

Reviewer 1.

The manuscript has changed substantially from the first submission, partly because the authors followed most of the reviewers' suggestions, making this nearly an entirely new manuscript. Thank you for taking my comments into account. All major problems raised in the first review have been addressed in an adequate manner.

The (new) manuscript describes data assimilation experiments with a sea ice model (neXtSIM) and deformation data in a realistic pan-Arctic configuration. The data assimilation scheme is simple, albeit effective, and the experiments clearly address potential predictability and predictive skill. The main "trick" of the assimilation scheme is the design of inverse observation operators that map the deformation data (derived from drift data) to the model variables sea ice concentration and damage (damage is a parameter unique to brittle models such as neXtSIM). The construction of these operators is described in the appendix (which was formerly not visible). My impression is that the most important result is stated in line 296: "In other words, assimilation of damage has little impact on forecast error, whereas assimilation of concentration plays a big role." The remaining results are also interesting (e.g., predictive skill of 3-4 days), but since the neXtSIM code is not publicly available, it is hard to make out how these results can be generalised and are relevant to the community except in a competitive way. The construction of the inverse observation operators depends on the neXtSIM simulations with a lot of tuning of parameters and this cannot be reproduced without access to the neXtSIM simulations (or code). The manuscript ends with a list of potential confounders and lays out ideas for future research addressing these confounders.

We are grateful for the positive evaluation of our revision! In the second revision we ran new and dedicated neXtSIM experiments with the mEVP rheology in order to illustrate that the proof-of-concept can be generalized for application in both brittle and viscous-plastic rheological frameworks. Moreover, the BBM is being implemented in open-source sea ice models such as neXtSIM-DG, SI³, etc. This makes the study relevant to the wide community of sea ice modelers who use either brittle, or viscous-plastic rheological frameworks.

Data and methods are described in a clear manner, the experimental design is sound and the result presentation is mostly clear (see smaller comments below). The language is clear (but could be improved, missing article, prepositions, etc, nothing that careful editing could not fix). The authors have reduced the market place tone of the manuscript, although, to my taste, this could be done even more to not distract from the science.

We have proofread the manuscript to improve grammar and we have further toned down the sentences that sounded like 'marketing', and focused more on the scientific discussion of our results.

The author do not make the model code of neXtSIM available. I am not sure how that complies with TC rules.

Copernicus Publications encourages authors to deposit software, algorithms, model code, etc, but this is not a strict requirement.

Main comment

In my understanding of the MS, the main result is that "assimilation of damage has little impact on forecast error, whereas assimilation of concentration plays a big role." (L296), because this statement gives other researchers a clear idea, which direction to follow when attempting to assimilate deformation data.

We agree that L296 is an important part of the findings in our paper, that it guides future efforts in assimilating deformation to improve sea ice LKF forecasts. However, the main result we want to show is the proof-of-concept that assimilation of observed deformation can improve LKF forecast,

which hasn't been demonstrated in the literature yet. The new experiments with the mEVP rheology reveal for the first time that the suggested DA-method is applicable for sea ice models based both on brittle and viscous-plastic rheologies, which makes the revised manuscript relevant to a wide community.

Connected to this, however, is the construction of the (inverse) observation operator, which could not be evaluated in the previous submission, as it is described in the appendix and the appendix was not part of the submission. I find the observation operators so elementary to this manuscript that I recommend moving their construction from the appendix to the main text.

We agree that the observational operators are an important part of our work and that they are not very complicated. However, we believe that the description of building the operators is quite technical, less related to the main scientific findings of the manuscript and is kept in the appendix.

From the appendix, in particular Figs 12 and 15, it becomes clear that the inverse operators are incomplete (and the authors describe this). Naively I would expect that if the observation operation H_A and its inverse H_A' were perfect, panels B and C would be the same, but the reconstructed fields (panels C) are basically rescaled deformation fields (by construction). In the damage case (Fig 12), this leads to a large underestimation of damage where deformation is small, and in the case of concentration to an overestimation of concentration minima (in LKFs) compared to the simulation data. In combination with the threshold eps_min in the weighting scheme (eq 24), this means that only the large (approximately correct) damage values and the low (lower than simulated) concentration values are used in the assimilation.

In the manuscript we made the assumption that the fraction of older ice (responsible for ice weakening) decreases with increase of total deformation. A physically intuitive relationship would be that the total sea ice concentration linearly decreases with divergence (neglecting refreezing of opened leads, etc). However, in our work we don't use this relation between these variables. In our method the inverse operator H' is built on the assumption that the total deformation is related to the fraction of older ice (A), since that is the fraction that impacts the ice strength. Some error is allowed in H' to depict this relationship, while our goal in assimilation is to minimise the error in forecasted deformation. Therefore, H' is deducted not from observations, but from minimisation of the forecast error. Please also see revision suggested in the reply below.

