We would like first to 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: *General Comments*

This paper evaluates the short-term (daily-to-weekly) dynamical sensitivity of Arctic Sea ice to initial positional errors under two rheological frameworks: Elasto-Visco-Plastic (aEVP) and Brittle Bingham-Maxwell (BBM). Using a "perfect model" framework, the study highlights that the BBM model exhibits stronger nonlinear sensitivity in pack ice, reaching a predictability limit for drift and deformation within 1-5 days. These findings provide valuable insights into the choice of rheological models and have practical implications for operational polar ice forecasting and search-and-rescue efforts.

Overall, the experimental design is interesting, the workload is solid and the language is generally fluent, certain parts need to be streamlined. My concerns regarding this paper primarily focus on the following aspects:

- the interpretation of the model configuration,

- the depth of the physical explanations,

- and the applicability of the findings to the modern, thinner Arctic ice.

If these issues can be fully discussed in this paper, I believe it will improve the contribution and research significance of this paper. Beyond that I think this work fits in the scope of TC and can be published after reasonably addressing the following concerns and recommending moderate to major revision.

Specific comments

1. ***Reviewer:*** *The authors selected January-March 1997 as the initial conditions for their simulations. While I understand that the year 1997 was chosen to take advantage of the RGPS assessment in Brodeau et al. (2024), the Arctic has transitioned into a "New Arctic" state characterized by thinner, more mobile, and more easily deformable first-year ice, alongside a warmer and more strongly stratified ocean. Readers will inevitably question how instructive predictability estimates based on the thick-ice conditions of 1997 are for contemporary Arctic forecasting. Given that the authors emphasize the operational relevance of this work, I strongly recommend they fully explain their reasons for selecting 1997 as the study*

period. I also recommend the author provide reasonable physical speculations in the Discussion section regarding how this shift to a "New Arctic" might alter their results. For instance, with a thinner and more fragile ice cover, the internal compressive strength of the ice decreases, thereby amplifying the dominant role of wind stress and ocean drag in the momentum balance. Consequently, is it highly likely that the "1-to-5-day predictability limit" derived from the 1997 thick-ice baseline essentially represents an upper bound under modern "New Arctic" conditions? And something like that…… Additionally, I also recommend clearly explain what are the substantive implications of this study for future model improvements and 'new Arctic' forecasting? I believe addressing these implications will significantly strengthen the paper.

We fully understand the reviewer's comment and agree that the 1997 winter corresponds to a winter season before the "regime shift" around 2007 in sea ice thickness described by, e.g. Sumata et al 2023 or Ludwig et al 2026. Those studies document a decrease in the observed sea ice thickness between 0.3 to 1 meter in average over the Arctic region, depending on the season (less in winter than summer) and on the type of observational product used to quantify the thickness trend.

As the Reviewer noticed, the primarily reason for our study to focus on 1997 is that we could build on the previous study by Brodeau et al, 2024 who had implemented the BBM rheology formulation in the NEMO-SI3 model and thoughtfully evaluated the model against RGPS deformation observations available for that year (as mentioned at the beginning of section 2.2 in our manuscript).

That said, we do acknowledge that a decrease of 0.3 to 1m of the sea ice thickness has presumably some consequences on the sea ice dynamics. In the elasto-brittle framework, a thinner sea ice breaks up more easily, as thickness influences the mechanical strength of the ice and thus its resistance to fragmentation (e.g. Ólason & Notz, 2014 and reference therein). However, we do not expect that the main conclusions of our predictability study would be drastically changed if based on a more recent year with a thinner ice: the very contrasted behavior of the system with either the BBM or aEVP rheology that we show with the CRPS analysis (that is, a more chaotic-like, less-predictable system with BBM), very likely persists in more recent years with thinner ice, as it reflects a fundamental characteristic of the rheology formulation. Even though not focused on predictability, the 2013 winter case study documented by Rheinlænder et al 2024 is for example an illustration of the contrasted appearance of the sea ice simulated with BBM and EVP rheology, with greater fracturing and a higher lead fraction in the BBM case, that remains within the "thinner-ice regime" established after the 2007 shift. This gives us confidence that our main conclusions regarding the contrasted predictability properties of the BBM and aEVP configurations would persist. However, the exact

values quantifying the predictability limits, as derived from the CRPS scores for the 1997 winter season, may vary somewhat for more recent years. Given that nonlinearity arises from the fracturing mechanical processes in the system (at least in winter, when the entire Arctic region is covered with pack ice and internal stress dominates over the other forces), a thinner, more easily breakable ice would probably result in even shorter predictability limits than those reported here with BBM (1–5 days for deformation and drift, 5–10 days for concentration).

Quantifying this statement in a robust manner would unfortunately require running additional ensemble forecast experiments for several 10-day periods for a more recent year. This would demand considerable additional work and computational resources. We propose instead (as suggested by the Reviewer) to discuss this aspect in the last section of the manuscript. We fully agree that it should be discussed, and more generally we have tried to re-arranged the section to emphasize the fact that we have focused, in this study, on an *upper bound* of the predictability limit (potential predictability) related to initial uncertainty. An "upper bound" because our estimations for this limit might even decrease in the context of thinner ice. But also an "upper bound" because in practice, predictability is also decreased by all the additional sources of uncertainties (i.e. in the atmospheric surface state, in model parameters, in the ocean model, etc…) that we exclude by design.

The revised text of the last section is too long to copy here entirely but is available in the revised manuscript attached. We hope that our modifications will clarify our interpretations of the results and acknowledge the various limitations of our approach pointed out by the Reviewer.

Ólason, E., and D. Notz (2014), Drivers of variability in Arctic sea-ice drift speed, J. Geophys. Res. Oceans, 119, 5755–5775, doi:10.1002/ 2014JC009897.

Ludwig et al (2026, preprint) Biases, Uncertainties, and Trends in Arctic Sea-Ice Thickness: A Cross-Product Analysis from 1995 to 2023, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-6201, 2026.

Sumata, H., de Steur, L., Divine, D.V. et al. Regime shift in Arctic Ocean sea ice thickness. Nature 615, 443–449 (2023). https://doi.org/10.1038/s41586-022-05686-x

Rheinlænder, J. W., Davy, R., Ólason, E.,Rampal, P., Spensberger, C., Williams, T. D., et al. (2022). Driving mechanisms of an extreme winter sea ice breakup event in the Beaufort Sea. Geophysical Research Letters, 49, e2022GL099024. https://doi.org/10.1029/2022GL099024

2. ***Reviewer:*** *The experimental description in this manuscript encompasses multiple stages (Initialization, single unperturbed reference simulation, extraction of states for P1-P8, spatial perturbation, ensemble branching for BBM and aEVP….). However, I found the narrative somewhat scattered while reading, requiring me to read through the entire text to fully piece the setup together. This is not reader-friendly, because the readers are not familiar with your experimental logic and components. Therefore,*

I strongly recommend that the authors restructure this section to improve its logical flow. Adding a clear table or a flowchart to illustrate experimental description would greatly help clarify the overall framework.

We appreciate the Reviewer's feedback. We propose to add the schematic below to summarize the experimental design, and we hope it can help clarify our approach.

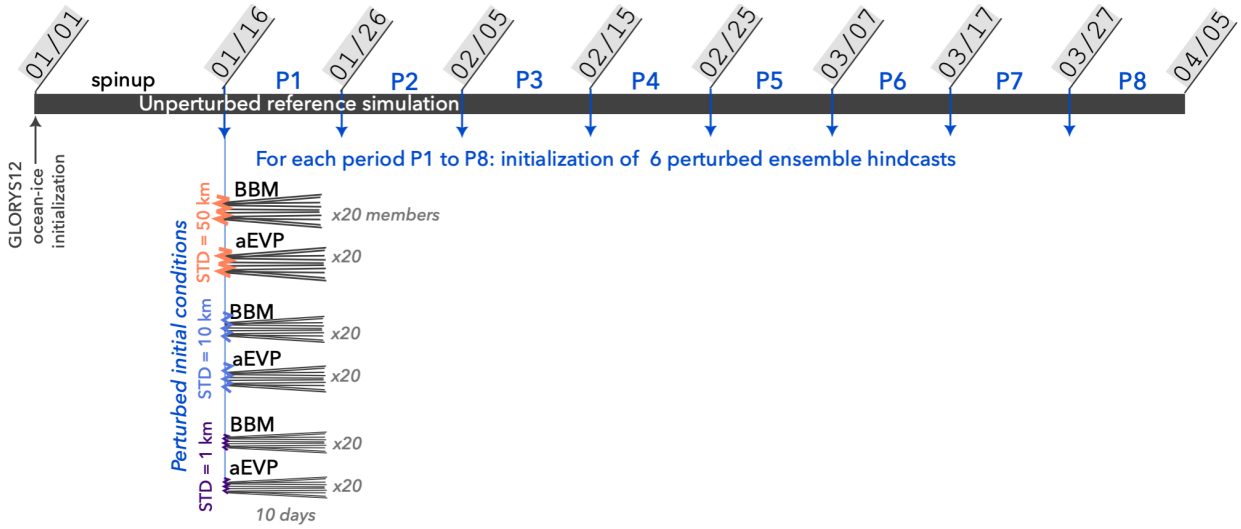

Figure: Summary of the experimental design, representing the unperturbed reference simulation from which are initialized, every 10 days, the 20-member ensemble hindcasts with perturbed ensemble conditions for both rheologies (BBM and aEVP) and three different magnitudes of the initial perturbation each.

3. ***Reviewer:*** *The authors briefly discuss the impact of the 15-day spin-up on sea ice thickness,this short spin-up period is also critical for the ocean module in the ice-ocean coupled model. When initializing the regional ocean model with GLORYS12 reanalysis data, the upper ocean mixed layer and surface currents inevitably undergo a geostrophic adjustment period to adapt to the new grid, bathymetry, and specific surface fluxes. Because ocean drag and ice-ocean flux is also an important term in the sea ice momentum/thermodynamics equation, any transient instabilities in the ocean surface currents during this initial phase could inadvertently force the sea ice, potentially affecting sea ice drift and LKFs. I am not asking that the authors rerun the computationally expensive ensemble simulations. However, to decrease the reader's doubts about this point, I recommend adding some sentences in the Discussion to explore how the transient state of the upper ocean might interact with sea ice rheology. This will greatly enrich the physical depth of the paper and provide better guidance for future modeling efforts.*

We agree that our short spinup might generate some adjustments of the ocean currents, however, we have made the assumption that those potential effects on the ice were of lower order than the sea ice dynamics itself and its response to wind forcing for the short-term predictability (<10 days) that we investigate in this study, focusing in the winter season. Our assumption is based on various previous studies which have documented by scaling analysis kinetic energy budget or observation analysis the fact that, in winter, when the sea ice fraction is high (>0.8) over the whole Arctic basin, the internal stress is a predominant force, and that wind forcing is the main external driver (explains most (>70%) of the sea ice drift variance (Ólason & Notz 2014, Leppäranta 2011,Thorndike & Colony 1982).

