

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).
- 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.
- 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).

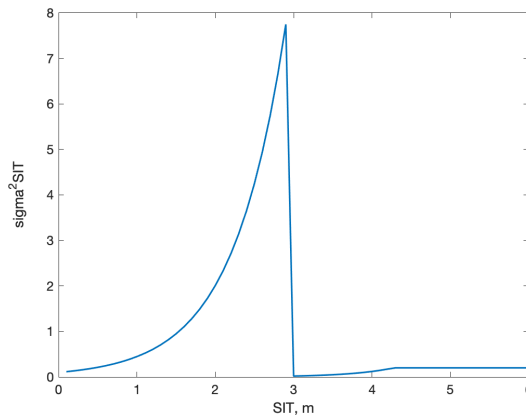


Figure R1: The assumed SIT observational variance as function of SIT reconstructed given Equation 2.

- 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).
- 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)?
  - 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  $0.5 \pm 0.2$ ; could it be one of the reasons of “moderate extent” of the SIC improvement?

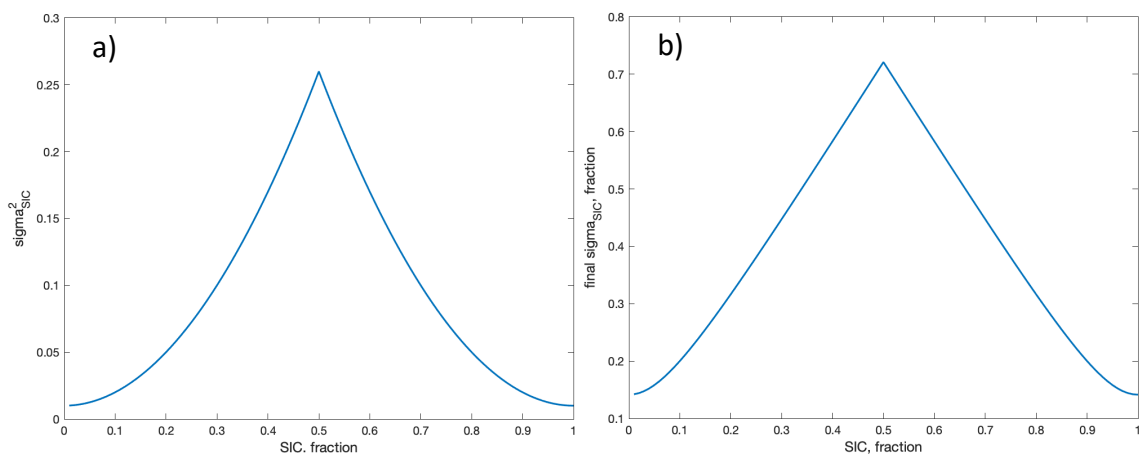


Figure R2: Assumed SIC observational variance as a function of SIC, reconstructed given equation 1 (a) and final SIC uncertainty as a result of doubled SIC observational variance (b).

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

## 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.

### **Typo/misprints**

Line 87: citation format issue – missing reference

Line 92: citation format issue – missing reference

Line 143: version 203

Line 278: a space required after the dot in “run.Especially”