Answers to Reviewer #2:

This paper presents a new reprocessing of the BEC SMOS SSS global L3 and L4 products v2.0 for a 9-years period comprising 2011 to 2019, with an updated methodology with respect to previous BEC version. The updates concern new filtering criteria, a new empirical correction for latitudinal and seasonal biases, an improved scheme for preserving small-scale gradients close to the coast, and the provision of a sea surface salinity uncertainty. Three BEC products are presented : a high resolution product, a low resolution product (smoothed with a radius of 50km) and a L4 product produced using a multifractal fusion technique applied to SMOS L3 using OSTIA SST fields.

After the methodology presentation, evaluation exercises are performed :
-by comparing the new BEC products with Argo salinities over the whole time series
-by comparing the characteristics of the new BEC product relative to other satellite salinity products and to a CMEMS salinity product. These comparisons are restricted to the 2017 year. They concern not only comparisons to Argo but also comparisons of singularity exponents, of spectral properties and results obtained with triple collocations.

This is a huge piece of work by one of the two data centers delivering SMOS Level 3/4 products. Given the unique length of SMOS satellite salinity, it is a very important topic. The paper is clearly written and contains many interesting results. Nevertheless, I think several aspects would be worth to be deepened before publication, as this paper will serve as reference for future work:
Thank you very much for all your comments and suggestions. We think they have helped to improve the quality of the manuscript.

- The 2017 year is chosen to intercompare the various products because it is close to a normal year ('there is not any large-scale geophysical phenomenon'). But given that the BEC products are calibrated using a climatology and that the methodology makes use of many filterings, it would be important to evaluate the product skills relative to others during a non normal year, e.g. 2016.
We have included in the manuscript the comparison with Argo also for the year 2016. The other metrics are not expected to change very much with the selected year.

- The metrics used in the comparisons with Argo are mean, standard deviation and rms difference. They should also include R2 : indeed, the BEC methodology employs many filterings and R2 would allow to measure their efficiency in maintaining the SSS variability.
We have included R2 in the Tables containing the statistics with respect to Argo, they correspond to Tables A4 and A5 in the new version of the manuscript.

- By construction, the singularity exponents of the L4 product are expected to be

close to the ones of OSTIA SST so I don't think that comparing BEC L4 and OSTIA SST properties provides an independent validation of BEC L4 products.
True. We have clarified this in the text. See the answer to the corresponding points below.

Detailed remarks :
Abstract : not sure doi are useful in the abstract.
This was suggested by the editor.

Abstract : last statement (iii) should be moderated (see below my comments about the triple collocation section) : I don't think the authors have an absolute reference allowing a measure of the error. Given the huge spatial variability in regions affected by rainfall and by continental freshwater discharges, I am surprised that a low resolution product is more accurate.
We have moderated this sentence by including that these are the results obtained from triple collocation:
Lines 19-21:
the results from triple collocation show the BEC SMOS level 4 product as the product with the lowest estimated salinity error in most of the global ocean and the BEC SMOS high-resolution level 3 as the one with the lowest estimated salinity error in regions strongly affected by rainfall and continental freshwater discharges.

In the section of triple collocation we have added a discussion, as suggested by the reviewer below (see also our answer in the corresponding point).

L20 : verb is missing
Included.

L126 : radiometric sensitivities wording is confusing (see Randa et al (Recommended Terminology for Microwave Radiometry, National Institute of Standards and Technology Technical Note 1551 (August 2008) : Radiometric sensitivity is often used to mean radiometric resolution, but this use is discouraged in view of the definition of sensitivity).
I guess the authors mean radiometric resolution.
Thank you very much for providing us with this reference! We have changed the text accordingly.

Section 2.2.3 : It is useful to recall all these filterings. It might be useful to point out more specifically the changes with respect to Olmedo et al 2017 and to display maps of the number of filtered points.

