The authors would like to thank Dr. Nils Hutter (the referee) for the invaluable comments. The following are the replies for each comment, together with specific revisions that are made. The original comments are in *green italic* font, and the revisions are highlighted in the revised manuscript in yellow.

Reply to comments of Referee #1:

The paper describes the creation of a multi-resolution suite of grids for CESM with a focus on their use for sea-ice modelling in the Arctic. The authors study the effect of grid resolution and number of EVP subcycling steps on the statististical properties of sea ice deformation as well as sea ice extent and volume. In particular, the localization of sea ice deformation in shear and failure lines such as leads and pressure ridges is studied. The authors present their model configuration as a starting point for more dedicated studies on sea ice dynamics and climate simulations and share the corresponding code and data. The simulations are analysed without optimising model parameters to the specific grid resolution and the evaluation of the simulations need to be improved. Therefore, I recommend the manuscript for publication in Geoscientific Model Development after consideration of my general and specific comments.

General comments:

1) The authors present untuned model runs with biased ice thickness fields (with too thick ice in the Beaufort Gyre) and also the ice volume in the simulations differs with the resolution used. However, the authors describe good agreement of sea ice coverage and volume of all simulations although a comparison with a sea ice thickness product, e.g. PIOMAS, is missing. So first, I suggest a thorough evaluation of the ice thickness and to study the differences in sea ice state between the different resolution simulations. I see two potential ways how to handle these different resolutions simulations that produce different sea ice state:

Reply to the general comment (1): the authors thank the referee's comment on the modeled sea ice thickness, and would like to make the following reply. Based on extra numerical experiments, we have discovered sensitivity of modeled ice thickness to the strength parameterization scheme. By default, CESM utilizes an ice strength parameterization in Rothrock (1979), detailed in Lipscomb et al. (2007). Instead of the traditional scheme in Hibler (1975) in which ice strength is related to mean ice thickness, the ice strength is closely related to the energy conversion and dissipation during the ridging process. Fig. 1 (below) compares the equilibrium Arctic sea ice thickness with the two ice strength schemes for TS045 (CESM D-type). With the scheme of Hibler (1975), the sea ice is considerable thinner in the Beaufort Gyre (BG), and arguable more reasonable with respect to observed sea ice climatology. We further confirm that thinner ice in BG is independent of the specific grid we use, showing similar differences for the default built-in grid of GX1V6 in CESM (Fig. 2). We have also found similar results with TS015 (CESM D-type experiment) as well as CESM G-type experiment (ocean-ice coupled run) with TS045 (results not shown).
Since Rothrock (1979) and Lipscomb et al. (2007) are used by CESM, including its scientifically validated experiments, we consider this choice of ice strength parameterization reasonable for the experiments and analysis within the realm of this study.
Furthermore, in a recent paper (Stewart et al., 2020), ocean-ice coupled experiments with both CORE2 NYF and normal-year based forcings from JRA55-do are compared. Fig. 18 of

the reference confirms our findings above, showing thick ice in BG for CESM (version 2), as well as drastically different ice thickness fields modeled when different years of JRA55-do are used for normal-year forcings. This indicates that NYF based experiments that produce reasonable sea ice distribution can serve the study of certain aspects of model performance, but not necessary representation of the mean status of sea ice in reality.
In summary, we consider the model output attained with CORE2 NYF in our experiments reasonable for the comparative study across the resolutions. Furthermore, we look forward to exploring the attribution of strength scheme dependent sea ice thickness in the future.

