Author comments

Ensemble design for seasonal climate predictions: Studying extreme Arctic sea ice lows with a rare event algorithm

J.Sauer, F. Massonnet, G. Zappa, F. Ragone

submitted to Earth System Dynamics (ESD) within the Special issue "Theoretical and computational aspects of ensemble design, implementation, and interpretation in climate science" (EGUSPHERE-2024-3082)

January 24, 2025

Reviewer #1

Reviewer. *This study presents an interesting next step from the Sauer et al 2024 paper. It focuses on using a rare event sampling algorithm to diagnose how initial conditions influence the extremely low summer sea ice conditions in the Arctic in an intermediate complexity model. This paper is well written, clear and relatively concise, and would be a good fit for the journal. It is an interesting progression from the original work and different enough to merit publication. The authors find that initial conditions are very important for determining the extent of extreme low sea ice conditions, and also present plausible mechanisms through which these low conditions form. I have some points the authors should consider before publication.*

Reply. We thank the reviewer for the very helpful comments to improve our manuscript. We will reply to these comments in the following. For the sake of clarity, we will copy/paste the individual comments before our replies and actions.

Reviewer. *Major comments:*

One concern is in balancing the needs to learn things from more idealised setups (such as this intermediate complexity model) so that the learnings can be transferred to the real world and comprehensive models, versus the question of how applicable these learnings will be in a much more complex model. Throughout my reading of this manuscript, I kept thinking to myself how many of these results will be applicable in state of the art models, and this is something the authors inherently cannot answer because they have restricted themselves to the intermediate complexity model from the previous study. I think this is ok for this study but would suggest that for further progress to be made, a step will need to be taken towards implementing this in a comprehensive model.

Related to this, the authors barely discuss the implications of the simplifications of the intermediate complexity model on the results (only alluded to in the last paragraph). To my mind there are a number of idealisations which could make a large impact:

- *Lack of sea ice dynamics,*
- *Atmosphere only influences sea ice through thermodynamics,*
- *Idealised ocean dynamics (I find it hard to tell how idealised),*
- *Extremely coarse atmospheric resolution,*
- *Lack of snow on sea ice (I think?)*
- *Binary sea ice concentration (i.e no sea ice fraction >0 or <1).*
- *Realism of Arctic cloud cover and how it responds to sea ice changes*
- *Unsure who well coupled modes of variability are represented in this model*

I think that each of these could have important implications for how you interpret your results. And as such I think looking at 0.001% events in this model is challenging to interpret for the real world.

Reply. For a relatively small computational cost, the Planet Simulator (PlaSim) allows to learn about the dominant large-scale thermodynamic process and time scales relevant for the formation of low sea ice states in the Arctic. This model and the present study provide an extremely useful knowledge required to carefully prepare the experimental design with the rare event algorithm when applied to computationally expensive state-of-the-art climate models. Certainly, as suggested by the reviewer, progress regarding our understanding of the precise characteristics, dynamics and statistics of extremely negative sea ice area anomalies will require the implementation of the approach to a comprehensive climate model. A similar rare event algorithm study to the one of this manuscript, albeit using EC-Earth model version 3.3.1, is currently in preparation.

While the Planet Simulator (PlaSim) includes snow on sea ice, all the other idealisations mentioned by the reviewer potentially affect the development and predictability of extreme late summer sea ice lows. The dominant modes of atmospheric circulation variability such as the Arctic Oscillation and the Arctic Dipole pattern are well represented in PlaSim. However, their association with anomalously low sea ice states is potentially underestimated in PlaSim due to the lack of sea ice dynamics and eventually due to a too coarse spatial resolution. For example, an anomalously positive summer phase of the Arctic Dipole pattern may favour anomalously low summer sea ice states via enhanced southerly winds pushing the sea ice edge from the Pacific side of the Arctic towards the central Arctic Ocean [Wang et al. 2009], while a negative summer Arctic Oscillation phase can lead to enhanced sea ice export from the marginal Arctic seas towards the central Arctic Ocean and out of the Arctic via Fram Strait [Ogi et al. 2016]. Likewise, the impact of synoptic-scale storms on anomalous reduction of sea ice area [Zhang et al. 2013] is not captured by the model. Consequently, the lack of sea ice dynamics potentially leads to an underestimation of the amplitudes and probabilities of the most extreme sea ice lows in PlaSim compared to the real world.

Similarly, regional biases in the sea ice thickness field certainly impact the comparability of the drivers of anomalously low sea ice area in PlaSim compared to observations and potentially also affect the estimated probabilities (cf. Figure R1-1(right) and page 207 of Chevallier et al. [2019]).

Action. We discuss these points in section 5 of the revised manuscript.

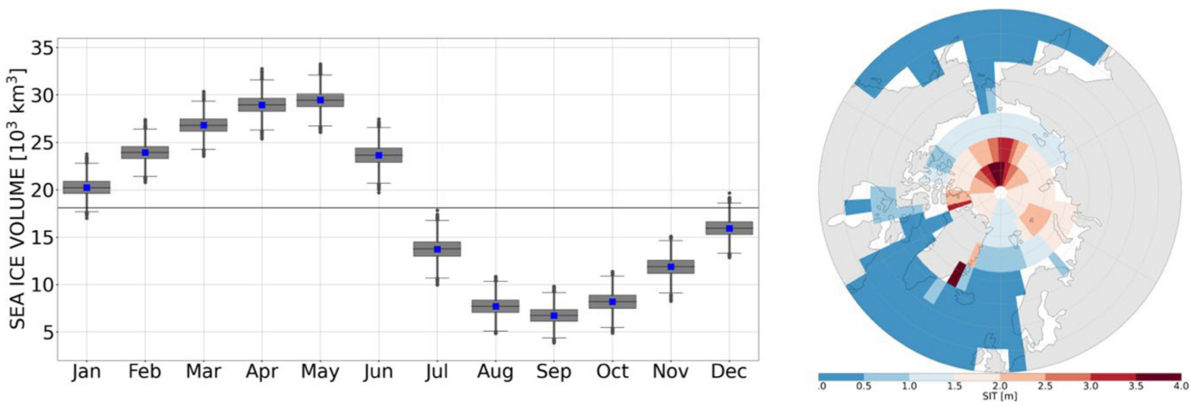

__Figure R1-1:__ Model years 501-3500 of PlaSim-LSG control run [Sauer et al. 2024]: (Left) Distributions of monthly mean pan-Arctic sea ice volume [10^3 km^3] with respect to the 3000 control run model years. The averages and medians are given by the blue squares and the horizontal lines in the boxes. The boxes denote interquartile ranges and the maximum whisker length is defined as 1.5 times the interquartile range. The horizontal gray line shows the annual mean pan-Arctic sea ice volume. (Right) Annual mean sea ice thickness [m].

Reviewer. *Another major point which needs to be addressed is that you spent a substantial portion of the introduction motivating the question of anthropogenic versus* *natural to the loss of sea ice in 2007 and 2012. However, you didn't link your work back up to this question in the conclusions. How has what you have learnt made progress with this question? If the results from the intermediate complexity model are difficult to translate into this question then I suggest you reduce the role of this question in the introduction. However, I think you could have a good paragraph in your conclusions addressing this question.*

Reply. The purpose of this study is to present an approach to study the probabilities/predictability and physical drivers of low sea ice states in a seasonal climate prediction set-up, thereby exploiting ensemble simulations starting from single initial conditions and reducing the computational cost required to sample a sufficiently robust statistics on the extremes. For example, the approach used in this study can be used in the future to quantify the impact of the initial condition in specific years (such as in 2007 and/or 2012) to the amplitudes and probabilities of extreme sea ice lows. Such an approach would primarily inform about how predictable the 2012 event would have been based on the knowledge of the initial condition. Learnings about the relative contribution of global warming to the probabilities of these events would only be indirect and qualitative (indirect via the impact of preconditioning of enhanced open-water-formation due to the ongoing winter sea ice thinning [Lindsay et al. 2009; Parkinson and Cosimo 2013; Guemas et al. 2013]).

The experimental design could, however, be easily adapted to make more quantitative statements about the relative contribution of global warming to extreme sea ice loss. This could for example be achieved by initializing an ensemble with late winter initial conditions sampled from e.g. the period 2016-2025 (i.e. one initial condition per year times 30 perturbed replicates yielding an ensemble of 300 trajectories) and comparing the statistics with another ensemble initialized with data e.g. 1986-1995 drawn from a historical simulation. That is the reason why we include the discussion about the relative contribution of global warming to the probabilities of the extremes.

Action. As our study is not designed to provide conclusions about the relative contribution of global warming to extreme sea ice loss events, we slightly shorten this part in the introduction. We keep this in the conclusion.

Reviewer. *L25: Potentially useful reference: England et al 2019 https://journals.ametsoc.org/view/journals/clim/32/13/jcli-d-18-0864.1.xml*

Reply. Thank you for this interesting reference. In the revised manuscript, we refer to this paper in lines 24 and 33.

Reviewer. L30: In this paragraph it would be useful to give the context that summer sea ice conditions are still above the 2012 minimum, which suggests that climate change is not the entire reason, because global warming has obviously continued considerably since 2012.

Reply and action. *Yes, thanks. We included this point in lines 38-40 of the revised manuscript.*

Reviewer. *L64: You also have the need to rely on multiple models rather than a single model which would increase the computational expense further. However ,wouldn't the existing large ensembles of CMIP5 and CMIP6 models be a useful resource here? https://www.cesm.ucar.edu/community-projects/mmlea There are multiple models with 100 ensemble members and if you are interested in a 2012-like event relative to the forced signal, rather than an event in 2012, you could search through the timeseries for each member which would give a much larger sample size.*

Reply and action. Yes, this is certainly a helpful data source. An analysis of the CMIP5/CMIP6 large ensemble would provide useful complementary information to a rare event simulation applied to one state-of-the-art climate model. Unfortunately, however, this analysis is beyond the scope of the paper. We mention the possibility to complement a future EC-Earth3 rare event algorithm study with an analysis of the CMIP large ensembles in section 5.

