Review of "Implementation of a brittle sea-ice rheology in an Eulerian, finite-difference, C-grid modeling framework: Impact on the simulated deformation of sea-ice in the Arctic" by Brodeau et al, (MS gmd-2023-231)

The manuscript describes the Brittle Bingham-Maxwell (BBM) rheology of Olason et al (2022) and its implementation into SI3, the sea ice component of NEMO. Particular emphasis is placed on the implementation of a staggered C-grid. A few simulations serve to evaluate the implementation relative to statistics of observations and the default VP-rheology with an aEVP solver. The structure of the text is mostly clear.

Clearly this is important work that will make it possible for a large community sea-ice/ocean/climate models to use the BBM model in their simulations. The manuscript, however, requires major revisions to turn this into a scientific paper.

As a preamble, we want to start by sincerely thanking the anonymous reviewer for his extensive review of our manuscript, both in terms of depth and length. It contains numerous good and constructive points, and we think it significantly contributed to making our manuscript way more accurate and solid than it was at the first submission stage. We tried to address every comment and suggestion carefully, which is also why our response to the reviews took quite a significant time.
Yet, we regret the tone used sometimes by the reviewer, which in our opinion was not always appropriate, not always respectful of our work, and suggesting an intentional lack of intellectual honesty from our end.

Before moving to a point-by-point reply format, here is some important information regarding the new simulations performed for the new version of the manuscript.

Based on a concern expressed by both the anonymous reviewer and Dr Plante on our initialization/spinup strategy, we have decided to use a simpler approach. As such, our simulations have been performed again: initialized with the new strategy (described in section 3.2 of the new manuscript), and consequently, all diagnostics, figures, and table values have been processed and generated again, based on these new simulations.

Also based on concerns raised by the reviewers, on what has been wrongly interpreted as unfairness towards the aEVP rheology, we now make it very clear, all over the paper, that what we use as the reference SI3 simulation is the default workhorse setup of SI3 as provided in the current version of NEMO. As such, for instance, experiment "SI3-aEVP" has been renamed to "SI3-default", and it is clearly stated in the paper that a better-tuned "aEVP" would likely perform better (L513-516).

Furthermore, following the suggestion of reviewer #2, we have included a new section (3.1) that presents results of idealized SI3 simulations run on the "cyclone" test-case of Mehlmann et al. 2021, using both our BBM implementation and the default aEVP setup of SI3 (new figure 5).

Another important information regarding the new simulations:
In the first version of the manuscript, our developments and simulations were based on the version 4.2.1 of NEMO. At the time, it was the current stable release of NEMO. But during the course of the review process of this paper, a bug related to the ocean-ice drag parameterization has been identified and fixed by the NEMO team. This bug has been judged sufficiently severe by the NEMO team to justify the release of a new stable version: the 4.2.2.
Link to release note: https://forge.nemo-ocean.eu/nemo/nemo/-/releases/4.2.2
Consequently we have switched to version 4.2.2 of NEMO for all the simulations presented in the new version of the manuscript.
The bug fix has significantly impacted the values of the simulated deformations and has therefore required a new tuning of the air-ice drag coefficients.

Major issues

The main problem with the manuscript is that the statements in the introduction, discussion and conclusion are often not backed by the presented work. Instead, the reader gets the impression that this is a text that tries to "sell" this implementation, where a scientific paper should describe the work and provide objective statements about the value of the work in the context of the scientific discussion. Here, most of the text implies that one should only use the current sea ice model with BBM, because everyone else is getting it work. I am exaggerating a little. Because is a general "impression" after reading the text, it is very hard to pinpoint individual issues, because the phrasing etc is scattered throughout the manuscript. Here are some more prominent example, and I have listed many places where I found the wording or phrasing to fit for scientific paper. For example:

- The statement about heterogeneity and coupled model in lines 17-20 is too strong. It should be clear that the heat fluxes are very strong in leads and that this is important, but it not as clear how important the transition from parameterising this (i.e. in terms of fraction ice cover) to resolving the leads is. As long as the atmosphere does not resolve the heterogeneity of the high resolution surface fields (typically in coupled system, atmosphere models have a much lower resolution than the ocean/sea ice system, because the characteristic eddy scale is so much larger), it is not clear if this heterogeneity has any impact on the atmosphere, because one needs to average over it anyway.

We toned down the sentence the reviewer is referring to. It now reads:

“This stresses the relevance of accurately representing sea-ice dynamics in simulations of the coupled/multi-component earth system, such as regional and global climate simulations, and even short-term sea-ice predictions.”. (L17)
We bring to the reviewer's attention the following sentence taken from the introduction of Hutter et al. (2022) that seems to point towards the same idea, and with whom we tend to agree:
“To directly simulate these processes and to provide a more detailed picture of the complex Arctic climate system, we need sea ice models that explicitly resolve LKFs.”

- Already in the model description section (2.3, page 7), the authors come to the conclusion that (l180) “using the C-grid is not the most appropriate choice”. Then (l183) “The spatial staggering between the point definition of the normal (diagonal) and shear (off-diagonal) elements of these tensors becomes an issue whenever the parameterization of the constitutive law requires e12 or sig12 to be known at a T-point.”

These are unbacked statement. It may be the result of the work presented here, but cannot be the starting point. It’s clear that the spatial staggering is not good for all terms. One may argue that the stress tensor divergence is the leading term in the momentum balance. But it is not clear, why this is worse for sea ice models than for ocean models, where the dominant geostrophic balance also suffers from the velocity staggering on C-grids. That’s why early ocean models were implemented on B-grids with co-located velocities. The C-grid model has been shown to violate wave dispersion relationships and the Coriolis terms leads to numerical noise. But this noise has been found to be less problematic than the issues of the B-grid, so that now ocean models (e.g. NEMO) mostly use the C-grid in spite of problematic discretisation of the main balance in the ocean (geostrophy).

Since this manuscript stresses the staggering issue so much, it should become clear to the reader, why it is worse for the stress tensor than for velocity vector to not be collocated at one grid point.

We have completely rewritten the parts the reviewer mentions, taking into consideration his remarks, and avoiding any judgmental or “negative” phrasing about any type of grid arrangement everywhere in the new version of the manuscript.

- Interpretation of figures, e.g.

— l372 “power-law tail (in Fig7)”

Power laws lead to a straight line in a loglog plot (as the reference with -3 shows). None of the curves, not even the RGPS curve show that. Clearly the aEVP solution differs more

from the RGPS solution than the BBM solution. But I think that this metric does not allow any strong conclusions (such as “suggests the advantage of BBM over aEVP”).
This is true that none of the pdfs are showing perfect power-law tails, and that they are indeed showing a slight curvature/departure from a straight line. This is, by the way, why we include the -3 power law exponent as a “reference” in the figure of the PDFs. The point of our statement about the tails of the PDF being close to a straight line in a loglog scale is another way to say that the variable follows a distribution of values whose moments are strongly impacted by the most extreme values when being computed. This would for example not be the case for a process characterized by Gaussian statistics.
In other words, what is important here is not the perfect alignment with a power law, but rather to note that the tail is clearly closer to a power law than what a Gaussian distribution would exhibit. That being said, we corrected the text to report that the tails of the PDFs can be “approximated” by a power-law. (L606)

Further the aEVP model is not tuned to give high deformation events. It is possible to do so (see Bouchat et al 2022 and Bouchat and Tremblay 2017), so this comparison does not “demonstrate higher skills” of the BBM model, but only the difference between two simulations with default parameters. I think that the authors need to tone down their conclusions.
As suggested by the reviewer, we toned down our conclusions, and modified the text accordingly, which now reads: ”We note that the extreme values of deformation rates are, if not absent, largely underestimated for SI3-default in our setup, as highlighted by the departure between the observed and simulated PDFs shown as color bars below each panel of figure 8. This may be improved by better tuning the parameters of the VP model, in particular the ratio between the ice compressive strength and the ice shear strength (Bouchat et al. 2017). However, the scope of this paper is not to perform such a tuning.” (L513-516)

— l389 “As illustrated in figure 8, …”

Same as before, clearly the agreement of the BBM simulation with RGPS is better, but the specific aEVP solution has not been tuned to give high deformation, so that designing a metric that emphasises high deformation (for which the brittle model family was designed) is a self-fullfilling prophecy and of little value.
It is not clear to us what the reviewer wishes to achieve with this comment. The figure demonstrates that we manage to do what we set out to do; namely to have a good representation of high deformation events in the model compared to observations. The standard aEVP solution is included as a reference.

