

Author's response

1. Reviewer #1

We thank the reviewer for their constructive feedback and suggestions. Below, we provide our responses to each comment and describe the revisions proposed to address them.

Review of

Next-generation Metop ASCAT Surface Soil Moisture datasets

by Hahn et al.

General comments:

This study provides a detailed account of the methods used to create a soil moisture dataset with a C-band spaceborne scatterometer. This dataset is publicly available. The paper is well written, and I only have a few minor comments.

Recommendation: minor revisions.

Thank you very much for the positive feedback!

Particular comments:

- Overuse of the word “generation”: “next-generation Metop”, “ERA5 ... fifth generation”, “generation of the backscatter”, “time series generation”, “Metop-Second Generation”, “generation of ASCAT SSM”. Where possible, I would suggest finding alternatives to the word 'generation'. In particular, I don't understand why the dataset associated with this paper should be described as 'next-generation'. In the title and abstract, I would replace 'next-generation' with 'novel'.

We will refine the manuscript to use the word "generation" more carefully. However, we recognize that there are established technical terms such as Metop Second Generation (SG) or ERA5 (fifth generation), which should remain unchanged.

- L. 204 (Eq. 2): “w” is undefined.

In this case "w" refers to the Hamming window weighting. We will modify the text accordingly.

- L. 223: The Fibonacci grid appears to be a new concept, at least in the context of geophysical files. The authors should justify this choice more clearly and explain how the ASCAT data users can read such data using standard GIS software or Python libraries.

The following sentence will be added to section 3.1.2: " In the context of the ASCAT SSM dataset, the Fibonacci grid is used as a discrete global grid (DGG) that offers uniform sampling of the Earth's surface. Data stored on this grid consist a collection of geolocated points (EPSG:4326), requiring no specialized software for processing or interpretation.

- L. 251-252 : "However, this format compromises the ease of future data extension due its strict sorting structure" is not very clear. Could you clarify and explain why?

We will modify the text and add the following description: "More specifically, a contiguous ragged array stores chronologically ordered data for multiple grid points as a single contiguous block. While this data structure allows efficient data access, inserting data out-of-order (e.g., from earlier grid points) will corrupt both. Avoiding this problem requires inserting data at its correct chronological position which, would require a large amount of I/O operations (i.e. shifting elements, rewriting datasets).

- L. 294 (Eq. 9): ":@" do you mean "="?

In this context, it is meant to "define" the term δ so that the expectation of the difference between Fore and Aft beam is zero.

- L. 394 (Eq. 21): Is there any justification for using a linear relationship to convert ASCAT sigma0 values into soil moisture values? Using a sigmoid function would easily solve the outlier issue.

Empirical studies have shown that radar backscatter and soil moisture are linearly related (e.g. <https://doi.org/10.1109/TGRS.1986.289585>). However, using a sigmoid function for scaling is an interesting idea, we will definitely investigate it in future work. Thanks for this suggestion. One remark though, we should probably remain cautious since it may hide critical calibration issues or instrument drifts normally presented as outlier.

- L. 433: What is the procedure for dealing with temporary standing water, for example during flood events?

Temporary standing water may change the backscatter signal, with its effects varying by location and flood characteristics. Unfortunately, information of temporary standing water information is not available in near real-time and also incorporating historic data is not straight forward because of its varying impact. Definitely a limitation to be addressed in future work. At the moment, ASCAT SSM data during (large-scale) flood events should ideally be cross-checked against auxiliary datasets (e.g., flood extent maps).

- L. 453: Is the likelihood of soil freezing indicated by a flag? Soil freezing and snow cover could be derived in real time from numerical weather forecast models.

Climatological probability flags for frozen soil and snow cover are provided as part of the ASCAT CDR/ICDR/NRT SSM dataset. The ASCAT NRT data stream has started to use ECMWF forecast information contributing to the $surface_{flag}$ (Table 14). We will mention this detail in section 4.1.

- L. 534-535: Replace by: "SWI is computed using an exponential filter of SSM. This mimics water infiltration into deeper soil layers. It does so in a purely statistical way."

