The sunspot number clarified

Which kinds of daily International Sunspot Numbers (ISN) are currently published by the SIDC?

Three daily indices are currently published by the SIDC:

the Estimated ISN (EISN),
the Provisional ISN (PISN),
the Definitive ISN (DISN).

Although they are all based on the same type of observations, i.e. visual sunspot counts as a measure of solar activity, they are obtained from different data sets, by different processing methods, with different accuracies and they serve different purposes.
In the ideal case the three numbers are equal, but in practice it is inevitable that there are small differences. Therefore, It is useful to explain what each index actually means.

What is the source of the Wolf number?

Once per day, observers worldwide communicate their measured Wolf number (= 10G + S), based on visual sunspot (S) and group (G) counts, to the World Data Center for the Sunspot Index. Presently, this is done through a dedicated password-protected WEB form: the Wolf interface. However, some observers still use the old way of reporting their observations on paper, through regular mail. This method slows down the final processing, but for the sake of continuity in the index it is important to keep these observers included, as some of them represent the longest individual datasets.
The observers give the exact observing time, the observing conditions, the total numbers (group count, sunspot count and Wolf number), and optionally, the North/South and central zone numbers. The observers using the web interface can provide their results to the SIDC immediately after each observation. However, some users prefer to insert their data as a block, creating some delay, but this must be done at least once a month.

The personal reduction coefficient K

A personal reduction coefficient K is calculated for each station. It has values typically in the range 0.4 and 1.7 and leads to a *normalized* sunspot number K(10G+S). In other words, the K factor rescales the raw sunspot counts of each observer to the ones of Rudolph Wolf, the astronomer who introduced the above Wolf formula, thus simulating the same eyes, same telescope and same conditions.
During the procedure for calculating the Provisional ISN, the K factor of every station is computed for every observation that passes the elimination procedure. Once a year, a personal K factor for every station and for every month of the previous year is calculated. A yearly mean K value per station is also computed.

The estimated sunspot index

The Estimated ISN (EISN) is calculated and issued on a daily basis for the day of calculation and the day before. Only the total EISN is calculated, with no hemispheric index. It is an automated procedure and no manual intervention is done. The EISN is meant as a quick temporary and approximate index for real-time applications.
We use the raw Wolf number from all observatories, that have provided counts early in the day, i.e. before 12:30 UT. Typically, there is between 10 and 20 stations available.
The calculation of the EISN is based on a simple and straightforward statistical method. The Wolf numbers are multiplied with the yearly mean K factor (see paragraph about the

personal reduction coefficient K). The EISN is a simple average of the K-scaled values after excluding any individual value that deviates abnormally from the other values.
Once the Provisional ISN (PISN) is calculated after the end of a month, the temporary EISN is dropped and replaced by the PISN.

The provisional sunspot Index

The Provisional ISN (PISN) is calculated and issued on the first of every month. The outcome is a daily total and hemispheric sunspot number for the entire elapsed month. An example: on the first of September, we calculate the daily PISN for each day of August. The PISN is computed semi-automatically with as little manual intervention as possible.
To calculate the PISN values for the concerned month, we use all Wolf numbers. Typically there is data from around 65 stations available.
The calculation of the PISN is based on an extensive statistical treatment. This treatment filters out any abnormal daily observation, or the complete data set from an erratic station (quality control). This prevents long-term drifts in the resulting index. One of the key elements derived from this processing is to update the K reduction coefficients for each station: the evolution of each station is monitored for each day relative to the entire network.
A subset of the contributing stations also provides hemispheric counts. A similar process is applied to the North and South counts to derive the hemispheric PISN. There is an additional constraint: the sum of the North and South ISN must match the total ISN.
The PISN replaces the EISN. And, once the Definitive ISN (DISN) is calculated, the DISN replaces the PISN

The definitive sunspot index

The Definitive ISN is calculated and issued on a quarterly basis, when we have collected data from all the contributing observatories. This is the final index and it is appended to the historical sunspot index time series. The same semi-automated treatment, as for the PISN, is now applied to the full data set. If the result agrees within 5% of the PISN value, the PISN is kept and becomes definitive. Otherwise, the new calculated value replaces the PISN and becomes the definitive ISN.
It is only at this stage, and only in peculiar situations that a manual verification by our scientific staff is required. For example, this happens when the observed raw Wolf values from the whole network are clustered around two distinct values. This special case can be caused by small short-lived isolated sunspots (lifetimes < 24 hours) or by the appearance or disappearance of a large group at the solar limb in the course of one UT day.
The DISN replaces the provisional and estimated International Sunspot Number.

