# Read-me-first note for the release of the SMOS Level 2 Soil Moisture data products

<table>
<thead>
<tr>
<th><strong>Processor version</strong></th>
<th>Level 2 Soil Moisture V650</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Release date by ESA</strong></td>
<td>November 15, 2017</td>
</tr>
<tr>
<td><strong>Authors</strong></td>
<td>Expert Support Laboratory (ESL) Level 2 Soil Moisture + Array Systems Computing Inc.</td>
</tr>
</tbody>
</table>

## Further information

Details on the processing algorithms can be found in the Algorithm Theoretical Baseline Document (ATBD)

- SO-TN-ARR-L2PP-0037 v3.10

available here: [https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/smos/content/-/asset_publisher/t5Py/content/data-processors-7632](https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/smos/content/-/asset_publisher/t5Py/content/data-processors-7632)

Information about the L2 Soil Moisture (L2SM) products structure can be found in the SMOS Level 2 and Auxiliary Data Products Specifications document:

- SO-TN-IDR-GS-0006 v8.5

available here: [https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/smos/content/-/asset_publisher/t5Py/content/data-types-levels-formats-7631](https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/smos/content/-/asset_publisher/t5Py/content/data-types-levels-formats-7631)

Additional information, including change history of L2SM algorithm and associated processor, documentation can be found on the SMOS L2SM CESBIO website: [www.cesbio.ups-tlse.fr/SMOS_blog/](http://www.cesbio.ups-tlse.fr/SMOS_blog/) and ARRAY website: [www.array.ca/smos](http://www.array.ca/smos)

## Contact for helpline

For all issues related to data access, formats and read/write, processors please contact ESA’s HelpDesk on [eohelp@esa.int](mailto:eohelp@esa.int).

## Comments to ESL Level 2 soil moisture team

The Level 2 Soil Moisture team would like to get your feedback on the product, either directly ([yann.kerr@cesbio.cnes.fr](mailto:yann.kerr@cesbio.cnes.fr), [rajesh@array.ca](mailto:rajesh@array.ca), [susanne.mecklenburg@esa.int](mailto:susanne.mecklenburg@esa.int)) or through the BLOG ([http://www.cesbio.ups-tlse.fr/SMOS_blog/](http://www.cesbio.ups-tlse.fr/SMOS_blog/)) where you can also find the latest news!
1. Introduction

This note summarises the quality of the SMOS Level 2 Soil Moisture data products generated by version 650 of the Level 2 Soil Moisture Processor (L2SM).

Version 650 of the Level 2 Soil Moisture data product is now available for the entire SMOS mission lifetime with the following file class and version:

<table>
<thead>
<tr>
<th>File class</th>
<th>File version</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>REPR</td>
<td>V650</td>
<td>1 June 2010</td>
<td>14 November 2017</td>
</tr>
<tr>
<td>OPER</td>
<td>V650</td>
<td>15 November 2017</td>
<td>present</td>
</tr>
</tbody>
</table>

Measurements from the commissioning phase (12 January 2010 - 31 May 2010) show drifts due to instrument tests taking place in this period. Even though data are available (upon request) it is not advisable to use them.

SMOS data users are invited to read this note carefully to ensure optimum exploitation of the version 650 dataset, which supersedes the previous version 620. Further information on the quality of the dataset can be found in the reprocessing verification and validations reports available here: https://earth.esa.int/web/guest/missions/esa-operational-co-missions/smos/content/-/asset_publisher/t5Py/content/data-quality-7059

This note is organised as follows:

- Main improvements in the L2SM version 650 dataset
- L2SM version 650 dataset performance and caveats
- Future algorithm evolution

2. Main improvements in the current L2SM version 650 dataset

The main improvements introduced in the current operational version V650 of the SMOS Level 2 Soil Moisture products as compared to the previous version 620 are related to processing algorithm updates, processing parameters configuration and auxiliary files changes.

In terms of algorithms, there are three changes between V650 and V620:

- The Chi2 values in V620 were too high with respect to the theoretical expectation. A Chi2 rescaling was therefore introduced to mitigate any possible mischaracterization of the observational error in the cost function used in the retrieval of soil moisture. As a consequence, the Chi2P (the probability of the Chi2) and the FL_Chi2_P flag are now usable for additional quality control checks.
- Current files retain the parameter values associated to the lowest DQX. However, in V620 a very low DQX values induced a risk of “locking” the variables (Tau, HR retrievals) on potentially wrong past retrieved values and propagate them. Minimum floor values for DQX stored in the current files have thus been introduced in V650 to limit the possibility of locking current files on wrong values.
- The last change is related to the V620 rain rate-based flood rule. This rule is considered to be no longer adequate for such purpose, particularly for strong events like hurricanes, where Numerical Weather Predictions for rain rate are intrinsically inaccurate. The flood flag has been redesigned and currently based on a configurable threshold depending on soil moisture itself and its uncertainty.