We propose to add these lines at the end of section 2.2 (after discussing the impact of the short spinup on the sea ice) to be more explicit on this point:

Regarding the ocean state, the short spinup may similarly result in some adjustment of ocean currents; however, given that our study focuses on short-term predictability (<10 days) in winter conditions with high sea ice concentration covering the Arctic, this is unlikely to significantly affect our results. Under such conditions, internal ice stress is known to dominate the momentum budget and wind forcing accounts for the majority of sea ice drift variance, making the ocean current adjustment a second-order effect compared to the sea ice dynamical response to atmospheric forcing \citep{Olason2014,Thorndike1982,lepparanta2011}.

- Ólason, E., and D. Notz (2014), Drivers of variability in Arctic sea-ice drift speed, J. Geophys. Res. Oceans, 119, 5755–5775, doi:10.1002/ 2014JC009897.
- Leppäranta, M. (2011), The Drift of Sea Ice, Second Edition, Springer Berlin, Heidelberg pp.350
- Thorndike, A. S., and R.Colony (1982), Sea ice motion in response to geostrophic winds, J. Geophys. Res., 87(C8), 5845–5852, doi:10.1029/JC087iC08p05845.
- Maeda , K., Kimura , N., & Yamaguchi , H. (2020). Temporal and spatial change in the relationship between sea-ice motion and wind in the Arctic. Polar Research, 39. https://doi.org/10.33265/polar.v39.3370

4. ***Reviewer:*** *Lines 160-171: The method employs random displacement maps to initialize positional errors. While I am not intimately familiar with the specific technical details of this procedure, I am curious as to how this approach handles local mass conservation and sub-grid ice thickness categories. I suggest a clearer explanation of this in the text.*

Our perturbation method does not enforce local mass conservation, but these perturbations are only applied to the initial state, where uncertainty in the initial mass is a physically reasonable assumption. Moreover, by applying the same random

displacements consistently to all sea ice variables, we ensure that a given physical feature (for example an open lead) is displaced consistently (i.e. with the same $\Delta x, \Delta y$ applied) in all variables from which the given feature is potentially seen (e.g. concentration, thickness, ice drift, etc.). Our approach differs in that respect from previous studies that had only perturbed a single sea ice variable (e.g. sea ice thickness in Mohammadi et al. 2018, sea ice cohesion is Cheng et al. 2020, Korosov et al. 2023 referenced in the manuscript). We propose to add a few lines in section 2.3 of the manuscript to clarify this point:

Our perturbation method does not enforce local mass conservation as it only applies local random displacements on the x and y directions. But these perturbations are only applied to the initial state, where uncertainty in the initial mass is a physically reasonable assumption. Moreover, those displacements are applied consistently across all sea ice variables, ensuring that a given physical feature (for example an open lead) is displaced or distorted identically in all variables in which it appears (e.g. sea ice concentration, thickness, ice drift, etc.). This stand in contrast to previous previous studies that perturbed only a single sea ice variable (e.g. sea ice thickness in \citet{MohammadiAragh2018}, sea ice cohesion in \citet{Cheng2020} and \citet{Korosov2023}).

5. ***Reviewer:*** *Lines 350-365: Extensive paragraphs are devoted to explaining the differences in CRPS curves between specific periods. This distracts from the core message. Please synthesize this section (this also applies to the rest of Sections 3 and 4).*

We thank the Reviewer for their feedback. We propose to remove the details about the different individual periods and simplify as:
"*Comparing the CRPS evolution of the three different amplitudes of initial perturbations (red, blue, and purple envelopes corresponding to the evolution of the large, medium, and small-amplitude perturbations, respectively, in Fig. 8a-b) helps to identify the lead time beyond which predictability is fully lost—that is, the point at which the three envelopes become indistinguishable and the initial errors no longer influence the forecast. For concentration, we find that on average, the three curves are not yet totally indistinguishable after 10 days, suggesting that some predictability remains at that time lag (on average). Individual forecast periods, however, show contrasting behaviors (Fig. 8b), reflecting varying sensitivity to initial errors depending on the pre-existing sea ice state and atmospheric forcing: positional perturbations applied to a heterogeneous initial field produce a larger initial CRPS than those applied to a uniform field, and the growth rate of the CRPS is further modulated by the intensity of wind forcing during the forecast period."*

Reviewer: Minor Comments

Reviewer: *Line 35: Please clearly define the metric or index used for "potential predictability" early in the text to help readers build an intuitive understanding.*

Estimating a potential predictability horizon in this study relies on our ensemble approach and on the probabilistic skill scores (CRPS) computed across several hindcasts initialized with initial conditions perturbed at different amplitudes (small, medium, large). However we feel that introducing these practical details as early as line 35 would disrupt the flow of the introduction, which was intended to remain general at that stage. Our methodology is nonetheless outlined slightly later in the introduction (lines 80–100). We acknowledge however that the connection to the notion of potential predictability introduced around line 35 was not made explicit enough. We therefore propose the following modification to this paragraph starting at line 80:

In contrast, we adopt a complementary approach, examining the potential predictability of the sea-ice system: within a "perfect-model" framework (e.g., Reifenberg and Goessling, 2022), we introduce positional perturbations at initialization into a sea-ice model while assuming both the model and atmospheric forcing are perfect. Using an ensemble approach, we then track how these initial positional errors grow over time and influence the forecast skill across multiple sea-ice-relevant metrics. By systematically varying the magnitude of the initial positional uncertainty from large to small, we assess the highest achievable forecast accuracy for a given magnitude of initial error. ***For a given sea ice variable and associated score, we can thus also estimate the lead time beyond which predictability is fully lost, that is the lead time at which the magnitude of the initial errors no longer influences the forecast accuracy (meaning the smallest initial perturbation has led to the same amount of forecast error as the larger ones). This practical estimate of the predictability limit, derived under the assumption that all sources of uncertainty other than the imposed initial positional error have been eliminated, can be viewed as equivalent to the notion of potential predictability introduced above.***

Reviewer: *Lines 45-46: The introduction mentions that the BBM framework was developed to more realistically capture linear deformation patterns. However, it lacks accessible physical explanation of the core differences between BBM and aEVP. A clearer physical explanation would make this section much more accessible.*

We fully understand the reviewer's comment and propose to clarify the difference between the two rheologies with the following text, added in the revised version of our manuscript (attached to our response):

> *" In elasto-visco-brittle frameworks (including the BBM formulation), unfragmented sea ice is treated as an elastic, damageable solid, while fragmented sea ice is represented as a visco-elastic material in which irreversible deformations act to relax internal stresses. Unlike viscous-plastic frameworks, elasticity is thus an inherent and physically meaningful component of these rheological models. Moreover, elasticity in those models is modulated through a time-evolving damage variable which retains memory of the fragmentation state of the sea-ice cover, a mechanics concept that is absent in standard viscous-plastic rheologies.*
>
> *We also note that the interplay between elasticity and damage — even under the assumption of an isotropic constitutive model — naturally gives rise to strong anisotropy and strain localization, down to the nominal spatial and temporal scales of the model (i.e. grid resolution and time step; see e.g. Rampal et al., 2019; Ólason et al., 2022). Consequently, mechanically coupled fields — including damage, concentration, thickness, and velocity — tend to exhibit pronounced spatial gradients."*

Reviewer: *Lines 108-110: Please clarify if the thermodynamic parameters are same between the two configurations.*

Yes, thank you for pointing out this need for clarification. We propose to modify the text in those lines as follows:

> *"We take advantage of this single modelling framework (i.e. NEMO-SI3) in which either the BBM or the aEVP rheology formulation can be set, to test the impact of rheology on the short-term predictability properties of the system. Hindcast experiments are thus run both with the BBM and aEVP rheologies, while all other settings are kept the same in the modelling framework –namely* ***same sea ice thermodynamic parameters****, same atmospheric forcing, same ocean model parameters, same grid resolution and time-step– allowing for a clean comparison."*

Reviewer: *Lines 103-104: The model configuration is 1/4 horizontal resolution. Given that actual sea ice deformation scales range from hundreds of meters to a few kilometers, I suggest the authors add some discussion about whether higher resolution models would make the discrepancy in sensitivity between aEVP and BBM even more pronounced.*

We are unsure to fully understand the Reviewer's comment.

But this is true that 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 to the resolution, and in particular of the discrepancy in sensitivity between aEVP and BBM, was not explored in this study. This would certainly merit further investigation.

We do not expect, pending further verification, much sensitivity of the BBM simulations to increasing the resolution as such brittle-based rheological framework is capable to reproduce the observed scaling invariance properties of sea ice deformation from the scale of the Arctic basin down to the nominal grid scale, whatever that nominal grid scale is (Rampal et al. 2019). However, the same statement may not be valid for the aEVP simulations since this rheological framework was shown sensitive to the grid resolution, with such scaling properties of sea ice deformation only being exhibited in simulations if run at resolution of about <4km (Bouchat et al. 2022).

Also, increasing the resolution might give the ocean dynamics a growing role: the better resolution of mesoscale eddies would potentially introduce additional variability beneath the ice, potentially acting as an extra source of uncertainty for the sea ice and thus further constraining its predictability limits. But this remains to be verified and quantified in future work.

We have added some lines of discussion on this point in the last section:

"Furthermore, 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 and would certainly merit further investigation, since each rheology is known to respond differently to resolution changes \citep[e.g.][]{Rampal2019, Bouchat2022}. At higher resolutions, ocean dynamics may also play a growing role: the better resolution of mesoscale eddies would introduce additional variability beneath the ice, potentially acting as an extra source of uncertainty for the sea ice and thus further constraining its predictability limits. This however remains to be verified and quantified in future work."

