Comment on essd-2021-461
Anonymous Referee #2

Referee comment on "First SMOS Sea Surface Salinity dedicated products over the Baltic Sea" by Verónica González-Gambau et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-461-RC2, 2022

Need for satellite-based sea surface salinity (SSS) determinations is very urgent in the Baltic Sea, because in the sea with large-scale estuarine gradients, salinity is important both for the physical and ecosystem dynamics. The MS is a major contribution in the development of regionally adjusted SMOS method, which is timely and technically sound. The products of SSS have been prepared as well. The work has been done within the project of ESA initiative Baltic+.

The MS should be published. I made some comments, which can be taken into account when preparing the revised version of the MS. The comments reflect an oceanographer’s opinion.

1) The introduction is lacking the description of SSS of the Baltic Sea. Since it is the target of present study, the main SSS features should be outlined and their forming mechanisms should be explained. Mentioning that the salinity dynamics is complex (for example, on lines 26-30) should be preceded by simple basic facts. For example, due to the geographical separation of locations of major freshwater sources and the straits channeling the saline water inflow, the SSS has persistent large-scale variations between and along the basins. I recommend rewriting the introduction accordingly.

It could be also interesting to compare the large-scale SSS variations (about 6 psu per 400-500 km in the Gulf of Finland and the Gulf of Bothnia) with other oceanic regions of SMOS applications.

2) The statement “the Baltic Sea is one of the most challenging regions for the SSS retrieval from L-band satellite measurements” (lines 43-44) is followed by a lengthy text, where the main peculiarities cannot be found easily. What I found from the text, the main challenges for SMOS application are vicinity of coasts, seasonal ice cover and low salinity
range. I would like to read (such) a short message somewhere before or after the lengthy presentation.

3) Chapter 2 “Generation of Baltic+ SSS products” has mutually dependent sub-chapters 2.1 “Data sets used in the generation of the products” and 2.2 “Algorithm description”. A reader interested in SSS results, but not familiar in details of SMOS method, would benefit if there will be a few introductory sentences before the title of 2.1. Presently, the data sets like “SMOS Brightness Temperatures” are listed perhaps too abruptly, without explanation of their role in the algorithms. Therefore, rewriting is recommended.

I understand from 2.1 that SMOS brightness temperature is the basic observed quantity for SSS retrieval, while auxiliary data like sea ice cover, rain rate, 10-meter wind speed, 10-meter neutral equivalent wind, significant wave height of wind waves, 2-meter air temperature, surface pressure, and vertically integrated total water vapour are used as well. SST fields are used for correction of systematic biases of SSS. Sea ice data are used to exclude the ice covered areas from SSS estimates. Error filtering and validation of SMOS products is done by comparing with independent SSS reanalysis and operational forecast products.

4) The chapter 2.2 “Algorithm description” is much broader than just description. It contains significant research to study the options of implementation and fine-tuning of algorithms. The title could be modified accordingly. It is like “Results” chapter in many of the papers.

5) For the attraction of readers, the chapter 2.2 could contain a flow diagram explaining basic steps of the data flow and calculations within the algorithm.

6) The sub-chapter ”2.2.1 Generation of brightness temperatures” starts from corrections to the basic algorithm. I recommend revision that the presentation start with a description how brightness temperature is generally calculated. Then details could follow.

7) The “Gkj correction” (line 132) has to be explained in present MS.

8) “Half first Stokes” parameter (line 167, Fig. 2 caption) has to be explained.

9) The statement “For very diluted solutions, the conductivity depends almost linearly on the salinity” is clear without Fig. 3. The figure could be omitted. For clarity, there could be a reference to the oceanic algorithms how salinity is calculated from the CTD data (EOS-80, TEOS-10).
10) Significant downward spike of TB around SSS=-7 in Fig. 2b has to be mentioned and interpreted. Negative values of SSS should be explained on the figure as well.

11) The title ”Definition of a SMOS-based climatology“ (line 189) is not clear in the context of the algorithm. Perhaps it is ”Estimation of SSS systematic errors of SMOS with respect to reanalysis“. If not, some rewriting could be useful in order to improve clarity of the title and the text. ”SMOS-based climatology (denoted as sssclim)” is defined only at the end of the sub-section (line 218), until that the reader is unclear what is meant under the term in the title of the sub-section.

12) The statement ”non-expected spatial gradient appeared close to the coasts“ (line 200) has to be explained. What was the situation, (a) high gradients found in the SMOS data that reflect real coastal dynamics, as seen by in situ data and reanalysis, or (b) high gradients not corresponding to the real gradients, (c) something else?

