Manuscript **TC-2019-219-RC1**

Parameter Optimization in Sea Ice Models with Elastic-Viscoplastic Rheology

by Panteleev et al.

Response to Reviewer 1 (2nd revision)

We took into account the comments provided by the Reviewer 1 and significantly modified our manuscript. In particular, two additional OSSEs (small $\mathbf{k_T}$ and weaker wind) are conducted and discussed. Below are detailed replies for reviewer suggestions and comments:

MAJOR COMMENTS:

Reviewer: *The motivation of the experiment choices is not clear (and when I said "somewhat random collection" I didn't mean the principle twin-experiment setups as the reply of the authors implies, but the choice of parameters, configurations etc):*

The most important term in the sea ice momentum equations is the stress divergence term which includes the rheology formulation. Finding appropriate parameters ***P**** *and* ***e*** *is a fundamental problem in sea ice modelling and has been the subject of many previous papers (many of which are now cited, others are missing, which could have been used for further motivation, btw I am sorry about the van Leuven comment, I didn't realise that this was never published) and future papers will follow. It is easy to motivate including these parameters into the control vector and the authors do so, but only after motivating extensively the inclusion of land-fast ice parameters.* ***P**** *and* ***e*** *are important and relevant for anyone working with sea ice models and should come first, in the motivation, in the presentation of the experiments, in the (basically absent) discussion, and in the conclusions.*

Reply: We may agree, that ***P**** and ***e*** are the most important parameters for sea ice modeling, but our major objectives are prioritized by NRL goals, which currently are:

a) Improve the sea ice forecast in the Land Fast areas.
b) Create new data assimilation approaches and improve the sea ice hindcast/forecast.

These goals establish *our priorities*, so we first consider the optimization of the landfast sea ice parameter. As we understand, potential readers who are interested in *fundamental* problems may always skip sections related to the optimization of k_2 and k_T and directly proceed to the section related to *P** and *e*. We would also like to note that there is a substantial difference between sea ice modeling and sea ice data assimilation**.** In particular, sea ice numerical modeling has the goal of developing an accurate model which simulates ice conditions for a given period under the assumption that model parameters are already known. Sea ice DA has the goal of reconstructing poorly-known model parameters from observations with the ultimate goal of improving the short/medium range forecast of a given model. As we wrote in lines 19, 72-75 (previous version of the manuscript) our goal is to find an appropriate way for the short-range forecast. In the current version of the manuscript, we also described our priorities for the optimization of the land fast ice parameters. Our results indicate that 4Dvar LFI optimization can be achieved in 4-15 iterations. This can be easily realized for 4Dvar nested sea ice data assimilation models or even for Pan-Arctic high resolution sea ice models. The optimization of *P** and *e* requires more iterations (~50) and will require significantly more computational resources.

Reviewer: *The discussion about land-fast ice parameterizations is relatively new (I am aware of some work in the early 2000, and then Rozman/Itkin around 2014/2015, and the cited papers by König-Beatty, Lemieux and the not cited Einar Olason, who was the first to do a realistic simulation of land fast ice following König-Beatty's idea of added tensile stress, the k_T parameter, to a Mohr-Coulumbic yield curve, he, btw did not need any tensile stress for the elliptical yield curve to get fast ice in the Kara Sea), and to my knowledge CICE and the sea ice model of the MITgcm are the only sea-models with a larger user community to use the parameters. In this sense, it is totally unclear to me, why this parameterization receives so much attention in this paper, even before the more central rheology parameters P* and e.*

Reply: We tried to include all (known to us) references related to the Land Fast ice parametrization, but, our manuscript is not about the developing a new approach on how to model land fast ice. We describe and validate how to optimize the spatially varying land fast ice parameters in the framework proposed by König-Beatty, Holland 2010 and Lemieaux et al., (2015, 2016) and the elliptical yield curve, so we do not understand why we should discuss the option of modeling Kara-Sea LFI in the framework of the Mohr-Coulomb yield curve? This has little (if any) relevance to the subject of the manuscript.
We would also like to note that *a)* the UAF ROMS-sea ice model does utilize the LFI parametrization proposed by Lemieux et al. (2015), and now they are working on including the LFI parametrization into the MOM6-sea ice model; *b)* your comments on various options to model LFI underline the difference between numerical modeling and data assimilation. Numerical modeling replies to the question: can we model "Kara ice" this way? For DA: "how can we improve the model state and/or forecast using a given model?"

