



WDC- Sunspot Index and Long-term Solar Observations

Preservation Plan



- **Storage:**

The technical aspects of the storage are described in the Technical Infrastructure document.

The small volume of the SILSO data allows easy storage on ordinary low-capacity media. The strategy is to save multiple local copies on independent computers and storage. In addition, we take advantage of the long-term backup strategy for the servers of the ROB (described in the *Technical Infrastructure document*). Outside the WDC itself, this thus only involves the responsibility of the hosting institution, which is committed to the long-term preservation of this data service through a formal letter of agreement with WDS/ICSU.

The depositors, in our case the stations/observers of the SILSO network, are actually entering a copy of their own observations to contribute to the primary index calculation. Therefore, there is no transfer of a unique original of the observations to the WDC-SILSO. Observers keep their own local original version of their observations. Entering their data in the WDC-SILSO database only provides an additional central storage giving better guarantees of the long-term preservation of their data (separate independent copy) and of course, it makes their data scientifically usable for the production of this solar activity index. Cf. the crowdsourcing nature of this process.

The storage of the main output series is in preparation on a general-purpose open-access repository. It will complement the data provided by the WDC-SILSO and can replace it in the unlikely event of a problem that would prevent our infrastructure from recovering. However, by the continuously-growing nature of the input database (updated every 5 minutes), the raw input observations will not be stored on this kind of platform, which is designed for complete versions of data sets.

- **Safety strategy - Migration plan - Continuity of access:**

To mitigate the risk associated with the minimal size of the team dedicated to the WDC-SILSO, in 2011 we started actions beyond the ROB:

- **Relocation of operations through a network of partners:** Few other teams have the scientific expertise to take over a WDC dedicated to the sunspot number. Therefore, personal informal contacts are maintained with a few competent colleagues, with whom we have collaborations. In this way, relocation candidates are identified and can be solicited if the continuation at the ROB proves impossible, similarly to the past Zurich-Brussels transition, which largely involved pre-existing personal contacts (M. Waldmeier and A. Zelenka in Zurich with A. Koeckelenbergh and P. Cugnon in Brussels).
- **External permanent storage:** While action 1 above would ensure the relocation of WDC operations (production of new numbers, coordination of the observing network), we also considered the survival of the data themselves, in particular following our recent initiative to re-calibrate the data series in the context of a community effort (cf. *other sections below*). As this involves the latest data tools, we are now preparing the permanent storage of master copies of the successive versions of our data set on a general-purpose open-access repository. A new revision of the sunspot series is in preparation for the course of 2019: on this

occasion, the chosen repository will be implemented and each new data set will thus also get a DOI number, ensuring both independent permanent storage and version tracking.

- **Perpetuating the expertise:** Over recent years, in the framework of the first re-calibration of the sunspot number series, we published an unprecedented number of scientific articles documenting the past production and the identified defects in the sunspot number (Cf. the ROB bibliography page <http://www.sidc.be/publications/index.php>, choosing Clette and Lefèvre as authors). A full technical documentation of the data-production software ("user manual") is in preparation.

- **Data format and structure – Methods:**

An entirely new version of our core production software will be created in 2019-2020, and will be stored externally through a web-based hosting service for version control allowing the traceability of the software, next to the data themselves.

With the modernization of all the data production software (documentation, porting codes from heritage-Fortran to Python, graphical products, porting data from ASCII files to databases) we plan on introducing headers and licences directly in the produced files. Thus the metadata will be much more developed.

All base data are provided as simple uncompressed ASCII text files, lowering the potential compatibility issues with future software. For the convenience of some of our users, we started to produce files in CVS (comma separated values) format in 2014.

- **Data quality control:**

- The quality of the methods and datasets is monitored by the scientific community through the use of the produced datasets published on the WDC-SILSO website.
- Scientists from the Sunspot workshops, largely represented in the ISSI team are working on the methods applied or to be applied to past and present sunspot numbers and are thus monitoring the data and methods quality.
- A more formal scientific overseeing committee is being created involving the IAU, IAGA and URSI entities.
- In addition to that, a BELSPO BRAIN.be project is progress, which aims at a complete monitoring of the observing stations and processing in almost real-time.

- **Preservation period of the data:**

Considering the infrastructure of the Royal Observatory of Belgium, the guaranteed period should be eternity. However, to be reasonable, we consider different options of storage and we prefer to guarantee a preservation period of 25 years (that is the guaranteed lifetime of most general-purpose open-access repositories).



- **Licenses/Data rights:**

Sunspot counts are "basic observations" (not creative) and therefore not copyright-able. The processing done by WDC-SILSO from basic sunspot counts to the sunspot index is a creative act and therefore copyright-able. The CC BY-NC License applies to the sunspot index.