https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

The impact of stochastic sea ice perturbations on seasonal forecasts

Kristian Strommen[1,2,*], Michael Mayer[1,3,*], Andrea Storto[4], Jonas Spaeth[1], and Steffen Tietsche[1]

[1]European Centre for Medium-Range Weather Forecasts, Reading, England and Bonn, Germany.
[2]Department of Physics, University of Oxford, United Kingdom.
[3]Department of Meteorology and Geophysics, University of Vienna, Vienna, Austria
[4]Institute of Marine Sciences (ISMAR), National Research Council of Italy (CNR), Rome, Italy.
[*]These authors contributed equally to this work.

Correspondence: Kristian Strommen (kristian.strommen@ecmwf.int)

Abstract. Sea ice ensemble forecasts can be highly underdispersive, meaning that the ensemble spread is notably lower than the average forecast error. One common strategy to address underdispersion is to add stochastic perturbations to the forecasts. We detail the implementation of a stochastically perturbed parameter (SPP) scheme for SI^3, the sea ice component used by the Integrated Forecast System (IFS), the forecast model used and developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). We then evaluate its impact on seasonal forecasts of northern hemisphere summer and winter. The inclusion of SPP is found to enhance ensemble spread for sea ice concentration (SIC) and sea ice thickness (SIT) forecasts by around 10% relative to an unperturbed forecast, which results in a better calibrated probabilistic forecast. Some small but robust changes to the mean state are also found, including a general decrease in the mean SIC and a redistribution of the winter ice from the central Arctic to the ice edge. These changes reduce or increase the mean bias depending on the region. Changes to the mean and spread of the sea ice result in changes to the mean and spread of air temperature up to at least 850hPa, altering the mean air temperature biases of the model. An apparent consequence of this is a significant increase in seasonal forecast skill of 500hPa geopotential (Z500) over the Euro-Atlantic domain in winter, which partially projects onto the North Atlantic Oscillation. We conclude that sea ice stochastic perturbations can be a valuable contribution to increased reliability of seasonal forecasts of the sea ice itself and can impact seasonal forecasts of the atmosphere at high and mid latitudes.

1 Introduction

In the context of numerical weather prediction (NWP) and climate modeling, a stochastic parameterization scheme is a way to insert random perturbations into the model in a controlled way. Such perturbations are appealing for two reasons. Firstly, no matter the model resolution, there will always be fluctuations in the system arising from unresolved sub-gridscale processes, and due to the non-linear nature of the governing equations they can still exert a non-zero net effect on the large-scale flow (Palmer, 2012). Even in cases where the physical parameterisation included to represent such processes is considered reasonable, there are inevitably free parameters in such parameterisations that cannot be tightly constrained from data, and the associated uncertainty can again ripple out to the large-scale. Unresolved sub-gridscale processes are therefore a potential source of model biases. Secondly, NWP models often produce ensemble forecasts that are underdispersive (i.e., the ensem-

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

ble distributions do not have enough spread relative to typical forecast errors). Stochastic parameterization schemes, or just stochastic schemes for short, address both of these concerns. The injected noise aims to represent the unresolved processes, with a practical effect being an often substantial increase in ensemble spread. The potential for such schemes to thereby substantially improve probabilistic weather forecasts is well known (Palmer et al., 2009; Sanchez et al., 2015; Berner et al., 2017; Lang et al., 2021; Xu et al., 2022). ECMWF has been using stochastic schemes for the atmosphere in operational ensemble forecasts since 1998 (Buizza et al., 1999). The original scheme, based on tendency perturbations, has recently been replaced by one based on parameter perturbations (Lang et al., 2021).

Many studies have shown that such schemes can also have a beneficial impact on climate model simulations, even in single simulations, due to the impact of the schemes on small-scale processes (Dawson and Palmer, 2015; Davini et al., 2017; Watson et al., 2017; Christensen et al., 2017; Juricke et al., 2017; Strommen et al., 2017). While most applications have focused on perturbing the atmospheric component of the model, stochastic schemes have also been implemented in the land (MacLeod et al., 2016), ocean (Juricke et al., 2017; Storto and Andriopoulos, 2021), sea ice (Juricke et al., 2013; Juricke and Jung, 2014) and dynamical core (Subramanian et al., 2019) components of NWP and climate models. The effect of combining multiple such schemes was documented in Strommen et al. (2019).

Here, we will focus on stochastically perturbed sea ice. The idea of adding stochastic perturbations to the sea ice is particularly attractive, as sea ice forecasts are often systematically underdispersive, even (in fact, especially) at the time of initialization (Dirkson et al., 2019, 2022; Peterson et al., 2022). This is likely related to the typically large biases models exhibit in sea ice (Notz and the SIMIP Community, 2020) and the importance of unresolved or poorly modelled sea ice and boundary layer processes. Juricke et al. (2014) found that ensemble spread on the sub-seasonal timescale can indeed be enhanced by including stochastic perturbations to the ice strength parameter P^*, and to a lesser extent on seasonal timescales. Juricke and Jung (2014) showed that such perturbations can also change the mean state in the polar regions, despite being mean-zero, which they ascribed to asymmetries in the response of sea ice to P^* perturbations. Their analysis generally highlighted some of the complex phenomena arising when adding sea-ice perturbations to a model, including potential differences in coupled versus uncoupled simulations (owing to the important effects of ice-ocean-atmosphere coupling) and the different timescales of the response (with a potentially multidecadal transient adjustment of the ice). Strommen et al. (2022) corroborated such changes to the mean and variability of sea ice, while also highlighting the prospect for stochastic perturbations to improve ice-ocean-atmosphere coupling and, as a result, seasonal timescale teleconnections from the Arctic to the mid-latitudes. As a result, it is possible that stochastic sea-ice perturbations could affect forecast skill, not just for sea-ice forecasts but also atmospheric forecasts. The potential importance of sea ice at driving midlatitude atmospheric variability has been highlighted in several past papers (Kretschmer et al., 2016; Wang et al., 2017).

The goal of this paper is to evaluate the implementation of a stochastic parameterization scheme to SI^3, the sea ice component used by the Integrated Forecasting System (IFS), the forecast model used and developed by the European Cdntre for Medium-Range Weather Forecasts (ECMWF). The scheme used is a modified version of the one described in Storto and Andriopoulos (2021), which is based on parameter perturbations, following the ideas of Juricke et al. (2013). The use of a parameter-perturbation approach ensures consistency with the approach taken to stochastic perturbations in the atmospheric component

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

of the IFS (Lang et al., 2021), which has the advantage of being conservative of tracers, momentum, and sea ice volume, unlike alternative perturbation schemes based on tendency or energy perturbations.

We will first of all detail the scheme used and report on experience gained from various sensitivity tests, which informed the choice of a default configuration. Then we will examine the impact of this default configuration on ensemble forecasts in the IFS, including the basic impact on the mean state and ensemble spread of sea ice concentration (SIC), sea ice thickness (SIT), and air temperature, as well as how forecast skill is affected in both the ice and atmosphere. We also include analysis concerning to what extent changes to the mean and spread of sea ice are related, which is relevant for applications where it is desirable to minimise mean state changes without compromising enhanced spread. Finally, we provide a discussion on what next steps are suggested by our results.

The focus of the analysis will be on the Arctic and northern midlatitude summer and winter. Impacts on Antarctica are included in Supporting Information.

2 Data and Methods

2.1 Model simulations and observational data

We examine the impact of stochastic sea ice perturbations using a development version of cycle 49R2 of the IFS. Cycle 49R2 will be used in the upcoming SEAS6 seasonal forecast system of ECMWF, which updates the former SEAS5 system (Johnson et al., 2019). It is also the cycle that will be used for the new ERA6 re-analysis product. Most of the differences between the development version we used and the final cycle 49R2 are small, and a counterpart set of simulations with SPP using a model set-up precisely matching SEAS6 system produced qualitatitvely similar results to the ones we show here. We therefore expect the impact of SPP to be the same in the final version of cycle 49R2.