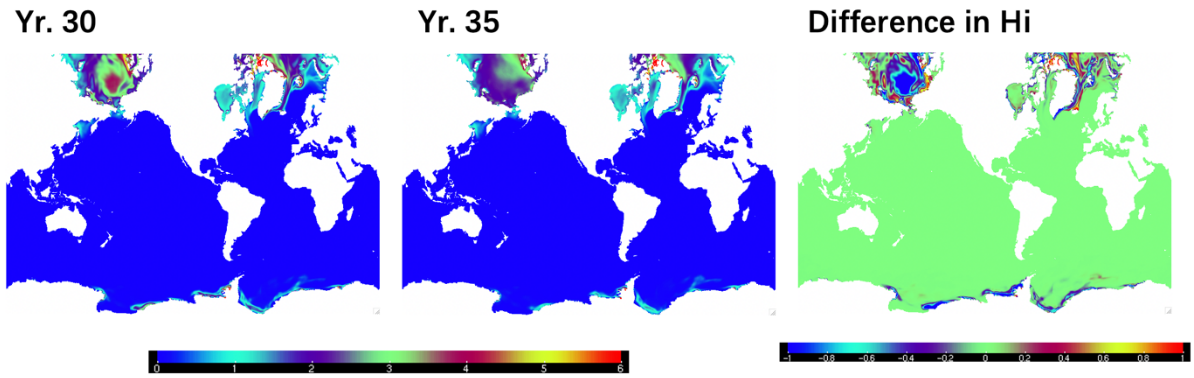

Fig. 1. March sea ice thickness difference for TS045 grid at equilibrium states under CESM D-type experiments. Left: Rothrock (1975) scheme; Middle: Hibler (1979) scheme; Right: difference (middle *minus* left). All values are in meters.

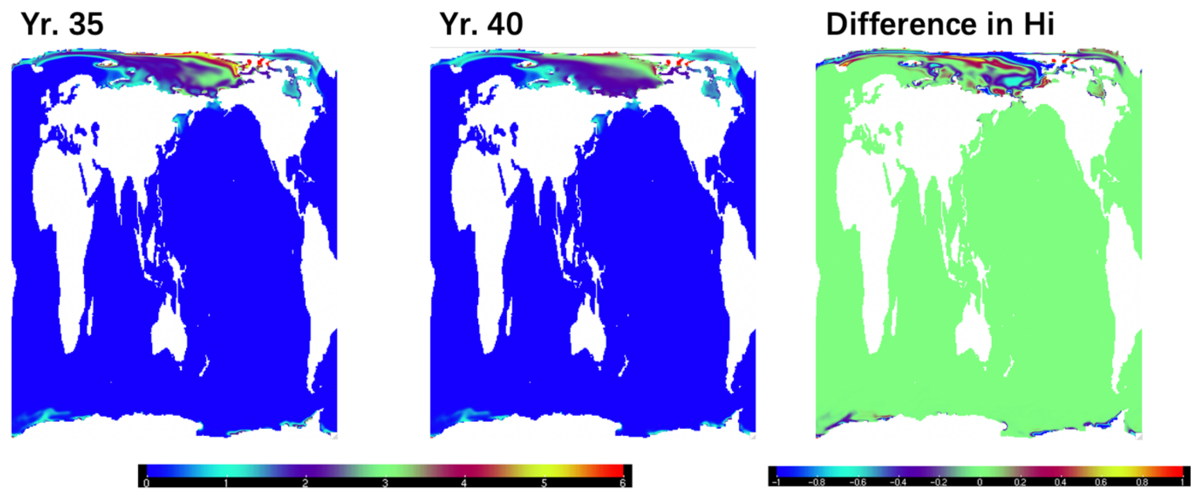

Fig. 2. March sea ice thickness difference for GX1V6 grid at equilibrium states under CESM D-type experiments. Left: Rothrock (1975) scheme; Middle: Hibler (1979) scheme; Right: difference (middle *minus* left). All values are in meters.

(a) If you are interested to study the effect of different resolution on sea ice dynamics (which is the topic of Section 3.3), all simulations should produce comparable sea ice distributions (concentration fields thickness as well volume and extent). Otherwise it is not possible to disentangle the effect of the change in resolution and the change in sea ice state on the dynamics. Systematic tuning methods (Massonnet et al., 2014; Ungermann et al., 2017; Sumata et al., 2019) could be used for all three simulations to optimize the parameter choices for each simulations by minimizing the model-observations misfit (for

instance concentration, thickness, and drift). To resolve the issue of too thick ice in the Beaufort Gyre, drag coefficients and the ice strength parameterization could be tuned. The tuned simulations are then a good starting point for further multi-resolution studies and also the various parameters determined in the optimization will provide insight in how model physics change with resolution.