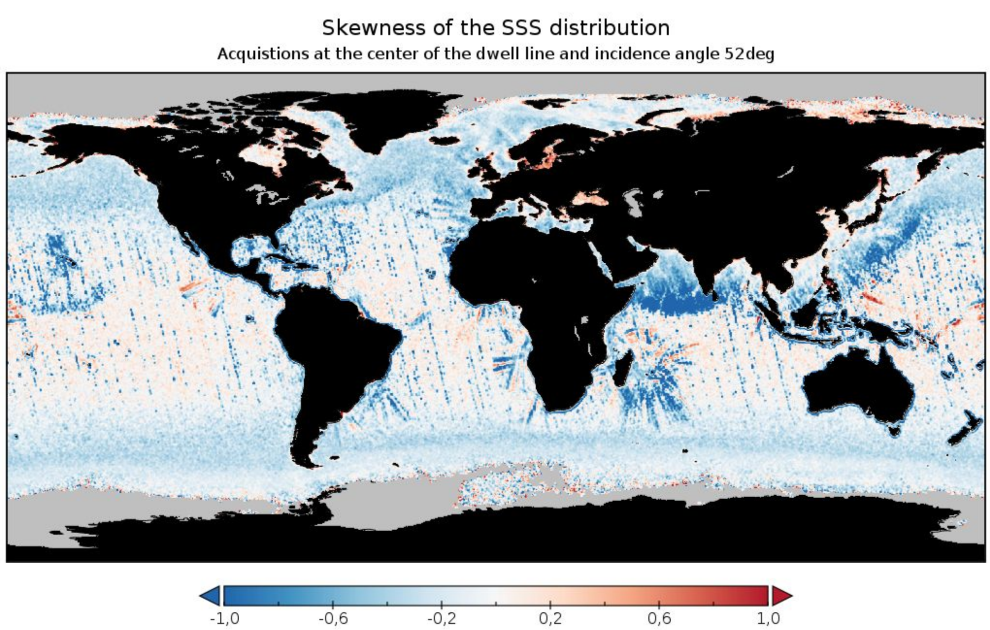
We have included in the text which are the differences with respect to the filtering criteria previously applied:
Lines 147-150:
These filtering criteria are the same as the ones introduced in (Olmedo et al.,

2017). The only difference is that now the criterion corresponding to the kurtosis is more relaxed: In (Olmedo et al., 2017) the set $\{s_n^{raw}(\gamma)\}$ was considered not valid and thus discarded out when the kurtosis of the distribution were larger than 4. Now we filter only platykurtotic distributions but not leptokurtotic ones.

Line 159:
This is new with respect to the criterion proposed in (Olmedo et al., 2017).

Line 166-170:
Generally, in the open ocean, $\sigma_{\varphi,\lambda}$ is small, so equation 2 is dominated by σ_{γ}. Therefore, in the open ocean, the new and the previous filtering criteria have similar performances. However, in those regions with strong salinity dynamics, such as coastal regions, $\sigma_{\varphi,\lambda}$ is not small and its contribution in equation 2 becomes dominant. Therefore, in those regions with strong salinity dynamics, the new filtering criteria is more relaxed and thus allows capturing better the salinity variability.

In particular, the distribution of the salinity is naturally skewed in rainy regions and in river plumes, what is the effect of such filterings in these regions ?
The main effect of this filtering criterion is to remove those acquisition conditions with historical time series regularly affected by RFI. In the following figure the values of the skewness for acquisition conditions at the center of the dwell line and at incidence angle values of 52 deg for ascending orbits are presented. The affectation of RFI close to the Euro-Asian coast is evident, besides the RFI tails close to Madagascar and the western coast of Africa. However, the reviewer is right in the point that the values of skewness lower than -1 close to the Northern Brazilian coast, are probably due to the advected plume of the Amazon by the North Brazilian current. We are currently revisiting this filtering criteria. In the regional products that we are developing in the framework of different ESA initiatives (Baltic+ Salinity dynamics and EO4SIBS projects) we are not applying any restriction on the skewness of the distributions. So, in future versions of the global product we will check the impact of not considering this filtering condition.



We have added a sentence to explain this in the manuscript:
Lines 150-157:
Regarding the impact of the filtering criterion corresponding to the skewness, this is the same as the one proposed in Olmedo et al. (2017). This criterion aims at discarding ocean regions affected by RFI
contamination. Although some geophysical events tend to be not symmetric and fresh, as continental discharge and ice melting, and this leads to negative skewed salinity distributions, the typical skewness in these cases is around -0.5. The skewness values lower than -1 correspond typically to distributions that are affected by non geophysical phenomenon. However, we continue revisiting this criterion and probably in the next version of the product we will analyze the impact of not including this criterion of the skewness.

