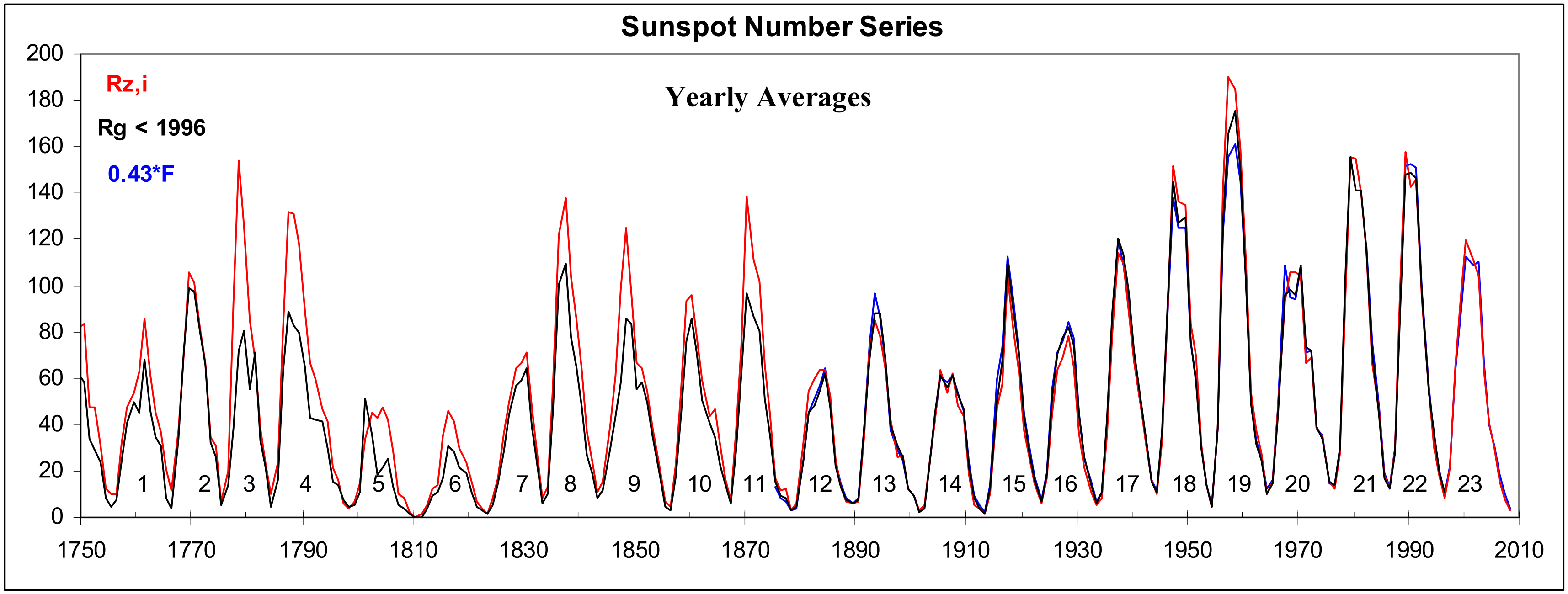
Recalibration of the Sunspot Number and Consequences for Predictions of Future Activity and Reconstructions of Past Solar Behavior

Leif Svalgaard
Stanford University

Solar activity during the Onset of Solar Cycle 24
Dec. 8, 2008, Napa, California

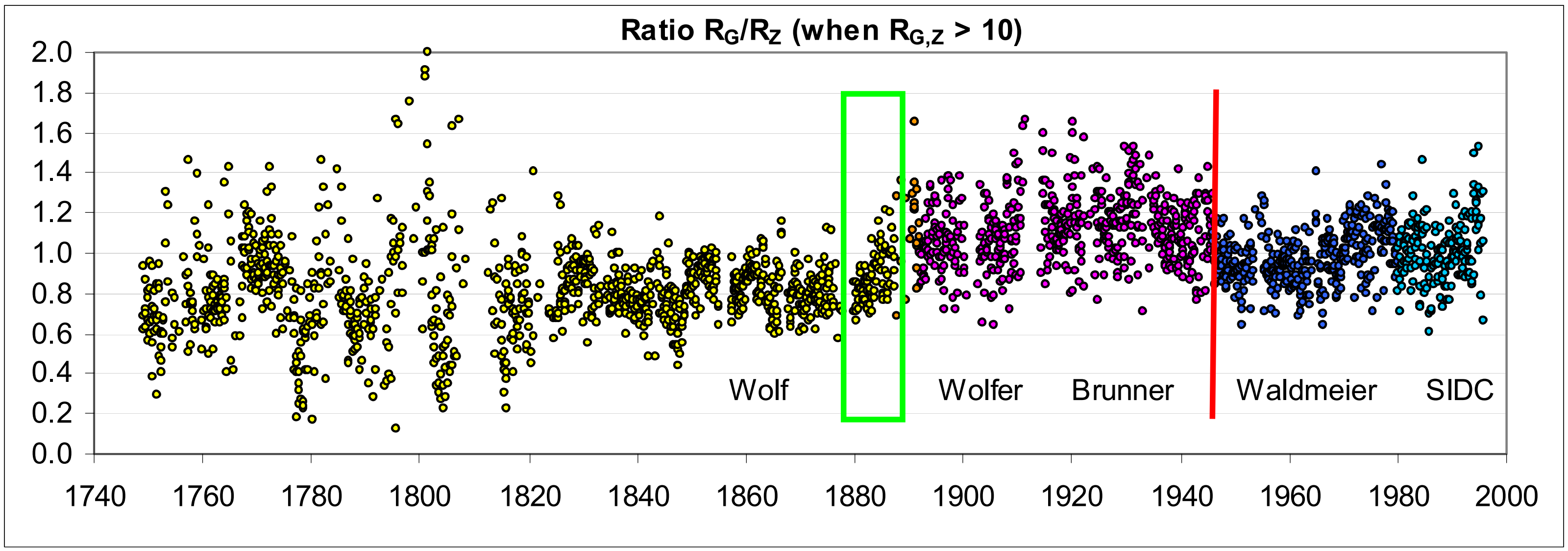
We have TWO really long-term Sunspot Number [SSN]Series:



Hoyt & Schatten's Group SSN since 1875 is largely derived from Greenwich Sunspot Observations [Area measurements] and fits the Zurich SSN reasonably well during time of overlap 1875-1995, but is consistently lower before that. Since 1996, The GSN, Rg, can be derived as 0.43 times the size [in kb] of the yearly sunspot region files maintained by David Hathaway [the files have a fixed length 80 byte daily record for each region!].

