The case of the missing solar cycle

Jan Janssens
Published in Heelal – January 2008

Introduction

According to the most recent NOAA-predictions, the next solar cycle would start in 2008. At the same time, this would end the ongoing 23rd solar cycle. Officially, these series of solar cycles start in 1755, but the first telescopic solar observations date already back from 1610. At the time, it concerned only sporadic observations, in contrast to the last 250 years when the sun was systematically observed. So there was quite some commotion when Ilya Usoskin announced in 2001 that he and his team had detected a new, short but weak cycle in this official series. Additional research in subsequent years has the existence of this extra cycle neither denied nor confirmed. However, it is certain that the end of the 23rd solar cycle will add a new piece to this mystery.

The solar cycle

The rhythmical variation of the solar activity with a period of about 11 years was discovered by Samuel Heinrich Schwabe in 1843. Though this discovery ranks among one of the most important in astronomy, it was actually only a spin-off of his solar observations he had started in 1826, as he was looking for the planet Vulcan. This was a hypothetical planet supposed to be located between the sun and Mercury, and should explain the observed deviation in Mercury's orbit. Thanks to Einstein and his Theory of Relativity, we know today that such a planet is not necessary. Schwabe however meticously noted all spots visible on the sun, and saw in the yearly variation already after only 17 years the existence of the solar cycle.

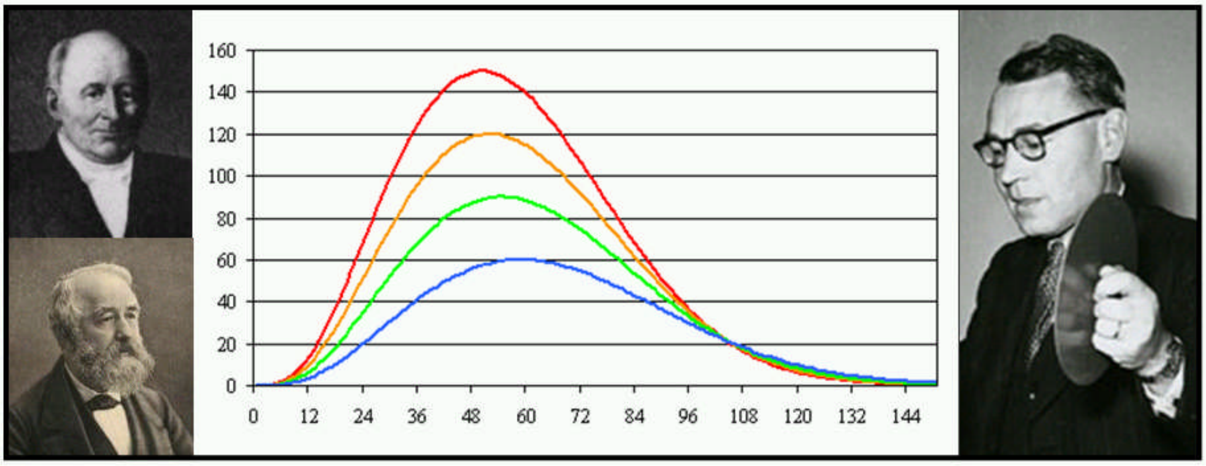

Figure 1: With their research, Samuel Heinrich Schwabe (upper left), Rudolf Wolf (lower left) and Max Waldmeier (right) laid the foundation for the solar cycle, the Wolfnumber and all the other features of solar cycles. The curves model solar cycles with maximum Wolfnumbers of 150, 120, 90 en 60. Active cycles reach their maximum faster than weak cycles.

Schwabe's results made a deep impression on Rudolf Wolf, who was director of the Bern Observatory at that time. He started a profound investigation into sunspots and introduced the sunspotnumber, which would later be named after him the Wolfnumber (R). This parameter is nothing else than the sum of the total number of spots and ten times the number of sunspot groups g. In other words: $R = 10 \times g + f$. 2 sunspot groups with a total of 12 spots give a Wolfnumber of 32. Wolf extended his datafiles by also incorporating historical observations into his research, initially till 1745, but later even all the way back to 1610. That way, he could finetune the 10 year value deduced by Schwabe to a more accurate average of 11,1 years.

