NOTES ON THE USE OF SMAP

Table of contents

1. Introduction

SMAP allows processing Sentinel-3 L1A files using the back-projection fully focused SAR (FFSAR) processing approaches.

In this documentation file, a general explanation to introduce the notions of FFSAR geometry, on orbit and on ground. A dedicated section explains then how to configure the processing options available in the SMAP processor. Finally, in the last section, some figures of the output SMAP variables from the provided test data set (TDS) are reported for illustration and validation.

2. Fully-Focussed SAR processing

By coherently focusing on subsequent along-track surface points, separated by a selectable distance, the Fully-Focused SAR (FFSAR) processing allows reaching an extremely high resolution not achievable with the Unfocused SAR (UFSAR) processing.

The FFSAR processing operates over a **"data block",** composed of a number of bursts, that is shifted along the bursts of the input file to process the data sequentially as shown in Figure 1. The size (in number of bursts) of the data block defines the **illumination time**, i.e. the extension of the synthetic aperture of the FF-SAR processing.

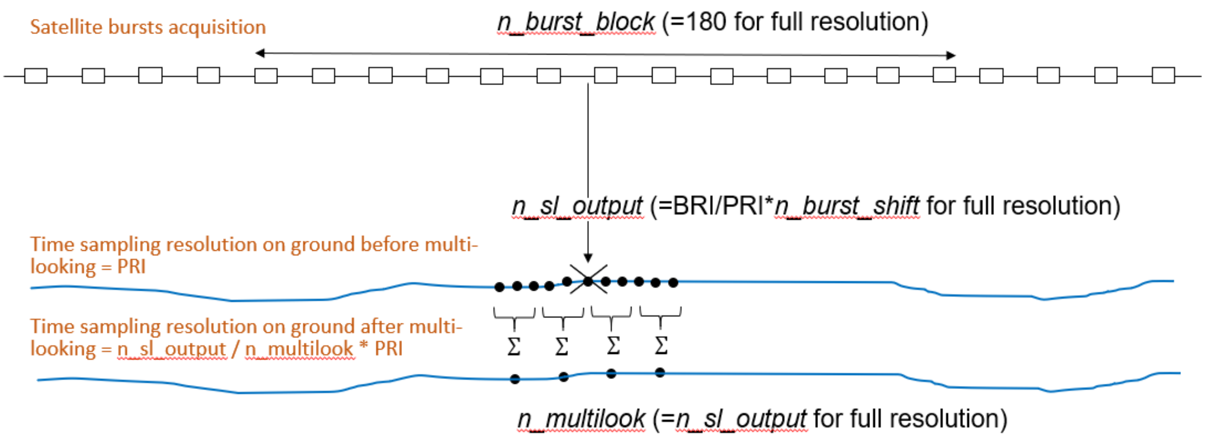


FIGURE 1. GENERATION OF MULTI-LOOKS WITH FFSAR FROM ONE DATA BLOCK.

During the focusing process, each data block leads to the generation of a **single-look** on ground that is aligned with the center of the data block. Subsequent single looks are equally spaced in time by the Pulse Repetition Interval (PRI).

The single-looks can be spatially averaged in power (multi-looked) to obtain **multi-looks**:

$$multilook = \frac{1}{n_{multilook}} \sum_{n_{multilook}} |singlelook|^2.$$

The **sampling frequency** (also called posting rate in the configuration file) defines the number of single-looks that are multi-looked in the multi-looking step as

$$n_{mutlilook} = \frac{1}{PRI * posting\_rate} = \frac{PRF}{posting\_rate}$$

Hence the maximum sampling frequency achievable is equal to the PRF (17825Hz for Sentinel-3/SRAL), in case there is no multi-looking. The final **along-track spacing** of output data (multi-looked L1b waveforms), that refers to the distance between two multi-looks, is formulated as:

$$spacing = \frac{v\_proj}{posting\_rate},$$

where $\boldsymbol{v\_proj}$ is the satellite velocity projected on ground, which is around 8441m/s. To calculate the projection on ground, we multiply the satellite velocity on flight by the factor $\alpha = \frac{r_E + altitude}{r_E} \simeq 1.2$, where $r_E = 6371e3m$ is the Earth radius and $\boldsymbol{altitude}$ is the satellite altitude.