Therefore, V650 algorithm specific changes are the following:

- Chi2 rescaling with a ratio 1/1.66 with acceptance [0.05 1] for FL_Chi2_P
- Current file minimum DQX cut-off value: TH_Curr_Min_DQXxxx TLV=TFO=0.01, ROU=0.02
- \( TH_{FLOW} = 0.65 \text{ [m}^3\text{/m}^3\text{]} \) (raise \( FL_{FLOW} \) when \( SM+DQX_{SM} > TH_{FLOW} \))

Regarding processing parameters configuration, the following changes were applied:
- \( TH_{DQX_{SM}} \) raised from 20% (V620) to 30% (V650)
- \( TH_{DQX_{TauNad}} \) raised from 0.4 (V620) to 0.6 (V650)

The auxiliary files changes are summarized in Table 1. These changes are indeed the most important drivers for soil moisture differences between V650 and V620. The main change relates to using the IGBP database instead of the ECOCLIMAP database to characterize land surface coverage. One motivation for this change was to make the retrieval compatible to the one used within the SMAP mission. Moreover, soil moisture CDF-matching between ECMWF forecast and SMOS measurements (V620 baseline) was done to improve retrievals in the case of mixed forest/nominal pixels.

Table 1 - Changes in auxiliary files between V650 and V620 dataset

<table>
<thead>
<tr>
<th>ADFs</th>
<th>V620</th>
<th>V650</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUX_DFFRA</td>
<td>ECOCLIMAP+GLOBCOVER</td>
<td>IGBP Simplified (Broxton)</td>
</tr>
<tr>
<td></td>
<td>218 classes</td>
<td>18 classes; no wetland, no barren</td>
</tr>
<tr>
<td>AUX_LANDCL</td>
<td>Designed for ECOCLIMAP</td>
<td>Adjusted for IGBP 18 classes but keeping the parameters close to V620</td>
</tr>
<tr>
<td>AUX_ECMWF</td>
<td>No soil moisture CDF-matching</td>
<td>Soil moisture CDF-matching at GRIB file level</td>
</tr>
<tr>
<td>AUX_DFFSNO</td>
<td>Activated since 19 Dec 2016</td>
<td>Activated</td>
</tr>
</tbody>
</table>

The change from the ECOCLIMAP to IGBP land cover impacts the redistribution of the nominal vegetated soil (NO) and forested areas (FO), as shown by Figure 1 comparing the two decision tree branches.

---

![Figure 1: Different repartition of land use classes between V620 (ECOCLIMAP, left) and V650 (IGBP right)](image)

If we focus only on changes related to the nominal low vegetated soil case (NO) and the forest case (FO), as shown in Figure 2, we observe that the IGBP has increased the number of NO and FO cases leading to more soil moisture and opacity retrievals. In the northern latitudes many past cardioids retrievals (brown, yellow) become either NO or FO and thus produce soil moisture estimates. We
also observe a redistribution NO/FO, NO becoming FO (green) and FO becoming NO (red) in certain areas.

Figure 2: changes from V620 to V650 in terms of decision tree branches

3. L2SM version 650 dataset - performances and caveats

3.1 Performances

The comparison of the new V650 with respect to the V620 against a collection of calibration-validation sites concatenating the time-series of 9 networks of in-situ measurements, namely SCAN (75 stations), OZNET (17 stations), AMMA (2 stations), HOBE (2 averaged stations), SMOSMANIA (9 stations), REMEDHUS (8 stations), VAS (1 Elbara retrieval), WATERSHEDS (3 averaged stations) and SODANKYLA (1 averaged station) is shown by the two following Taylor plots diagrams.

The average performance of V650 retrieved soil moisture in term of correlation with in-situ soil moisture is equivalent to V620 for descending orbits and slightly improved for ascending orbits. The biases of soil moisture V650 are significantly reduced for both orbits (see Figure 3 and 4).