The difference between Rg and Rz before ~1875 is problematic for prediction and assessment of long-term solar activity. Was the past century really *that* much more active than previous centuries?

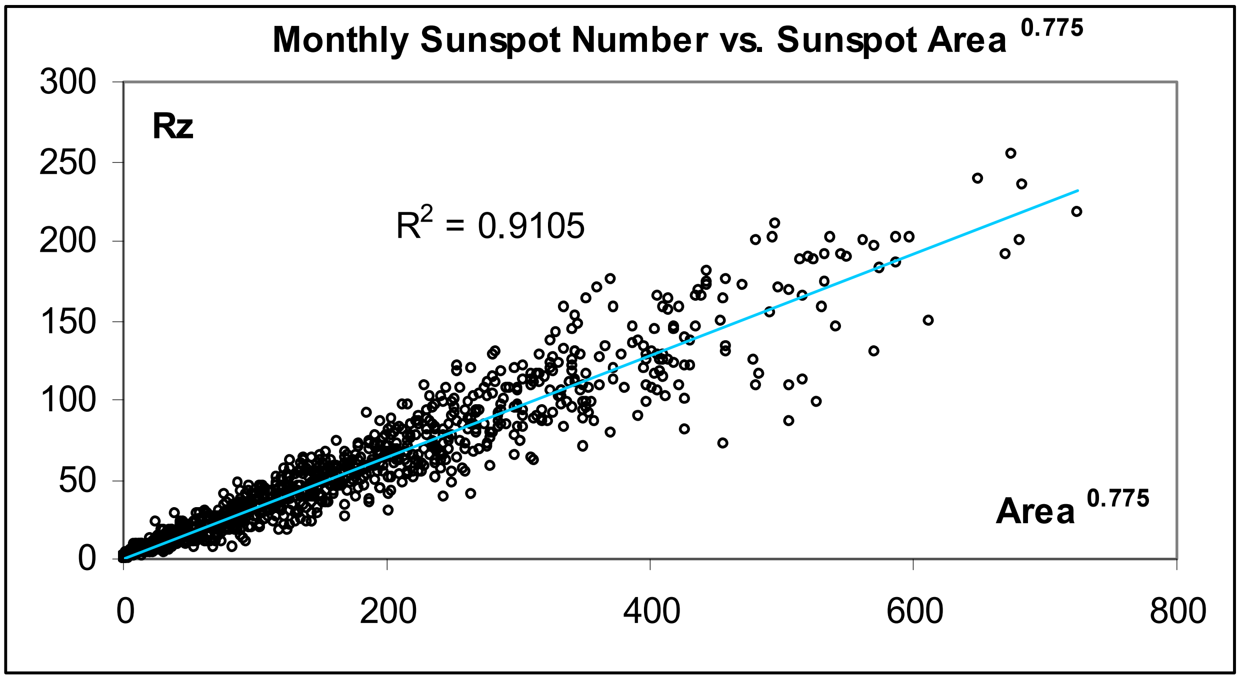
The ratio between monthly means of the Rg and Rz [except when the SSN is less than 11] shows clear jumps in ~1945 when Max Waldmeier took over and a 'creep' ~1880-93, when Wolf's assistant Alfred Wolfer was performing more and more of the observations:



Wolfer disagreed with Wolf about how to count small spots [Wolf didn't count them as their visibility depended too much on the seeing and even advocated counting very large spots twice]. The 'conversion factor', 0.6, between Wolfer and Wolf may not completely remove the discontinuity caused by changes of observing procedure and observer, although we'll argue later that the fault is mainly with Rg.

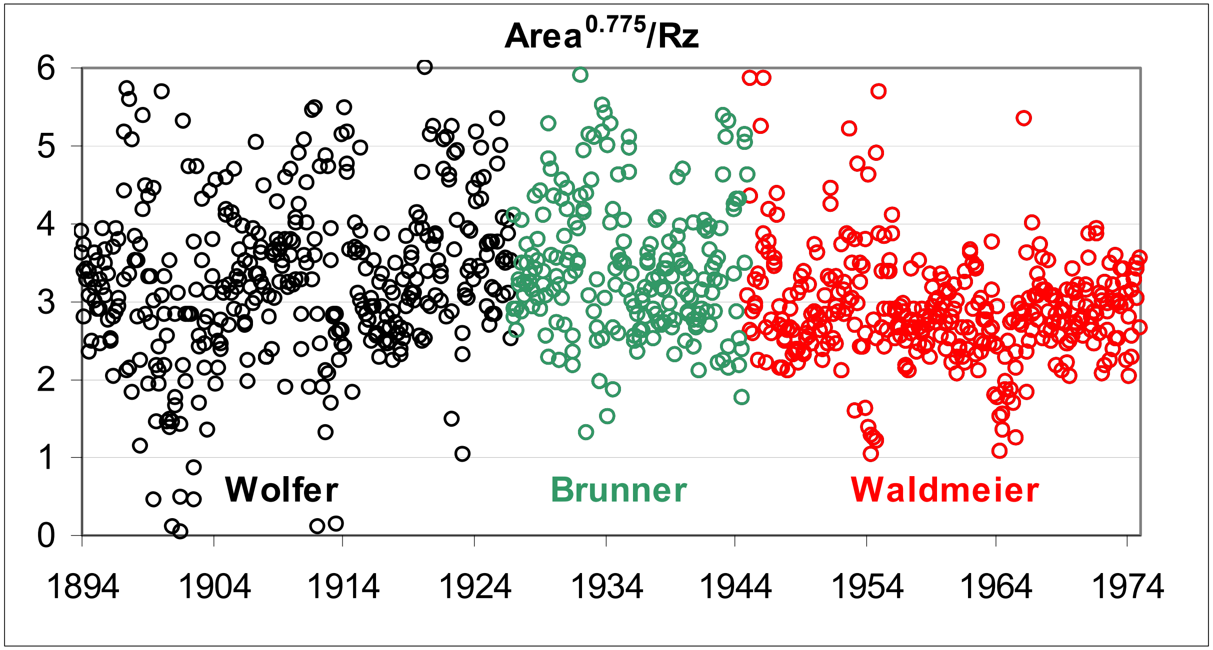
But we can directly compare Rz to the Greenwich Sunspot Area measurements, SA [1894 (to avoid the ‘creep’)-1975]. The relationship between the two quantities is slightly non-linear:

$$\mathbf{Rz = (1/\mathit{r}) \cdot SA^{0.775}}$$



Since there is no zero-point offset in the fit, we can meaningfully plot the ratio $r = SA^{0775} / Rz$:

