

Comment on essd-2021-92

Anonymous Referee #1

Referee comment on "Polar maps of C-band backscatter parameters from the Advanced Scatterometer" by Jessica Cartwright et al., Earth Syst. Sci. Data Discuss.,
<https://doi.org/10.5194/essd-2021-92-RC1>, 2021

The authors present a dataset that is a derivation of existing and freely available ASCAT remote sensing data. Its main purpose is to save others from having to go through extensive processing in order to extract parameters from the raw data, using published methods (by the authors). These parameters relate more directly to physical information and can thus be analysed directly. As the authors write themselves, other such derivates from the same raw ASCAT remote sensing data already exist, however with different focus and/or using different and less advanced/extensive methods.
The article has a good structure and flow. It provides a well-balanced introductory part which covers sensors/physics/products/applications and thus provides useful background also for users new to such data. The description of the methods and in particular the dataset itself are rather short. Given that this is not the only ASCAT derivative that exists out there, this part of the data description would deserve more weight.

General/major comments:

- Novelty: Given there are existing derived ASCAT products, the new dataset could be more clearly distinguished from existing ones (section 2.3).
- Verification: The article should include data quality figures (statistics, error assessment) and a comparison of the three generated products. Comparing the three timescales using a concrete example/application would help the user to choose and know the limits of the product they chose compared to the other products. I am also missing some kind of a validation/comparison to comparable existing products that authors themselves mention in section 2.3.
- Verification/applications: The applications section is very shallow. Users who are not familiar with ASCAT data don't get any help here - I would have expected one or several case studies with results/maps for different uses, e.g. those you mention in section 2.1 - glacier surface mass balance, sea ice, ice bergs... Currently, applications are only mentioned in a theoretical way without any concrete examples. It would be much more

useful for the reader to have an actual application result/example, and ideally a comparison of some of the mentioned "such studies" (P8,L217) with the parameters provided within the three new datasets.

- The temporal coverage of the dataset (various starting dates, until 2019 only) is not sufficiently described in the article, in particular it is not mentioned that the datasets are available through 2019 only. This is especially important because this means that the flagship 1-day-dataset only covers a few months in 2019. Would it be possible to include 2020?

- Projection information should be available in the dataset/metadata: an unambiguous CRS/EPSG string/code is not provided in the article, nor in the dataset DOI page or the NetCDF files themselves. Not all users will be familiar with the polar stereographic projection and unclear attribute names (both X/Y and latitude/longitude are used in the NetCDF files) and ambiguous projection information means users might struggle to read the data or use it together with their own data.

Minor/detailed comments, sections 1&2

P2,L42: double brackets in citation; also at P4,L111

P3,L64: typo, missing verb

P3,L74: complicated phrasing. Note that C-band penetration varies considerably and may be only a few meters or even more than 20m even for dry snow and clean glacier ice. Consider: For dry snow and glacial ice, C-band penetration depths can reach 20m or even more

P4,L103-104: wrong dash (use --- instead - also in P4,L122 and many other places)

Here also in the wrong place, disturbing the reading flow. Use commas/brackets for both examples of sensor types.

P5,L124: what are "simple anisotropy parametrisations" (why aren't these sufficient)?

P5,L129: The distinction of your dataset to a) and c) is obvious, but why b) is not good enough is not immediately clear and the distinction deserves more explanation. Also, the term "balance" isn't very clear in the light of a-c). The paragraph about your own dataset might fit better in the end of this section?

P5,L135f: Within the context, it is not entirely clear whether these datasets (CERSAT, NASA SCP) are somehow part of your product or a competing/complementary product. Rather call them "products" to clarify (maybe also refer to b) in the list above?)

P5,L143: it is a little unclear what you mean with this sentence. You mean that OSI SAF provides some of the same information as your dataset, but yours is better (in what sense?)

P5,L45: temporal or spatial resolution-enhanced? Best to be specific as both types of resolution are mentioned throughout the text.

section 3

P6, L157: misleading, as this sounds like full coverage of the entire polar areas every 50 minutes. Please clarify.

Table 1: The star * footnote is a little confusing (and it is questionable whether the table caption is the right place for it, given some of it are technical/processing choices you

made). March or July 2007...? Also, the following explanation for MetOp-B doesn't fit here (would need another footnote). Suggestion: Add another column (or replace "Data Available") with "Data Used".

