Review of: The impact of stochastic sea ice perturbations on seasonal forecasts

by Kristian Strommen, Michael Mayer, Andrea Storto, Jonas Spaeth, and Steffen Tietsche

March 31, 2026

Manuscript Synopsis

This paper is an overview of using stochastic parameter perturbations (SPP) in the SI3 sea ice component of ECMWF's seasonal forecast system, SEAS6 (latest version). The manuscript describes the SPP scheme, based on the Storto and Andriopoulos [2021] STOPACK scheme, and then goes on to describe the schemes impact on spread within the sea ice and in the atmosphere. Finally, it describes the impact on various skill diagnostics for both sea ice and the atmosphere, further detailing some of the possible teleconnection impacts in the atmosphere.

In general, the manuscript is well written and an enjoyable read. It presents a good overview of stochastic parameterizations in a sea ice model for use in ensemble forecasting. The system analyzed is a forecast system with stochastics perturbations added to the forecast integration — with the initial spread of the sea ice system not well documented (major comment #1). As such it does not directly address the often underdispersive sea ice initial conditions (ll. 39-40 of Introduction), but it does comprehensively describe their usefulness in the subsequent forecast to increase spread and more adequately represent forecast uncertainty. The paper would be a huge benefit to subseasonal to seasonal forecasters of Arctic (and Antarctic – although the focus here is on Arctic) sea ice and its interaction with atmospheric teleconnections and polar oceanographic processes — although, no effects of the stochastics on the sub-surface ocean are investigated in this manuscript.

My primary concerns with the manuscript are concerns of information communication. More complete documentation (or links to documentions) of the (existing – as the innovations here do not affect these) initial sea ice conditions supplied by the ocean and sea ice analysis ORAS6 would be in order, as otherwise it is very hard to interpret and understand some of the comments regarding the growth of ensemble spread in the coupled versus uncoupled simulations (Section 4.3). Secondly, I could use a further talking through of some of the points in the discussion of atmospheric teleconnection (Major Comment #2). The effect of introducing stochastics, which the authors show can influence the mean state (despite the neutrality of the stochastics itself), seems to introduce a tendency towards a more negative North Atlantic Oscillation (NAO) (Z500 anomaly in Figure 6c, and wind anomalies in Figure 8a) type behaviour, but they then describe (Section 6.2) the improvements to the temporal variability (correlations) in Z500 (Figure 16a) in terms of a strengthening of the meridional temperature gradient (their words, l. 437; actually stated as the non-SPP is biased weak), which I then ascribe to a strengthened, more zonal (my words; their words – modulated, l. 440) jet stream, which I further associated with a tendency towards a more positive NAO. There are a lot of subtleties here, some of which I may understand, while others may be outside my zone of expertise, and the authors make abundant use of qualifiers, for example, the pattern of increased correlation only partially overlaps the NAO pattern (ll. 427-429). Nevertheless, I could use to be guided out of my confusion. I acknowledge I have conflated the increased temporal coherence/correlation with a change in mean state — and that may be my entire problem. However, I am also confused that the bias in the mean Z500 state (looks like negative NAO) seems to be contrary to the increased meridional temperature gradient, but perhaps I am missing something, or have linked things I should not have.

My recommendation is Minor Revisions for additional explanatory material, although admittedly, addressing all my abundance of comments may constitute major undertakings.

Major Comments

1. Additional information, particularly, whether there are multiple members (an initial spread) of sea ice initial states, plus the ORAS6/ORAS5 citations to provide further understanding [At a minimum Zuo et al., 2017b,a, Browne et al., 2025]. It is difficult to understand the statement (ll. 239-240): "We did not run an uncoupled control (i.e., without sea ice perturbations), as this would be expected to have negligible ensemble spread." My initial reaction was that if you turn off the sea ice perturbations you would have **no** spread, but then I happened to know that there would be an (small perturbation) ensemble of sea ice initial conditions from ORAS6 to initialized from — others readers may not have this information. Furthermore, providing further background to the possibility of stochastic perturbations to enhance the existing spread of initial conditions (as pointed out in the conclusions/future work) would benefit the motivation of this study.