Line 170: Only the sea ice state is perturbed, while the ocean state remains unperturbed. Perturbing only the sea ice immediately breaks the local atm-ice-ocean momentum balance. This maybe small in winter when the area of open water is very small, but I would still

suggest that the authors add a few sentences discussing the potential impact of this initial change on the error growth rate.

We agree that this requires some explanation. In principle, our perturbation method could have been applied consistently to ocean variables in the same way as for sea ice. In practice, however, this would have introduced additional constraints, notably the need to avoid displacing ocean variables beyond bathymetric boundaries, and thus to arbitrarily scale down perturbations with depth and near bathymetry (similarly to what we have done to prevent sea ice displacements at the coast - see Section 2.3). We therefore made the simpler choice of representing only the uncertainty in the location of sea ice relative to the atmospheric forcing structures and the underlying ocean state. The resulting local disruption of the atmosphere–ice–ocean momentum balance is arguably representative of the artificial adjustments induced by ad hoc sea ice data assimilation in operational systems, where the ocean and atmosphere states are not necessarily corrected consistently.

We propose to add those comments in the second paragraph of section 2.3:
In this study, only the sea ice state is perturbed; the ocean state is left unperturbed. While the same displacement maps could in principle be applied consistently to ocean variables, this would require additional constraints to avoid displacing water masses beyond bathymetric boundaries (analogously to the coastal masking applied to sea ice, as described below). The simpler choice that is made here is to only perturb the position of the sea ice features with respect to the atmospheric forcing and the ocean underneath can be viewed as representative of the inconsistencies that may arise in the atmosphere–ice–ocean momentum balance in operational systems, when sea ice and ocean states are not corrected consistently by data assimilation.

And this sentence in the last section:

By design, only the initial sea ice state is perturbed (the ocean state and atmospheric forcing are not), in order to isolate the response of the coupled ice–ocean system to initial positional uncertainty in the sea ice features alone. These initial errors can be viewed as representative of the inconsistencies that may arise in the atmosphere–ice–ocean momentum balance in operational systems when sea ice and ocean states are not corrected consistently by data assimilation.

Reviewer: *Lines 222-224: There is a persistent 10°bias in the aEVP trajectories, but it lacks some physical explain. Could you add sentences to discuss if this is related to aEVP's strong viscous forces interfering with the ice pack's geostrophic adjustment to the Coriolis force? or is it due to other processes?*

This is an interesting point indeed. We have also noticed this statistical ~10° bias in the angle of the Lagrangian trajectories shown in Fig. 3b with the aEVP rheology. However,

some preliminary additional investigations (briefly described below) make us believe that its physical interpretation is non-trivial and requires dedicated investigations to be fully understood and would fall beyond the scope of the current paper. But importantly, this bias does not affect the predictability diagnostics and results presented in this manuscript, since those are computed within a "perfect model" framework to quantify the intrinsic limits of predictability in each model configuration. In the case of the Lagrangian trajectories specifically, predictability (Fig. 14) is evaluated by the ensemble spread of the trajectories within a given model at a given time lag, and is therefore insensitive to any systematic model bias against true observations. The purpose of Fig. 3 in our paper was only to illustrate to the reader that both configurations produce a sufficiently realistic and well-tuned representation of the sea-ice drift at local scales before focusing on their predictability properties in a “perfect-model” framework. We propose to modify the text to clarify this point at the end of section 3:

> *Overall, this evaluation confirms that both configurations simulate the observed drift with sufficient realism to be used in the predictability analysis that follows. Note also that, since the predictability diagnostics of this study are designed within a "perfect model" framework, they are insensitive to any systematic model bias against true observations. In particular, the ensemble spread of the simulated Lagrangian trajectories is evaluated for each model configuration and then inter-compared in Sect. 4.3 to illustrate the difference in their predictability properties, but independently from true observations.*

Regarding our preliminary additional investigations to better understand the ~10° angle bias seen with the aEVP rheology in Fig. 3b:

- First, we have double-checked that the wind-stress vectors were well aligned between the two simulations, as expected, since forced by the same ERA5 wind forcing.
- Then we have found that the angular difference between BBM and aEVP is present directly in the hourly sea ice drift vectors on the model's native grid. More specifically, we find that sea ice drift vectors in both aEVP and BBM simulations are deflected to the right of the wind stress, but the deflection angle seems systematically slightly larger in the aEVP simulations than in the BBM simulations in winter, and hence the angle bias found in Fig. 3b from the lagrangian trajectory diagnostics. Interestingly, the value of the deflection angle to the right appears to vary between winter and summer seasons.

  -> Taken together, these elements seem to point to the angular difference originating from a distinct dynamical response to wind stress through the rheology. We tentatively attribute this bias to a difference in how the two rheologies mediate the balance between wind stress, internal ice stress, and Coriolis forcing. An additional contributing factor may be the integrated response of sea ice drift direction within a

semi-enclosed basin, where coastal geometry constraints and the mean large-scale wind field may favor particular fracturing -and hence drift- directions or deflection angles. Disentangling these contributions robustly should be the subject of a dedicated work.

We propose to complement our initial statement (l. 223) as follows:

"while a 10° clockwise bias remains in the aEVP trajectories relative to the observations ***(likely denoting a distinct dynamical response to wind stress through each rheology).*** *"*

Reviewer: *Lines 430-434: Defining high-deformation events using the 95th percentile is statistically sound. I suggest also providing the absolute physical threshold values would be highly beneficial for readers' intuition.*

Good point. We propose to clarify the text as follows:

High-deformation events are defined as those when the hourly deformation in a model cell exceeds a threshold set at the 95th percentile of the hourly deformation distribution ***(respectively 0.06 day^{-1} and 0.05 day^{-1} for BBM and aEVP rheology)*** *during the winter season (January–March 1997) in the pack-ice region outlined in Fig. 1. This threshold approximately corresponds to the deformation value above which the ice material enters the plastic regime in the aEVP case and the visco-elastic regime in the BBM case.*

Reviewer: *Lines 378-382: The authors found that thickness behaves entirely differently from other variables, retaining predictability beyond 10 days. Could the authors explain why thickness memory persists so long?*

We interpret this contrasting behavior in thickness predictability, relative to the other variables, as a consequence of the physical processes governing thickness, which are primarily thermodynamic processes, operating on longer timescales than the dynamical processes that instead drive sea ice drift, deformation, and concentration at first order.

We propose to add this line in section 4.2.1:

We interpret this longer predictability limit of the thickness relative to the other sea ice variables as a consequence of the physical processes governing thickness, which are primarily thermodynamic processes, operating on longer timescales than the dynamical processes that drive sea ice drift, deformation, and concentration at first order.

We also propose to add this in the last section:
For these quantities, predictability rapidly degrades, and the CRPS evolution loses memory of the initial perturbations after only a few days: five to ten days for concentration, one to five days for deformation and drift, and longer than ten days for thickness ***(as its main driven is thermodynamic, associated with longer timescales, rather than dynamics)***.

Reviewer: *Technical corrections*

Line 7: We investigates -- We investigate

The correction has been made.

Line 131: Janurary -- January

The correction has been made.

Line 144: heterogenities -- heterogeneities

The correction has been made.

Line 421: rapidely -- rapidly

The correction has been made.

Figure 8 caption: hincasts -- hindcasts

The correction has been made.

Figure 12 caption: intialized -- initialized

The correction has been made.

Most Figures' labels and legends are much too small. Some require more than 200% magnification to read. Please increase the font sizes.

Point taken. The other Reviewer also made a comment on the resolution of the figures. We have provided improved figures (resolution and font sizes) for the revised manuscript attached.

While improving the readability of the figures, we have also noticed that we had made typos in naming the panels of the last row of Figures 9 and 11 (m, n, o and p instead of i, j, k, l). We have now corrected them on the figure and in the legend.

Figure A1, A2: axis lacks units.

Good point, the information is in fact missing on all the maps shown in the manuscript. Those numbers are grid points of the native model grid. We will mention it in all the captions where this is relevant. We have also noticed that the legend "x" of the x-axes was lacking on Figures A1 and A2. It is fixed in the revised manuscript.

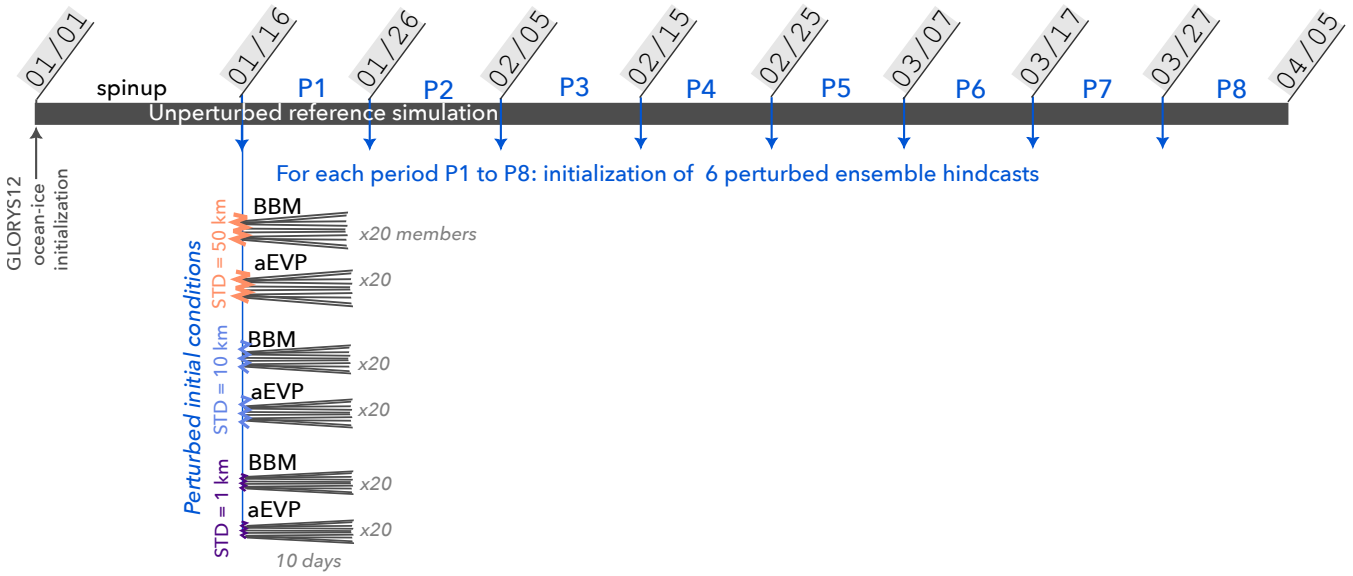

Figure 1. Summary of the experimental design, representing the unperturbed reference simulation from which are initialized, every 10 days, the 20-member ensemble hindcasts with perturbed ensemble conditions for both rheologies (BBM and aEVP) and three different magnitudes of the initial perturbation each.