Reviewer. *I am not convinced by the statement in L155, given that the mean state could be readily tuned to give a reasonable state. This doesn't necessarily tell us about how suitable the variability or extremes in the model are.*

Reply and action. We agree. The winter to spring pan-Arctic sea ice volume matches relatively well the climatology obtained from the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) (cf. Figure R1-1 and page 207 of Chevallier et al. [2019]). Regarding the spatial distribution of the annual mean sea ice thickness, substantial differences are visible between PlaSim and PIOMAS, with the former overestimating sea ice thickness in the central Arctic Ocean and in the Barents Sea and underestimating sea ice thickness northern of Greenland, compared to the latter (cf. Figure R1-1 and page 207 of Chevallier et al. [2019]). The regions of overestimated sea ice thickness correspond to regions relevant to extremely negative late summer sea ice area anomalies shown in Figure 4 of the main manuscript. Thus, the trajectories with the most extremely negative sea ice concentration anomalies available from the ensembles show no substantial sea ice concentration anomalies around the North Pole, while strongly negative sea ice concentration anomalies are present in the Barents and Kara Seas (Figure 4 of the main manuscript). In observations, the Barents and Kara Seas are sea ice-free in late summer [Kapsch et al. 2019]. While the overestimation of sea ice in the Barents and Kara Seas in PlaSim compared to observations likely contributes to an overestimation of the amplitudes of extremely negative pan-Arctic sea ice area anomalies compared to the real world, an opposing effect is to be expected from the positive sea ice thickness bias around the North Pole in PlaSim compared to observations.

Action. In the new manuscript, we mention the regional sea ice thickness biases in section 2.1 and discuss its implications for the results in section 5.

Reviewer. *L211: How do you know that M=7 is sufficient to explore the space?*

Reply and action. The purpose of the manuscript is to present and exploit an experimental design to study probabilities and drivers of low sea ice states in a seasonal climate prediction set-up, using single initial conditions. For example, the approach used in this study can be used in the future to quantify the impact of the initial condition in specific years (such as in 2007 and/or 2012) to the amplitudes and probabilities of extreme sea ice lows. Following the comments of reviewer #3, we produced nine additional realizations of the “SIA-NEUTRAL SIT1.93-NEUTRAL”, “SIA-NEUTRAL SIT1.93-LOW”, “SIA-LOW SIT1.93-NEUTRAL” and “SIA-LOW SIT1.93-LOW” initial conditions on top of the existing ones of the original manuscript (i.e. we produced a total of 4 x 2 x 10 ensemble simulations with 600 trajectories running from February-September). This allows us to estimate on a statistically robust basis for a few selected initial conditions the impact of winter sea ice area and cumulative area with sea ice thickness equal/larger than 1.93m on the probability of extremely low late summer sea ice area. We agree, however, that conclusions about the impact of low winter sea ice area/thickness on late summer sea ice lows could be drawn in a more robust way against the influence of one particular single initial condition if the ensembles were initialized with different initial conditions sampled from a subspace (e.g. all initial conditions exceeding +/-2 sigma levels with respect to sea ice area/thickness).

Actions.

- Section 2.2: We updated the description of the experimental design. We are now studying four instead of seven initial conditions and we are producing ten realizations instead of one realization of each control and rare event simulation. Table 1 is updated accordingly.
- Figure 2 of the original manuscript is updated by Figure R1-2: For the baseline of the statistics, we show the probability distribution functions for the sea ice area obtained from a 6000 trajectory “master” control ensemble with all ten control realizations merged together. Likewise, the original realization of the “NEUTRAL SIA – LOW SIT1.93m” experiment is replaced by “REALIZATION #9” where the most extreme late summer sea ice low is more extreme than in the original realization.
- Figure 3 of the original manuscript is updated by Figure R1-3 showing probability estimates including confidence intervals based on multiple realizations of the rare event simulations.
- The text in section 3.1 and in section 5 is updated according to the new data and according to the new probability estimates.
- Table 2 is reduced from seven to four initial conditions.
- Section 4: Analysis of the physics for the “NEUTRAL SIA – LOW SIT1.93m” experiment: “REALIZATION #9” replaces the orignal “REALIZATION #1” in Figures 4-7. The text is adapted according to the new realization.
- Line 536: We emphasize that our conclusions about the preconditioning of low summer sea ice area through the winter state refers to four specific initial conditions.
- A description of the computation of the confidence intervals is given in the Supplementary Information S4 of the revised manuscript.

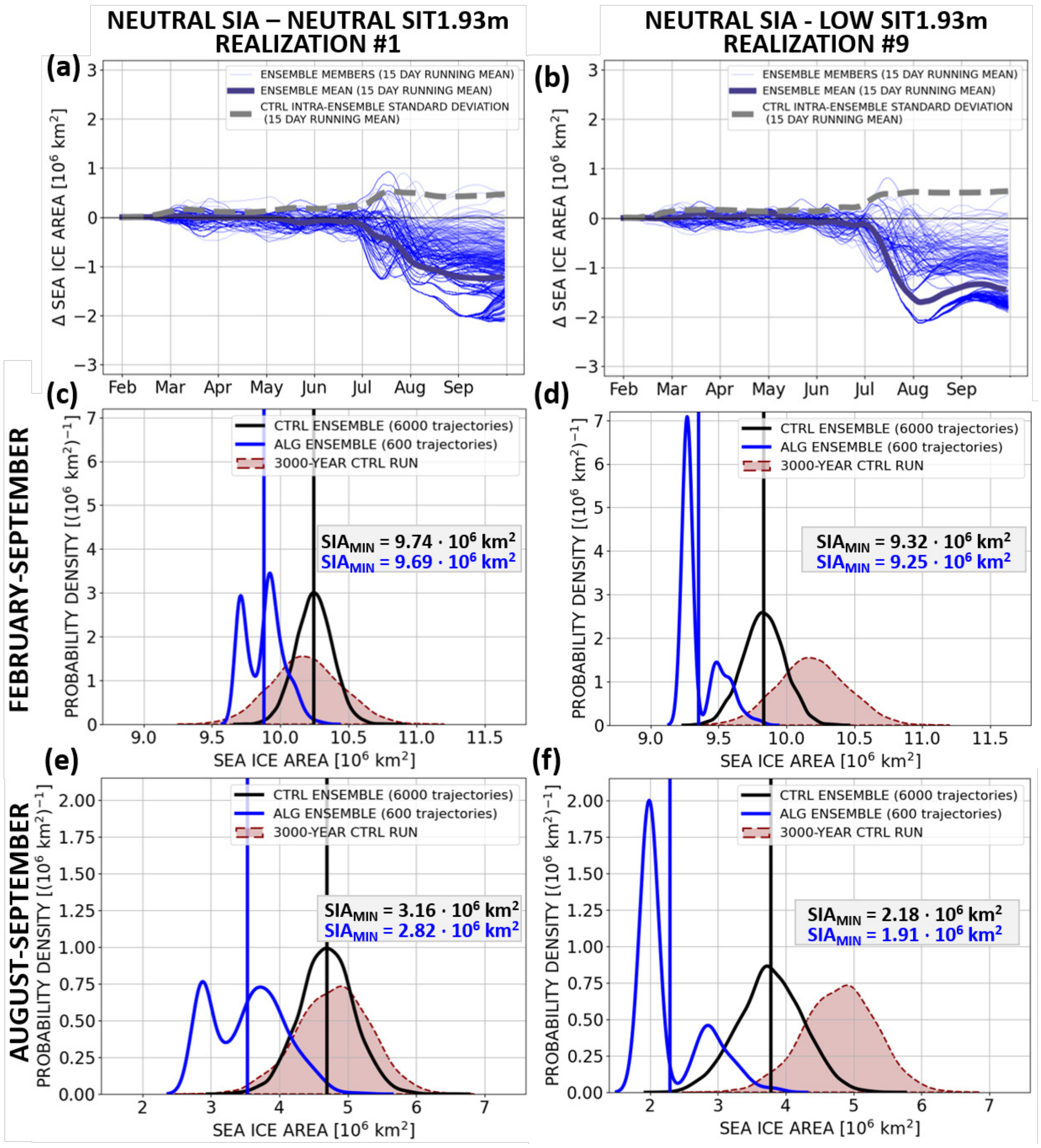

Figure R1-2: *Ensemble simulations initialized on 01 February (a,c,e) 1930 and (b,d,f) 1037 of the control run. (a-b) Rare event simulations: N=600 trajectories (thin blue lines) and ensemble mean (thick blue line) of daily pan-Arctic sea ice area anomalies relative to the daily climatology of the corresponding control ensemble means. The gray dashed lines show the intra-ensemble standard deviations in the control ensembles. All lines are presented as 15-day running means. (c-f) Probability distribution functions of (c-d) February-September and (e-f) August-September mean pan-Arctic sea ice area for (blue) the rare event simulation, (black) the control ensembles (N=6000 trajectories) and (red) the 3000-year control run. The vertical lines show the mean of the distributions. The black and blue values indicate the smallest February-September and August-September mean sea ice area value in the control and rare event ensemble simulations respectively.*