Reviewer: *Further, while the tensile strength parameter k_T is part of the "rheology parameters", the other parameters k_1 and k_2 of Lemieux's scheme are definitely not "rheology", but parameters of an "ad-hoc" (and very effective) parameterisation of grounding. For comparison, I haven't heard anyone calling the wind stress or ocean stress term, which have the same form as the basal/bottomstress term for grounding, part of the rheology.*

Reply: This is a question of the terminology. In our understanding, the term "rheology" describes the behavior of non-Newtonian fluids or solid materials. We doubt that sea ice is completely equivalent to a non-Newtonian fluid, at least at small scales, but, nevertheless the term "rheology" is used for sea ice. To address your comment, we provide an explanation that $k_{1,2}$ are not really rheological parameters (lines 48-49), but for the convenience of presentation, we refer to all control parameters as rheological parameters or (RPs).

Reviewer: *The generation of the TLA and regularization scheme is now described in the appendix, but the description is in an incomprehensible and unacceptable form. Instead of throwing most/all of the available information at the reader it would have been useful to sketch the derivation and the form of the TLA equations in a compact form. Then it would have also been possible to show and discuss the additional Newtonian damping term, which is now hidden somewhere in eq. A18-20 (I guess) with no discussion of the coefficient values etc. I would have tried to do this in a compact form, maybe more explicitly for a 1D example. In fact, if this were my paper, I would try to find a compact form with symbolic equations including the regularisation for the text in section 2 (the reason for this being the regularization term, not the TLA derivation, which one could in principle find elsewhere), and then maybe have everything spelled out in detail in the appendix. There's still text in section 2 that's unclear (i.e. the transpose of the sparse model matrix of the TL operator implies that this matrix is explicitly formed, which I don't believe is what the authors meant) and that is not supported by the* *information in the appendix. The description of the adjoint of the Lax-Wendroff scheme is confusing, because it suddenly involves non-linear terms that are not introduced in A9 and 10. Again, I don't think that there's anything wrong with the presented material, but the presentation does not help much to understand the derivation of the model.*

Reply: The derivation of the TL and ADJ models is actually based on the differential of the Lagrangian function. So, given an explicit formulation of the finite difference forward model, the TL and ADJ models are defined *automatically* as soon as you specify the set of control variables and the cost function. The math explaining the adjoint is calculus 3 level (see, for example, https://tutorial.math.lamar.edu/classes/calciii/lagrangemultipliers.aspx). Actually, it is because of this "formal differentiation" property, that the TL and ADJ can be obtained by TAMC or other TL/adjoint compilers. Since 1970, there have been numerous publications on the subject. In a symbolic/ compact form, the concept of adjoint is best described in the publication by Le Dimet and Talagrand (1986), and in a more popular form by Errico (1997). So we refer a potential reader to these and others publications. We assume that the finite-difference form of the TL and ADJ equations, given in Appendix A, would be useful to a potential reader who has no hands-on experience in its derivation from the parent numerical model. To address all of these issues, we moved the finite-difference description of the model from Appendix A to Section 2.1, and, following the Reviewer's request, provided a short description of the assimilation method in symbolic form in Section 2.2.1. In this exposition, we tried to follow the notation used by Ungermann et al. (2017) in the explanation of the data assimilation approach using the Green function method. Taking into account that other reviewers already approved our manuscript and did not object to the detailed presentation of the TLA (Appendix A), we do not feel comfortable in removing it completely.

Reviewer: *The model in this work is an EVP model that does not have much to do with CICE, except for the EVP scheme and the B-grid. The default CICE strength parameteriZation is different from what is discussed here (and yes, CICE does have the option of using the P* parameterization, but I have not seen many papers explicitly doing so except for the ECCO-group. Also, the "Cf" parameter in the Rothrock 1975 formulation is has a similar scaling function as P*, but that needs to be argued for in the text, if you want to relate to CICE*

Reply: We mentioned the CICE6 model solely because all CICE applications at NRL use P^* parametrization. We do not think that an overview of all CICE parameterization options is a necessary requirement in better understanding the manuscript. To avoid possible confusion, in the revised version of the manuscript, we tried to minimize discussion relevant to the "similarity" between our simple model and CICE6.

Reviewer: *there is not nearly as much complexity here as there is in CICE in all aspects (the advection scheme in CICE is totally different from the Lax-Wendroff scheme used here.*

Relating to CICE so often makes very little sense and should be dropped in most places. The presentation still requires work, mainly: (1) giving appropriate weight to different parts of the manuscript, (2) proper description of TLA generation and regularisation in a comprehensible compact form that has enough details so that it can be reproduced

Reply: We commented on the use of CICE (lines 79-80) and removed excessive mentioning of the model from the text. The TLA generation and step by step description of the 4Dvar concept in symbolic terms are now given in the modified Section 2.2.

MINOR COMMENTS:

Reviewer: *"Taking into account that sea ice observations are available daily, the experiments are configured for a 3-day data assimilation window in a rectangular basin" Unclear, why this is connected. I would think a 1 day window would be the natural choice if there's new data every data. Also Cryosat2 data is not available in gridded form every day. Needs a better explanation.*
Reply: The sentence was modified. See lines 7-8 of the current version of the manuscript.