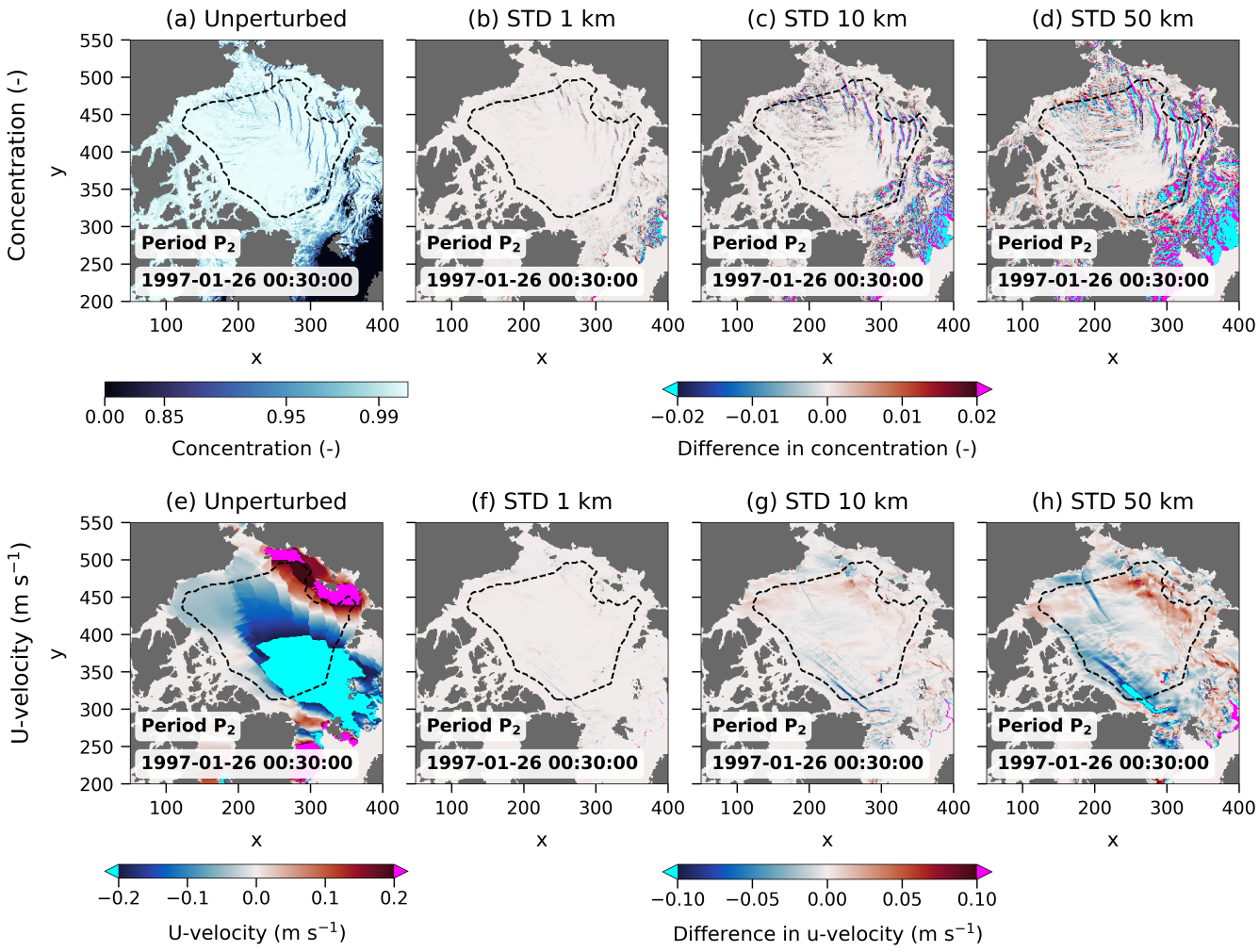

Figure 2. Unperturbed initial states in (a) sea ice concentration and (e) the U-component of sea ice velocity at the beginning of period P_2 (t_0: 1997-01-26) from the unperturbed reference experiment with the BBM model. Panels (b-d) and (f-h) show the difference in the initial states of two perturbed ensemble members for concentration and the U-component of sea ice velocity, respectively. Panels (b,f), (c,g), and (d,h) correspond to initial perturbations scaled for a standard deviation (STD) of 1 km, 10 km and 50 km, respectively. The black dashed line defines the boundaries of the region over which the ensemble scores for the pack ice (Sect. **??**) are computed. Axes indicate grid point indices on the model's native grid.

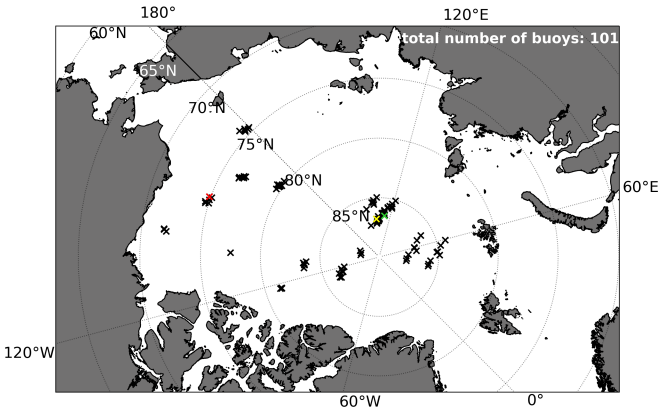

Figure 3. Location of the observed IABP buoys at initial time of the P1,...,P8 hindcast periods (start dates from the 16th of January to the 27th of March 1997, see Table **??**). 101 IABP buoys are considered in total. The colored "x" markers highlight the initial position of the few buoys taken as examples in the following (Figs. 5 and 14).

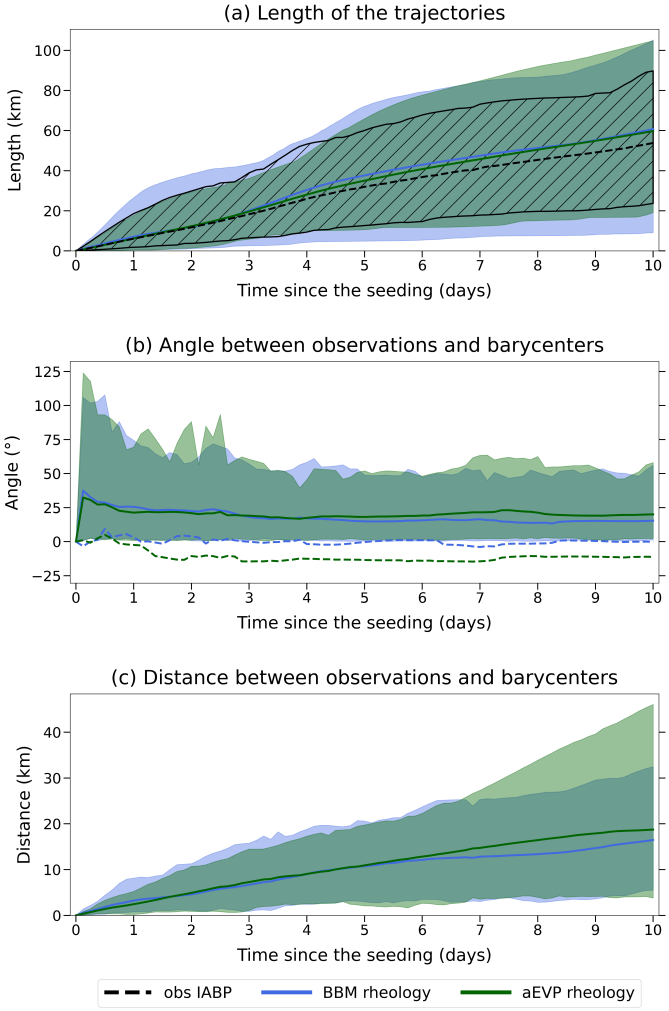

Figure 4. Temporal evolution of different properties of the simulated buoys compared to the observed IABP buoys: (a) length, i.e., the distance covered by the buoys, (b) angle between the two lines drawn from the seeding position to, respectively, the observed buoy and the barycentre of the N = 20 simulated buoys of each ensemble hindcast, and (c) the direct distance between the observed buoy and the barycentre. The results of BBM and aEVP are colored blue and green, respectively, with the thick lines for the ensemble means of all the buoys over all eight 10-day periods, while the shaded envelopes indicate the 5%-to-95% percentile ensemble distribution. In (a) the distance covered by the observed buoys is with the black line (mean) and the hatched envelope (5%-to-95% percentile). In (b), the envelopes and solid lines are computed considering the absolute value of the angle, while the dashed lines correspond to the mean value of the angle.