Wolf also started an international cooperation program between the observers. Their observations were used when the weather in Switzerland did not allow for reliable observations. Evidently, a correction-factor k was applied to these observations to get them comparable to Wolf's. Though the current Wolfnumbers are coordinated from Belgium (SIDC in Uccle), a similar methodology is still in use.

Waldmeier, one of Wolf's successors as director of the Swiss Federal Observatory in Zürich, deduced from the observational records many important features of solar cycles. One of the most important states that active solar cycles, thus having a high maximum Wolfnumber, reach this maximum faster than weak cycles. Waldmeier formulated this law already in his doctoral thesis in 1935, but still today it is being used extensively in solar cycle research. The relationship was named after him the Waldmeier law.

In the decades that followed, scientists found even more "rules" to which solar cycles seemed to adhere to. For example, a cycle starting with a high minimum is more likely to become very active. Active cycles also tend to live not as long as inactive ones, and a short cycle is often followed by an active solar cycle. However, the connection between all the parameters is for most of these so-called rules a lot lower than for the Waldmeier effect.

In search of old observations

The absence of quite some observations between 1610 and 1850 certainly contributed to these poor relationships between the different solar parameters. Wolf only disposed of continuous, daily solar observations as from 1848. For the period between 1818 and 1848, he had to use averages of the other days to fill in the sporadic "gaps". Moreover, between 1749 and 1818, the number of days without solar observations was so large that he could only publish monthly Wolfnumbers. But even then the missing data often forced him to interpolate. For some months he even used geomagnetic observations to deduce the Wolfnumber.

Thus till 1848, the official Wolfnumbers R_z ("z" from Zürich) are a mixture of direct solar observations and deduced (calculated) values. According to some scientists, the result was that old cycles often had a "different" look than the ones we observe today. Consequently, this had a negative influence when one tried to establish the solar cycle laws.

In the early nineties, 2 American researchers -Douglas Hoyt and Kenneth Schatten- started the phenomenal task to gather more solar observations and to found a more solid base for the evolution of the solar activity. Not only did they want to fill in the "gaps" by using real solar observations, they also wanted a clear confirmation or denial of the Maunder minimum. This period, during which there was hardly any solar activity at all, took place so to say between 1645 and 1715.

Hoyt and Schatten developped a new parameter R_g based on the number of sunspot groups. They had noted that the old archives often gave descriptions or indications on sunspot groups, without detailing the number of sunspots. Moreover, 90% of the change in the Wolfnumber occurs because of variations in the number of groups. R_g also took into account the k-factor of the individual observers. It was tuned to R_z – values from the period 1874-1976, during which the Royal Greenwich Observatory actively performed solar observations.