What’s intersting is that the aEVP solution is not always worse than the BBM-solution (e.g. 03-15 to 04-01). It would be interesting to understand, why BBM, from which we always expect higher deformation, underestimates the high deformation events here.

We agree that there is still much to learn from the analysis presented in (now) figure 9. Unfortunately, a deep analysis of a single metric, including the investigation of single events, is outside the scope of the current paper.

— l416/Fig9: " … but seems to break for scales larger than about 300 km"

Hutter et al 2018, doi:10.1002/2017JC013119 found that the scaling for VP models breaks down at around 10 times the grid spacing, marking the effective resolution of grid-point model where the VP model dynamics are resolved (i.e. you need 5-10 grid points to represent a sharp transition without inflicting numerical issues). With dx=10km in this simulation, it means that anything **below** 100km should be considered as unusable, i.e. the fit should be applied to the 3 points above 100km, which would lead to larger slopes. Brittle models scale down the grid scale, as they are designed to produce heterogeneity at the grid scale. The interpretation of the fig9 hence needs to be revised.

The reviewer is correct when suggesting that our interpretation of (now) figure 10 could have been more extensive. The important point here is not the slopes of the fit, but the fact that BBM in SI3 scales to the resolution of the grid. This has only been shown in more "exotic" modeling frameworks, so it's a new result. The reviewer suggests that we fit only the coarser spatial scales, but whether or not we do so is immaterial to the point above, so we prefer to keep the same, statistically rigorous approach for both models. In the revised paragraph we try to emphasize these points. The new text now reads:

"Figure 10 suggests that the total deformation rates simulated by SI3-default ceases to follow the expected power-law for scales larger than typically 100 km. This is in line with published results (e.g. Hutter et al., 2018; Bouchat et al., 2022). Hutter et al. (2018) argue that the VP model needs approximately ten grid cells to be able to resolve features, which suggests that the "effective resolution" of the model is ten times coarser than that of the numerical grid on which it is run. This implies that one should instead consider fitting the deformation rates at a resolution ten times coarser than that used by the model, i.e. 130 km in our case. This would yield power-law slopes that are in better agreement with those derived from observations. We argue that since sea-ice deformation is a scale-invariant process at the geophysical scale, a sea-ice model should be able to represent this scaling down to the model grid cell. Figure 10 suggests that our BBM implementation allows SI3 to achieve this despite the use of the Eulerian framework.."

- In the conclusions we read (l520) "Based on our results, we conclude that the ability of a continuous sea-ice model to simulate the complex sea-ice dynamics across scales, as observed from satellites, depends primarily on the type of rheology used rather than on the type of modeling formalism chosen (i.e. Eulerian versus Lagrangian)."

It is very likely that this is the case, but the statement is way too strong. I think, one can conclude that the difference between neXtSIM and SI3-BBM is smaller than between the

SI3-BBM and aEVP implementation (although no direct comparison with neXtSIM was made here, only indirectly), in particular in the observed statistics that are attributed to “complex dynamics” of deformation. There are no large scale comparisons about sea-ice distribution, thickness etc. so that the concluding statement can only be about the deformation properties.
Due to the major re-working of the conclusion, this sentence does not exist anymore.

Cited literature is quoted only where it fits the story of superior brittle models, but the same literature is not used to discuss the pros and cons, e.g.

- l32: “poses a fundamental and major challenge (e.g. Bouchat et al., 2022; Hutter et al., 2022)”.

Interestingly, “challenge” appears only once in Hutter et al, and never in Bouchat et al. Instead, both papers point to the ambiguity of scaling metrics and multi-fractality as a metric for evaluation different sea ice model. These references do not seem support the statement in this paragraph.
Hutter et al. (2022) conclude that “In this comparison between different sea ice models and model configurations, only very few models reproduce some statistics of LKF properties, namely density, number, length, and growth rate, within an acceptable range as defined by the interannual variability of satellite-based RGPS deformation data. Most models, however, simulate unrealistic LKF distributions”. This can surely be called a “challenge”, even if Hutter et al. do not use this exact phrase.

Bouchat et al. (2022) conclude that “the VP/EVP rheologies implemented in an Eulerian framework need to be run at higher resolution than that of the observations to yield spatial scaling exponents as high as those observed”. Again, it is clearly challenging to reproduce the observed spatial scaling, since very high resolution is a prerequisite.

In our opinion, it is not necessary to go to such lengths in explaining exactly how we interpret that the results of Hutter et al. (2022) and Bouchat et al. (2022) constitute a challenge in the introduction. Specifically, because we do not intend to address these challenges directly, but rather to present a new implementation of a rheology that does attempt to address these challenges.

Instead, I think that Bouchat and Hutter especially highlight the need for more than one MEB model to compare to and introduce more and different comparison and evaluation tools. This manuscript introduces an additional brittle model implementation, but uses the scaling analysis that was shown by Bouchat and Hutter to be not sensitive enough to discriminate between models.

In this paper, we oblige Bouchat and Hutter by presenting a new implementation of the successor of the MEB model. We hope that the community will take advantage of it and use it in future model comparison studies!
The point of our analysis is not to discriminate between models, but to show that our BBM implementation can reproduce the observed spatial scaling. The fact that the default SI3 setup does not show that despite Bouchat et al's conclusions, then doing so is not a foregone conclusion. Unfortunately, Bouchat et al. never show the higher order moments in their analysis, so we don't know how well the models in their study reproduce this important part of the observed statistics or whether they scale down to the grid resolution. By showing the scaling of all three moments, we demonstrate that BBM in SI3 does scale down to the resolution, which is an important goal with the development of the brittle rheologies.

- Plante et al. 2020 is probably the first implementation of MEB on a C-grid; it's not BBM, so the author's can claim that their implementation is the first BBM model the C-grid, but as far as I understand the differences between MEB and BBM (limit the maximum allowed compressive strength P), the C-grid plays no role in this so we can assume that Plante et al have experienced the same issues with the C-grid report here. It is true that (l63) "The idealized nature of these simulations prevented their results from being assessed against observations of sea-ice drift and deformation.", but equally one could say that avoiding idealised configurations makes it impossible to detect (small) implementation errors, that would, e.g., break symmetry etc. In l70, it is written that "[spatial interpolation the C-grid] is not well-suited for brittle rheologies", but it is not mentioned that Plante et al (2020) did find a solution to obtain stable solutions by a mix of double defined variable akin to the E-grid solution presented here and averaging. In line 196, Plante et al is cited to have observed checkerboard (not chessboard) instabilities, but the authors fail to write, that these patterns never appear in the presented solution, because Plante et al designed (and described) a the numerical scheme to avoid them. Instead, the manuscript gives the impression that their E-grid solution is the only way to solve this, thereby neglecting previous successful methods.