Documentation for cycle 49R1 can be found here: https://www.ecmwf.int/en/publications/ifs-documentation. Cycle 49R2 differs from 49R1 primarily by coupling to version 4 of NEMO and the new SI^3 sea ice model (Gurvan et al., 2019; Vancoppenolle et al., 2023). SI^3 is a state-of-the-art sea ice modelling system including a subgrid-scale ice thickness distribution with five categories, prognostic sea ice salinity, simulation of melt ponds, and better representation of snow on sea ice. An informal discussion of the use of NEMO-SI^3 with the IFS can be found in Keeley et al. (2024).

In our experiments, the IFS is run using the Tco199 cubic octahedral grid, corresponding to a grid spacing of around 50km at the equator, with 137 levels in the vertical. NEMO-SI^3 is run with a horizontal resolution of $0.25°$, with 75 levels in the vertical. Initial conditions in the atmosphere use ERA5 reanalysis (Hersbach et al., 2020), while initial conditions in the ocean and sea ice make use of a close-to-final version of the new ORAS6 ocean/ice reanalysis developed at ECMWF (Zuo et al., 2025); the differences with the final version are deemed unimportant for the present work. Both the IFS and NEMO-SI^3 are run using single-precision arithmetic, and the IFS has the stochastically perturbed parameter scheme of Lang et al. (2021) active.

With this model configuration, we generate two 50-member seasonal hindcasts covering the winter and summer seasons of 1993-2023. The first is our 'control' run (referred to as CTRL), and the second differs from this control only in its inclusion of stochastic sea ice perturbations (referred to as SPP). These are described in Section 3. For both configurations, a 50-member

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

ensemble is initialised on November 1st and May 1st for every year between 1993 and 2023 inclusive, and run for 6 months. In our analysis we will primarily be comparing the CTRL and SPP forecasts for the northern hemisphere winter (December-January-February, DJF) and summer (June-July-August, JJA) seasons.

For verification of sea ice concentration, we use the Global Sea Ice Concentration climate data record (SMMR/SSMI/SSMIS) from EUMETSAT OSI-SAF (https://osi-saf.eumetsat.int/products/osi-450-a) until 2020 and its extension OSI-430-a (https://osi-saf.eumetsat.int/products/osi-430-a) from 2021 onwards. For verification of atmospheric variables, we use ERA5 reanalysis.

2.2 Metrics and statistical methods

Let $x_{t,k}$ denote an ensemble forecast at lead-time t with $k = 1, \ldots, K$ ensemble members, and let y_t denote the verifying observations. Let $\overline{x_t}$ denote the ensemble mean forecast. Given N such distinct forecast, we evaluate the forecast skill using several metrics. The ensemble spread is the standard deviation σ_{x_k} computed over all members. The root-mean-square-error (RMSE) is

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_t (\overline{x_t} - y_t)^2}. \tag{1}$$

The spread-error ratio is the RMSE divided by square root of the average ensemble variance (computed using the sample variance formulation with 1 degree of freedom), with an additional correction factor to address the finite ensemble size (Roberts and Leutbecher, 2025):

$$\text{Spread/Error} = \sqrt{\frac{K+1}{K}}\,\frac{\sqrt{\frac{1}{N}\sum_t \sigma^2_{x_k,t}}}{RMSE}. \tag{2}$$

If the forecast is reliable in a probabilistic sense, i.e., if ensemble members and verification are exchangeable, then the expected value of this ratio is 1 (Fortin et al., 2014). Values close to 1 therefore indicate high reliability and values close to 0 indicate low reliability. For an additional measure of probabilistic forecast skill, we use the continuous ranked probability score (CRPS); see, e.g., Hersbach (2000). The CRPS for a single forecast at time t is

$$\text{CRPS}_t = \frac{1}{K}\sum_k |x_{t,k} - y_t| - \frac{1}{2K^2}\sum_i\sum_j |x_{t,i} - x_{t,j}|, \tag{3}$$

and the overall CRPS is CRPS_t averaged over all cases t. Informally, the CRPS is a generalisation of the mean-squared error which assesses the performance of the ensemble as a whole, as opposed to just of the ensemble mean. Lower values indicate better performance. In this paper a primary focus is on assessing improvements to the forecast distribution when including SPP, and hence in analysis we show rather the bias-corrected CRPS. This is done by first computing the mean bias over all times t, $\frac{1}{N}\sum_t(\overline{x_t} - y_t)$, and then subtracting this from every ensemble member before computing the CRPS. Finally, we will also measure forecast skill using the Pearson correlation between $\overline{x_t}$ and y_t.

To test statistical significance of changes between two forecasts, we use a two-tailed t-test when testing changes to the ensemble mean, and a Bartlett test to test changes to the ensemble spread. To assess significance of changes to the Pearson correlation, we use the "three-way-correlation test" from Siegert et al. (2017), which accounts for the presence of covariance

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

NEMO parameter name	Default value	Description
rn_pstar	1.5e+4	Ice strength parameter
rn_csrdg	0.5	Fraction of shearing energy contributing to ridging
rn_craft	5.0	Squeezing coefficient for rafting
rn_cio	7.5e-3	Ice-ocean drag coefficient
rn_cnd_s	0.35	Thermal conductivity of snow
rn_oiht	0.03	Ocean-ice heat transfer parameter
rn_time_fl	8.64e+5	Salinity restoring time scale for gravity (growth)
rn_time_gd	1.73e+6	Salinity restoring time scale for flushing in summer
rn_pnd_flush	0.01	Melt pond flushing efficiency

Table 1. Parameters being perturbed. Column 1: name of parameter in the NEMO namelist. Column 2: the default value of the parameter. Column 3: description of the parameter.

between two different forecasts of the same event. To test the significance of a single correlation we use a two-tailed t-test. These tests were primarily implemented using the python package scipy (Virtanen et al., 2020).

All data used has been regridded to the same regular 1 degree grid prior to analysis being carried out.

3 The stochastic perturbation scheme

The scheme used is the stochastically perturbed parameter (SPP) scheme implemented in the STOPACK package described in Storto and Andriopoulos (2021), extended to allow more sea ice parameters to be perturbed. A similar SPP scheme is already used in the atmospheric component of the IFS (Ollinaho et al., 2017; Lang et al., 2021). While STOPACK allows for several different types of parameter perturbation, the default one used here is based on log-normal perturbations obtained by drawing numbers from a random field evolving in space and time. This is similar to the scheme originally developed by Stephan Juricke and others (Juricke et al., 2013).

Concretely, a random noise field covering all gridpoints is first generated by repeatedly applying a Shapiro filter (Shapiro, 1970) to an initial random white noise field $\mathcal{W}$. The filter introduces lengthscales (i.e., correlations in space), the size of which is a function of the number of iterations: Figure S1 in Supporting Information shows this function. Call the resulting noise field $\mathcal{P}$. Then $\mathcal{P}$ evolves in time according to an autoregressive process with a lag of 1 day. At every time step, given a parameter p, a perturbed parameter $\hat{p}$ is then generated at each grid-point according to the formula

$$\hat{p} = p \cdot \exp(\xi_t), \tag{4}$$

where ξ_t is value of the random field

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

Table 1 gives a summary of which sea ice parameters are perturbed in the current implementation. The rn_pstar parameter is the same P^* parameter perturbed in Juricke et al. (2013) and follow-up work, including Strommen et al. (2022). All parameters except rn_csrd are perturbed log-normally with $\sigma = 0.5p$, where p is the default (i.e., unperturbed) value of the parameter. The rn_csrd parameter, which is a fraction, is perturbed using normally distributed noise by multiplying it by the value of the random noise field at that gridpoint; the perturbed value is then clipped to the range $[0, 1]$. We also explored perturbing the prognostic albedo value, as computed by NEMO, but this was ultimately not pursued due to the large impact such perturbations had on the mean ice state. Further exploration of how to generate more moderate albedo perturbations is of clear interest, given the potential for amplifying spread further.