Reviewer: *line 21: (e.g. Heimbach et al, 2010; Zhang and Rothrock, 2003; Vancoppenolle et al. 2009; Massonnet et al. 2015) not clear if you refer to model or to models with DA capacities. Heimbach et al describes an adjoint sea ice model simulation (but not DA), Zhang+Rothrock, and Vancoppenolle describe a forward model, Massonnet et al describe solutions of models with some sort of DA or state estimation.*
Reply: We do not see any controversy in our text. See lines 20-25 of the manuscript after the first revision. From our point of view, it is obvious that on lines 21-22 we refer to both models and models with DA capabilities. To further clarify the point, we changed "and" to "and/or" on line 20 (current version).

Reviewer: *l39: P* and α are not part of a typical (default) CICE simulation. See major comment.*
Reply: The sentence is corrected and *P*/e* option is directly linked to the NRL applications of the CICE5/6 model (see lines 39-43, current version)

Reviewer: *l45: (RPs) unnecessary abbreviation? I guess the authors have a different opinion.*
Reply: We agree with the Reviewer: we have a different opinion and prefer to use this abbreviation. A clarifying statement was made in the Introduction (lines 47-49.

Reviewer: *l82: Note, that optimization of the RPs through the 4Dvar DA approach allows us to efficiently use all available sea ice observations including sea ice velocity, that are rarely used for assimilation in sea ice DA systems. The latter is due to weak sensitivity of the sea ice state with respect of the ice velocity (e.g. Kauker, et al., 2009). Roughly speaking, the 4dVar DA approach allows us to use sea ice velocity observations for adjustment of the RPs and/or atmospheric forcing in an appropriate manner resulting in a better sea ice forecast (Stroh et al, 2019). Not clear why the "weak sensitivity" can be brushed aside for this approach, which is essentially the same as that of Kauker et al 2009.*
Reply: We did not brush aside the *"weak sensitivity"*, we wanted to state that attempts to use velocities are rare because of the weak sensitivity of the model state with respect to the initial velocity conditions (Kauker et al., 2009). Augmenting the 4dVar control vector with RPs allows us to use sea ice velocity observations more efficiently. We modified the paragraph to improve the clarity of presentation (lines 89-91).

Reviewer: *l89: I find OSSE a bit of an overstatement for the type of idealized experiments that are presented here. I would use this term OSSE only for realistic applications when the design of the observing systems is the subject, for example, resolution or accuracy requirements for a new satellite system to be designed, etc. This work is not about observations, but about TLA model development and testing.*
Reply: Indeed, the major goal of our manuscript is to investigate options on how to improve the short range sea ice forecast and we do that using the OSSEs with a simplified model. We consider the development of the TL and ADJ models is a technical goal, which formally can be resolved through automatic compilers in future. Note, that the proposed stabilization of the TLA models through the Newtonian dumping term is only a minor part of the study. We consider that

these idealized experiments are necessary because development of the TLA for the CICE is not easy and proper understanding of the problems/possibilities/advantages and disadvantages is a necessary step that should be made before proceeding in the development of the TL/ADJ code for CICE which may (or may not) include the respective RP controls. We do not see a principal difference between "OSSE" and the currently obsolete term "twin data experiment". Indeed, 10-15 years ago, we would be happy to use the term "twin data experiment". During the last decades, however, the term OSSEs became well established and, in our opinion, completely replaced the term "twin-data experiment". Note also, that the early OSSE references to "OSSEs" were made by Nitta (1975), and Arnold and Dey (1986) in studies which use simplified models. We have added these two references to the manuscript to clarify the issue of "inadequate terminology".

Reviewer: *ll96: Currently, satellite sea ice observations are typically available daily with a reasonably dense spatial resolution. Analysis of the SAR images (e.g. Panteleev et al., 2019) indicates that in the marginal sea ice zone, the pancake/cake ice with floe sizes of ~1-20 m may be easily replaced by floes exceeding 1 km in size in one week. As a consequence, we configured the OSSEs with a 3-day DA window assuming that such approach should have more impact on short therm sea ice forecast. Unclear reasoning.*
Reply: From our point of view, sea ice with floes 1-20 m and 1-2 km should have a different P^* and e. In the OSSEs, we assume that P^* and e are time invariant, i.e. P^* always has a maximum in the eastern and western parts of domain. But, indeed, P^* and e are the property of the sea ice, which moves. A 3-day DA window is reasonably "short enough" to assume that ice does not move "very far" from its original position. The DA window also cannot be very long for computational cost considerations. To prove this point, we conducted experiments with a 5-day long DA window. Results were similar but took almost twice as much computational time. We modified the sentence at lines 105-106. and added some discussion in Section 5.