It is also not discussed that in Plante's model, the heterogeneity does not appear to be as chaotic and noisy as in the neXtSIM publications. Maybe it makes sense to actually try to reproduce Plante's results?
As you will see, in the new version of the manuscript, we have completely rewritten the parts (and we have added new parts) related to the work achieved by Plante et al. 2020.
To reply to your question, in the early phase of our development, before choosing to use the E-grid, we have tried to implement the method of Plante et al. 2020.
Unfortunately we have not been able to get rid of the checkerboard instabilities, we see two possible reasons for this:

- A/ we have not done it properly

- B/ the numerical scheme of SI3 used to solve the equation of momentum is explicit (i.e. $u^{t+1} = f(u^t)$, apart from the use of u^{t+1} in the ocean-ice stress term), whereas with the McGill model, Plante et al. 2020 used an implicit solver (i.e. $u^{t+1} = f(u^{t+1})$), an approach that is more sophisticated and more numerically sound.

Even if we would have been able to correctly implement the method of Plante et al. 2020, we would have chosen to go for the E-grid approach for the following reasons:

- i) the damage variable at the corners of the grid cells is never averaged, so σ_{12} does not use a smoothed damage while σ_{11} and σ_{22} use a non-smoothed damage
- ii) we can advect all the components of the stress tensors using the exact same numerical scheme as that used for all the tracers
- iii) we obviously liked the idea of testing a new approach

Reasons (i) and (ii) are now clearly stated in the new version of the manuscript (end of section 3.2). We also added some discussion on why we think it is important to include the advection terms of the stress tensor components in a brittle rheology that relies on a damage tracer (section 2.2, L159-168).

“Cross-nudging” is a new filter introduced in the MS and this seems to be a crucial part of the E-grid it requires more attention. The authors (l263)) “conclude that the right compromise is achieved when gamma_C typically lies between 1 and 3, with 2 being the value used in our experiments”.

This is a central issue. In order to avoid the staggering issues of the C-grid, a new issue is introduced: the E-grid solutions diverge and need to be coupled explicitly. The coupling parameter is found by trial and error (fine with me), but there is no reference as to which metric is used other than “the right compromise [between smoothing and coupling]”, which is somehow related to Fig5. What is this “right compromise”? I need to know if I want to, as suggested, reject the C-grid (without attempting to fix the issues on it) in favour of an E-grid with different issues that need to be fixed. “spatial consistency”, “smoothness” (or rather the absence of it) are more soft metrics, which I cannot evaluate. I do not see a qualitative difference between 5b and 5c, but 5d appears even smoother hence it is rejected. The structure in 5a (no cross nudging) shows the checkerboard like patterns that Plante describe, but I don’t see a figure with gamma_C below 1 (but >0), where the solution is supposed to become increasingly noisy/

We think that we have now significantly improved the parts dealing with the cross-nudging (CN), providing more discussion behind its logic (section 2.3.2).
With respect to the first version of the manuscript, the method has also been improved after realizing that the CN should be applied on the vertically-integrated stress components (i.e. $h\sigma$) rather than on σ alone (little reasoning with a pen and a sheet of paper easily demonstrates that not using $h\sigma$ in equation 16 may introduce errors in regions with strong gradient of ice thickness). Some discussion has also been added about this point.

Also, as you suggest we have added more maps showing the effect of using different values of the CN parameter. Since we now use a value of 1 (rather than 2 as in our old simulations) we show figures for values: 0, 0.1, 0.5, 1., 2., and 10.
We make clear, in the discussion about this CN parameter, that it is up to the user to choose their own value, as a tuning parameter, but we also indicate that based on the various experiments we have performed, and as suggested by the new figure 4, the value of the CN should be of the order of 1. (L302-314)

For choosing gamma_C, I only see the possibility of generating some reference (observational data? NextSIM output?) and somehow define a "goodness of fit" to this reference data when you tune the cross-nudging parameter. It would be very interesting to see, if one couldn't achieve something similar with a C-grid and some averaging as done in Plante et al (2020).
Yes, it would be interesting to do what you suggest, yet we still have hope to find a method more elegant than the cross nudging, and more physical, or at least more numerically sound. We think that the message we convey in the new version of the paper clearly underlines the fact that the CN is an approach that is quite *ad-hoc*, but an approach that has the benefit of demonstrating that the E-grid approach works.

There are many judgemental statements that support the general "salesman" tone of the manuscript, for example,
- l204: "To avoid the problems related to the staggering of the C-grid"

- l234: "Thanks to"

that suggest that something is problematic, or better or worse, without any support in the text (or proof).

The language is often sloppy, there are many unnecessary repetitions. Often it sounds like listening to an informal talk about the subject (where the language would be OK to my mind). There are a few grammar problems, and many places where the formulations could be made much more concise (by removing unnecessary words and phrases).
We apologize for the sloppy language. As for the repetitions and grammar problems, we tried to remove them in the revised version of the manuscript. We also definitely tried to follow the reviewer's recommendation of making our text more concise. And again, we have done our best to strip all the judgmental/negative wording off the text.

Smaller problems, typos, technical issues:

Abstract:

Taking into consideration both the major and abstract-specific remarks of the reviewer, our abstract has undergone a substantial re-writing.

page 1

l2 new spatial discretisation framework
The E-grid is only new to NEMO, rephrase
This sentence no longer exists.

l3 well adapted to solve the equations of sea-ice dynamics
What is “well adapted” in this context. Can be removed
This sentence no longer exists.

l3: the numerical issues posed by the use of the staggered C-grid.
What are these issues?
This sentence no longer exists.

l6: "when using the newly-implemented BBM rheology and when"
grammar? Main clause is missing
This part has been completely rewritten.

l8: "the relevance of the use of this newly-implemented rheology for future modeling": Awkward, rephrase, no need to emphasise the usefulness of the present work.
We agree, we have completely removed this sentence.

l10: "This includes, in particular, coupled climate simulations performed with CMIP-class Earth System Models at coarse to moderate spatial resolution.": There is no information in this sentence.
Same here, sentence removed.

Introduction:

l14 Sea-ice is one of the most important physical interfaces

Not sure if sea-ice can be reduced to an “interface” (especially since this “interfaces” uses about 50% of the computer time, Table 3)
“interface” has been replaced with “component”. (L13)

page 2

l24: "the abundance …": Please add some references

We now have added the references to both the IABP (buoys trajectories) and RGPS (satellite-derived ice trajectories) datasets (L24-25).

l34: "Following the work of Girard et al.": I do not think that we need the history of brittle models again.

We removed this part and now get directly to the point (L34).

l45: These two constraints have proved to be impossible to respect with MEB because of an incomplete treatment of the convergence of highly damaged sea-ice, which results in unrealistic sea-ice thicknesses after a couple of years of model integration.

Does that mean, that MEB, while always being superior to other models (according to cited references), cannot even get the fundamentals right? Also, how much of this can be attributed to MEB, and how much to the specific implementation in neXtSIM?

Our answer, starting with a short historical recap:
This limitation of the MEB rheology, as introduced in Dansereau et al (2016) and implemented into neXtSIM in Rampal et al. (2016), was not documented in these two papers simply because it was at that time tested in the context of winter simulations only, typically for the Jan-Apr time period. Over these relatively short winter time periods, the bias in the thickness of the ice pack cannot be clearly seen. This is only after running the model for several years that the bias was discovered, and its origin identified/interpreted in terms of missing physics. This motivated the development of the BBM rheology to fix this problem, among others (Olason et al. 2022).
We can confirm to the reviewer that this issue should not be attributed to the implementation in neXtSIM, but to the lack of any ice resistance to converging motion when the ice concentration is close to 1 and the ice is highly damaged.

l49: pure -> purely

The sentence is no longer present.

l56: "excellent": What does make the scalability “excellent”? Reference? I think the use of superlative adjectives needs to be re-considered.

