Review for “Barents-2.5km v2.0: An operational data-assimilative coupled ocean and sea ice ensemble prediction model for the Barents Sea and Svalbard”

Recommendations: Major Revision

General Comments and summary

This paper presents the version 2.0 of the operational ocean and sea-ice forecast Barents-2.5km model. This version includes an ensemble prediction system (EPS) with an off-line ensemble-based data assimilation (DA) component. The system routinely assimilates sea ice concentration (SIC), sea surface temperature (SST), and in-situ hydrography observations. With the DA component, the Barents-2.5 km model shows better SST forecast skills, e.g., improvement over the persistence forecast during spring and summer. Although the predictive skill for SIC is not as high, the model still displays skillful performance with DA, as it reduces the model drift away from the truth state. Furthermore, the EPS also provides the uncertainty estimate of the model state. The ensemble spread for SST is generally reasonable, although it may miss some extreme values, whereas the ensemble spread for SIC is too small. Overall, the Barents-2.5km model with its EPS and DA component is a valuable tool for forecasting ocean and sea-ice conditions.

After carefully reviewing the paper, I am impressed with the technical details presented. The work is certainly worthy of publication in GMD. However, I have some concerns with the DA part. In particular, some of the context regarding the ensemble Kalman filter (section 3.1) appears to be incorrect, and in my opinion, some important DA details seem to be missing. Although the paper does include some analysis on the performance of DA, the issue of non-Gaussianity, which can be especially important for sea ice observation, is not much addressed or discussed. While I appreciate the manuscript's primary focus on documenting and demonstrating technical expertise, it's equally important to ensure that the context of the DA component is accurately and completely presented. Therefore, I recommend a major revision of the DA section before publication in GMD.

Specific line-by-line comments

- Line 178

The citation here can be a little misleading, as the EnKF in (Evensen 1994; Burgers et al. 1998) are not usually referred to as the deterministic version of EnKF. I suggest putting the citation (Sakov and Oke 2008) here, and move the citations (Evensen 1994; Burgers et al. 1998) to line 180.

For the EnKF reference, in addition to (Evensen 1994; Burgers et al. 1998), I recommend that also include another reference (Houtekamer and Mitchell, 1998)

Houtekamer, P.L. and Mitchell, H.L. (1998) Data assimilation using an ensemble Kalman filter technique. Monthly Weather Review, 126, 796–811.

- Line 183-184

Although it is tangent to the main thread of the paragraph here, van Leeuwen (2020) notes that in the original stochastic EnKF, the perturbations should be added to the ensemble equivalence of the observation $H(x)$ instead of the observation y. This distinction becomes significant when the observation error is non-symmetric (e.g., skewed), which can have important implications for, e.g., bounded observations.

van Leeuwen, PJ. (2020) A consistent interpretation of the stochastic version of the Ensemble Kalman Filter. QJR Meteorol Soc., 146: 2815– 2825. https://doi.org/10.1002/qj.3819

- Line 185-190

Equation (2) is the stochastic version of EnKF, while this paper uses the deterministic version of EnKF. Using Equation (2) here can be confusing to the readers. Therefore, I suggest, e.g., replacing Equation (2) with the deterministic transform equation in Sakov and Oke (2008) and replacing this paragraph with a new one (or add a new one) for the deterministic EnKF in Sakov and Oke (2008).

- Line 197-198

Is this a reasonable assumption for the observations assimilated in this work? This assumption will introduce larger representation error to the observations that are taken at time points more distant from the analysis time. Although this issue is discussed in Section 6.4, I suggest adding one or two sentences commenting on this assumption here.

- Line 199-202

I suggest extending this paragraph by adding more DA details. Specifically,

(1) Including some details about what the “spread reduction factor” and the “global moderation factor” mean, and how they work.

(2) It seems that only horizontal localization is applied. Do the sea surface observations have impact on the state variables in the ocean (e.g., the ocean current at 30-m deep)?

(3) I suggest providing more information on how the observation errors are moderated, since it is one of the key part in DA system. I wonder if the adaptive tuning of the observation error could somehow partially compensate the issue of the time-dependent representation error for the observations.

(4) Are the observations the same type of variables as the model states in ROMS/CICE? i.e., are the observation operators just interpolating the model state to the observation location?

- Line 291

Why are the SST validated for these regions (defined in Fig. 6) separately? Are there any important implications from Fig. 7? I suggest adding one or two sentences briefly discussing the results from Fig. 7.

- Line 304

What kind of failures in DA are specifically referred to here?