2. I would like to be pulled through the explanation of Section 6.2 slightly better, particularly in light of the changes in bias seen with the addition of SPP. The change in mean of DJF Z500 (Figure 6c) and 700hPa winds (Figure 8a) seems to indicate a tendency toward a more negative(ish) NAO. Meanwhile the discussion of Section 5.2 shows that SPP increases the T850 meridional gradient (Figure 16b), which (I believe) is associated with a stronger, more zonal jet stream. This is turn (with less conviction; it is complicated) is associated with a tendency toward a more positive(ish) NAO. None of this is directly responsible for the increased skill seen in correlation (which is independent of the mean state) in DJF Z500 (Figure 16a) and through projection onto the NAO pattern, increased skill in NAO index. However, as the authors and some of their supporting literature have argued, the changes in mean state can modulate (l. 440) the variability – although it is here my knowledge, and therefore my logic may breakdown. Possible graphical improvements to discussion:

   - Include the mean Z500 change (a repeat of Figure 6c) as a third spatial plot in Figure 16 – but this time as a full colour contour, instead of the secondary lines on Figure 6c. Or possibly a no SPP bias versus change in bias Z500 plot as done for T850 in Figure 16b. This allows to see the changes in bias and change in skill side by side. **Also** a dedicated plot with smaller threshold in lowest (absolute) contour value would *perhaps* show an anomalous Z500 low over the Ural Mountains (l. 213) that *must* exist for the anomalous circulation seen previously in Figure 8 (technically, it must appear in a plot of Z700 to give the anomalous U700 fields, but presumably it should be present in Z500, or mean sea level pressure as well).

3. The skill of the forecast system (Section 6) is evaluated by root mean square error (RMSE), including the spread to RMSE ratio, the bias corrected continuous ranked probability score (CRPS), and finally by correlation. My concern is that the root mean squared error is not a bias-corrected RMSE (or standard deviation error), while by construct, their CRPS score is. I am not looking for the authors to necessarily change these, but I do need them to justify this accordingly, especially as changes in bias with SPP are definitely present (Section 4). In particular, it is often argued that the spread/error relationship should be between a bias-corrected RMSE and spread – as spread should not be used to correct for bias, as this leads to decreased reliability. Furthermore, the CRPS score is often decomposed into reliability and potential [Hersbach, 2000]. I am somewhat knowledgable, but certainly do not have the access to the experts on the subject that the authors do. I *believe* that removing the bias before computing the CRPS will mostly (but not exclusively, at least not without additional constraints) affect the potential component. The authors should likely address this in some form. Ultimately, however, this all has an effect on how to interpret the results: If RMSE was calculated after removing the mean bias, not only might it have a better interpretation within the spread/error relationship, but it would also have a near 1-to-1 correspondence with correlation results (reliability). But as it is now, the RMSE results are likely more related to bias. Conversely, the separation between bias

and reliability of the CRPS, which is usually handled through the computation of the reliability and potential components, is now only seen through the calculated bias corrected CRPS. **Note:** Calculation of the CRPS components when validating against an analysis, as has been done here, can be complicated due to the increased memory requirements, so it is not necessarily the route to be taken.

4. Section 5: I am in total agreement that the changes in mean state with respect to changes in spread and the introduction of stochastics is a consequence of the bounded nature of sea ice and stochastics pushing the mean state away from the bounds due to the 1-direction limitations at the bounds. Indeed I was convinced at its first mention at the beginning of Section 4. Nevertheless, I applaud the authors for a detailed analysis, including toy models to demonstrate this important point. However, I would suggest one further addition to strengthen the point, and allow me to suggest a possible 2nd (secondary) mechanism due to irreversible dynamics (like ridging).

   (a) The thickness plots (Figures 3 & 4), seem to indicate an increase of thickness along coastal northern boundaries of the Canadian Archipelago and Greenland, along with the east Siberian coast, particularly in the winter (Figure 3b). This I suggest might be down to the stochastics leading to not easily reversible processes such as ridging, that could lead to a build up of thickness, especially in areas where ice is constrained by physical boundaries. Note: Dynamic processes would not necessarily change the grid cell average thickness, as it will just redistribute the ice within the grid cell, however, the accompanying decrease in ice concentration, would then allow for more ice to be advected in and actually increase the mean grid cell thickness. This could account for the increase in mean thickness in these areas, along with some of the decreased sea ice concentration. The decrease of sea ice thickness near the thickest ice right along the north coast of Greenland might then indicate a threshold where the stochastics does not favour increased ridging, but enough disturbance to the state to facilitate its breakdown. But also note: I am not a sea ice dynamicist.