The configuration described above was determined based on a number of sensitivity tests, which tested perturbing individual parameters versus multiple, varying σ and varying the spatial and temporal decorrelation scales. These tests suggested that each individual parameter, when perturbed, produced an enhancement in the ensemble spread with a particular spatial pattern and amplitude. Furthermore, when multiple parameters are perturbed, the resulting spatial pattern was approximately a combination of the individual patterns, with an amplitude which was higher then that achieved by only perturbing a single parameter. Examples of some of these single-parameter tests are shown in Figures S2 and S3; note that these sensitivity tests were performed with the earlier cycle 48R2 of the IFS. This motivated our decision to perturb a large number of parameters, since this resulted in a more uniform increase in spread across the Arctic/Antarctic domain. These sensitivity tests also suggested that the parameter perturbation that has the single biggest effect is the P^* perturbation in winter and the ice-ocean drag coefficient in summer (not shown). We found little sensitivity to the spatial and temporal decorrelation scales, but a larger spatial scale generally resulted in somewhat smoother changes to spread. Varying σ scales the overall amplitude of the spread up and down as expected. These tests motivated the choice of the relatively high value of $\sigma = 0.5$, chosen in order to get an appreciable increase in spread. We comment further on this in the Discussion.

A final comment, which may be of particular interest to anyone aiming to implement similar schemes in operational forecasts, is that when tested at a higher Tco319 resolution, there were occasional stability issues related to the build-up of sea ice at coastlines in the Arctic. These stability issues, which only appeared when the forecast was run for at least 3 months, were removed by tapering the stochastic perturbations near the coastlines, which is a feature provided by STOPACK.

4 Impact of SPP on ensemble mean and ensemble spread

4.1 Impact on the sea ice

Figure 1 shows how SPP affects the ensemble mean and ensemble spread of the seasonal mean SIC in boreal winter. The mean SIC is broadly reduced in the central Arctic, and increased around the ice edge, with the average change being a small net reduction. More detailed analysis (not shown) suggests this change is not large enough to constitute a change in the location of the ice edge, and is rather a compactification of the existing ice edge in CTRL. We discuss the reason for this mean state change in Section 5.2. The ensemble spread is enhanced almost everywhere, with the greatest change (of around 1-2%) seen around the ice edge, which is where virtually all the sea ice variability takes place in winter. The increased spread with SPP amounts

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

to a relative increase of around 10%. Figure 2 shows the equivalent impact in JJA. Here too SPP induces a net decrease of the mean SIC and an increase in ensemble spread almost everywhere. The increased spread is again roughly a 10% relative increase. Note that we have generally not highlighted gridpoints where the changes in the spread are statistically significant because the vast majority of the gridpoint changes seen with SPP are significant. The same is generally true for changes to the mean, with a few notable exceptions. The increased SIC in the Chuckhi and Bering seas in DJF are not significant ($p > 0.05$), and in JJA the only significant changes are the larger SIC decreases of 2-4% extending around the coastlines: these have been highlighted with stipling. The increased spread is significant ($p < 0.05$) at virtually every gridpoint for both seasons.

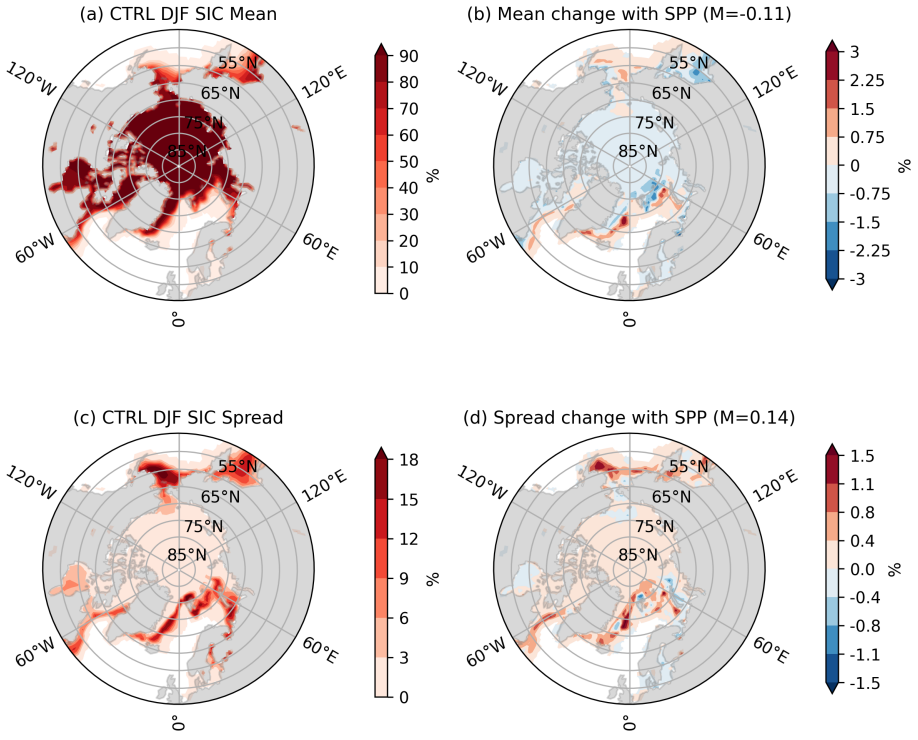

Figure 1. Impact of SPP on DJF SIC mean and spread. In (a): DJF SIC ensemble mean in CTRL. In (b): DJF SIC ensemble in SPP minus that of CTRL. In (c) and (d): the same but for DJF SIC ensemble spread. The value of M where it appears is the weighted area-average across all gridpoints in the domain where SIC is non-zero.

A very similar impact is seen in the southern hemisphere: there is a net decrease in the mean SIC and an increase in spread almost everywhere (see Figures S4 and S5). The consistent increase in ensemble spread is expected, but the consistent decrease in the mean is not. We discuss this further in Section 5.1. As will be seen later, the changes to the mean sometimes increase the model bias.

Figures 3 and 4 show the impact of SPP on sea ice thickness. There is again a consistent increase in SIT ensemble spread, amounting to a relative increase of around 10%. In summer there is no net change in the mean, but in winter there is a small

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

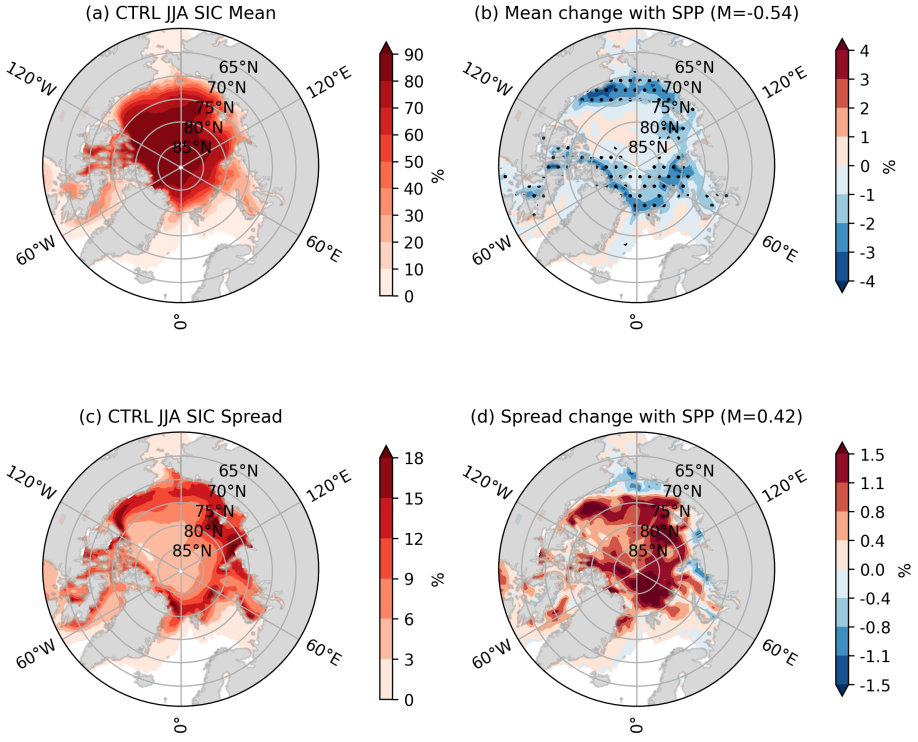

Figure 2. The same as Figure 1 but for JJA. Stipling in (b) highlights where changes are statistically significant ($p < 0.05$).

net increase. We discuss this further in Section 5.1. Corresponding results for the southern hemisphere are given in Figures S6 and S7.