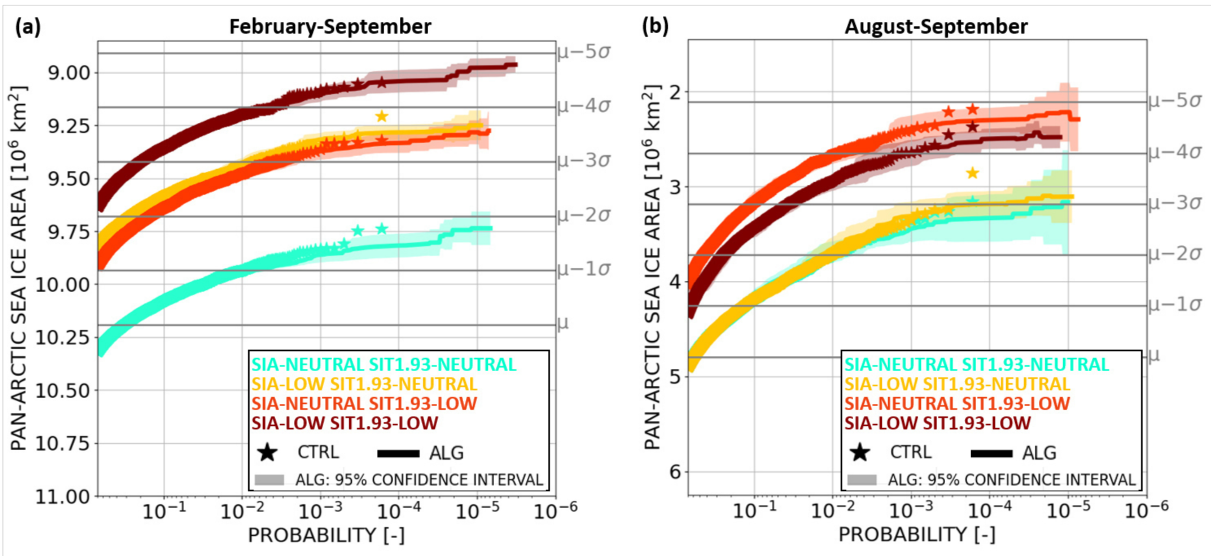

Figure R1-3: *Probabilities (x-axes) of (a) February-September and (b) August-September mean sea ice area equal or smaller than a given threshold (y-axes) as a function of four different initial conditions for (stars) the N=6000 trajectory control ensembles (i.e. ten realizations are merged into one single ensemble) and (solid lines) the average over four to ten algorithm estimates. The shading are 95% confidence intervals derived from the t-distribution of the four to ten estimates (see Supplementary Information S4 of the revised manuscript). Note that the y-axes are displayed in reverse order. In the legend, "SIA" and "SIT1.93" indicate the state of the January-February mean anomaly of the pan-Arctic sea ice area and the cumulative area with sea ice thickness equal or larger than 1.93 m respectively. The grey labels on the right of each panel show how many standard deviations a sea ice area value is below the mean of the 3000-year control run.*

Reviewer. *L355: Are the coupled modes of internal variability sufficiently realistic to make this conclusion for the real world or more complex models? I would suggest comparing this with CMIP6 preindustrial simulations with reasonable sea ice mean states. I would assume that there are no ice free summers in these simulations. Also on this point, is there anyone suggesting that this is possible.*

Reply and action. While the coupled modes of atmospheric circulation variability are well represented in PlaSim, their impact on low sea ice states is certainly underestimated due to the lack of sea ice dynamics. Therefore, and due to the relatively small number of initial conditions used in this study, transferring the conclusions about a sea ice-free Arctic to the real world is challenging. Moreover, the question about the probability of a sea ice-free Arctic is more meaningful for set of initial conditions taken e.g. from the period 2040-2049 from a climate change projection than for a stationary pre-industrial climate. Consequently, we decided to remove the conclusions about the sea ice-free Arctic in the manuscript.

Reviewer. *How does this fit in the context of these studies: https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019GL082947 and https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020GL088335*

Reply. Qualitatively consistent with the findings with comprehensive climate models, PlaSim captures a predictability barrier-like structure in spring where late summer sea ice area anomalies are much larger correlated with the sea ice area in July than with the sea ice area in May (cf. Figure R1-4 and Blanchard-Wrigglesworth et al., [2011], Tietsche et al. [2014], Day et al. [2014]). Accordingly, the analysis of the four experiments in our manuscript indicates that late winter sea ice area does not provide a source of predictability for extremely negative late summer sea ice area anomalies. While Bonan et al. [2019] indicates that the skill of sea ice volume to predict late summer sea ice area increases from winter/spring to summer, the presented correlations between late summer sea ice area and late winter/early spring sea ice volume are still larger than zero (Figure 1(a) in Bonan et al. [2019]). The fact that we find a source of predictability of late summer sea ice area anomalies in the late winter cumulative area with sea ice thickness equal or larger than 1.93 metres is therefore not inconsistent with Figure 1(a) in Bonan et al. [2019]. Moreover, Figure 4 of the original manuscript shows that extremely negative late summer sea ice area anomalies relative to the control ensemble means can be anticipated by negative sea ice volume anomalies prior to the development of negative sea ice area anomalies. Therefore, spring sea ice volume provides an additional source of predictability of extremely negative late summer sea ice area anomalies relative to the long control run in addition to the predictability associated with the late winter initial condition. This is consistent with Bushuk et al. [2020] and Bonan et al. [2019] demonstrating that early summer sea ice volume is a substantially better predictor of late summer sea ice area than spring sea ice volume.

Action. We mention in section 2.1 that PlaSim captures a spring predictability-like structure in the persistence of pan-Arctic sea ice area anomalies and refer to a lagged correlation analysis applied to sea ice area anomalies in the control run presented in Figure S1 in the Supplementary Information of the revised manuscript (Figure R1-4). We also include the above-mentioned explanation in section 4.1.

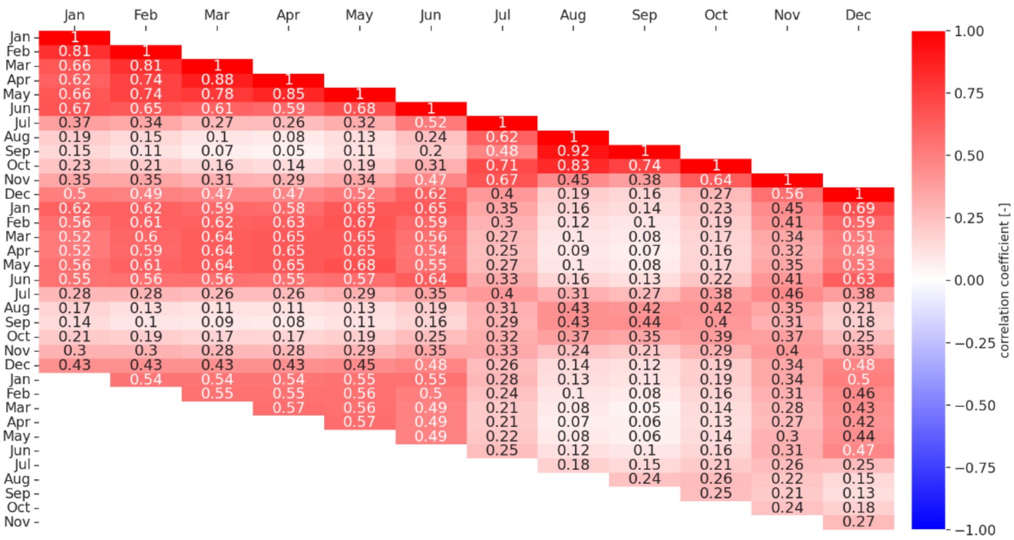

__Figure R1-4:__ Model years 501-3500 of PlaSim-LSG control run [Sauer et al. 2024]: (a) Monthly mean pan-Arctic sea ice area anomalies in the PlaSim-T21-LSG control run with fixed pre-industrial greenhouse gas conditions. Lagged correlations applied to monthly mean pan-Arctic sea ice area anomalies in PlaSim-T21-LSG with increasing lag from top to bottom up to a lag of 23 months.

References

Wang, J., Zhang, J.,Watanabe, E., Ikeda, M., Mizobata, K.,Walsh, J. E., Bai, X., and Wu, B.: Is the Dipole Anomaly a major driver to record lows in Arctic summer sea ice extent?, Geophysical Research Letters, 36, L05 706, https://doi.org/10.1029/2008GL036706, 2009.

Ogi, M., Rysgaard, S., and Barber, D. G.: Importance of combined winter and summer Arctic Oscillation (AO) on September sea ice extent, Environmental Research Letters, 11, 034 019, 2016.

Zhang, J., Lindsay, R., Schweiger, A., and Steele, M.: The impact of an intense summer cyclone on 2012 Arctic sea ice retreat, Geophysical Research Letters, 40, 720–726, https://doi.org/10.1002/grl.50190, 2013.

Chevallier, M., F. Massonnet, H. Goessling, V. Guémas, and J. Thomas: The Role of Sea Ice in Sub-seasonal Predictability: *SUB-SEASONAL TO SEASONAL PREDICTION: The Gap Between weather and Climate Foreasting*, A. W. Robertson and F. Vitart, eds., Elsevier Inc., 201-221, 2019

Sauer, J., Demaeyer, J., Zappa, G., Massonnet, F., and Ragone, F.: Extremes of summer Arctic sea ice reduction investigated with a rare event algorithm, Climate Dynamics, pp. 1–19, https://doi.org/10.1007/s00382-024-07160-y, 2024.