Equation 4 corresponds to a one sigma sorting, which seems very stringent, what is the rationale for this choice ? how sensitive is the result to this ?
At this point, we have salinity values that have been previously debiased and we collocate all of them in 9-days and a rectangular grid. We accumulate all the acquisitions, the acquisitions that we have per each incidence angle and per each dwell line, during the 9 days at a given lat-lon. For this set of debiased salinity values, we compute the mean and then, we apply a one-sigma criterion with respect to this mean. We have included the following sentence in the text:
Lines 181-183:
This criterion was also applied in the previous version of the product. Since, at this step, the salinity retrievals are already debiased and they are temporally and spatially collocated, the criterion of one-sigma applied here is expected to reduce the noise of the level 3 salinity maps only.

Figure A2 : I suggest to add a figure displaying the polynomial correction.
We have included the polynomial in Figure A2, as well as the description of the figure:

Lines 216-218:
In the bottom plots of Figure A2 the monthly interpolating polynomials $p(m,\varphi)$ are presented (in blue), as well as the mean difference $\Delta s_1(m,\varphi,\lambda)$ (in green). As observed in the plots, the approach of this correction has some limitations at high latitudes, where the sea ice dynamics also induce ice-sea contamination.

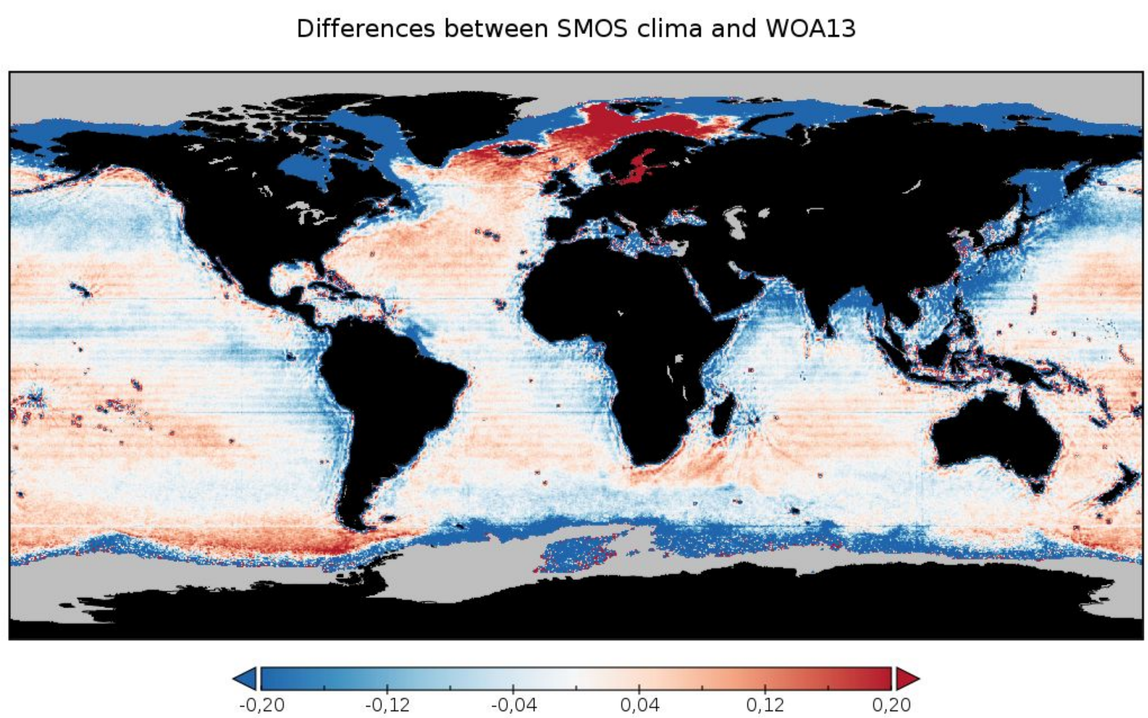
Section 2.2.6 : A major difficulty is how to filter SMOS retrieved salinity given that
salinity distribution is very skewed and that RFI contamination might lead to artificial skewed distribution too. In the updated methodology, more stringent filterings than in Olmedo et al. (2017) are applied.

The filtering criteria regarding the skewed distribution is the same as in Olmedo 2017. Only the filtering of kurtosis has been modified.

After having performed the serie of filterings and corrections, an inconsistency between the WOA reference and the mean corrected field appears in river plumes which is likely an effect of the skewed salinity distribution (Figure A3) but it seems that there is also a global north-south difference : could the authors refine the color scale of Figure A3 to allow a better display (e.g. with a 0.02psu resolution)?