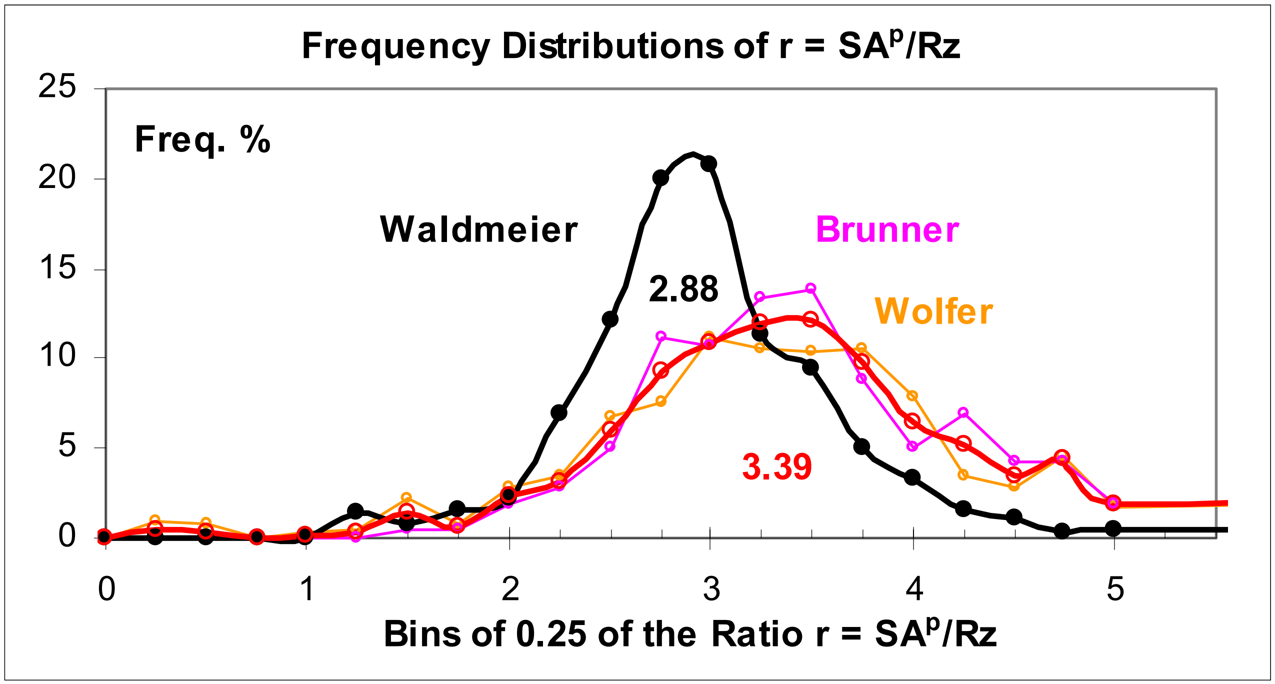
We see the same discontinuity in 1945 [apparently Brunner did not change the procedure – a true civil servant…]:



We can quantify the jump:

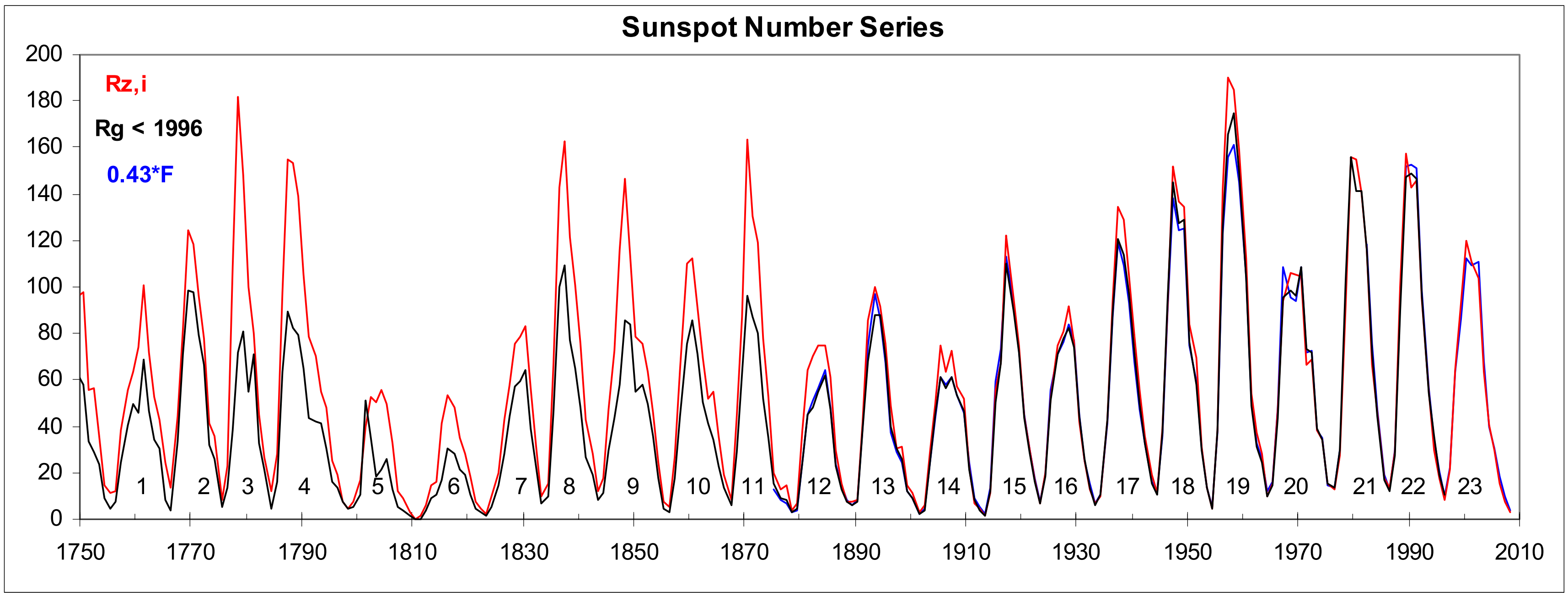
Observer	<$SA^{0.775}$>	<Rz>	<r>	1/<r>	Corr. Factor
Wolfer	123.71	36.50	3.3898	0.2950	1.175
Brunner	168.24	49.69	3.3859	0.2953	1.174
Waldmeier	223.45	77.46	2.8847	0.3467	1.000

Or show it as a distribution of values of the ratio, *r*:



The ‘correction factor’ is defined as that number by which to multiply Rz for an observer to match that of Rz by Waldmeier (*i.e.* corr. $factor_{observer} = \langle r \rangle_{observer} / \langle r \rangle_{Waldmeier}$). In other words, Rz for Wolfer-Brunner is too small by 17.5% compared to Rz observed by Waldmeier.