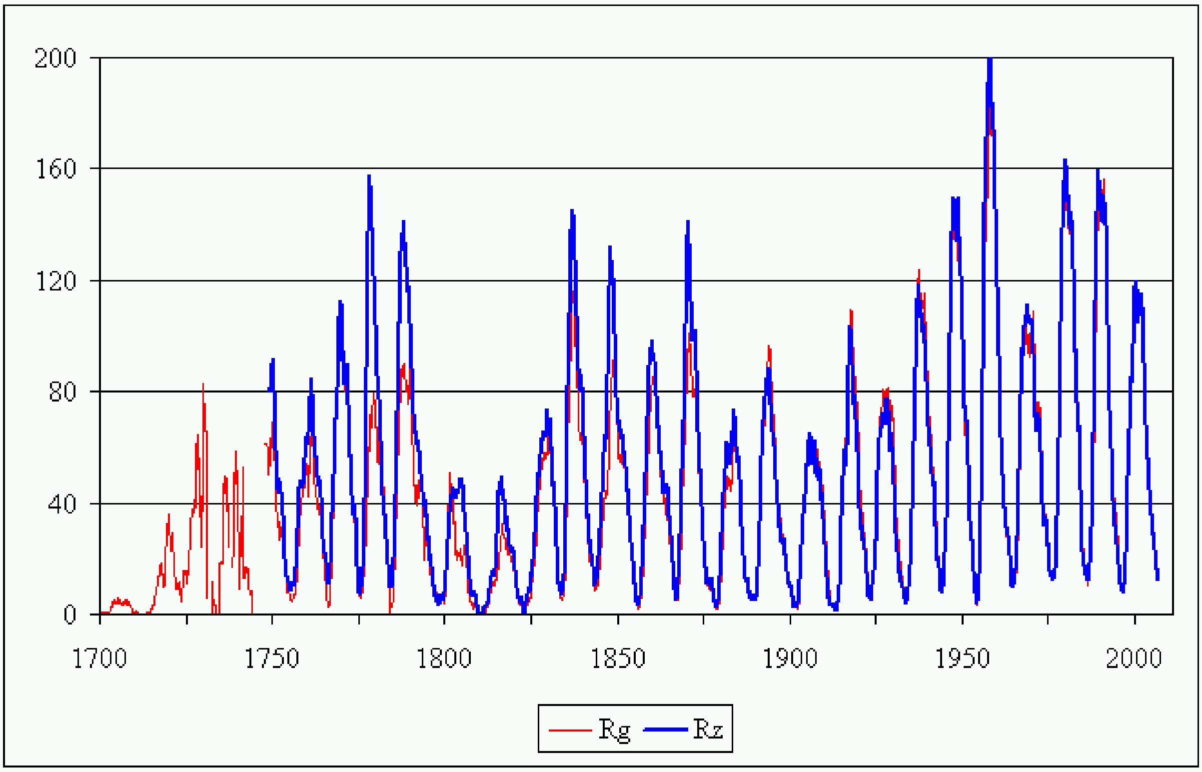

Figure 2: Evolution of the 13-monthly averages of the groupsunspotnumber R_g and the official (Zürich) sunspotnumber R_z. Prior to 1880, solar activity according to R_g is significantly lower than according to R_z. On the left (1700-1715), the end of the Maunder minimum can be seen. Despite the many additional observations gathered by Hoyt and Schatten, there are still relatively long periods without solar observations, like for example in the period 1740-1750.

In 1998, they published the results of their research in the renowned journal Solar Physics under the title "Group Sunspot Numbers: A new Solar Activity Reconstruction". The entire article, including the daily, monthly and yearly data, the standard deviations and the number of observers for the period 1610-1995, was not printed though, because it concerned over 9.700 pages! The data were digitized and can be consulted on the website of the National Geophysical Data Center. For the period 1610 to 1850, there were now 5 times as many days with solar observations available than Wolf could dispose of, and they confirmed the existence of the Maunder minimum. Moreover, the solar activity prior till 1882 seemed to be lower than what could be concluded from Wolf's data.

Of course, the usefulness and reliability of this new parameter needed to be tested thoroughly. 3 scientists from NASA's Marshall Space Flight Center took up the gauntlet and published their results in 2002. Hathaway, Wilson and Reichmann concluded that the general characteristics of a solar cycle were slightly better represented by the R_z-data. There existed also a better relationship with physical parameters like the solar flux and the sunspot area. The general evolution of solar activity prior to 1850 was nonetheless better represented by the R_g-data. More cycles and data (till 1610) were thus added to the series, resulting in an increase in the general reliability of the statistics.

The missing solar cycle

The average solar cycle lasts for approximately 11 years and reaches a maximum Wolfnumber R_z of 117. It takes a bit more than 4 years to rise from the minimum to the maximum (time of rise), and a bit more than 7 years to decrease from that maximum back to the new minimum (time of fall). Except for the recent cycles 17 and 20, most solar cycles do not satisfy this ideal picture, witness of which are the big uncertainty margins on the maximum (+/- 40) and the length of the cycle (+/- 15 months).