In the open ocean differences with respect to the reference is about 0.04 psu.

We have investigated the causes of these discrepancies in the framework of the ESA regional initiatives Baltic+ Salinity Dynamics and EO4SIBS. In coastal regions, the main contribution to these discrepancies is because of the lat-lon resolution in which the climatological salinity distributions are computed. In the computation of the climatological representant that is used for the debiasing, we accumulate all the retrievals in a rectangular grid of 0.25ºx0.25º in lat-lon and we use the eight neighbours of the given gridpoint to increase the statistics and thus obtain the central estimator of the distribution more accurate. However, the systematic errors close to the coast change rapidly in the eight neighbours and then this procedure is introducing inaccuracies in the estimation of the central estimator. In the open ocean, the order of the discrepancies is much smaller. We think that the reason for these discrepancies is probably due to the numerical truncations of the histograms we use in the climatological distributions (they are computed with salinity bins of 0.5 psu).



L243-247 : the (Boutin et al., 2016, 2018; Kolodziejczyk et al., 2016) publications refer to earlier versions of CATDS CEC products than version 4. A summary of the version 4 updates with respect to earlier versions is available on https://www.catds.fr/Products/Available-products-from-CEC-OS/CEC-Locean-L3-Debiased-v4.

Thanks, we have also included this reference.

Section 3.1.3 : It is indicated line 407 that CMEMS product assimilates SMOS CATDS SSS observations. This could be indicated in this section.

We have included this information in this section:

Line 299:

This *product assimilates SMOS SSS generated and distributed by CATDS.*

L294 : Figure A11 is mentionned in the text just after Figure A3. Figures might be reordered.

Changed

L351 : hsst mentionned twice.

Corrected.

Table A1 : CMEMS product

We do not understand your point here.

L415 : see my general comment about making the validation for the 'normal' year 2017. What would be obtained for year 2016 ?

We have included a Table (Table A6 in the new version of the manuscript) with the comparison of all the satellite products and Argo for the year 2016. Although all the BEC products present an increase in the standard deviation of the differences, this increase is observed in all the regions and it is consistent with the results of Table A3 for the first six years of the time series (2011-2016). We do not associate this increase to el Niño event (because it is present in all the regions). We have included this in the text:

Lines 471-475:

We have also computed the statistics of the comparison with Argo and all the satellite products for the year 2016 (see Table A6 to assess the performance in one year in which "el Niño" event occurred. The results are similar to the ones obtained in 2017. The BEC products present a slight increase in the standard deviation of the differences with respect to Argo in all the regions that is consistent with the general increase observed at the first six years of the mission and that have been above mentioned.

L437-439 : I don't understand. CATDS CEC v4 fields use the combination of all ISAS fields over the 2012-2018 period to calibrate the full time series
(see

https://www.catds.fr/Products/Available-products-from-CEC-OS/CEC-Locean-L3-Debiased-v4: 'These relative salinity variations are then converted, in a last step, to salinities by adding a single constant determined, in each pixel, from SSS statistical distribution over the whole period (SMOS SSS distribution compared to ISAS SSS).
This last step only determines the absolute SSS calibration in each grid point; the SMOS SSS temporal variation is independent of this adjustment.' In other words, both BEC and CATDS are calibrated using a climatology. CATDS does not use ISAS SSS to perform an adjustment to a specific year.
We have removed this sentence.

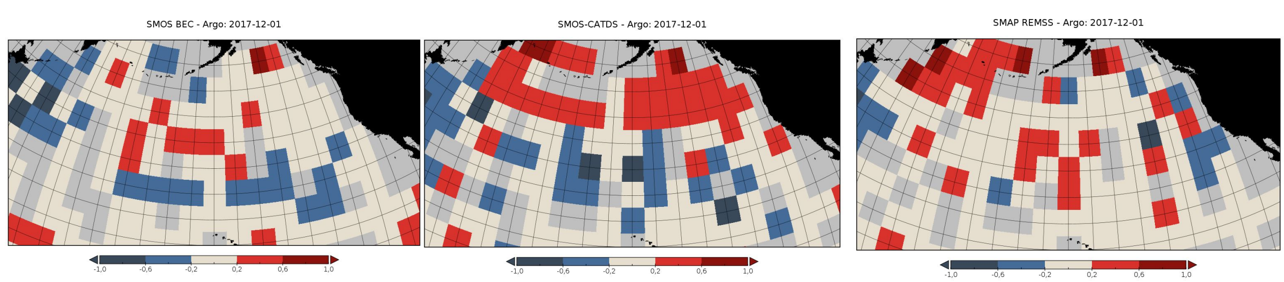
Figures A5 to A10 and Table 4: I recommend to add information about R2.
We have included the information about R2 in the Tables A4 and A5.