***Reviewer*:** *I agree that Eq(6) is correct, but it's unfortunate, because it hides the form of Δ^2 as the sum of ice divergence + ice shear/e^2*
Reply: We do not think that this form is "unfortunate" as it emphasizes the invariance of Δ wrt to coordinate transformations.

Reviewer: *ll145: the explicit advective time step -> an explicit time stepping scheme why not say "a Lax-Wendroff time stepping scheme" here, as you don't say more in the appendix, either. Linear or non-linear?*
Reply: We moved the entire description of the finite difference scheme to Section 2.1. We used simplified Lax-Wendroff in order to simplify TL and adjoint code derivations. The applied isotropic diffusion produces a certain loss in accuracy in the approximation of advection, which has negligible impact on the results of the manuscript. Additional discussion of the advection scheme can be found at lines 569-570.

Reviewer: *l150: "The adjoint code was obtained by transposition of the sparse matrix in the code simulating the action of the TL operator on a perturbed state vector". In spite of the large amount of information in the appendix, this is still not clear. Is this matrix ever formed explicitly so that you can transpose it? Or is this done only symbolically by re-ordering the operations as in "automatic differentiation"?*
Reply: The matrix was neither formed explicitly, nor "symbolically obtained by re-ordering the operations". The code for computing its action on a vector was derived by taking the derivatives of the cost function (21) wrt to the state variables. We added a new Section 2.2.1 and provided more details in Section 2.2.2 (lines 234-236).

Reviewer: *Tab1: turning angle 0.4343 rad = 24.88 deg. Why such a number?*
Reply: The misprint is corrected to 0.436332 (25º). But we do not think this is essential from a scientific point of view.

Reviewer: *the proper way would be something like this: A similar instability of the TL EVP solver has been observed in the MITgcm sea ice model (M. Losch, personal communication).*
Reply: Corrected. Lines 251-252.

Reviewer: *l170: (e.g., Yaremchuk et al, (2009));*
Reply: Corrected.

Reviewer:
l180: more simple-> simpler? or just simple?
Reply: This part of the text was rewritten.

Reviewer: *l180: time scale T_d: is this the same as in eq(2)? That doesn't make sense*
Reply: Misprint corrected and more details describing regularization terms are added (lines 267-269).

Reviewer: *l182: a Newtonian*
Reply: This part of the text was rewritten. Lines 265-272.

Reviewer: *l200: MIT the model is called "MITgcm", not MIT, please correct everywhere*
Reply: Corrected.

Reviewer: *l216: missing parentheses around https://icesat-2.gsfc.nasa.gov/*
Reply: Corrected. Line 305.

Reviewer : *l219: A similar error level was …*
Reply: Corrected. Line 308.

Reviewer: *l225: "with spatial decorrelation scale of 150 km and temporal decorrelation scale of 7 days" two missing articles?*
Reply: The sentence was rephrased to improve clarity (lines 314-316 of the revised manuscript)

Reviewer: *l228: where do these numbers come from?*
Reply: These numbers were removed to prevent distraction and improve the conciseness.

Reviewer: *I maintain that the systematics between the first group and there rest remain unclear. These are two very different experiments and it's not clear why the LF scheme of Lemieux et al 2016 receives similar attention as the optimization of the more universal parameters P* and e*
Reply: See our comment above. Attention to optimizing the parameterization of the LFI scheme is caused by current priorities of NRL research. Potential readers may always skip sections related to the LFI and proceed to *P** and *e* experiments without any loss in understanding the material.

Reviewer: *ll247: The maximum number of control variables associated with the initial conditions (the number of ice model grid points occupied by the SIT, SIC and SIV fields) was*

about 9000. The RP control fields were defined on coarser (δxp=15 (or 7)δx) grids with bilinear interpolation on the model grid of the respective OSSEs. Thus, the maximum dimension of the RP control vector never exceeded 36 elements. This remains very unclear.
Reply: The paragraph was rewritten (see lines 337-340). Additional details are now given in the new subsection 2.1.2 (lines 193-194)

Reviewer: *261: k_T was set to 0.6. I maintain that this is an unusually large value, except for explicit land fast ice simulations. You wouldn't use that universally in a Pan-Arctic simulations. (see Olason 2016, A dynamical model of Kara Sea land-fast ice, JGR, doi:10.1002/2016JC011638, where k_T=0 with the elliptical yield curve in spite of the focus on land fast ice). Tremblay and Hakakian (2006) report upper and lower bounds for compressive and tensile stress, I am not sure if this can be used to infer k_T=0.6. In König-Beatty and Holland, the noisy EVP solver prohibited land fast ice for smaller values of k_T, Lemieux et al (2016) discuss the value of k_T and obtain better agreement with observations with small values of k_T = O(0.1).*
Reply: You are correct, Tremblay and Hakakian (2006) reported estimates of maximum and medium tensile stress. However, as we understand, König-Beatty and Holland (2010), used Tremblay and Hakakian (2006) results and derived estimates k_T~0.5-0.6 (page 188 from König-Beatty and Holland, 2010). In order to resolve this discussion, we conducted additional experiments with k_T =0.2. See the new Figure 2 and discussion on lines 373-384 of the current version of the manuscript.