Further, if one believes in the model dynamics (that have been used to generate the pairs in panels A and B to construct the observation operator), this method overly emphasises the low concentration in leads; assimilating these reconstructed fields will then lead to strong weaknesses and immediately allows more deformation (because ice strength/elasticity depends on concentration exponentially). This gives an (alternative) algorithmic explanation for the result. I think that this aspect requires a discussion in the text as it relates directly to the main result (to my mind) that the concentration proxy is more effective.

One should also include a discussion how realistic this effect is and if one could achieve this also in a different way. After all, there is concentration data available each day at a good resolution and I can imagine that with a similar "boost" or "enhancement" of the LKF structure in these data (i.e. a direct map between panel B and C in Fig 15) they could already lead to a similar effect.

In the observations the total ice concentration is reduced due to divergence. However, as seen on Figure 15.A, the divergence is observed only in ~ 50% of all LKFs, whereas other LKFs are formed by convergence and/or shear. Since during the assimilation we decrease not the total sea ice concentration but the fraction of older ice (as one of the ice strength components), and since we want to reflect a reduction of ice strength in all LKFs, we instead use the total deformation as the predictor. Our experiments show that accounting for concentration reduction due to divergence alone decreases the effect of assimilation by 30 ~ 40 %.

The following clarifications were added to the manuscript to address these comments:

“Since reliable simultaneous observations of concentration and deformation at scales of 1 day / 10 km are not available, and damage is not an observable variable, we cannot use an empirical inverse operator. I.e., $H' \neq H'_o$, where H' is the inverse operator in question, and H'_o is the empirical inverse operator between the observed total sea ice concentration (x^o) and sea ice deformation (y^o): $x^o = H'_o(y^o)$. Ultimately, the purpose of the inverse operator is not to describe a physically realistic process (i.e., linear decrease of concentration due to divergence) but to minimize the error of the deformation forecast: $E_n^F = y_n^o - y_n = y_n^o - H \circ M \circ A \circ H'(y_{n-1}^o)$, where H is the forward operator to compute deformation from ice drift, M is the numerical model to forecast ice drift, A is the assimilation of the inverse operator applied to the observed total deformation at the previous time step.”

On lines 190 – 195 it was emphasized that the relationship between e_{tot} and A_o is a reasonable assumption for improving the LKF forecasts:

“Since A and d are the components of sea ice strength, and the total deformation is a good proxy for the presence of weak ice, we suggest building the inverse operator H' on the assumption that A and d are related to the total deformation.”

Minor comments, suggestions, typos:

page 5

eq(9), not clear if the noise is added to the initial conditions ore actually to the model operator (i.e. to the solutions after integration).

The noise is added to the model operator. The sentence is rewritten as follows:

where ψ_t is a random noise added to the model operator due to uncertainties in model numerics that cause the forecasts run to differ from the truth run.

1141: “The first forecast is initiated from t_0 ”

I guess I do not understand the difference to “forecasts initiated from truth”, also there’s some confusion about t_0 and t_1 , when does this (set of) run(s) actually start?

As described on lines 120 – 131 the “truth run” is initialized on t_0 (1 December 2020), the model is spun up until 1 January 2021, and this data is not used. Then the “truth run” continues until 31 January 2021. The ‘Forecasts initiated from truth’ are initiated from the output of the “truth run” on January 1st, January 2nd, January 3rd, etc.

Notation t_i is replaced with t_i and the following clarifications are added to eq. 14:

$$\mathbf{x}_{t_i \rightarrow t}^T = M_{t_i \rightarrow t}(\mathbf{x}_{t_i}^{tr}) + \psi_t$$

where i denotes days in January 2021 (e.g., January 1st, January 2nd, January 3rd, etc), ...

page 6

Eq16: not consistent with other eq, e.g. eq14

The equation is replaced with a set:

$$\mathbf{x}_{t_i}^{ar} = A(\mathbf{x}_{t_i}, \mathbf{y}_{t_i}^o; H', \mathbf{w})$$

$$\mathbf{x}_{t_i \rightarrow t}^{ar} = M_{t_i \rightarrow t}(\mathbf{x}_{t_i}^{ar}) + \psi_t$$

L161: “ $t_n + 2$ ” latex problem. t_{n+2}

Corrected.

Eq17 and L162, confusion about $\epsilon^A_{\Delta t}$ and $\epsilon^A_{t_n \rightarrow t_{n+1}}$. I guess they are the same?