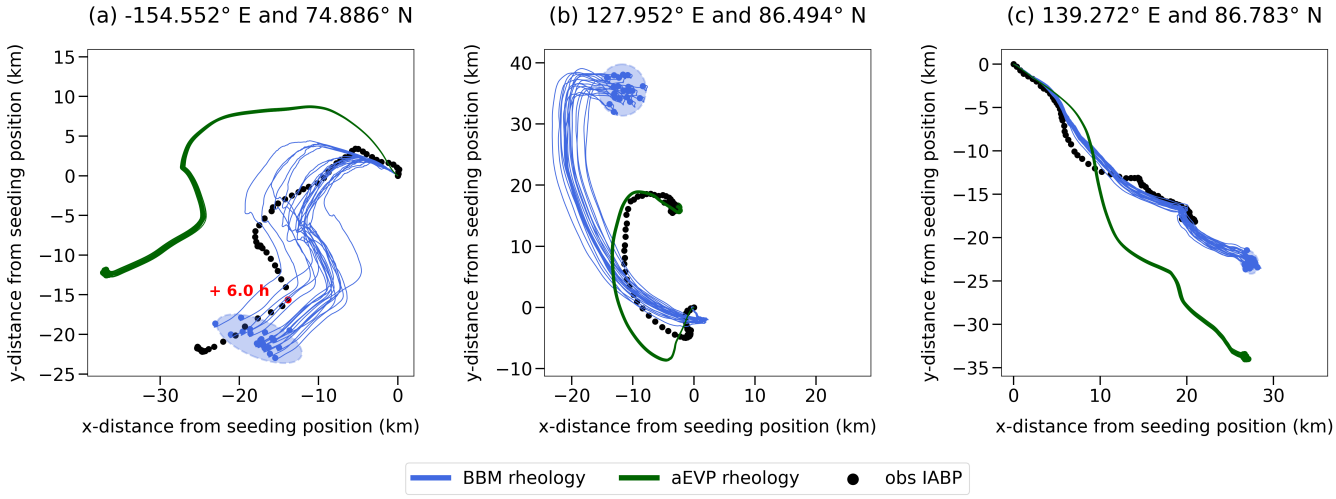

Figure 5. Three examples of observed IABP trajectories (black) and corresponding virtual trajectories simulated from the BBM ensemble hindcast (blue, 20 members) and aEVP ensemble hindcast (green, 20 members) seeded at midnight on: (a) March 7th, at -154.552° E;74.886° N (red cross in Fig. 3), (b) February 15th, at 127.952° E;86.494° N (green cross in Fig. 3), and (c) February 5th, at 139.272° E; 86.783° N (yellow cross in Fig. 3). The observed IABP trajectories are plotted with a plain black circle every 3 hours. In case of missing values, the circle following the gap is shown in red, with the corresponding time gap. The simulated ensemble trajectories are plotted as colored curves, and only the final positions after 10 days are marked as plain circles. The shaded ellipses represent the 95% confidence regions of the final positions, assuming a bivariate normal distribution.

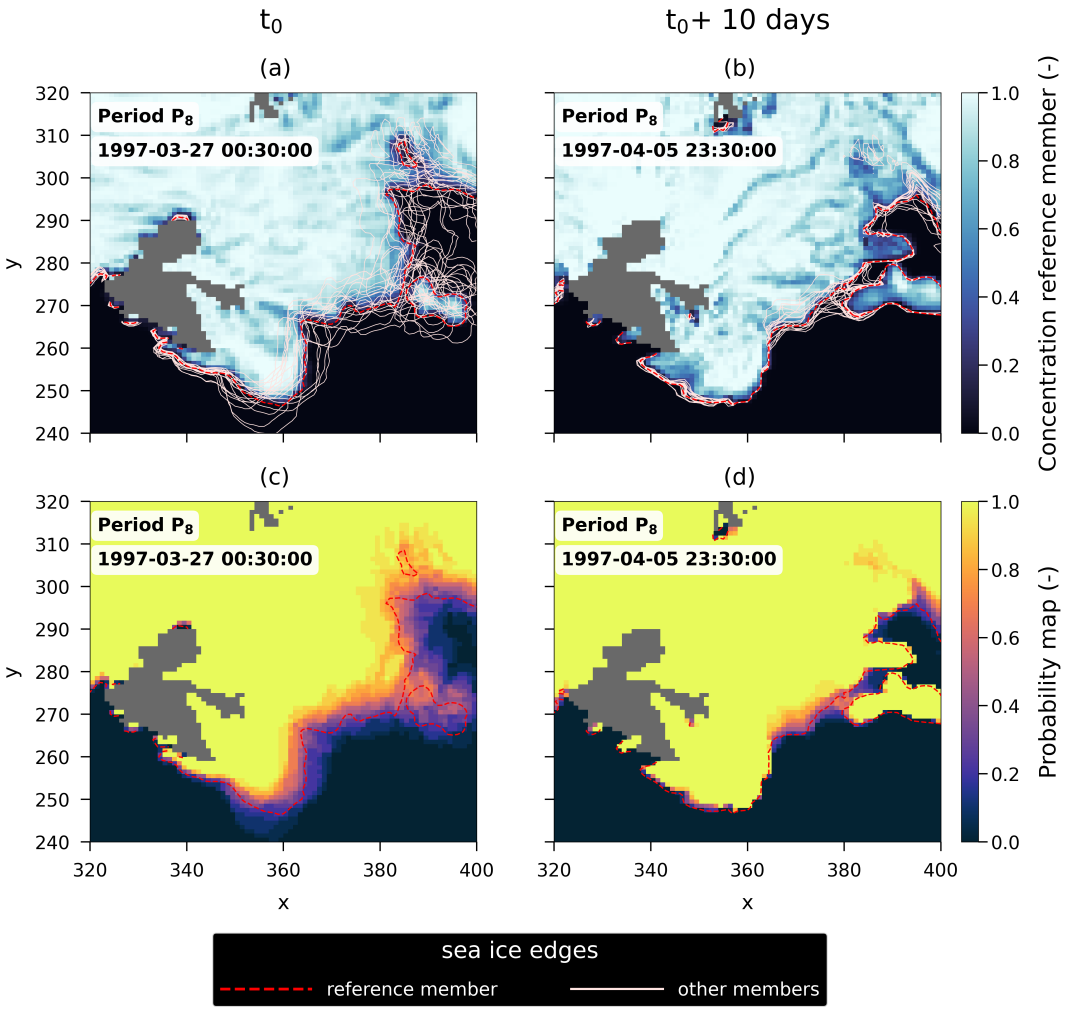

Figure 6. Example of the dispersion of the sea ice edge position (pink and red lines) in the Svalbard region in the BBM ensemble members for period P_8 and from the 50-km-scaled perturbation, at initial time t_0 and after 10 days (a,b respectively). The hourly sea ice concentration field of the reference member is shown in the background for their respective lead time (shading in a, b). The corresponding maps of probability $P[c > 0.15]$, computed from the 19 members excluding the chosen reference member (whose ice edge is shown in red) are also plotted at initial time and after 10 days (c,d resp.) to illustrate the methodology to compute the SPS metrics (see text for more details). Axes indicate grid point indices on the model's native grid.

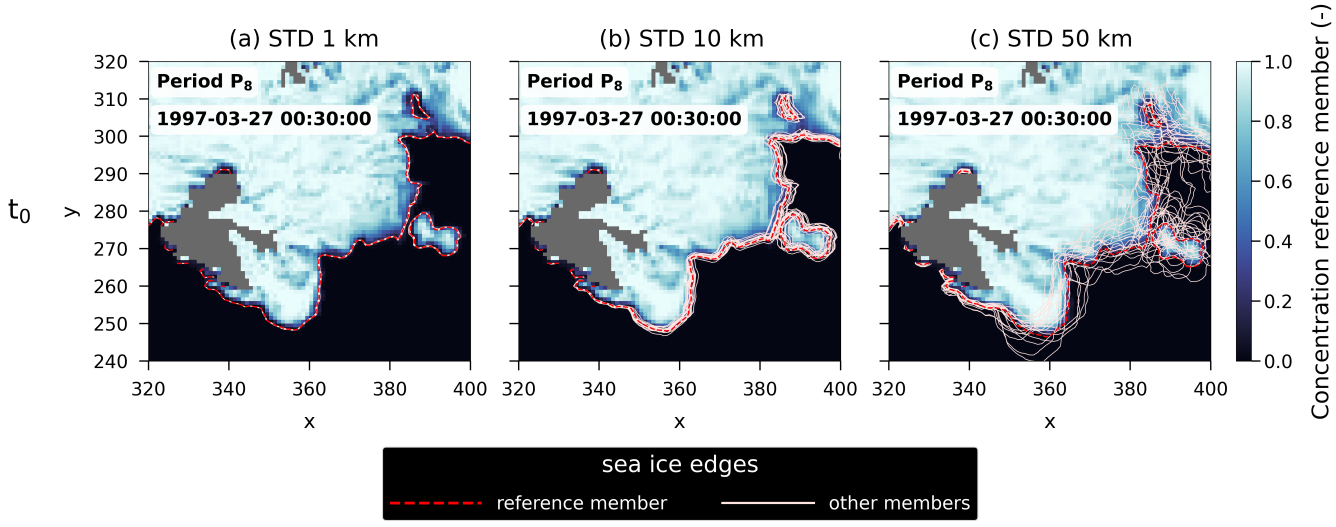

Figure 7. Initial spread of the local sea ice edge position in the Svalbard region during period P_8 (1997-03-27) for an initial perturbation scaled to a standard deviation (STD) of (a) 1 km, (b) 10 km, and (c) 50 km. The ice edge of the reference ensemble member (used here to illustrate the method) is indicated by a thick dashed red line, while the remaining 19 members are shown as thin pink lines. The corresponding hourly sea ice concentration field of the reference member is displayed in the background (shading). Axes indicate grid point indices on the model's native grid.

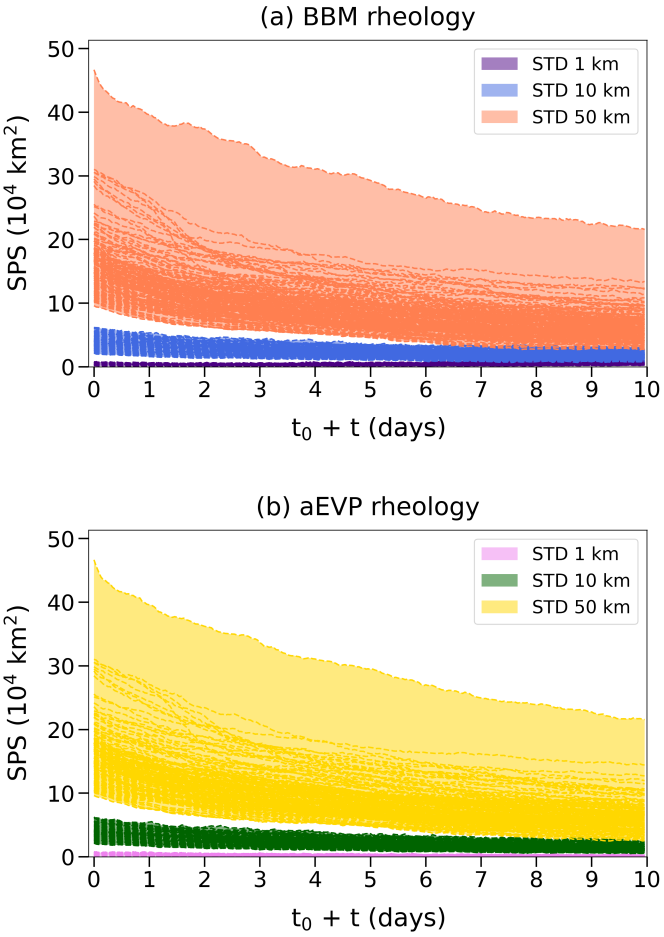

Figure 8. Temporal evolution of the SPS score computed over the whole Arctic, for BBM and aEVP hindcasts (a,b resp.). The plotted curves correspond to the SPS scores computed for each of the eight forecast periods, and for each ensemble member taken alternatively as the pseudo-truth, resulting in total in 8 Periods $\times$ 20 scores = 160 curves. The colors correspond to the three different amplitudes of initial perturbation.