We are aware that the exponential filter is mathematically simple but we respectfully disagree that it is purely statistical method. It was derived by solving a first-order differential equation that connects the soil surface layer and second layer below, yielding a spectral behavior as produced by models such as the Geophysical Fluid Dynamics Laboratory (GFDL) model as e.g. described in Delworth and Manabe (1988). Please see the paper by Ceballos et al. (2005) published in Hydrological Processes (<https://doi.org/10.1002/hyp.5585>).

- L. 551: I would replace 'best performance' with 'best agreement', given that passive microwave SM is not perfect.

Agreed, the term 'best agreement' will be used.

Editorial comments:

- The Pearson correlation coefficient (R) is defined multiple times. The first definition is sufficient. Then, use 'R'. Same for 'SNR'.

We will reduce the number of R and SNR definitions.

- Is it "H SAF" or "HSAF"? 'H SAF' looks strange.

The official acronym is defined as "H SAF" (see <https://hsaf.meteoam.it/>)

2. Reviewer #2

We thank the reviewer for their constructive feedback and suggestions. Below, we provide our responses to each comment and describe the revisions proposed to address them.

General comments:

The manuscript presents a new ASCAT surface soil moisture (SSM) dataset with a nominal sampling of 6.25 km alongside the standard 12.5 km product. The development of higher-resolution scatterometer-based soil moisture datasets is valuable for a wide range of hydrological, agricultural, and climate applications, and the effort to provide publicly accessible datasets through the H SAF framework is highly appreciated. The manuscript is generally well written and provides a detailed description of the processing chain and validation procedures.

Thank you very much for the positive feedback!

However, in its current form the manuscript does not clearly communicate some key aspects of the dataset development and positioning relative to previous ASCAT SSM products. Therefore, I recommend major revisions before the manuscript can be considered for publication.

- First, the title refers to the dataset as a “Next-generation” ASCAT soil moisture product. While the abstract briefly mentions several methodological updates to the ASCAT SSM retrieval algorithm, these improvements are not clearly highlighted in the methodology section. The methods are described in great detail, but it remains difficult for readers to identify the key algorithmic advancements compared to previous ASCAT SSM releases. It would therefore be helpful if the authors could more clearly summarize the main updates and distinguish them from earlier ASCAT SSM datasets in the methodology section.

We recognize that the main updates are not clearly summarized and we will add a new table to the methodology section.

- Second, I suggest that the validation section include a comparison with previous versions of the ASCAT SSM datasets. This would help to more clearly quantify the improvements in performance or data quality introduced by the updated processing chain, and would make the characterization of this dataset as a “Next-generation” product more convincing.

We will provide an ISMN validation using the previous ASCAT SSM NRT 12.5 km product compared against the new ASCAT SSM 12.5 km product.

Specific comments:

- The abstract should provide clearer information about the dataset. In particular, the temporal coverage of the dataset is not specified, which is essential information for a data description paper. In addition, although a DOI is provided, it is not clear which DOI corresponds to the 6.25 km dataset and which corresponds to the 12.5 km dataset. The version identifier of the dataset should also be specified.

More information on the temporal coverage will be provided in the abstract, as well as a correct mapping of the DOI for 6.25 and 12.5 km dataset.

- The introduction does not clearly highlight the limitations of the previous ASCAT SSM products and the motivation for the new dataset release.

The following paragraph will be added to the introduction: "The aforementioned improvements of the soil moisture retrieval algorithm have been gradually implemented and published as offline ASCAT SSM data record products provided by the Satellite Application Facility on Support to Operational Hydrology and Water Management (H SAF) lead by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)~\citep{h113-ssm-dr, h115-ssm-dr, h119-ssm-dr}. However, the ASCAT SSM NRT dataset~\citep{somo12} produced as part of the EUMETSAT ground segment has never benefited from these algorithmic advancements, due to a separate implementation of the NRT and offline processing chains. The EUMETSAT NRT processing chain has been discontinued on 14 July 2025 and replaced by a new EUMETSAT H SAF processing chain, migrating the latest offline soil moisture retrieval improvements to both data streams. The algorithmic updates of the newly released EUMETSAT H SAF ASCAT SSM datasets address the following limitations, which were present in the EUMETSAT NRT processing chain:"

- **Model parameter estimation only based on Metop-A ASCAT for the time period 2007–2014.**
 - **Only using a land/water ratio flag during spatial resampling, instead of filtering individual backscatter echos over cities and open water.**
 - **Globally static vegetation correction parameters.**
 - **Long-term land cover changes affecting backscatter trends, propagating into the soil moisture retrieval.**
 - **No information on the effect of subsurface scattering.**
- The title of Table 1 should include the term “microwave”.