As specified on line 138 Δt is forecast lead time. For $t_n \rightarrow t_{n+1}$ lead time would be equal 1. The following clarifications are added:

Then the forecast is compared with deformation computed at $t_n \rightarrow t_{n+1}$ (corresponding to lead time $\Delta t = 1$), at $t_n \rightarrow t_{n+2}$ ($\Delta t = 1$), etc.

page 8

Eq22 needs some reformatting (“*” should probably not appear).

Corrected.

page 9

L224: “setting the weight as 0 or 1 for damage and concentration”

Table 1 implies that values of 0, 0.5, and 1 are used.

Rewritten as follows:

The variable-dependency is tested by setting the weight w_d or w_a to 0, i.e., letting assimilation to update either damage, or concentration, or both to see the impact of the update.

Table1: units for ϵ_{\min} and a_1 ?

The units for ϵ_{\min} (d^{-1}) and a_1 (% d) are added to the table.

L229: “Over 30 experiments” how were these chosen? According to table1 there are $7 \times 3 \times 3 \times 3 = 189$ parameter combinations.

That was a typo in the manuscript: 64 experiments were run. All effective combinations of the parameters were tested. Some ineffective combination of parameters (e.g., $w_c = 0$ and $w_d = 0$, or $w_c = 0$ and any value of a_1) were not tested. The tested values of parameters are also updated in the table.

page 10

L259 “that should not shock the model.”

Unclear, what this is supposed to mean, please rephrase. “Shocking the model” was not a problem introduced so far.

The sentence is removed.

page 11

Figure 2. It is not quite clear what we see here. The observations (derived from drift) in panel A,B,E area compared to simulation (I am guessing without DA) in C and F, and D and G are the analysis without further integration? I.e. directly after applying eq (22) and giving the model no time to adjust to these new initial conditions?

If that’s the case, this should be better described in the text.

Yes, that’s the case. The caption is rewritten:

Figure 2. A: sea ice deformation (d^{-1}) computed from observed sea ice drift between 15 and 16 January 2022. B and E: sea ice concentration and damage reconstructed from the observed deformation using Eqs. 20 and 21. C and F: concentration and damaged simulated by neXtSIM on 16 January. D and G are the analysis (Eq. 22) without further integration.

In the “analysis” of SIC (panel D), it is quite obvious that the observations replace the model where available, but otherwise there’s not impact. The effects on damage (G) are probably also small (and hardly visible) because the reconstructed damage is so “featureless” (most values between 0.9 and 1)

Yes, low values of deformation, or absence of observations (i.e., due to gaps in satellite data) do not impact the analysis.

l270: “Visually the position of cracks correspond well to the observations,” that is not so obvious to me.

On Jan16, the Forecast with DA looks more similar to the Forecast without DA than to the observations which have been just inserted, maybe near the Fram Strait the DA has an effect. Most strikingly the region of large deformation in the obs from the centre of the Arctic towards Alaska/N.Canada does not show up at all.

In fact, these observations were not used in assimilation. As described on lines 160 – 163, we use the observations from $t_{n-1} \rightarrow t_n$ for assimilation and from $t_n \rightarrow t_{n+1}$ for verification. It is therefore not surprising that the forecast looks similar to the Forecast without DA. Nevertheless, the differences in the forecasts are clearly visible and can be attributed to the assimilation. The sentence is rewritten: Visually the position of some cracks corresponds well to the observations, ...

l276: “On the third day the improvement introduced by DA is obvious only in the central Arctic, on intersection of two large cracks crossing the entire ocean.”

That appears to be completely fortuitous as the two large LKFs that from a semicircle are already visible in both forecasts (with and without DA) on the day before they appear in the observations (which do not have any direct impact anymore).

It is impossible to say if this particular improvement is completely fortuitous or not. The sentence is rewritten:

On the third day the improvement which is likely introduced by DA is obvious only in the central Arctic, on the intersection of two large cracks crossing the entire ocean.

page 12

l283: “but also outside it.” According to Fig4 and 5 the impact is much smaller outside of the data coverage (the color range is different from Fig 3, column 5). This should be mentioned in the text.

The sentence is rewritten:

Visual comparison of forecasts and observations, as well as the maps of MCC increase show that the correlation has improved not only in the area covered by the assimilated observations but also outside it, although, to a lesser degree.

page 14

l293 It does not become clear (to me) from the description how the parameters are varied. Since this is a multivariate system (4 parameters), the sensitivity of one parameter may depend on the values of the other 3. How is that handled?