The V650 and breadboard configuration (named D51) are extremely similar, if not equal, which indicates that the entire reprocessing campaign over the entire SMOS mission lifetime performed as expected.

---

1 The products released are prone to be improved. Feedback to yann.kerr@cesbio.cnrs.fr and rajesh@array.ca on any issues you may identify while using or validating the data is most welcome.
At global scale, the V650 provides more (between 5% and 10%) retrievals for soil moisture and opacity. Wetter soil is expected at northern latitudes because of IGBP classification (less cardioids retrievals and NO/FO redistribution). Wetter soil is also obtained for heavily vegetated soil, and
particularly for forest categories, along with an increase of retrieved opacity. For other low vegetation to bare surfaces, V620 and V650 are similar.

Figure 5 to Figure 6 illustrate these changes with respect to soil moisture, while Figures 7 and 8 refer to the vegetation opacity. They were obtained by averaging 28 months of retrieved soil moisture and opacity for the specific months of January, April, July and October from January 2010 to April 2017.

Figure 5 Averaged soil moisture (4 months per year - January; April, July and October) during 7 years for V650 (top) and V620 (bottom) – black dot reports systematic failures.
Figure 6: Difference (V650-V620) of averaged soil moisture (4 months per year - January, April, July and October) during 7 years.
Figure 7: Averaged vegetation opacity (4 months per year - January; April, July and October) during 7 years for V650 (top) and V620 (bottom) – black dot reports systematic failures.
Figure 8: Difference (V650-V620) of averaged vegetation opacity (4 months per year - January; April, July and October) during 7 years.

With the addition of the Chi2 rescaling, the probability of Chi2 is more reliable. The associated flag FL_Chi2_P can now be used in association to the FL_NO_PROD flag for filtering data:

- FL_NO_PROD = 0 and FL_Chi2_P = 0 – successful retrieval with good fit quality.
- FL_NO_PROD = 0 and FL_Chi2_P = 1 – successful retrieval with degraded fit quality.
- FL_NO_PROD = 1 – rejected retrieval as usual.

One minor issue has been found in the retrieved dielectric constant over Antarctica due to ice temperature not correctly used. Potential users of the cardioid retrievals are advised not to use these data for this specific area.

3.2 Caveats (please read)

To help the user to understand the soil moisture product content, some basic information about the algorithm decision trees, the used radiative transfer model and the input L1C brightness temperature data are given below. Further algorithm details are illustrated in the soil moisture retrieval Algorithm Theoretical Basis Document (ATBD) available here (https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/smos/content/-/asset_publisher/t5Py/content/data-processors-7632).

- **Primary decision tree.** For each node of the SMOS grid, the concerned area (the working area is a 123 km x 123 km box when including the minor contributions) is described as including several fractions: e.g., vegetated soil, ice, forest, open water. Then, the first decision tree (17 branches) selects which of the fraction(s) will be considered for the retrieval. This depends on the respective weights of the various land...
use types, including also consideration on topography. Let us for example assume that part of the scene consists of vegetated soil while another fraction consists of open sea. Since the sea fraction is not relevant for retrieving soil moisture, the radiometric contribution due to sea is computed using auxiliary data and it is considered as a default contribution. This is only an example; indeed, the forward modelled brightness temperatures used in the retrieval will integrate most of the time default contributions.

Note that the retrieved soil moisture corresponds only to the area where the dominant land use is present.

- **Radiative models.** There are basically 3 radiative models in the L2 algorithm, depending on how the dielectric constant of the surface is computed. They can be used either in the retrieval iterative loop, or simply in order to build default contributions.
  a) The **nominal model** is the standard soil vegetation radiative transfer model and includes soil moisture (L-MEB, L-band Microwave emission of the Biosphere);
  b) The **water surface model** is used for the sea in coastal pixels, wetland and lakes;
  c) The so-called **cardioid model** is used for retrieval whenever it can only aim at providing information about the dielectric constant itself (e.g. ice, barren surfaces).

The nominal model will be used over vegetated soil and forest. While these cases are the only ones of direct interest as far as soil moisture is concerned, it may be mentioned that preliminary results using the cardioid model suggest there is indeed physical meaning in them.

It is important to note that a value of -999 for a geophysical parameter and its associated DQX in the L2SM UDP product implies that either the retrieval for that parameter failed or was not attempted. When all retrieval attempts fail for a node the FL_NOPROD is set to 1, otherwise it is set to 0.