Reply to the general comment (1.a): the authors fully agree with the referee that it is necessary to tune the model in order to improve modeled sea ice distribution and comparability. In our study, our original intention is to align the parameterization across the resolution to ensure good comparability, at least the thermodynamics should be kept the same across the resolution range. As indicated by the referee, we have found sea ice thickness in BG greatly dependent on the ice strength parameterization scheme, as shown above. Although the scheme we have adopted is used by default in CESM, we consider it to be replaced with the more commonly used scheme in Hibler (1979) in future studies.

(b) If such systematic tuning is not possible due to limited computational resources, the authors should be more cautions with statements regarding the good agreement of ice thickness fields and agreement of all three simulations. The differences between the three simulations should be described and interpreted in details. Possible reasons for the different sea ice distributions should be provided along with guidance what limitations with regard to the kinematic studies originate from the different sea ice distributions.

Reply to the general comment (1.b): according to the referee's suggestions, we further compare the Arctic sea ice climatology among the 3 grids. However, based on experiments that evaluate the effects of sea ice strength parameterization (shown above), we argue that ice strength parameterization causes much more uncertainty than the differences among the 3 grid resolutions. Since we have limited computational resource for the model runs, we follow the criteria below for the new experiments and related analyses. We align the thermodynamic parameterization schemes as well as parameter values across the 3 resolutions. We examine in detail the ice strength parameterization (including its parameters) in the modeled climatology. Furthermore, we include a new appendix showing all the parameters that are utilized in our study.

2) The analysis of deformation rates in the manuscript is limited to two 3-day intervals. Since the scaling properties of sea-ice deformation are highly variable (Stern & Lindsay, 2009) and strongly impacted by atmospheric winds (Herman & Glowacki, 2010), limiting the analysis to such a short time interval does not allow robust conclusions on the model capability to simulate multi-fractal deformation rates. It can not be excluded that the two dates chosen for the analysis mainly highlight the imprint of the atmospheric forcing. Another problem with the too short interval are the CDF of deformation rates that do not show power-law tails due to strong fluctuations (although stated differently by the authors). I suggest to extend this analysis to at least one entire winter. This will reduce the impact of specific wind conditions, smoothen the CDFs, and allow a more robust interpretation of the presented results with regard to the models ability to simulated strongly localized deformation rates along leads and pressure ridges. In addition I recommend to remove all statements on temporal scaling based on these two 3-day intervals from this manuscript, as now temporal scaling analysis is performed by the authors.

Reply to the general comment (2): regarding to the referee's comment on the analysis of deformation fields, we have made the following two revisions. First, we extend the study to the full winter, including the study of C-CDF and scaling analysis. The discussion on these two representative days are further carried out on daily deformation fields during the winter months. Second, as pointed out by the referee, the temporal scaling has not been carried out in strict manner, and therefore we remove the statements involving temporal scaling. Specifically, we only present the 3-day results without relating them to (or drawing any conclusion on) temporal scaling properties.

3) The good agreement of deformation fields between the different resolutions surprised and impressed me. In your simulations only the degree of detail in deformation feature increases, but the general patterns agree across the different resolutions. Knowing that ice fracture is a chaotic process that is very sensitive to small variations in ice strength these results puzzles me, as I was expecting that the deformation fields diverge very fast due to the different deformation history. At high resolution, a deformation event which is associated with divergence reduces concentration and thickness, and thereby the ice strength, such that deformation is more likely to appear in the same spot again. This effect should not be so effective in coarse resolution simulations as the reduction in concentration and thickness is much smaller due to the size of the grid box. This different memory should cause different reactions to the same atmospheric forcing. Do you see a reduction in concentration and thickness along the simulated LKFs in all your simulations? Do you see reoccuring deformation lines in all simulations? Your results indicate rather that in general this described feedback is not so strong and that fracture is mainly driven or better prescribed by the forcing, which would be an interesting result. This aspect of your results is definitely worth more discussion in the paper and maybe some additional analysis.

