Response to Referee 3:

We are greatly appreciated for your comments and constructive suggestions which are quite important to further improve our paper from quality to scientific dissemination. In the following text, we will answer all the questions or comments (in italic with black color) one by one in blue color.

General Evaluation:

This work presents a good assessment of the usefulness of SMOS sea surface salinity data in Arctic Ocean modeling. The control vs SSS assimilation experimental design is valid, statistical error diagnostic is standard and the evaluation is well done by comparing with independent source of observations. Discussion is also informative and interesting. In general, I found it being a nice piece of research without major flaws. However, the following attentions need to be paid for improving the manuscript. 1) there are a few mistakes in the equations, although it seems that the authors did the correct diagnostics according to the figures presented.

-**A**: Sorry for the related errors in Eqs. like Eq. 5 missing the sqrt. Further corrects in the Eqs will follow the common mathematical rules. For instance, Eqs. 4 and 5 will be corrected as:

$$Bias = \frac{1}{\sum_{i=1}^{N} O_i} \sum_{i=1}^{N} \sum_{1}^{O_i} (H\overline{X}_i - y_i) \qquad (4),$$

$$RMSD = \sqrt{\frac{1}{\sum_{i=1}^{N} O_i} \sum_{i=1}^{N} \sum_{1}^{O_i} (H\overline{X}_i - y_i)^2} \qquad (5),$$

Where i is the ith day, O_i represents the number of observations on this day, and N represents the total number of days depending on the source of observations. Then Xi represents the model daily average at the observation time as the ensemble mean of 100 model members. H is an operator to extract the SSS simulation from the model at the observed location.

2) some acronyms, data and analysis method are not clearly defined in the text - please see my specific comments.

-**A**: Thanks for this suggestion. The consistency in text will be checked more strictly in the revision, especially to the related comments.

3) the English writing must be improved as some parts of the article reads awkward and confusing which makes it hard for readers to follow.

-**A**: Thanks for this suggestion. We will improve the language more fluent and ask a native speaker to proofread the revision.

Other suggestions about making changes to figures and writing are included in the annotated document. Overall, I would suggest a minor revision recommendation for this article.

-**A**: We work on these comments and give one-by-one responses as followed.

<u>*Line by line comments from PDF file :*</u>

*Line 8-9: "*two assimilation runs assimilated two successive versions of the SMOS SSS product,*"* this sentence is not very clearly written. please rephrase and expand if necessary.
-A: It is changed into "Using the Deterministic Ensemble Kalman filter from July to December 2016, two assimilation runs respectively assimilated two successive versions of the SMOS SSS product, on top of a pre-existing reanalysis run. "

Line 17: "FWC" full acronym
-A: It is Freshwater Content (FWC) and its calculation is defined by Eq. 6.

Line 60: "(Yueh et al., 2001) (e.g, the sensitivity drops from 0.5 to 0.3 K PSU^{-1} when the sea surface temperature decreases from 15 to 5°C);*"* no need a separate parenthesis
 -A: It is changed into "...larger SSS error (Yueh et al., 2001; e.g, the sensitivity drops from 0.5 to 0.3 K PSU^{-1} when sea surface temperature decreases from 15 to 5°C); "

Line 73: "six" what are the six products?
-A: Thanks for pointing this out, it is deleted and changed to as follows:

Xie et al. (2019) evaluated the V2.0 SSS product and another gridded Arctic SMOS SSS product developed by LOCEAN (Boutin et al., 2018) during the years 2011-2013. These two SSS observations, together with an Arctic reanalysis (Xie et al., 2017) and one objective analysis product (Verbrugge et al., 2018), were validated against in-situ observations and compared with two climatology datasets: the World Ocean Atlas of 2013 (WOA2013; *ref.*,

Zweng et al., 2013) and the Polar science center Hydrographic Climatology (PHC 3.0; *ref.*, Steele et al., 2001).

Line 105: "TOPAZ" define full acronym
-A: Thanks for this comment, the system name has been branded as TOPAZ for more than 20 years and the reference to the original acronym is no longer used.

Line 106: "consistently" what does consistently mean technically?
-A: One of advantages using EnKF is the multivariable adjustment even assimilates one type of observation in which other related model variables can be updated depending on the error covariance relationships. To avoid possible misunderstanding, it is replaced by "simultaneously".