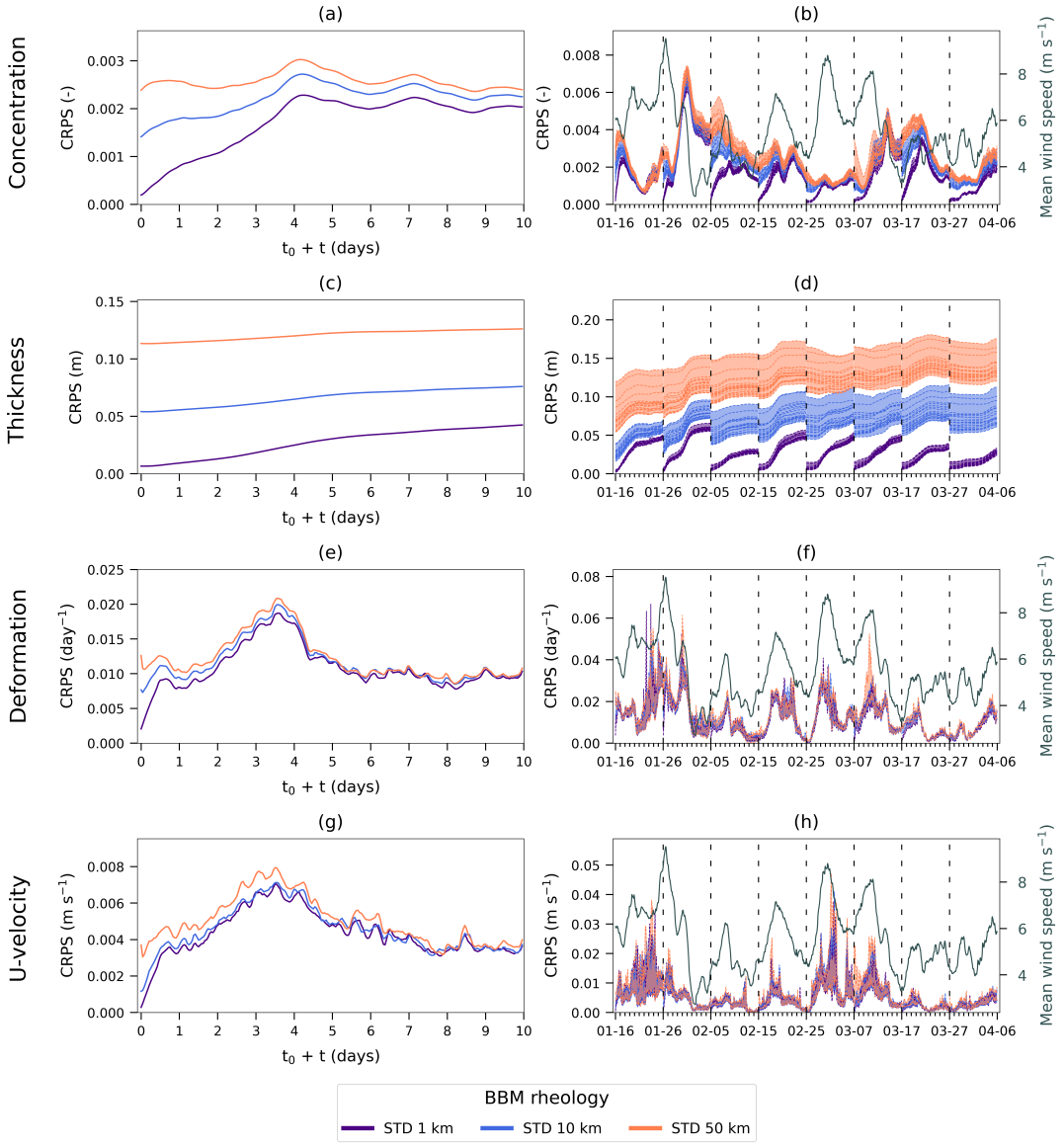

Figure 9. Temporal evolution of the CRPS metric in the BBM hindcasts for the three types of initial perturbations, shown in average for the eight periods P_1,...P_8 (left column) and separately for each period (right column). The CRPS metric is computed for (a,b) sea ice concentration, (c,d) thickness, (e,f) deformation and (g,h) velocity along the x-axis. The colors correspond to the three amplitudes of initial perturbation. The solid line on (b, f, h) stands for the mean wind speed averaged over the same domain as the CRPS metrics (cf Fig. 2). The colored envelopes (right column) correspond to the min-to-max of the CRPS scores computed for each member taken alternatively as the pseudo-truth. The thin curves on panels (b) and (d) show the scores corresponding to each member taken alternatively as the pseudo-truth.

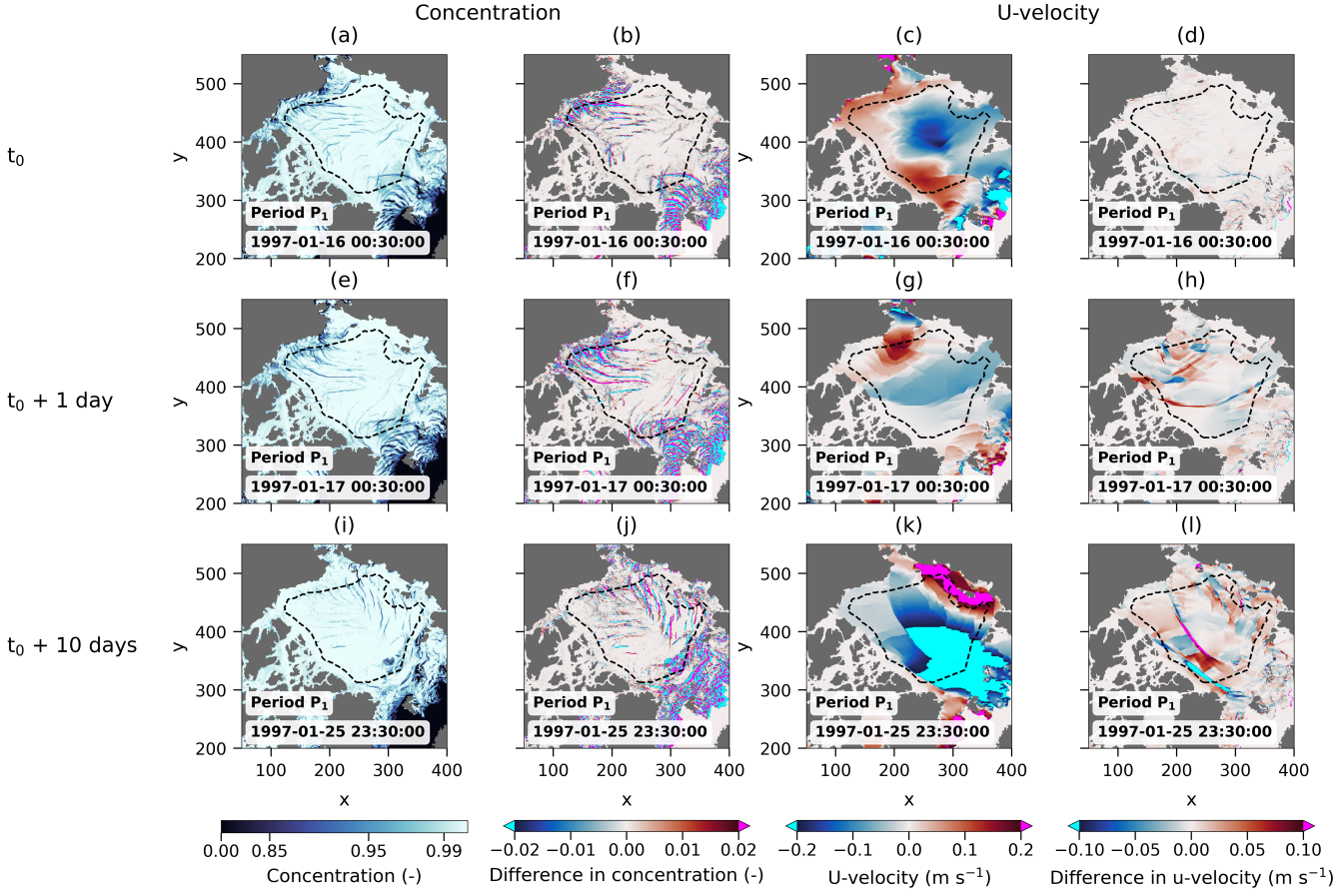

Figure 10. Sea-ice concentration and U-component of sea-ice velocity (a,e,i and c,g,k, resp.) from the BBM hindcasts initialized with the 10-km perturbation at three lead times during period P_1: (a–d) initial time, (e–h) after one day, and (i–l) after ten days. Panels (b,f,j) and (d,h,l) show, respectively, the differences in concentration and in the U-component of velocity between two perturbed ensemble members. The black dashed line defines the boundaries of the region over which the CRPS metric is computed. Axes indicate grid point indices on the model's native grid.