Reviewer: *l270: "due to the absence of tensile strength in ice (kT =0)" and due to the non-converged EVP solution, see König-Beatty and Holland*
Reply: We do not agree. We conducted the experiment with a much larger number (~2000) of sub-cycling iterations and got the same solution with ice moving eastward in the entire region. So, we attribute the absence of the LFI solely to the condition k_T=0.

Reviewer: *Caption of Fig2 does not refer to (e) and and (f) explicitly (but to a-d)*
Reply: Corrected.

Reviewer: *l284: 3m -> 3 m? In fact, the use of a space before a unit is inconsistent throughout the text*
Reply: Corrected.

Reviewer: *l295: and SIV that's not different than the "true solution", where "initial velocities [...] were set to zero" (l288)*
Reply:
Corrected. Lines 397-398.

Reviewer: *paragraph starting at line 319: Rather than using the numerical difficulties associated with the Heaviside function to waive the optimization of k1, one could use this analysis to argue for a smooth parameteri**z**ation, that would also be more physical, because it is very unlikely that in a grid cell of a finite (usually large) extend of order (km) all ice ground at the same time and instantaneously (like in a cloud scheme not the entire grid cell of a AGCM will suddenly go through a phase transition). Instead a smooth transition is more appropriate as we expect the physics of sea ice to be smooth. I can see that this is beyond the scope of the paper, but I also think that the scope of the paper is fuzzy with these two types of experiments. It's not clear why the optimization of well established and very important parameters P* and "e" is juxtaposed with two parameters of a relatively new parameterisations that are hardly used.*

Reply: As we outlined above, the structure of the paper reflects our (and NRL's) current research interests and goals. In the future, we will definitely address the problem of k_1 optimization. Also, it is not clear what you mean by "a smooth transition is more appropriate as we expect the physics of sea ice to be smooth." The LFI regions (with characteristic scales of 30-100 km) and polynyas (~10-20km), near the coastline, are typical phenomena observed in NRL's pan-Arctic CICE model runs at a 2-km resolution, that are characterized by discontinuities in ice thickness and concentration fields (i.e. cannot be considered as "smooth" in our opinion).

Reviewer: *Section 4 "eccentricity" is not what "e" is. It's has been defined correctly before.*
Reply: Corrected.

Reviewer: *l335: had the form of a Gaussian-shaped cyclone.*
Reply: Corrected.

Reviewer: *ll339: the trace of the stress tensor $P_{tr} = -tr\sigma/2$ this is commonly called σ_I, the first invariant of the stress tensor, or divergence of stress. Not sure why this needs a new name here.*
Reply: We thought that the subscript "I" could be mixed with "1". An appropriate correction was made (lines 445)

Reviewer: *l348 "as most of the currently available observations". Rephrase so that it says what you mean (bias free observations are a common assumption in DA systems)*
Reply: Corrected. Lines 449-450.

Reviewer: *l364: definition of S_u: why did you use this norm and not the more common RMSE? Or is this implied? Unclear*.
Reply: We utilize different norms for scalars and vectors. Since only relative changes of these quantities are important, we do not see a problem here.

Reviewer: *l368: std($h_{opt} - h_{true}$). In what sense is this form different from the definition of S_u? If it is different, why would you use different forms?*
Reply: We used STD for scalar values and 1-norm for the velocities (see previous comment).

Reviewer: *ll369: I find it much more plausible, that the impact on velocities is largest by optimizing rheology parameters that directly affect the momentum equations. The effect on the thickness field will take longer assimilation windows because the changed dynamics needs time to advect thickness (and concentration). [And later you say so, why not here?]*
Reply: We agree with the Reviewer: on lines 370-371 of the previous version of the manuscript, we actually made a similar statement: "Another possible reason is that $\mathbf{C}_{rh}$ is not well controlled by the initial SIT distribution at the time scale of the relatively short 3-day assimilation window." Note that under the simplifying assumptions of Kohl and Stammer (2004), (see also Panteleev et al., 2013) sensitivity is proportional to the accuracy of SIT observations and, therefore, will be large if the respective error variance is small enough.

Reviewer: *l380: "that spatial locations of extrema agrees well with true distribution" fix grammar*
Reply: Corrected.

Reviewer: *ll387: "that inaccurate position of the atmospheric cyclone" there's an article missing*

Reply: Corrected.

