Replies to Reviewer 2

In this paper, damage and concentration – assumed to be a function of satellite-derived total sea ice deformation – are assimilated using a simple nudging method in the Brittle-Bingham- Maxwell (BBM) model – an off-spring of the EB/MEB model – for short-term sea ice forecast of Linear Kinematic Features. Results show that assimilation of sea ice concentration improves the skill of a short-term forecast for up to five days lead-time in comparison with persistence or a free simulation. The assimilation of damage on the other hand does not improve the skill of the forecast. It is argued that that the reason for the lack of sensitivity to damage assimilation is because the healing of ice (that reduces the damage) has a timescale that is too fast (24 hrs) compared with the desired lead time of the forecast (O(days)).
The paper is well written. The organisation and depth of the paper however is lacking. The text refers to a non-existent appendix on several occasions; the sensitivity to assimilation parameters consists of four 1 or 2-line paragraphs; there are multiple unsubstantiated statements; the results section presents inconsistent results (sea ice concentration larger than one) and focuses on small scale features in order to justify broad conclusions. The authors state that neXtSIM is not ready to assimilate deformation yet and that the paper is a proof-of-concept. There is no problem with this approach, but the authors have pushed this paper out for review much too early. I recommend that the paper be rejected for the moment and that authors be encouraged to resubmit later, a substantially different version of the paper that addresses (not rebut) the comments below. I am recommending a "reject with encouragement to resubmit" simply to give the author time to properly address the comments.

We appreciate the reviewer's constructive comments. As requested by the reviewers, more experiments for evaluating predictability of sea ice deformation were run (please see the replies to the first reviewer for details) in the extra time provided by the editors. The manuscript is undergoing a significant revision that can be accomplished without a rejection. All the requirements and suggestions are duly addressed without a rebuttal.

Overstatements:

Line 304: "The viscous-plastic (VP) rheology used in MITgcm is known to have a less realistic slower time evolution of LKFs (Hutter et al., 2018) than the BBM rheology in neXtSIM (Olason et al., 2022). As a result, the sea ice simulated by the BBM rheology has more rapid error growth (loses skill faster) due to the correctly resolved intermittent ice motion and localised ice deformation." If the intermittent motion is correctly resolved, why is the error growing faster?

Generally speaking, the higher is the intermittency of a stochastic process, the lower is the predictability. Therefore, a model with a low intermittency will have a tendency to overestimate intrinsic predictability.

However, our case neXtSIM is probably overestimating the intermittency. Indeed, looking in details on Fig. 13 in (Bouchat et al.,2021) we can see that the spatial structure function of neXtSIM matches the RGPS observations very well, whereas the temporal one is overestimated by 3 - 5 %. The MITgcm at 4.5 km (same or similar to as used in (Mohammadi-Aragh et al., 2018)) is worse in the spatial domain and seems better in temporal domain.

Line 44: "In a recent model intercomparison paper (Bouchat et al., 2021), only one model, neXtSIM (neXt Generation Sea Ice Model, Bouillon and Rampal, 2015a; Rampal 45 et al., 2016), proved to be capable when run at the same spatial resolution as the available observations (i.e. $\sim$ 10km) to simulate the observed probability distribution, spatial distribution and fractal properties of sea ice deformation." This is incorrect. One 10-km McGill run was able to reproduce the PDF of divergence and multiple 10-km models were able to reproduce temporal scaling and spatio-temporal coupling. The only place where neXtSIM did perform better compared to other 10-km models is the spatial scaling. Intermittency does not come from the brittle parameterization in the EB/MEB/BBM. All models participating in SIREX (VP, EVP, EB) show intermittency (Bouchat et al., 2021). The source of intermittency in observed deformation is still an unresolved issue. The authors must clarify what they mean by "intermittency. If the temporal scaling exponent is used to discuss the LKFs intermittency, this is incorrect. If instead the authors are referring to LKF growth rate and lifetime, this is also incorrect. In SIREX2, it is shown that no apparent link is present between the LKF growth rates, lifetimes and the temporal scaling/multi-fractal parameters. The intermittency is revealed by the quadratic nature of the structure function.

Following (Rampal et al., 2008) we consider the exponent α as a measure of the degree of intermittency.