All effective combinations of the parameters were tested. The spread of the metric for a given parameter value indicates dependence of the metric on other parameters. We have analyzed the scatterplots of each parameter against other parameters and the metric; however, it is impossible to show all these results in the manuscript. We, therefore, limit the presentation to the mean value of the metric (Fig 7) and scatterplot of KS vs A_{MCC} metric (Fig 8), which allows us to draw a conclusion on the overall variability of the metrics and the best combinations of parameters.

Fig7+8: I don't see how the presented data allows the conclusion that " $w_c=1$ provides the best restyle when a_1 is 0.9 or 1.2 (l312). For that you'd need a (scatter) plot with w_c and a_1 on the x and y axes, wouldn't you?

A circle is added to Fig. 8 in the revised manuscript to show two groups of points with the lowest KS and lowest $1-AMCC$ that correspond to the values of parameters specified on lines 314, 315.

page 15

Figure 7. Drawing lines through 2 or 3 points is a bit misleading, bar plots or points with error bars would be more "honest", similar in Fig8 using a colorscale/range, even a discrete one, implies that more than just 2 or 3 values have been tested. Here I would use a legend with color labels.

Barplots are used instead of lines on Fig. 7. A note is added to the caption of Fig. 8 indicating that the discrete colorbar denotes a limited number of values.

l314: "Based on these observations the following values were chosen as the recommended ones" I don't think that this "recommendation" is helpful for anyone but the authors. Given that we do not have access to the code of the model, no-one will be able to follow this "recommendation" for the sub-optimal DA procedure. I would rephrase this as "Based on these results, the best parameter choice for our configuration appears to be:"

The sentence is rephrased accordingly.

page 16

l318: "successful" I find this word in this context inappropriate. It's not clear what is "success" in this context. It sounds (again) like (over-)selling a product.

The sentence is rephrased as follows:

We present a proof-of-concept of assimilation of satellite-derived sea ice deformation which increases the accuracy of deformation prediction for the first 2 – 3 days.

l319: remove "relatively"

The term 'simple' cannot be used in the absolute manner. Our method is simpler than, e.g., EnKF, but is more complex than, e.g., direct insertion of observed concentration, therefore 'relatively simple' seems more appropriate here.

L322: "for the sake of confirming several hypotheses"

Not clear what these hypotheses are, and in the language of hypotheses testing this is impossible, you can only test and reject hypotheses.

The sentence is removed.

page 17

l356: exits -> exceeds

Corrected.

L357 "at very short time scales" -> "on very ..."

Corrected.

l358: neighbours -> neighbourhood, neighbouring elements?

Replaced with "neighbor elements."

page 19

L386: “2–5 days” in Fig6 is it 3-4 days. Where do these numbers come from?

Corrected to “3 – 4”.

L388 applications (missing “s”). What are “real world applications”? They would still use models, right? I would write: “more generally”

Corrected accordingly.

page 20

L433 now it is “2 – 3 days”, why not stick with one set of numbers (3-4 days according to Fig6)

Corrected accordingly.

L433: remove “relatively”

The term ‘simple’ cannot be used in the absolute manner. Our method is simpler than, e.g., EnKF, but is more complex than, e.g., direct insertion of observed concentration, therefore ‘relatively simple’ seems more appropriate here.

L433: The entire sentence is weird. I would rewrite this as:

The simple data insertion approach may lead to inconsistent initial conditions because of spatial gaps in the data. Still the dynamics of the neXtSIM extrapolates the spatially discontinuous satellite observations of deformation ...

The sentence is rewritten as follows:

The relatively simple data insertion approach may lead to inconsistent initial conditions because of spatial gaps in the data. The spatially discontinuous satellite observations of deformation are extrapolated by the model, connecting the elements of linear kinematic features in a realistic manner.

page 24

L545 “(in $\log_{10}(1-d)$ space)” figure11 axes labels say “ $\lg(1-d)$ ”, please fix.

Corrected accordingly.

Reviewer 2

This is a second review of the paper entitled “Towards improving short-term sea ice predictability using deformation observations”. Results from the paper show that data assimilation of (reconstructed) concentration and damage derived from observed total deformation improved LKFs forecast up to two days and that the model is able to fill the gaps in the Sentinel coverage (used to calculate the total deformation) with LKFs that are qualitatively reasonable. The authors implicitly assume that the ability of the model to fill the Sentinel gaps stem from the BBM rheology by suggesting that an mEVP rheology would not be able to achieve this.

The paper is much improved, and the authors have addressed satisfactorily most of the comments from the reviewers. There is still, however, some unsubstantiated statements, misleading statements and a circular argument in the approach used in the assimilation procedure in the revised manuscript. The paper attempts two things a proof of concept that (A) assimilation can improve LKFs forecast and (B) that the BBM rheology is responsible for this improved forecast including being able to fill in the gap in Sentinel coverage, but neither is done convincingly. Both goals would be welcome contributions to the community, but in its present state the manuscript is not satisfactory. The authors should choose one option and present it in a convincing manner (see points below for details). I recommend that the paper be accepted for publications after the comments below are fully addressed (i.e., not rebutted).