About 200 years ago, a most curious series of solar cycles occurred. The 2 shortest cycles – both of them did not last 9 years- were followed by cycle 4, which would itself last for over 14 years and became so the longest lasting of the entire official series. This trio also belongs to the active solar cycles, with maximum Wolfnumbers between 125 and 165. It was therefore all the more remarkable that they were followed by 2 cycles with a maximum R_z that hardly reached 50. This period is called the Dalton minimum, and it lasted approximately from 1795 till 1825. The somewhat more active cycle 7 ended these interesting series, with the peculiarity that the time of rise was 50% longer than the time of fall (respectively 6 and 4 years).

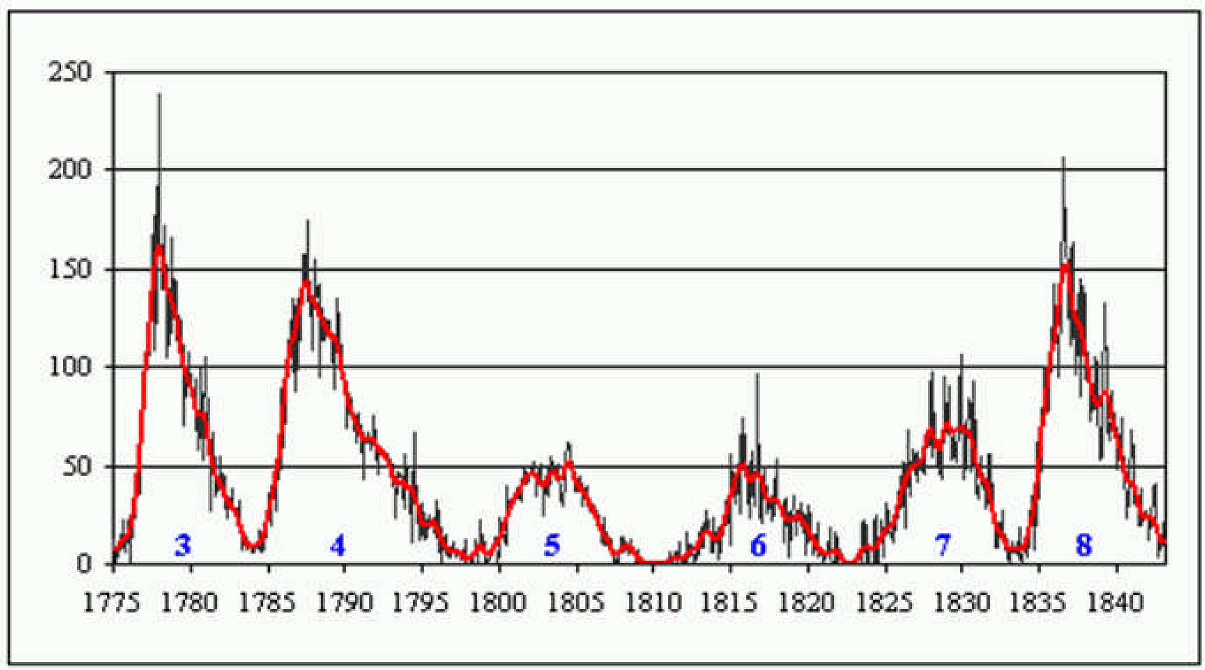

Figure 3: The quartet of abnormal cycles flanked by the somewhat more normal cycles 3 and 8. Of the 23 official cycles, cycle number 4 was the longest with over 14 years, cycles 5 and 6 had the lowest amplitude, and cycle 7 had a time of rise lasting 2 years longer than its time of fall. The Dalton minimum covers the cycles 5 and 6 and lasted approximately from 1795 tot 1825.

Especially the very long cycle 4 drew the attention of the Fin Ilya Usoskin. He and his collaborators did not fail to notice that a large part of the descending branch of this cycle (1790-1795) was badly supported by observations. For example, in the entire year 1792 there were only 4 (four) days with solar observations available. Moreover, the reliability of these observations was poor: On 3 April 1791, 6 different European observers saw between 1 and 6 sunspot groups!