Figure A7 : Since the paper contains many figures, I suggest to remove Figure A7 which contains information in some way redundant with Figure A9. In case the authors prefer to keep Figure A7, it would be interesting, in addition to the longitudinal mean, to display standard deviation.
We have decided to keep Figure A7, because it validates one of the main novelties introduced in the new version of the product: the latitudinal correction.
We have not included in the paper the plots corresponding to the standard deviation because of the excessive number of Figures in the paper. However, the information of the std could be extracted from Figure A10.

Line 447 : 'In the case of satellite products,' could be removed.
Done.

Line 463 : Remove 'ESA is funding'.
Done.

Line 478 : This -0.4psu difference observed in December with all satellite products is surprising : is it homogeneously spread over the region or concentrated on a few grid points?
The largest difference appears on the first of December. We have plotted spatially the differences between the different satellite products and Argo:



The largest negative differences that dominate the spatial average reflected in Figure A.10 are not concentrated in a few gridpoints, they are located in the central Pacific (latitude 40-30N). In the beginning, we thought that the differences could be due to the contamination during eclipse periods, but the latitudes are too low. We still do not know which is the reason behind this negative bias observed in the different products.

Section 3.3.2 : The singularity analysis is interesting as it is a way to check how natural gradients are retained by the various processings. However, I think the 'best performance' obtained with BEC L4 is artificial because BEC L4 products are built assuming that SE of OSTIA SST and of BEC L4 SSS should be similar. Hence the results obtained with BEC L4 on Figures A12 and A13 are not independent validation (section 3.3 is validation) but rather a verification of the proper behaviour of the L4 construction methodology (which is good to show). Hence the results presentation needs to be revised. L508 'For instance' might be replaced by 'By construction'. In the following, the performances obtained with BEC L4 should not be compared with the ones obtained with other products, but they might be presented as an 'ideal' case.
We have modified the text accordingly with this suggestion:
Lines 539-545
BEC L4 (top right plot in Figure A12) presents a good correspondence between h_{SSS} and h_{SST} even at the most negative values of h_{SST}. This is expected because the BEC L4 product is computed from OSTIA SST by applying multifractal fusion. The method improves the quality of the salinity maps by using the spatial structures of the OSTIA SST. Therefore, the singularity exponents of the BEC L4 are expected to be close to the ones of the OSTIA SST. Therefore, among this section, BEC L4 is expected to provide the best performance.

OSTIA SST is taken as a reference but OSTIA SST is not perfect. I guess SST SE could be different if derived with another SST field. This could be discussed a bit and maybe an order of magnitude of the uncertainty associated with OSTIA uncertainty could be derived?
The reviewer is right and it is a very interesting point. However, to fully answer this point a similar analysis to the one done in this section should be done for all the available SST products. Moreover, we should consider other ocean scalar as a reference, and honestly we do not know which could be used (ocean colour could be strongly affected by clouds, altimetric products have corser resolution than SST, as well as SSS…). We think this requires a huge piece of work that is out of the scope of this paper.
We have included the following sentence at the beginning of the section:
In this section we take the SE of OSTIA SST as a reference to assess the effective spatial resolution of the salinity products. OSTIA SST product is not perfect and it also has some limitations in describing the small spatial gradients of SST, which could be reflected in the results of this comparison.

Section 3.3.3 : what are the confidence levels on the spectra (Fig A14 & A15) and on their slopes ((Fig A16) ?
We have included in Fig A14 & A15 the information about the standard deviation of the set of individual spectra that have been used in the mean spectra represented in these figures.
The vertical line represented in the plots applies to all the frequencies in the log-log scale in which the spectra are represented. We have added the following sentence in the text:
Lines 569-571:
We have also included the information about the standard deviation of the set of individual spectra that have been used in the mean spectra represented in these figures. The vertical line represented in the plots applies to all the frequencies in the log log scale in which the spectra are represented.