      - **Aside:** It is never mentioned that the sea ice thickness used in your analysis is the grid cell mean thickness, presuming that is the SI3 model variable. I find it useful to always remind readers of this.

   (b) An additional simple piece of information I would think would further confirm some of the manuscripts suggestions would be to tabulate (they are single numbers – unless you produce a yearly timeseries) the hemispheric change in Sea Ice Extent, Sea Ice Area and Sea Ice Volume with and without sea ice SPP. The sea ice extent should increase with SPP, as the decreasing high concentration ice does not affect the sea ice edge, but increasing the low concentration sea ice should increase the sea ice extent. If the stochastics is essentially working on free ice, these changes in concentration at either end may spread out the ice edge, but leave the sea ice area the same. Finally, if sea ice dynamical processes start to be important, the sea ice area might change, but the sea ice volume would (maybe) be left unchanged (just re-distributed). This likely might only work in the winter, as summer time thermodyanamics, like sea ice being ejected into warmer waters, is likely to affect results. Nevertheless, this would likely be a simple useful diagnostic to have, which is not necessarily easy to ascertain from just viewing changes in the spatial distribution of sea ice concentration and thickness.

5. Figure 10 (and discussion): I mostly agree with everything discussed with regards to Figure 10. However, I have a few comments.

   - Scatter plots are easy and convenient, but they have a tendency to sometimes give false impressions due to the overlaying of points on the plot. In this context, I would think a 2-dimensional probability distribution (or 2-D binned plot) would be better. I.e. Bin the number of occurrences as a function of spread and mean change, and then plot frequency of (normalized) occurence as colour or gray shading on 2-D plot. Instead of repeating the full (black dots) scatter plots for rows 2 and 3, you would then simply replace them with 2-D PDF's for $> 80\%$ and $< 20\%$ concentration.
   - This would negate the need of the 90% of points contour. **Note:** A contour encircling 90% of points is **not** unique (otherwise it would not be possible to gerrymander election districts).

Enclosing 90% of points using the smallest area, however, is unique – and I suspect what is shown.

- The scatter plot for JJA is **not** more Gaussian (ll. 303-304: Actual statement "this non-Gaussianity is less pronounced"), as it is not centered around the zero mean change. Rather it is *missing* the positive (blue) node of the bi-modal distribution due to thermodynamics reducing low concentrations that may have increased due to stochastics. I.e. Thermodynamics suppresses this mode in the summer – and likely(?) elongates this mode in the winter, as thermodynamics further increases the low concentrations increased by stochastics. More bluntly, all points get (on average) pulled to the left (lower concentrations) in JJA and to the right (higher concentrations) in DJF due to thermodynamics.

Minor Comments

1. l. 7: "(enhance) by around 10% relative to an unperturbed forecast" This is somewhat confusing and harks back to Major Comment #1. A reader reading "unperturbed" forecast will naturally(?) assume a zero-spread (deterministic) forecast and wonder how you can increase zero by 10%. I'm afraid you may have to expand the explanation somewhat ("relative to a forecast where pertubations are not applied directly to the sea ice").

2. l. 18: "tendency perturbations" Although I have no desire for additional acronyms to be added unnecessarily, I wonder if it would not be advantageous to somehow mention this is what is commonly referred to as Stochastic Perturbed Parametrization Tendencies (SPPT)

3. l. 79: Should mention what NEMO4/SI3 is an upgrade from (I believe NEMO3.4/LIM2 based on ORAS5 documentation).

4. l. 81. Technically the 5 category ice is not a **property** of SI3, but an adjustable **option** (i.e. could equally have chosen 1 category or 10 category). I do not see any perturbed sea ice parameter other than snow conductivity that might be affected by the number of thermodynamic layers, but it might be useful to document the number of ice and snow thermodynamic layers as well (ORAS6 has 4 ice and 1 snow thermodynamic layers, which I suspect would be the same here), along with whether a melt pond scheme has been activated (yes, presumably if you perturb the meltpond parameter *rn_pnd_flush*).

5. ll. 83–86. I would think the properties of the sea ice grid would be far more important to discuss than the atmospheric grid. Beyond mentioning that the ocean grid is 1/4°, I think it would be wise to mention it is (based on Browne et al. [2025]) the eORCA025 tripolar grid, with grid sizes between 3 and 12 km [Subich et al., 2020] in the Arctic, and down to 4 km [Smith et al., 2021] in the Antarctic. This would be of importance to the covariance length scales generated by the Shapiro Storto and Andriopoulos [2021] or Perlin Noise Perlin [2002] filters.