Handling very low activity levels

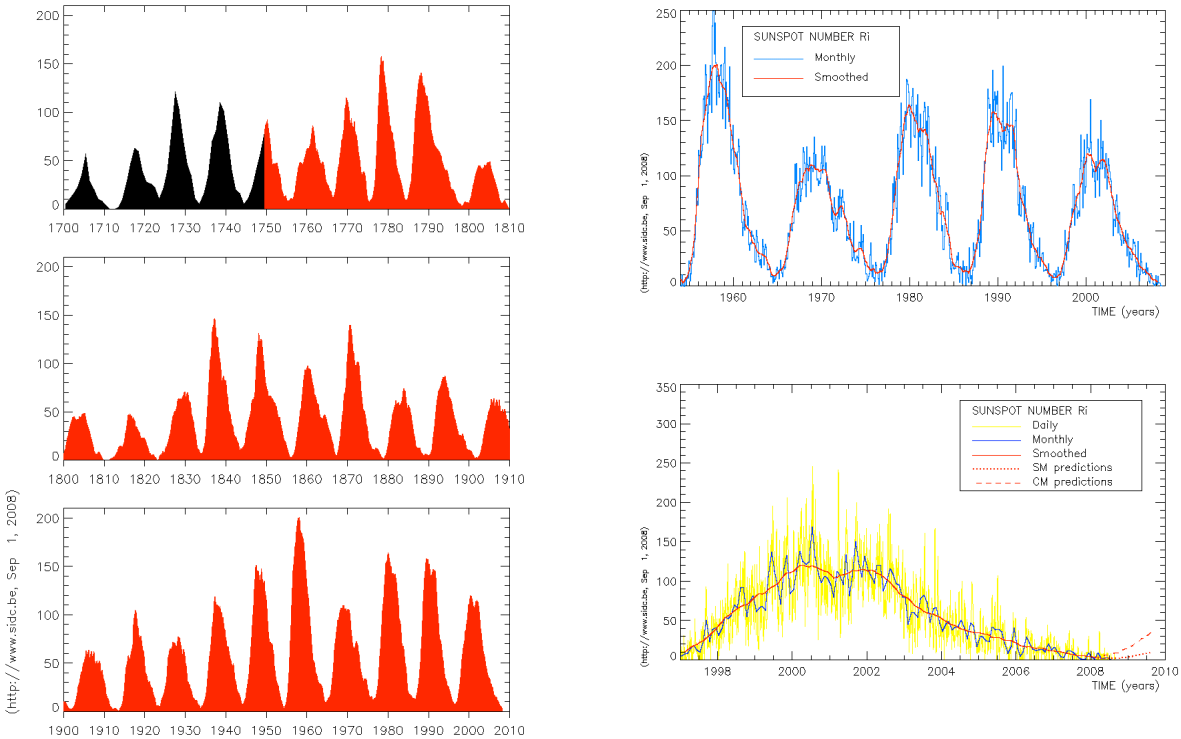
One of the situations requiring a human arbitration is associated with single small isolated and short-lived sunspots. This is most noticeable around the minimum of the solar activity cycle, like the one occurring now in 2008. Due to the 24hour binning (observations are grouped between 0hUT and 24hUT) and the variable ability of each observer to detect the smallest sunspots, we end up with some stations reporting one sunspot and others who did no see any sunspot. Let us explain further. *No sunspot* may mean that the sunspot was present but the observer was unable to see it because of poor observational conditions, such as weather, a small telescope... On the other hand, it may also indicate that the sunspot had actually vanished by that time, while it was present earlier on the same day. Individual stations do not observe at exactly the same time.
In order to validate the existence of such a marginal reported sunspot, a qualified SIDC scientist must then check the detailed chronology within that day and consider the overall observer capability (value of K coefficient).
In common practice, the fact that a sunspot is reported by a significant group of observers

leads to the inclusion of the sunspot on that date, thus neglecting the no-sunspot-observations. The rationale is: multiple observations of a sunspot exclude the possibility that the reported sunspot was an independent false detection. So, a sunspot was really present on that day, although it may have existed only for part of the day or it was small enough to be missed by part of the observers with smaller instruments or imperfect atmospheric conditions.
Also keep in mind that the sunspot index is derived with a limited precision, just like any other index. If the monthly mean sunspot number is 0 or 0.5, you can definitely say that activity was low.

What index should be used?

If you want to perform long-term investigations, the definitive ISN series is definitely the most suitable. However, there is a delay of a few months. So, when investigating the last cycle and recent evolution, use the definitive numbers in combination with the provisional ISN to be up to date to the last month. The provisional numbers are also used in models forecasting the sunspot number for the coming months. Now, if you need a proxy for solar activity in a model that runs in real-time, then you may use the estimated ISN.

Improvements and rethinking of the processing method is an ongoing project. We are currently developing an alternative program for calculating the PISN. Of course, long-term consistency is vital. Therefore, a cross-analysis between the output from the old and new software must be applied over an extended period.



Above graphs show different kinds of the International Sunspot Number. The left plot gives an overview from 1700 up to now from the yearly (black) and monthly smoothed Sunspot Number. The graph in the right top corner shows the monthly and monthly smoothed sunspot number, while the graph in the right bottom corner show the daily, monthly, smoothed monthly sunspot number. For every the time range of investigation, there is an appropriate sunspot number.