Federica Guarnaccia
July 14, 2026

SMOS: The Unexpected Sun Sentinel – A guide to the SMOS Space Weather Prototype Data

LIST OF CONTENTS

1. GENERAL INFORMATION 3

1.1 Instrument Description 3
1.2 Image Reconstruction 4
1.3 Stokes Vector Reconstruction 6
1.4 THE SMOS Space weather products 6

2. THE SMOS SOLAR FLUX 11

2.2 Processing Pipeline 11
2.3 Product Format 12

3. THE SMOS SOLAR RADIO BURST DATABASE 16

3.1 Processing Pipeline 16
3.2 Product Format 19

4. THE SUN BT ADF 25

4.1 Processing Pipeline 25
4.2 Product Format 26

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

Federica Guarnaccia
July 14, 2026

Version / Date	Content
1.0 / 20260401	First Release
1.1 / 20260508	Changed ADF creator ID from 3 to 7 for L1B AUX and improved file naming convention description.
1.2 / 20260526	Changed folder name of the solar radio burst prototype dataset.
1.3 / 20260706	Recalibrated DoCP in SRB products and streamlined nomenclature
1.4 / 20260706	**Minor text change to better explain OPER/REPR differences**

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

1. General information

1.1 INSTRUMENT DESCRIPTION

The Soil Moisture and Ocean Salinity (SMOS) mission is led by ESA as part of the Earth Explorer missions and consists of a microwave imaging satellite originally developed to retrieve soil moisture and ocean salinity maps. The mission's instrument is a single passive microwave payload called MIRAS: Microwave Radiometer using Aperture Synthesis, meant to collect pixel brightness temperatures. It uses 69 Lightweight Cost-Effective Front-end (LICEF) receivers placed in a three-arm Y configuration to measure the phase difference of incident radiation from an area at different angles, allowing for a much more detailed image compared to that of a single receiver.

Frequency/Wavelength	1.413 GHz / 21 cm
Swath width	1050 km
Spatial resolution	35-50 km
Orbit	sun-synchronous (polar)
Altitude	758 km
Inclination	98.45°
Orbital period	100.1 min
Repeat cycle	149 days
Boresight incidence angle	32.5° (along-track)

Table 1: SMOS and MIRAS main properties.

Figure 1: SMOS artistic view, credit: ESA, AOES Medialab

The instrument thus uses three coplanar arms consisting of an elementary antenna regularly spaced (0.875 λ, maximum redundancy) in a Y-shaped configuration with each deployable antenna arm having a length of 4 m. An oversampling by a factor of about 2 with respect to the Nyquist criterion is achieved in each receiver, which improves the radiometric sensitivity. Images are retrieved with an interferometric Fourier synthesis over the correlations between each pair of antenna elements. Each LICEF is an antenna-receiver integrated unit measuring radiation in the1404-1423 MHz band.

Each of the receivers can be commanded to measure either H and V polarization or an internal load. A calibration network enables correlative measurements for the entire array. This includes:

• A central redundant noise diode that is distributed to all receivers of the hub.

• Distributed noise diodes on each arm segment to measure both segment receivers and the adjacent ones on either side of the hub or the next element.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

Three NIR (Noise Injection Radiometer) at the hub measure the H and V polarization simultaneously. All individual receivers digitize the measured I and Q signals and send them, via a point-to-point optical fiber connection, to DICOS (Digital Correlator System), in which the correlations between all possible receiver pairs are made and accumulated for the duration of the integration period (1.2 s).

MIRAS can work in different configurations:

1. Dual-polarization mode, in which all receivers are switched to either Horizontal (H) or Vertical (V) polarization.
2. Full polarization mode, in which segments of the array are switched according to a predefined sequence between H and V. Through co and cross-pol acquisitions, the full Stokes vector can be reconstructed. This is the default operational mode, running since June 2010.
3. Calibration mode, in which internal systems are calibrated.

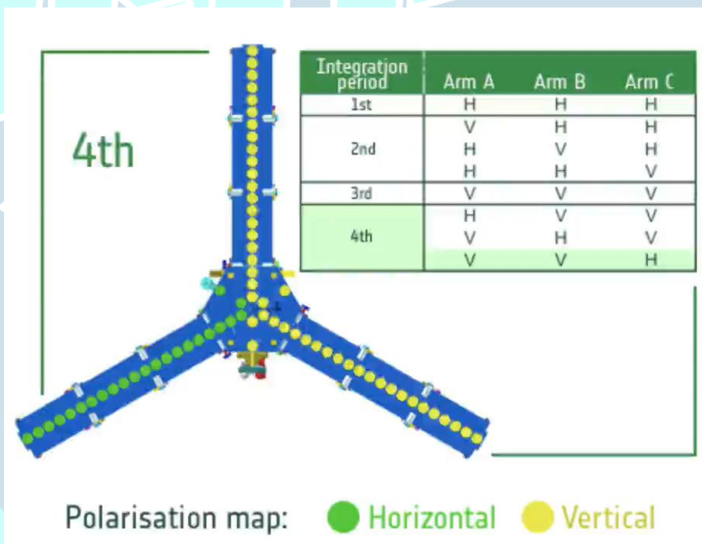

Figure 2: Acquisition sequence to reconstruct the full Stokes vector in full mode.

1.2 IMAGE RECONSTRUCTION

Aperture synthesis radiometers measure both the antenna's temperature (thermal noise power collected by a single antenna) and visibilities (complex cross correlation between signals collected by two antennas).

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

Measuring the visibility with different pairs of antennas (baselines) means measuring the Fourier transform between the visibility function and the brightness temperature, weighted by the antenna pattern as a function of antenna spacing[1].

As the brightness temperature is a 2D function defined inside the unit cycle ($\eta^2 + \xi^2 \leq 1$), the Fourier transform can be sampled over a hexagonal grid. However, since the antenna size and spacing do not satisfy the Nyquist criterion, the alias-free field of view (FOV) is limited by periodic repetitions of the unit circle. The alias-free FOV can still be extended (EAF-FOV) up to the Earth-Sky border since a large part of the unit circle is occupied by the sky, and its temperature is well known.

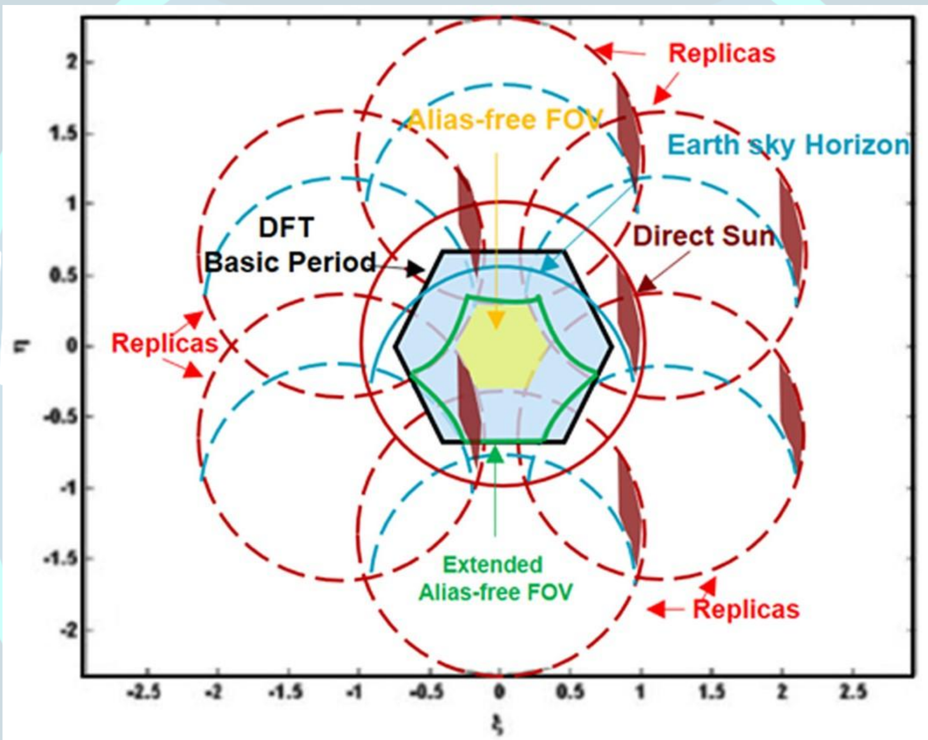

Figure 3: Antenna's FOV: hexagonal sampling and image replicas.