Indeed, SIREX indicate that all models have some degree of intermittency (a slope of the temporal scaling line in log-log space) but only two low resolution models (RASM-WRF-EVP and RASM-WRF-EAP) are as close to the RGPS observations of total deformation as neXtSIM (Fig. 11.a in Bouchat et al., 2021). It is interesting to note that RASM-POP-EAP, which has exactly the same setup as RASM-WRF-EAP except for using atmospheric forcing with much lower temporal resolution, has much lower deformation rates, probably indicating the role of the atmospheric forcing in intermittency.

Nevertheless, the slope of neXtSIM temporal scaling is higher than the RGPS observations by 3 - 5 % (Fig 11.c and Fig 12.c), probably indicating overestimation of intermittency by neXtSIM. The text is rewritten as follows:

"In a recent model intercomparison paper (Bouchat et al., 2021), neXtSIM simulations (neXt Generation Sea Ice Model, Bouillon and Rampal, 2015a; Rampal 45 et al., 2016), ranked among the best for simulating the observed probability distribution, spatial distribution and fractal properties of sea ice deformation, even though it operates on a low resolution grid of 10km. Analysis of spatial and temporal scaling (Fig. 13 in Bouchat et al., 2021) shows that the spatial structure function of neXtSIM matches the RGPS observations very well, whereas the temporal one is overestimated by 3 - 5 %, probably indicating overestimation of the intermittency by neXtSIM."

Line 209: "Thus, due to its rheology, neXtSIM is able to extrapolate and create realistic connections between the observed and assimilated pieces of LKFs." This is not demonstrated in the manuscript. I believe any rheology that assimilates sea ice contration (A) will show LKFs in line with observations. See Major Points below.

We agree and the sentence is removed.

Major Points:

On two occasions, the authors are referring to an Appendix that is not included in the paper. The appendix must be included.

Unfortunately, the Appendix was not added due to technical reasons. It is added in the revised manuscript.

Equation 7-10: Sea ice concentration (A) increases in convergence (until A=1) and decreases in divergence; sea ice concentration can also increase or decrease in shear. A single dependency of A on eps_tot is therefore missing events where convergence is present along LKFs. The damage (d) dependency on eps_tot is more realistic, but the authors argue that assimilation of A is useful to increase predictive skill, whereas d is not. See below for more on this topic. The single dependence of A on (eps_tot) must be justified.

We assume that all deformation events (convergence, divergence and shear) indicate presence of weaker ice that may continue to be deformed. Ice weakness is simulated in neXtSIM by decreased concentration or increased damage. Observation of any deformation components (including convergence) is interpreted in the assimilation procedure as an increase in ice weakness and, therefore, decrease in concentration or increase in damage. We cannot find reasoning for why weakening of ice due to convergence should be different from weakening due to divergence or shear and therefore suggest that the total deformation is a good proxy for detection of weak ice and a single dependence of A (and d) on the total deformation can be used.

It should be added, that we have two ice categories in the model: older ice, whose concentration is used in rheology and younger ice, which is formed during water freezing and is converted to older ice only after exceeding a threshold in thickness. Only the older ice is updated in the assimilation procedure and the total ice concentration remains the same.

Corresponding explanations are added in Section 3.3.

Line 145: The value of a1 in F_A(eps_tot) is found through sensitivity experiments using the same BBM model. This is a circular argument. This functional dependence must be derived from sea ice concentration and total deformation derived from passive microwave and Sentinel. I believe the author will find that the functional dependency is not a simple linear relationship. The author must at least show this relationship from observations and acknowledge the simplicity/caveat of the approach.

This is not a circular argument. Indeed, a_1 is found in optimisation of:

$$\min(\varepsilon^{obs}_{i+1} - \varepsilon^{sim}_{i+1}) \Rightarrow a_1 \quad (1)$$

where $\varepsilon_{i+1}^{sim} = f(a_1, A_i)$. Therefore a_1 is a function of observed total deformation and previous concentration and is not a function of a_1.

We must use simulations of the same model for tuning a_1 for two reasons: First, reliable simultaneous observations of ice concentration and deformation at scales of 1d / 10km are not available. Second, weakening of ice (by decreased A or increased d) must be consistent with the rheological model parametrisation. For example, if observations (in case it existed) would show a higher rate of concentration decrease per deformation rate than we obtain from simulations, then the assimilation would decrease the concentration too much, and the model would predict higher deformation than was actually observed.

Corresponding explanations are added in Section 3.4.

Fig 5a: The skill of the model is assessed using the fraction of points where the correlation between observed and simulated deformation is significant. Statistically speaking, there will always be some points that will remain significantly correlated. The statistical significance of the signal must be shown in the figure. I also see no spatial structure in the regions of high correlations which suggest that the high-correlation points are just random events.