Reviewer: *l389: "(0.1 m/s)" the extra space after "(" happens very often, but not always. I am not sure if this is intentional or not, but I guess, the space should be removed here and everywhere else.*
Reply: Corrected.

Reviewer: *l396: the wind appears to dominate the optimization. Without proper wind, the optimization appears meaningless pointing to severe problems in the sea ice model parameterizations (why should they depend on the wind?). It thinks that's worth pointing out, maybe in the discussion.*
Reply: We agree that wind is a dominant forcing in sea ice dynamics and, therefore, it is a major factor in both sea ice modeling and data assimilation. You may have perfect P^* and e but your sea ice model solution will be incorrect if it is forced by incorrect wind. Note, however, that without proper wind, any sea ice model solution becomes meaningless, but the DA solution may still be reasonable because wind forcing can be adjusted to the correct values through the assimilation of the SIV/SIC/SIT observations if they are accurate enough (e.g. Stroh et al., 2019). We added a short discussion on the importance of accurate wind forcing for obtaining reasonable results. See lines 640-644 of the revised manuscript.

Reviewer: *l400: to a less degree than e. Fix grammar*
Reply: Corrected.

Reviewer: *ll403: "To mimic these conditions, we conducted another OSSE with spatially and temporally invariant sea ice concentration A=1. Numerically this was achieved by removing the advection equation from eq. (4), and removing initial A0 from the control vector C0."That changes the model and makes it difficult to understand the generality of the results. Why not construct an experiment where the ice strength is strong enough (relative to the wind) so that the system does not move? That would be more "realistic".*
Reply: Actually, we do not see the problem here. A=1 may be interpreted as an additional very strong cooling, which removes areas where A<1. We actually tried different winds and always had some regions with A<0.9 somewhere. Your suggestion with the experiment with zero ice velocities is meaningless, because: a) If ice does not move from the wind, P^* can take any value above some critical point which is actually defined by the wind amplitude; b) As sea ice data shows, ice is always in motion in the central Arctic. An extra sentence was added (line 511-512)

Reviewer: *l413 "about 0.5 N/m2" that's a lot!*
Reply: We agree. This is a rather strong wind corresponding to intense Arctic cyclones that are rather frequent phenomena in the Arctic Ocean. Note, this wind was applied only near the western and open boundaries. We also found a minor inaccuracy here: the wind actually increased from 16 m/s up to 20 m/s, which yields ~0.5 N/m^2 by the end of the data assimilation period. Thus *wind stress* increased 1.5 times, *not wind*. To resolve the issue with wind stress amplitude, we conducted an additional experiment with a twice as small wind (~10 m/s) and stress about 0.15 N/m^2. See new Figure 9 and lines 513-518, 548-559- related to the results of the new experiment.

Reviewer: *l415: "SIT increases over almost everywhere" this raises the question of volume conservation. Does the model conserve volume, and is this important for the optimization.*
Reply: As mentioned above, the condition A=1 can be physically interpreted as cooling followed by immediate sea ice production. So, formally, sea ice volume is not conserved. But we

do not understand how this is important for DA with a 3-day window. Theoretically, this problem can be important for very long DA windows, but again, this will be a problem with the model. Not the DA algorithm.

Reviewer: *l418: maximuma-> maxima (a spell checker would have found this)*
Reply: Corrected. Line 503.

Reviewer: *l431: "because sub-optimal Ptr distribution fails" missing article*
Reply: Corrected.

Reviewer: *l431: maximums → maxima (used previously)*
Reply: Corrected.

Reviewer: *l435 "std[P 2 (opt) − P (true)]" Ptr squared?*
Reply: Corrected.

Reviewer: *l437: in Figure 9e,f*
Reply: Corrected.

Reviewer: *ll439: "The effect could probably be attributed to the region with zero convergence along the western boundary where the rheology does not play a significant role." rephrase to be more specific.*
Reply: The term **"**zero convergence" has been changed to "divergence": when the ice field diverges, internal stress becomes less important. A comment has been added to lines 544-545.

Reviewer: *Section 5 Conclusions and discussion" promises the unusual order of conclusions at the beginning followed by a discussion (why not do it conventionally with discussion first, followed by conclusions?), but then starts with a long summary of the results where the actual conclusions are hidden in the details, and at the end there are two paragraphs of discussion. I suggest to rewrite this section for better clarity.*
Reply: We do not see a problem here. Actually, the title of the section is "Conclusion and Discussion", not vice versa. For example, the first conclusion is that the common way to stabilize TL is not efficient, however the Newtonian scheme allows us to do this, and after, we provide some discussion that the VP solver allows us to avoid this problem. We do not understand, how this can be outlined in the opposite order. Similarly, the "conclusion" is related to the optimization of Land Fast ice and P*/e parameters. Perhaps the title of this section is confusing, and so we have re-titled the section as "Summary and Discussion". We also include some discussion related to the additional experiments.