Lindsay, R. W., Zhang, J., Schweiger, A., Steele, M., and Stern, H.: Arctic Sea Ice Retreat in 2007 Follows Thinning Trend, Journal of Climate, 22, 165–176, https://doi.org/10.1175/2008JCLI2521.1, 2009.

Parkinson, C. L. and Comiso, J.: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm, Geophysical Research Letters, 40, 1356–1361, https://doi.org/10.1002/grl.50349, 2013.

Guemas, V., Doblas Reyes, F., Germe, A., Chevallier, M., and y Melia, D. S.: September 2012 Arctic sea ice minimum: Discriminating between sea ice memory, the August 2012 extreme storm, and prevailing warm conditions [in “Explaining Extreme Events of 2012 from a Climate Perspective”], Bull. Amer. Meteor. Soc, 94, S20–S22, 2013.

Kapsch, M.-L., Skific, N., Graversen, R. G., Tjernström, M., and Francis, J. A.: Summers with low Arctic sea ice linked to persistence spring atmospheric circulation patterns, Climate Dynamics, 52, 2497–2512, https://doi.org/10.1007/s00382-018-4279-z, 2019.

Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C. M., and DeWeaver, E.: Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations, Journal of Climate, 24, 231–250, https://doi.org/10.1175/2010JCLI3775.1, 2011.

Tietsche, S., Day, J. J., Guemas, V., W. J. Hurlin, S. P. E. K., Matei, D., Msadek, R., Collins, M., and Hawkins, E.: Seasonal to interannual Arctic sea ice predictability in current global climate models, Geophysical Research Letters, 41(3), 1035–1043, https://doi.org/https://doi.org/10.1002/2013GL058755, 2014.

Day, J. J., Tietsche, S., and Hawkins, E.: Pan-Arctic and regional sea ice predictability: Initialization month dependence, Journal of Climate,27, 4371–4390, https://doi.org/10.1175/JCLI-D-13-00614.1, 2014

Bonan, D. B., Bushuk, M., and Winton, M.: A spring barrier for regional predictions of summer Arctic sea ice, Geophysical Research Letters, 46(11), 5937–5947, https://doi.org/10.1029/2019GL082947, 2019

Bushuk, M., Winton, M., Bonan, D. B., Blanchard-Wrigglesworth, E., and Delworth, T. L.: A mechanism for the Arctic sea ice spring predictability barrier, Geophysical Research Letters, 47, 1–13, https://doi.org/10.1029/2020GL088335, 2020.

Author comments

Ensemble design for seasonal climate predictions: Studying extreme Arctic sea ice lows with a rare event algorithm

J.Sauer, F. Massonnet, G. Zappa, F. Ragone

submitted to Earth System Dynamics (ESD) within the Special issue "Theoretical and computational aspects of ensemble design, implementation, and interpretation in climate science" (EGUSPHERE-2024-3082)

January 24, 2025

Reviewer #2

Reviewer. *The authors present a novel simulation ensemble design to analyze statistics of sea ice area lows in the Arctic using a rare event algorithm. The work is an extension of Sauer 2024 and demonstrates the ability to look at seasonal sea ice predictions conditioned on particular initial winter sea ice states. Seven sets of ensemble simulations are generated from seven initial states taken from a long pre-industrial control simulation of the PlaSim-LSG coupled climate model. For each ensemble, 600 trajectories from February 1 to Sepember 30 are run with perturbed initial conditions for both a control and rare event algorithm case with resampling time of 5 days. The rare event ensembles increase the number of extreme lows compared to the control simulations and the analysis demonstrates that the initial state of the ice has a strong influence on the distribution of minimum area. For two particular trajectories, atmosphere and ocean conditions are investigated for contributions to the extreme ice area loss.*

Overall, the manuscript is well-written and introduces a novel approach for estimating the probablity of extreme sea ice loss over a season given the winter ice state. The method has the potential to provide new insight into physical drivers of Arctic sea ice loss. The main limitation of the study is the relatively simple sea ice model configuration in combination with the coarse resolution of the coupled climate model. Given that the sea ice model is purely thermodynamic, neglecting drift and deformation, with only a single layer, it is not clear how broadly the conclusions would hold for more sophisticated models. The authors do address these limitations in the manuscript and state their plans to apply the methodology to a more complex model.

Reply. We thank the reviewer for the important comments allowing to improve our manuscript. We will reply to these comments in the following. For the sake of clarity, we will copy/paste the individual comments before our replies and actions.

Reviewer. *1. The introduction focuses on the historical sea ice loss events in 2007 and 2012, which are clearly a motivation for the research in this manuscript. However, the analysis in the paper cannot provide insight into questions of how anthopogenic greenhouse gas forcing or sea ice dynamical drivers may have contributed to these events. The introduction also includes a good bit of general descriptive text that matches (at times word-for-word) parts of the introduction in Sauer 2024. It would be worthwhile to edit the introduction to provide a more specific motivation of the analysis in the paper and clarify what questions relating to sea ice loss are being addressed.*

Reply. The purpose of this study is to present an approach to study the probabilities/predictability and physical drivers of low sea ice states in a seasonal climate prediction set-up, thereby exploiting ensemble simulations starting from single initial conditions and reducing the computational cost required to sample a sufficiently robust statistics on the extremes. For example, the approach used in this study can be used in the future to quantify the impact of the initial condition in specific years (such as in 2007 and/or 2012) to the amplitudes and probabilities of extreme sea ice lows. Such an approach would primarily inform about how predictable the 2012 event would have been based on the knowledge of the initial condition. Learnings about the relative contribution of global warming to the probabilities of these events would only be indirect and qualitative (indirect via the impact of preconditioning of enhanced open-water-formation due to the ongoing winter sea ice thinning [Lindsay et al. 2009; Parkinson and Cosimo 2013; Guemas et al. 2013]).

The experimental design could, however, easily adapted to make more quantitative statements about the relative contribution of global warming to extreme sea ice loss. This could for example be achieved by initializing an ensemble with late winter initial conditions sampled from e.g. the period 2016-2025 (i.e. one initial condition per year times 30 perturbed replicates yielding an ensemble of 300 trajectories) and comparing the statistic with another ensemble initialized with data e.g. 1986-1995 drawn from a historical simulation. That is the reason why we include the discussion about the relative contribution of global warming to the probabilities of the extremes.

Action. As our study is not designed to provide conclusions about the relative contribution of global warming to extreme sea ice loss events, we slightly shorten this part in the introduction. We keep this for the conclusion.

Reviewer. *Lines 145-155, there is a discussion of the model's ability to capture the average seasonal cycle of sea ice area with enough fidelity for the analysis. Given that the winter volume of ice is more predictive of the seasonal minimum sea ice area, it would be good to include information on how well the model matches expected seasonal ice volume and spatially resolved ice thickness.*

Reply. We agree. The winter to spring pan-Arctic sea ice volume matches relatively well the climatology obtained from the Pan-Arctic Ice Ocean Modeling and Assimilation System (PIOMAS) (cf. Figure R2-1(left) and page 207 of Chevallier et al. [2019]). Regarding the spatial distribution of the annual mean sea ice thickness, substantial differences are visible between PlaSim and PIOMAS, with the former overestimating sea ice thickness in the central Arctic Ocean and in the Barents Sea and underestimating sea ice thickness northern of Greenland, compared to the latter (cf. Figure R2-1(right) and page 207 of Chevallier et al. [2019]). The regions of overestimated sea ice thickness correspond to regions relevant to extremely negative late summer sea ice area anomalies shown in Figure 4 of the main manuscript. Thus, the trajectories with the most extremely negative sea ice concentration anomalies available from the ensembles show no substantial sea ice concentration anomalies around the North Pole, while strongly negative sea ice concentration anomalies are present in the Barents and Kara Seas (Figure 4). In observations, the Barents and Kara Seas are sea ice-free in late summer [Kapsch et al. 2019]. While the overestimation of sea ice in the Barents and Kara Seas in PlaSim compared to observations likely contributes to an overestimation of the amplitudes of extremely negative pan-Arctic sea ice area anomalies compared to the real world, an opposing effect is to be expected from the positive sea ice thickness bias around the North Pole in PlaSim compared to observations.

Action. In the new manuscript**,** we mention the regional sea ice thickness biases in section 2.1 and discuss its implications for the results in section 5.