6. ll. 95-97: The SMMR/SMMI/SMMIS OSI-450-a products the authors use for validation may easily vary from other sea ice concentration products by (approximately) amounts near to the changes in spread seen in this manuscript [Renfrew et al., 2021, Peterson et al., 2022, Niraula, 2023], at least in the marginal ice zone. As SEA6 would be initialized by this product, it would remain the best product by which to estimate forecast error. However, it might be a good idea to state that there is some uncertainty in the sea ice concentration "observations."

7. ll. 117-118. You might want to mention you only used correlation in the context of verification of Z500? [I kept expecting correlation results that never came.]

8. Section 2.2: I will have to admit to not knowing Weather and Climate Dynamics requirements. Is a open code respository of analysis tools required for publication.

9. ll. 132-135: The stochastic perturbation time scale (1 day) is given, but the covariance length scale (number of iterations of Shapiro, and presumably also the Perlin Noise filter) is not given. The actual length scale would be approximate and geo-spatially dependent owing to the grid cell variation in the eORCA025 grid, particularly in the Arctic, but an approximate range could still be provided.

10. l. 168. I suspect this statement is relatively independent of the definition of ice edge, but please define the definition of ice edge (sea ice concentration = 0.15?)

11. Figure. 1-4: Fram Strait seems to be a relatively active area (likely to always be an active area) in changes to mean state and spread of both sea ice concentration and thickness. Were any calculations performed to diagnose changes to mean and spread of sea ice transport out of the Arctic choke points (i.e. besides Fram, Bering/Davis/others [see e.g. Smith et al., 2024]) – or even just Fram Strait? This would seem to have implications for Section 5.

12. Figure 6, 7 & 17 (others?) It is not indicated whether the solid or dashed lines correspond to positive contour levels. I assume it is the solid lines – but that is an assumption without context in the caption.

13. Figures 6 & 7: This is the first and *only* time that the change in spread is expressed as a percentage change. You explicitly write this out in the "units" of the colour bar, so as to differentiate with the usage of % for concentration and concentration changes (good). But perhaps this should also be expressed in the caption explanation for Figure 6b – which now only indicates it is a change in spread (i.e. not a fractional/percentage change in spread relative to the control spread – I assume).

14. ll. 239-240: This is were I missed an explanation of whether there are perturbations or not in the initial conditions. If there are no perturbations to the initial conditions there will not be negligible spread, there would be **no** spread. If there were perturbations to the initial conditions, then there should be spread, but not necessarily negligible spread — unless the spread in the initial conditions is quite undispersive.

15. ll. 241–244: Perlin Noise, Does this change the characteristics of the covariance lengthscale in any appreciable fashion?

16. Figure 9b-c: The spatial plots are quite obviously masked to correspond to the hemispheric mean in Figure 9a. While I do not think it is necessary to do this masking, if done, if should probably be documented in the caption.

17. Figure 9b: **Note:** The *month-to-month temporal variability* is essentially (or at least strongly dominated by) the seasonal variability. I.e. You are seeing where the sea ice changes most between summer and winter – and not month to month timescale variability as you might see in atmospheric fields such as the Z500.

18. ll. 249–255: An observation. I am not particularly surprised independent perturbations do not lead to as much increased spread as in the uncoupled simulations. The increased spread in the uncoupled simulations is presumably coming from non-linear or non-orthogonal interactions between the perturbed fields. My impression is that the coupling with the atmosphere would more than adequately generate that. Plus, I would assume that the atmospheric perturbations are not provided by the same perturbation fields either, nor with the same temporal and lengthscale coveraiance characteristics, so there is already an existing non-orthogonal independence built in.

19. Figure 11 (JJA): I seem to recall that the 60% sea ice concentration is important for air/sea flux exchange. Is the peak in JJA spread and mean change (peak in amplitude of change, not neccesarily frequency) possibly related to thermodynamic considerations?

20. ll. 390–391: "At the coastlines, where validation is more robust, there is a mixed impact." I am not sure I completely agree with this statement. Firstly, do you mean at the coastlines, or at the sea ice edge / marginal ice zone? In both cases, however, the SMMR/SMMI/SSMIS satellite product has a relatively large footprint that gives the observational product a rather large uncertainty in these regions. Indeed, the OSI-SAF product has been shown to have a large degree of uncertainty in the marginal ice zone [Renfrew et al., 2021].