To analyse the maximum achievable FF-SAR focusing over a transponder, the full along-track resolution (2.3 s illumination time) and posting rate (= PRF) can be exploited**.** As no multi-looking is performed, the spacing between subsequent L1b single-look waveforms on ground is around 0.5 meters while the along-resolution resolution of the waveforms will be given by:

$$along\_track_{resolution}(single\ look) = \frac{0.886\ *\ \text{wavelength}\ *\ \text{altitude}}{(2\ *\ v\_proj\ *\text{illumination\_time})}$$

Assuming the average altitude equal to 815 km and a maximum illumination time of 2.3 s, the along-track resolution is about 0.5 m. When more than 5 single-looks are multi-looked, the along-track spacing can be approximated with the **along-track resolution**.

For hydrological targets, multi-looking is mandatory to reduce waveform noise. As an example, a **posting_rate of 640 Hz** (8441/640= 13.18 meters separation between two multi-looks) corresponds to the multi-looking of PRF/posting_rate=27 single-looks.

The **magnitude squared coherence (MSC)** gives a number between 0 and 1 characteristic of the coherence of the single-looks into each multi-look. For one multi-look, the MSC is formulated as the square of the sum the single-looks divided by the sum of the square of the single-looks, i.e.

$$msc = \frac{\left|\frac{1}{n_{multilook}} \sum_{n_{multilook}} slc\right|^2}{\frac{1}{n_{multilook}} \sum_{n\_multilook} |slc|^2}.$$

Thanks to the Cauchy-Schwarz inequality the “msc” is bounded by one:

$$msc \leq \frac{\frac{1}{n_{multilook}^2} \left|\sqrt{\sum_{n_{multilook}} slc} * \sqrt{\sum_{n_{multilook}} 1}\right|^2}{\frac{1}{n_{multilook}} \sum_{n_{multilook}} |slc|^2} = 1.$$

Note that MSC is given at level 1b in two-dimension for each multi-look and range gate and at level 2 in one-dimension for each multi-look but at the epoch range gate.

For given Level-1A product, blocks of data composed of maximum 180 bursts (correspondent to an illumination time = 2.3 s, see section 4.5) are processed sequentially until all bursts of the product are processed. Two subsequent blocks are shifted one another by a predefined amount named **shifting**, which is set equal to 4 bursts. The single-looks generated for each data block correspond to the $4 * \frac{BRI}{PRI} = 906$ pulses which compose one radar cycle, in terms of bursts and pulses, for Sentinel-3. The single-looks are centred around the central burst of data block. They do not no overlap between subsequent data blocks and a continuity in ground coverage is guaranteed as shown in Figure 2.

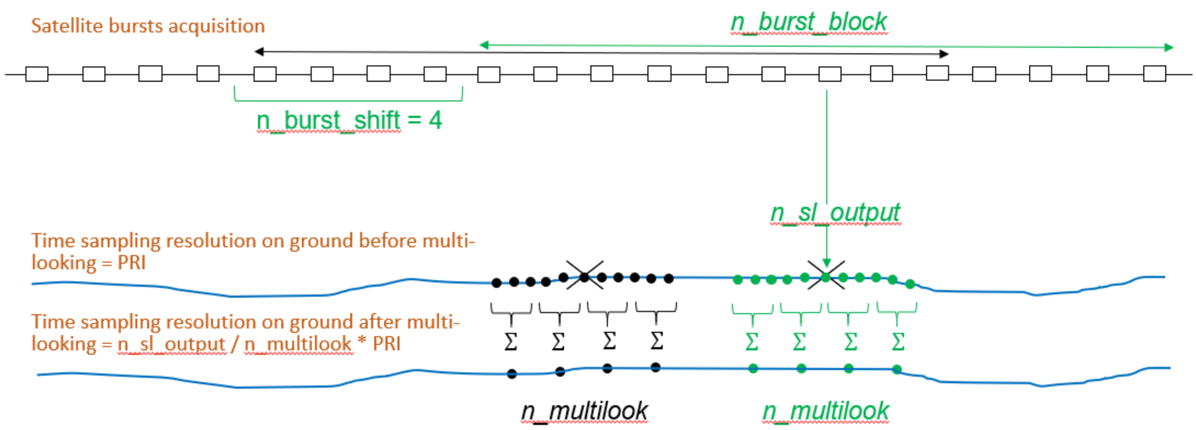


FIGURE 2. DATA BLOCK SHIFT AND GENERATION OF MULTI-LOOKS WITH FF-SAR.