While one could argue that the scalability of NEMO is “excellent” with respect to that of neXtSIM, we agree that the adjective “excellent” is too strong in a broader GCM-related context. Therefore, we have replaced “excellent” by “good” and added the reference to Tintó Pims et al. 2019 (L55).
In the same sentence, we have also replaced “SI3” by “NEMO”, because mentioning SI3 rather than NEMO was an inaccuracy from our side. SI3, just as its predecessor LIM3,

entirely relies on NEMO's MPI horizontal partitioning, and cannot be run in a standalone mode outside of the NEMO environment.

page 3

l60: double "))"
Corrected.

l74: In this paper, we propose a solution to this problem and provide a detailed description of the implementation of BBM into 75 an Eulerian, finite-difference, staggered-grid modeling framework; namely that of SI3, the sea-ice component of the NEMO modeling system.

There is a lot of repetition of previous paragraphs in this paragraph. It may be worth it to try to streamline the introduction to avoid unnecessary repetitions.
We have done our best to do so.

l84: some important aspects

It's always good to "discuss some important aspects", but what are they? Rewrite to be more specific.
The "important aspects" are now clearly stated in our new sentence:
"In section 4, we discuss some numerical aspects of our implementation and some limitations of the BBM rheology." (L81)

page 4

eq(1) sign error in Coriolis term? Unless \vec{k} points downward.
It was an error, well spotted. Corrected.

l91: Appendix A1

It is tedious, but I think the notation needs to be introduced where it appears (on top of Appendix A1, to which I do not always want to refer, when reading the manuscript)
The notation is now introduced after Eq. 1, and other equations. (L86)

l92: writes -> is written as
Corrected. (L91)

l98: where the underbar notation indicates that the tensors are expressed in their pseudovector form, and K is the elastic stiffness tensor

I guess “K” is also in its pseudovector form? Why no underbar? The pseudo vector form is also called Voigt notation, maybe add to make it clearer to more readers.
Yes, K is in its pseudovector form, we have added the underbar. And we now use the term “Voigt form” rather than “pseudo-vector form” throughout the text. (e.g. L99)

l109: I don’t think it “happens to differentiate”, but it

differentiates, also since it is specific to BBM, this is a bit of repetition. Please rewrite.
The sentence has been rewritten as follows:
“The BBM constitutive equation (4) only differs from that of MEB through the inclusion of the term $\tilde{P}$.” (L106)

page 5

l110: the excessive convergence of ice when damaged

I am sure that the physical motivation for the form of this term (eq8) is described at length in Olason et al. Still it would be a courtesy to the reader to repeat the reasoning here, because it appears to be so fundamental to BBM. Otherwise it just appears to be a quick fix to solve a severe MEB problem.
We’ve added new information summarizing the physical motivation for this term (L106-110).

l114: Ólason et al. (2022) follow Dansereau et al. (2016)

That’s nice of Ólason et al., but what do you do in this paper?
We changed the sentence to “We follow Dansereau et al. (2016) and Ólason et al. (2022) in using a two-step approach to solve equation 4.” (L116)

Also, what is a “two-step approach” in this context? I guess this follows in after l115, but it’s not clear from this sentence.
The next sentence now starts with “As the first step, …”, which helps the reader understand that we are explaining what these two steps are. (L116)

Don’t get me wrong, I am fine with omitting details and referring to previous papers for them, but here the mix is strange: Many (all?) details of the equations of the model are repeated, but some steps in the solution method are omitted. As a reader, I would be fine with saying: The BBM model is described in Olason et al, we do everything in the same way, please refer to Olason et al. Or put all or most of this into the appendix, as it is not new. The only information that I need as a reader is the stress tensor and its discretisation (according to the introduction and abstract).

Thanks to your feedback and some reshaping, we think this part is now easier to follow in the new version of the manuscript and we wish to keep it.

l127: "In the case of the BBM framework, Ólason et al. (2022) and the damage criterion shown in Fig. 1 and dcrit expresses as follows"

Something is wrong with this sentence, please fix

The sentence has been simplified to "In the BBM framework, d_{crit} is expressed as follows:" (L128)

l132: "for the healing the ice": for the healing of ice OR for healing the ice

For the healing of the ice. Corrected (L133).

l132: which is associated with refreezing within open leads and which is therefore based on a rate of decrease of the damage that depends on the temperature of the ice.

Please rewrite to disentangle the various relative clauses and the not always correct usage of "which" and "that".

We have rewritten the two following sentences instead (L133):

"As suggested by Rampal et al. (2016), a slow restoring process is applied to the damage to account for the healing of ice under refreezing conditions. The rate of decrease of the damage associated with this refreezing is taken proportional to ΔTh , the temperature difference between basal and surface ice:"

page 6

l145: "which" Start a new sentence, too many relative clauses

We now start a new sentence (L144).

l146: "In non-regularized frameworks" Not clear what this means in this context. What is a regularised framework, in contrast?

A regularized framework would be one in which we would introduce, via the constitutive equation for instance, a characteristic length scale for the process of interest: here, deformation (or damage). In such a framework, deformation (or damage) would localize at this characteristic length scale. We do not introduce such regularization in our model and in sea ice models in general. One good reason is the fact that no characteristic length scale for the deformations seems to stand out in the observations, over a wide range of space scales. Even if we would consider a mean, characteristic width for leads, this length scale would be much smaller than the model resolution.
The mention of the term “non-regularized” was not really necessary in the context, and confusing, therefore we have rephrased this sentence as follows:
“The combination of elasticity and damage, even if treated in an isotropic manner, naturally simulates a strong anisotropy and localization of the deformation, down to the nominal spatial and temporal scale…” (L146)

l150: tend to exhibit very sharp gradients, or "near-discontinuities"

It needs to be shown that this is specific to brittle models. The neXtSIM MEB papers always imply this, but e.g. the fields in Plante et al (2020) are generally much smoother. Also these “near-discontinues” also appear in high-resolution VP simulations (e.g. Ringeisen et al 2019, 2021).
We removed the term “near-discontinuities” from the sentence (L149-150)

L152 remove the “,”
This sentence is no longer present.

l159: "In BBM ..."

This seems to be a specific issue with the specific BBM implementation of Olason. In theory (Dansereau et al 2016 describe an iterative procedure with a tolerance, their page 1350, rhs column, Plante et al update damage, E, \lambda during an iteration), it should be possible to iterate the “two-step approach” until convergence.
As noted by the reviewer, Dansereau et al. use an iterative approach to solve the MEB problem. This is not the choice made by Olason et al. . In order to fix this inaccurate statement in which we were opposing the brittle rheologies in general to the viscous-plastic ones. This part have been heavily re-worked and instead at the end of the section 2.2, we have a more developed about the time splitting approach (starting L169):
“Finally, note that in their numerical implementation of BBM, Ólason et al. (2022) chose to solve the dynamics explicitly using a time-step sufficiently small to account for the propagation of damage in the ice in a physically realistic manner.”