We are grateful to the reviewer for the valuable comments. We have introduced corresponding revisions that are relevant for the manuscript while staying within the scope of the study.

The reviewer found that the two major findings in this paper are not yet convincing. The comment on point (A) is very general and the reviewer hasn't provided further details or suggestions regarding this criticism. We believe that the experimental results presented in this paper constitute a good proof-of-concept that assimilation of deformation improves LKF forecasts. Although there is still room for improvement for both the model and the assimilation method we hope the reviewer will agree that the results support point A (in the literature, there are plenty of OSSE works based on models or DA-schemes that are imperfect, e.g., Fritzner et al., 2019; Zhang et al., 2018, etc.).

For point (B), we agree that attribution of the forecast improvements solely to the choice of model rheology is not supported by the experimental evidence. To address this, in the revised paper we ran new dedicated neXtSIM experiments with the mEVP rheology as suggested by the reviewer. We found that forecast improvements occur for both rheologies, which generalises our findings about the assimilation impact in point (A) for both brittle and viscous-plastic rheological frameworks. We also found differences in forecast performance when comparing the two rheologies that are described in the revised manuscript.

Major comment:

General: Irrespective of the goal of the paper (Option A or B), the following points must be addressed prior to publications.

1. The study is simple. The ice strength is set to nearly zero along observed lines of high deformations (Linear Kinematic Features, LKFs), and the model in turn produces large deformation along the observed LKFs where strength is set to ~zero. Not surprisingly, the model is also able to link observed LKFs when there are observational gaps along a given LKFs (this was shown in idealized experiment by Ringeisen et al 2019). The ice strength in the BBM model has an exponential dependence on ice concentration (or open water fraction) and a linear dependence on damage. Given that total deformation cannot be assimilated in a simple manner – because it is not a prognostic variable – transfer functions (operator) are developed between simulated total deformation and simulated concentration and damage. These transfer functions are then used with observed total deformation in order to derive

reconstructed damage and concentration that are used in the assimilation. This approach implicitly assumes that the relationship between deformation, sea-ice concentration and damage in the model is correct. This assumed equivalence between simulated and observed total deformation is problematic.

In the manuscript we made the assumption that the fraction of older ice (responsible for ice weakening) decreases with increase of total deformation. We are by no means implying this is the ‘correct’ relationship (i.e., physically realistic; physically intuitive). Apparently, a realistic relationship would be that the total sea ice concentration linearly decreases with divergence (and increases with convergence), if we neglect refreezing of opened leads, rafting of thinner ice, and other processes. However, in our work we don’t use this relation between these variables. In other words, $H' \neq H'_o$, where H' is the inverse operator in question, and H'_o is the empirical inverse operator between the observed total sea ice concentration (x^o) and sea ice divergence (y^o): $x^o = H'_o(y^o)$.

In our method the inverse operator H' is built on the assumption that the total deformation is related to the fraction of older ice (A_o), since that is the fraction that impacts the ice strength. Some error is allowed in H' to depict this relationship, while our goal in assimilation is to minimise the error in forecasted deformation. Therefore, H' is deducted not from observations, but from minimisation of the forecast error E_n^F :

$$H' = \underset{H'}{\operatorname{argmin}}(E_n^F)$$

$$E_n^F = y_n^o - y_n = y_n^o - H \circ M \circ A \circ H'(y_{n-1}^o),$$

where y_n is forecasted total deformation at step $t=n$, y_n^o is observed total deformation, H is the forward operator to compute deformation from ice drift, M is the numerical model to forecast ice drift and A is the assimilation of the inverse operator H' applied to the observed total deformation y_{n-1}^o at the previous time step. As seen from the above formulae, there are no circular arguments.

In Olason et al. (2022) we showed that both the magnitude, the spatial pattern, the temporal evolution, and the spatial scaling of the simulated deformation fields have quite high accuracy when compared to the RGPS observations. This provided the basis for assimilation, since the model simulation can reproduce a lot of the important features in the observation. Although the simulated relationship is not perfect, it is accurate to a first order to allow the assimilation approach to improve forecasts.

The following clarifications were added to the manuscript to address the reviewer’s comment:

Line 88: “It should be noted that there are two ice categories in the model: young ice, which is formed during water freezing, and older ice which is formed after young ice exceeds a threshold in thickness (Rampal et al., 2019). Only the older ice concentration (referred to as A) is used in the rheological equations.”