An even more important reason to take a close look at that 4th solar cycle, was the existence of an often used rule concerning the solar cycle's intensity (this is the *sum* of all its monthly Wolfnumbers). The intensity of an uneven cycle always tends to be greater than the intensity of the preceding, even cycle. This is called the Gnevyshev-Ohl rule (1948), and it worked perfectly, except for… the 4th cycle.

Trying to find out what was going on, Usoskin used the groupsunspotnumber R_g as determined by Hoyt and Schatten. He used R_g and not R_z because the reliability prior to 1850 was higher. Next, he eliminated all single and unreliable observations (big k-factor). Finally, he drew the best fitting curve through the remainder of the observations. This new curve clearly showed an extra minimum early 1793 and an extra maximum in 1795.

The extra cycle had important consequences. The 4th cycle, originally lasting for over 14 years, could now be split up in a cycle of almost 9 years and one of almost 7 years. The 5th cycle would become 1 year shorter for the benefit of this extra cycle. Usoskin noticed also a much better symmetry in the distribution of the length of the solar cycles compared to the original series. Moreover, the Gnevyshev-Ohl rule behaved correctly for the 4th cycle, as well as all the way back to the beginning of the telescopic solar observations in 1610. Finally, the introduction of the new cycle resulted in a similar behaviour in solar activity both during the Dalton as during the Maunder minimum: a very sudden decrease of a normal cycle towards very low solar activity followed by a gradual restoration of this activity.

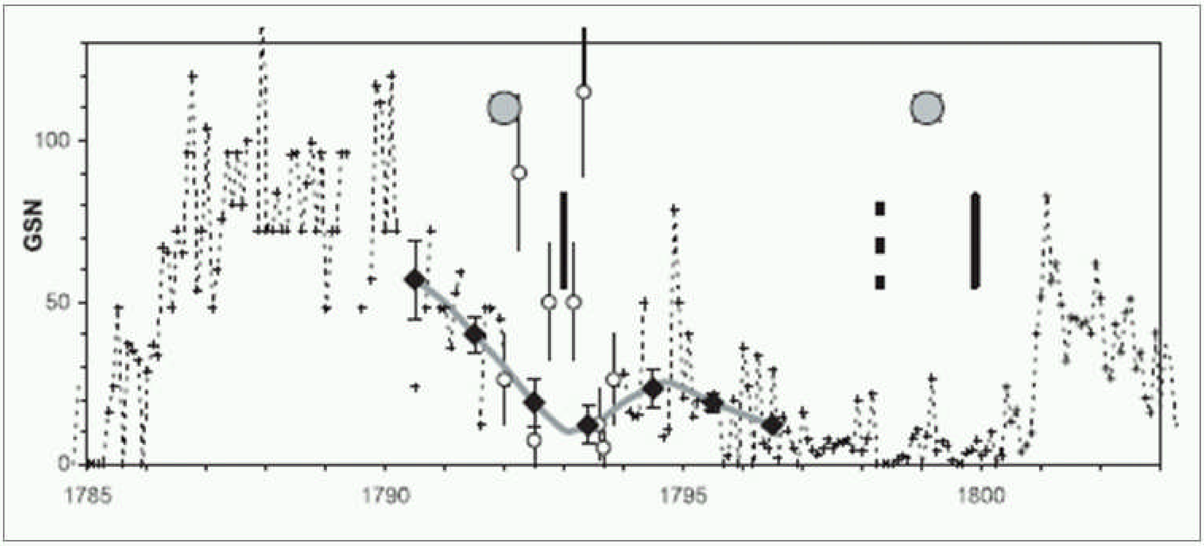

Figure 4: According to Usoskin (2003), the new cycle is situated between the 2 thick vertical bars: from 1793,1 till 1799,8, with a maximum of 35 +/-15 in 1795,0. Vertically the groupsunspotnumber(GSN, R_g), except for the period 1792-1793 during which Usoskin used his own estimated values (open circles with error bars). The big grey circles indicate naked eye observations of sunspots. The thick dashed vertical line indicates the end of the official cycle. [Usoskin, 2003]