Finally, Figure 5 shows the change in mean sea ice drift due to SPP, overlaid on the change in mean SIC. In DJF there is a clear suggestion that perturbations result in ice moving from the central Arctic to the ice edge, consistent with the changes to mean SIC. In JJA the changes in drift are far less clear, and give the impression that perturbations are resulting in random motions. The magnitude of the velocity changes are on the order of 20% relative to the baseline velocities in CTRL. We return to this in Section 5.1.

4.2 Impact on the atmosphere

Figure 6 shows changes to the mean and ensemble spread of T2M and 850hPa air temperature (T850) in winter. There is a broad warming across the central Arctic in winter, extending out to Hudson Bay, contrasted by a cooling along the ice edge of the Greenland, Barents, Kara and Chukchi seas. The direction of change is consistent at both surface and 850hPa, though only the change around Greenland, Barents and Kara is significant at 850hPa. A significant increase of ensemble spread is also seen in many regions of the Arctic, with a suggestion that some of the change is reaching the midlatitudes. The increased air temperature spread is an expected consequence of the increased SIC and SIT spread, as fluctuations in these quantities modulate

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

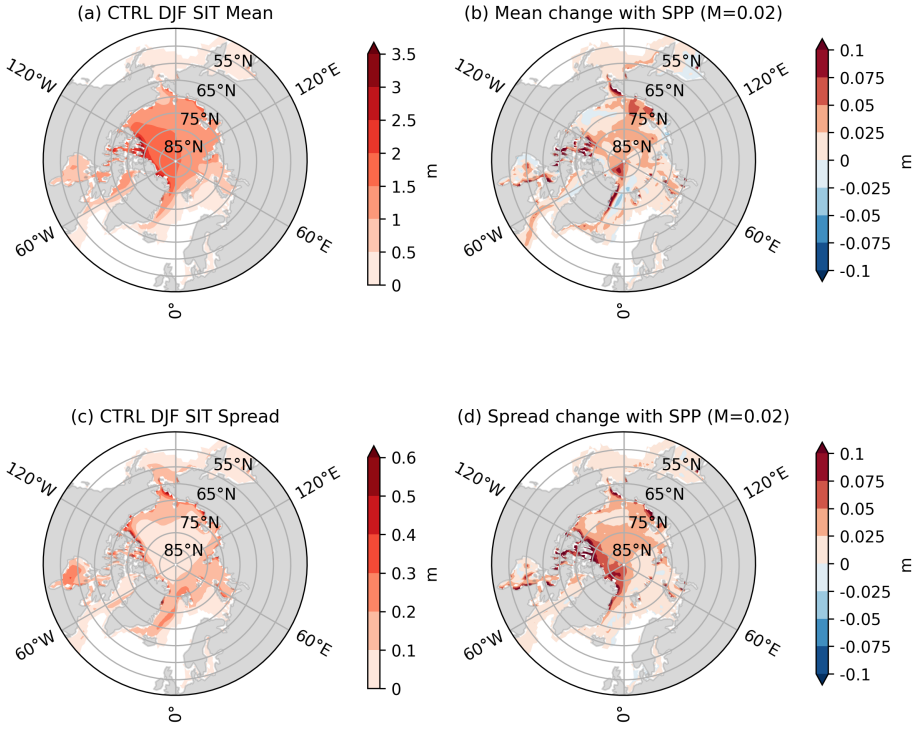

Figure 3. Impact of SPP on DJF SIT mean and spread. In (a): DJF SIT ensemble mean in CTRL. In (b): DJF SIT ensemble in SPP minus that of CTRL. In (c) and (d): the same but for DJF SIT ensemble spread. The value of M where it appears is the weighted area-average across all gridpoints in the domain where SIT is non-zero.

surface heat fluxes and albedo, thereby directly altering near-surface temperatures. The changes in summer air temperature are shown in Figure 7, and are notably smaller. This is unsurprising, since the Arctic surface air temperature in summer is tied to the freezing point and so there is little scope for the ice perturbations to affect the atmosphere. The changes to summer T2M spread appear to be closely connected to SIC and SIT changes (Figure 2 and 4), with increased T2M spread in regions where SIC/SIT decrease with SPP and conversely.

In winter, the increased mean temperatures seen in the central Arctic are consistent with the reduced SIC, as this reduction allows a greater flux of heat from the warm ocean to the cold atmosphere. The decreased mean temperatures around the ice edge are consistent with the increased SIC there, for the converse reason. In summer, the changes to SIC and air temperature do not align as clearly: there is a broad decrease of SIC around the perimeter of the Arctic circle, but while the temperature goes up around the Chukchi sea, it goes down in the Barents and Kara seas. To understand these changes further we now look at circulation changes. Figures 6 and 7 also show the mean change in 500hPa geopotential height (Z500) arising from SPP. A positive and undulating Z500 anomaly of around 2-3m extends across most of the Arctic in winter, while in summer the Z500 anomalies are small (less than 2m) everywhere except the Chukchi sea. To further visualise these circulation changes, we

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

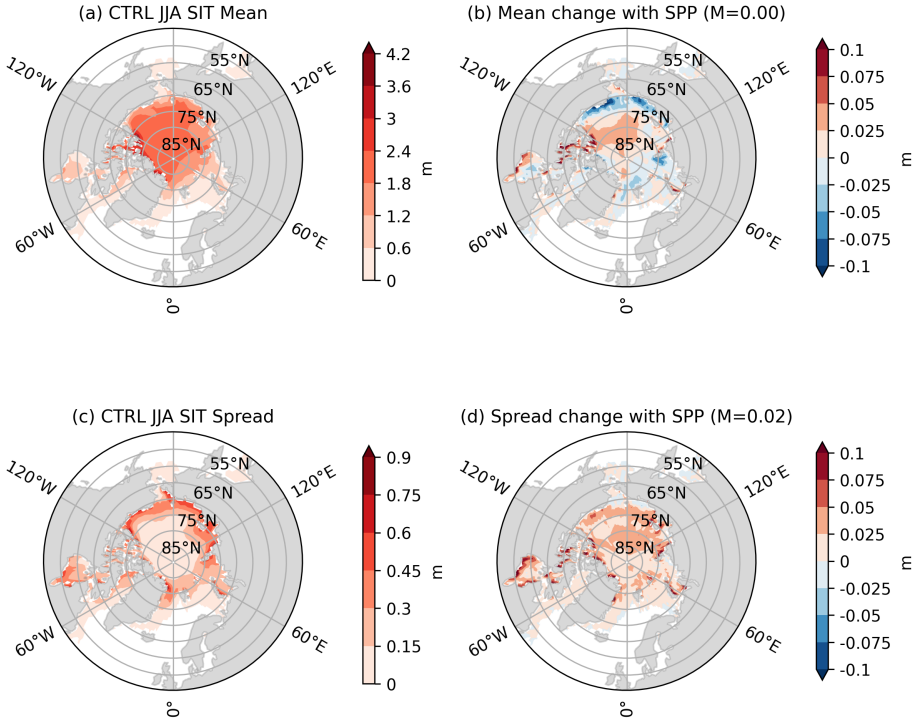

Figure 4. The same as Figure 3 but for JJA.