It is indeed difficult to illustrate the evaluation method on an example with gaps in observations. Below (Fig. 1) is an example of maximum cross correlation (MCC) maps computed for a free run and a forecast with assimilation of synthetic observations. Only statistically significant (p=0.01) values of MCC are shown. There is an obvious spatial pattern in the MCC values linked with co-alignment of LKFs on the deformation maps. This example also shows that MCC increases in some regions (e.g. in Beaufort Sea) after assimilation where co-alignment is visually better. And in some regions it remains the same (high or low), which indicates that MCC can be used for evaluation of assimilation impact. Examples with masking of insignificant MCC and a more contrast colormap are added in the manuscript.

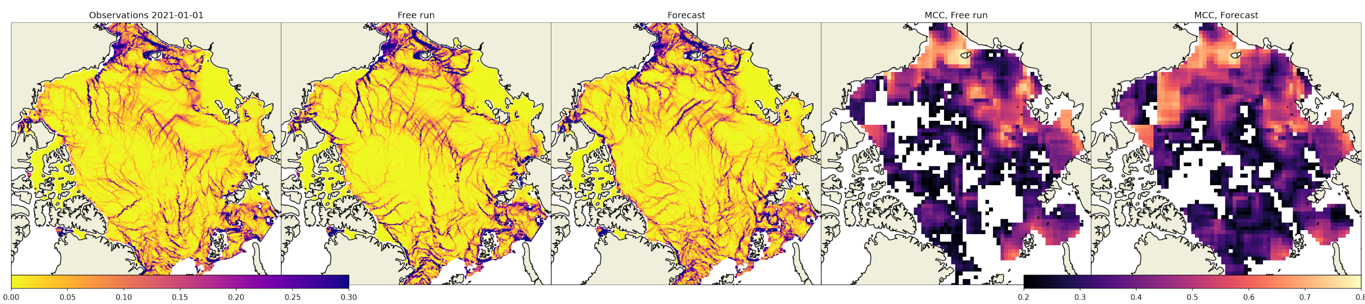

Figure 1: Impact of assimilation of synthetic observations.

Fig 5b: I would have expected that the root mean square difference of the forecast run would asymptote to the free run. The fact that it does not is suspicious. This must be explained.

As we are more interested in improvement of LKF localisation, and the MCC method suits better these needs, and many more experiments (free runs, assimilation of synthetic and real observations) have to be evaluated; it was decided to exclude the D_{P90} metric from evaluation of the predictability. In the sensitivity experiments the D_{P90} metric was replaced by the Kolmogorov-Smirnov test applied to PDFs of forecasted and observed deformation. Corresponding explanations are added to the Methodology section.

The discussion of the sensitivity of the forecast on epsilon_min and the weighting factor w_d appears in two one-line paragraph. I suggest removing them, or a more in-depth discussion should be provided.

The section on sensitivity to assimilation parameters is extended as new results obtained from the assimilation of synthetic observations are achieved.

Figure 6. The constant a1 is negative. This means that the sea ice concentration is larger than 1 (see Equ 10 and since eps_tot > 0). This is not physical. The results in the figure cannot be correct.

The sign of a1 on Figure 6 was incorrect and is corrected in the revised manuscript.

I am assuming that the sign of a1 is incorrect. If so, A = F_A(eps_tot) = 1 – a1 * eps_tot =0.76 for a1 = -1.2 (Line 232) and eps_tot = 0.2 (Fig 3). For A=0.76, P* is scaled down two orders of magnitude and the ice has no strength. Any rheological model where the ice strength is set to nearly zero along a

line (an observed LKF in this case) will deform along that line – this can be tested simply. The author instead argue that the brittle rheology is key to the correct simulated location of the LKFs. This is another unsubstantiated statement.

The statement on relation to rhelology is removed.

Results: The simulated concentration and thickness fields after assimilation should be presented. Reading from the deformation fields and the a1 constant derived from sensitivity experiments, we should see concentration of 0.8 along LKFs, something that is not accord with RGPS observations at 10km scale resolution. I suspect assimilating damage would help producing more realistic fields. The authors give reasons for why damage assimilation is not successful, but those are not convincing. See below for more details on this topic.

The fields of concentration and damage after assimilation are added.