This discussion about details of the time stepping and convergence is a little akin to the EVP evolution, where the somewhat naive iterative process lead to a (noisy) solution

that was not intended, until papers like Lemieux et al 2012, Boullion et al 2013, Kimmritz et al 2016 came up with a solution for this (revised, modified, adaptive EVP).
We agree.

page 8

l204: To avoid the problems related to the staggering of the C-grid, namely the interpolation

I strongly suggest avoiding this type of judgemental phrasing here and elsewhere, write instead: To avoid the interpolation of … due to the staggering of the C-grid …

I am pretty sure that the E-grid is not without issues, and I am waiting for similar statements about the "problems related to the E-grid".
We have modified the sentence accordingly, it now reads:
"To avoid the interpolation of the damage and the stress components between the center and the corner points of the grid cell, and allow the consistent advection of all the components of the stress tensor, an additional sea-ice velocity vector, noted $(\hat{u}, \hat{v})$, is introduced." (L225)

l209: Arakawa E-grid

BTW, there was a successful ocean model that used the E-grid: The Hamburg Large-Scale Geostrophic (LSG) model (by Ernst Maier-Reimer). Here, the E-grid was chosen, because the dominant balance (geostrophy) can be expressed more accurately and without noise while retaining some of the "nice" properties of the C-grid (representation of divergence). Maybe that's an analogy worth mentioning.
Absolutely, we were not aware of this and we thus thank the reviewer for sharing the reference. We have added a reference to Maier-Reimer et al., 1993 when first discussing the E-grid. (L230)
In section 2.3.2, that deals with the problem of the decoupling of the 2 solutions with the E-grid, we also added a couple of sentences mentioning the reason behind the choice of this grid in LSG and how this decoupling, which primarily affects "short gravity waves" is not a problem based on the large time step of the model, and how their use of "moderate" viscosity and diffusion prevents the "split-modes" for the slow modes that are resolved. (L264)

page 9

L234 "Thanks to" -> "With"

Again this is judgemental, this time in the “positive” sense. Statements like these set a certain tone that appears biased and scientifically non-objective.
Done (L255).

l237 “and no interpolation is needed to solve the equations” But this advantage comes with the disadvantage of the cross-nudging
Yes. The following sentence starting with “It does, however, result in an …” we think the reader is made rapidly aware that the E-grid approach is not perfect (L266).
We also mention this in the conclusion (L641-L644).

l248: upper-convected time derivative

Later this term is introduced properly (l289), maybe do it here already. Or just use “advection” for simplicity here and introduce the upper-convected time derivative later.
We have completely re-written this part. (L270)

L259 and eq16: denoted by interpF@T and interpT@F,

As this is a math expression, why not use more “math-like” symbols to denote the interpolation, e.g. \overline{\hat{\sigma}_{11}}^T, where \overline{...}^T means interpolation to T points, etc.
Equation completely re-written following your advice (L283).

page 10

l264: Don’t refer to figure 5 before figure 4?
Figure ordering should be fine now.

l266: horizontally and vertically aligned with the grid cells. -> aligned with the grid
There is no “vertical” in a 2D horizontal grid.
The corrected sentence is now: “... that are aligned along the x- or y-axis of the grid.” (L308)

L289 upper-convected time derivative

It would be good to add a reference here, as this terminology of complex fluids is probably not common knowledge of the GMD reader.

After reading up on this (in a book about polymer flow!) it is not even clear, why we have to use the upper-convected time derivative, and not the lower-convected time derivative or a linear combination of the two. In Danserau et al (2016) something similar is called the Gordon-Schowalter derivative (it’s not the same but some linear combination of the upper and lower convected derivative), and obviously it is not entirely clear what is the correct form to use, as any frame-invariant time derivative is formally allowed. So some discussion with appropriate reference seems in place here.

Section 2.4.1 has been largely re-written taking into account your suggestions, it now includes:

- appropriate references
- discussion about the existence of the 2 formulations (upper- and lower-convected), and about the resulting dilemma faced by the modeler
- inclusion of the equations for the lower-convected formulation in the manuscript, that we have also coded in our implementation to perform sensitivity experiments

We also added a paragraph to justify our choice to use the upper-convected form in our simulations at the end of section 2.4.1. Also a new figure (C3) in Appendix C, at the very end of the manuscript, shows the impact, on the PDFs of simulated deformations, of using or not these two different formulations.

Eq17 and 18 do not describe the upper-convected time derivative. WIth L in eq18 and a plus-sign in Eq17, this would be LOWER convected time derivative.

The component form (eq19) is correct (and consistent with the form of L = (\nabla\vec{u})^T \cdot \sigma + \sigma\cdot\nabla\vec{u}) for the upper convected time derivative)

Eq18 would give components with \partial_y U and \partial_x V exchanged (w.r.t eq19).

Yes, we have corrected our mistake, and now equations for both lower- and upper-convected time derivatives are provided (Eq. 18,19,20,21).

page 11

l299: Then, the tensor-specific contribution −L is added.

Is this done successively, I.e. use the sigmas after advection with the material derivative D/DT to compute L (i.e. some sort of split operator method), or do you use the sigmas before applying D/DT? Please be more precise.

Thank you for raising this point, it helped us find a bug: we were wrongly using the sigmas that had already undergone the material derivative update!

This has been corrected in the new version of our implementation and all the simulations discussed in the manuscript have used the correct version of the code.

So thanks to your input, we can definitively write that:

`L` is computed using the stresses that have not been updated by the `D/DT` yet. It is indeed important to mention this, and the text has been modified accordingly. (L358-359) l305: It largely inherits from LIM3 Rousset et al. (2015), to which it succeeds

Please simplify and fix in-line citations.
We have removed this sentence. Just as you point out below, this information is not important for the paper.

What are “significant inclusions”? Are they important for this manuscript? If not I wouldn’t mention that (again, we do need a full history of the model components), if important, then we need more information.
Sentence removed, see previous reply.

page 12

l324: “We carried out a twin coupled ocean/sea-ice hindcast,”

That is a lot of words for just saying “we compared two simulations”
This sentence has been removed (see beginning of section 3.2)

l329: “while”, wrong connector, -> and
This sentence is gone (re-working due to new spinup strategy).

l331: For the second spin-up segment,

Why does SI3 need this “re-initialisation”? The model should be more or less in balance with the ocean state.

Also the initialisation is short. In my experience, a sea ice model needs some 2-3 years to spin-up (the ocean model much longer, but that’s not really necessary for this paper).
Based on your remarks and the concern expressed by reviewer #3 about our original spinup strategy, we have chosen to use a simpler, yet shorter, spinup strategy and the two simulations have been run again (see the introductory part about the major changes undergone by the paper at the beginning of this document).
We agree that it would take years of simulations to spinup the Arctic ocean in our model, but, as you underline, based on the scope of the paper, we think that it is not a problem to use a short spinup as we do. Note that we initialize the model using the 3D ocean and sea-ice data of a reanalysis (GLORYS2) that used an earlier version of NEMO on almost the same numerical grid and bathymetry (global 1/4° ORCA025 configuration).
See beginning of section 3.2 for details about the new spinup strategy.
We have added a sentence to stress that while our spinup is much too short to obtain a spun up ocean circulation, its duration is acceptable for the scope of the paper (L482).

l333: were extracted from a coupled OCE-neXtSIM simulation

Why do you need the solution from a different simulation to restart the model?
As mentioned in the previous reply we now use a different spinup strategy, and no longer use fields from a different model/simulation. Also see the introductory part about the major changes undergone by the paper at the beginning of this document.

l335: a duration sufficiently long for the coupled system to recover from the ad-hoc reinitialization.
I doubt, that this is long enough. The fast waves will have left the domain, but everything else ...?
We agree, see our two previous replies.

l339: the tuning of SI3 is kept as close as possible to the default namelist

rewrite: tuning is a process and you cannot kept a process close to a namelist.
Our new sentence is: "For these experiments, the adjustable tuning parameters of SI3 are kept as close as possible to those of the reference configuration of NEMO." (L489).

l340: "thermodynamics features"; Grammar?
The sentence is now: "As such, the thermodynamic component uses 5 ice-categories." (L490)

L343 :Table A1 in appendix C.