The term "microwave" will be added in the caption.

- Table 1 currently provides only limited information (spatial resolution, temporal coverage and references). It would be useful to include additional characteristics of the different soil moisture products, such as sensor type (active/passive), observation frequency band, and whether the product provides absolute or relative soil moisture estimates.

Two new tables with more information will be provided.

- Table 2 could be removed and the information integrated into the text.

Information will be merged into the corresponding paragraph.

- The datasets used in this study have different spatial resolutions, but the manuscript does not clearly describe how they were spatially matched with the ASCAT grid. Please clarify the resampling or spatial matching method used. In addition, it would be helpful to provide download links or access information for the datasets introduced in this section.

Information on spatial matching (mostly nearest-neighbor) will be provided, as well as download links.

- Figure 1 provides only a very high-level overview of the workflow. Given the extensive

description in the methodology section, the figure could be expanded to better illustrate the key processing steps and algorithm components, which would help readers more easily understand the overall processing chain.

Figure 1 will be updated to include a more detailed illustration of the model parameter estimation and application of the TU Wien change detection method.

- The methodology section is very detailed and, in some places, somewhat verbose. The authors may consider condensing less essential descriptions so that the key processing steps and methodological aspects are easier to follow.

We will try to condense the methodology section keeping core technical descriptions. However, it would be helpful to provide more information on which subsection(s) appear to be verbose.

- L196: The manuscript states that the theoretical spatial resolution of the 6.25 km product is about 15 km, while the actual spatial resolution cannot be determined precisely. It would be helpful to clarify the typical effective resolution or its expected range.

We will clarify the effective resolution by adding the following text to the manuscript in section 3.1.1: "The effective spatial resolution closely approximates the theoretical resolution due to the application and size of the Hamming window."

- The caption of Figure 4 is too brief.

Modified caption: "A visualization of the cross-over concept: Backscatter-incidence angle curves for constant soil moisture and varying vegetation intersect at characteristic angles where the influence of vegetation on backscatter is minimized. These angles, called the dry and wet cross-over angles, are chosen to minimize the effect of vegetation during the estimation of the dry and wet backscatter reference."

- L465: The ASCAT SSM product is expressed in degree of saturation, while the reference datasets (e.g., GLDAS, ESA CCI, and ISMN) are typically provided in volumetric soil moisture ($\text{m}^3 \text{m}^{-3}$). The manuscript does not clearly describe how this unit mismatch is handled in the validation procedure.

A description will be added in section 3.5, clarifying that the mismatch in units does not affect the computation of Pearson correlation (R) or the signal-to-noise ratio (SNR).

3. Reviewer #3

We thank the reviewer for their constructive feedback and suggestions. Below, we provide our responses to each comment and describe the revisions proposed to address them.

Review

Title: Next-generation Metop ASCAT Surface Soil Moisture datasets

Authors: Sebastian Hahn et al.

MS No.: essd-2025-746

The paper describes the new version of the ASCAT surface soil moisture data. The changes in the algorithm w.r.t. the current version are substantial, making a full-length data description article necessary. The article topic is definitely of interest to ESSD readers.

The paper is generally well written. The ASCAT soil moisture data have been used very widely for many years, and the new version will be an important dataset for the community. Error estimates are

included with the dataset, and the error estimates as well as the data quality are adequately described in the article. The length of the article is appropriate, and the article is well-structured and clear.

Thank you very much for the positive feedback!

Having said that, I was missing a one-to-one comparison of the current (old) and new data versions. The paper is suitable for publication after the authors address the comments below.