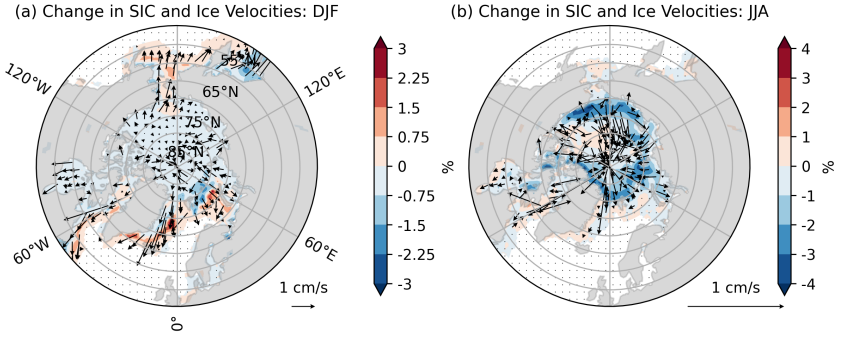

Figure 5. Impact of SPP on sea ice drift. In (a): the ensemble mean change in DJF SIC with SPP (filled contour) and the ensemble mean change in sea ice velocities with SPP (black arrows). In (b): the same but for JJA. The length of the arrows are proportional to the magnitude of the change. A reference arrow shows the length of a 1 cm/s change.

show in Figure 8 the change in the ensemble mean 700hPa wind vectors, overlaid onto the T850 ensemble mean change already shown. These suggest that circulation anomalies play an important role in determining the final temperature response to SPP. In particular, in winter we highlight the cyclonic anomaly centered on the Ural mountains, which results in cold Siberian air being

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

advected over the Barents and Kara seas, consistent with the net cooling. The cooling over Greenland in winter also appears to receive a contribution from this advection. Similarly, comparatively warm marine air is advected towards the Chukchi sea in summer, consistent with the net heating there. The cooling seen over Barents-Kara in summer appears to be unrelated to circulation changes, and may rather be a result of the wintertime cooling persisting through to summer via ocean memory.

To summarise, the inclusion of SPP leads to small but robust changes in the mean atmospheric circulation up to at least 850hPa, especially in winter. In winter we also see a significant increase in air temperature ensemble spread with SPP. The mean state changes can likely be understood using the same framework as Petoukhov and Semenov (2010), namely as a combination of a local convective response, due to changes in heat fluxes, and a baroclinic response, due to changes in temperature gradients. Because the net temperature change is generally negatively correlated with the local SIC change, these effects seem to generally be acting coherently in our case, with the circulation changes reinforcing the local thermodynamic responses.

We emphasise here that the only difference between the CTRL and SPP experiment is the inclusion of sea ice SPP. Thus, the prime causal driver of these changes is necessarily the sea ice, not the atmosphere, even if the latter determines the final response via feedbacks. Though we note it is also possible that part of the changes to SIC are related to snow conductivity perturbations, as these could alter surface temperatures which in turn alter SIC.

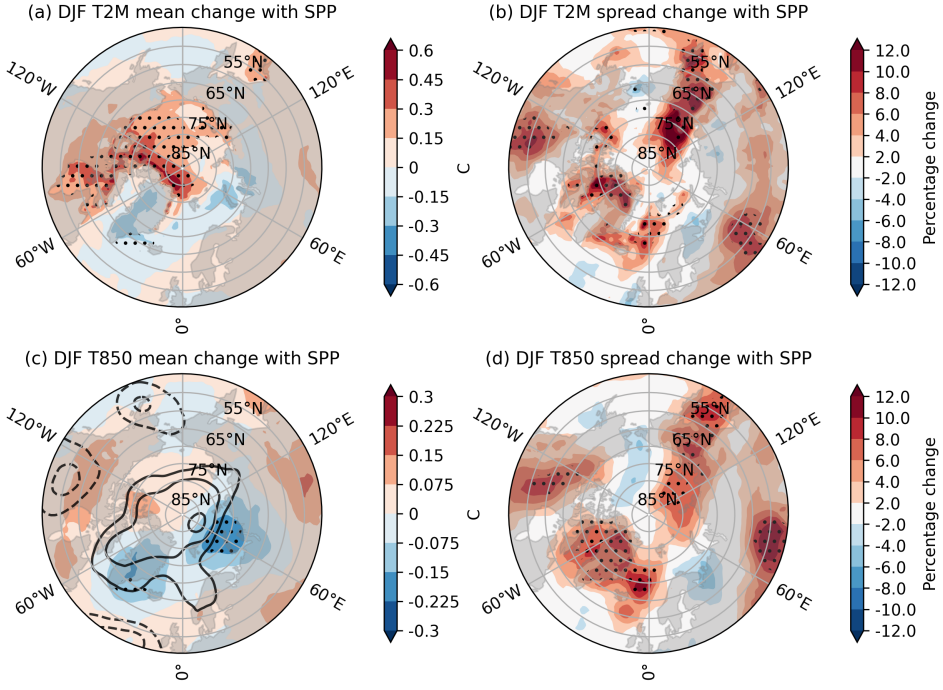

Figure 6. Impact of SPP on DJF air temperature. In (a): the DJF T2M ensemble mean in SPP minus that of CTRL. In (b): the DJF T2M ensemble spread in SPP minus that of CTRL. In (c) and (d): the same but for T850. In (c), the line contour shows the difference in the DJF Z500 ensemble mean, with contours drawn at $\pm 2, \pm 3$ and ± 4 meters.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

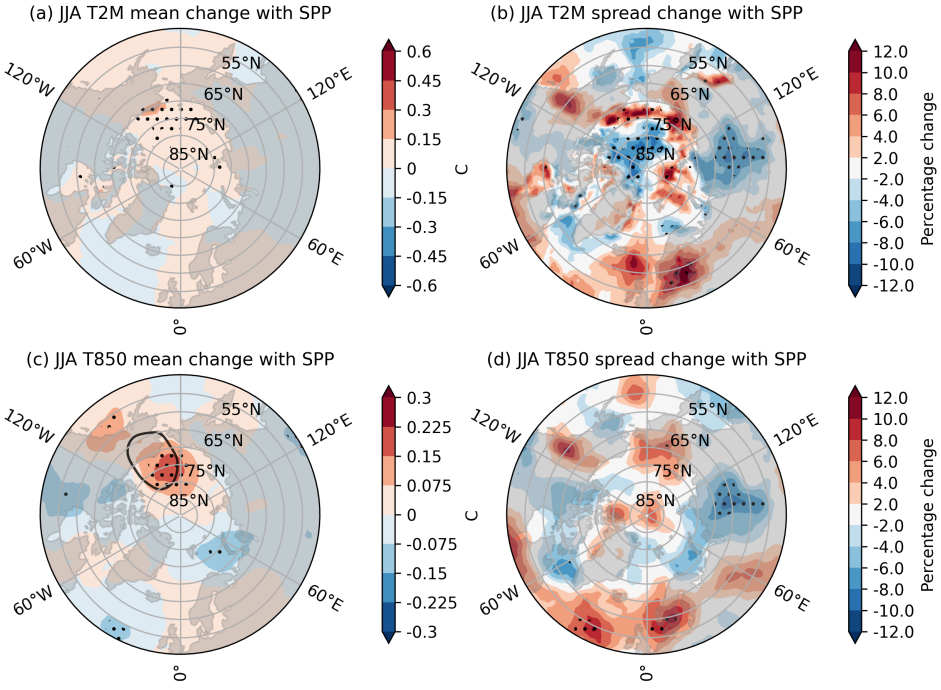

Figure 7. Impact of SPP on JJA air temperature. As Figure 6 but for JJA. The line contours are drawn at the same levels.