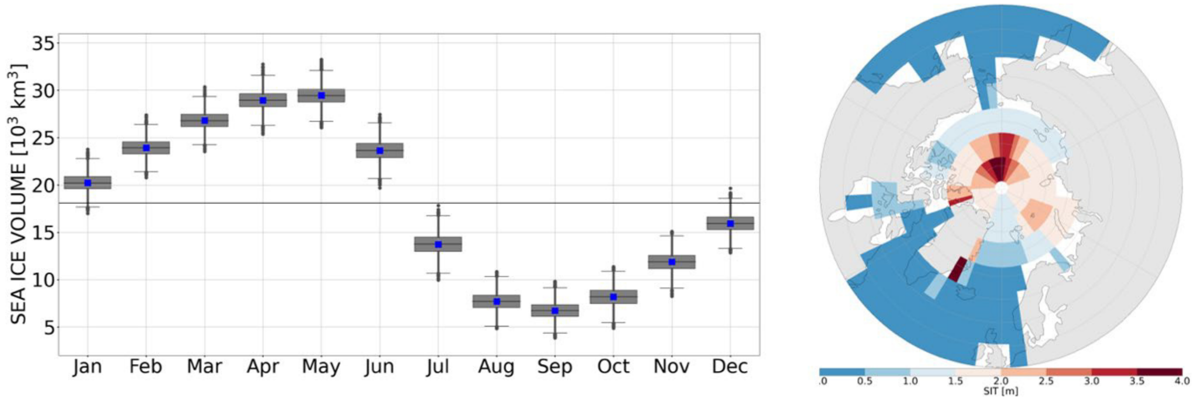

Figure R2-1: *Model years 501-3500 of PlaSim-LSG control run [Sauer et al. 2024]: (Left) Distributions of monthly mean pan-Arctic sea ice volume [10^3 km^3] with respect to the 3000 control run model years. The averages and medians are given by the blue squares and the horizontal lines in the boxes. The boxes denote interquartile ranges and the maximum whisker length is defined as 1.5 times the interquartile range. The horizontal gray line shows the annual mean pan-Arctic sea ice volume. (Right) Annual mean sea ice thickness [m].*

Reviewer. *3. The table in Figure 1a could use additional explanation. What is the composite analysis (Line 243) and what is the significance of the thickness bins (Line 245)?*

Reply. We agree. We explain this more carefully in the revised manuscript. The composite analysis is based on a 3000-year control run with PlaSim-LSG presented in Sauer et al. [2024]. The 3000-year control run is subdivided into five 600-member ensembles (initial conditions shifted by five years respectively). Afterwards, we estimated for each of the ensembles the mean anomalies of different quantities (e.g. the cumulative area of sea ice thickness equal or larger than a given threshold) for those years in which the February-September mean of August-September mean sea ice area anomaly is below -2.5 standard deviations. We normalized those anomalies by the standard deviations of the control run to be able to compare the predictive skill of for example sea ice thickness with area-related quantities. For the choice of the bins, we used values of the studies of Lindsay et al. [2008] and Chevallier et al. [2012]. The remaining values are interpolated.

Action. We explain the estimation of statistical significance of the composite mean values in the Supplementary Information S4 of the revised manuscript.

Reviewer. *4. Figure 2 c-f. It is notable that the rare event ensembles have a bimodal probability distribution. However, this is not mentioned in the discussion in the text. Is there a physical reason for this behavior?*

Reply. Two factors may lead to a bimodality in the rare event algorithm distribution of the target observable. One possibility is that the bimodality is purely related to sampling and does not reflect a real model bimodal distribution. A second possibility is that two types of distinct dynamical pathways lead to a “physical” bimodality of the distribution obtained with the rare event algorithm (see Ragone and Bouchet [2021]).

In the case of extremes of pan-Arctic sea ice area reduction, a bimodality in the distribution of the sea ice area may be favoured by the particular physical properties of the coupled atmosphere-sea ice-ocean system in the Arctic, including preconditioning and amplifying feedback mechanisms. Potentially, there is a region-dependent threshold in spring sea ice thickness where the late summer open-water-formation-efficiency is substantially increased in trajectories with a sea ice thickness below that threshold compared to trajectories with a sea ice thickness above the threshold. Once an anomalously large amount of open water area is formed, local feedback mechanisms (sea ice-albedo and water vapour) would likely maintain or even further amplify the anomalously large area of open water.

In the "SIA-NEUTRAL SIA-NEUTRAL" experiment, i.e. where both initial SIA and SIT1.93 are taken to be aligned with climatology, the peak of the probability density function at an August-September mean sea ice area value of 2.9 10^6 km^2 is related to hundreds of replicated trajectories originating from two ancestors (Figure 2(a,e) of the original manuscript and Figure R2-2). The bimodality may therefore simply emerge from the fact that trajectories leading to sea ice area values corresponding to the local minimum of the probability density function are undersampled in this particular simulation. Out of ten "SIA-NEUTRAL SIA-NEUTRAL" rare event simulations, a pronounced bimodality as in Figure 2(e) of the original manuscript occurs in a total number of three realizations (cf. Figure 2(e) of the original manuscript and Figure R2-3). The August-September mean sea ice area values at which the two peaks and the local minimum of the probability distribution functions occur, however, vary from one realization to the next (cf. Figure 2(e) of the original manuscript and Figure R2-3). Therefore, we do not expect that the real model probability density function has a bimodality explained by two peaks at specific sea ice area values.

In order to get a first understanding, nevertheless, of the physical particularity of the lower peak in “REALIZATION #1” of the "SIA-NEUTRAL SIA-NEUTRAL" rare event simulation, we compute composite mean anomaly values of different quantities both 1) conditional on the occurrence of trajectories belonging to the lower peak (magenta in Figure R2-2) and 2) for all trajectories leading to an August-September mean sea ice area anomaly smaller than -2.5 control ensemble standard deviations excluding the lower peak (referred to as “moderately extreme trajectories”; black in Figure R2-2; note that these composites based on the output with the algorithm are computed as described in Sauer et al. [2024]; Figure R2-4). Trajectories corresponding to the lower peak of the probability distribution function are characterized by a clearly more pronounced sea ice thinning during late spring than the moderately extreme trajectories before the two types of trajectories start to differentiate with respect to the sea ice area in July (Figure R2-4(a-b)). Furthermore, substantial differences in the thermodynamic states of the atmosphere between both types of trajectories occur. The lowest bundle shows stronger and more persistent positive atmospheric water vapour and cloud cover anomalies and thus a stronger downward longwave radiative forcing during spring than the moderately extreme trajectories (Figure R2-4(d,e,f)). Likewise, enhanced downward sensible heat flux anomalies in April-May and positive integrated water vapour anomalies in late summer, the latter potentially occurring as a response to enhanced sea ice loss, exclusively occur in the trajectories of the lowest peak of the distribution.

Action. We mention and formulate a hypothesis the regarding the bimodality in section 5 of the new manuscript, including Figures R2-2, R2-3 and R2-4 in the Supplementary Information. A more precise understanding about the physical mechanisms favouring the bimodalities and about systematic characteristics among different realizations of the experiment requires, however, a more sophisticated analysis than presented here and is beyond the scope of this paper.

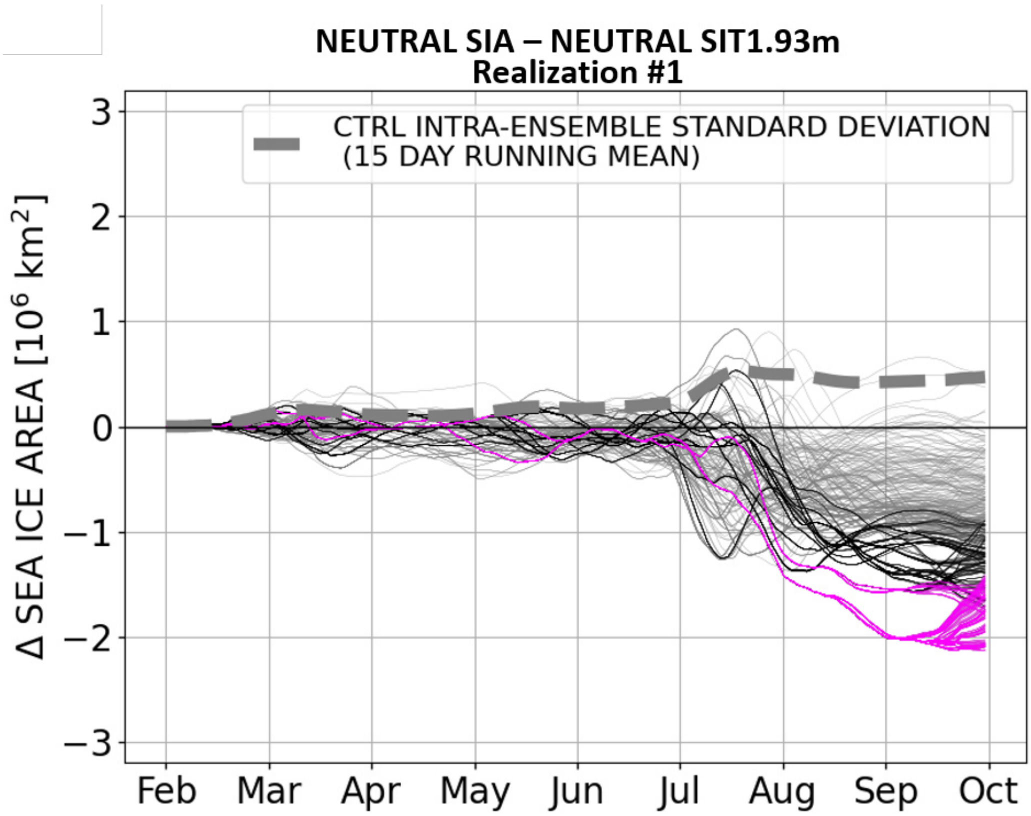

Figure R2-2: *Rare event ensemble simulations initialized on 01 February 1930 of the control run. 15-day running mean of daily pan-Arctic sea ice area anomalies [10^6 km^2] relative to the control ensemble mean. The magenta lines show trajectories leading to the lower peak of the bimodal distribution of Figure 2(e) of the main manuscript and the dark black lines the trajectories leading to August-September mean sea ice area anomalies below -2.5 control ensemble standard deviations excluding the lowest peak. The gray dashed line shows the 15-day running mean of the intra-ensemble standard deviation.*