The RGPS observations don't provide sea ice concentration estimates, and the current PMW observations cannot be used for accurate estimation of sea ice concentration variations at 10 km. Nevertheless, we agree that such reduction in total ice concentration is not realistic at these scales. As noted above there are two ice categories in neXtSIM: old ice and young ice, with the former being used in rheology and with the sum of both being used in thermodynamics. Reducing the older-ice concentration for increasing of ice weakness is compensated by increasing younger-ice concentration for keeping the thermodynamic balance. Thus, older-ice concentration plays a role of sub-grid parametrisation of ice weakness that is preserved on longer time scales. As opposed to sea ice damage that is working on shorter time scales. Please also see the answer below.

It is argued that the damage does not increase predictive skill of LKFs because ice heals too rapidly (24 hours). In the real world, ice heals through thermodynamic processes on much longer time scale. This choice of short healing time scale must be justified. Perhaps this is the cause of the lack of predictive associated with the assimilation of the damage.

According to the current neXtSIM parametrisation the damage increases due to healing from 0 to 1 in 15 days. This parameter was found as optimal in experiments (Rampal et al., 2019).

In the current study, in the new experiments requested by reviewers we found that, in fact, it is not the fast damage healing that reduces the efficiency of damage assimilation. We hypothesise that concentration and damage act on different time scales. This hypothesis was tested in an idealised twin-experiment: an intact initial ice field with damage=0 and concentration=1 everywhere was 'broken-up' along realistic LKFs. In one experiment, the elements in the LKFs were initiated by reducing concentration to 0.65 and in another one - by increasing damage to 1. Variation of damage and concentration in several thousand elements of broken-up and intact ice was studied (see Fig. 2).

The study shows that in case when LKFs are initiated by reduced concentration the situation is quite simple: concentration of ice in the unbroken elements is stably high, and in the broken elements it is first low and then stably increasing due to freezing (and convergence).

For damage the situation is quite different: in the initially unbroken elements the average damage remains relatively low (0.7 - 0.85), but damage variations are very large with standard deviation reaching 0.2. This happens because in some unbroken elements, that surround the initiated cracks, the internal stress exits the Mohr-Coulomb envelope and damage increases up to 1 at very short time scales (few time steps as discussed on Fig. 7 in the manuscript). Further, a cascade of damage events occurs in the neighbours of these newly broken elements. Probability of a break up (damage increase) is higher in directions of high internal stress. Thus, the information about the initiated damage is almost instantly forgotten - it is masked by many newly damaged elements.

Large scale observations of deformation at hourly frequency could probably confirm or reject the hypothesis of how damage propagates in reality, and illustrate whether or not assimilation of damage indeed leads to a more accurate deformation field **on small time scales**. However, we assimilate and validate against daily observations that show only long term memory in ice weakness expressed in reduced ice concentration.

The details on sea ice categories, the discussion of time scales at which concentration and damage act with the accompanying figures, and the maps showing impact of assimilation on damage and older-ice concentration are added to the manuscript.

The error is shown as 1/4 sigma. This is highly unusual. Typically, one would show an error envelope equal to one sigma (four times larger than what is shown in Figure 5).

The figure is re-plotted to show errors of several forecasts (see Fig **??**).