The visibilities are first corrected from the contribution of the physical temperature of the receivers by using a set of measured calibrated visibilities obtained from deep Sky observations. Then, these are corrected from the contribution of the foreign sources of brightness present in the FOV, including: direct Sun, Moon, and Sky radiation, and Sun glint effects.

[1] The time delay from a given source point to each pair of antennas introduces a complex exponential term that is a function of the antenna spacing.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

The Sun is present in roughly 97% of the SMOS images in the Sky region outside the reconstructed hexagon. Since one of the six aliases is usually located inside the EAF-FOV, the Sun alias contaminates the SMOS images, entering either above or below the antenna plane:

- If the Sun is above the antenna plane, the solar radiation is captured by the front-lobes antenna patterns, resulting in high contamination of the images.
- If it is below the antenna plane, the radiation is captured by the back-lobes antenna patterns. Since their gain is very small with respect to those in front, this results in a minimum and mostly local contribution.

Dedicated algorithms have been developed to retrieve this signal and remove it from the SMOS images. This is stored inside the SMOS L1B dataset.

1.3 STOKES VECTOR RECONSTRUCTION

As LICEFs are scheduled to operate in a pre-defined sequence, the H and V components of the solar flux are retrieved in the first and third integration period, respectively, as well as in the second and fourth alongside the HV (VH) component. In the end, the full Stokes vector is reconstructed as:

$$I = T_{VV} + T_{HH}$$
$$Q = T_{VV} - T_{HH}$$
$$U = -2 \cdot Re(T_{HV})$$
$$V = -2 \cdot Im(T_{HV})$$

These BT are originally evaluated in the antenna's best fit plane and are rotated in the Sun frame, a coordinate system where:

- Y – axis: the orientation of the Sun's rotation axis (Sun's North Pole direction from the Carrington Elements) rotated in J2000 and projected on the plane perpendicular to the Earth-Sun vector.
- Z – axis: Earth-Sun direction, oriented forward to the Earth.
- X – axis: computed using the Cartesian rule.

Note that only I and V are independent of the system's rotation.

1.4 THE SMOS SPACE WEATHER PRODUCTS

An Ubuntu virtual machine is in charge of the generation of the SMOS Space Weather products, freely available in the `Space Weather Products` folder, whose content is updated by running a pre-defined pipeline every day, which is orchestrated by crontab to start at 09:00 UTC.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

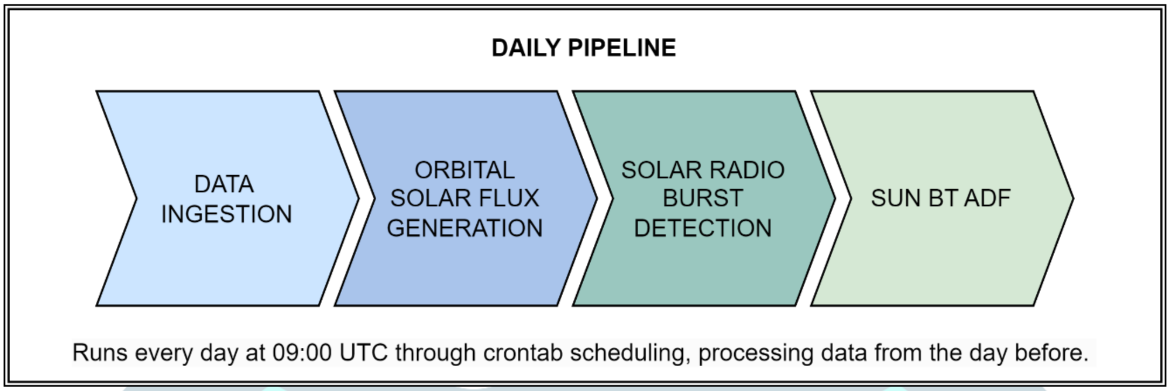

Figure 4: Daily processing pipeline main processes.

The daily pipeline executes the following tasks in the order below:

1. SMOS L1B data ingestion from the SMOS Dissemination server. Data are distributed in the standard Earth Explorer format with a binary data block file and a header.
2. SMOS orbital solar flux generation and dissemination.
3. Daily solar radio burst screening aimed at the generation of the operative bulletin and the dissemination of the plots and high time resolution data of the occurring solar radio bursts.
4. Sun BT Auxiliary Data File (ADF) update, aimed at Sun glint modelling for the Level 2 Ocean Salinity (L2OS) SMOS product.

The output of this processing pipeline is public for the solar radio burst detector and the orbital solar flux generator, while access to the Sun BT ADFs is private, and data are pushed to the agreed DPGS directory.

Additional products available for dissemination are:

- Solar radio burst entries injected after revision, including files for the false negatives of the daily routine updated monthly with data from the previous calendar month.
- Additional test files which may be generated ad hoc outside of the virtual machine with limited access to selected users. These folders are intended to share products generated for collaboration with external partners and may be shared with a wider audience after the test stage.

Both the orbital solar flux and the radio burst archive are sorted by year.

See the dedicated sections for each of the main SMOS Space Weather products for additional information about the generation pipeline and the data format and content.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

Data format of the output files and the specifics of the generation pipelines are outlined in the subsequent sections of this document.

Access to the products is provided upon request by generating named users and through

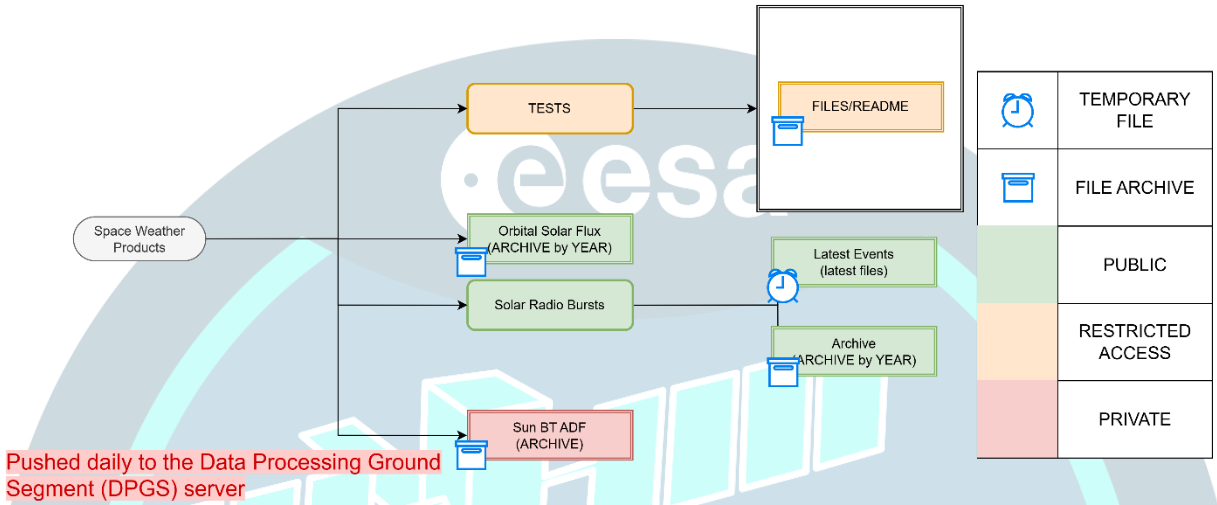

Figure 5: FTP structure of the virtual machine. Green is used for path destinations available for every user, while yellow is chosen for destinations meant to only be accessed by selected users, red is used for data uploaded to the DPGS server but stored locally in paths not accessible outside of the RedLab team.

public – private key exchange.

For more detailed information, feel free to contact the following e-mail addresses:

- ✉ federica.guarnaccia1@serco.com
- ✉ roberta.forte@ext.esa.int
- ✉ raffaele.crapolicchio@ext.esa.int
- ✉ redlab-oper@eo-sppa.org

List of References

Anterrieu, E., Khazaal, A., 2008. Regularized image reconstruction for the SMOS space mission. Journal of Physics: Conference Series 135. https://doi.org/10.1088/1742-6596/135/1/012008

Borges, A., SMOS-PLM Team, 2004. SMOS Mission & MIRAS Instrument. Synthetic Aperture Radiometer in Space. 12th GAAS Symposium, Amsterdam.