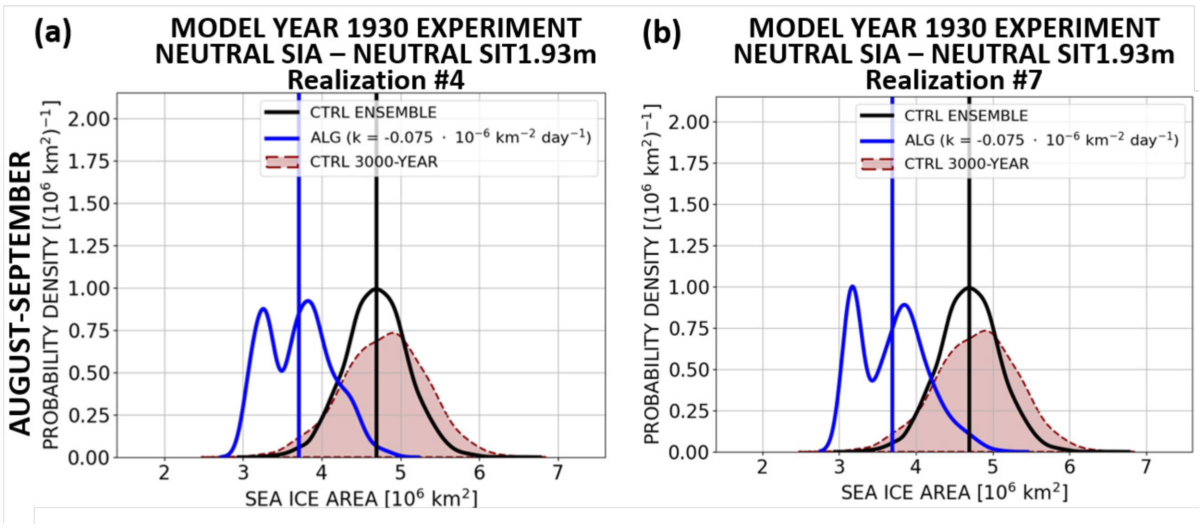

Figure R2-3: *Ensemble simulations initialized on 01 February 1930 of the control run. August-September mean pan-Arctic sea ice area for (blue) the rare event simulation, (black) the control ensembles (N=6000 trajectories) and (red) the 3000-year control run for (a) rare event simulation realization #4 and (b) rare event simulation realization #7. The vertical lines show the mean of the distributions. The black and blue values indicate the smallest August-September mean sea ice area value in the control and rare event ensemble simulations respectively.*

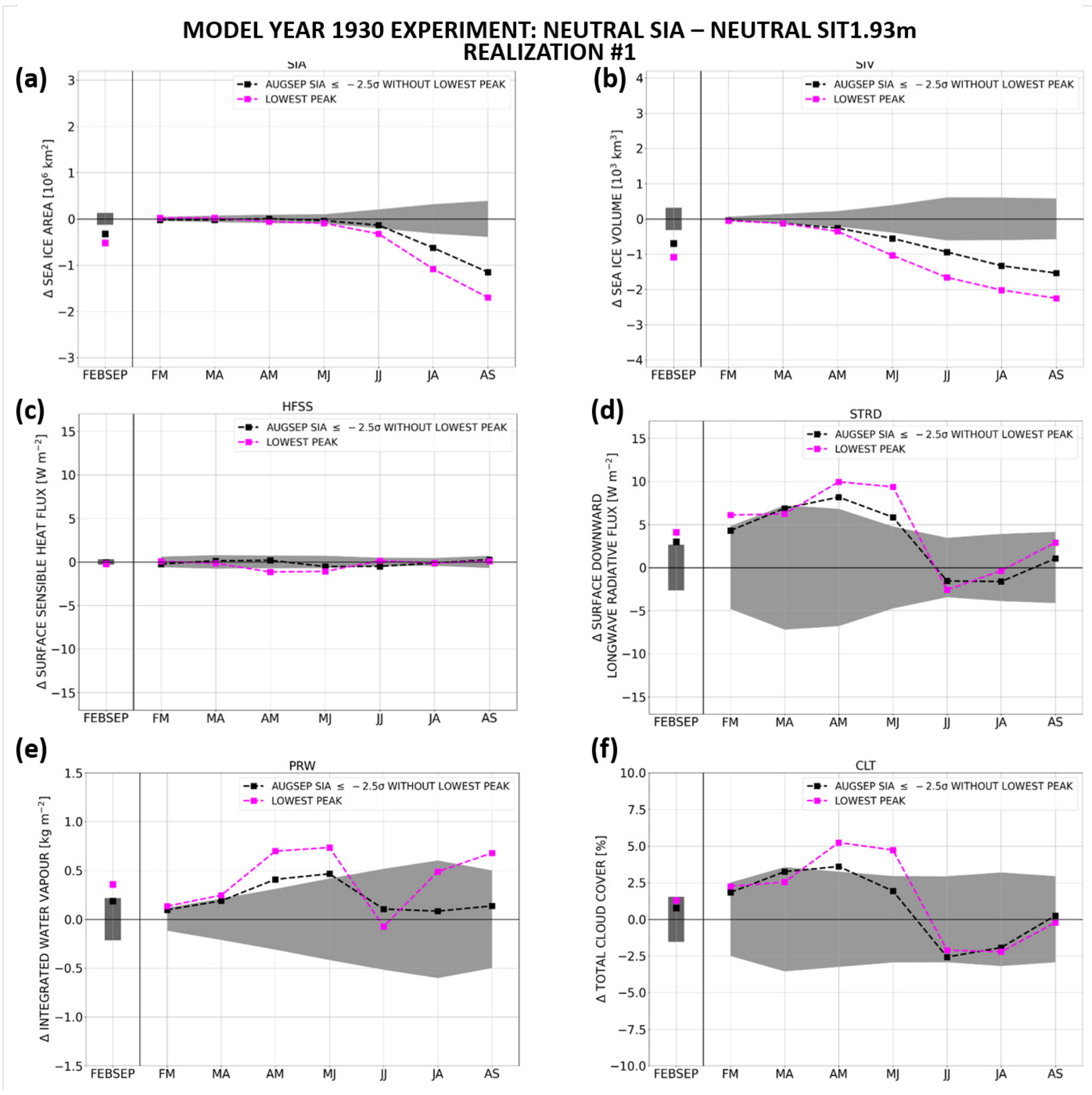

Figure R2-4: *Rare event simulation realization #1 starting from a neutral winter initial condition. Composite mean anomalies of different quantities for all trajectories leading to an August-September mean sea ice area value belonging to (magenta) the "lower peak" of the rare event algorithm probability distribution function of Figure 2(e) and (black) for all trajectories leading to an August-September mean sea ice area anomaly equal or smaller than -2.5 standard deviations of the control ensemble excluding the magenta trajectories (see Figure S4 for the selected trajectories). The composites are computed using the estimator in Equation (3) of the main manuscript (see Sauer et al. [2024] for details about the computation of composites with the output of the algorithm). (a-b) Pan-Arctic sea ice (a) area and (b) volume. (c-d) Spatially averaged composite mean anomalies over all ocean grid boxes northern of 70° N. (c) Surface sensible heat flux and ([W m^{-2}]; positive upwards) (d) downward longwave radiative flux anomalies [W m^{-2}] (Direction-independent absolute values of the downward and upward fluxes are considered, i.e., a positive (negative) anomaly indicates a radiative flux that is stronger (weaker) in magnitude than the climatology). (e) Integrated water vapour [kg m-2] and (f) total cloud cover [%] anomalies. (a-f) Shading indicates the intra-ensemble standard deviation of the control ensemble and anomalies are computed with respect to the climatology of the control ensemble.*

Reviewer. Line 460: cloduiness -> cloudiness

Reply. Thank you for spotting this typo.

References

Lindsay, R. W., Zhang, J., Schweiger, A., Steele, M., and Stern, H.: Arctic Sea Ice Retreat in 2007 Follows Thinning Trend, Journal of Climate, 22, 165–176, https://doi.org/10.1175/2008JCLI2521.1, 2009.

Parkinson, C. L. and Comiso, J.: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm, Geophysical Research Letters, 40, 1356–1361, https://doi.org/10.1002/grl.50349, 2013.

Guemas, V., Doblas Reyes, F., Germe, A., Chevallier, M., and y Melia, D. S.: September 2012 Arctic sea ice minimum: Discriminating between sea ice memory, the August 2012 extreme storm, and prevailing warm conditions [in "Explaining Extreme Events of 2012 from a Climate Perspective"], Bull. Amer. Meteor. Soc, 94, S20–S22, 2013.

Chevallier, M., F. Massonnet, H. Goessling, V. Guémas, and J. Thomas: The Role of Sea Ice in Sub-seasonal Predictability: *SUB-SEASONAL TO SEASONAL PREDICTION: The Gap Between weather and Climate Foreasting*, A. W. Robertson and F. Vitart, eds., Elsevier Inc., 201-221, 2019

Kapsch, M.-L., Skific, N., Graversen, R. G., Tjernström, M., and Francis, J. A.: Summers with low Arctic sea ice linked to persistence spring atmospheric circulation patterns, Climate Dynamics, 52, 2497–2512, https://doi.org/10.1007/s00382-018-4279-z, 2019.

Sauer, J., Demaeyer, J., Zappa, G., Massonnet, F., and Ragone, F.: Extremes of summer Arctic sea ice reduction investigated with a rare event algorithm, Climate Dynamics, pp. 1–19, https://doi.org/10.1007/s00382-024-07160-y, 2024.

Lindsay, R.W., Zhang, J., Schweiger, A. J., and Steele, M. A.: Seasonal predictions of ice extent in the Arctic Ocean, Journal of Geophysical Research: Oceans, 113, 1–11, https://doi.org/10.1029/2007JC004259, 2008.

Chevallier, M. and Salas-Melia, D.: The role of sea ice thickness distribution in the Arctic sea ice potential predictability: A diagnostic approach with a coupled GCM, Journal of Climate, 25, 3025–3038, https://doi.org/10.1175/JCLI-D-11-00209.1, 2012.