*Line 153: "*discrepancies were a decreasing function of salinity*"* not a clear expression. please rephrase
-A: Thanks for this comment. The related sentences are rephrased as "In a recent study, Xie et al. (2019) evaluated the SMOS-based SSS products using in-situ observations and revealed strong regional dependence for the V2.0 product errors: smaller than 0.4 psu in the Northern Atlantic but increasing dramatically to 1 psu in the Nordic seas and over 2 psu in the central Arctic. Undoubtedly, the salinity observation errors from Passive Microwaves are higher in high latitudes than elsewhere. Furthermore, in the Beaufort Sea (as Fig. 12a in Xie et al., 2019), the error of the SSS V2.0 product and the Arctic reanalysis product from TOPAZ (same as Exp0 used in this study) both show an inverse relationship between SSS values and SSS errors. "

Line 170: "in situ" all "in situ" in this manuscript should be in italic form
-A: All are corrected with italic in the text.

Line 230: "RMSD" incorrect equation. this should be a root mean.
-A: Yes, it is wrong and corrected.

Line 246: "(Eq. 1)" should be equation 4
-A: To avoid this misunderstanding, more explanation is added as "but the mean of SSS innovation, calculated as the observation minus the model simulation (cf. the bracket in Eq.1), shows the saline bias of 0.4 psu, highest in September."

Line 257: "Figure 3" it's hard to observe the difference between the experiments based on Fig. 3 which shows total SSS value. to facilitate comparison, one should show difference maps instead.

-A: It is also recommended by other referees so this figure has been partly replaced as follows.
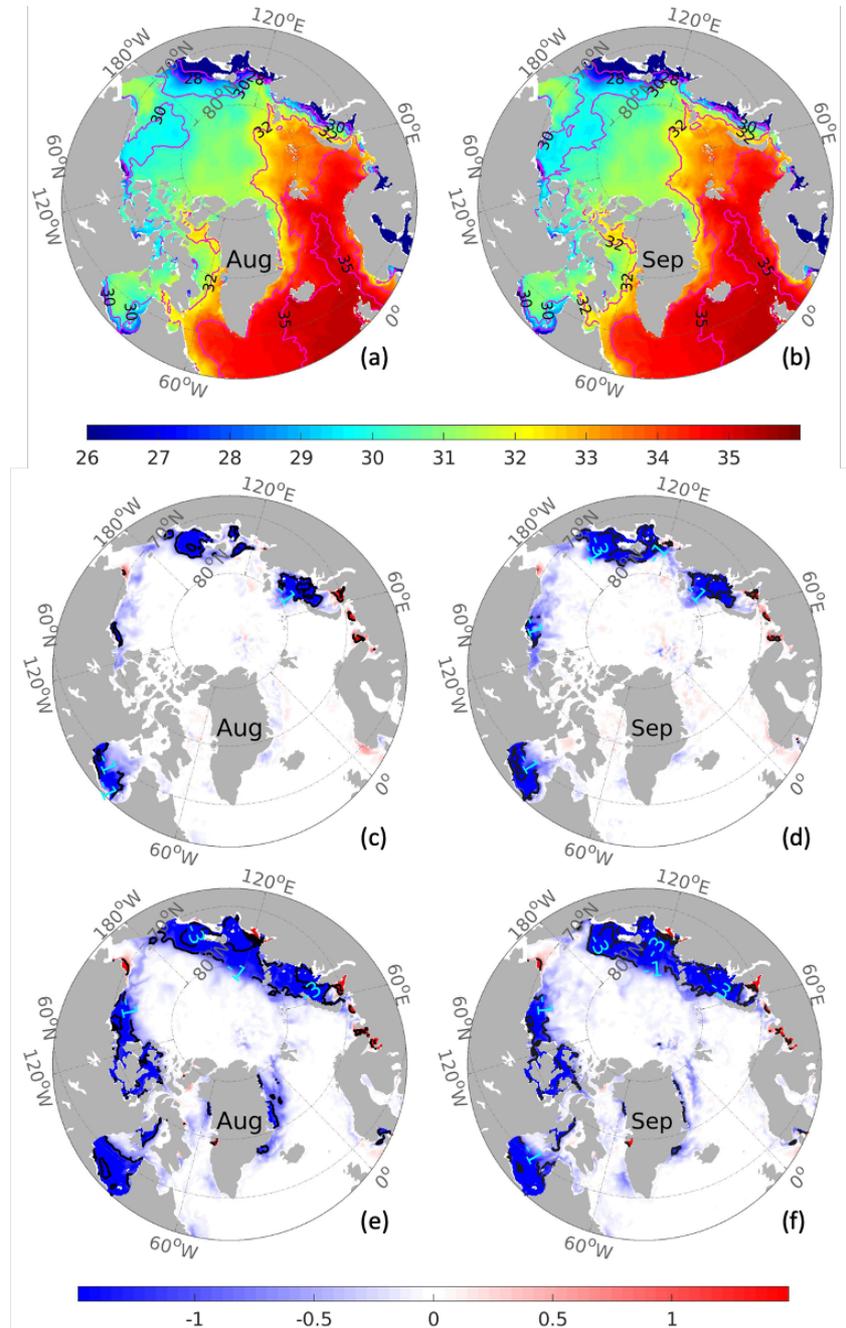