If we accept the fidelity of the RGO sunspot observations [at least for a few decades around 1945] we must ascribe the artificial increase of Rz after 1945 to Waldmeier's inexperience [Friedli, 2005] as he struggled with learning how to construct the sunspot number. Subsequent observers have strived to match Waldmeier, so in order to remove the 1945 discontinuity [and be consistent with modern counts] we must increase the pre-1945 Rz by 17.5%:



This, of course, just makes the discrepancy with the Group Spot Number worse [<Rz> ~ 1.4 <Rg> before 1875].

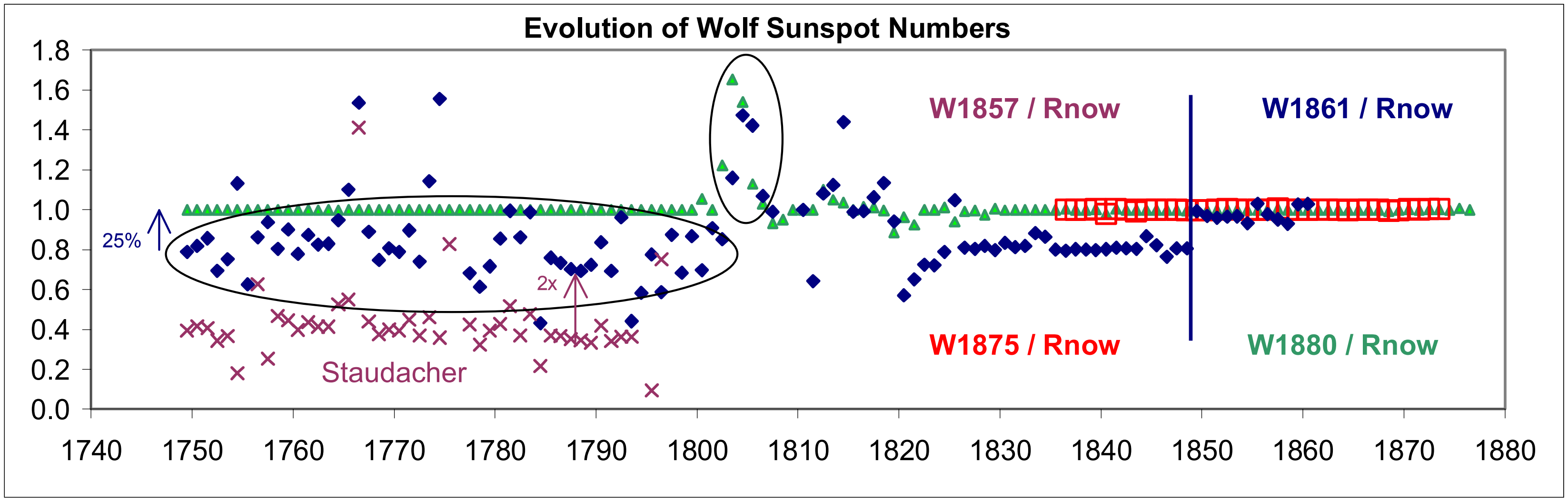
Friedli [2005] writes:

> "Zum Einen war Waldmeier vor seiner Ernennung zum Direktor der Eidgenössischen Sternwarte jahrelang der Leiter der Aussenstation in Arosa und hatte kaum Beobachtungserfahrung am Wolfschen Normalrefraktor in Zürich, zum Anderen hat Waldmeiers Vorgänger William Brunner nach seiner Emeritierung nicht mehr weiterbeobachtet und auch dessen langjähriger Assistent hat die Eidgenössische Sternwarte schon ein Jahr nach Waldmeiers Amtsantritt verlassen. Das neue Beobachtungsteam in Zürich war also relativ unerfahren und musste zudem noch während der Minimumsphase beginnen. Erschwerend kam hinzu, dass die beiden nachfolgenden Zyklen die intensivsten je direct beobachteten waren, mit bis zu 100 Einzelgruppen pro Sonnenrotation in den Maximumsphasen. Waldmeier hat denn auch selber befürchtet, sein als konstant angenommener Skalenfaktor könnte variieren."

"The new observer-team in Zurich was thus relatively inexperienced" and "Waldmeier himself feared that his scale factor could vary". We now know that his fear was not unfounded.

BTW, Friedli still uses the original Zurich Refractor used by Wolf and he and his friend, Keller [who was Waldmeier's assistant] have continued the observations and production of their own 'private' Zurich sunspot number up to the present day. [I'm trying to lay my hand on that data. No joy yet…]

So where did Wolf get his sunspot numbers from? He actually published several versions of his 'Wolf Numbers'. Here we compare these versions with the modern 'official' list, by forming the ratio between the earlier values and the 'official' values [staying away from Rz = 0]:



Note that in the 1875 version, almost all values [as given in the 1861 list] before 1849 were increased by 25%] and that the pre-1800 values [based on Staudacher's drawings] were doubled between the 1857 and 1861 lists [before being further increased by the 25%.]. Note also that the Dalton Minimum was considerably less deep in the earlier versions [Wolf originally had no data and was just guessing…].

Here we compare the 1861 list with the official list:

Abstract of his latest Results. By Prof. Wolf.

(*Translation communicated by Mr. Carrington.*)

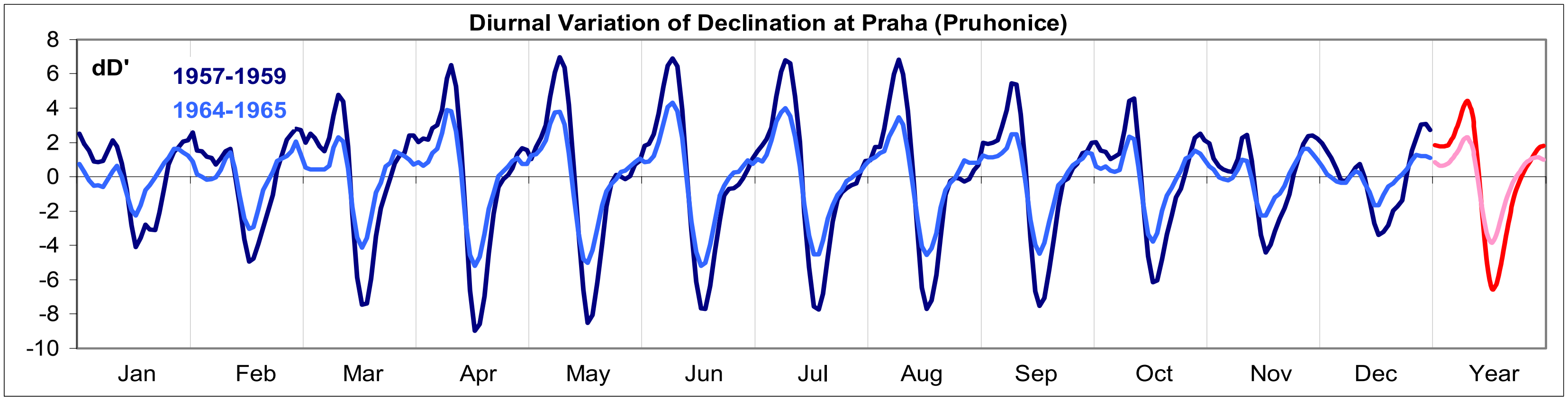
Some fine series of observations of Flaugergues, Adams, Arago, and others, have enabled me to fill in previous breaks, and to express in the same unit my Relative numbers (for the abundance of Solar Spots in successive years) for the years from 1749 to 1860. They are as follows:—