To **prevent aliasing artefacts**, the processing can be done over an **extended receiving window** using the **range_ext_factor** parameter in the configuration file (see section 4.5) at the cost of an increased computation time.

2. SMAP architecture

The architecture of the SMAP code, as shown in Figure 4, can be described through 4 main blocks related to: inputs, processing levels (Level-1b and -2), outputs and documentation.

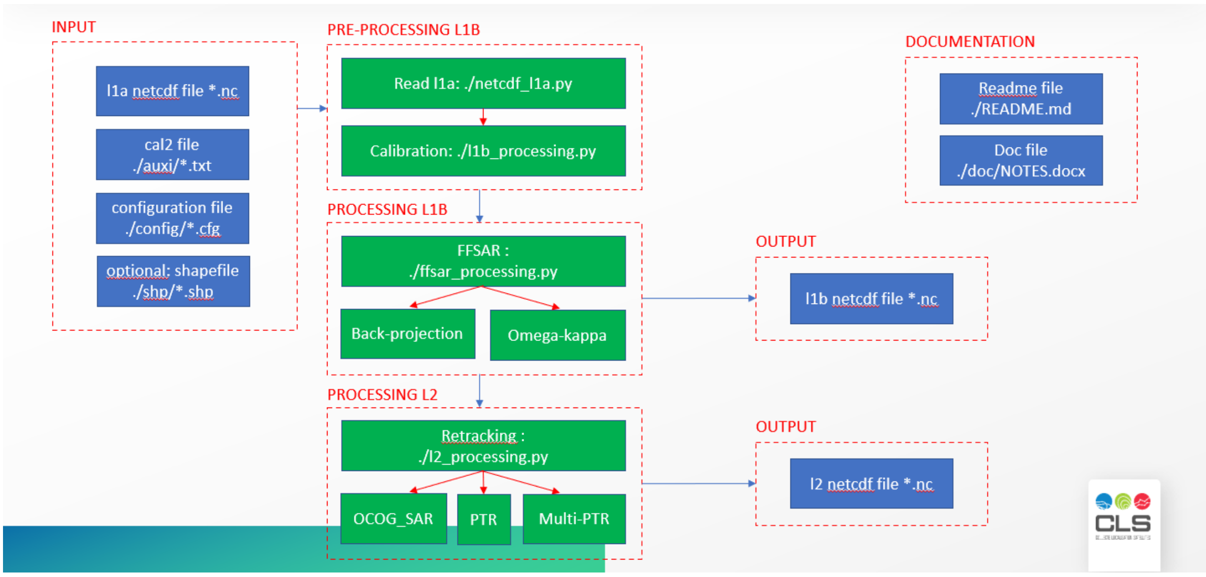


FIGURE 3. ARCHITECTURE OF THE SMAP CODE.

2.1. Inputs

The inputs of the SMAP code are:

- **One or more Official Sentinel-3 Level 1A product(s)** (available on the ESA SciHub website https://scihub.copernicus.eu/dhus/), to be put into a predefined folder into the sub-directory "**input**".
- **A Configuration file** (examples can be found in the sub-directory "**config**").
- **A CAL2 ASCII file** (available in the sub-directory "**auxi**").

An **optional** input file is the:

- **Shapefile** to select specific regions of the globe to process, like small water bodies (an example can be found in the sub-directory "**shp**", section 4.1 for more information).

Paths to the **CAL2** and **shapefile** files shall be indicated in the configuration file (see section 4.1).