Usoskin was of course very much aware he was skating over thin ice. In fact, the existence of this new cycle could only be proved based on the heliographic latitudes of the sunspot groups, but this kind of data did not exist for that timeframe. Alternatively, Usoskin went searching for indirect evidence by gathering data on aurora and natural radioactive isotopes. From the evolution of the Beryllium isotope, nothing conclusively could be deduced, but the aurora observations clearly indicated a maximum in 1796-1797. This would fit perfectly in the descending branch of the new cycle. Usoskin considered it very unlikely that this peak in geomagnetic activity would really belong to the original 4th cycle because in that case, it would have taken place 10 years after the maximum, which seemed really too late. At the very least, the additional data and analysis did not contradict the possible existence of an additional solar cycle.

Great doubts
Natalia Krivova, a Russian solar researcher of the Max Planck Institute, summarized all the objections against Usoskin's new solar cycle in just one sentence: Absence of solar observations does not automatically mean an absence of sunspots.

In an article in Astronomy&Astrophysics (2002), she and her collaborators proved that there did exist quite a bit more observations with sunspot groups in the period 1792-1793. Moreover, according to Chinese sources, there appeared a naked eye sunspot group in 1792: an event that is likely to happen 4 times more often when R_g is above the cycle average than under it. Periods with many consecutive spotless days are also not uncommon so many years after solar cycle maximum. Even the average solar activity during those 2 years corresponded very well with that during other, better observed solar cycles (3 to 5 years after the maximum). All this indicated that the sun at that time certainly did not behave as during a period of minimum activity, in contrast to Usoskin's claims.

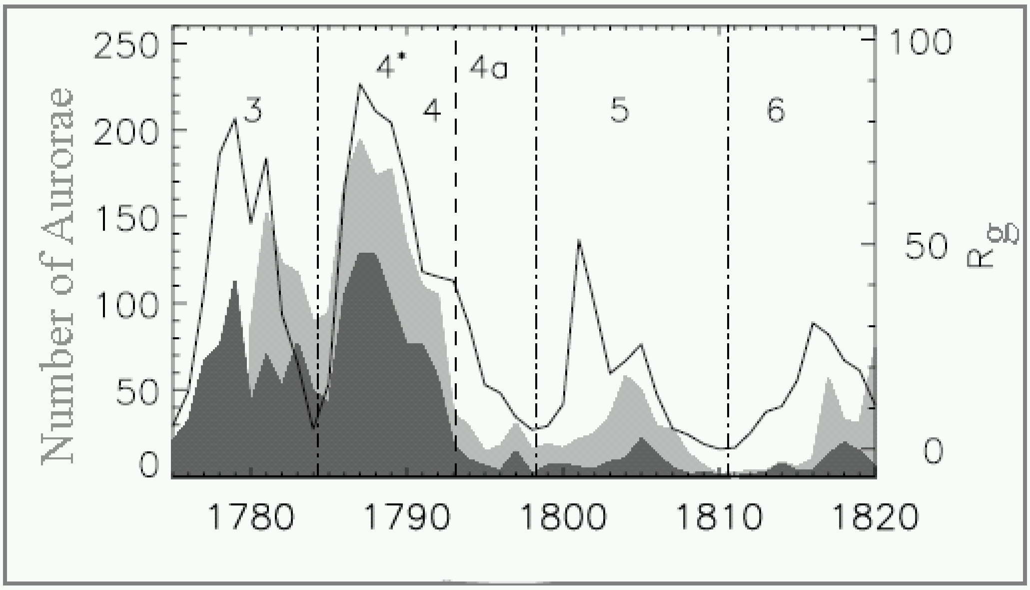

Figure 5: According to Krivova (2002), a few things simply don't add up when the new cycle is being examined in the light of the number of aurora observations (left axis). The light grey area represents the number of worldwide aurora observations beneath the geomagnetic latitude of 62°, and the dark grey area the number beneath 55°. Though a small aurora- maximum can be seen in the descending branch of the new cycle (1797), this activity is higher at the beginning of this cycle (1793), and the cycle maximum corresponds with the aurora minimum (1795). As a comparison, the evolution of the groupsunspotnumber R_g is given and represented by the full line (right axis). The minima of the official cycles are indicated by the dot-dashed vertical lines. There does not seem to be any need for a new cycle to explain this aurora evolution. [Krivova, 2002 – adapted]