Reply to the general comment (3): regarding to the referee's comments, we in the first place were also a little bit surprised about the consistency of deformation features across the 3 vastly different resolutions. We conjecture that this is due to the specific experiment design of using NYF forcing, as well as the specific strength parameterization scheme we have adopted. We are yet to explore deeper and carry out attribution study of why the kinematic features agree and find the factors can break such consistency. Specifically, we have 2 aspects to look into. First, we want to explore the effect of inter-annually changing forcings on the system. A constantly alternating forcing may deviate the pathway of different resolutions in terms of spatial distribution of thick/thin ice, and may cause differences in the deformation events (or strengths). Also with the ice strength parameterization scheme of Hibler (1979), we also want to carry out similar analysis comparing the deformation fields. Regarding the two specific questions, first, we do witness concentration changes with LFKs if divergence is present. Second, we witness some cases of re-occurrence of the LKFs, but we haven't carried out systematic analysis yet.

Specific comments:

P1, Line 3, "multi-fractality": of what? Please add scaling of sea-ice deformation

Reply: revised as indicated by the referee.

P1 Line 19, "kilometer-scale" satellite observations: SAR images have a resolution in the range of tens of meters. The drift and deformation products derived from consecutive SAR images have a kilometer-scale resolution. Please be more specific.

Reply: according to the referee's comment, we have revised the sentence as follows: "… kilometer-scale sea ice drift and deformation estimates with Synthetic Aperture Radars …"

P2 Line 1, "Linear kinematic features": You have not described what these linear kinematic features are. Please describe once what they are (failure and shear lines where deformation is localized).

Reply: revised by replacing "Linear kinematic features …" as "Linear kinematic features, including local deformation regions of sea ice failures and shearing, …"

P2 Line 7: In the VP framework, the transition between viscous and plastic deformation depends on the stress states and not the concentration. The concentration influences the stress states by scaling the ice strength, but there is no direct link as suggested by your description. Please clarify.

Reply: the sentence is revised to be more precise, as follows: "it describes the sea ice as a two dimensional continuum with nonlinear viscosity and plastic deformations under high stress conditions of compression and shear".

P2 Line 14: CMI -> CMIP (here and elsewhere in the manuscript)

Reply: replaced here and every use throughout the text.

P2 Line 15: This is true for VP/EVP models. For other rheologies that include memory of past deformation, as the Maxwell elasto-brittle rheology, also coarser grid resolution might produce similar deformation statistics.

Reply: we revise the sentence to be more precise, as follows "With VP rheology, the capability of sea ice models to resolve fine-scale deformations is inherently bounded by the resolution of the models' grid."

P2 Line 18: The continuum assumption is part of all continuum sea-ice models regardless what rheology they use. Please consider not explicitly mentioning the rheology here.

Reply: the sentence is revised by removing the specific rheology of VP, as follows "Although the continuum assumption of the sea ice cover does not necessarily hold at these resolutions, …"

P2 Line 22, "main driver": It is not clear to me what you mean with main driver. Please clarify this sentence.

Reply: the sentence is revised as "… adopted by various research groups in the world for climate studies."

P4, Table1: Please be more specific with the grid descriptions in the "Notes: column, such that the table is understandable without reading the text. There is enough space for that.

Reply: the last column (Notes) of the table is extended for a more understandable description.

P4 Line 21-22: at the grid location and 60 vertical layers,. . .

Reply: revised according to the suggestion.

P4 Line 25: Please rewrite sentence.