Line 103: “The “observed” model variables damage d_o and concentration of older ice A_o

Lines 173 – 179 were rewritten:

“Since reliable simultaneous observations of concentration and deformation at scales of 1 day / 10 km are not available, and damage is not an observable variable, we cannot use an empirical inverse operator. I.e., $H' \neq H'_o$, where H' is the inverse operator in question, and H'_o is the empirical inverse operator between the observed total sea ice concentration (x^o) and sea ice deformation (y^o): $x^o = H'_o(y^o)$. Ultimately, the purpose of the inverse operator is not to describe a physically realistic process (i.e., linear decrease of concentration due to divergence) but to minimize the error of the deformation forecast:

$E_n^F = y_n^o - y_n = y_n^o - H \circ M \circ A \circ H'(y_{n-1}^o)$, where H is the forward operator to compute deformation from ice drift, M is the numerical model to forecast ice drift, A is the assimilation of the inverse operator applied to the observed total deformation at the previous time step.”

On lines 190 – 195 it was emphasized that the relationship between e_{tot} and A is a reasonable assumption for improving the LKF forecasts:

“Since A and d are the components of sea ice strength, and the total deformation is a good proxy for the presence of weak ice, we suggest building the inverse operator H' on the assumption that A and d are related to the total deformation.”

I suggest instead:

- a. Damage cannot be observed and therefore, a transfer function must be developed from model output, as done by the authors.
- b. Concentration however can be calculated directly from Sentinel-derived divergence (eps_I) contrary to what is done in the paper – see for instance work by Kwok who has produced a dataset of thin ice based on RGPS divergence as an example of how this is done. The reconstructed concentrations from observed divergence should be assimilated in the model, instead of the reconstructed concentrations obtained with the model-trained operator.

The approach of Kwok et al. for computing the total sea ice concentration is based on the aforementioned ‘correct’ linear relation between divergence (divergence rate multiplied by forecast time) and decrease of the total sea ice concentration. However, as seen on Figure 15.A, the divergence is observed only in ~ 50% of all LKFs, whereas other LKFs are formed by convergence and/or shear. Since we decrease not the total sea ice concentration but the fraction of older ice (as one of the ice strength components), and since we want to reflect a reduction of ice strength in all LKFs, we instead use the total deformation as the predictor. Our experiments show that accounting for concentration reduction due to divergence alone decreases the effect of assimilation by 30 ~ 40 %.

2. Another choice made by the authors is to associate total deformation with a reduction in sea ice concentration irrespective of the sign of the divergence – which is clearly not realistic. Radarsat-derived divergence PDFs are nearly symmetrical around zero with an equal number of scenes with positive and negative divergence. Presumably, the authors make this assumption because the assimilation of damage (present in both divergence and convergence) does not improve the predictability of LKFs. Instead, a lengthy and unclear discussion related to timescale is included to explain why damage does not improve predictability. This may suggest instead that the functional dependence of ice strength on damage is not appropriate. See item #3 below for more on this issue.

In our method the inverse operator H' is built on the assumption that the total deformation is related to the fraction of older ice (A), since that is the fraction that impacts the ice strength. Our experiments

show that accounting for older ice concentration reduction due to divergence alone, decreases the effect of assimilation by 30 ~ 40 %.

With a scalar damage unaffected by other sea ice processes (unlike the concentration), the dependence of the ice strength on damage is completely arbitrary, since the damage and stress must evolve in a manner that is consistent with the constitutive relation (i.e. the effective strength after damaging should be the same regardless of whether we have a factor $1-d$ or $1-f(d)$, where f is any function that increases monotonically from 0 to 1 with d). What may be sub-optimal in our model is either the constitutive relation itself or the numerical implementation of the co-evolution of stress, damage and concentration, and we now say this in our revised discussion section. Having said this, the goal of the manuscript is to show that despite having an imperfect model and an imperfect DA-method we can still improve the LKF forecast.

3. The two stated reasons for why damage does not improve LKFs forecast (different timescale; linear vs exponential dependence) are not convincing. I would argue instead that this may be an indication that damage is not correctly parameterized in the EB/MEB/BBM model.
 - a. **Timescale:** The important timescale for damage to consider is not the timescale for damage growth but rather the timescale for damage healing. If damage grows and propagate in the correct direction – along the line of maximum stress (Line XX) – the benefit of that should be seen one day later.
 - i. The authors states that hourly Sentinel observations would allow to test this hypothesis, but they do not exist. A feasible way to test the hypothesis would be to slow down the propagation of damage in the model and see if assimilation improves LKFs forecasts.
 - b. **Exponential vs linear dependence:** The fact that assimilation of sea ice concentration has a positive impact on LKFs forecast and damage has not, suggest that perhaps ice strength should also have an exponential dependence on damage.