Camps, A., Corbella, I., Torres, F., Duffo, N., Vall-llosera, M., Martin-Neira, M., 2005a. The impact of antenna pattern frequency dependence in aperture synthesis microwave radiometers. IEEE Transactions on Geoscience and Remote Sensing 43, 10. https://doi.org/10.1109/TGRS.2005.854309

Camps, A., Corbella, I., Torres, F., Vall-llossera, M., Duffo, N., 2005b. Polarimetric Formulation of the Visibility Function Equation Including Cross-Polar Antenna Patterns. IEEE Geoscience and Remote Sensing Lett. 2, 3. https://doi.org/10.1109/LGRS.2005.846885

Camps, A., Vall-llossera, M., Corbella, I., Duffo, Nú., Torres, F., 2008. Improved Image Reconstruction Algorithms for Aperture Synthesis Radiometers. IEEE Transactions on Geoscience and Remote Sensing 46, 1. https://doi.org/10.1109/TGRS.2007.907603

Camps, A., Vall-llossera, M., Duffo, N., Zapata, M., Corbella, I., Torres, F., Barrena, V., 2004. Sun effects in 2-D aperture synthesis radiometry imaging and their cancellationIEEE Transactions on Geoscience and Remote Sensing 42, 6. https://doi.org/10.1109/TGRS.2004.826561

Colliander, A., Tauriainen, S., Auer, T.I., Kainulainen, J., Uusitalo, J., Toikka, M., Hallikainen, T., 2005. MIRAS Reference Radiometer: A Fully Polarimetric Noise Injection Radiometer. IEEE Transactions on Geoscience and Remote Sensing 43, 5. https://doi.org/10.1109/TGRS.2004.840667

Corbella, I., Duffo, N., Vall-llossera, M., Camps, A., Torres, F., 2004. The visibility function in interferometric aperture synthesis radiometry. IEEE Transactions on Geoscience and Remote Sensing 42, 8. https://doi.org/10.1109/TGRS.2004.830641

Corbella, I., Torres, F., Camps, A., Duffo, N., Vall-llossera, M., 2009. Brightness-Temperature Retrieval Methods in Synthetic Aperture Radiometers. IEEE Transactions on Geoscience and Remote Sensing 47, 1. https://doi.org/10.1109/TGRS.2008.2002911

Corbella, I., Torres, F., Duffo, N., González-Gambau, V., Pablos, M., Duran, I., Martín-Neira, M., 2011. MIRAS Calibration and Performance: Results From the SMOS In-Orbit Commissioning Phase. IEEE Transactions on Geoscience and Remote Sensing 49, 9. https://doi.org/10.1109/TGRS.2010.2102769

Corbella, I., Torres, F., Wu, L., Duffo, N., Duran, I., Martin-Neira, M., 2013. Spatial biases analysis and mitigation methods in SMOS images, in: 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS. Presented at the IGARSS 2013 - 2013 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Melbourne, Australia, pp. 3415–3418. https://doi.org/10.1109/IGARSS.2013.6723562

Crapolicchio, R., Forte, R., Guarnaccia, F., Di Ciolo, L., 2025. The L-Band solar flux retrieval from the Soil Moisture and Ocean Salinity (SMOS) mission: Theoretical algorithm and its validation. Journal of Space Weather and Space Climate 15, 47. https://doi.org/10.1051/swsc/2025042

Kerr, Font, Waldteufel, Camps, Bara, Corbella, Torres, Duffo, Vallossera, 2000. New radiometers SMOS a dual pol L-band 2D apertura synthesis radiometer, in: Proceedings (Cat. No.00TH8484). Presented at the 2000 IEEE Aerospace Conference, IEEE, Big Sky, MT, USA. https://doi.org/10.1109/AERO.2000.878481

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

Khazaal, A., Cabot, F., Anterrieu, E., Kerr, Y.H., 2020. A New Direct Sun Correction Algorithm for the Soil Moisture and Ocean Salinity Space Mission. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13. https://doi.org/10.1109/JSTARS.2020.2971063

Khazaal, A., Richaume, P., Cabot, F., Anterrieu, E., Mialon, A., Kerr, Y.H., 2019. Improving the Spatial Bias Correction Algorithm in SMOS Image Reconstruction Processor: Validation of Soil Moisture Retrievals With *In Situ* Data. IEEE Transactions on. Geoscience and. Remote Sensing 57, 1. https://doi.org/10.1109/TGRS.2018.2853619

Wu, L., Torres, F., Corbella, I., Duffo, N., Duran, I., Vall-llossera, M., Camps, A., Delwart, S., Martin-Neira, M., 2013. Radiometric Performance of SMOS Full Polarimetric Imaging. IEEE Geoscience and Remote Sensing 10, 6. https://doi.org/10.1109/LGRS.2013.2260128

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

2. The SMOS Solar Flux

2.2 PROCESSING PIPELINE

The SMOS orbital solar flux is generated out of the L1B products by:

1. Date-based data selection.
2. Data filtering to remove low-quality data points, due to radio frequency interference (RFI), Sun-eclipse contaminated points, Sun elevation angle position, and possible additional issues leading to failed Sun alias reconstruction.
3. Sun BT values are corrected for the obliquity factor and residual antenna plane biases.
4. Horizontal and vertical components of the Sun BTs are resampled to the same timestamps and averaged to further improve data quality.
5. Using moving averages and the standard deviation, outliers are removed (3σ filter). Orbital medians are then evaluated.
6. Sun BT values are turned into solar flux estimates in solar flux units (sfu) and adjusted to a constant Sun-Earth distance set to 1 AU.
7. Finally, the orbital solar flux product is generated as the solar flux and additional orbital information are stored in the output files.

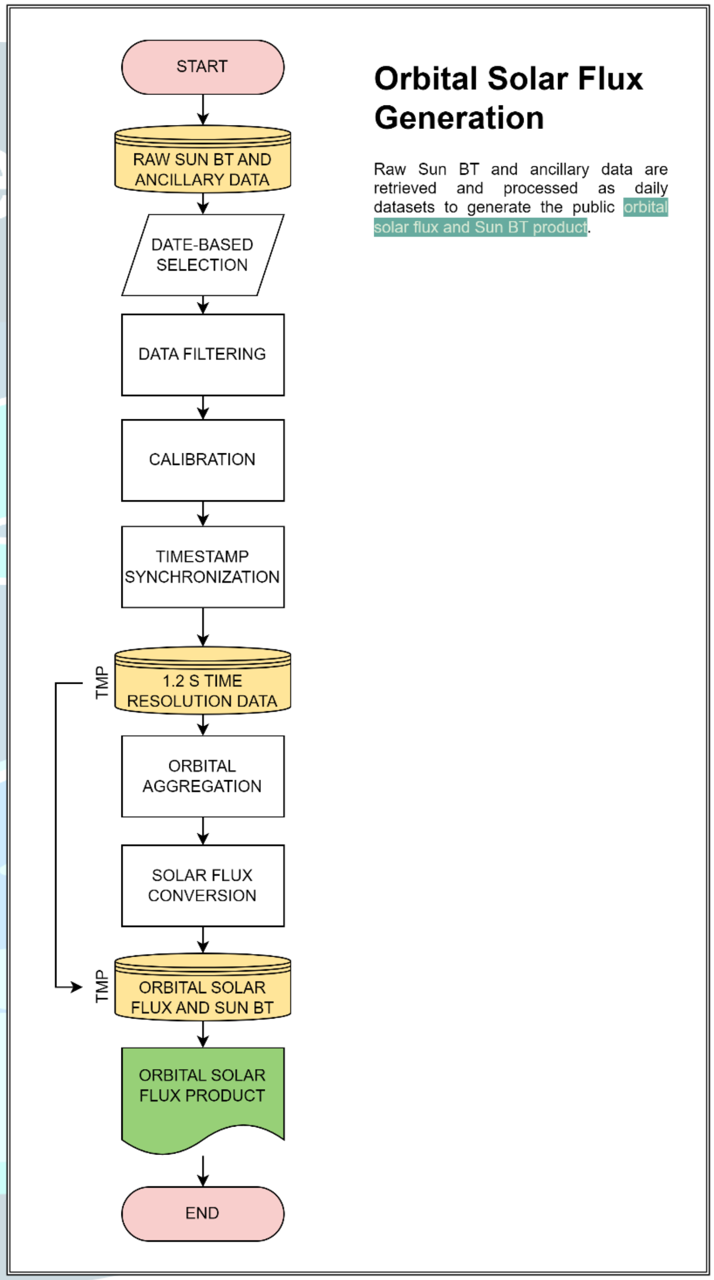

Figure 6: Orbital solar flux generation pipeline.