Reply: we reformulate the sentence as "Second, we configure the model according to the grid resolution, including the choice of parameterization schemes and related parameters that are used."

P6 Figure 3: Please think about using the same limits for the contour plots for both grids. This would make it easier to see the difference between them. The contour lines are also hardly visible, you might also want to use a brighter red instead.

Reply: the figure is revised to improve clarity and readability, according to the referee's suggestion.

P8 Line 6-8: The thickness anomaly in Beaufort Gyre could also be caused by too weak ice and not properly tuned ice strength parameterization. The thick ice north of CAA and Greenland is then advected by the ice drift and accumulates within the Gyre.

Reply: the authors have carried out experiments which show very large sensitivity of equilibrium sea ice thickness (including BG) with respect to the ice strength parameterization. By replacing the default scheme adopted by CESM with Hibler (1979), the anomalously thick ice in BG is now gone. Considering the fact that the default scheme adopted by CESM is widely used in many scientifically validated experiments, including OMIP and CMIPs, we plan to carry out detailed attribution study in the future on this issue.

P8 Line 13, "With the warm start-up, the experiments with TS005 approaches equilibrium towards year 42.": Only for the extent, the volume is still decreasing. Please clarify.

Reply: we further make analysis and attribution of the differences among the 3 grid resolutions. Before that, the experiments are also continued to allow TS005 to reach an equilibrium after year 45.

P8 Line 17, "The overall sea ice coverage and volume of TS005 is also in good agreement with satellite observations and PIOMAS dataset.": I would not describe the strong overestimation of sea ice extent in winter as a good agreement. In addition, I miss the comparison with the PIOMAS dataset in the figure. Please state where to find this comparison.

Reply: we revise the sentence as "In general, the overall sea ice coverage and volume as modeled with TS005 are consistent with satellite observations and PIOMAS dataset". In Fig. 5.b we also add the climatological seasonal cycle from PIOMAS dataset. The years of 1979-2000 are adopted, same as the SIE climatology from NSIDC data.

P8 Line 17-19: I do not understand why using the same parameterizations for all three grids is a reason for reasonable results. It is known that model parameters need to be adapted to different grid resolutions to show similar physics (e.g. Williams & Tremblay, 2018). Please clarify or rewrite.

Reply: regarding the comment, we would like to make the following clarifications. First, we expect the thermodynamic parameterization schemes, along with the parameter values, to be consistent across the resolutions. The dynamics across the resolutions definitely will cause differences in the distribution of ice (both local ridging and spatial re-distribution), but we do not want to compensate these errors (or uncertainties) with thermodynamics. Second, we fully agree that in order to quantitatively improve the model's performance, the modeler (or model user) should tune the model to available observations. However, since under NYF there is no exact sea ice climatology for us to match, we therefore focus on the consistency among the 3 resolutions as long as they attain reasonable results. For future studies we look forward to carrying out IAF based experiments, for which a full suite of model tuning to match observed sea ice historical changes is planned.

P9 Line 8, "removed of seasonal cycle": -> and the seasonal cycle is removed

Reply: revised.

P10 Figure5, "satellite-observed": Please state which satellite product is used for this comparison

Reply: the mean annual cycle of NSIDC SIE product of year 1979-2000 from SSMI/SMMR sensors is retrieved and used as climatological seasonal cycle. The figure caption is revised to include this information.

P10-11 AO index analysis It is not clear why this analysis is needed here. As the corresponding explanation is rather complex, please consider to remove them from manuscript for clarity.

Reply: regarding the comment on AO analysis, the authors want to clarify that the inclusion of AO indices is to ensure that the NYF dataset is not untypical in terms of wintertime atmospheric forcings. If prominent negative or positive AO is present in the NYF dataset, we would expect much different sea ice circulation and thickness distribution.

P11 Line 8: sybcycle count -> subcycles

Reply: revised as indicated.

P11 Line 17-18: "The kinematic features with TS005 are richer and much narrower, such as the network of shearing in Beaufort Sea." Do you want to say that in TS005 more and finer features are simulated?