- **Secondary decision tree.** Depending on the content of the working area, one of 3 radiative models is used for retrieval. For each of them, one must then define which parameters are to be retrieved, and what are the constraints assigned to the initial values. The secondary decision tree lists (for each of the 3 models) 3 options depending on the expected vegetation optical thickness and 3 options depending on the “information richness” expected from the data, which is estimated from the incidence angle coverage. As expected, the commissioning phase evidenced that this scheme with 27 options was too complicated and it was meant to be simplified eventually. Currently there are only 2 retrieved parameters for instance.

Note that the SMOS SM products are not produced as are many other similar products, so read carefully the ATBD to see exactly what is being done (as it is not always standard practice) so as to take advantage of SMOS characteristics

The user should also consider the following remarks on the usage of the Level 1 brightness temperature in the retrieval of the soil moisture parameter:

- **SMOS L1 product (L1C).** The L1C consists of brightness temperatures reconstructed from interferometric data in the reference frame of the SMOS antenna plane. Hence
these radiometric data are associated to upwelling Stokes parameters through a transformation combining the Faraday effect and geometrical factors.

_SMOs brightness temperatures are NOT TBH, TBV, third or fourth Stokes parameters and CANNOT be compared directly with Earth surface observation or modelling._

- **Soil moisture retrieval efficiency.** The SMOS soil moisture retrieval is based on matching measured and modelled (surface emission) brightness temperatures, with the modelled values varying as a function of the incidence angle and depending on soil moisture as well as other physical parameters. There is definitely, for a very large fraction of nodes of the SMOS grid, a robust ability of the radiative model to match the angular signature of brightness temperature while producing realistic values of retrieved parameters.

- **Polarisation mode.** According to the End-of-Commissioning review decisions, the full polarization acquisition mode has been selected for the operational phase, accounting for the potential information provided by this mode. However, until full understanding of third and fourth Stokes parameters has been achieved, “pseudo dual pol” is used for retrievals over land surfaces; i.e. only the antenna level brightness temperatures corresponding to first and second Stokes parameters will be used in the retrieval. Note that third and fourth Stokes are much improved with respect to initial releases but not yet totally up to our expectations. The Level 2 algorithm does not yet use the cross polarisation terms in the retrieval process.

  *Note that brightness temperature polarisations are always given in the antenna and can only be transformed into ground values (i.e., $H$ and $V$) through a transformation related to the instantaneous view angles. Only points located on the satellite subtrack are (almost) in $H$ and $V$. A Matlab tool is provided on the blog ([http://www.cesbio.ups-tlse.fr/SMOS_blog/](http://www.cesbio.ups-tlse.fr/SMOS_blog/)) to perform this task.*

- **Time reference.** We remind that the acquisition time all SMOS products is provided in EE CFI Transport Time format and the reference time in all SMOS products is MJD2000. The MJD2000 counts the time elapsed since 2010/01/01 00:00:00 UTC.

  *Take care MJD2000 is not J2000 standard reference time that starts at 2010/01/01 12:00:00 UTC*

### 4. Further algorithm evolution

Two main areas where progress is needed concern the correction of biases and the detection/mitigation of spurious signals.

V7xx should benefit from the next version of the L1 processor which will have some significant improvements.

For L2SM we will pursue the improvements done with respect to the parametrisation of the surface in the context of the new IGBP map, capitalizing on the work done for V650 and SMOS-IC outputs as well as a tentative roughness map (derived from SMOS) inclusion.

It is also expected to have an improved forest model and to include the organic soil model of dielectric constant at high latitudes. We will also endeavour to improve the water bodies mask.
Time permitting, we will investigate single soil moisture retrieval underneath both low vegetation and forest canopies.

Some caveats were already identified, as detailed above, and are under investigation.

Version 650 of the soil moisture Level 2 processor, which includes substantial corrections and improvements as a result of the evolution work done since the Commissioning Phase, is now delivering the best soil moisture products available. Nevertheless, we believe there is still room for improvements!

The data is now yours to evaluate. Please let us know of any issues encountered (on scientific points!) so that we can try to fix them for the next – general - release. We are looking forward to collaborating with you on these topics and make SMOS products even better.

Yann, Philippe and Philippe, Ali and Ali,
Ahmad, Arnaud, Beatriz, Cecilia, Cristina, Eric, François,
Jean Pierre, Jennifer, Marie, Mike, Nemesio, Paolo,
Raffaele, Rajesh, Roberto, Steven, Susanne, Tim.