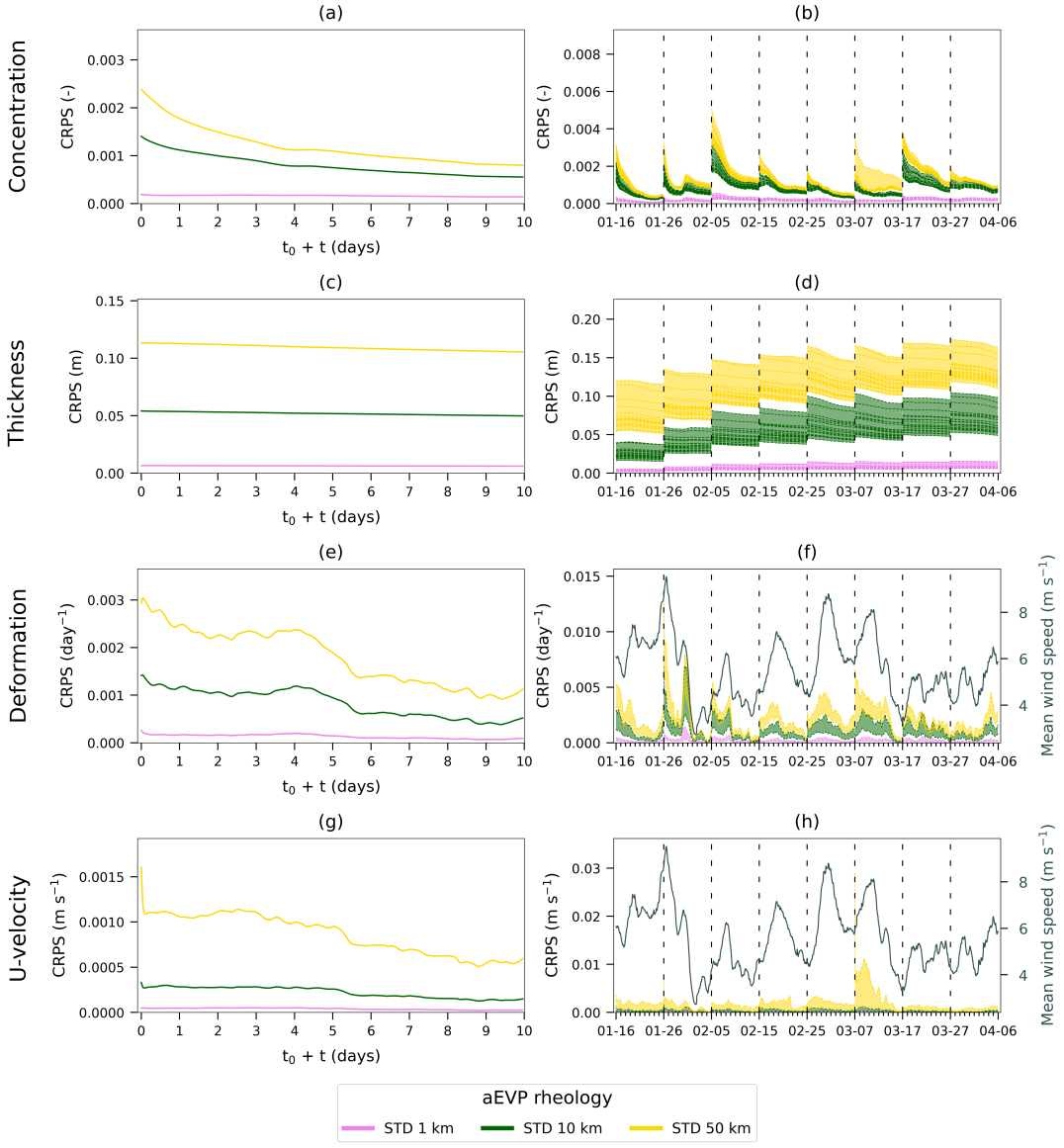

Figure 11. Same as Fig. 9 but for the CRPS metric from the aEVP hindcasts initialized with the 10-km perturbation.

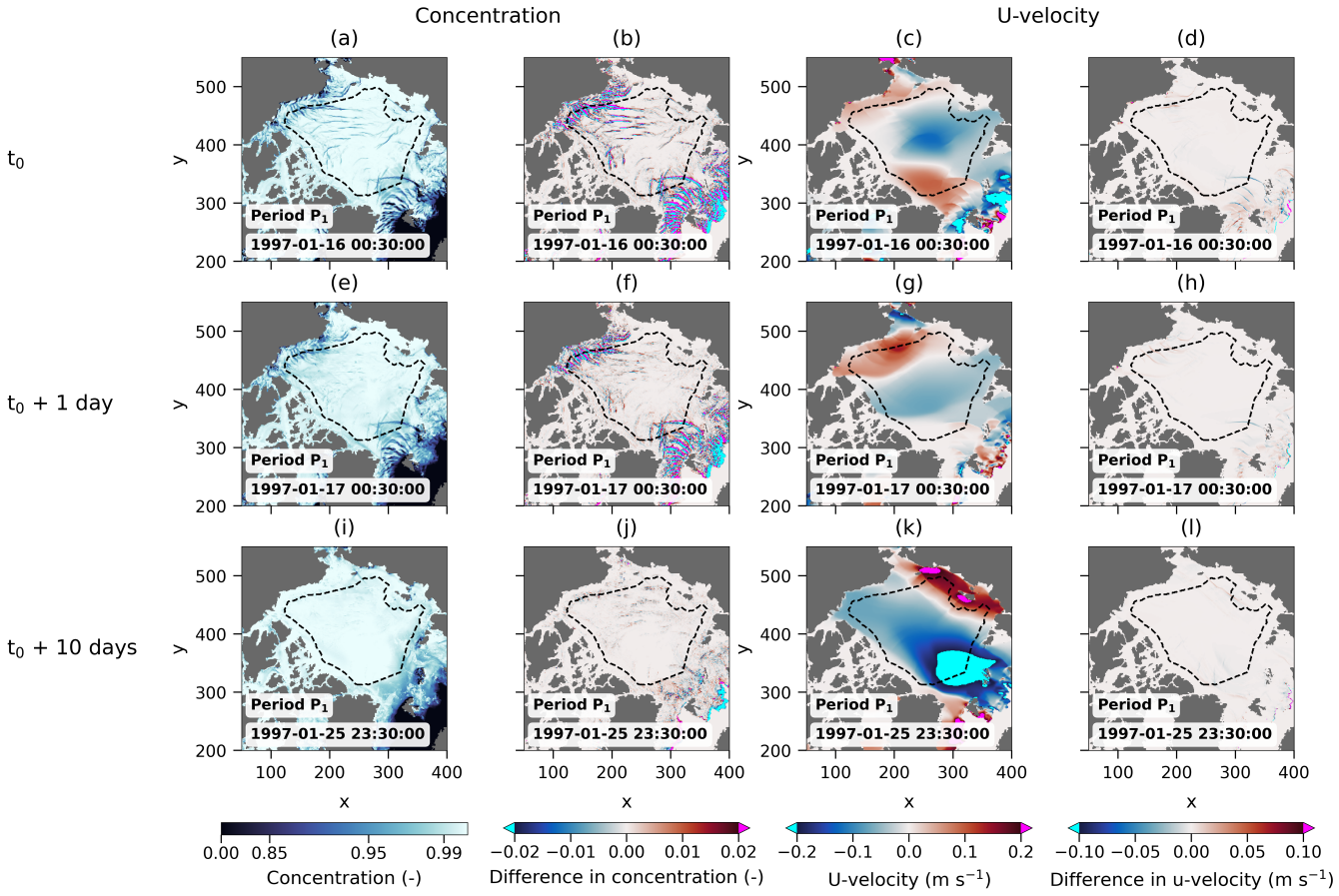

Figure 12. Same as Fig. 10 but from the aEVP hindcasts initialized with the 10-km perturbation.

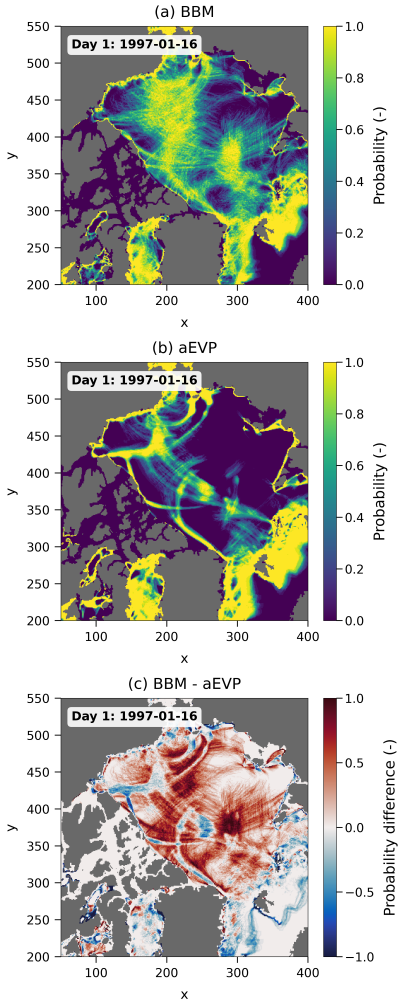

Figure 13. Maps illustrating the likelihood of experiencing at least one high-deformation event within 24 hours (here plotted for the first day of period P_1 as an illustration) in each model cell in the BBM and aEVP ensemble experiments (a,b resp.) initialized with STD = 50 km positional perturbations. The high-deformation events are defined as those when the hourly deformation in a model grid cell exceeds a threshold set at the 95th percentile of the hourly deformation distribution over the winter season in the pack-ice region (see text for details). Panel (c) shows the difference in probability between panels (a) and (b). Axes indicate grid point indices on the model's native grid.

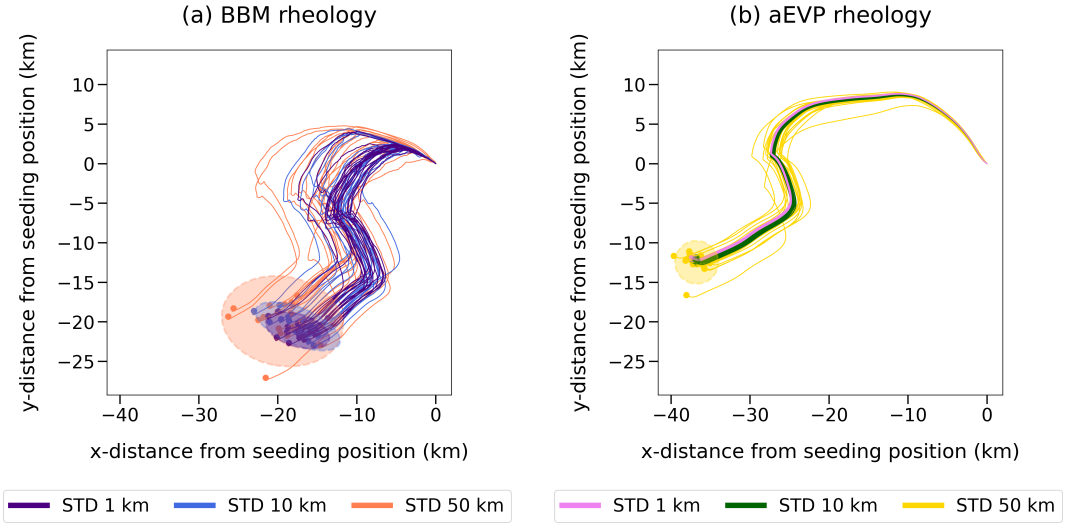

Figure 14. Example of virtual buoy trajectories simulated from the BBM and aEVP ensemble hindcasts (a,b, respectively) from an seeding on 7th March 1997 at midnight at -154.552° E and 74.886° N (red cross in Fig. 3). The coloured circles highlight the final positions after 10 days and the shaded ellipses represent the 95% confidence regions of the final positions, assuming a bivariate normal distribution. The colors correspond to the rheology and amplitude of the initial perturbation of the experiments.