There are many problems: Why is called A1 if it belongs to section C? Section C1 is one sentence, and is not needed. Table A1 does not contain any parameter values, only descriptions; according to SI standards, units should not be in brackets "[]"; especially in this context the brackets make no sense (same for Table A5).
Section C1 has been removed. The table in question (old table "A1") has also been removed because it is not needed anymore: experiment SI3-default is now 100% in agreement with the default setup of SI3. The fix of the "ocean-ice drag" bug when switching from NEMO 4.2.1 to 4.2.2 (see the introductory part of this reply) allows to use the default ice-atmosphere drag coefficient of 1.4 in SI3-default, in place of the previously used value of 1.15, with the same results when it comes to the PDFs of deformation. And we are now using the default number of aEVP iterations in SI3-default: 100 (180 in previous version).
All brackets surrounding units have been removed from the text.

Paragraph l339 to l347 could much clearer, including the extra information, supposedly in tables in the appendix. The tables could be in the main text, so that the reader does not have to flip back and forth in the paper (from here to appendix C, then back to Table

A1, etc.). Since the time-splitting was already introduced, no need for phrases like “As mentioned in … “ (better: For the time-splitting approach (Section 2.2), we use a small timestep of …)

Reference to appendix C and table A1 no longer exist (see reply to former remark).
There is now Table 1 that lists SI3 default parameters relevant to both experiments SI3_default (in particular those related to the aEVP rheology) and SI3-BBM. And there is Table 2, that lists the BBM-related parameters used in SI3-BBM and their value. These tables are no longer to be found in an Appendix C but in the text.
We reckon that the changes and simplifications made to this paragraph, based on this point and the previous one, have made it clearer.
The Copernicus Latex template for GMD encourages authors to gather Tables and Figures at the end of the manuscript so we keep it this way.

Section 3.3 should be part of a data and methods section. Now it is strangely split between the model evaluation section and the appendix. It would be much nicer for the reader to have everything in one place.

Everything is now gathered at the end of section 2, in the new subsection 2.5, and Appendix C is now used to present some additional figures.

L351 “(RGPS hereafter)” Unnecessary; it’s enough to introduce the abbreviation RGPS earlier in the line.

Done. (L375)

• Highlight, page 13

?

l358: "(see the Code and data availability section." Closing “)” is missing

We have removed this sentence.

L366 in the literature

I guess it’s fair to cite Ron Kwok’s paper about this.

We added the reference to Kwok, 2001. (LX)

L367 quite realistic

What’s the meaning of “quite” in this context?

We have replaced “quite realistic” with “appear somewhat realistic”. (L500)

L368 very smooth fields of deformation with no such localized features

Definitely true, but the figures, where the quadrilateral data is plotted on triangles with gaps in between makes it very hard to read the figures. For example, the aEVP solution seems to be noisy, but I cannot tell if this is an effect of the plotting. Convergence is an issue with aEVP (it's very slow) and one can only expect "smooth" fields at all times if the aEVP parameters are tuned properly, see Kimmritz et al 2016. The smoother the solution, the slower the convergence.
We can assure the reviewer that what he sees as noise is just/and only the result of the plotting using the mask defined by the presence of RGPS observations. There is thus not much we can do to prevent these gaps in these figures, because our deformations are calculated using quadrangles, which, based on our construction process (see new section 2.5.2), tend to exclude some "left-over" triangles that were not merged into quadrangles. Some "empty" space is thus left in between.
For your information, here is a snapshot of the (instantaneous) total deformation (computed online with the *Eulerian* velocities), at roughly the same period, for the two simulations.

L369 consistent

I am not sure, if you can say that, because there are also "coarse" (ie. 10km) runs in Bouchat et al 2022 the have quite some LKFs (their Fig10, the McGill model with smaller e). EVP models tend to have fewer LKFs in that paper (see previous comment about convergence), there's even a comment about EVP models, convergence and deformation rates in Bouchat et al (their section 4.1.1)
It is true that the number and other metrics for LKFs is affected by different model parameters, including the *e* parameter. Our point here is that our aEVP solution is quite comparable with what other people get when using standard (or commonly used) parameter values. We have, therefore, rephrased this sentence as: "... this is consistent with the findings of recent studies that evaluate VP-driven sea-ice simulations run with a horizontal grid size larger than a few kilometers and standard parameter values."

page 15

L429 "A critical requirement for the consistent implementation of the brittle rheology"

This is phrased as a well known fact, whereas this is just what the authors find. An ill-meaning reader could conclude: The authors did not manage to succeed in stabilising the model on a C-grid. Please tone down these statements.
This sentence is no longer present.

L434 The "Leap Frog scheme" may be a good analogy, but by the same analogy, the leap frog scheme is very much outdated and more stable schemes are commonly used in

general circulation models nowadays (e.g. 2nd or 3rd order Adams-Bashforth in FESOM, MITgcm, ROMS, 2nd order Runge-Kutta in MOM6). I am surprised to learn that NEMO still uses a Leap Frog scheme.
We agree. The NEMO development team has been working on the implementation of the RK3 scheme for the last few years. RK3 should normally become the default in version 5 of NEMO, which should be released before the summer.

In this light, introducing yet another filter like the Asselin filter does not "serve[] a useful purpose" (l441)
The sentence has been changed to:
"As of now, our cross-nudging approach clearly lacks physical and numerical consistency, but it somehow allows to demonstrate that the implementation of a brittle rheology, along with the advection of the internal stress tensor, is feasible onto an E-augmented C-grid." (L583)

page 16

l462: that -> which
Corrected. (L596)

l464: When SI3 is coupled to OCE, however, the cost increase is somehow dissolved by the overwhelming cost of OCE and falls below 30%.

It is interesting to note that other groups find that the sea ice dynamics, especially at high resolution, can become the most expensive part of a sea ice-ocean model (e.g. Koldunov et al 2019, doi:10.1029/2018MS001485). I do not share the relief, that the sea ice model does not get "that much" more expensive when coupled to an ocean model. The cost increase definitely does not "dissolve".
In NEMO, the relative cost of the ice component is a bit influenced by the number of vertical ocean levels used by OCE, but with the NEMO reference number of 75 levels (we use 31) the cost of SI3 is reported to be of typically 40% (Clément Rousset, lead developer of SI3, personal communication). This number has increased with the years as more advanced parameterizations, categories, etc, have been added in LIM and then SI3.
Now about our use of the term "dissolved" (that we have removed) and the fact that the numbers are smaller in coupled mode: in regular ocean/sea-ice coupled mode, NEMO couples the ocean and sea-ice modules in a sequential, and not in a parallel, way. This means that the cost of SI3 simply adds up to that of OCE.
The way we formulated this was confusing and we have rephrased the sentence:
"This lower value is explained by the fact that by default, the coupling between OCE and SI3 is done sequentially. As such, the cost of SI3 simply adds up to that of OCE, and the cost of

OCE is expected to be independent of the mode used (in our case: 113 and 114 cpu h for SI3-default and SI3-BBM, respectively).” (L607)
Also, based on a remark of RW2, these “performance” diagnostics have been performed again with the default value of N_{EVP}=100. Also note that in all our new BBM simulations (including those on which the deformation analysis is performed) we now use a number of sub-time-steps of N_S=100 for the time-splitting, and not N_S=180 as before; because based on sensitivity experiments that we have performed during the course of the review we have come to the conclusion that N_S=100 in our BBM implementation is largely sufficient. So now, both aEVP and BBM simulations perform 100 subcycles under one advective time-step, in all the results presented. In that regard, we have also added two new figures in the new Appendix C (figures C1 & C2) that compare the cyclone test-case results for aEVP and BBM using more sub-cycles, N_{EVP}=1000 and N_S=200, respectively.