21. Most ensemble verification of sea ice is done bypassing complete knowledge of the sea ice concentration, and focusing on more important changes mostly to the marginal sea ice by evaluating skill on the predictability of the probability of ice (fraction of ensemble members that exceeds some sea ice

concentration threshold, most commonly 15%) as in Zampieri et al. [2018], Peterson et al. [2022] and others. I would be curious if any of the conclusions regarding sea ice predictability would differ from taking this standpoint.

22. There is no thickness assimilation in the initial sea ice state, so there are likely significant biases in the sea ice thickness. It would be an interesting topic (not necessarily for this manuscript) to evaluate whether introducing stochastic perturbations to the sea ice increases or decreases these biases – and if then actually assimilated increases/decreases the ability to properly forecast ice thickness evolution over a season. The Copernicus combined Cryosat/SMOS reprocessed thickness dataset is available from 2010 onward (from 1995, but I would not necessarily recommend the pre-CryoSat2 period) which should give a suitable window to produce (DJF) sea ice thickness forecast errors (`https://data.marine.copernicus.eu/product/SEAICE_GLO_PHY_CLIMATE_L3_MY_011_013/description`.

23. There is no mention of any effects on the sub-surface ocean in the manuscript. While there is unlikely enough sub-surface observations to meaningfully test the changes to the sub-surface forecast skill, it might have been interesting to at least show pan-Arctic average changes to both spread and mean T/S profiles. This could at least give a broad view of how deep the stochastics might affect the sub-surface solution.

24. The Perlin [2002] paper was not easy to find. Please provide the DOI: `https://doi.org/10.1145/566570.566636`. Also, I believe "noise" should be capitalized as "Noise." It is not generic noise — but a proper name "Noise algorithm." That being said, I did not find the given citation particularly useful for the context I was looking to answer: What covariance lengthscales are generated by (iterations of) the Noise algorithm.

Typos and Grammar

- Centre is mis-typed "cdntre" (l. 55). I am surprised that one got past internal review!
- Figure 3: Not an error — just an observation: The sideways unit m is easily misinterpreted as an upward oriented E. I was originally thinking that units were neglected on the colour bar – and wasn't sure what E stood for, until I realized the colour bar title (units) was rotated. I would suspect any remedy might not improve the possibility for confusion (and it could only be me than is confused), so I will not suggest or insist on any.
- l. 236: Considering the subject matter, I would consider writing ocean and sea ice only forecast.
- Figure 19: My eye is drawn to a dipole which somewhat unnaturally seems to be split across the 60°E/120°W meridian, at least north of Svalbard and then south again to the Canadian Archipelago. I do not see this in any of the other plots, but given that this is a comparison between a coupled and uncoupled simulation, I wonder if this is the only plot based upon the output from NEMO/SI3, with all other plots being based upon sea ice output from (and interpolated to) the atmospheric grid. The default assumption must be that this is real, but I do want to beware of possible *plotting* artifacts. The most obvious candidate might be the north fold in the ORCA025 grid (and I have produced many plots, some published with discontinuities across that line), but that is **not** along that meridian — as the north fold is along the 73°E / 107°W meridian. Nevertheless, I just want to check whether there might be some other possible plotting artifact.
- l. 335: 9/nine. I do not normally pay attention to this, but this may be an instance where typographical conventions would suggest that 9 be spelled out as nine.
- l. 558. I believe the reference type should be [Hersbach et al., 2023] and not Hersbach et al. [2023]. [\citep and not \cite if using LaTeX].
- ll. 575, 577, 586, 589, 591, 602, 608, 615, 626, 630, 642, 644, 646, 650, 656, 663, 665, 673, 681, and 697. There are extra `https://doi.org` in the DOI of many references — and clicking on the front part will take you to non-existing pages.

References

P. Browne, E. de Boisseson, S. Keeley, C. Pelletier, and H. Zuo. Sea ice data assimilation in ORAS6. *EGUsphere*, 2025:1–21, 2025. doi: 10.5194/egusphere-2025-3991. URL https://egusphere.copernicus.org/preprints/2025/egusphere-2025-3991/.

H. Hersbach, B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J-N. Thépaut. ERA5 monthly averaged data on single levels from 1940 to present, 2023. URL https://doi.org/10.24381/cds.f17050d7.