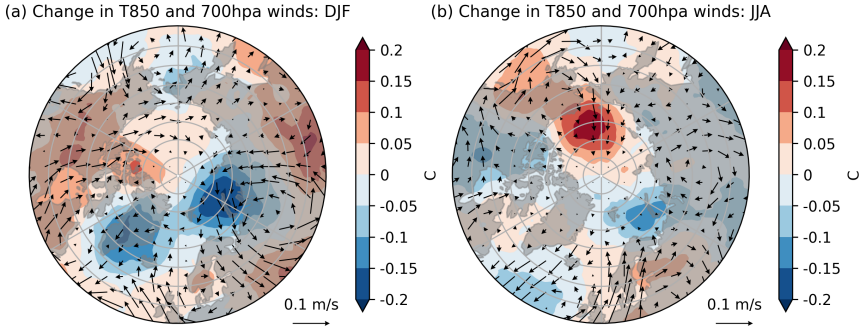

Figure 8. Impact of SPP on the circulation. In (a): the change (SPP minus CTRL) in DJF mean T850 (filled contour) with the change in DJF mean 700hPa winds overlaid as vectors. In (b): the same but for JJA. The length of the wind vectors is proportional to the magnitude of the change.

4.3 Differences between coupled and uncoupled simulations

Juricke and Jung (2014) showed that the effect of sea ice perturbations on the mean ice evolution over time can differ between coupled and uncoupled simulations, which they attributed to the interactive atmosphere exerting a net negative feedback on the ice. Here we highlight that there are also large differences in how perturbations affect ensemble spread. This has implications

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

for the use of SPP in uncoupled data assimilation which we return to in the Discussion. To assess these differences we focus on a smaller set of 11-member ensemble forecasts, each of which is initialised on January 1st of 2013 and run for one year. We run such forecasts using the following model configurations: coupled CTRL (no sea ice perturbations) forecast; coupled default-SPP forecast; coupled SPP forecast but using a different random field for each of the 9 perturbed parameters; uncoupled CTRL ocean-only forecast; uncoupled default-SPP ocean-only forecast. Note that here "uncoupled" means using prescribed ERA5 data as boundary forcing for a freely evolving NEMO4-SI^3 simulation. Furthermore, to make the coupled forecasts more directly comparable to the uncoupled forecasts we turned off both atmospheric SPP and atmospheric initial condition spread in the coupled forecasts, so that the only source of spread comes from the sea ice SPP. We did not run an uncoupled control (i.e., without sea ice perturbations), as this would be expected to have negligible ensemble spread, due to the shared atmospheric forcing and slow timescales of the ocean and ice. A final technical point is that the random fields used in the simulation with independent parameter perturbations are generated using Perlin noise (Perlin, 2002), rather than the iterated Shapiro filter. The impact of these differing noise generators were found to be minimal (not shown), but the Perlin noise generator is faster, and thus preferable when generating multiple random fields.

Figure 9(a) shows the ensemble spread for the Arctic-average SIC (60-90N) for the various forecasts. In the uncoupled default SPP forecast, substantial ensemble spread is added by the perturbations (relative to the baseline spread of ≈ 0 expected without any atmospheric spread or ice perturbations), peaking in August. When changing SPP to have independent random fields for each perturbed parameter, we see a notable and consistent increase in the spread, peaking at a roughly 60% increase in August relative to the default SPP spread. It is not a priori clear that independent perturbations should result in greater spread, as compensating effects from different physical parameters could in theory result in a net reduction in spread. However, it turns out that in this case introducing a greater number of perturbations results in much larger ensemble spread. However, in the coupled simulations the picture is far less clear. As expected, the baseline spread in the coupled CTRL is higher than in the uncoupled simulations due to the strong forcing from the atmosphere with its fast error growth, and the default SPP scheme amplifies this spread. The relative increase in spread is, however, much more modest in the coupled simulations, and there is no discernible difference between using the same random field or individual random fields.

These results were further confirmed using additional nudging experiments. In these experiments, coupled simulations were run in which the atmospheric vorticity, divergence and temperatures were relaxed back towards ERA5 using a 12-hour relaxation time scale. They can therefore be thought of as intermediate between a fully coupled simulation and a completely uncoupled simulation. In these nudged simulations (not shown), the baseline ensemble spread of Arctic averaged SIC sits between that of the uncoupled and coupled curves in Figure 9(a): the inclusion of SPP enhances spread considerably, and switching to independent perturbations considerably amplifies this effect.

We interpret these results as saying that sea ice perturbations are far more effective at generating spread in uncoupled ocean forecasts compared to coupled atmosphere-sea-ice-ocean forecasts. We hypothesise that this is due to two effects. Firstly, in the coupled simulations the ensemble spread is dominated by the atmospheric spread, meaning that ice perturbations will, relatively speaking, be less effective at amplifying spread compared to uncoupled simulations where ensemble spread is purely ice driven. In fact, the ensemble spread may be close to saturated in the coupled simulations. We computed that the ensemble

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

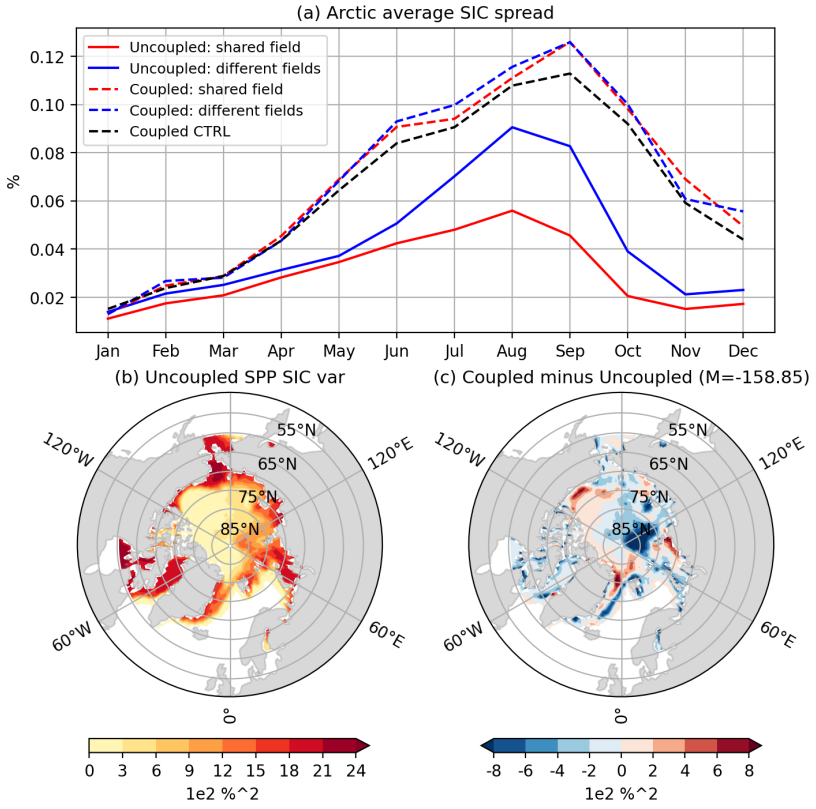

Figure 9. Impact of SPP in coupled vs uncoupled simulations. In (a): forecast ensemble spread of the Arctic-average SIC for each month of 2013 for coupled CTRL (black dotted), coupled regular SPP (red dotted), coupled SPP with independent random fields (blue dotted), uncoupled regular SPP (red) and uncoupled SPP with independent random fields (blue). In (b): the temporal variability, measured as the standard deviation of month-to-month SIC changes at each gridpoint, for the uncoupled SPP simulation. In (c): the difference in temporal variability between the coupled and uncoupled SPP simulation.

spread for the coupled CTRL in August is approximately the same as the standard deviation of detrended interannual variations in August. If it is assumed that the variation in monthly sea ice from one year to the next is an approximate upper bound on how much the sea ice is capable of changing within a single month, then this suggests that the spread may already be close to the maximum of what is achievable given the inherent timescales of the ice. Further increases to spread are likely to be difficult to achieve as a result. Secondly, as suggested in Juricke and Jung (2014), the atmosphere may be acting to damp the SPP-induced ice anomalies via negative feedbacks.