- Line 347-349

I suggest including the reference(s) for the rank histogram, e.g., (Hamill 2001).

Hamill, T.M. (2001) Interpretation of rank histograms for verifying ensemble forecasts. Monthly Weather Review, 129, 550–560. https://doi.org/10.1175/15200493(2001)129 2.0.CO;2.

- Lines 353-355

I suggest incorporating more descriptions on how the reliability diagram is generated, and the way to interpret the reliability diagram.

- Line 366-367

Similar to the previous comment, e.g., why does the reversed S-shape indicate low ensemble spread?

- Line 374-396 (general comment for section 5.4)

It is interesting to also discuss the analysis increment for the unobserved variables, e.g., ocean current.

- Line 384

I suggest being more specific here. For example, revise "The correlation for SST is…" to "the correlation between … and … is"

- Line 481-483

While I do not insist on conducting more DA experiments to address the following issues in this paper, it would be helpful to include some of the discussions regarding the following questions:

(1) Could a larger inflation factor lead to a better DA and forecast performance?
(2) Exploring whether techniques to address insufficient ensemble spread, e.g., see the list below, can be effective would be an interesting experiment in the future work as well.

Zhang F. Q., Snyder C., Sun J. Z. (2004), Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. Mon. Weather Rev. 132: 1238–1253.

Anderson, J. L. (2007), An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus A, 59: 210-224. https://doi.org/10.1111/j.1600-0870.2006.00216.x

Anderson, J. L. (2009), Spatially and temporally varying adaptive covariance inflation for ensemble filters. Tellus A, 61: 72-83. https://doi.org/10.1111/j.1600-0870.2008.00361.x

Whitaker J. S., and Hamill T.M. (2012), Evaluating methods to account for system errors in ensemble data assimilation. Mon. Weather Rev. 140:3078–3089.

Ying, Y., and Zhang, F. (2015). An adaptive covariance relaxation method for ensemble data assimilation. Quarterly Journal of the Royal Meteorological Society, 141(692), 2898-2906.

El Gharamti, M. (2018). Enhanced adaptive inflation algorithm for ensemble filters. Monthly Weather Review, 146(2), 623-640.

(3) Since SIC is a bounded variable, assuming Gaussian error for SIC is inappropriate especially when SIC value is close to the boundary (i.e., zero or one). This non-Gaussianity can make Gaussian DA method, like EnKF, sub-optimal. It would be helpful to check the ensemble distribution of SIC at a single location before and after DA in a single DA cycle, (1) when SIC observation is close to the boundary, e.g., SIC = 0 or 1, (2) when SIC observation is away from the boundary, e.g., SIC = 0.5. Using some non-Gaussian DA techniques (e.g., Bishop 2016; Poterjoy 2016; Hu and van Leeuwen 2021; Anderson 2022; Chan et al. 2023, etc) can alleviate this problem and may improve the assimilation and the forecast of SIC.

Bishop, C.H. (2016), The GIGG-EnKF: ensemble Kalman filtering for highly skewed non-negative uncertainty distributions. Q.J.R. Meteorol. Soc., 142: 1395-1412. https://doi.org/10.1002/qj.2742

Poterjoy, J. (2016). A localized particle filter for high-dimensional nonlinear systems. Monthly Weather Review, 144(1), 59-76. https://doi.org/10.1175/MWR-D-15-0163.1

Hu, C. C., & Van Leeuwen, P. J. (2021). A particle flow filter for high-dimensional system applications. Quarterly Journal of the Royal Meteorological Society, 147(737), 2352-2374. https://doi.org/10.1002/qj.4028

Anderson, J. L. (2022). A Quantile-Conserving Ensemble Filter Framework. Part I: Updating an Observed Variable, Monthly Weather Review, 150(5), 1061-1074. https://doi.org/10.1175/MWR-D-21-0229.1

Chan, M.-Y., Chen, X., & Anderson, J. L. (2023). The potential benefits of handling mixture statistics via a bi-Gaussian EnKF: Tests with all-sky satellite infrared radiances. Journal of Advances in Modeling Earth Systems, 15, e2022MS003357. https://doi.org/10.1029/2022MS003357

- Figure 9

In the caption, I suggest also adding the meaning of the other lines (blue, blue dashed, orange, orange dashed).

- Figure 11

(1) Why there are more than 4 lines in each panel (there are only 4 in the legend)?
(2) What are the values in x-axis? Why are they not chosen uniformly between [0,1]?
(3) line 2 in the caption: … stating hos well … -> stating how well

- Figure 13

I recommend adding the unit to the variables in the figure.