Reply: yes, and we revise the sentence as "With TS005, the model simulates more and finer sea ice kinematic features".

P11 Line 35: The region for the analysis you have chosen is problematic as it mixes pack-ice regions with coastal regions. In coastal regions stable deformation features, like flaw lead, are found that show nearly constantly very high deformation rates, which impacts the presented CDFs. I suggest to use the entire Arctic Ocean as study region and filter all grid points that are closer than 150-200km to the coast as done in other scaling studies.

Reply: we agree with the referee's comment on the effect of coastal region on the PDF and scaling analysis. We therefore update the results by limiting the analysis within the basin, to the common regions according to TS045 (the coarsest resolution among the three).

P12 Line 3: Please be cautious for two reasons: (1) just because the PDF/CDF of sea ice deformation shows a power-law tail does not mean it is multi-fractal. To show mulit-fractality a scaling analysis of the moments of sea-ice deformation need to be performed that shows a non-linear convex structure function (you do this analysis but it is described later). (2) The distributions shown in Figure 9 show hardly power-law distributions. I suggest to use the methodology of Clauset et al. (2009) to test for power-law distributions.

Reply: the authors fully agree with the referee. We have made a mistake that we have included the texts describing the spatial scaling results to the analysis of C-CDF. We have removed the description of multi-fractality from this part of the paragraph.
Regarding the power-law distribution, since for the end of the C-CDF tail (which corresponds to relatively larger deformation events) we are hit with very small sample count, therefore, the determination of slopes is carried out for the range of 0.05 and 0.25 for the C-CDFs (noted in the figure caption).

P12 Line 6: What do you mean with "spatial scaling"? Are you coarse-graining the high resolution simulation to coarser grid resolution? Please clarify.

Reply: yes, and we revise the sentence as " … we carry out: (1) the spatial coarsening of the model output of TS005 onto TS015 and TS045, and (2) that of TS015 onto TS045."

P12 Line 6-21: (1) The CDF in Figure 9 hardly show power-law tails and deviate from strait tails. It is not clear how you determine the power-law slopes. I recommend to use larger time intervals for the analysis to reduce the imprint of certain atmospheric forcing conditions and second to use the methodology presented in Clauset et al. (2009) to test for power-law distributions.

Reply: the authors have extended the analysis to winter months and limited the analysis to strictly within the basin (away from coast) to attenuate atmospheric noises and avoid potential problems on coastal regions. Details and results are presented separately in the revised Sec. 3.2.

(2) It is unclear to me how you relate the slopes the CDF-tails of coarse-grained deformation rates to the nominal resolution of the grid. Please describe this concept more in detail.

Reply: the authors have rewritten the paragraph to include a more formal introduction of the methodology we have adopted. In brief, we use the simulation results from TS005 as a reference, by spatial coarsening of the results with TS005 to a coarser grid such as TS045, we compute the difference between the tail slopes at the grid scale of TS015. As shown, the slopes are flatter for TS005, indicating that with TS015 the model cannot simulate reasonable sea ice deformation rates with respect to the reference of TS005. We consider this difference due to that the grid-cell scale deformation is actually not realistic for TS015, and therefore a coarser spatial scale can be determined with coarsening, when the slopes from TS015 finally matches that of TS005. This scale we consider is the effective resolution of TS015 for simulating sea ice deformations.

P12 Line 22-26: Given the limitations of your analysis (short-time interval, no clear power) I do not recommend a direct comparison with observations or at least mention these limitations.

Reply: according to the referee's comment, we revise the paragraph to avoid direct comparison with Marsan et al. (2004) or any observational dataset. Afterall, due to the use of NYF, there is no direct comparability of these data.

P12 Line 32, "about 1.3 on Feb. 6th for all three grids": I see values from 1.2 to 1.3.

Reply: according to the referee's comment, we revised the sentence as "and between 1.2 and 1.3 on Feb. 6th for all three grids."