Ragone, F. and Bouchet, F.: Rare event algorithm study of extreme warm summers and heatwaves over Europe, Geophysical Research Letters, 48, e2020GL091 197, https://doi.org/10.1029/2020GL091197, 2021.

Author comments

Ensemble design for seasonal climate predictions: Studying extreme Arctic sea ice lows with a rare event algorithm

J.Sauer, F. Massonnet, G. Zappa, F. Ragone

submitted to Earth System Dynamics (ESD) within the Special issue "Theoretical and computational aspects of ensemble design, implementation, and interpretation in climate science" (EGUSPHERE-2024-3082)

January 24, 2025

Reviewer #3

Reviewer. *This paper takes an important next step in the sampling of extreme low Arctic sea ice extent. Rare event simulation techniques are used to study the influence of initial conditions on the rare event of interest.*

Reply. We thank the reviewer for the important comments allowing to improve our manuscript. We will reply to these comments in the following. For the sake of clarity, we will copy/paste the individual comments before our replies and actions.

Reviewer. *A similar investigation on the influence of initial conditions on an extreme event, albeit using a very different method (action minimization), is performed in Plotkin et al. 2019, which should be mentioned more explicitly.*

Reply and action. Yes, we agree. We explicitly mention the study of Plotkin et al. 2019 in lines 100-102 of the revised manuscript.

Reviewer. *My main concern about the paper is that the RES runs are performed only once for every initial condition. Variance estimators are also not applied, so the sampling uncertainty in RES cannot be quantified. Nevertheless, statements about the tails of the RES ensembles are made that could be the result of statistical fluctuations. For example, reference is made to "probabilities of less than 0.001%" and "the trajectory with the most extremely negative sea ice area anomaly obtained with the rare event algorithm". It is apparent in Fig.2 that the most extreme realisations in the RES runs originate from just a few ancestors, so there may be a lot of sampling variability involved here. This uncertainty should be investigated, or the statements should be weakened, perhaps as an investigation of specific storylines that may not be general.*

Reply. We agree that the trajectories leading to the most extremely negative late summer sea ice area anomalies among the given ensembles originate from just a few ancestors (Figure 2(a,b) of the original manuscript). Therefore, the amplitudes of the most extremely negative sea ice area anomalies available within an ensemble and the estimated probabilities of exceeding the amplitudes of these extremes may substantially vary from one realization of the experiment to the next.

To account for this issue, we produced nine additional realizations of the control and rare event ensemble simulations originating from the “SIA-NEUTRAL SIT1.93-NEUTRAL”, “SIA-NEUTRAL SIT1.93-LOW”, “SIA-LOW SIT1.93-NEUTRAL” and “SIA-LOW SIT1.93-LOW” initial conditions respectively in addition to the existing ones of the original manuscript (i.e. we produced a total of 4 x 2 x 10 ensemble simulations with 600 trajectories running from February-September). This allows us to obtain a better estimate of the probabilities of extremely negative sea ice area anomalies compared to the original manuscript, and to derive 95% confidence intervals around the estimates of sea ice area anomaly amplitudes for a set of probability levels. For a given probability level, we derive the confidence intervals from $M \in \{4,5,6,7,8,9,10\}$ t-distributed sea ice area amplitude estimates available from M rare event simulations. The variable sample size emerges from the fact that the smallest probabilities reached by a rare event simulation varies from one realization to the next.

The two trajectories studied in section 4 of the original manuscript lead to August-September mean sea ice area values of $2.82 \cdot 10^6$ km^2 and 2.07 10^6 km^2, and we assessed the conditional probabilities of observing a sea ice area below those values as smaller than 0.001%. The probability-amplitude curves obtained from the average over the M experiments show that our probability estimate was even too conservative, i.e., the two trajectories studied in the original manuscript had even smaller probabilities than indicated in the manuscript.

Compared to the original manuscript, the confidence intervals allow us to compare on a statistically more robust basis the amplitudes and probabilities of late summer sea ice area between the four experiments. Based on four initial conditions, the results suggest that the amplitudes and probabilities of negative late summer sea ice area anomalies depend on the late winter cumulative area of sea ice thickness equal or larger than 1.93 metres, but not on the late winter sea ice area.

Actions.

- Section 2.2: We updated the description of the experimental design. We are now studying four instead of seven initial conditions and we are producing ten realizations instead of one realization of each control and rare event simulation. Table 1 is updated accordingly.
- Figure 2 of the original manuscript is updated by Figure R3-1: For the baseline of the statistics, we show the probability distribution functions for the sea ice area obtained from a 6000 trajectory "master" control ensemble with all ten control realizations merged together. Likewise, the original realization of the "NEUTRAL SIA – LOW SIT1.93m" experiment is replaced by "REALIZATION #9" where the most extreme late summer sea ice low is more extreme than in the original realization.
- Figure 3 of the original manuscript is updated by Figure R3-2 showing probability estimates including confidence intervals based on multiple realizations of the rare event simulations.
- The text in section 3.1 and in section 5 is updated according to the new data and according to the new probability estimates.
- Table 2 is reduced from seven to four initial conditions.
- Section 4: Analysis of the physics for the "NEUTRAL SIA – LOW SIT1.93m" experiment: "REALIZATION #9" replaces the orignal "REALIZATION #1" in Figures 4-7. The text is adapted according to the new realization.
- A description of the computation of the confidence intervals is given in the Supplementary Information S4 of the revised manuscript.

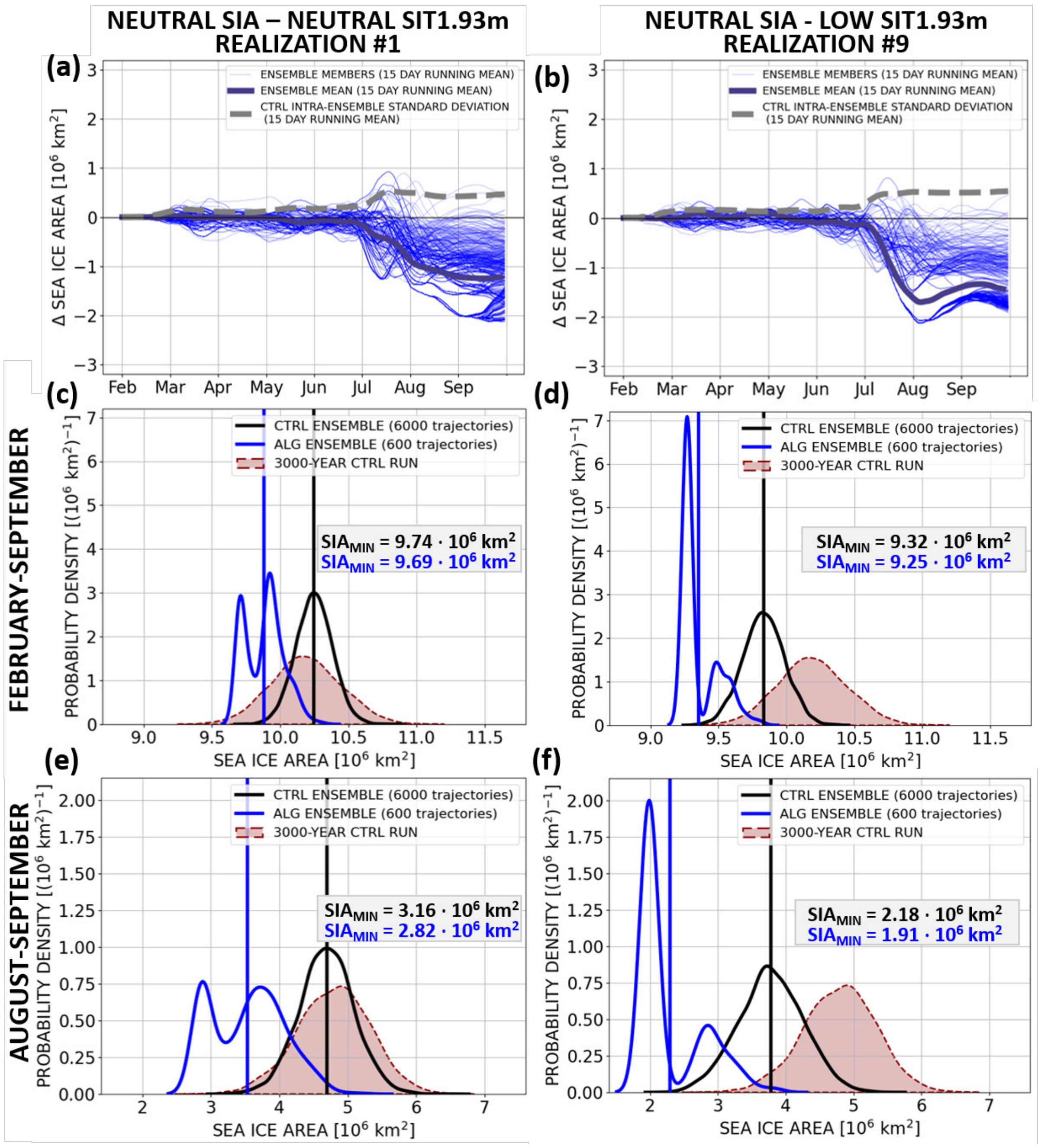

Figure R3-1: *Ensemble simulations initialized on 01 February (a,c,e) 1930 and (b,d,f) 1037 of the control run. (a-b) Rare event simulations: N=600 trajectories (thin blue lines) and ensemble mean (thick blue line) of daily pan-Arctic sea ice area anomalies relative to the daily climatology of the corresponding control ensemble means. The gray dashed lines show the intra-ensemble standard deviations in the control ensembles. All lines are presented as 15-day running means. (c-f) Probability distribution functions of (c-d) February-September and (e-f) August-September mean pan-Arctic sea ice area for (blue) the rare event simulation, (black) the control ensembles (N=6000 trajectories) and (red) the 3000-year control run. The vertical lines show the mean of the distributions. The black and blue values indicate the smallest February-September and August-September mean sea ice area value in the control and rare event ensemble simulations respectively.*

