Author's response:
We thank Reviewer 1 for her/his dedicated comments. We have revised the manuscript accordingly and we hope the current version reaches the high standard expected. We respond point-by-point to the Reviewer's remarks in the following. Our replies are written in blue. Line numbers mentioned in this document are from the revised manuscript with track changes

Review of the manuscript "Novel Arctic sea ice data assimilation combining ensemble Kalman filter with a Lagrangian sea ice model" by Sukun Cheng et al.

The paper presents an experience of implementing a Kalman-type ensemble-based filter to combine sea ice concentration (SIC) and thickness (SIT) observational information with a Lagrangian sea ice model. In particular, the authors assimilate SIC from the Ocean and Sea Ice Satellite Application Facility (OSI-SAF) and the merged SIT product from CryoSat-2 and SMOS satellite missions into the Lagrangian sea ice model neXtSIM with the deterministic Ensemble Kalman filter (DEnKF). The filter analysis is performed on the ensemble of Lagrangian model states individually interpolated to a reference grid. The updated states are projected back onto the temporarily variable model mesh to reinitialize the model for the next forecast phase. The sea ice forecasting system is evaluated for the Arctic Ocean over the 2019/2020 winter time period. The OSI-SAF sea ice drift (SID) observations are used as independent information for the evaluation, additionally, to the assimilated OSI-SAF SIC and CS2SMOS SIT data. The subject of the paper is well within the frames of the journal. Generally, the paper is well structured and detailed, and clearly written; the figures are of a good quality; the method used is well justified. However, I have got few comments (*e.g.* on the system settings), which the authors might still want to clarify and further discuss in the manuscript before publishing.

Specific comments:

1) Abstract (Lines 13 -14): Please consider rephrasing the last sentence of the Abstract, I am not sure it can be stated in such a generalized context that the "model ... demonstrates comparable skills to operational forecasting models that use DA", since there was no explicit comparison to "operational DA forecasting models" carried out in this study and discussed in the paper (except for references to TOPAZ system).
We agree with the Reviewer that comparisons with operational DA forecasting models are not quantitatively carried out in the manuscript. This was concluded based on a qualitative comparison with the TOPAZ4 output which we are familiar with.
Following the Reviewer's suggestion, we rephrase the last sentence in the abstract.

2) I am a bit concerned about the definition of 'bias' in line 320. In line 320 the bias is defined as "model-minus-observations $d(t)=H(x(t)) - y(t)$", while "$d(t)=H(x(t)) - y(t)$" is indeed innovation (a difference between modeled and observed variable). The bias, from the statistical point of view, is an analyzed systematic feature of the innovation after averaging spatially or/and temporally as shown, for instance, in Figure 2b and presented in Table 2.
Thank you for pointing out this potential confusion.

Our definition of bias follows from Williams et al. (2021). As we always use bias as spatial or temporal means, we add an average operator in our definition. See line 380 in the revised manuscript.

3) Equation 2: Please double check whether the formulated in the equation is correct and whether it is what has been implemented in the study to approximate the SIT uncertainties. Given equation 2 the observational error variance is a discontinues function of sea ice thickness (SIT, h_{ice}): with too (unrealistically) strong increase with h_{ice} for the h_{ice} less than 3 m and saturated small (too small?) values for h_{ice} larger than 3 m (see Figure R1).

We apologise but there was a typo in the equation for the condition of thickness < 3m. We are thankful to the Reviewer for noticing it. The typo has now been corrected by adding a negative sign in the exponent in the second term.

$$\sigma^2_{\text{obs,SIT}} = \begin{cases} \min(0.2, 0.02e^{1.8(h_{ice}-3)}), & h_{ice} > 3m, \\ \max(0.02, 0.1e^{-1.5h_{ice}}), & \text{otherwise.} \end{cases} \quad (2)$$

4) Inflation (Lines 251 - 255). Necessity of inflation was emphasized also in other studies dealing with real sea ice thickness observations. Especially, it was required when no forcing perturbation was used. (More references could be added). Please elaborate a bit more on this ("Inflation") step of the data assimilation: if/how it relates to forcing perturbation; how the regular inflation within DEnKF works; and why it was additionally required to increase by factor of two the observational variance (it means all the assumed data uncertainties (Eq.1 and Eq.2) were further increased).

Prompted by the Reviewer's suggestion, we address the inflation step more extensively in the revised section 4.5.2 (lines 290-303).

a. Are there any other arguments to increase the assumed observational errors? Representation error? Possible misrepresentation of observational errors by Eq. 1 and Eq. 2 (Figure R1 and Figure R2a)?

Observational errors generally include measurement (instrument) errors and representation errors. The representation errors cover several error sources including unresolved scales and processes, observation-operator errors, and quality-control errors (Janjic et al., 2017).

However, the sea ice products we used only provide parts of the observation errors, which are not sufficient to represent the entire observational errors. In particular, CS2SMOS thickness data provides errors in the merging and interpolation of CryoSat-2 and SMOS products. While the OSISAF ice concentration product provides a concentration algorithm and tie-point uncertainties and smearing uncertainty due to satellite footprint mismatches. This situation motivates the increase of observational errors to properly represent the associated uncertainty.

We quoted a discussion from William et al. (2021) below as a reference: "The error levels in the CS2-SMOS product are only the interpolation error and are thus a lower bound as they do not include uncertainties in the individual CS2 and SMOS products. CS2 in particular is sensitive to the ice and snow densities used or the snow thickness which affect the conversion from freeboard to thickness (Zygmuntowska et al., 2014)."

b. Whether the finally considered SIC uncertainty (as a result of doubled SIC observational variance, Figure R2B) is not too large to properly constraint the model if the observed SIC is 0.5±0.2; could it be one of the reasons of "moderate extent" of the SIC improvement?