P13 L1-2, "Furthermore, no positive value of β is detected at q = 0.5, which is consistent with Marsan et al. (2004) (Fig. 4 of the reference).": Please clarify. In Marsan et al. (2004) beta is positive for q=0.5. Also in your Figure 9 beta seems to be positive for q=0.5. What would be the physical interpretation of negative scaling exponents if you find them in your model?

Reply: we correct the mistake in citing the value of beta from Marsan et al (2004), by changing "positive" to "negative". For reported data with observations [such as Marsan et al. (2004)], the value of Beta is still positive for q=0.5. For the model to report negative values for Beta, statistical issues might be the cause. Since the usual practice of using linear fittings in the analysis in Fig. 10 (left side of each panel), we usually ignore the different statistical confidence on different scales. For larger spatial scales, the sample count is significantly smaller than small scales, given that the analysis is carried out on the same dataset. Uncertainty in estimating the mean deformation rate could play an important role for the estimation of Beta, causing statistically insignificant negative values for Beta.

P13 Line 6-8: Please be more specific. Do you mean that with increasing resolution, deformation rates are more localized with yields to more pronounced scaling?

Reply: regarding the referee's comment, we consider this description applies to the contrast between the two days (Dec. 20th and Feb. 6th). But this description is not very strict, since we lack the analysis of relating it to general cases and convincing attribution studies. For revision, since we have the model results from the whole winter, we relate the difference in scaling properties to various factors including circulation and forcing data.

P13 Line 8-9: Temporal scaling is indicated by the decrease in β for the daily field and 3-day field for Dec. 20th, and not evident for Feb. 6th. This could be also caused by just smoothening of deformation fields due to advection. To test for for temporal scaling a Lagrangian analysis is needed that follows the ice deformation with the drift. Please remove this sentence or add analysis.

Reply: the authors agree with the referee that the temporal scaling analysis formally involves Lagrangian based analysis. Besides, we want to clarify that a pure Lagrangian diagnosis with our model (which is based on Eulerian grids) is not practical, since we cannot attain the Eulerian drift speed at every time step. Therefore, we use daily mean sea ice drift speed for all the spatial analysis. We project that using smaller time interval for the computation of mean sea ice drift speed, the analysis results would converge to a pure Lagrangian based analysis.

P13 Line 14, "indicating less dominant large-scale 15 features on Feb. 6th.": or a more heterogenous distribution of deformation rates along the LKFs.

Reply: the authors agree with the referee's insight on this issue. We consider this scaling property inherent to the specific scenario of circulation/forcing. The LKFs on Feb. 6th, as a result of the thickness distribution and circulation, are indeed more localized and heterogeneous when compared with Dec. 20th.

P13 Line 15-16, "Furthermore, there is more effective temporal scaling on Dec. 20th than Feb. 6th, as shown for C-CDFs in Figure 9 and structure functions in Figure 10.": Please remove this sentence since no temporal scaling analysis is done.

Reply: we agree with the referee that due to the lack of Lagrangian perspective, the analysis we carried out here is not strictly a temporal scaling analysis. We have revised the sentence to remove the statement of "temporal scaling".

P14 Line 4, "Figure 7": Do you mean Figure 12?

Reply: corrected.

P14 Line 6, "noisier": It is really hard to spot the noise in Fig. 12 except you zoom very strongly in certain regions. Could you find better ways to show this? For instance, plot or average the difference between the deformation rates in a grid cell and its local surrounding (couple of grid cells). This would shift the focus on the noise. Or just zoom on a certain subdomain where the noise is seen.

Reply: according to the referee's comment, we add a supplementary figure showing the details in region around and north of CAA for the experiments with TS005. This will highlight the noise due to non-convergent EVP solutions.

P17 Figure 9: This figure needs more explanation in the caption: What do the colors refer to (NDTE?)? Are 0.05 ◦ , 0.15 ◦ , and 0.45 ◦ the grid resolution and why do you not use the names T005, etc. here?