Hans Hersbach. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5):559 – 570, 2000. doi: 10.1175/1520-0434(2000)015⟨0559:DOTCRP⟩2.0.CO;2. URL https://journals.ametsoc.org/view/journals/wefo/15/5/1520-0434_2000_015_0559_dotcrp_2_0_co_2.xml.

B. Niraula. *Ice Edge Verification – Measuring the skill in our forecasts and disagreement in our observations.* PhD thesis, Universität Bremen, 2023. URL https://doi.org/10.26092/elib/2298.

K. Perlin. Improving Noise. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 681–682, 2002. doi: 10.1145/566570.566636. URL https://doi.org/10.1145/566570.566636.

K. Andrew Peterson, Gregory C. Smith, Jean-François Lemieux, François Roy, Mark Buehner, Alain Caya, Pieter L. Houtekamer, Hai Lin, Ryan Muncaster, Xingxiu Deng, Frédéric Dupont, Normand Gagnon, Yukie Hata, Yosvany Martinez, Juan Sebastian Fontecilla, and Dorina Surcel-Colan. Understanding sources of northern hemisphere uncertainty and forecast error in a medium-range coupled ensemble sea-ice prediction system. *Quarterly Journal of the Royal Meteorological Society*, 148(747):2877–2902, 2022. doi: https://doi.org/10.1002/qj.4340. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4340.

I. A. Renfrew, C. Barrell, A. D. Elvidge, J. K. Brooke, C. Duscha, J. C. King, J. Kristiansen, T. Lachlan Cope, G. W. K. Moore, R. S. Pickart, J. Reuder, I. Sandu, D. Sergeev, A. Terpstra, K. Våge, and A. Weiss. An evaluation of surface meteorology and fluxes over the iceland and greenland seas in ERA5 reanalysis: The impact of sea ice distribution. *Quarterly Journal of the Royal Meteorological Society*, 147(734):691–712, 2021. doi: https://doi.org/10.1002/qj.3941. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3941.

Gregory C. Smith, Charlie Hébert-Pinard, Audrey-Anne Gauthier, François Roy, Kenneth Andrew Peterson, Pierre Veillard, Yannice Faugère, Sandrine Mulet, and Miguel Morales Maqueda. Impact of assimilation of absolute dynamic topography on arctic ocean circulation. *Frontiers in Marine Science*, 11, 2024. ISSN 2296-7745. doi: 10.3389/fmars.2024.1390781. URL https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2024.1390781.

Robin S. Smith, Pierre Mathiot, Antony Siahaan, Victoria Lee, Stephen L. Cornford, Jonathan M. Gregory, Antony J. Payne, Adrian Jenkins, Paul R. Holland, Jeff K. Ridley, and Colin G. Jones. Coupling the u.k. earth system model to dynamic models of the greenland and antarctic ice sheets. *Journal of Advances in Modeling Earth Systems*, 13(10):e2021MS002520, 2021. doi: https://doi.org/10.1029/2021MS002520. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021MS002520. e2021MS002520 2021MS002520.

Andrea Storto and Panagiotis Andriopoulos. A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 147(736):1691–1725, 2021. doi: https://doi.org/10.1002/qj.3990. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3990.

C. Subich, P. Pellerin, G. Smith, and F. Dupont. Development of a semi-lagrangian advection scheme for the nemo ocean model (3.1). *Geoscientific Model Development*, 13(9):4379–4398, 2020. doi: 10.5194/gmd-13-4379-2020. URL https://gmd.copernicus.org/articles/13/4379/2020/.

Lorenzo Zampieri, Helge F. Goessling, and Thomas Jung. Bright prospects for Arctic sea ice prediction on subseasonal time scales. *Geophysical Research Letters*, 45(18):9731–9738, 2018. doi: 10.1029/2018GL079394. URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL079394`.

Hao Zuo, Magdalena Alonso-Balmaseda, Eric de Boisseson, S Hirahara, Marcin Chrust, and Patricia de Rosnay. A generic ensemble generation scheme for data assimilation and ocean analysis. Technical report, ECMWF, 2017a. URL `https://www.ecmwf.int/node/17831`.

Hao Zuo, Magdalena A. Balmaseda, and Kristian Mogensen. The new eddy-permitting orap5 ocean reanalysis: description, evaluation and uncertainties in climate signals. *Climate Dynamics*, 49:791–811, 2017b. doi: 10.1007/s00382-015-2675-1. URL `https://doi.org/10.1007/s00382-015-2675-1`.