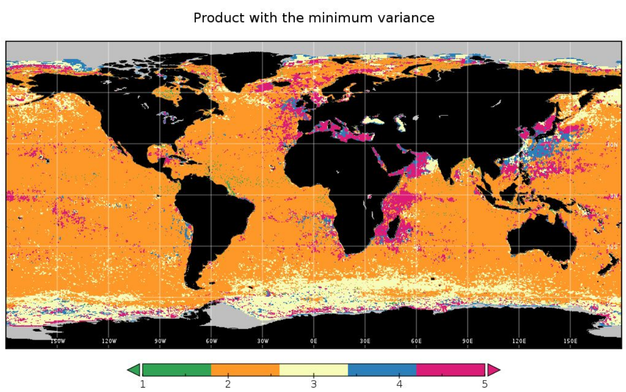
In rainy regions like ITCZ can we really expect PDS of SST and SSS to be similar ?
The reviewer is right that rain and other phenomena can produce different PDS for SSS and SST. However, in this section we are working with temporal and spatial scales in which this kind of phenomena are expected to have a negligible effect: we are averaging all the spectra available in 2017 and comparing the slopes in the spatial scales of 100-1000km. We have included the following sentence in the text:
Lines 573-575:
Although the effect of rain and other geophysical phenomena can lead to differences in the PDS for SSS and SST, in the spatial and temporal scales in which we have computed the slopes, these differences are expected to be negligible.

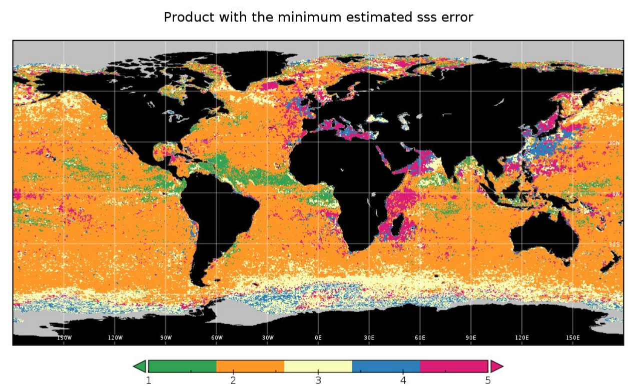
L559-560 : I don't understand the meaning of this sentence: SMOS and SMAP are not the same instrument.
We have removed this point. It was confusing.

L563-565 : isn't it by construction ?
We have included this sentence:
Lines 600-603:
That is partially because of the fusion method used in the generation of the BEC L4, which improves the salinity by introducing the spatial correlation consistency with respect to OSTIA SST.

Section 3.3.4: I am puzzled with the results (or their interpretation) obtained with the triple collocation. According to Figure A17 and A19, BEC products would have the lowest errors in very variable SSS areas with respect to other products. Why ? Could it be an effect of filtering ?
To clarify this point, we have made the following experiment. We have computed

the variance of the salinity in 2017 for all the L3 satellite products. In the figure below, we display a map with the product that presents the lowest salinity variance (left) and the product with the lowest estimated error from triple collocation (right). In the plots, 1 (green) corresponds to SMOS BEC HR, 2 (orange) SMOS BEC LR, 3 (yellow) SMOS CATDS, 4 (blue) SMAP JPL and 5 (pink) SMAP REMSS. Both maps are similar and the product with the lowest variance is also the one with the lowest estimated error in the most of the ocean. However, precisely in the Equatorial Atlantic, which is affected by the Amazon plume salinity dynamics and the Congo plume dynamics, both maps are providing different information. This is because the estimation of the salinity error is computed not only with a single triplet, but with several ones.





However, the reviewer is right that the computation of the product of minimum error implies several simplifications, that have to be taken into account in the interpretation of the results, We have added this in the text:

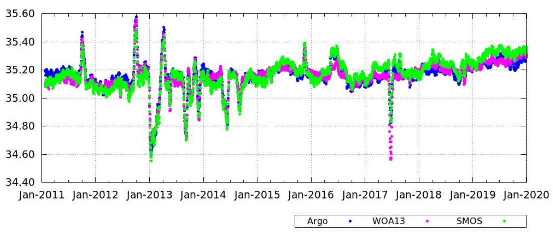
Lines 620-624:

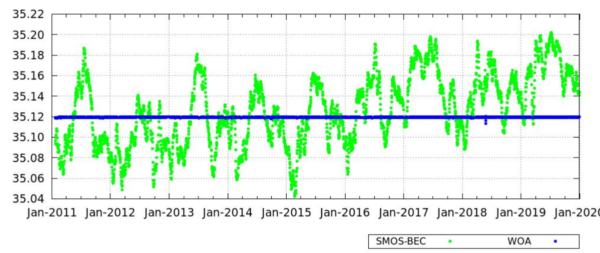
We have not included the uncertainty associated with the estimation of the salinity error in this computation. This implies that, although we represent one map with a single product only at each grid point, after considering the uncertainty of each estimation, several products could provide similar performance from a statistical point of view. In particular, the estimated error in regions of large salinity variability presents larger uncertainty than in regions of low salinity variability (see Figure A18). Therefore, the following maps are less accurate in these regions.