Reply: the figure as well as the caption are revised to be more precise. The lines are colored by the grid name, while different shapes represent different spatial scales (0.05-deg, 0.15-deg, and 0.45-deg).

P17 Line 5, "equilibrium in sea ice thickness and volume": But volume is still increasing, please clarify how this fits to the claimed equilibrium.

Reply: we further carry out continued experiments to ensure that a quasi-equilibrium status is attained for TS005.

P22 Line 17: MITGcm -> MITgcm

Reply: corrected.

P22 Line 20, "with initial study with temporal scaling analysis with 3-day mean drift fields": Please remove, since you have not done a temporal scaling analysis.

Reply: corrected as "with initial study of 3-day mean sea ice drift fields".

P23 Line 22-23: Remove one "in our study".

Reply: corrected by removing one of the "in our study"

P22 Line 28 - P23 Line 3: This paragraph is rather a summary of on going research in the sea ice modelling community and your future plans and not a conclusion of your study. Please remove it here or move to the state of research in the introduction.

Reply: the whole paragraph is moved into Sec. 1 (on page 4), as suggested by the referee.

P23 Line 5: Which efforts? Please add citations.

Reply: citations are added for EVP convergence, including Lemieux et al. (2012), Kimmritz et al. (2015), and Koldunov et al. (2019).

P23 Line 30 is -> are

Reply: revised.

P24 Line 3-4, "Given that the modeled sea ice climatology is reasonable and consistent among the three resolutions": In the high resolution run, the sea ice climatology is distinctively different from the two coarser runs, which indicated that parameters of the sea-ice model need to be tuned for each specific resolution to reach the same climatology. I agree that using a slab-ocean in this study is fine, but further tuning of sea ice model parameters would be required to obtain runs with comparable sea ice climatology. Please elaborate on this.

Reply: the authors agree with the referee's comment on the model output presented in the original manuscript. Two aspects are improved with add-on experiments. First, we further carry out continued experiments to ensure that a quasi-equilibrium status is attained for TS005. Second, in this study we have used NYF to force the sea ice model, and this hinders the comparison with observational datasets and the ensuing tuning process, since we do not expect the model exactly match any climatology we choose. Therefore, in this study we focus on the consistency and inter-comparison of the modeled sea ice states across the resolutions, given that the climatological sea ice status is reasonable. We intend to carry out IAF based experiments for future studies, in which a formal model tuning process is planned by using observed sea ice changes as reference datasets.

References:

Flato, G.M. and W. D. Hibler (1995). Ridging and strength in modeling the thickness distribution of Arctic sea ice. J. Geophys. Res.–Oceans, 100:18611–18626.

Hibler, W.D. (1979). A dynamic thermodynamic sea ice model. J. Phys. Oceanogr., 9:817–846.
Kwok, R. (2010). Satellite remote sensing of sea-ice thickness and kinematics: a review, Journal of Glaciology, 56(200):1129-1140
Lipscomb, W.H., E. C. Hunke, W. Maslowski, and J. Jakacki (2007). Improving ridging schemes for high-resolution sea ice models. J. Geophys. Res.–Oceans, 112:C03S91, doi:10.1029/2005JC003355
Marsan, D., Stern, H., Lindsay, R., and Weiss, J. (2004). Scale Dependence and Localization of the Deformation of Arctic Sea Ice, Phys. Rev. Lett., 93, 178 501, doi:10.1103/PhysRevLett.93.178501.
Rampal, P., Dansereau, V., Olason, E., Bouillon, S., Williams, T., Korosov, A., and Samaké, A. (2019). On the multi-fractal scaling properties of sea ice deformation, The Cryosphere, 13, 2457–2474, doi:10.5194/tc-13-2457-2019.
Rothrock, D.A. (1975). The energetics of the plastic deformation of pack ice by ridging. J. Geophys. Res., 80:4514–4519