Reviewer: *l445 with respect to*
Reply: Corrected.

Reviewer: *l445: "all rheological parameters" k2 (and k1) is not a rheological parameter and "k_T" is not a parameter that is commonly used (i.e. different from zero)*
Reply: We agree that k_2 and k_T are not the rheological parameters, but from a physical point of view, sea ice rheology is also not "real rheology" which describes a Non-Newtonian liquid. See our explanation above.

Reviewer: *l448: "Lemieux et al., (2015, 2016) and Konig Beatty and Holland, (2010)" (and elsewhere) remove "," before "("*

Reply: Corrected.

Reviewer: *l455: "Analysis of the TL approximation accuracy has shown that Newtonian stabilization has errors similar to the ones observed in the case of diffusion-based stabilization, and thus the Newtonian scheme can be successfully used in sea ice models based on the EVP solvers." This analysis/discussion would be very important to understand in detail, because it is potentially something that readers may want to apply themselves. However, the "analysis" is extremely short, leaving out important aspects, like a stability analysis based on the value of the damping coefficients, etc. It would be interesting to understand, why this term successfully stabilizes the system etc*

Reply: We do agree that it would be interesting to analyze the effects of the proposed realization and we plan to work on this problem in future. Note, however, that the application of the conventional scale-selective Laplacian regularization was first published as an *ad hoc* method by Hoteit et al. (2005) (actually, it was applied earlier, but was not published). Some additional analysis was done by Hoteit et al. (2006), and the theoretical interpretation was done by Yaremchuk et al. (2015). So, it is typical when some technique is initially applied *ad hoc*, and after some time, it finds some theoretical background. We did provide some speculation why scale-selective regularization poorly works with ice (see lines 261-263 of the manuscript). From the viewpoint of applications, the Newtonian damping should be kept as small as possible and its value is usually defined by trial and error.

Reviewer: *l471: Since you have cited Losch et al (2014) already, they could be cited here again, because there AD tools were already used to compute the matrix times vector operation necessary for the matrix free JFNK with FGMRES in a sea ice model, however without significant improvement over simpler finite difference schemes.*

Reply: Corrected.

Reviewer: *l480: "kT and k2, responsible for grounding and arching phenomena" unless intended as a "chiasms", I would turn the order around so that the first symbol corresponds to the to the first explanation (k2: grounding, kT: arching)*

Reply: Corrected.

Reviewer: *ll491: "In the second group of OSSEs, we analyzed the possibility of reconstructing spatially varying sea ice strength P* and ellipse axes ratio e distributions." I would argue that this is the more important part of the paper and should come first.*

Reply: As explained above, optimizing k_2 and k_T was the top priority of the current research.

Reviewer: *l500: collocated in the regions of strong… collocated with regions of strong … I am also not sure if "collocate" is the correct word here. I would use "co-locate"*

Reply: Corrected.

Reviewer: *l502: provides a slightly more accurate reconstruction*

Reply: Corrected.

Reviewer: *l504: "Accurate forecasting of Ptr is very important because it better informs avoidance of regions with excessive compressive stress." Not the right context.*

Reply: Corrected. Moved to Summary and Conclusions section.

Reviewer: *ll523: This is a totally generic and even inaccurate statement (in spite of its generality) that should be removed or adequately modified: On the one hand, it is not clear to me, how the MOSAiC observations that are very local should help to constrain a local model,*

where open boundaries and boundary conditions as control parameters should be an important aspect that is not even mentioned in this work. On the other hand, the MOSAiC observations are regionally too confined to serve as a serious data source for a large scale model. I do agree that MOSAiC will be substantial to future sea ice parameter studies, but mostly for slow thermodynamic process and for fast stress parameters, because local drift and deformation as well as stress measurements are obtained.
Reply: The statement was removed at the Reviewer's request. However, from our point of view MOSAIC data can be helpful for *P*/e* analysis related to the 2-kilometer resolution in the NRL model.

Reviewer: *l541: the slightly different form of \epsilon vs. \varepsilon is unfortunate, because both are "epsilon" and easy to confuse this with strain rate tensor.*
Reply: We modified the notation and put the dot over all the quantities related to the components of the strain rate tensor. So now, they are easily distinguished from time scale ratios.

Reviewer: *Are Eq A7+8 solved implicitly? u^s and v^s appear in both equations on the left hand side*
Reply: Yes, they are. The respective matrices are block-diagonal with 2 x 2 cells and can be inverted explicitly.

Reviewer: *l550: the Lax-Wendorff scheme the scheme is called "Lax-Wendroff"!!*
Reply: Misspelling corrected.

Reviewer7: *l559: MIT -> MITgcm (Heimbach et al 2010)*
Reply: Corrected.

Reviewer: *TAMS -> TAMC*
Reply: Corrected.