This result is surprising given the Argo comparisons presented in Table A4. For instance, from Table A4 in the Amazone region, the std difference and the RMS difference between satellite SSS and Argo salinity obtained with BEC products are the highest, which suggests a higher error with BEC products. This is opposite to what is indicated with Figure A17 and Figure A19. In the equatorial band, in Table A4, the REMSS products display the best comparisons to Argo which seems again opposite to the errrors displayed on Figures A17 and A19. The apparent contradiction between results obtained with triple colocation and Argo comparisons should be further explored before a conclusion on the error is drawn.

The results of the comparison with Argo floats, although they are informative and provide insights of the quality of the salinity products, they have also some limitations:

a) The Argo and satellite data sampling is different. In the following figure we can see how different is the salinity dynamics described only from information in the locations and times of Argo, with respect to the salinity described from the locations and times of available satellite data. The two graphics represent a daily average of the salinity in the region comprising between 30ºS and 30ºN. On the right, we display the mean salinity value computed using only the locations where Argo measurements are available. In the left plot the mean salinity value described by Argo is represented in blue, by SMOS BEC LR is represented in green and the mean provided by a constant annual reference (WOA13) in pink. From the left plot we could say that the three salinity fields describe quite well the salinity dynamics. However, when we look at the mean salinity obtained by averaging over the full region (the locations where satellite data is available) we can see that one of the fields is constant on time (WOA13) while the other one presents some variations: a seasonal behaviour and a trend in the last years.





b) Other limitations of the comparison with Argo is that the temporal and spatial scales that the satellite measurements represent (typically 7-9-day integrated measurements in an area of 25 km x 25 km (approx)) are not the same as the ones that the in situ measurement represent (they are instantaneous and punctual). This could lead to significant differences between satellite and in situ data (due to the intra pixel variability) that have not to be associated with an error of the satellite measurements.

c) Last, but not least, we also have to consider the differences in the depth of the acquisitions. Typically, Argo measurements are considered in 5 m-10 m, while satellite data is acquired at a few centimeters depth.

We have added the following discussion in the manuscript:

Lines 646-659:

The results obtained from triple collocation provide a complementary view to the comparisons with Argo floats (see section 3.3.1). Although the comparison between in situ and satellite data provides very valuable information about the quality of the satellite products, this comparison has several limitations that triple collocation method has not:

–Sampling: the in situ measurements are provided over few samples while the satellite data is synoptic. The dynamics displayed by the in situ measurement could be strongly conditioned by its sampling. Therefore, the results from the comparison could not be completely representative of the quality of the satellite

product in the considered region.
–The spatial and temporal scales of the in situ and satellite measurements are different: The in situ measurements provide punctual and instantaneous measurements while satellite measurements correspond to an integrated measure of several days and a footprint of several squared kilometers.
–The in situ measurements are typically given at several meters depth while the satellite data are providing the salinity in few centimeters depth.
The above mentioned reasons could lead to apparent contradictory results between the results corresponding with Argo comparison and the results corresponding with triple collocation. Indeed, both comparisons constitute different metrics that are providing different pieces of a complex puzzle.

L584 : a verb is missing
Corrected

L617-623 : a caveat for very variable regions (Amazone) should be added.
We have already included several caveats in the corresponding previous sections.

L626-628 : I expect this to be by construction (see previous comments)
We have included the following sentence:
Lines 685-686:
This is partially because of the multifractal fusion method used in the generation of the BEC L4 product.

On the figures displaying maps (A3, A5, A6, A17, A18, A19), latitude and longitude
labels are missing.
We have included the corresponding labels in latitude and longitude.

Figure A13 : The y axis of last figure should be std(SSS).
Done

Figure A13-A16 (in particular A14-A15) : colors are hard to see (especially yellow)
Changed

Table A2 : I guess Indic should be Indian ? (same remark for Figures A8-A10)
All corrected