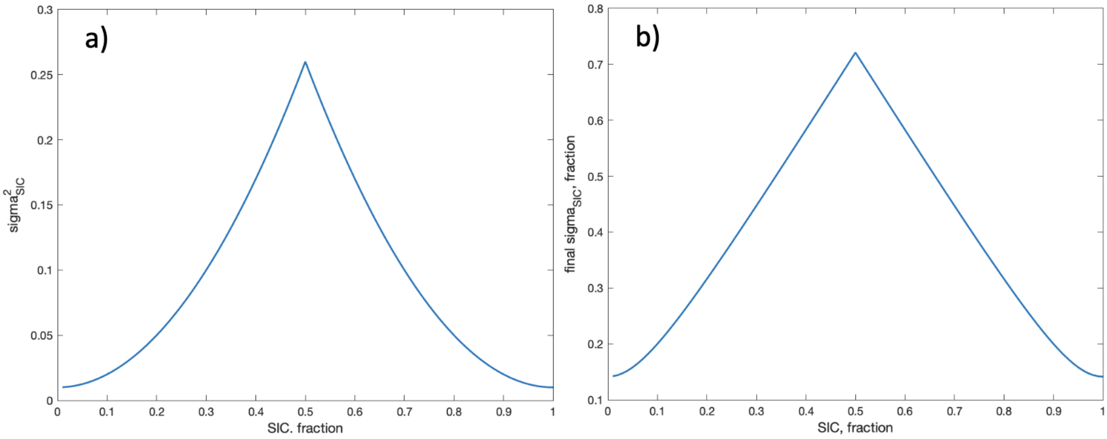

Figure R2: Assumed SIC observational variance as a function of SIC, reconstructed given equation 1 (a) and final SIC uncertainty as a result of doubled SIC observational variance (b).

This is possible, but not very likely in our case. The spatially averaged sea ice concentration is above 0.75 around mid-November 2019 in Fig. 2(a), which has a smaller observation error compared to the start of November as shown in Fig. R2. With higher sea ice concentration and lower observation uncertainties, the improvements get less significant. This suggests that, instead of misspecified observation uncertainties, the moderate improvements are related to the high sea ice concentration where the majority of the Arctic ocean is covered by ice without much need for improvements by infrequent assimilations where the uncertainties in the forcing propagate to the model forecast.

c. Were there any sensitivity experiments carried out with respect to the original ensemble spread due to perturbation of the atmospheric and oceanic forcing and the internal model parameter?

The sensitivity experiments on the impact of perturbation of the atmospheric forcing and the internal model parameter (ice cohesion) are reported by Cheng et al. (2020). Those experiments have indeed guided and motivated the experimental setup in this current work.

Minor comments

Line 161: why 2.5 km/2days not 1.25 km/day, could it be better to convert to and use m/s units

Line 483, 485: similar comment on the "km/2days" used as units for velocity while m/s is used few lines above. I understand that the authors would like to refer somehow to the decorrelation time scale, nevertheless, I still think that m/s would be a more meaningful unit.

In this study, we evaluate the model data on the observation spaces rather than the other way around. Presenting the ice drift data km/2days is to be consistent with the OSISAF ice drift observations settings for comparison. According to the OSI-SAF ice drift introduction https://osisaf-hl.met.no/osi-405-c-desc, the time span of the OSISAF ice drift observations is

48 hours or 2 days. It is the time delay between the start time and the stop time of the ice motion described by one vector.
More importantly, the ice drift velocity varies over time, thus its displacement is nonlinear over 2 days. Averaging the quantities over unit time could cover this fact and cause potentially missing interpolation in the evaluation.

Typo/misprints

Line 87: citation format issue – missing reference fixed

Line 92: citation format issue – missing reference fixed

Line 143: version 2o3 It seems fine to use CS2SMOS version 203 instead of 2o3, according to https://spaces.awi.de/display/CS2SMOS/CryoSat-SMOS+Merged+Sea+Ice+Thickness

Line 278: a space required after the dot in "run. Especially" fixed

Reference

Anderson, J. L., & Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Monthly weather review, 127(12), 2741-2758.

Anderson, J. L. (2007). Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. Physica D: Nonlinear Phenomena, 230(1-2), 99-111.

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). The data assimilation research testbed: A community facility. Bulletin of the American Meteorological Society, 90(9), 1283-1296.

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., & Korablev, A. (2012). TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic. Ocean Science, 8(4), 633-656.

Zygmuntowska, M., Rampal, P., Ivanova, N., & Smedsrud, L. H. (2014). Uncertainties in Arctic sea ice thickness and volume: new estimates and implications for trends. The Cryosphere, 8(2), 705-720.

Xie, J., Bertino, L., Counillon, F., Lisæter, K. A., & Sakov, P. (2017). Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013. Ocean Science, 13(1), 123-144.

Janjić, T., Bormann, N., Bocquet, M., Carton, J. A., Cohn, S. E., Dance, S. L., ... & Weston, P. (2018). On the representation error in data assimilation. Quarterly Journal of the Royal Meteorological Society, 144(713), 1257-1278.

Cheng, S., Aydoğdu, A., Rampal, P., Carrassi, A., & Bertino, L. (2020, December). Probabilistic forecasts of sea ice trajectories in the Arctic: impact of uncertainties in surface wind and ice cohesion. In Oceans (Vol. 1, No. 4, pp. 326-342). MDPI.

Williams, T., Korosov, A., Rampal, P., & Ólason, E. (2021). Presentation and evaluation of the Arctic sea ice forecasting system neXtSIM-F. The Cryosphere, 15(7), 3207-3227.