The numbers in Table 3 tell me that the ocean model uses 157/159 cpu h in this configuration (not sure where the difference of 2h between the setups comes from), the sea ice model 139 or 223 cpu h, so that with BBM, the sea ice model already uses more than 50% of the total time. Even the 139 cpu h of aEVE appear long in this context (47% of the total run time). That’s where Koldunov start to worry about overall performance. I think that this needs to be discussed in more general terms, i.e. how much time to allocate to a small part of our coupled model.
Please see our reply to the previous point, and note that these numbers have been updated due to modifications mentioned in our previous reply.
The cost of the ocean component is now included in the new version of our Table 4.
We will not discuss what you suggest as it is quite out of the scope of this paper, and it is something that should be addressed by the NEMO development team. We plan to work together with them on this matter when merging our implementation into a dedicated branch of NEMO, which should be done before the end of the year we hope.

L472 “an insufficient number of iterations”

aEVP was designed to lead to smooth solution even when the solver is not fully converged. If there are checkerboard patterns in the solutions, then the choice of aEVP parameters is poor and should be improved (“In practice, the value of c depends on forcing, geometry of boundaries and on resolution and has to be selected experimentally”, Kimmritz et al 2016).
Since we now use the default N_{EVP}=100, and based on the fact that a new sensitivity experiment we have performed with N_{EVP}=500 would show the same results (for the diagnostics presented in the paper, see for instance Figure 2 at the end of this document) as the one with N_{EVP}=100, we have removed these 2 last sentences of section 4.1.
We also show the impact of having N_{EVP}=1000 instead of N_{EVP}=100 on the solution of the cyclone test-case (figure C1 in Appendix C).

l477: like for instance about the Arctic sea-ice thickness distribution.

Please rephrase. Also, what are “promising results”? I think it would serve the manuscript, if this were formulated more concisely.
We have completely rewritten the sentence in a more concise way:
“Based on comparisons against various types of observations, recent studies suggest that large-scale models using BBM can realistically simulate the dynamics and properties of sea-ice (Ólason et al., 2022; Rheinlænder et al., 2022; Boutin et al., 2023; Regan et al., 2023).” (L613)

If this were my manuscript, I’d rewrite the entire parargraph along these lines:

Large-scale realistic sea-ice simulations with a model using the BBM rheology showed encouraging agreement of, for instance, the Artic sea-ice thickness distribution with observations (see also, Ólason et al., 2022; Boutin et al., 2023). Yet, deformation in convergence and sub grid-scale processes related to sea-ice ridging are not represented by BBM with the same degree of accuracy. The model overestimates the number of converging events with magnitudes of about 1 to 5% per day, and underestimates, although not as much as the aEVP solution, the most extreme events (Fig 7c, and Ólason et al. (2022). So far, parameter tuning, in particular the BBM-specific ridging threshold parameter P_{max}, did not help to improve agreement with observed convergence PDFs (not shown), so that we conclude that some fundamental processes need to be reconsidered in BBM (and aEVP) [or now some other educated guess/speculation about the BBM (and maybe aEVP) equations that makes it impossible in principle, to get the convergence right].
Thank you! This was helpful. Apart from the first sentence, the paragraph has been re-written accordingly with only minor modifications. (L613-620)

page 17

L488 “are doing better in this particular matter” , Rephrase: “agree better with observations” or similar.
The sentence has been shortened and re-written:
“The fact that the deformation fields simulated by neXtSIM in Ólason et al. (2022) are in better agreement with RGPS in this regard, suggests that this problem is linked to some numerical aspects of our BBM implementation rather than the BBM rheology itself.” (L622)

l490: best -> most likely (and remove the following “likely”)
Done. (L624)

L491 “as these steps are absent”

Rephrase, “numerical dispersion and diffusion” are not “steps”.
The sentence has been improved:
“This is most likely the consequence of the introduction of some additional numerical dispersion and diffusion by the advection scheme and the cross-nudging treatment, respectively, as these two features are absent in neXtSIM.” (L624)

Also you could test that by using other, more diffusive/less dispersive advection schemes.
Yes we should test this at some point. The Eulerian version of neXtSIM that we are currently developing (neXtSIM-DG) uses the Discrete-Galerkin method for the advection, in this regard, it will be interesting to compare BBM simulations run with neXtSIM-DG and SI3.

But what about the cross-nudging process? That was not part of neXtSIM either and could very well be a likely candidate for differences.
Definitely. The cross-nudging was indeed already mentioned in the 1st version of the manuscript (see old L91). It is still mentioned in the updated version, see re-worked sentence above.

L495 “relevant” not the right word here. -> promising?
Yes, promising it is, corrected. (L629)

Conclusions:

In general, I don't think that the tone of the conclusions is appropriate (see main points). As if there is only this solution and everything else is wrong from the beginning. Many conclusions are drawn from trial and errors (as described here and in the text). I think it is good to show and discuss failures, so that others do not stumble into the same problems again, but there needs to be a more systematic list of things that did not work and why. The way this manuscript is written, I get the impression, we are presented with the result of something that finally works in spite of all the hardships encountered along the way, told by the fireside.
Based on your remarks, the conclusion has been significatively reworked. We also put more emphasis on the “ad-hoc” aspects of our implementation, such as the cross-nudging, and those that remain problematic.

L501 The use of the Arakawa C-grid, as used in SI3, has proven to be poorly fitted for brittle rheologies.

The paper does not show this. It states that there are fundamental issues related to the staggering of grid points (but that is a problem not only for brittle rheologies) and describes a method to overcome noise issues that were not even demonstrated. By no means there is “proof” that it cannot work on a C-grid.
You are right. This part has been largely re-written taking your remark into consideration.

“poorly fitted” -> I don’t think that the C-grid has been “fitted” to anything.
Gone with the re-writting.

L505 This approach prevents the numerical schemes at play in the rheology

What “numerical schemes”? A numerical scheme is designed by a person (or soon by AI), but is not “at play in the rheology”. The paper does not present any stability analysis and or show development of noise due to numerical instability, grid staggering or whatever, so all you can say that with your scheme you are able to suppress any noise that may appear.
Yes, you are right. This sentence is no longer present.

L512 deformations -> deformation statistics
Corrected. (L655)

L513 “with respect to the viscous-plastic rheology”, “the aEVP simulation with admittedly not properly tuned parameters (reported checkboard noise)”.
The part about the checkerboard noise in the aEVP of SI3 has been removed. We use the exact default and recommended setup of SI3, which is the result of years of collaboration between the NEMO development team and the Sea-ice Working Group of NEMO. As such, we cannot do better than moderating our conclusions, and making it clear in the text that our reference simulation is the workhorse setup as provided by NEMO and not the most “appropriately-tuned aEVP”. This is also why we have changed the name from “SI3-aEVP” to “SI3-default”.
This particular sentence has been re-written, and aEVP is not mentioned anymore:
“Based on a comparison with satellite observations, this analysis demonstrates that the use of the newly implemented BBM rheology results in simulated sea-ice deformation statistics that are realistic.” (L654)

The sentence should include, that this happens on the same grid, with the same grid spacing etc.

Not relevant anymore in the re-writtent sentence.

The aEVP solution also has LKFs, just very few.
Yes, and one can also see them well in the figure showing the Mehlmann test-case that has been added to the paper (new Figure 5).

page 20

L550 “The average of the four surrounding points is used” What is done near the boundaries?
We have added the following sentence right after eq. A1:
“Note: surrounding points located on land or open-boundary cells are excluded from the averaging.” (L688)

A1 I would avoid “if \phi @F (@T)”, but instead use superscripts as in A2
We have done so, but we still need to keep a “(if \phi defined @X)” after each equation.

page 22

A5 Table of symbols related to the numerical implementation

The text promise values for the parameters but there are none. Remove [] around units
Parameter values are now provided in Table 1 & 2. “[]” around units have been removed everywhere in the text.