We agree with these comments and thank the reviewer for sharing his thoughts. Slowing down the damage propagation or changing the dependence of ice strength on damage are feasible ways forward that may potentially improve the effect of assimilation. However, testing of these approaches goes outside the scope of the manuscript, which aim is to provide a proof-of-concept for improving prediction of LKFs by assimilation of satellite-derived deformations using an existing state-of-the-art sea ice model and the method presented in this manuscript. We added to the list of possible future steps in the Discussion section.

4. Another result that suggests that P_{max} may not have the proper functional dependence on damage is the fact that the authors must resort to decreasing sea ice concentration (and therefore ice strength P_{max}) whether divergence or convergence is present in the observations. The reason for implementing damage in the first place in sea ice models (Girard et al., Rampal et al., Dansereau et al., etc) is to weaken the ice early when ice deforms without the need to rely on a decrease in sea ice concentration that operate on longer a timescale, i.e. for divergent or convergent flow. The fact that sea ice concentration must be reduced even when convergence is present suggest that damage does not do its job (see Line 368). The authors appear to be missing a good opportunity to suggest a different functional dependence of P_{max} on damage based on results from this assimilation procedure.

We thank the reviewer for this comment. We agree with the reviewer that the functional dependence of P_{max} on damage can potentially be improved and that our study allowed the potential impact of this improvement to be identified. However, we think that addressing this properly would require a significant and therefore dedicated study. We thus decided to add to the discussion section this interesting point as a potential for future studies.

- Figure 11: The colorbar must be defined. I suspect it shows the density of points for a given damage and eps_tot . More importantly, there is a dark yellow line (highest density of points) on the top right of the figure that shows a large number of undamaged ice ($\log(1-d) = 0$) for very large total deformation ($-1 < \log(\text{eps_tot}) < 0$). These points are contrary to the argument presented in the paper (see Equ 5). In panel (b), these points are omitted (the near zero $\log(1-d)$ and $\log(\text{eps_tot}) \approx 0$ are removed). This feature must be explained, not deleted from the panel.

We thank the reviewer for spotting this. There were indeed a few elements that were erroneously selected for training the damage-deformation relation. After fixing a bug in the selection code these elements disappeared without changing the results of the fitting. A colorbar was added to the figure.

- Line 563-564: “Note, that unlike Eq. 19, the optimisation is performed here in the space of observations and using the observed total deformation.” This must be clarified. The optimisation is done in the space of reconstructed diagnostics (concentration and damage) from observed total deformation (see Figure 14), not “in the space of observations”.

The sentence is rewritten: “... the optimization is performed here in the space of observed and

simulated total deformation: $\max_{H'}(A_{MCC}(\varepsilon_{tot}, \varepsilon_{tot}^o))$.

Option A: Proof of concept that assimilation can improve LKFs forecast

Remove all statements from the paper that imply that the simulated LKFs shown are the result of the BBM rheology: any sea-ice model (with both compressive and shear strength) where we set the ice strength to zero along a line will deform along that line.

Option B: BBM is responsible for improved LKF forecast and correctly filling the gaps in sentinel coverage.

If the authors want to pursue that route, the following should be addressed

Line 407: “We expect that the model equipped with the mEVP rheology will not be capable of spatial extrapolation of the assimilated ice weakness (lowered A or enhanced d), and that further tuning of the BBM rheology can improve the practical predictability of LKFs”. This is an unsubstantiated statement, i.e., the statement of a belief rather than a statement supported by facts. This does not have its place in a scientific paper. The authors mention on line 405-406 that the mEVP is an available option in neXtSIM: “The BBM rheology can be further tuned and compared to the modified Elasto-Visco-Plastic rheology (mEVP, Bouillon et al., 2013) that is already an available option in neXtSIM (Olason et al., 2022) for estimating the impact of rheology on the sea ice predictability.” I propose that the authors run the same experiment using mEVP in order to support this statement. My own “belief” is that they will find that this is not the case. My “belief” is based on results from Ringeisen et al. (2019) showing how a standard VP rheology joins random weakness in a sea ice slab together into single LKFs feature even when the weaknesses are not oriented along lines where the maximum internal stress are present.

As suggested by the reviewer we ran assimilation experiments with the mEVP rheology in place of the BBM, using the same numerical implementation and parameters’ values as in Olason et al. 2022. In the experiments we assimilate deformation only through reduction of sea ice concentration. Two experiments were run: one with the same values of parameters of the assimilation method as in the “best” experiment with BBM, and another experiment with a higher value of a_1 . Experiments show (see Figure 11 in the revised manuscript) that with the same assimilation setup the error of the EVP-based forecast for the first day is the same as for the BBM-based forecast. However, for larger lead

time the error grows faster and saturates at a higher level in the mEVP simulation (see Figure 12 in the revised manuscript).