Please note that no **CAL1** file is needed as the related information is directly extracted from Level-1A files (from variables named "burst_power_cor_ku_l1a_echo_sar_ku" and "burst_phase_cor_ku_l1a_echo_sar_ku".

2.2. Processing

Given the input files, the processing is divided into three blocks running sequentially:

- **Level 1B pre-processing**: in which Level-1A files are read and the waveforms calibrated (CAL1, CAL2, π-shift, experimental corrections),
- **Level 1B processing**: in which the fully focussed SAR processing (back-projection) is performed to obtain fully-focused L1b multi-looked waveforms,
- **Level 2 processing**: in which the retracking of fully-focused L1b multi-looked waveforms is performed to obtain amplitude, epoch, range and sigma0 estimates. Three retracking methods are can be selected in the configuration file (OCOG, PTR and Multi-PTR, see section 4.5).

2.3. Outputs

The SMAP code automatically generates an output directory (path to be indicated by the users in the configuration file, see section 4.2) including the following sub-directories:

- **cmd/**: it contains the bash command file created and launched,
- **log/**: it contains the log files (one for Level 1b and one for Level 2) with all the performed processing steps. If an error occurs during the processing (e.g. in case a wrong configuration is set), this is reported in the log file,
- **l1b/**: it contains the output netCDF Level 1b file,
- **l2/**: it contains the output netCDF Level 2 file.

2.4. Documentation

It is composed of both the present file and the readme file (**README.md**). These explain how to correctly launch SMAP. To avoid issues with the processor and its output products, the users shall install the environment and configure the processing as explained in the **README.md** file.

For further questions about information not provided in the documentation file, please feel free to contact us by sending an email to the CLS team, Samira Amraoui (*samraoui@groupcls.com*) or Thomas Moreau (tmoreau@groupcls.com).

3. Configuration file

3.1. Section FILES

- **cal2_file**: absolute path to the .txt file containing 128 floating point values representing the gain profile of the low-pass filter.

- **shapefile**: absolute path to the .shp shapefile.

The shapefile should be structured like the shapefile '**shapefile_smap.shp**' provided in the '**shp**' folder and contain polygons with an integer attribute '**pass_nb'**. SMAP will select the polygons with a '**pass_nb'** attribute corresponding to the **relative pass number** of the input L1A file(s) that intersect(s) the polygon. Polygons are particularly suited for small water bodies (containing only one pass number), for large water bodies (containing several pass numbers) it is preferable to use a bounding box selection. This field can be left blank in the configuration file if no shapefile has to be used.

3.2. Section DIR

- **output_dir:** absolute path to the directory that will contain all the outputs of SMAP (L1b and L2 netCDF files, log files, command files).

3.3. Section GENERAL_PARAMETERS

- **lat_min, lat_max, lon_min, lon_max**: bounding box parameters for **geographical selection**. Latitude shall be between -90 and 90, longitude between 0 and 360. Please insert these maximum values when no bounding box is desired and all the records of the Level-1A input product shall be processed.

- **surf_type**: a combination of 0,1,2 and 3 (separated with commas without spaces) shall be set to perform a **geographical selection** based on the Level-1A surface type flags (open_ocean_or_semi-enclosed_seas : 0 / enclosed_seas_or_lakes : 1 / continental_ice : 2 / land : 3).

3.4. Section PROCESSING

- **l1b_processing**: selection of the L1b processing **FFSAR back-projection**: please insert **ffsar_bp**.

- **l2_processing**: selection of L2 processing (retracker) to be applied to the waveforms. It can be stacked (separated with commas, without spaces): OCOG_SAR,PTR,MultiPTR. **OCOG_SAR** is the standard threshold peak retracker [Wingham, (1986)]. **PTR** retracker fits the main peak of the waveform with the range PTR (least-square criterion). **MultiPTR** retracker fits multiple peaks of the waveform with PTR retracking, with a number of peaks to retrack indicated in the configuration field **MultiPTR_n_estimates** and return retrack parameters for each individual peak. **PTR** (or **MultiPTR**) is an experimental algorithm which, as for numerical retrackers [Tourain et al., 2021], expresses the radar point target response as a square sinc function (that mimics an ideal CAL1 PTR shape). This model is suited for retracking single-point waveforms (e.g. over transponder or highly reflective surface such as leads and some inland waters).