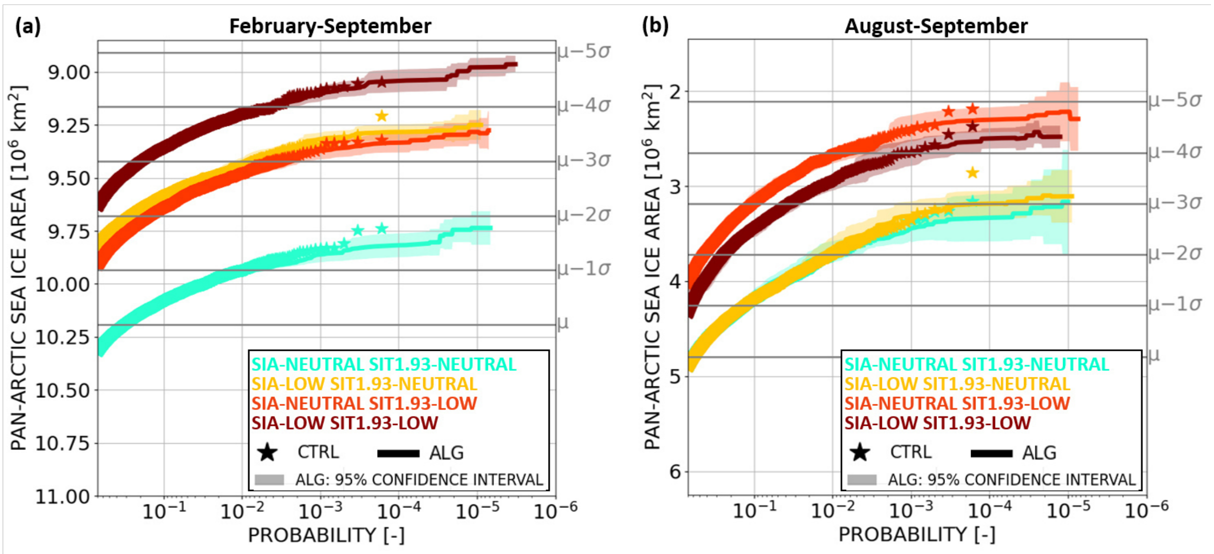

Figure R3-2: *Probabilities (x-axes) of (a) February-September and (b) August-September mean sea ice area equal or smaller than a given threshold (y-axes) as a function of four different initial conditions for (stars) the N=6000 trajectory control ensembles (i.e. ten realizations are merged into one single ensemble) and (solid lines) the average over four to ten algorithm estimates. The shading are 95% confidence intervals derived from the t-distribution of the four to ten estimates (see Supplementary Information S4 of the revised manuscript). Note that the y-axes are displayed in reverse order. In the legend, "SIA" and "SIT1.93" indicate the state of the January-February mean anomaly of the pan-Arctic sea ice area and the cumulative area with sea ice thickness equal or larger than 1.93 m respectively. The grey labels on the right of each panel show how many standard deviations a sea ice area value is below the mean of the 3000-year control run.*

Reviewer. *As the aim of the paper is to investigate the influence of climatic drivers, I wonder why RES sampling is performed from a single initial condition (IC) and not from a set of extreme states, for example all those that are the exceeding +/- 2 sigma levels in Fig. 1 right. This would make statements about the influence of drivers more robust against the choice of one particular IC.*

Reply. We agree that initializing an ensemble with a set of different initial conditions drawn from a chosen subspace (e.g. all initial conditions exceeding +/-2 sigma levels) would be an interesting approach to study climatic drivers in a more robust way against the influence of one particular single initial condition. The purpose of the manuscript, however, is to present an experimental design to study probabilities and drivers of low sea ice states in a seasonal climate prediction set-up. For example, the approach used in this study can be used in the future to quantify the impact of the initial condition in specific years (such as in 2007 and/or 2012) to the amplitudes and probabiltiites of extreme sea ice lows. This is the reason why we initialized the ensembles with single initial condiitons.

Reviewer. *line 41: do not managed -> do not manage*

Reply and action. Thank you. The paragraph is, however, written in past tense and we will use "did not manage".

Reviewer. *line 110: less than 0.001% -> a probability of less than 0.001%*

Reply and action. Thank you. We have adapted this in the text.

Reviewer. *Fig. 1: the spacing in the x-axis label of "delta pan-Arctic sea ice area" makes the hyphen look like a minus sign.*

Reply and action. We agree. We replaced Figure 1 from the original manuscript by Figure R3-3.

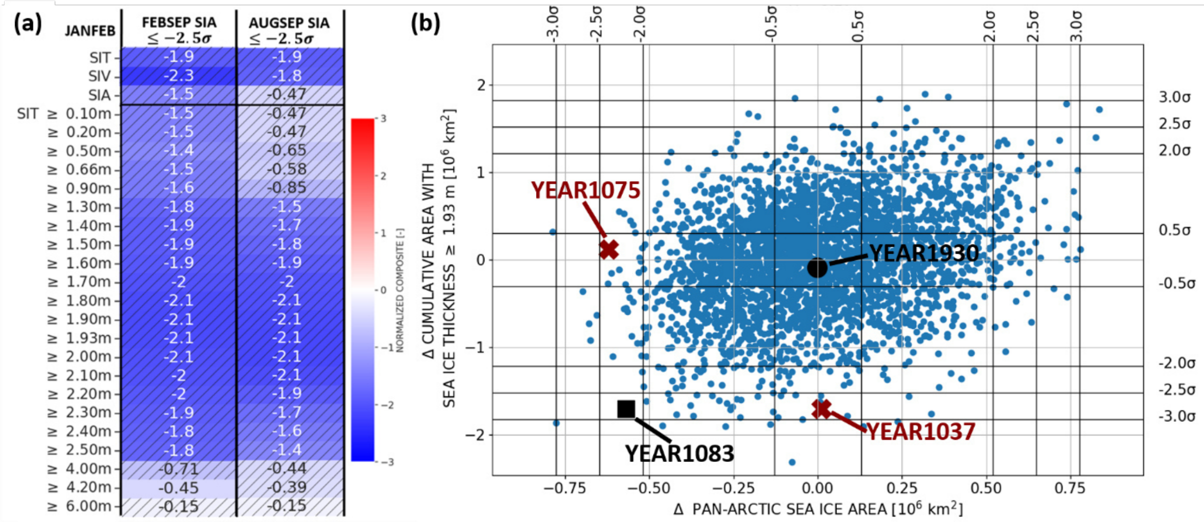

Figure R3-3: *PlaSim-LSG 3000-year control run (Sauer et al. [2024]): (a) Normalized mean anomalies of January-February mean quantities conditional on extreme negative (left) February-September (FEBSEP) and (right) August-September (AUGSEP) mean pan-Arctic sea ice area anomalies equal or smaller than -2.5 standard deviations. "SIV", "SIA" and "SIT" are the pan-Arctic sea ice volume, sea ice area and mean sea ice thickness. "SIT ≤ threshold" are anomalies in the cumulative area with sea ice thickness equal or larger than a critical threshold. Hatching denotes statistical significance at the 5% level assessed from a two-sided t-test applied to five composite estimates after subdividing the 3000-year control run into five 600-member ensembles (see Supplementary Information S4). (b) Scatter plot of January-February mean anomalies of SIT1.93 vs. pan-Arctic sea ice area including the years from the selected initial conditions.*

Reviewer. line 313: in what unit is the computational cost expressed?

Reply. *Thank you for spotting the missing unit. We are writing"10^3-10^4 years".*

Reviewer. *Fig. 2: The text says simulations are performed until September, but plots continue until October.*

Reply. Yes, we removed the label "Oct" and we shortened the thick blue line of the ensemble mean such that it definitely stops before 01 October (Figure 3-1).

Reviewer. *Table 2: "control" is used to refer to two different types of runs. Could these be differentiated?*

Reply and action. We updated table 2 and its labelling to differentiate between "control run" and "control ensemble".

Reviewer. *Line 496: The most extreme negative sea ice anomaly in the RES ensemble is dependent on the RES ensemble size, hence quantitative statements about the % of impact are not very meaningful. A high percentile would be more informative here.*

Reply and action. Yes thank you for this advice. We replaced the sea ice area value of the most extremely negative sea ice area anomaly by the 1% percentile value computed as the average of the 1% percentile values obtained from the ten realizations with the rare event algorithm. For the computation of the percentiles with the algorithm, we weight the contribution of each trajectory to the percentile by the product of Z_k and the exponential factor shown in Equation (3) of the main manuscript. We updated table 2 according to these values.

Reviewer. *Line 635: minimizing -> minimization*

Reply. Thank you. We adapted this error in the revised manuscript.

References

Plotkin, D. A., Webber, R. J., O'Neill, M. E., Waere, J., and Abbot, D. S.: Maximizing simulated tropical cyclone intensity with action minimization, Journal of Advances in Modeling Earth Systems, 11, 863–891, https://doi.org/10.1029/2018ms001419, 2019.

Sauer, J., Demaeyer, J., Zappa, G., Massonnet, F., and Ragone, F.: Extremes of summer Arctic sea ice reduction investigated with a rare event algorithm, Climate Dynamics, pp. 1–19, https://doi.org/10.1007/s00382-024-07160-y, 2024.