Note that the Sun elevation angle is a key quality indicator of the SMOS acquisitions and is defined as the elevation of the Sun alias in the antenna plane. In the front-lobe, high angles are associated with high data quality of the raw L1B Sun BT, since the signal enters

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

the lobe with less slant. On the other hand, low values (higher inclination) require more incisive calibration. In the currently operative L1 processor version 724, data acquisitions obtained at angles below 0.032 rad need to be discarded due to a bug. This constraint affects every day with a highly variable impact, going from a couple of minutes per orbit in most months, to severe degradation during the central part of April, and the end of August. This lower threshold will be removed in the subsequent L1OP version 781, which is expected to become operative by the end of 2026.

On the other hand, when the Sun alias appears in the antenna's back-lobe, a lower Sun elevation angle means that the Sun is closer to the main lobe and thus results in higher data quality compared to greater angle positions. Data collected in the back-lobe at Sun elevation angles above 0.2 rad are discarded due to the too low signal-to-noise ratio. This constraint is more relevant during the winter months, since the Sun only appears in those positions during some months of the year.

Data calibration is performed by the evaluation of linear correction factors for distinct bins of the antenna plane and the different polarizations and lobes, defined by comparison with a dedicated ground reference in the period from 2010/06/01 to 2021/12/31. This ground reference relies on input data collected in the 1 – 3.7 GHz frequency range, intercalibrated and resampled to match the SMOS frequency and time resolution. These coefficients are stored in a Look Up Table (LUT), which has been validated and is applied to the operative dataset. For additional details, see the references below.

2.3 PRODUCT FORMAT

Data are delivered in `ASCII` format, as `txt` files of 7 – 9 kB each on average. These are named with the following convention:

```
SM_ssss_MIR_SUNFLX_yyyyMMddTHHmmss_yyyyMMddTHHmmss_xxx_zzz_1
```

With:

- `ssss`: string set to either REPR or OPER, depending on whether an older version of the dataset was already delivered using a previous version of the algorithm or not.
- `xxx`: baseline version with padding leading zeros, changing with L1OP and calibration LUT versions.
- `zzz`: file version with padding leading zeros.

The first date corresponds to the validity start of the first L1B file used for the generation, and the last date is the validity stop time of the last L1B file featured in the UTC time zone.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

Each file includes a header:

Data Center	Product Generation Time (UTC)	Processor Version	N_Front	N_Back	N_L1B
RedLab	2025-12-19T11:49:21	724_2.0	15	15	29

- `Data Center`: generator, always set as RedLab.
- `Product Generation Time (UTC)`: time of the output file creation.
- `Processor Version`: L1P version and LUT version separated by underscore.
- `N_Front`: number of front-lobe orbital estimates.
- `N_Back`: number of back-lobe orbital estimates.
- `N_L1B`: number of input L1B files.

Note that the expected number of L1B files is in the 27-29 range. Different values may be present occasionally, depending on many different factors, and are investigated daily to check for the integrity of the files. Additionally, the sum of the front and back-lobe estimates is bound to be close to the number of L1B files inputted, but is not necessarily the same, depending on how the orbital coverage is distributed across each product, and how the semi-orbit start time compares to the date start.

Then, data points are listed in tabular form with an initial header row.

- `Median Time (UTC)`: median estimation time of each half-orbital estimate according to ISO 8601 format (i.e.,: `yyyy-MM-ddTHH:mm:ss.SS`).
- `Sun BT (K)`: orbital Sun BT provided as absolute measurements.
- `Solar Flux 1AU (sfu)`: corresponding solar flux adjusted to a constant Sun-Earth distance.
- `N_Snapshots`: number of valid snapshots used for each estimate evaluation.
- `Sun BT STD (K)`: standard deviation of the orbital Sun BT estimation.
- `Mean Sun Elevation (rad)`: average Sun elevation of the valid Sun BT data points.
- `Sun-Earth Distance Factor`: correction factor used to normalize the absolute flux.
- `SMOS Orbit Number`: unique ID associated with SMOS orbits.
- `ANX Start Time (UTC)`: orbit's ascending node start time in the ISO 8601 format.
- `Start Time (UTC)`: validity start time in the ISO 8601 format.
- `Stop Time (UTC)`: validity stop time in the ISO 8601 format.
- `Antenna Plane`: antenna plane flag. Set to 1 for front-lobe and to 0 for back-lobe data points.

Data points are listed by antenna plane (front first) and sorted in ascending order by time. Since data are filtered at different stages of the processing pipeline, the actual delta between

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

each data point is not exactly 50 min (i.e., half orbital period). Resampling is thus necessary to achieve homogeneous spacing. For data from L1P v724, the baseline string is set to `101`. Data from the L1P v783 will be called `2xx`, with the last characters representing the LUT version, and starting from `00`.

Below is an example of a simple MATLAB® script to parse orbital solar flux files.

```
function [data,header] = smos_flux_parser(file_name)
%% ======================================================================
% ingest the SMOS orbital solar flux product and return the orbital
% observations and header information.
%
% argument(s):
% - file_name: char array or string with full path to the input file
%
% output(s):
% - data     : table containing the content of each orbital estimation
% - header   : structure containing the header of the file (values as char)
%
% CC: FGuarnaccia, 2026.
%% ======================================================================
% format specifications:
hdr_fmt = '%-12c%-30c%-20c%-10c%-10c%-10c';
data_fmt = '%-23c%-15c%-20c%-13c%-17c%-25c%-26c%-20c%-23c%-23c%-23c%-13c';
% open file_name
f_id = fopen(file_name,"r");
try
    % read header
    C = textscan(f_id,hdr_fmt,2);
    header = struct;
    % for each field:
    for i_f = 1:numel(C)
        name = replace(strtrim(C{i_f}(1,:)),["(",")"," "],["","","_"]);
        header.(name) = strtrim(C{i_f}(2,:));
    end
    % skip blank line
    fgetl(f_id);
    % read data
    cols = textscan(f_id,data_fmt);
    data = table;
    for i_c = 1:numel(cols)
        name = replace(strtrim(cols{i_c}(1,:)),["(",")"," "],["","","_"]);
        data.(name) = strtrim(cols{i_c}(2:end,:));
    end
    fclose(f_id);
catch
    % enforce closing
    warning("ERROR: data parser failed")
    fclose(f_id);
end
end
```


Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

2.4 ADDITIONAL NOTES

The SMOS orbital solar flux dataset provides a highly stable dataset, effectively filtering out outliers and reducing orbital effects. However:

- The data quality of the SMOS back-lobe dataset is lower than that of the front-lobe dataset. The coefficient of variation[2] in the front-lobe mostly falls in the 0.05-0.10 range, while values in the back-lobe are doubled. As such, usage of the back-lobe dataset is not advised.
- Solar radio bursts' influence is mostly removed from the products unless in the presence of strong solar storms.

The suggested approach for daily data aggregation is to evaluate the median of each orbital flux output. Precision of the proposed method is estimated to be above 95%, comparing favorably to the Nobeyama (NoRP), Radio Solar Telescope Network (RSTN), and the Penticton (F10.7) radio-telescopes.

List of References

Achard, F, D'Souza, G., 1994. Collection and pre-processing of NOAA-AVHRR 1 km resolution data for tropical forest resource assessment. In: Technical Report EUR 16055 EN, European Commission, ECSC-EC-EAEC, Brussels. https://op.europa.eu/en/publication-detail/-/publication/15b97b3e-cd63-4370-9fbd-bab130 5a9eb1

Crapolicchio, R., Forte, R., Guarnaccia, F., Di Ciolo, L., 2025. The L-Band solar flux retrieval from the Soil Moisture and Ocean Salinity (SMOS) mission: Theoretical algorithm and its validation. Journal of Space Weather and Space Climate 15, 47. https://doi.org/10.1051/swsc/2025042

Crapolicchio, R., Guarnaccia, F. and Forte, R., 2025 The XXV Solar Cycle Progression Seen at L-band by the SMOS Mission. European Space Weather Week 2025 (ESWW25), Zenodo. https://doi.org/10.5281/zenodo.17960474

Shimojo, M., Iwai, K., Asai, A., Nozawa, S., Minamidani, T., Saito, M., 2017. Variation of the Solar Microwave Spectrum in the Last Half Century. The Astrophysical Journal 848, 62. https://doi.org/10.3847/1538-4357/aa8c75

[2] Orbital standard deviation normalized by the orbital average.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

3. The SMOS Solar Radio Burst Database

3.1 PROCESSING PIPELINE

Solar radio bursts detection is carried out in two different stages. First, the preliminary data assignment is done through an automatic screening procedure. Then, with a monthly cadence, data entries are further checked, and flags are revised. Every day:

1. Date-based selection[3].
2. Input data are processed similarly to the orbital solar flux pipeline, except for the filtering step, where the RFI filter is initially omitted. The orbital solar flux procedure is followed through to orbital aggregation, and the corresponding high-time-resolution solar flux is thus evaluated.