3.5. Section PROCESSING_OPTIONS

- **illumination_time:** illumination time in seconds, it defines the integration time associated with the synthetic aperture. Please insert a **floating** number between **0.08** to **2.3**.

- **posting_rate:** posting rate in Hz related to the FFSAR along-track multi-looking of the single-looks, the spacing on ground between two multi-looks is equal to \`**satellite_velocity/posting_rate**\` meters. Please insert a **floating** number between **20** and **17825**.

- **zp:** oversampling factor in range. Please insert an **integer** between **1** and **2**.

- **hamming_range:** whether to apply Hamming weighting in range. Please insert either **yes** or **no**.

- **hamming_az:** whether to apply Hamming weighting in azimuth, along the burst. Please insert either **yes** or **no**.

- **range_ext_factor:** extension factor in range, processing is done on \`**128 * range_ext_factor * zp**\` range gates to be truncated into \`**128 * zp**\` central range gates at the end of the processing. Please insert an **integer** between **1** and **2**.

- **OCOG_threshold_sar:** threshold for the OCOG retracker. Please insert a **floating** number between **0.0** and **1.0**. The parameter is applied when the field '**l2_processing**' includes '**OCOG_SAR**'.

- **MultiPTR_n_estimates:** number of peaks to be retracked with the Multi PTR retracker. Please insert an **integer** between **1** and **10**. The parameter is applied when the field 'l2_processing' includes '**MultiPTR**'.

3.6. Other comments

- The tracker range and range values in the output product are corrected for **USO drift** and **internal path delay**. The COG distance is also included.
- The **AGC** correction is included in the calibration of the data, all the other calibrations applied (like CAL1, CAL2, π-shift between bursts,…) can be found in the **l1b_processing.py** file.
- The **surface height** is computed as '**alt_ffsar - range_ffsar'** (range_ffsar depends on the retracker that is used). It does not include any geophysical correction on the range. The user

can find geophysical corrections in official L2 PDGS products. **alt_ffsar** is relative to the reference **ellipsoid WGS84** (not to the geoid).

- Constants and useful variables for processing can be found in the **cst.py** file.
- Users can apply **empirical** (e.g. the Threshold Peak Retracker, [Davis, (1997)]) **& physical retrackers** by using data fields in the L1b products. These also include the tracker range and the scaling factor for sigma0 evaluation. The nominal tracking gate number for Sentinel 3 is 44 as indicated in the constant file cst.py (variable *self.abs_ref_track*), for more details please check the website FAQ - Sentinel-3 Alitmetry Technical Guide - Sentinel (esa.int).

4. Test Data Set (TDS)

Test data set (TDS) output files generated by the SMAP processor are provided in the '**output**' folder to illustrate some results and allow users to check that SMAP is correctly running on their machine. If the correct environment has been installed (see **README.md**), only minor discrepancies could occur.

Two products are considered: one collected over a hydrological target and one over a transponder located in Crete. Both products have been run with the FFSAR back-projection and the FFSAR omega-kappa algorithms.

4.1. Collect and Launch the TDS products

Users shall download the following products from the ESA SciHub website (https://scihub.copernicus.eu/dhus/):

- S3A_SR_1_SRA_A__20180628T192512_20180628T201541_20180723T214218_3029_033_013______LN3_O_NT_003 for the transponder TDS

- S3A_SR_1_SRA_A__20190730T101715_20190730T110744_20190824T200439_3029_047_279______LN3_O_NT_003 for the hydrology TDS

Once downloaded, the 2018 file containing data over the Crete transponder shall be put in the folder '**input/transponder**' and the 2019 file containing data over an hydrological target shall be put in the folder '**input/hydro**'.