Even more remarkable were her conclusions on the aurora. Krivova confirmed a small maximum in 1797, but noted immediately that the aurora activity in 1792-1793, during the so-called minimum of the new cycle, was much higher than during this aurora maximum. Moreover, the maximum of the new cycle (1795) fell together with a local minimum aurora activity, what has never been seen with any other solar cycle. Finally, the aurora activity in 1797 was very concentrated and confined to the beginning and ending of the year. All these conclusions undermined any existence of a new solar cycle.

In her article, Krivova spent an entire article on the unusual length of the old and new cycles. Unfortunately, she made an error in her calculations thereby nullifying her entire argumentation. Moreover, while determining the average solar activity in 1792-1793, she had not taken into account the (un)reliability factors of the concerned observers. As a result, her average R_g-value was twice as high as Usoskin's and in error. Usoskin saw his conclusions on the new cycle even supported as new findings turned up another naked eye sunspot in February 1799, exactly halfway the minimum of (old) cycle number 4. Thus, the appearance of the naked eye spot in 1792 did not necessarily have to be considered as a sign of high activity. Only on the refutation of Krivova's remarks concerning the aurora activity he remained vague, but of course it concerned only an indirect parameter.

The 23rd solar cycle

The existence of the new solar cycle remains very debatable. As long as no new proofs from additional observations or latitude determinations of sunspots become available, Usoskin's cycle will not be included in the series of official solar cycles.

More importantly, it seems that an indirect reason for the existence of the new cycle will soon be no longer valid. The Gnevyshev-Ohl rule states that the intensity of an uneven cycle is always higher than the preceding even cycle. This means that the intensity of the ongoing 23rd cycle would be greater than that of the 22nd cycle. By the way, this law was one of the reasons for NOAA to predict an active 23rd cycle, possibly even more powerful than the already very impressive 22nd cycle.

In reality, cycle 23 remained well below the activity level of its predecessor. With a maximum Wolfnumber of about 120, it is considered as a moderate cycle only. Of course, this doesn't say anything on the intensity of this solar cycle. The fact is that scientists have been watching in awe the very high levels of activity far after the maximum. For example, none will forget the "Halloween" groups that appeared in October and November 2003. Not only did 3 of the top-10 of largest sunspot groups of this cycle transit the solar surface, NOAA 0486 was also responsible for producing the most powerful solar flare since the start of satellite observations. But also in the summer of 2004, in January and September 2005, up till December 2006 (NOAA 0930), complex and very active sunspotgroups continued to appear.

So, the intensity of the 23rd cycle is becoming higher than what one could expect based on the maximum Wolfnumber. Nonetheless, all indications are that the Gnevyshev-Ohl rule will not be respected this time. The intensity is at the moment still about 10% lower than that of the foregoing cycle. Though an international team of solar specialists stated in April 2007 that the beginning of the new cycle is not to be expected before March 2008 (with an uncertainty of a half year), the remaining time and current solar activity level seem not high enough to lift the intensity over that of the 22nd cycle. Interesting to note is that if cycle 23 would end in March 2008, it would have last almost 12 years. That's a long cycle, but not as extreme as the 4th cycle. And maybe that if the intensity is calculated using R_g , it will still be greater than that of solar cycle 22…

Whatever the outcome, the end of the 23rd cycle will add a new piece to the mystery of the missing solar cycle.