At the moment, steps 3 to 9 are only performed over data acquired when the Sun appears in the antenna's front-lobe. As such, since data in the back-lobe are not automatically scanned for solar radio bursts, the availability of the algorithm is 50%.

3. Local fluctuations are removed through a first-order median filter.
4. Consecutive periods of high flux are identified through the comparison with a daily flux reference.

If any event lasts at least 45 seconds:

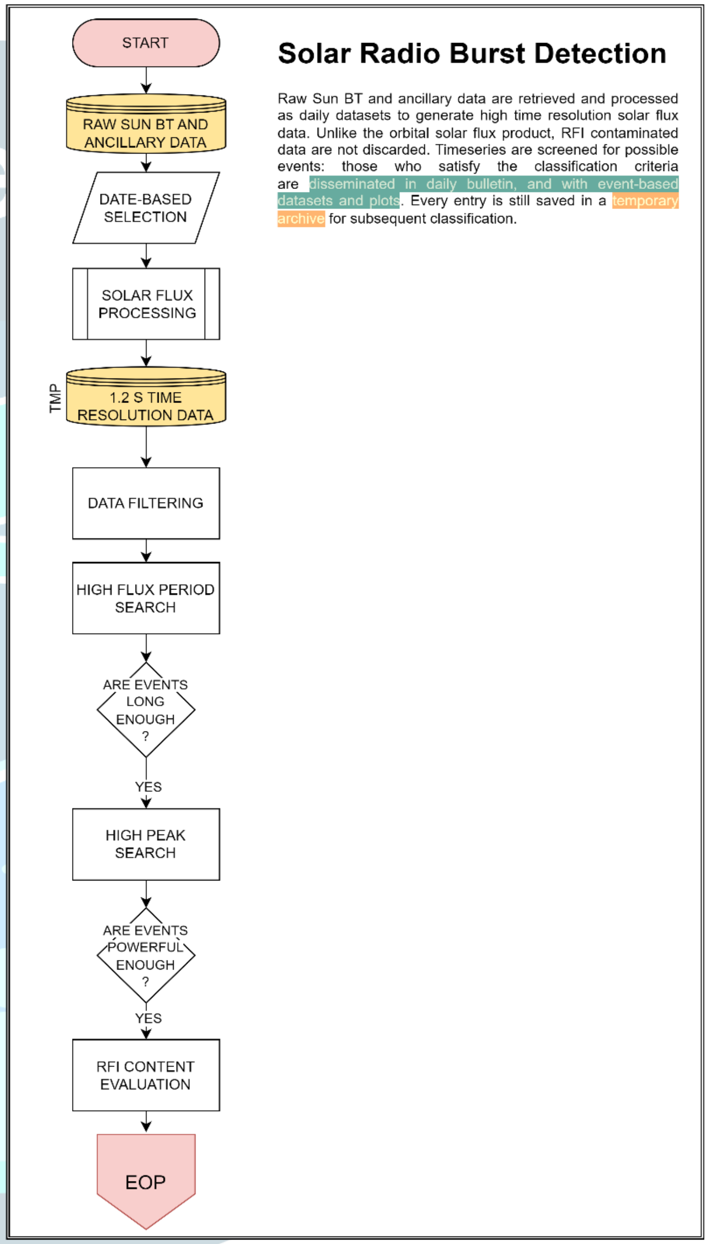

Figure 7: Solar radio burst screening pipeline (pt.1).

[3] Alternative solutions will be tested in the future to push SMOS solar radio burst detection into near-real time, by increasing the number of daily processing and replacing calendar date selection with a sliding time coverage of fixed width.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

5. Peak prominence is evaluated on a statistical basis relying on quantile representation.

If any event found has a peak prominence of at least 15 sfu:

6. The content of RFI contamination is evaluated for each entry.

If any entry passes this threshold, it is stored in a private database alongside a provisional flag. Flag assignment is carried out by:

7. Extraction of the main parameters for each event. These are transformed through standard scaling and principal component analysis.
8. Data classification, performed with a random forest algorithm.

If at least one event is classified as a burst:

9. The solar radio burst bulletin is generated alongside the plot and high time resolution binary dataset for each event.

Note that the algorithm automatically classifies entries with a duration of 10 minutes or higher as solar radio bursts.

The events found by the daily processing pipeline are stored locally on a private repository, regardless of the original flag assignment. These events are later reviewed once per month with a pre-defined routine.

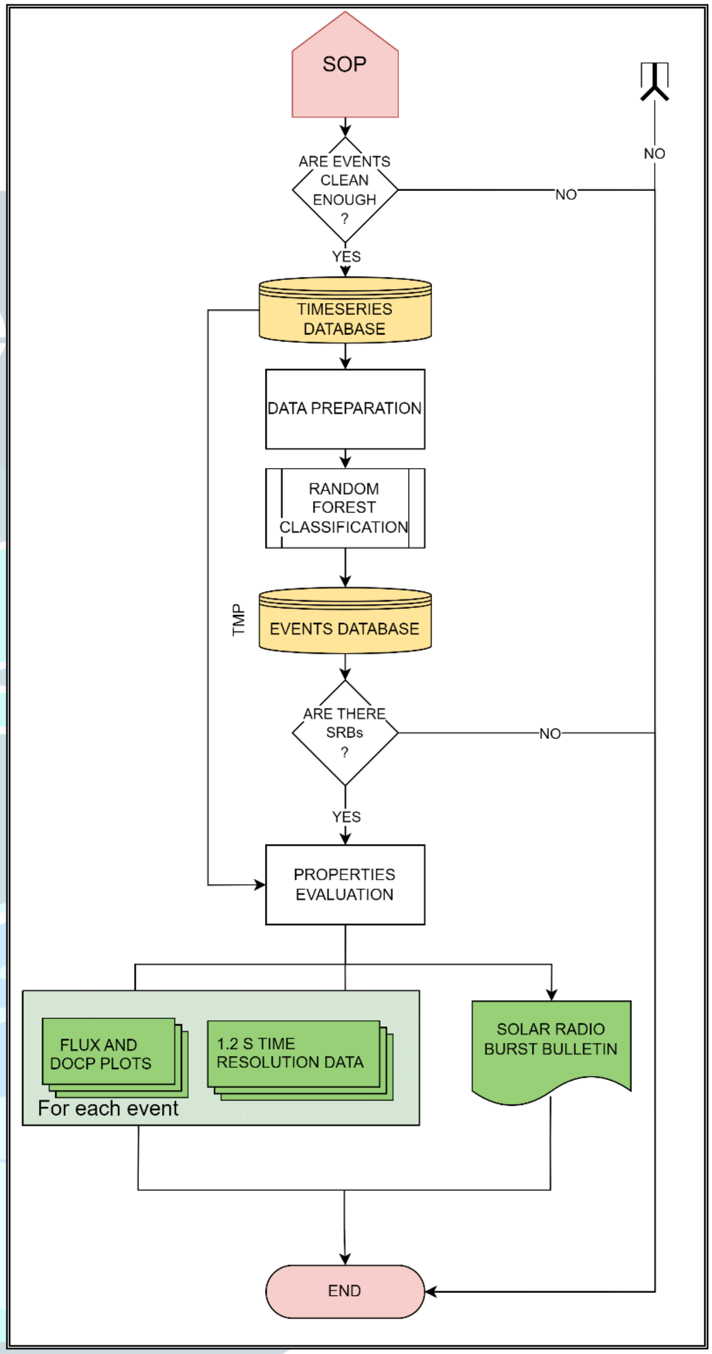

Figure 8: Solar radio burst screening pipeline (pt.2).