Major comment:

The one big missing piece is an apples-to-apples comparison and validation of the current and new versions of the ASCAT data. This includes two aspects:

1. The implicit assumption is that the new product is better, or at least no worse, than the current product version. The article only touches on this in Lines 608-611. I recognize that the current data come in two flavors (time-series product and NRT product), which are generated with different algorithms and might thus have different quality. I also recognize that the current data and the new data are on different grids. Finally, I recognize that the processing changed significantly from the current version to the new version. But from a user's perspective, the most important piece of information is the quality of the new data compared to the current data. None of the above gridding and processing differences is truly an obstacle to a comparison of the current (old) and new versions. The authors should add at least one of the current products to at least one of the validation results. Since the ISMN validation used the QA4SM tool, it should be very straightforward to add the current (i.e., old) time-series product to the ISMN validation in Fig 15.

In order to provide a direct comparison between the "old" and "new" version, we will provide an ISMN validation using the previous ASCAT SSM NRT 12.5 km product compared against the new ASCAT SSM 12.5 km product.

1. Another burning question for this reader is whether the new version of the ASCAT data is suitable as an extension of the current (old) version. ASCAT data are routinely used in the land analysis of NWP and reanalysis products, where rescaling is baked into the system. Do the current (old) and new data have sufficiently similar climatologies to allow for simply replacing the input ASCAT obs? Or do the assimilation systems need to be recalibrated for the new ASCAT version? At a minimum, a comparison of the climatological mean and std-dev of the current (old) and new products would be needed to start answering this question.

This is certainly a good point to discuss and can be answered without a comparison of the climatological statistics. We will add the following description in the beginning of the methodology section (after the *new* summary of the key algorithmic advancements): "Updating the soil moisture retrieval model parameters modify both the climatological mean and the temporal dynamics of the ASCAT SSM datasets, resulting in inconsistencies with the previous ASCAT SSM NRT version. Consequently, data assimilation systems cannot transition to the new ASCAT SSM dataset without prior recalibration."

Minor comments:

1. The mathematical notation needs minor improvement and clarification:
 1. Equation (3): Please clarify subscript "i". I assume the meaning of "i" here is different from that of $i=1,12[13]$ used in section 3.3.1, and it is also different from that of "i" in (5). Perhaps best to use a different index variable in (3).
"i" refers to "all observations". The index will be changed to "j", also updating the

following sentence: "In addition, an outlier detection method based on the Median Absolute Deviation (MAD) is applied per beam for all echos (σ^0) to further refine the spatially resampled dataset ..."

2. Equation (8): I understand that $\hat{\sigma}$ here indicates the backscatter after normalization of the azimuth angle ϕ . But I did not see $\hat{\sigma}$ again in the remainder of the paper. (Perhaps I missed it.) I expected to see $\hat{\sigma}$ used in section 3.3.3 ("Incidence angle dependency"). It remains unclear to me if there is an order to the azimuth and incidence angle correction or normalization. Please clarify.

We will add the following sentence at the end of section 3.3.1: "All following processing steps are using the azimuth angle corrected backscatter, but for simplicity we use the usual notation of backscatter σ^0 rather than the hat notation $\hat{\sigma}$."

3. Equation (12): Subscript "m" has not been introduced. I understand it stands for mid-beam. Line 290 introduces subscripts "a" and "f", although somewhat implicitly, and Line 274 introduces the triplet "Fore, Mid, or Aft" (but not the lower-case subscripts "f", "m", and "a"). Please clarify.

Fore/Mid/Aft will be referenced to σ^0_f , σ^0_m , σ^0_a in the text.

4. Equation (15) introduces the symbol "d" for day-of-year. This is easily confused with the subscript "d" for "dry (reference)". I recommend changing the symbol for day-of-year to "DOY", or at least an uppercase "D", to more clearly distinguish the "day-of-year" and "dry" notation.

Symbol "d" will be changed to "DOY"

5. Equations (17-18): The (percentile) symbols " P_2 " and " P_{98} " have not been defined explicitly. The meaning can be inferred from the context, but the reader should not be asked to sort this out.

Symbols " P_2 " and " P_{98} " will be defined in the text.