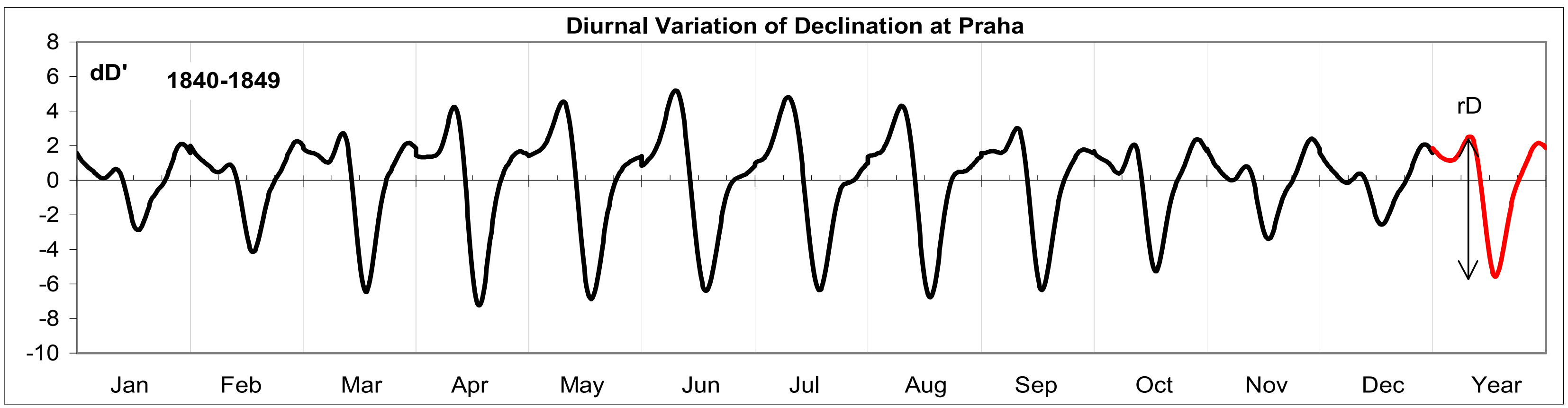
1749	63·8	1777	63·0	1805	50·0?	1833	7·5 m
1750	68·2 M	78	94·8	06	30·0?	34	11·4
51	40·9	1779	99·2 M	07	10·0?	35	45·5
52	33·2	1780	72·6	08	2·2	36	96·7
53	23·1	81	67·7	1809	0·8	37	111·0 M
54	13·8	82	33·2	1810	0·0 m	38	82·6
55	6·0 m	83	22·5	11	0·9	1839	68·5
56	8·8	84	4·4 m	12	5·4	1840	51·8

1749	80.9	1777	92.5	1805	42.2	1833	8.5 m
1750	83.4 M	78	154.4	06	28.1	34	13.2
51	47.7	1779	125.9 M	07	10.1	35	56.9
52	47.8	1780	84.8	08	8.1	36	121.5
53	30.7	81	68.1	1809	2.5	37	138.3 M
54	12.2	82	38.5	1810	0.0 m	38	103.2
55	9.6 m	83	22.8	11	1.4	1839	85.7
56	10.2	84	10.2 m	12	5.0	1840	64.6

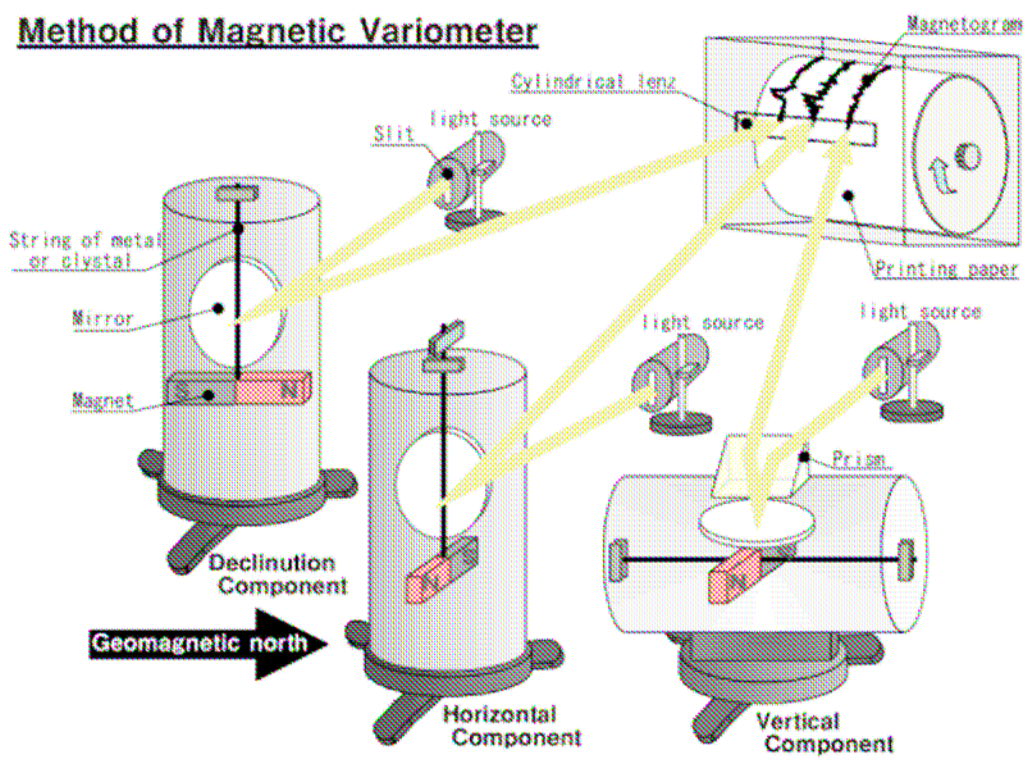
From MNRAS, 1861 and from the current dataset at SIDC in Brussels

How did Wolf justify all these adjustments? And are they still valid today? He [and others] had noted as early as the 1850s that the direction of the Compass Needle [today called the Declination] varied systematically during the day:





This variation [~10 arcmin] was easily measured in the 1840s. And clearly depends on solar activity.