The periodic review includes the following steps:

1. Each entry is individually checked by analysis of the 1.2s resolution dataset, comparison with NOAA's solar activity daily bulletin, with the RSTN or the NoRP 1s time resolution data, or any combination of these datasets.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

2. An event ID is given to each of the N unique monthly SRBs events starting from 1 to N. Multiple entries may be associated with the same event ID.
3. The monthly database is disseminated, adding to the shared folder also the missing events files associated with false negatives of the daily procedure. The monthly overview also includes false positives, if present, with a dedicated flag.

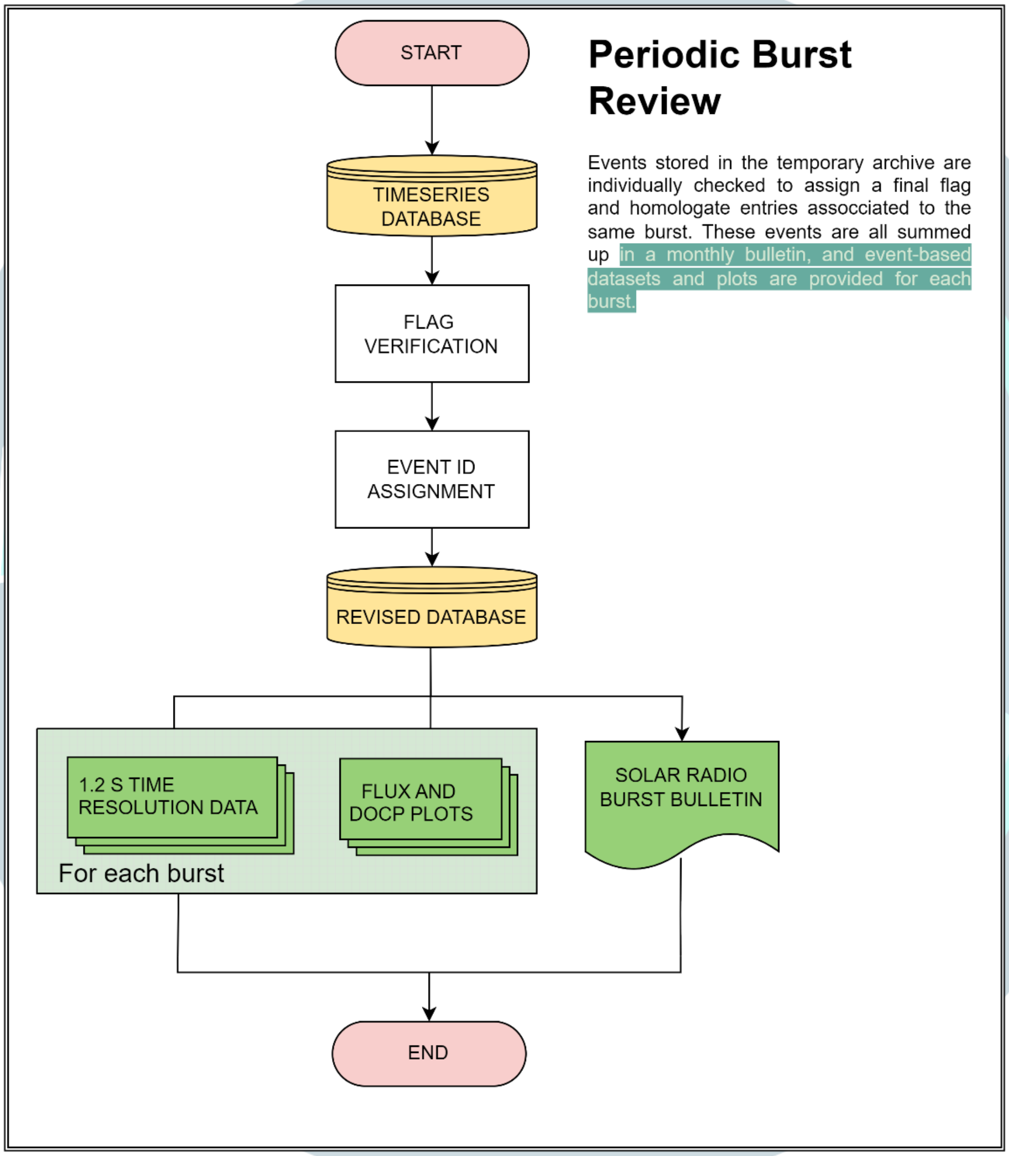

Figure 9: Solar radio bursts periodic review pipeline.

The monthly overview of the SMOS solar radio burst detections is scheduled to take place in the first half of the following calendar month.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

3.2 PRODUCT FORMAT

The SMOS solar radio bursts detection product is disseminated at two different stages, as reflected in the folder structure:

- `Latest_events`: operative products generated daily by the automatic flagging procedure. These include daily bulletins (`ASCII` format) and individual events representation (`netCDF` and `png`). Entries will thus include true solar radio bursts as well as potential false positives. As the estimated algorithm precision is >99%, very few false positive detections are expected.
- `archive`: reprocessed products generated monthly through the review procedure illustrated. These include monthly summaries (`ASCII` format) and individual events representation (`netCDF` and `png`). Entries will thus include both true solar radio bursts and potential false positives, as found in the `Latest_events` folder, alongside false negatives. As the estimated classification algorithm sensitivity is ~77% with reference to the individual burst's ID[4] false negatives are expected. This sensitivity value still compares favorably to NOAA's ~50%, evaluated in the same reference period at times when the Sun alias appears in the MIRAS front-lobe. Overall, as the screening algorithm (daily processing pipeline step 1 to 6) is ~94%, some events may still be missing from the final repository.

The content of the `Latest_events` folder is meant to be temporary: daily bulletins are meant to be replaced by the more complete monthly files, and the events data are moved into the `archive` folder alongside the additional files covering false negatives of the screening procedures.

The `netCDF` and `png` products are named with the following convention:

`SM_ssss_MIR_SUNRBR_yyyyMMddTHHmmss_yyyyMMddTHHmmss_Enn_xxx_zzz_1`

With:

- `ssss`: string set to either REPR or OPER, depending on whether an older version of the dataset was already delivered using a previous version of the algorithm or not.
- `nn`: event number of the day with padding leading zeros. Events in the OPER folder are chronologically sorted; to keep consistency, events introduced in the monthly review are sorted chronologically and serialized after the last OPER file of the day.

[4] This means that at least one entry associated with each burst ID is properly flagged as a burst. Additional entries providing a more extensive coverage of the event may be initially flagged as negatives.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

- xxx: baseline version with padding leading zeros, changing with L1OP and calibration LUT versions.
- zzz: file version with padding leading zeros.

The first date corresponds to the validity start of the first L1B file used for the generation, and the last date is the validity stop time of the last L1B file featured in the UTC time zone. The current baseline version for data from L1P v724 using the LUT v2.0 is 102.

Both the png and netCDF files cover the duration of each event, plus 1 h before the start and 1 h after the end of each entry. These are two-panel figures with the first one providing information about the solar flux in sfu and normalized to a constant Sun-Earth distance of 1 AU.