6. Equations (17-18): I'm unsure what exactly is included in the set of data months in the argument of P_2 and P_{98} . For example, if time "t" falls into August 2015, the set of months here could include all months within the period June 2012 through November 2018 (that is, Aug 2015 plus/minus 42 months, and assuming my math is correct). Or it could include on the August months, that is {Aug 2012, Aug 2013, Aug 2014, Aug 2015, Aug 2016, Aug 2017, Aug 2018}. The text in Line 355 ("account for seasonal vegetation effects") suggests that only August data are included. But the notation suggests to me that all months in the 85-month period are included. This requires a bit more clarification in the text and/or equations.

The following example will be included in the text: "e.g. the target month August 2015 would include the time period from February 2012 to February 2019."

7. Line 389: Replace "(i.e., a=5, b=95)" with "(i.e., a=5%, b=95%)". (Missing units.)

The % sign will be added.

2. Lines 55-56: "... this gives users the advantage..." The "relative" soil moisture units of ASCAT soil moisture are not an "advantage" over the "absolute" soil moisture units of SMAP and SMOS soil moisture. Users can just as readily rescale SMAP and SMOS data, if that is what they want or need to do. Please revise the sentence in question accordingly.

Rescaling data using techniques such as min-max normalization or CDF matching is

always an option. However, in this specific case (converting SMAP/SMOS vol.SM in $m^3 m^{-3}$ to % saturation), the scaling is based on the relationship "vol.SM / porosity = % saturation". Therefore, the original soil parameters (i.e. porosity) used to derive vol.SM would be required to correctly convert back to % saturation and thus allow the use of custom soil parameters for scaling. To avoid confusion, the sentence will be revised to: "This allows users to derive absolute soil moisture themselves by combining ASCAT data with their own soil parameters."

1. Table 2: The "observation periods" of Metop-B and Metop-C do not end on 2024-12-31. Both satellites remain operational as of this writing. Perhaps this end date refers to the Metop-B/C data used in the derivation of the algorithm constants? Please clarify. Also, Line 543 notes an end date of 2023-12-31 for data used in the SNR validation, but I could not find an end date for the ISMN validation.

Yes, the time periods refer to the ASCAT data used to compute the model parameters. Table 2 will be removed and dates will be included in text: "ASCAT backscatter observations from all three Metop satellites are used to compute the model parameters (Metop-A: 2007-01-01 – 2021-11-15, Metop-B: 2013-06-01 – 2024-12-31, Metop-C: 2019-04-01 – 2024-12-31)."

1. Line 104: "generate a frozen soil and snow cover probability flag". Based on later text, I understand this flag to be "climatological". I suggest changing this to "generate a climatological frozen soil and snow cover probability flag" (or similar clarification)

Sentence will be modified.

2. Line 396: Add explicit reference to section 3.3.2, where "ESD" is defined.

Reference will be added.

3. Line 445: replace "While in these regions observation noise is typically higher, ..." with "While observation noise is typically higher in topographically complex regions, ...". (The preceding sentence is about all land regions, not just topographically complex regions.)

Sentence will be modified.

4. Table 3: I recommend adding additional information to this table. This might require formatting the table as a "sideways" table or transposing the table. But I think the additional clarity is worth the formatting complications.

1. Provide quantitative information about the latency of each product.
2. Provide explicit information about the swath "duration" (PDU). g., H130 should have 60-minute PDUs, and H122 should have 3-minute PDUs. (It's possible that by adding this info the "Category" information becomes redundant; but there's nothing wrong with some level of redundancy if it helps the reader's understanding.)
3. The data format column could be dropped by adding the following to the caption: "All products are available in netCDF format. H121 and H29 are also available in BUFR format."

A new table with more information will be provided.

Editorial comments:

- Line 49 (typo): delete "with" (or otherwise fix grammar)
"with" will be deleted.
- Line 81: delete "currently" (redundant)

"currently" will be deleted.

- Line 166 (typo): replace “scatters” with “scatter” or “scattering” or “scatterers”
"scatters" will be replaced by "scatterers".
- Line 213 (typo): replace “observation” with “observations”
"observation" will be replaced.
- Line 499 (typo): Replace “application” with “applications”
"application" will be replaced.