Response to Anonymous Referee #1 of **manuscript egusphere-2025-6379 (second review)**

Title: Arctic sea ice predictability on daily-to-weekly timescales: sensitivity to initial positional errors under different rheology formulations

Author(s): Lohenn Fiol et al.

Reviewer 1: The authors have objectively addressed all my questions and concerns and incorporated a sufficient discussion regarding the model results into the revised manuscript. I believe this manuscript is now suitable for publication in The Cryosphere, subject to a few minor typo corrections.

For example, Line 218: duplicate word "...previous previous..."; Lines 416-417: "heterogenities" to "heterogeneities"; Line 590: "winder" to "winter" etc...

We thank the reviewer for their positive assessment and for confirming that our first revisions adequately addressed their previous comments. We have now corrected all the typos identified by the reviewer, and have carefully proofread the manuscript again to catch any remaining similar errors.

Response to Anonymous Referee #3 of **manuscript egusphere-2025-6379**

Title: Arctic sea ice predictability on daily-to-weekly timescales: sensitivity to initial positional errors under different rheology formulations

Author(s): Lohenn Fiol et al.

We thank the Reviewer for their careful evaluation of our manuscript and for the constructive comments provided. Below are our responses to the points raised (in black font) and *our propositions for the revised manuscript in blue* *while the reviewer's comments and questions are reported in gray.*

Reviewer: *Summary:*

This manuscript examines the short-term predictability of Arctic sea ice under two rheological formulations, BBM and aEVP, using a perfect-model ensemble framework with initial positional perturbations. The study shows that BBM produces stronger nonlinear amplification of initial errors and much larger Lagrangian spread, whereas aEVP tends to smooth initial heterogeneities and yields more convergent, and therefore more predictable, solutions. The topic is timely, the methodology is generally solid, and the results are relevant to operational short-range sea-ice forecasting.

Recommandation:

Overall, I find this manuscript to be clear and well-written, and I recommend considering its publication after making a few minor revisions. My comments below are intended to strengthen robustness, interpretability, and operational relevance.

Major comments:

1. I have two reservations about the fairness of the comparison between the two rheological methods.

---

a) Both BBM and aEVP hindcasts are initialized from states extracted from a BBM reference simulation (then perturbed). While this is understandable for the sake of comparability, I wonder if it might lead to an initial imbalance in the aEVP, such as causing the early smoothing process to proceed too rapidly. It might be helpful to add some clarification in Section 2.2.

→ Indeed we use the same initial states extracted from a BBM reference simulation for the sake of comparability, but also because we aim to investigate the sensitivity of the forecast to initial errors **representing misplaced sea ice features, especially LKFs**. And those are more realistically-represented in a BBM simulation than aEVP at this 1/4° resolution. At the beginning of section 4.2.2 we have already commented: *"This choice ensures that the extracted initial states include a realistic level of heterogeneities and LKFs as it would be in an operational context where data assimilation would assimilate high-resolution observations as envisaged by, e.g. Korosov et al. (2023) or Moro et al.(2024)."*

Initializing an aEVP simulation with a BBM-derived state may indeed introduce an initial imbalance, on top of the perturbations we applied by design. However, it is difficult to conceive how such an imbalance would **reduce** non-linear amplification. On the contrary, it could be seen as an additional perturbation and would thus be expected to *favor* non-linear amplification, which is not what we observe.

---

b) Using different drag coefficients to match mean drift is reasonable and often necessary. Still, I wonder whether some of the contrast observed in the report might have been affected by this calibration choice. Please add a short sensitivity statement documenting that the principal conclusions are not artifacts of drag tuning.

→ Thanks for the suggestion. We propose to add this sentence at the end of section 2.1:

Note that since our predictability analysis presented in Sect. 3 follows a "perfect-model" approach, in which the predictability scores quantify the evolution of ensemble members relative to one another rather than against observations, the tuning of the drag coefficient does not directly affect these scores, as the same value is used across all members of a given ensemble. The qualitative results presented in this study are therefore expected to hold for different values of the drag coefficient. Our choice of values was primarily guided by the criteria of realism and fairness of comparison between the BBM and aEVP configurations.

---

2. The perfect-model setup is clearly described, but the operational interpretation can be sharpened. I suggest adding a concise synthesis that clearly distinguishes between which conclusions are considered intrinsic upper bounds and which results are expected to deviate due to real-world uncertainties (such as wind, ocean initial state, model parameters, etc.). This would make the practical value of the results more immediately usable by forecasting groups.