Reviewer: *eq A18-24: I don't understand, why the time stepping algorithm has been used here to illustrate how the TLA works. It makes the entire presentation far more complicated and difficult to understand. The four of eq. A18-A24 is unacceptable. I don't think that any of the reviewers meant this, when asking for a better description of the TLA*
Reply: In your previous review, you requested of us to provide more details for potential reproducibility. From our understanding, reproducibility can be achieved either by accurate differentiation of the Fortran code as it is done in TAMC (or by hand) or by providing a relatively accurate description of the TLA models, as it is done in the Appendix. Because other reviewers considered the Appendix acceptable (had no objections against it), we do not feel comfortable in removing it. However, following your request, we added some schematic explanation in the new Section 2.2.1.

Reviewer: *eq. A18-20 shouldn't the damping term be highlighted somehow? I am not even sure if -eps delta sig1 is the term*
Reply: It is one of the terms. A more detailed explanation is now given in the new Section 2.1.2. The terms are now highlighted in the Appendix A of the revised manuscript.

Reviewer: *l584: "variational assimilation experiments we used the strong constraint state-space formulation of the problem, minimizing the cost function" doesn't the adjoint method start from the cost function defining the scalar product with respect to which the entire equations are*

derived? The presentation appears backwards through the eyes of the practitioner, but for understanding the principles of what has been done it's not very useful.
Reply: The basic principles of 4Dvar DA theory in geophysical applications has been extensively published since the 1970s. So, all details of the technique can be easily found. It takes significant space to explain everything and we do not feel that the format of the journal publication is the proper place for providing this information. For this reason, we outline only the basic principles and steps of the 4Dvar data assimilation in the revised Section 2, and provide additional references (Le Dimet and Talagrand, 1986 and Lewis and Derber, 1985).

Reviewer: *l601: "with the additional range constraints for the selected control fields (Section 2.3)" If you are using bounds, a better choice would have been an L-BFGS algorithm (as in m1qn3) with bounds, e.g. L-BFGS-B*
Reply: We do not agree. We are familiar with the L-BFGS routine and actually tried it. We found that M1QN3, with bound constraints, was computationally more efficient and converged faster in most of the experiments.

Reviewer: *l612: paralell -> parallel*
Reply: Corrected.

Reviewer: *Eq A27-31: the symbol $\mathcal{I}^T$ is not introduced. I have no idea what this is. It should be the "adjoint" model, i.e. the transpose of the tangent linear model matrix, which is never explicitly formed*
Reply: The superscript ^{T} is a standard notation for the transposition of a matrix. The respective matrix I is actually introduced on line 599 of the original manuscript, so I^T denotes the transpose of the matrix I which performs bilinear grid-to-grid interpolation from the (coarse) control grid (where the control fields and their gradients are defined) onto the (finer) model grid.

Reviewer: *l623: adjont -> adjoint*
Reply: Corrected.

Reviewer: *l624: "the Newtonian damping given by the terms −εδσ1,2,3 in eq. (A18-A20)" this information should've come much earlier*
Reply: We agree. Now, these terms are discussed in the modified Section 2 (subsection 2.2.2). See lines 267-269.

Reviewer A77: *l723: Lemieux,J.F., missing space*
Reply: Corrected.

References:

Köhl and Stammer, D., 2004. Optimal observations for variational data assimilation. J. Phys. Oceanogr. 34, 529-541

Panteleev et al, 2014 Configuring high frequency radar observations in the Southern Chukchi Sea, Polar Science, (7), 2, P77-81.

Ungermann, M., L. B. Tremblay, T. Martin, and M. Losch (2017), Impact of the ice strength formulation on the performance of a sea ice thickness distribution model in the Arctic, J. Geophys. Res. Oceans, 122, 2090–2107, doi:10.1002/2016JC012128.

Manuscript **TC-2019-219-RC1**

Parameter Optimization in Sea Ice Models with Elastic-Viscoplastic Rheology
by Panteleev et al

Response to Reviewer 2 (2nd revision)

We would like to thank the Reviewer for useful comments that helped us to improve the manuscript.

MINOR COMMENTS:

Reviewer:* *B1 *L18,19: AO is specifically used for 'Arctic Oscillation', I would suggest not using this abbreviation..*
Reply: Corrected.

Reviewer:* *B2 *L42: reference for Konig Beatty and Holland is missing*.
Reply: Corrected.

Reviewer:* *B3 *Please check again all the citiation form, especially 'et al.'. For example, L65, L66*
Reply: Corrected

Reviewer:* *B4 *Eq(1,3), please confirm if h is the effective thickness*
Reply: Corrected

Reviewer: B5 *L200, using MITgcm instead of MIT*
Reply: Corrected

Reviewer: B6 *Eq (A25), extra plus '+'*
Reply: Corrected.