P6, L175: typo (it)

P6, L169: Figure order? Nr. 5 is first...? Would these figures be exactly the same (mirrored) for the Northern hemisphere, and MetOp-A and -C, respectively? Could be worthwhile to add this to the caption.

section 4

P7, L184ff: It is unclear/misleading what you mean with "version 3" - table 3 on the website you refer to? Table 3 corresponds to the "Northern Hemisphere Projection Based on WGS 1984", defined by the EPSG code 3413. Rewrite, and please provide the (unique) EPSG codes - assuming you used the projection described in table 4 on the NSIDC page for the southern hemisphere? --- After analysing the datasets, I suggest adding here that the dataset also provides lat/lon values for each pixel.

P7, L189: It might be good to remind the reader here how your approach differs from (and is much better than?) the one of existing products (in section 2.3).

P7, L208: "observations" sounds like an independent dataset, do you mean your own raw data? Please clarify. In particular for users not so familiar with ASCAR data / anisotropy, the residual parameter ought to be explained more. "Residual" is commonly interpreted a quality measure, and in this context this also seems correspond with the occurrence/percentage of invalid pixels - all of this is easy to confuse.

P7, L211: Consider changing to something like this: "In addition to the 9 parameters, latitude/longitude coordinates and a data quality flag parameter are provided for each grid location..." to clarify that the data flag is an additional layer, and different to the residual parameter. It would then be nice to see a plot of the quality flag mask as well.

P7,L211ff: some statistics would be nice here? What is the percentage of invalid pixels of the different products (e.g. per map/year, over land/sea, within the polar circle...)? How do the 3 timescale products differ, in particular for the years where you provide all products?

P7,L213: typo (use)

P7,L214: What about the northern hemisphere? Fig. 6 only covers the south. Some simple statistics like the nr of invalid pixels over land (Antarctica)

sections 5&6

This section is more a theoretical comparison of the new product with existing ones, plus some potential future developments. No actual applications...?

P8, L223: Please clarify: "These data" = your own products or the NOAA/CERSAT? The description of other products would fit better in section 2.3.

P8,L226: Table 1 suggests 2006 to present (2021) = 15-16 years? Not 12 years as stated here.

P8, L227ff: This is an outlook rather than an application. Move to conclusions? Missing timeframe - currently, it seems that these satellites will be launched ca. 2024 (<https://www.eumetsat.int/metop-sg>)? As it is written now, the 2012 reference is misleading for readers who are not familiar with the upcoming MetOp-SG missions.

P8,L237: 2021 seems to be outdated (see link above)?

Figures 1-4: A description of what can be seen in these datasets (applications) would be

desirable, to help the reader see the potential. What about the data quality flag layer? For comparison of the three different products, a side-by-side view of the same parameter and area might be more useful. In turn, in addition to a few such comparisons, it should be sufficient to show one complete set of parameters for one product rather than four full-size figures with little different information content.

References: please include DOIs to facilitate finding the references.

datasets

Access:

- From the dataset doi page, it is possible to a) download the entire dataset, which requires providing an email address, or b) view the dataset contents - where parts of the dataset can be downloaded without providing an email address. Is this difference intentional?
- Clicking on "view dataset contents" leads to a folder structure containing subfolders. This is well organised, but a link/README file linking back to the doi page and/or how to cite the data might be useful to add to ensure correct referencing, especially given that the dataset DOI is not included in the attributes of the individual .nc files (consider changing this?).
- I experienced that downloading the data was extremely slow (several hours for 2-3 GB despite my very fast internet connection, several attempts). Was this a temporary problem of the data server, or is it always the case? Limited download capacity will prevent users from using this data.

Data content:

- Attributes: I got confused by the mix of dimensions X/Y and the variables "latitude" and "longitude" in the NetCDF dataset. While not an expert on NetCDF standards, I found it rather misleading to use a polar stereographic grid, but not defining the X/Y dimensions (they are in km), yet link to that polar stereographic grid in the latitude/longitude variables (which are in degrees) description. To avoid confusion, it should be made clearer to the user that there are two coordinate systems in the same dataset, sort of. Also, note the typo in the description field of the latitude/longitude variables attributes (steTreographic).
- The definition of the projection (dataset attributes) is insufficient - the provided link to the NSIDC page describes several projections (see comment above) - better to provide the projection/EPSG code directly. (The NH/SW projections are not stated on the dataset

DOI page, so no help to find there.) It seems best practice would be to define a grid mapping variable, as described on this paragraph on coordinate reference system (CRS) standards for NetCDF climate and forecast data: <http://cfconventions.org/Data/cf-conventions/cf-conventions-1.8/cf-conventions.html#grid-mappings-and-projections>

- There is a ASCAT2day_NH_2012.nc (and SH) dataset despite the article stating that 2-day-processing only started in 2013...? I checked some parameters and it seems the entire NH file only contains fill values (-999)?