These two hypothetical effects can be qualitatively supported using idealized models of the coupled ice-atmosphere system based on stochastic differential equations: see, e.g., Strommen et al. (2022). We refer to such an idealized model as a "LIM", short for "linear inverse model", following typical conventions: see Supporting Text S1 for details. One can show that (a) coupled LIM simulations with a negative atmospheric feedback show greater overall ensemble spread compared to corre-

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

sponding uncoupled LIM simulations, but that (b) increased stochastic forcing is much more effective at increasing ensemble spread in uncoupled LIM simulations: see Supporting Text S1 and Figure S8. Both the negative feedback and the dominance of atmospheric forcing contribute to this effect in the LIM, though the latter effect seems to be more important in the range of parameter values we considered. It is challenging to quantitatively estimate the sign of the atmospheric feedback in NWP models, because the net feedback is a combination of both immediate local thermodynamic feedbacks and slower timescale dynamical adjustments which can originate further afield. However, in the LIM, the clearest signature of a negative atmospheric feedback is reduced temporal variability in coupled versus uncoupled simulations. In Figure 9(b) we show the month-to-month SIC variability across all ensemble members for the uncoupled SPP forecast, and in (c) we show how this temporal variability differs in the coupled SPP forecast. The sign is broadly negative, consistent with the atmosphere exerting a net negative feedback.

5 Mechanisms underlying the impact of SPP

In Section 4 we have seen that SPP alters the mean and spread of SIC and SIT in a particular way. The aim of this Section is to shed light on why we see these particular impacts. In Section 5.1, we show that changes to the ensemble mean and ensemble spread are strongly related, and in Section 5.2 we present a simple conceptual model which explains most of the observed features.

5.1 How are changes to the sea ice mean and spread related?

Figures 1 and 5 show a consistent picture of SPP resulting in ice being pushed outwards with associated changes to the mean. This suggests the stochastic perturbations may be at least partially exerting a consistent change to the system despite being random and mean zero. Furthermore, such systematic changes to the mean may not be independent of the changes to the ensemble spread, due to the strong asymmetries present in the sea ice. In this section we aim to shed further light on how the mean state, mean state changes and ensemble spread changes are related, where changes are always understood as changes due to SPP.

Figure 10(a) and (b) show scatter plots of seasonally averaged SIC mean state changes versus ensemble spread changes at each gridpoint in the Arctic for winter and summer. Thus, for DJF, we take the values at each gridpoint in 1(b) as the x-coordinates and the values in Figure 1(d) at the corresponding gridpoints as the y-coordinates. We also show histograms of the spread change across all gridpoints. In winter, while most points experience a small reduction in mean SIC and a small increase in the spread (red contour and histogram), there is clear evidence of non-Gaussian structure in the point cloud. In summer, this non-Gaussianity is less pronounced, but the majority of points show the same clear tendency for reduced mean and enhanced spread (red contour). We can understand the non-Gaussian features by highlighting points where the baseline CTRL mean SIC is high (SIC>80%) and low (SIC<20%): see Figure 10. It can be seen that these two categories of ice, the ice pack on the one hand and the ice edge zone on the other, occupy distinct 'regimes' as far as impacts of SPP go. In the ice pack regime, there is a visually clear negative correlation between changes to the mean and the spread: points where the mean go *down* more

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

experience a greater increase in spread. The linear correlation ranges between -0.5 to -0.7 in the two seasons ($p < 0.05$). On the other hand, in the ice edge zone the correlation is positive: points where the mean goes *up* more experience a greater increase in spread. The linear correlation is smaller in this case, ranging between 0.3 to 0.4 in the two seasons ($p < 0.05$).

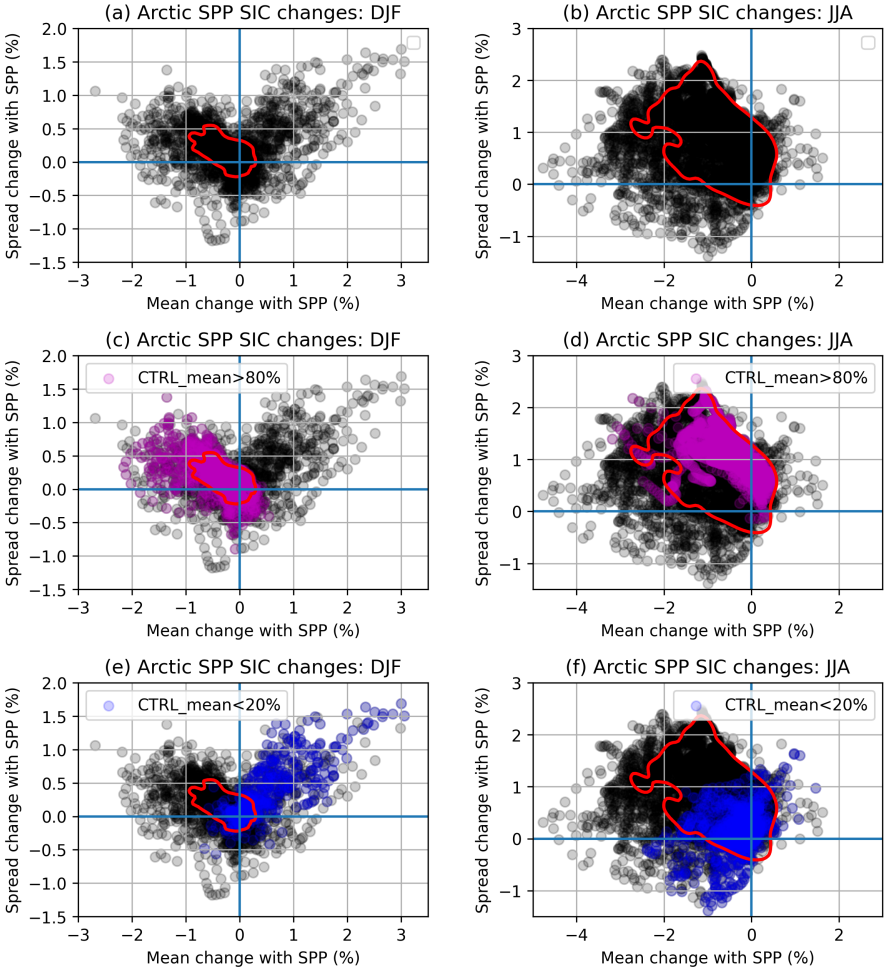

Figure 10. Relationship between changes to the mean and the spread. In (a): scatter plot of the SIC ensemble mean change versus the ensemble spread change at each gridpoint in the Arctic (60-90N) for DJF averages. The red contour encloses 90% of all the points. In (c) and (d): the same as (a) and (b) but gridpoints where the CTRL ensemble mean SIC exceeds 80% are marked in mauve. In (e) and (f): the same as (a) and (b) but gridpoints where the CTRL ensemble mean SIC is less than 20% are marked in blue.

The spatial asymmetry of the changes in winter (decreased SIC in the ice pack and increased SIC at the ice edge) suggests that some of the relationships highlighted above may be related to how the impact of perturbations differs depending on the underlying mean SIC state in the CTRL experiment. To assess how changes to the SIC mean and spread depend on the CTRL mean state, we show in Figure 11 2D histograms of these quantities, normalised so each column sums to 1. This normalisation

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

is for visual clarity: the raw histograms are bimodal, especially in winter, being dominated by gridpoints that are close to 100% SIC or close to 0%. The normalisation makes it easier to assess changes in the less common intermediate values. From this figure we can see that in winter, the mean SIC tends to increase in regions where the initial SIC mean is low (<10%), and to decrease in all other regions. The decrease is especially large in regions where the initial SIC is close to 100%. The spread is broadly enhanced independently of the initial SIC mean. This is consistent with the impression of Figure 1, since the gridpoints where the initial SIC is low in winter are precisely those along the ice edge, and the gridpoints where the initial SIC is high are those in the central ice pack. In summer the patterns are less clear. Spread increases almost everywhere, while the mean only increases in regions where the initial SIC mean is very low or very high, and decreases otherwise. Again, this is consistent with Figure 2, which shows changes that are not immediately obviously related to the baseline CTRL mean. We note that the picture is largely the same in the Antarctic (not shown).