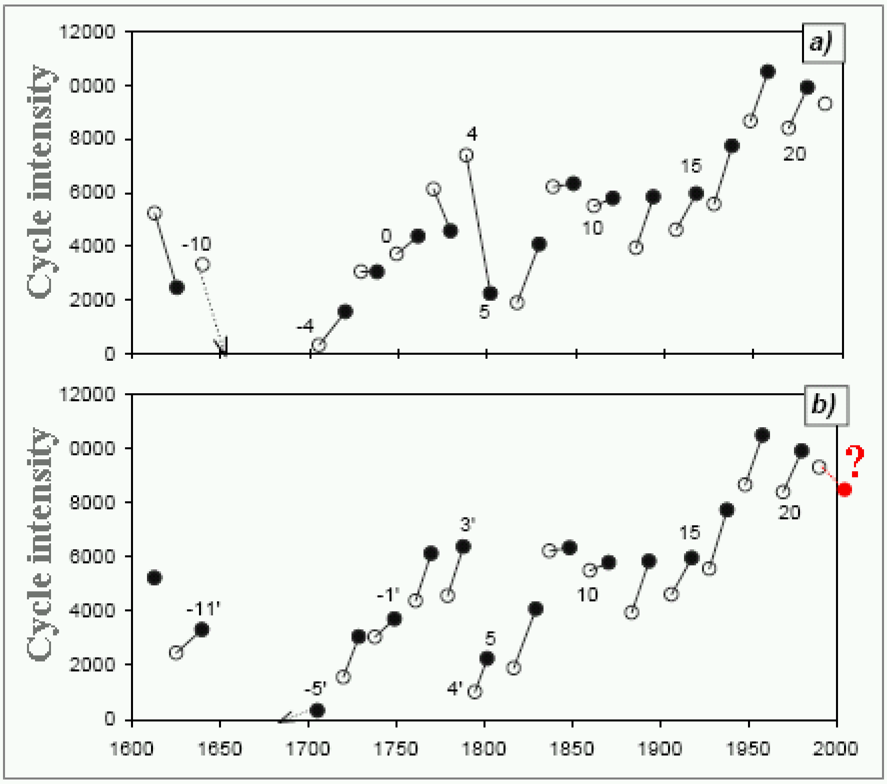

Figure 6: The Gnevyshev-Ohl rule states that the intensity of an uneven solar cycle (full circles) is higher than that of the preceding even cycle (open circles). In the top figure, the official numbering is used. It is clear that the G-O rule isn't valid for the solar cycles 2-3 and especially 4-5.In the bottom figure the new solar cycle is introduced, as well as the new numbering (prior to cycle 5). The G-O rule is now restored and is valid for all cycles since 1610. However, a problem seems to appear for the ongoing 23rd solar cycle, which has still an intensity lower than that of the preceding cycle 22, and thus looks like it is going to violate the G-O rule. [Usoskin, 2001 – adapted]

References and literature

- Hathaway, D. H., Wilson R. M., & Reichmann, E. J. 2002, Sol. Phys., 211, 357
- Hoyt, D. V., & Schatten, K. 1998, Sol. Phys., 179, 189
- Krivova, N. A., Solanki, S. K., & Beer, J. 2002, A&A, 396, 235
- Usoskin, I. G., Mursula, K., & Kovaltsov, G. A. 2001, A&A, 370, L31
- Usoskin, I. G., Mursula, K., & Kovaltsov, G. A. 2002, Geophys. Res. Lett., 29, 36
- Usoskin, I. G., Mursula, K., & Kovaltsov, G. A. 2003, A&A, 403, 743

- Beck, R., Hilbrecht, H., Reinsch, K., & Völker, P., Solar Astronomy Handbook, 1995
- Janssens, J., Zon en Aarde: Een unieke Relatie, 2003
- Taylor, P. O., Observing the Sun, 1991

- NGDC (R_g) http://www.ngdc.noaa.gov/stp/SOLAR/ftpsunspotnumber.html#hoyt
- SIDC (R_z) http://sidc.oma.be/sunspot-data/