To process the transponder FFSAR back-projection TDS from the python shell, load the directory including the processor & input files and do:

python launch_smap.py -c config/transponder/config_trp_ffsar_bp.cfg

To process the hydrology FFSAR back-projection TDS from the python shell, , load the directory including the processor & input files and do:

python launch_smap.py -c config/hydro/config_hydro_ffsar_bp.cfg

In general, users can create a sub-folder (e.g. "river") in the 'input' directory and place inside it all the tracks to be processed. After having created the associated configuration file (e.g.

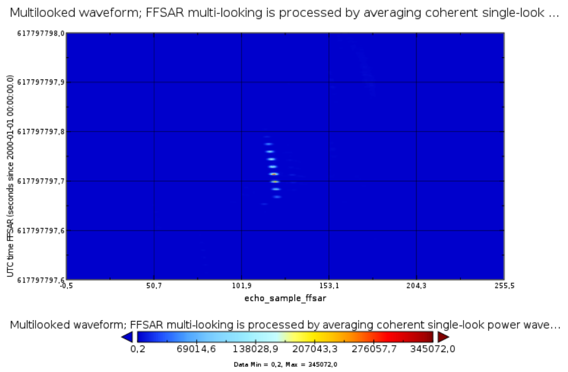
"***config_river_ffsar_omegak.cfg***", to be placed into the ***config/river*** directory), including the field input dir set as "**input_dir = ./input/river**", the processing can be launched as follows:

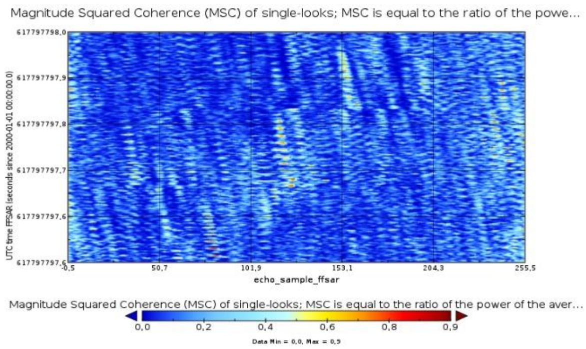
python launch_smap.py -c config/river/config_river_ffsar_omegak.cfg

4.2. Results of the Level 1b

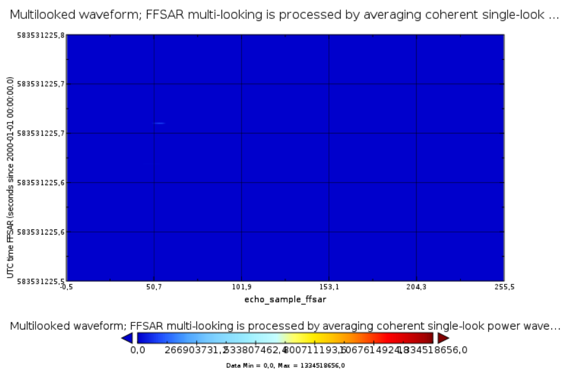
Panoply (https://www.giss.nasa.gov/tools/panoply/download/) has been adopted to visualise the TDS netCDF files. Below are reported the **multi-looks** (output netCDF variable "**multilook_ffsar**") and the **magnitude squared coherence** (output netCDF variable "**msc_ffsar**") of the Level 1B TDS files generated by SMAP.

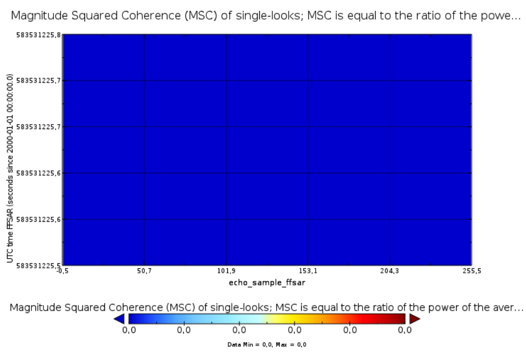
- The multi-looks and the magnitude squared coherence of the hydrology FFSAR back-projection TDS are:





- The multi-looks and the magnitude squared coherence of the transponder FFSAR back-projection TDS are:



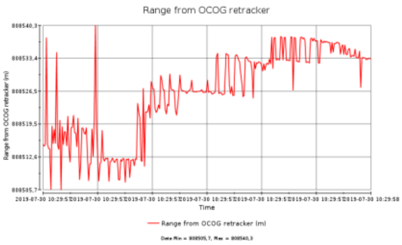


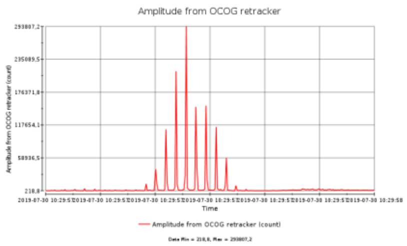
4.3. Results of the Level 2

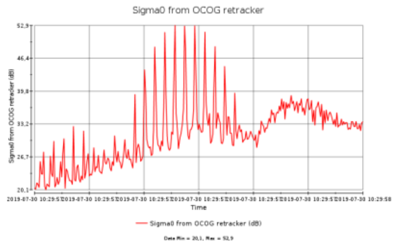
Panoply (https://www.giss.nasa.gov/tools/panoply/download/) has been adopted to check the provided netCDF files. Below are reported the retracked parameters **range** (output netcdf variable “**range_ocog_ffsar**”), **amplitude** (output netcdf variable “**amplitude_ocog_ffsar**”) and **sigma0** (output netcdf variable “**sig0_ocog_ffsar**”) of the Level 2 TDS files generated by SMAP.

Users shall keep in mind that **small discrepancies** between the FFSAR back-projection and omega-kappa processing approaches at Level 1B can lead to small discrepancies in the output L2 retracker variables.

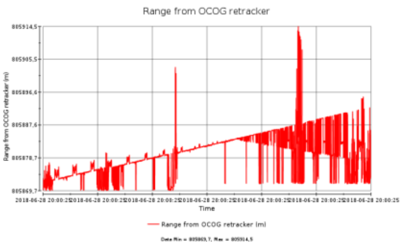
- The range, amplitude and sigma0 of the hydrology FFSAR back-projection TDS are:

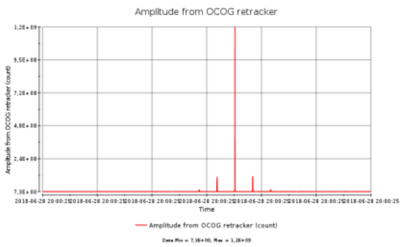


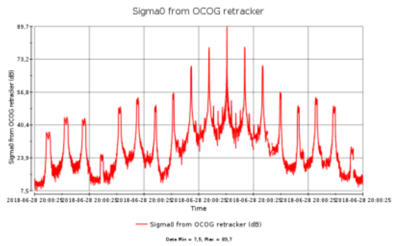




- The range, amplitude and sigma0 of the transponder FFSAR back-projection TDS are:







Reference

- Davis C. H., "A robust threshold retracking algorithm for measuring ice-sheet surface elevation change from satellite radar altimeters," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 4, pp. 974-979, July 1997, doi: 10.1109/36.602540.

- Egido, A., & Smith, W. H. (2016). Fully focused SAR altimetry: theory and applications. IEEE Transactions on Geoscience and Remote Sensing, 55(1), 392-406.] (https://ieeexplore.ieee.org/abstract/document/7579570).

- Guccione, P.; Scagliola, M.; Giudici, D. 2D Frequency Domain Fully Focused SAR Processing for High PRF Radar Altimeters. Remote Sens. 2018, 10, 1943. https://doi.org/10.3390/rs10121943.

- Tourain C. *et al.*, "Benefits of the Adaptive Algorithm for Retracking Altimeter Nadir Echoes: Results From Simulations and CFOSAT/SWIM Observations," in *IEEE Transactions on Geoscience and Remote Sensing*, doi: 10.1109/TGRS.2021.3064236.

- Wingham D. J., Rapley C. G. and Griffiths H. "New techniques in satellite altimeter tracking systems." ESA Proceedings of the 1986 International Geoscience and Remote Sensing

Symposium(IGARSS'86) on Remote Sensing: Today's Solutions for Tomorrow's Information Needs,. Vol. 3. 1986.