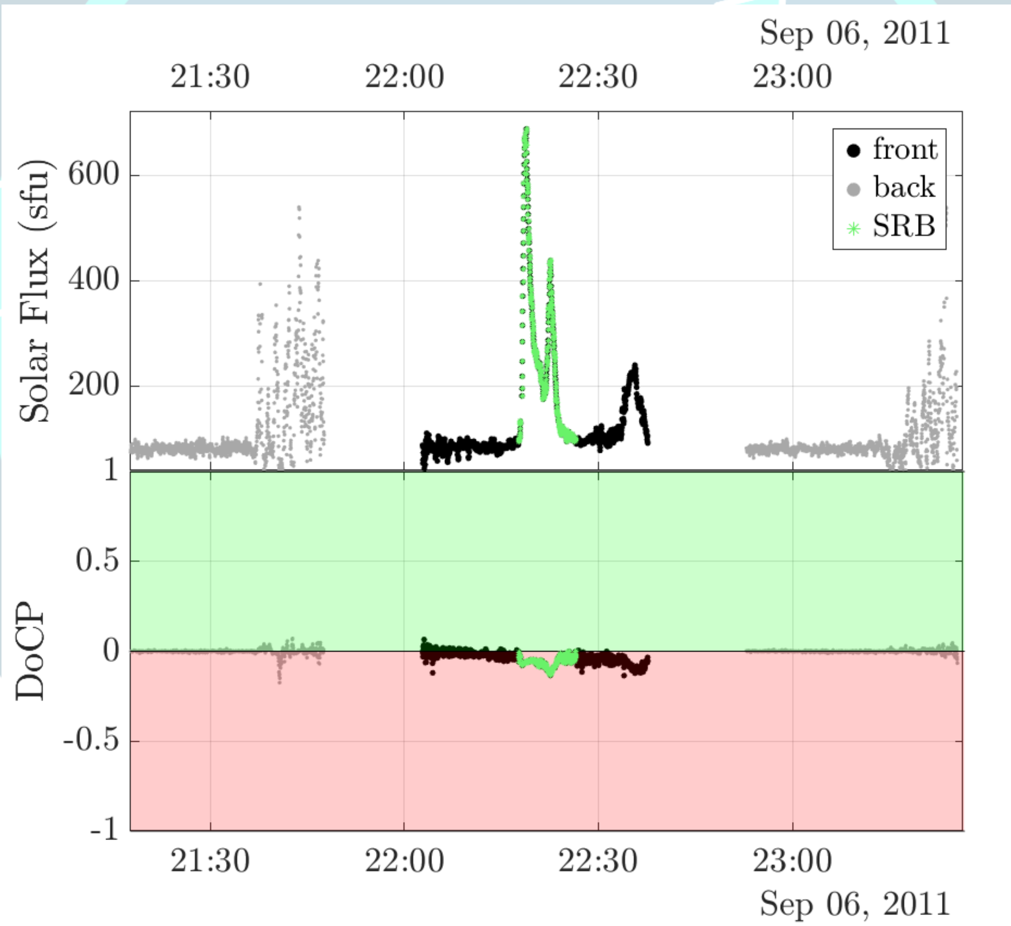

Figure 10: Example of png output automatically generated. Red shading is used for RHCP and green for LHCP.

Back-lobe data are reported in grey, however this are mainly meant for context overview as the automatic procedure only runs on the front-lobe dataset. The DoCP is shown in the

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

second panel; note that out-of-scale values ($|DoCP| > 1$) may still be present due to RFI contamination. Additionally, the DoCP in the back-lobe is not always accurate in L1OP v724, as information about the polarization status is often drowned by noise.

The `netCDF` files are written in the `netCDF4` standard, including the variables:

- Time: string in `yyyy-MM-ddTHH:mm:ss.SS` format (UTC).
- Flux: double in solar flux units adjusted to a constant Sun-Earth distance of 1AU.
- DoCP: double, defined as negative for RHCP and positive for LHCP.
- Front: string set to "F" for front-lobe acquisitions, and "B" for back-lobe data.

The size of each figure falls in the 40 – 100 kB range on average, while the size of the binary data is usually in the 200 – 900 kB range, with fluctuations depending on data size.

The two types of bulletins are named after the extremes of the first and last input. Namely, the daily bulletins follow the structure:

`SM_ssss_MIR_SUNRBR_yyyyMMddTHHmmss_yyyyMMddTHHmmss_xxx_zzz_1`

With:

- `ssss`: string set to either REPR or OPER, depending on whether an older version of the dataset was already delivered using a previous version of the algorithm or not.
- `xxx`: baseline version with padding leading zeros, changing with L1OP and calibration LUT versions.
- `zzz`: file version with padding leading zeros.

The first date corresponds to the validity start of the first L1B file used for the generation, and the last date is the validity stop time of the last L1B file featured in the UTC time zone.

While the monthly bulletins are named:

`SM_ssss_MIR_SUNRBR_yyyyMMdd_yyyyMMdd_xxx_zzz_1`

With:

- `ssss`: string set to either REPR or OPER, depending on whether an older version of the dataset was already delivered using a previous version of the algorithm or not.
- `xxx`: baseline version with padding leading zeros, changing with L1OP and calibration LUT versions.
- `zzz`: file version with padding leading zeros.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

The first date corresponds to the first and last day of the calendar month included in the file. The release of the monthly bulletins is meant to replace the previous daily bulletins. File size fluctuates from 1 to 11 kB for monthly files, and 0.5 – 2 kB for daily files, depending on the number of entries.

Files share the header structure, which is reported below.

Data Center	Product Generation Time (UTC)	Processor Version	N_Entries
RedLab	2026-01-14T10:17:37	724_2.0	7

- `Data Center`: generator, always set as RedLab.
- `Product Generation Time (UTC)`: time of the output file creation.
- `Processor Version`: L1OP version and LUT version separated by underscore.
- `N_Events`: number of events found.

The body of text is provided in tabular form with one row for each detection and a header. Entries associated with false positives of the operative pipeline do not feature data for most fields, unless the `Entry Number`, `Start` and `Stop Time`, and `Event Type`.

- `Entry ID`: event serial number of the month with zero padding. This may be associated with multiple entries. If the entry is noise, it corresponds to "---" [Monthly only]
- `Entry Number`: entry serial number of the day with zero padding. These are unique IDs in the daily bulletin file, but not in the monthly bulletin[5].
- `Start Time (UTC)`: automatic entry start time according to ISO 8601 format (i.e.,: `yyyy-MM-ddTHH:mm:ss.SS`).
- `Stop Time (UTC)`: automatic entry stop time according to ISO 8601 format (i.e.,: `yyyy-MM-ddTHH:mm:ss.SS`).
- `Peak Flux Time (UTC)`: timestamp associated with the peak solar flux according to ISO 8601 format (i.e.,: `yyyy-MM-ddTHH:mm:ss.SS`).
- `Peak DoCP Time(UTC)`: timestamp associated with the peak DoCP according to ISO 8601 format (i.e.,: `yyyy-MM-ddTHH:mm:ss.SS`), set to a dash (–) if captured during dual mode.
- `Peak Flux 1AU(sfu)`: corresponding to the peak solar flux adjusted to a constant Sun-Earth distance of 1 AU.

[5] It is possible that one event may be split into multiple entries taking place on two different days (end/start of the day). In that case, there is the possibility of more than one event sharing both the event and entry number, since the sequencing of the latter resets every day. In that case, plot and binary data files identification needs to account for the event start and end datetime strings to provide a unique match.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

- `Peak DoCP`: corresponding to the peak DoCP set to a dash (–) if captured during dual mode.
- `Energy (10-19erg*cm-2*Hz-1)`: approximate energy associated with the detection, defined as the time integral of the normalized solar flux timeseries.
- `Entry Type`: entry type identifier, set to 1 for solar radio bursts and 0 for noise.

Note that both bulletins are generated only if at least one event is flagged as positive in the operative pipeline, one solar radio burst is identified from the archive, or both. The size of the monthly bulletins is usually in the 0.5 – 6 kB range, and 0.5 – 2 kB range for the daily files.

Find below an example of a simple script to parse the daily and monthly solar radio burst bulletins (ASCII files) in MATLAB®.

```
function [data,header] = smos_burst_parser(file_name)
%% ========================================================================
% ingest the SMOS solar radio burst bulletin product and return the entry
% catalogue and header information for both the monthly and daily files.
%
% argument(s):
% - file_name: char array or string with full path to the input file
%
% output(s):
% - data     : table containing the content of each entry
% - header   : structure containing the header of the file (values as char)
%
% CC: FGuarnaccia, 2026.
%% ========================================================================
% format specifications:
hdr_fmt = '%-12c%-30c%-20c%-10c';
% define file type enforcing file_name as char:
file_name = sprintf('%s',file_name);
[~,base] = fileparts(file_name);
switch length(base)
    % monthly bulletin: SM_ssss_MIR_SUNRBR_yyyyMMdd_yyyyMMdd_xxx_zzz_1
    case 46
        data_fmt = '%-10c%-20c%-23c%-23c%-23c%-23c%-20c%-20c%-30c%-10c';
    % daily bulletin: SM_ssss_MIR_SUNRBR_yyyyMMddTHHmmss_yyyyMMddTHHmmss_xxx_zzz_1
    case 60
        data_fmt = '%-20c%-23c%-23c%-23c%-23c%-20c%-20c%-30c'; % 30
    otherwise
        msgbox('ERROR: do not change the file name to use this parser!')
        error("Invalid file name")
end
% open file_name
f_id = fopen(file_name,"r");
try
    % read header
```