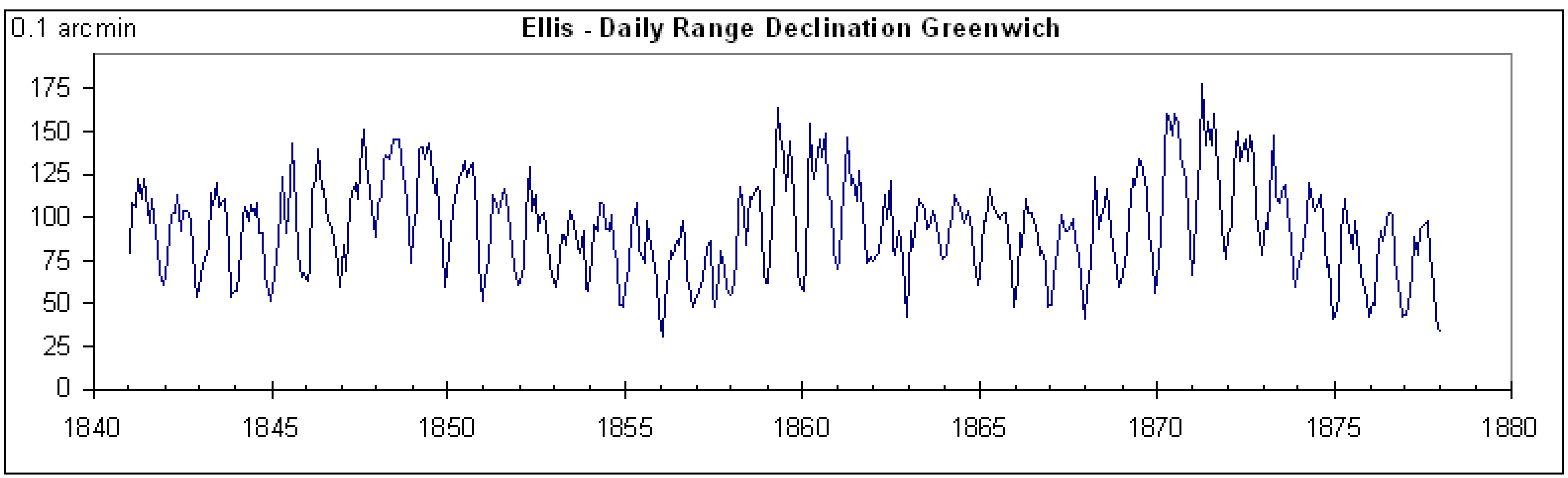


The classical instruments were simple and were in use for more than 150 years [now we do it in other ways – fluxgates, proton magnetometers, digital recording – but the instruments of old were reliable and easily calibrated]. Even without photographic paper, a small telescope could be focused on the suspended magnets and their movements recorded [‘eye readings’].

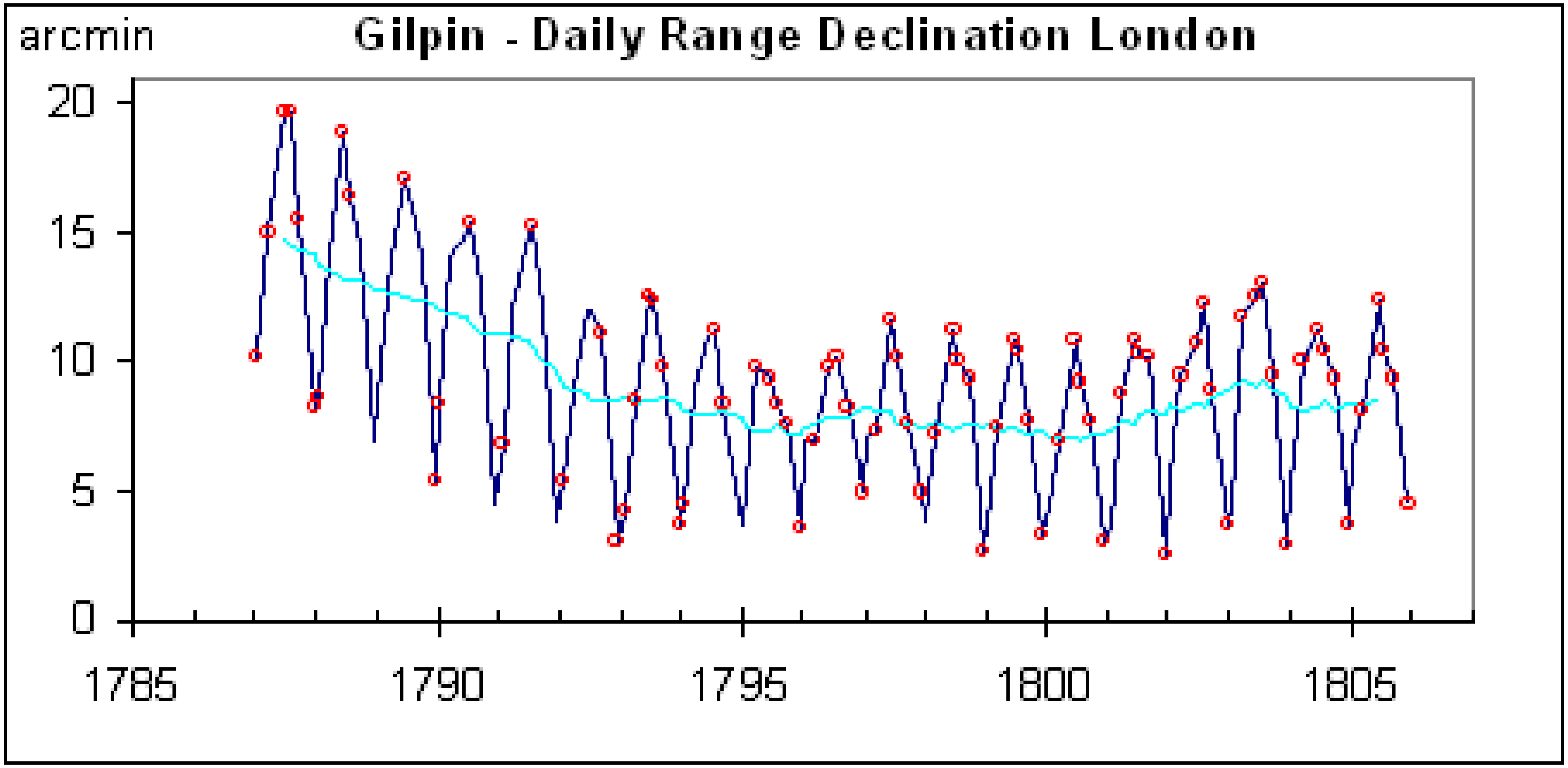
Wolf posited the following linear relationship:

$$\boldsymbol{rD = a + b\, R_W}$$

where $\boldsymbol{rD}$ is the range of the geomagnetic Declination from its extremum in the morning to its extremum in the afternoon and $\boldsymbol{R}_W$ is Wolf's newly defined Sunspot 'Relative' Number. Wolf labored the rest of his life to determine the 'constants' $\boldsymbol{a}$ and $\boldsymbol{b}$, several times lamenting that "by now the last of the doubting Thomases would have to give in and accept my results". Unfortunately they never really did and the 'constants' varied from station to station and with time and with season, and Wolf's relationship was eventually abandoned and forgotten. But as you can see, Wolf was clearly on to something:



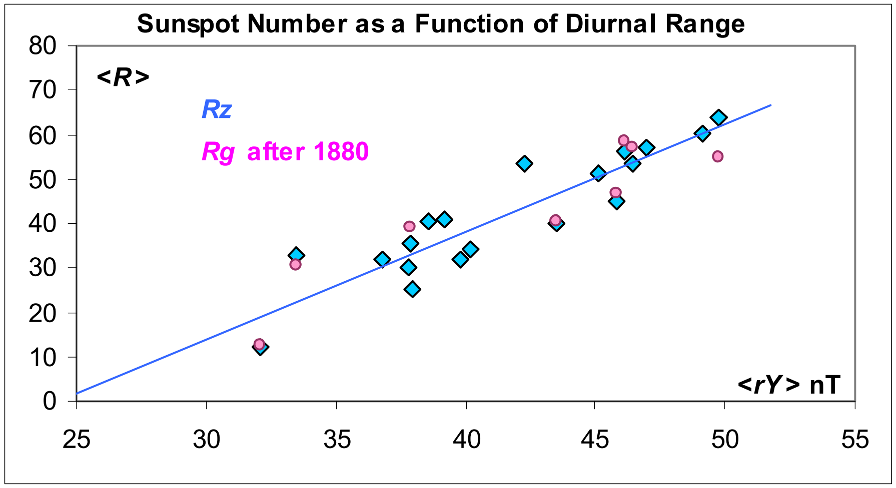
Even going further back in time:

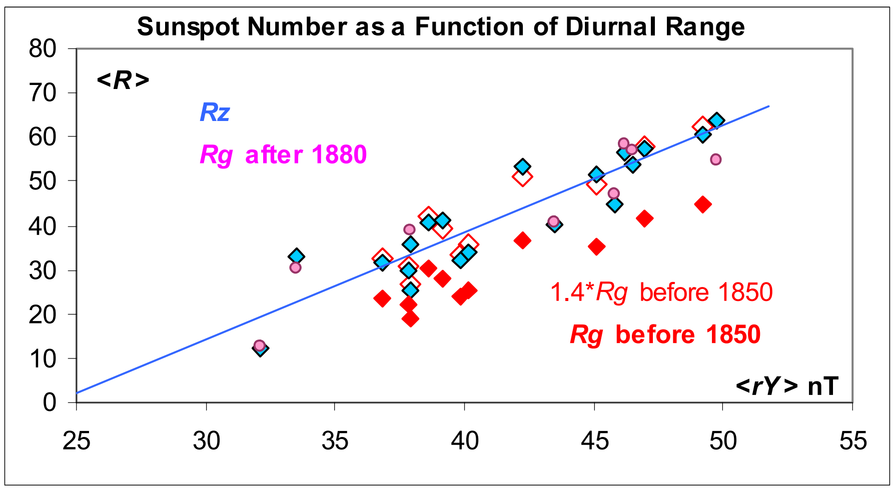


Note that there does not seem to be much sign of the ‘lost sunspot cycle’ between cycle 4 and cycle 5, supposedly peaking in 1793. It was the size of the variation shown here [and a similar one for Paris by Cassini] that caused Wolf to double the Staudacher-based SSNs between 1749 and 1795.

Adolf Schmidt [~1900] had analyzed the data for two intervals on either side of the start [1875] of the RGO sunspot observations for several stations and determined the diurnal ranges:

obs	name	lat	long	interval
WDC	Washington D.C.	38.9	283.0	1840-1842
DUB	Dublin	53.4	353.7	1840-1843
MNH	Munchen	48.2	11.6	1841-1842
PGC	Philadelphia	40.0	284.8	1840-1845
SPE	St. Peterburg	60.0	30.3	1841-1845
GRW	Greenwich	51.5	0.0	1841-1847
PRA	Praha	50.1	14.4	1840-1849
HBT	Hobarton	-42.9	147.5	1841-1848
MAK	Makerstoun	55.6	357.5	1843-1846
KRE	Kremsmunster	48.1	14.1	1839-1850
TOR	Toronto	43.7	280.6	1842-1848
WLH	*Wilhelmshaven*	53.7	7.8	*1883-1883*
GRW	*Greenwich*	51.5	0.0	*1883-1889*
WDC	*Washington D.C.*	38.9	283.0	*1891-1891*
PSM	*Parc Saint-Maur*	48.8	0.2	*1883-1899*
POT	*Potsdam*	52.4	13.1	*1890-1899*
COP	*Kobenhavn*	55.7	12.6	*1892-1898*
UTR	*Utrecht*	52.1	5.1	*1893-1898*
IRT	*Irkutsk*	52.3	104.3	*1899-1899*





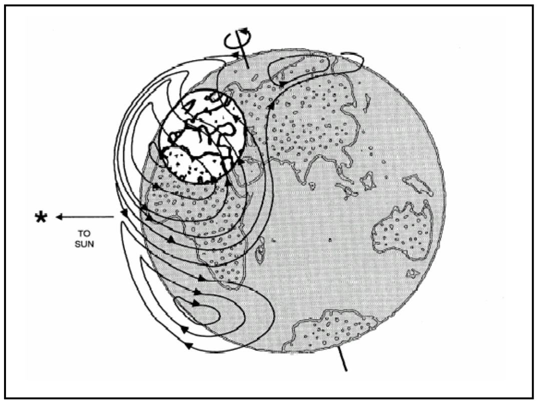
For each station we now compute the average <Rz>, <Rg>, and diurnal range [converted to force units, nT, from arc minutes] and plot <Rz> against the range <rY> [blue diamonds, left]. You can see Wolf's linear relationship in action. For the eight stations with data after 1880, the <Rg>s are also plotted [pink dots] and they match the <Rz> points reasonably well.