→ We agree that distinguishing intrinsic upper bounds from other sources of operational degradation is important, and we have aimed to do so at several points throughout the manuscript rather than in a single section. In particular, we already state explicitly that, by design, only the initial sea ice state is perturbed while the ocean state and atmospheric forcing are kept unperturbed, so as to isolate the response of the coupled ice–ocean system to initial positional uncertainty alone. We note that this corresponds to an upper bound for predictability (or potential predictability), which will, in practice, be lower once additional sources of uncertainty—such as wind forcing, ocean initial state, or model parameters—are taken into account, and we have now made this point more explicit by adding a sentence on the qualitative-versus-quantitative distinction directly to this paragraph in the Summary and Conclusions section:

"In that sense, we focus on an upper bound for predictability (or a potential predictability) that will, in practice, be lower when additional sources of uncertainty—such as atmospheric forcing, ocean initial state, and model parameters—are taken into account. This upper bound is nonetheless relevant in itself to characterize the behaviour of the modelled sea ice, and we

expect the qualitative contrasts identified between rheologies to remain relevant for operational forecasting even as quantitative predictability limits are reduced."

Beyond this general framing, we also discuss, for specific results, how the corresponding conclusions connect to operational practice. For example, in the conclusion section:

- *In that regard, improving operational forecasts of the sea-ice edge is unlikely to be achieved through a changed rheology or through more accurate assimilation of its observed initial position alone, but rather through better initialization of the underlying ocean and more realistic atmospheric forcing—at least at the present model resolution. Our results do not exclude, however, that at higher resolution, a more turbulent ocean circulation that includes better-resolved mesoscale features may introduce additional nonlinearity in the coupled ice–ocean response, potentially increasing the sensitivity to initial position errors.*
-
- *These contrasted behaviors are not merely academic, they also matter for practical applications, such as field operations or search-and-rescue response, where understanding uncertainty is often as important as the forecast itself (e.g. Tietsche et al., 2020). A system that strongly amplifies small initial errors (as documented here for BBM) may offer lower deterministic predictability yet provides more realistic—and therefore more useful—information about the range of plausible trajectories or the local likelihood of severe deformation. Conversely, a system that smooths heterogeneity (as in aEVP) may appear more predictable and also easier to handle for technical operational purposes (e.g. Data Assimilation), but it risks underestimating extreme outcomes. Such contrasts reinforce the value of ensemble forecasting for sea ice: in a highly nonlinear and only weakly predictable medium, ensembles provide not only an estimate of the forecast state but also a quantification of the risk associated with user-relevant events.*

Given that the Discussion and Conclusion sections are already quite extensive, we would prefer to keep these clarifications distributed alongside their corresponding results rather than consolidating them into a separate synthesis, which we worry could feel repetitive. We hope the additional sentence above, together with these examples, sufficiently clarifies the point you raised.

---

3. The paper appropriately acknowledges the 1/4° resolution constraint. I encourage one additional step: explicitly separate which conclusions are likely resolution-robust (e.g., qualitative rheology contrast) and which are configuration-dependent (e.g., exact day ranges of predictability loss).

→ We agree this is an important distinction, but we would like to remain cautious on this point rather than overstate what our single-resolution study can support, since we have not

investigated higher-resolution configurations with predictability metrics yet. This is why we would like to keep the sentence we had already formulated in the discussion as is:
"... our results are representative of the difference in the predictability behavior between the BBM and aEVP rheology in a model configuration at a horizontal resolution of 1/4°. The dependence of the results on resolution, and in particular the discrepancy in sensitivity between aEVP and BBM, was not explored in this study and would certainly merit further investigation, since each rheology is known to respond differently to resolution changes. "

That said, we can add some precision on what is likely to hold at higher resolution:

"Nonetheless, we expect the qualitative contrast between the two rheologies (i.e., the fact that BBM exhibits non-linear error amplification associated with LKFs while aEVP does not) to persist across resolutions, as it likely stems from a structural difference in how the two rheologies represent deformation. In contrast, quantitative metrics, such as the specific timescales of predictability loss, are likely to be more sensitive to resolution change."

---

4. I would also like to point out that the current final section combines both the Discussion and the Conclusions into a single, rather long section, which feels somewhat heavy. Splitting this into separate sections would likely improve clarity and make the manuscript more reader-friendly.

→ Thank you for the suggestion. It is true that our last section has grown after the first round of reviews and breaking it into 2 parts could help the reader. We propose to break it up into a *Summary and Conclusions* part, and then a Discussion part (starting with the sentence: *"While these results carry clear practical relevance, a number of important considerations regarding the scope and generalizability of our findings merit discussion.". We also propose to split the Summary and conclusions part into 3 subparts:*

- *Ice-edge position results*
- *CRPS results in the pack ice region*
- *Practical consequences on predictability of high-deformation events and Lagrangian trajectories*