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

```
    C = textscan(f_id,hdr_fmt,2);
    header = struct;
    % for each field:
    for i_f = 1:numel(C)
        name = replace(strtrim(C{i_f}(1,:)),["(",")"," "],["","","_"]);
        header.(name) = strtrim(C{i_f}(2,:));
    end
    % read data
    cols = textscan(f_id,data_fmt);
    data = table;
    for i_c = 1:numel(cols)
        name = replace(strtrim(cols{i_c}(1,:)),["(",")"," "],["","","_"]);
        data.(name) = strtrim(cols{i_c}(2:end,:));
    end
    fclose(f_id);
catch
    % enforce closing
    warning("ERROR: data parser failed")
    fclose(f_id);
end
end
```

List of References

Flores-Soriano, M., Cid, C., and Crapolicchio, R., 2021. Validation of the SMOS mission for space weather operations: The potential of near real-time solar observation at 1.4 GHz. Space Weather 19, 3. https://doi.org/10.1029/2020SW002649

Guarnaccia, F., Forte, R. and Crapolicchio, R., 2025. The XXV Sun Cycle Solar Radio Burst Database by the SMOS mission. European Space Weather Week 2025 (ESWW25), Zenodo https://doi.org/10.5281/zenodo.17960494

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

4. THE SUN BT ADF

4.1 PROCESSING PIPELINE

The SMOS Sun BT ADF is daily updated through the following steps:

1. The orbital Sun BT ASCII file covering the previous day is retrieved and ingested. If this file was not generated, the routine skips to step 4.
2. The daily Sun BT is generated by aggregation of the orbital front-lobe dataset. The median operator is used to filter out the influence of strong solar radio storms.
3. The Sun BT is rescaled with the agreed scaling factor. For v724, data needs to be multiplied by a 0.9 factor.
4. The last valid Sun BT ADF is retrieved and ingested.
5. Data are merged. If the daily file was successfully retrieved, the last seven data points of the ADF (i.e., the forecast window) are trimmed, and the daily measure is appended. Otherwise, after the forecast window is trimmed, the last valid measurement of the ADF is propagated up until the expected end, that is the previous day date.
6. Forecast is applied to extend the dataset to a forecast window of seven days. Current forecast is performed through naïve forecast (i.e., “last valid measure”).
7. The header and the data block files are generated and compressed in the agreed format.
8. The output of the routine is disseminated by uploading through SFTP to the DPGS server.

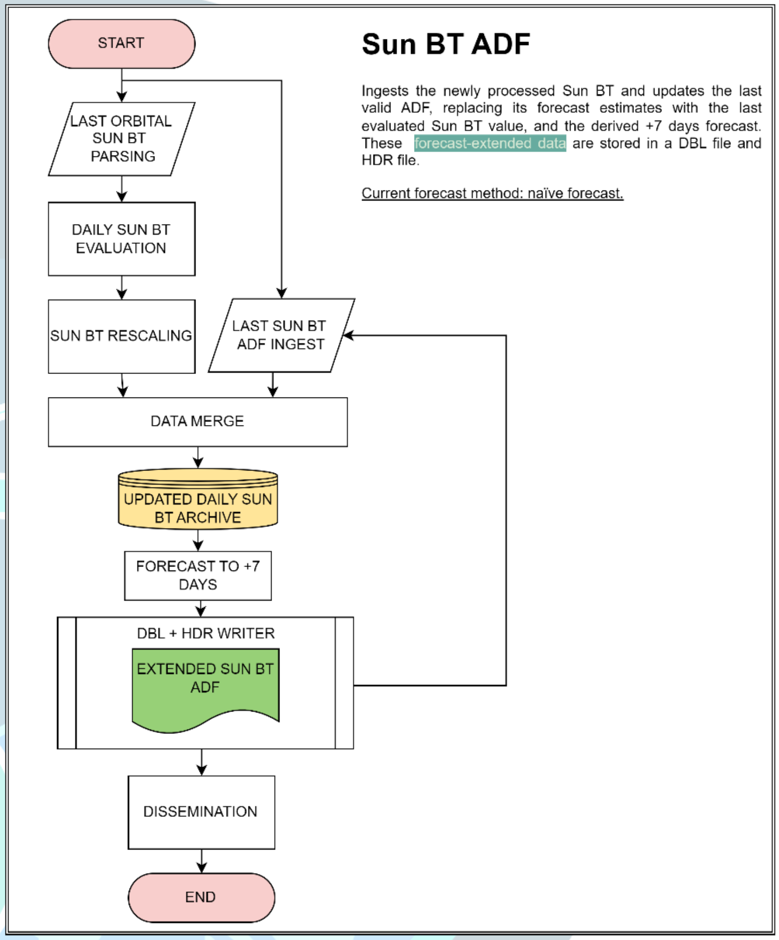

Figure 11: Sun BT Auxiliary Data File generation pipeline.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

4.2 PRODUCT FORMAT

The Sun BT ADFs are named:

`SM_ssss_AUX_SUN_BT_20100101T000000_yyyyMMddT000000_xxx_zzz_7`

With:

- `ssss`: string set OPER for daily generated files, TEST/REPR substrings may be feature in other files generated ad hoc depending on whether an older version of the dataset was already delivered using a previous L1OP version or not, or the file was generated with a modified pipeline, or both.
- `xxx`: baseline version with padding leading zeros, changing with L1OP and calibration LUT versions.
- `zzz`: file version with padding leading zeros.

The second date string corresponds to the last covered day, that is the final measurement day plus the forecast window of seven days. Under nominal operation, this is the current date plus six days.

The header and data block files share the same base name and are stored inside a `TAR` archive compressed through `Gzip` into a `tgz` file. Every new file that is distributed to the DPGS will have a unique name, by updating progressively the file version if a new reprocessing is needed.

The data block file is a `DBL` binary file including the following items:

- The number of dataset records, that is the number of days elapsed from the validity start and stop date.
- Date UTC in `yyyyMMdd` format stored as unsigned 32 bits integers.
- Time UTC in `HHmmss` format stored as unsigned 32 bits integers and set to a default value of the start of the day (i.e., 000000).
- Sun BT in Kelvin stored as a float. This is the rescaled absolute daily Sun BT.

The total number of items is equal to the number of dataset records multiplied by three (Date, Time, Sun BT triplets) plus one (the first item, i.e., the number of records). Information about date, time, and Sun BT is stored by cycling between each of the three variables for each date before moving to the next date.

Note that as the validity start date is the 1st of January 2010 while the first available measurement is from the 12th of January, the first eleven days are filled with the Sun BT value evaluated during the latter date.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

The header file comes with the `HDR` extension and is an `XML` file with the following structure:

```
Earth_Explorer_Header
├── Fixed_Header
│     ├── File_Name
│     ├── File_Description
│     ├── Notes
│     ├── Mission
│     ├── File_Class
│     ├── File_Type
│     ├── Validity_Period
│     │     ├── Validity_Start
│     │     └── Validity_Stop
│     ├── File_Version
│     └── Source
│           ├── System
│           ├── Creator
│           ├── Creator_Version
│           └── Creation_Date
└── Variable_Header
      └── Specific_Product_Header
            ├── Main_SPH
            │     ├── SPH_Descriptor
            │     ├── Ref_Doc
            │     ├── Precise_Validity_Start
            │     ├── Precise_Validity_Stop
            │     ├── Checksum
            │     ├── Header_Schema
            │     ├── Datablock_Schema
            │     ├── Header_Size
            │     ├── Datablock_Size
            │     └── HW_Identifier
            └── List_of_Data_Sets
                  ├── Data_Set
                  │     ├── DS_Name
                  │     ├── DS_Type
                  │     ├── DS_Size
                  │     ├── DS_Offset
                  │     ├── Ref_Filename
                  │     ├── Num_DSR
                  │     ├── DSR_Size
```


Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org

```
│   └── Byte_Order
└── Data_Set
    ├── DS_Name
    ├── DS_Type
    ├── DS_Size
    ├── DS_Offset
    ├── Ref_Filename
    ├── Num_DSR
    ├── DSR_Size
    └── Byte_Order
```

The header is encoded with UTF-8 using the `HDR_SM_XXXX_AUX_SUN_BT_0400.xsd` scheme.

List of References

Guarnaccia, F., Forte, R. , and Crapolicchio, R., 2026. SMOS AUX Sun BT Product TGRD. RedLab-SER-SM-REP-008.

Contacts:
federica.guarnaccia1@serco.com
roberta.forte@ext.esa.int
raffaele.crapolicchio@ext.esa.int
redlab-oper@eo-sppa.org