Fig. A1 Top: Monthly simulated SSS (unit: psu) from Exp0 in August (left column) and September 2016. The black isolines indicate the 26, 28, 30, 32, 34, and 35 psu isolines. **Middle and bottom**: the monthly SSS differences in ExpV2 (middle line) and ExpV3 (bottom line) with respect to that in Exp0. The black lines are -3, -1, 1, and 3 psu for SSS.

4

*Line 337-338: "*which suggests that the V2.0 SSS product loses the benefit around there by DA in this system.*"* is it due to the number of obs asimilated is small? or simply the data quality is not good? any further explanation?

-A: it is due to its quality limit around Greenland Island. For example, no more innovative information can be dug by the DA system if compared to the climatology of SSS.

*Line 344: "*the fresh Arctic water and the fresh coast water converge with the saltier Atlantic Water.*"* is this the reason why Expv2 and EvpV3 have no improvement?

-A: Due to this dynamic complexity, the SSS quality requires the resolution of the concerned SMOS SSS products both on space and time. The relative fine resolution of the V3.1 verifies its advantage compared to V2.0 here.

*Line 349-353: "*As indicated from SSS scatterplots of the three runs in BB (S6 in Table 1, also shown in bottom panels of Fig. 7), …, even has no significant reduction of the RMSD in GS.*"* what caused the unsatisfactory performance of ExpV2 and ExpV3 is still not clearly discussed.

-A: It may be because the SSS RMSD has been close to the smallest around 1.2 psu (as shown in Table 2), which seemly indicates the level of SMOS SSS uncertainty at high latitude areas at the current approaches. The SSS scatterplots in BB clearly show the simulated SSS between 30-32 psu in the model contains larger uncertainty compared to the salty SSS. Consequently, near the coastal regions like in the BB and the Nordic, the retrieving process of the V3.1 SSS product still has no big difference compared to the previous V2.0.

Line 408: in Eq. 6, the acronym is confusing. what does L mean? please define.

-A: Thanks for pointing this out, it is replaced by FWC.

*Line 434-437: "*However, so far with the limited amount of in-situ data, it is not fair to conclude whether this is a change for the better or the worse. Significantly different from sparse in-situ observations in the Arctic, the reanalysis product can better represent the characteristics of FWC variations in space and time.*",* these two sentences read like contradicting one another. do you mean the changes due to DA do not guarantee an improvement, but it might be useful for describing spatiotemporal variation of the FWC?

-A: Not exactly. Thanks for this comment. The FWC figure is changed by adding the difference between ExpV2/ExpV3 and Exp0 so the related sentence will be rephrased in the revision.

Line 493: "other version of SSS product"

-A: It is corrected by "the V2.0 SSS product"

Line 505: "indexes" is this the same as "overall score" in Table 2? how do you define this index? this should be included in the methodology section. The plural form of index is indices, not indexes

-A: Right, the basic idea is trying to help us to distinguish the affected areas. So it can be replaced by "overall score" to avoid too many definition equations.

*Table 2: "*Overall score*"* same as the impact index mentioned in discussion?

-A: It is a good idea to use "overall score".

Figure 3: suggest to keep Exp0 maps as they are but show difference maps for Expv2 and v3

-A: Thanks for this suggestion. This figure was changed as the above mentioned.