13) Caption of Fig. 4 should indicate that the monthly mean differences are presented for the year 2013. Why this year has been selected, some reason should be given in the text.

14) The statement ”To avoid lack of statistics“ (line 209) is not clear. In addition, the content and need for Table 1 could be better explained.

15) It is not clear, why the constant exclusion criteria (”larger than 2 psu are also discarded“, line 250) was applied for SMOS SSS anomaly instead of location-dependent variable criteria. The SSS variance in the Baltic Sea (can be simply determined from CMEMS reanalysis) is rather variable, as can be seen from the studies of fronts in the Baltic Sea. For example, the Baltic Proper is much more homogeneous regarding SSS than the Belt Sea and the Gulf of Finland, and the strait areas to the Gulf of Bothnia and the Gulf of Riga. Therefore, it seems that the Baltic Proper could have much smaller exclusion criteria than regions of high SSS variability. A more detailed reasoning could be useful.

16) The statement in line 240 ”Any raw SSS out of the range [-150, 100] psu is not considered as part of the valid raw SSS values“ does not follow the oceanographic point of view. In my understanding, ”raw“ data mean almost real data but they need to be filtered and corrected. If the defined range of raw SSS is completely out of range of the real sea surface salinity in the Baltic Sea, the quantity should have another name than ”raw SSS“, even if this term has been published for other seas with higher salinity.

To clarify, when studying the living room temperature, if some intermediate result is 300 degrees Celsius, I would not call this as raw temperature but something else.
17) The filtering criteria (lines 239-255) should give some examples of bad and good data.

18) It is not clear why the oceanic limits [0, 35] (line 253) are used in the Baltic Sea for the exclusion criteria since in the major part of the sea SSS will never exceed 10 psu. The problems with dielectric constant models (they have been tested on oceanic salinity range) were mentioned earlier in lines 52-54. Is the criteria in line 253 related to this or is there some other reasoning?

19) In the title “2.2.7 Mitigation of time-dependent biases” (line 267) the term “mitigation” is not in the right context, see https://www.britannica.com/dictionary/mitigating or similar.

20) The statement “However, due to the lack of Argo floats...” (line 277) should be rephrased, since Argo floats are used in the Baltic Sea (https://www.euro-argo.eu/News-Meetings/News/News-archives/2019/Argo-floats-complement-the-Baltic-Sea-observation-network).

21) For assessing the temporal corrections to SMOS salinity, in situ measurements are taken from SeaDataNet (lines 278-279), which, to my knowledge, is not collecting and disseminating the whole set of Baltic-wide FerryBox measurements in operational time scales. Why this extensive valuable data set is excluded here?

22) A spike in Fig. 8 must be explained. Is it due to the problems of in situ data or satellite retrieval? Or is it due to the fragmentary spatial distribution of in situ data compared to the regular reanalysis data?

23) In the section “3.1.2 FerryBox lines in situ salinity” the source of data has to be specified. It is not clear, from where the data quality flag definitions are taken from. The flag PSAL_QC is used only on line 340, perhaps to omit the notation.

24) The Table 2 presenting the FerryBox ship routes and periods is not complete. The data are collected and disseminated by CMEMS. Comparing with my downloads, data from Baltic Princess and MS Romantika are not included.

25) The title “3.1.3 SeaDataNet in situ salinity” is not correct, since two data collections are used: SeaDataNet and ICES.
26) It could be interesting to read (perhaps in 3.4.4 Description of salinity dynamics), what could be the reasons of SMOS SSS overestimation relative to FerryBox results (positive bias) in the Gulf of Bothnia in years 2012 and 2017, both in the products of L3 (Fig. 13) and L4 (Fig. 16).

27) The tables 3-8 have similar structure and perhaps some of them could be combined. Besides yearly statistics, summary values for the whole period could be useful as well.

28) Technical issues with figures (legends, units, color scales) and abbreviations in the text etc should be corrected.

Additional remark. The MS uses extensively the term "climatology", with an interpretation as "climatological data" (https://community.wmo.int/wmo-climatological-normals). Climatology is “the description and scientific study of climate” (https://glossary.ametsoc.org/wiki/Climatology) or a “branch of the atmospheric sciences concerned with both the description of climate and the analysis of the causes of climatic differences and changes and their practical consequences” (https://www.britannica.com/science/climatology). I would prefer using the term "climatological data” and leave “climatology” for the classical approach as a branch of science, although data-oriented jargon as in the present MS is used sometimes in the papers dealing with technical aspects of oceanographic data processing and model development.