This is, however, not the case for the eleven stations from 1850 and before. Their <Rg> [red diamonds] lie well below the fitted line. To make them fit it suffices to multiply their values by 1.4 [giving red open diamonds].

It might not have escaped your attention that the factor 1.4 is just what would be required to remove the difference between Rz and Rg. So, using the Earth itself as an instrument resolves the issue.

How Does This Work?

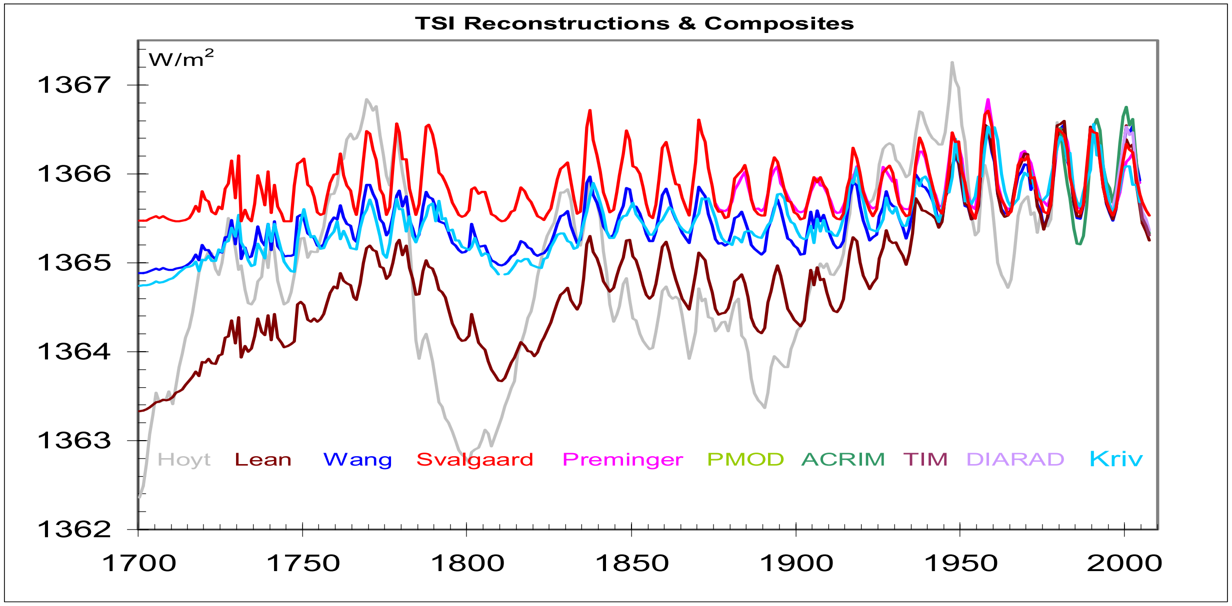
Solar activity is the source of EUV and FUV causing and maintaining the ionosphere. Ionospheric winds and tides move a conducting plasma across the Earth's magnetic field inducing a current, the magnetic effect of which we see at the surface, so the magnitude of this effect is a measure of the conductance and hence of EUV/FUV and thus of their proxy, the sunspot number. Wolf's 'linear relationship' is thus physically sound and the vast storehouse of geomagnetic 19th century data [as yet not digitized] can with our modern understanding serve as the medium for a solid calibration of the Sunspot Number, so we can get a list we can all agree on, and use as a common base for work on the temporal evolution of solar activity. Such work is ongoing.

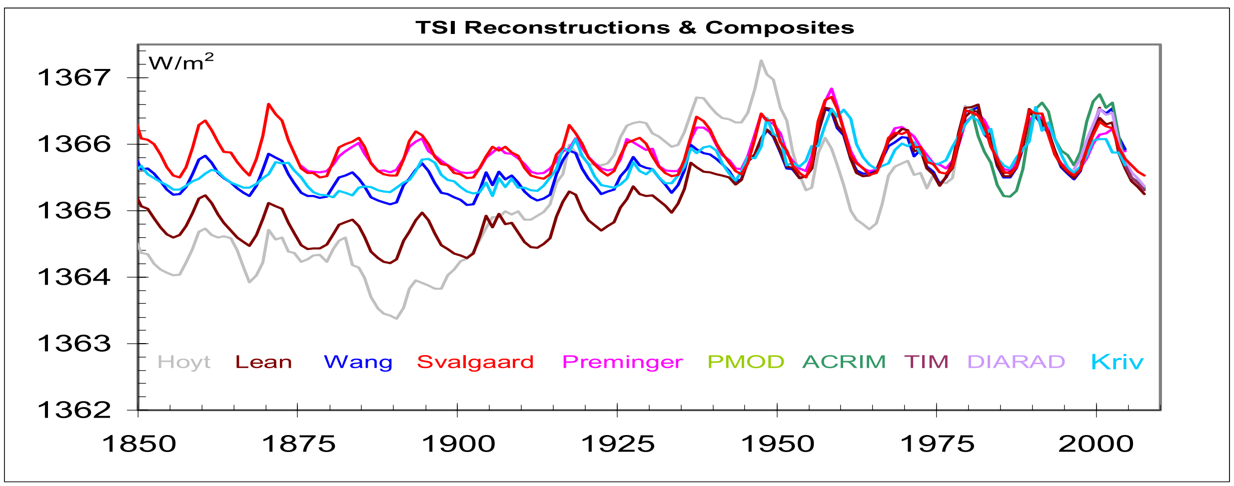


Consequences

Since past behavior is the key to successful prediction, the possibility of a revision of the sunspot series has an obvious impact on the tuning and calibration of models and correlative works. The sunspot number is input to a great many 'models' and methods used to predict solar activity, climate, and even the stock market [as is sometimes claimed]

As an example, we offer what TSI might look like with a revised Sunspot Series [matches the reconstruction by Dora Preminger et al. nicely]:





We expect and have already felt serious resistance to the suggestion of any revision of any of the Sunspot Series, because of the obvious impact on so much other work and studies, but we feel that Rudolf Wolf would agree and perhaps be delighted.