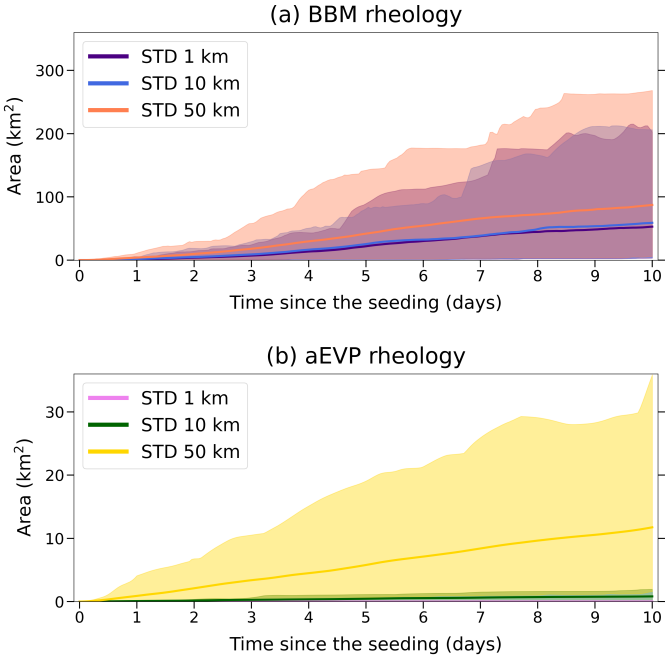

Figure 15. Temporal evolution of the spread of the virtual buoys from the BBM and aEVP ensemble hindcasts (a,b, resp.), as measured by the area of the ellipse defined by the 95% confidence contour of the distribution (assumed to be bivariate normal). The colors correspond to the rheology and amplitude of the initial perturbation of the experiment. The thick lines show the ensemble mean of all the buoys over all the eight 10-day periods, while the shaded envelopes indicate the 5%-to-95% percentile ensemble distribution. Panel (a) uses a y-axis range 10 times larger (area in km^2) than panel (b).

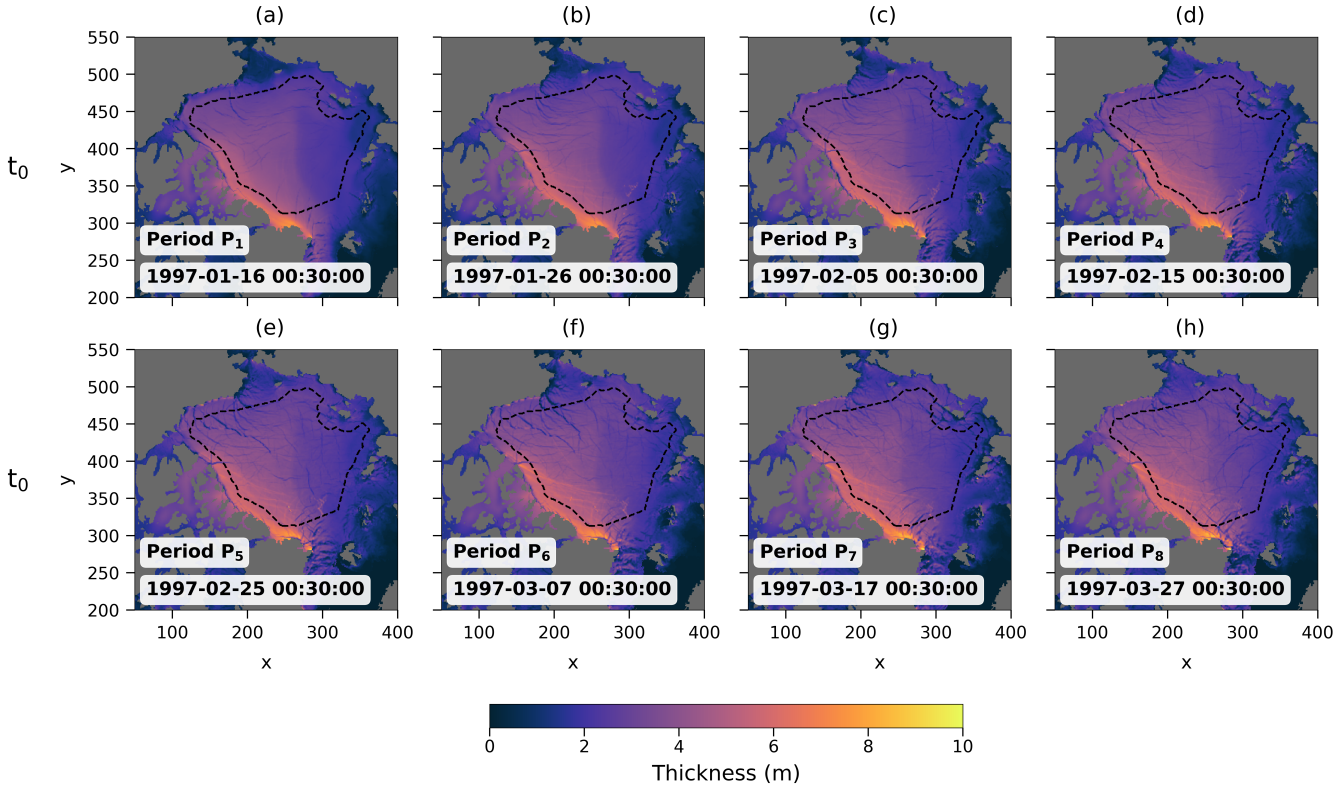

Figure 1. Sea ice thickness maps at initial time (t_0) of each period P_1,...,P_8 from the reference (unperturbed) experiment (see Sect. **??** for more details). The black dashed line defines the boundaries of the region over which the CRPS metric is computed in Sect. **??**. Axes indicate grid point indices on the model's native grid.

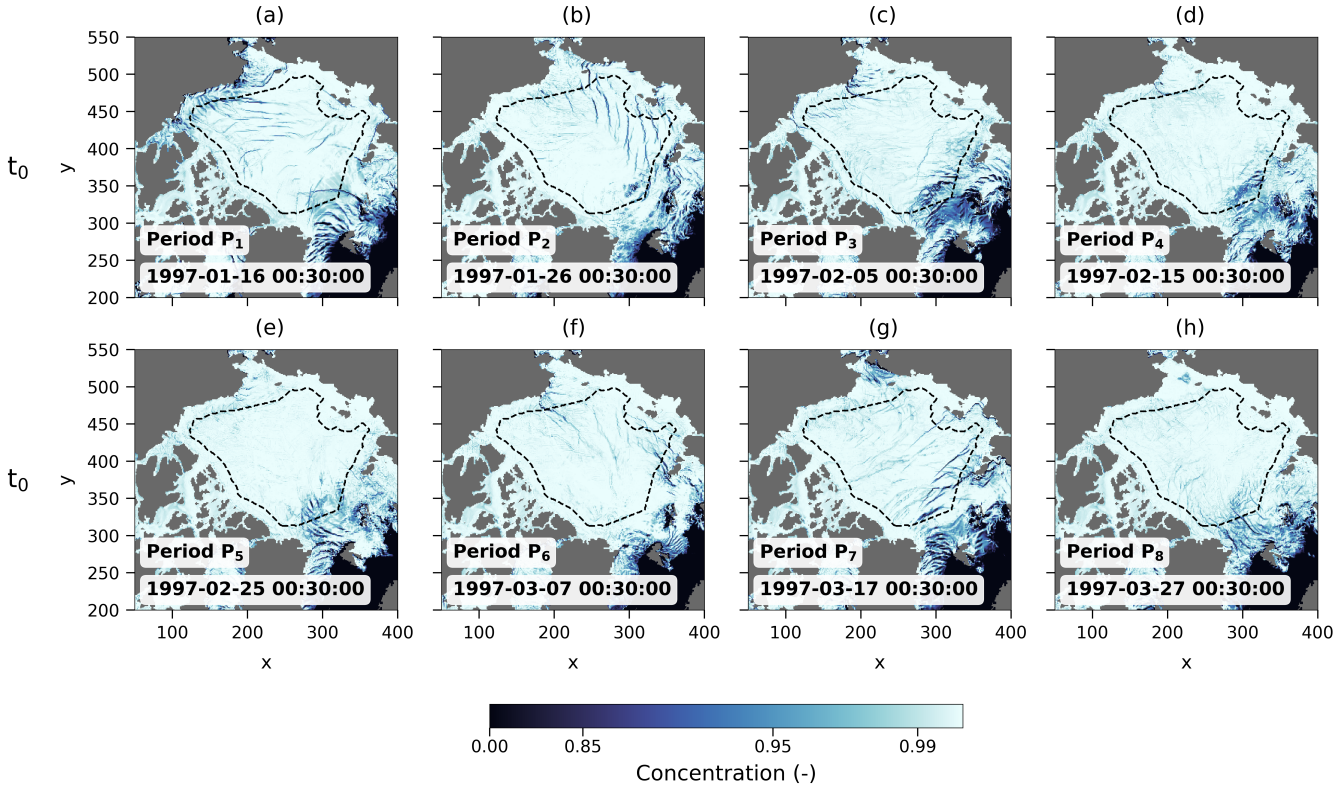

Figure 2. Sea ice concentration maps at initial time (t_0) of each period P_1,...,P_8 from the reference (unperturbed) experiment (see Sect. **??** for more details). The black dashed line defines the boundaries of the region over which the CRPS metric is computed in Sect. **??**. Axes indicate grid point indices on the model's native grid.