A5: e1t, etc. why not use \delta x ^{U} for this (or similar). This looks like it’s a Fortran variable from SI3. I think this can appear in the paper, but not in the list of symbols, and I would refrain from using the variable names in equations and text (as done in B2), because for a non-NEMO user/developer it makes the expressions impossible to read and check.
We agree, all the et1, etc. are now renamed as you suggest.

page 23

l565: I think that equations, here “(Eq. B5, B6)”, should be introduced before they are referenced. Now I have to read appendix B2, etc before Appendix B1.
We don’t agree. B1, the algorithm block, serves as a summary of what comes next, in a more detailed way, in B2. So we think it should stay this way.

page 24

B1 (@T), etc not really necessary, as the notation has defined this already

We agree, they have all been removed.

page 25

B7: it should be pointed out in the text that \bar{h} is used and not \hat{h}

Yes, we have added a sentence about this, same for the two previous equations (B5,B6) that use \bar{A}. (L748 & L752).

page 26

B12: what is “N”?

N is the upper limit for compressive stress, it is declared in the parameter table of Appendix A. It is also mentioned right after Eq. 13.

Eq B13, t_d is not a free parameter?

The damage time scale, t_d, is not a free parameter, neither in Ólason et al. (2022) nor in Dansereau et al. (2016), where it is first introduced. In Dansereau et al. (2016), it depends on the grid spacing and the speed of propagation of elastic waves - which is a free parameter. Ólason et al. (2022) calculate this propagation speed from the damaged elasticity, which is locally influenced by the damage, as well as using the local grid size. We follow the approach of Ólason et al. (2022) here.

elsewhere: (B14) is implied by B13 when d_crit=1

Yes. But d_crit can be > 1 (when the current stress state is within the Mohr-Coulomb yield envelope), and <0 under weak shearing and compression (see Eq.13).

page 28

L645 "The value of the BBM-specific parameters ..."

Unfortunately they are not listed in the table

They are now listed in Table 2.

L660 What are the criteria for “reasonably well shaped”?

We are now more specific, this part has been rewritten as follows:
“Triangles with an area smaller than 25%, or larger than 75% of the nominal area of the quadrangles to be constructed (i.e. the square of the spatial scale under consideration), or with an angle below 5° or above 160°, are excluded. Neighboring pairs of remaining triangles are then merged into quadrangles in order to transform the triangular mesh into a quadrangular mesh.” (now in section 2, L397)

page 29

l666: "consistent with the spatial scale of interest," what does that mean here?

This sentence was indeed vague, we have rephrased it adding our criterion, it now reads:
“- the square root of the area of the quadrangle falls within a ±12.5% range of agreement with the horizontal scale under consideration”
(now in section 2, L405)

L667: “almost identical”? Aren’t they within a 3day bin anyway? What does this mean in addition?

Yes they are within a 3-day bin. But each RGPS buoy position has its own time-position, with an accuracy at the level of the second. And these individual time positions can slightly differ between neighboring buoys, in particular when considering large quadrangles, typically when dealing with scales from 40 km and above. We tolerate a time gap of 60 seconds between the vertices of an aspiring quadrangle, hence the initial use of “almost”. We have rewritten this sentence that was indeed confusing in a more accurate way:
“the time position of each of the four points defining the vertices of the quadrangle should not differ from that of any of the other three points by more than 60s”
(now in section 2, L408)

L678 To prevent computational waste

I don’t think that’s an appropriate term here. “To save computer time/resources” … or similar

We now use your suggestion with “resources”. (now in section 2, L419)

L681 “feeds on”

I don’t think that an algorithm can “feed on” something (animals do)

After consulting fellow scientists whose mother tongue is English, it turns out that it is completely acceptable to write so in a scientific paper. We have changed “algorithm” to “software” though. (L422)

L685 please check the language of the author contributions for grammar, use of vocabulary ...
This part has been re-written properly. (L796)

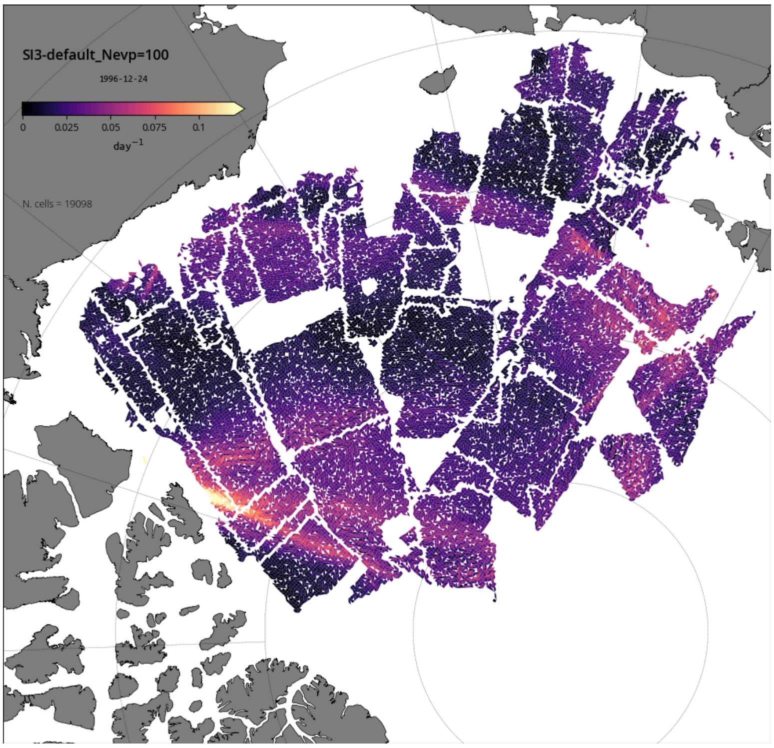

Fig.1/ Figure 7.c of manuscript, total deformation for experiment SI3-default, (N_{EVP}=100).

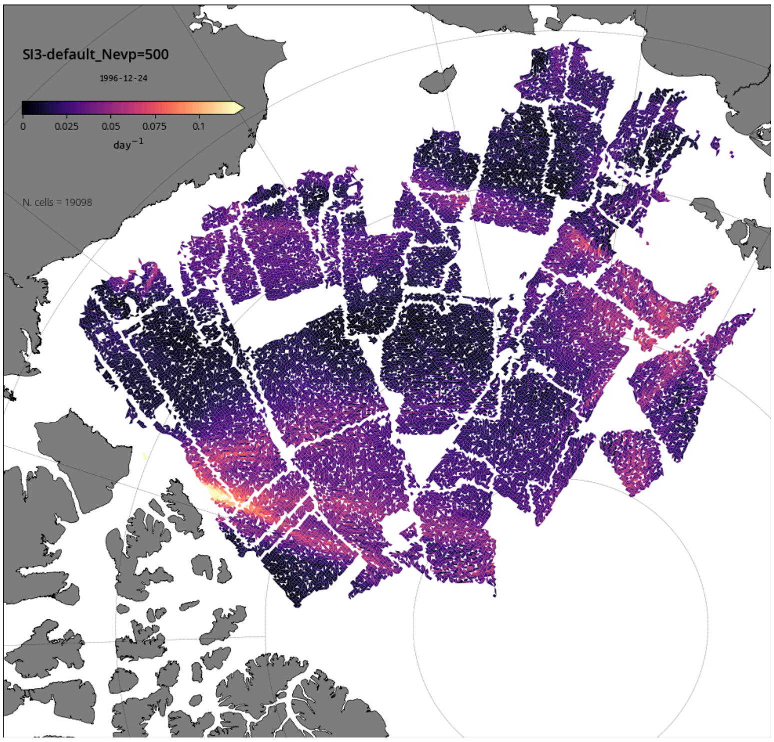

Fig.2/ Same as Figure 7.c of manuscript, but for SI3-default using N_{EVP}=500 instead of N_{EVP}=100