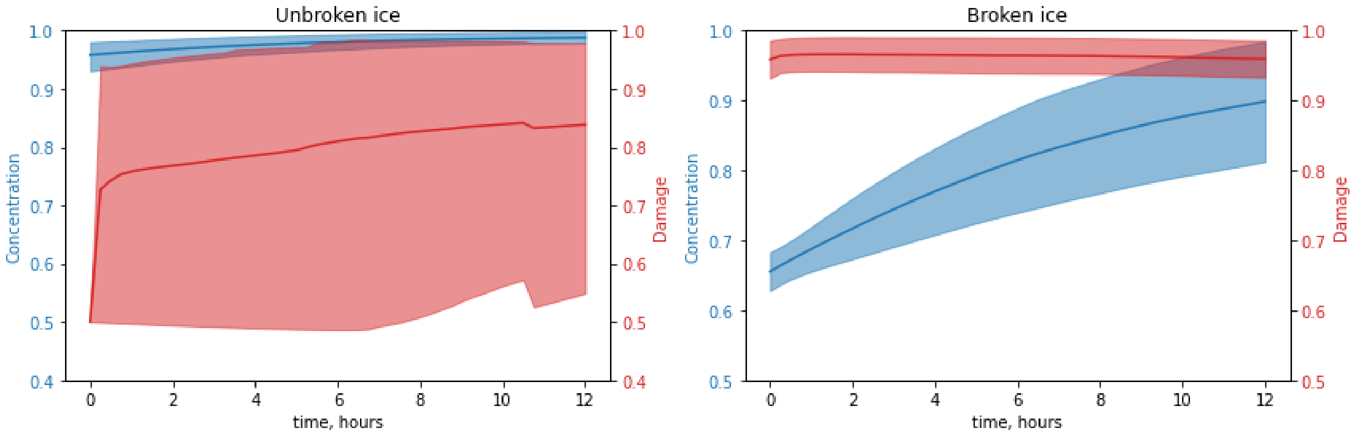

Figure 2: Mean and standard deviation of damage (red) and concentration (blue) in intact (left) and broken-up (right) elements.

Minor Points:

Line 23: "Under external forcing the ice deforms primarily as an elastic material." Most deformations in the pack ice are plastic and occurs along LKFs. This sentence also contradicts the next sentence.

The sentences are rewritten:

"As a consequence, sea ice does not drift freely anymore, but instead exhibits an intermittent and localised drift resulting from complex dynamics driven by a brittle mechanical behaviour, which can be described the following way. First, under increasing external forcing the undamaged ice deforms primarily as an elastic material. Internal stresses gradually accumulate in the material until a failure criterion is reached, which corresponds to a limit when sea ice fractures, and then the ice starts deforming along the multiple narrow and elongated cracks, and does so until these later refreeze or when the load (winds and currents) on the ice changes."

Line 26: "...start deforming along multiple narrow and elongated cracks formed and does so until these later refreeze". Or when the load (winds) on the ice changes. This should be added.

Please see the previous reply.

Line 48: :"... the exact timing and position of strong deformation zones, or LKFs, is not yet predicted precisely". The exact position of LKFs will never be located precisely because it depends on unresolved weaknesses within the ice pack. What we can hope to reproduce is the timing, the orientation of the LKFs with respect to the large-scale forcing and their statistical distribution (width, length, density, angle of fracture, etc.). This should be corrected. This is another sentence that suggests incorrectly that BBM could eventually simulate LKFs position correctly.

When we stated the fact that errors still exist in small-scale features in our study, we did not intend to comment on BBM's ability to eventually reproduce these features correctly. The end of the sentence is rewritten as follows:

"...the exact timing and spatial distribution (including orientation, width, length and angle of fracture) of strong deformation zones, or LKFs, is not yet predicted precisely."

Line 93: "The observed variables vo (damage and concentration) is computed...". "Damage" is not observed. The word "is" should read "are".

The sentence is rewritten as follows:

"The 'observed' model variables damage d_o and concentration A_o are derived from the observed deformation ε_o using the following experimental formulations:"

Line 94: Sometimes, the Greek lowercase epsilon symbol is used and sometimes the lunate epsilon symbols is used. The author needs to choose only one form for consistency. See Line 94 and Equations 6-7 for examples. This needs to be corrected everywhere.

The Greek lowercase symbol (ε) is used everywhere.

Line 102: It is said that CNEMS has a temporal resolution of 12 hours on Line 102; and it is said in the next sentence that it is "observed nearly every day". This sounds contradictory. This should be clarified. See Line 115 as well where 24 hours is specified.

The sentences in Section 3.1 are rewritten as follows:

"The dataset comprises gridded products derived from Sentinel-1 synthetic aperture radar (SAR) images, with 10 km spatial resolution. Ice drift is computed from pairs of images separated by approximately 24 hours and the product is delivered every 12 hours. The spatial coverage of the product is irregular - the East Siberian, Laptev and Kara sea and the polar gap (north of 87^oN) are never covered, while other Arctic regions are observed at least once nearly every 24 hours."

Line 105: "The model is forced with the European Center for Medium-Range Weather Forecasts (ECMWF)". The version of ERA should be specified: e.g. ERA5, ERA- interim, etc.

Version is added:

"The model is forced with the latest version (Cycle 45r1) of the Integrated Forecast System European Center for Medium-Range Weather Forecasts (ECMWF) (Owens and Hewson, 2018). "

Line 117: "total" should read "total deformation".

Corrected.

Line 203: The units of eps_tot must be given the first time it is introduced (Table 1, Line 197). At present it is only given on Line 203.

The units are specified in Section 2.

Line 235: eps_tot and divergence is used interchangeably; yet they are very different. One includes shear the other not.

The term "divergence" is used only three times: on lines 117, 144 and 233 and every time a correct relation to total deformation is given.