In light of these new results, we thus think that our manuscript now could be **Option C**: it is a proof of concept of a generalized approach that observed deformation can be assimilated to improve forecast of sea ice deformation in both brittle and viscous-plastic sea ice models. Lines 405 - 409 are rewritten as follows:

“Tuning of the BBM rheology regarding the speed of damage propagation and changing the dependence of ice strength or P_{max} on damage could improve the practical predictability of LKFs. The mEVP rheology could also be further tuned by changing the number of sub-cycles or adding a damage parameter (as in Savard et al., 2023).”

A sub-section “Experiments with mEVP rheology” was also added to the Discussions section in order to describe these new findings.

Minor Points:

Line 6: “We show that high values of ice deformation can be interpreted as reduced ice concentration and increased ice damage - scalar variables of neXtSIM”. It can be interpreted/simulated this way but this is not the case in reality and this may be hiding a defect in the formulation of the damage parameterization. See major point above.

We agree that there is room for improvement of the damage parametrization. However, that does not prevent the assimilation of deformation through modifying either or both concentration and damage as the two components of sea ice strength. Given that reduction of concentration can be used in a more generalized approach in both EB and EVP rheologies the sentence is rewritten as follows:

“A new method for assimilation of satellite-derived sea ice deformation into numerical sea ice models is presented. ... We show that high values of ice deformation can be interpreted as reduced ice concentration or increased ice damage - scalar variables responsible for ice strength in brittle or viscous-plastic sea ice dynamical models. This method is tested as a proof-of-concept with the neXt-generation Sea Ice Model (neXtSIM), where the assimilation scheme uses a data insertion approach and forecasting with one member.”

Line 45-47: “In a recent model intercomparison paper (Bouchat et al., 2021) neXtSIM results 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 10 km.” The spatial resolution is lower, but the effective spatial resolution is higher given that a finite element model can resolve discontinuities at the model grid scale, contrary to a finite difference model where 7-8 grid points are needed to resolve a discontinuity.

The sentence is rewritten as follows: “... even though the dynamic equations are solved on a low-resolution (10 km) triangular mesh using the finite element method.”

Equation 5: Are the results sensitive to the choice of exponent in the P_{max} equation? I.e. power $3/2$ as opposed to linear.

The power $3/2$ is for thickness in Eq. 5, it has no impact on the results where concentration and damage are modified.

Line 196: This sentence should read: “Only the older [sea] ice [concentration] is updated in the assimilation...”. The word “concentration” is missing.

The sentence is updated accordingly.

Line 234-235: “The effect of assimilation on the prediction skill is evaluated by comparison of the simulated and observed total deformation fields as it is crucial information for safe navigation, ecological and climate studies.” Ecological and climate studies have much longer time scale than the 2-days improved predictability from an assimilation approach. Only safe navigation should be kept as a motivation in this sentence.

The sentence is updated accordingly.

Line 244: The Kolmogorov-Smirnov Distance is the difference between cumulative distribution functions (CDFs) not PDFs. This should be corrected. It is however correct that the KS distance can be used to compare two distributions and assess whether two PDFs were drawn from the same underlying probability distribution.

The sentence is updated accordingly.

Line 371: The free run was not defined.

“Free run” is replaced by “truth run (x^t)”

Line 378: “In the second element (orange lines on Fig. 10) the initial deformation at the break-up event is larger, the concentration decreases rapidly and, as a result, the deformation on later steps reaches much higher values.” Define “rapidly”. A lead opening over a 6-day time scales does not seem particularly rapid.

The scale of the X-axis is given in format MM-DD HH and concentration decreases by almost 10% in 6 hours. The sentence is rewritten: “... the concentration decreases rapidly (at a rate of $\sim 1\%/hour$) ...” and the note on axis format is added to the figure caption: “The scale of the X-axis is given in format MM-DD HH.”

Bouchat et al. reference. Update this reference from the ESSOAR to the published version.

The reference was updated.

Line 561-562: “However, in Eq. 21, concentration is a function of ϵ_{tot} assuming that ice breaks and becomes weaker due to both convergence and shear.” But in observations ice become weaker when it deforms whether there is a decrease of concentration (divergence) or not (convergence).

The sentence is rewritten as follows:

However, in Eq. 21, the concentration of older sea ice is a function of ϵ_{tot} assuming that ice breaks and becomes weaker due to both convergence, divergence and shear.