To summarise, changes to the mean sea ice and changes to the sea ice ensemble spread are not independent of each other, and there is clear evidence of distinct regimes within which the covariance of these two changes is strongly constrained. In the next section we discuss the likely reason for this.

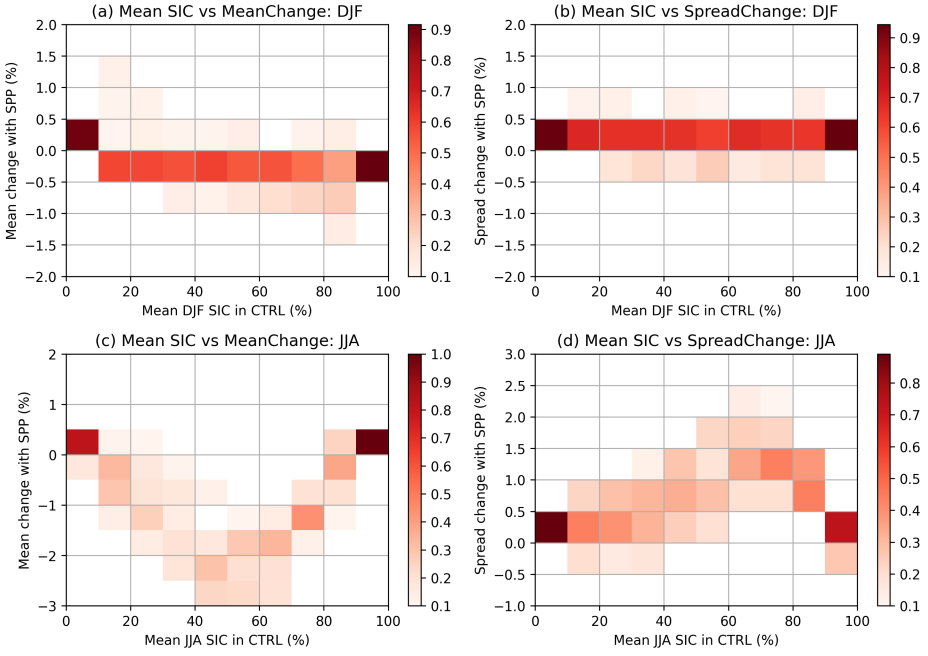

Figure 11. Dependence on the basic mean state. In (a): 2D histogram of the mean DJF SIC in CTRL against the mean SIC change with SPP at each gridpoint in the Arctic (60-90N). For visual clarity each column has been normalised to sum to 1. In (b): the same but showing mean DJF SIC in CTRL against the spread change with SPP. In (c) and (d): equivalent histograms for JJA.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

5.2 Conceptual model of the impact of SPP

It is natural to speculate that many of the impacts of SPP, including the impact on the mean SIC and the relationship between the mean and spread, can be explained by the strong asymmetries and non-linearities present when perturbing sea ice. Some of these asymmetries exist at the level of the parameters themselves. For example, Juricke et al. (2013) argued that low values of P^* (rn_pstar in Table 1) are more effective at reducing the ice strength than high values are at increasing it, meaning that mean zero perturbations to this parameter can systematically push the system towards weaker ice. This in turn can promote enhanced drift (e.g., from the ice pack to the edge in winter) and break-up of ice. Because we perturb 9 parameters which all interact with each other, understanding such parameter-level asymmetries in detail is not feasible. However, we argue that a simple conceptual model based on more fundamental asymmetries can already explain many of the results.

One obvious fundamental asymmetry in the system comes from the fact that SIC is bounded above and below. Thus, if one perturbs a gridpoint which is close to 100% SIC, one will only see an effect if the perturbation is acting to decrease the SIC. Furthermore, the SPP scheme only perturbs gridpoints where there is *some* sea ice. Thus, if a perturbation pushes a gridpoint with very low SIC down to zero SIC, that gridpoint will in effect become invisible to SPP. In particular, SPP will not be able to re-freeze the water at that gridpoint in and of itself. These two heuristic boundary effects both suggest that SPP will lead to an overall reduction in SIC. This suggestion can be confirmed in a simplified statistical model. We start by drawing a sample from a strongly bimodal distribution with peaks close to 0 and 100, which is then clipped to be bounded by 0 and 100. We then repeatedly 'perturb' this distribution by repeatedly adding random white noise with a specified standard deviation to the samples: samples that end up below zero are removed, while samples that end up above 100 are clipped to 100. Repeating this multiple times for multiple choices of standard deviation results in Figure 12, which confirms that this procedure produces a consistent decrease in the mean which scales relatively linearly with the magnitude of the perturbations. We emphasise that this is purely a toy-model and the specific numbers in the figure are not meant to be taken at face value. In particular, the complete removal of samples going below zero is likely too strong an assumption, because if spurious melting occurs in SI^3 due to a perturbation the model will typically push towards forming new ice at the next time step. We also note that because SPP perturbs sea ice parameters and not the sea ice itself, it is hard to know how to best construct such toy-models; e.g., is the net effect of SPP on the SIC tendencies themselves actually white noise? But Figure 12 at least supports the basic heuristic. The same heuristic also implies that gridpoints where the SIC is very low will tend to show a net increase with perturbations, and the gridpoints with very high SIC will show a clear decrease. This is what was seen in Figure 11(a) for winter; the same effect is less clear in summer because there are far less gridpoints with very high SIC.

This first fundamental asymmetry can also explain the relationships found in the previous section between the original mean state, mean state changes and ensemble spread changes. When gridpoints close to 100% SIC experience a reduction in SIC due to SPP, the variability will also increase, because points where the ice is particularly dense are constrained from moving or changing. We therefore expect a negative correlation between mean state changes and ensemble spread changes across such gridpoints. On the other hand, for gridpoints close to 0% SIC the opposite occurs: increasing the SIC is now expected

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

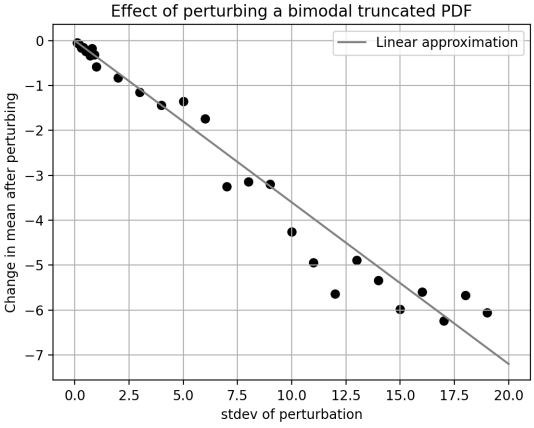

Figure 12. Toy-model for SPP's mean state impact. The change in the mean of a truncated bimodal distribution when repeatedly applying white noise with a given standard deviation, as a function of this standard deviation. The grey line is a linear fit to the points.

to enhance spread, as there is more ice to move around, while reducing it can bring the spread down to zero by removing it entirely. We therefore expect a positive correlation for such gridpoints. These expectations are realised in Figure 10.

Another fundamental asymmetry in winter is that between the ice pack and the ice edge. When the perturbations affect ice advection in winter, there is little scope for moving the ice towards the central Arctic, but plenty of scope for moving the ice towards the edge. This is consistent with what is seen in Figure 5(a). In summer, this constraint does not exist, and the ice is much more free to move, implying that perturbations should push the ice around more randomly. This is again consistent with Figure 5(b), which gives a clear impression of random changes to the drift.

The first asymmetry is present in both winter and summer and suggests the mean SIC should decrease with SPP in both seasons. The second asymmetry is only present in winter, and suggests an increase in SIC at the edge and a decrease in the center. Thus, these asymmetries, and the purely statistical or mechanical consequences of them, already likely explain much of the SIC mean state change seen with SPP. Our confidence in this explanation is further bolstered by the fact that SPP has essentially the same impact in the southern hemisphere: there is a net decrease in SIC in both seasons, but in winter there is a net transport of ice away from the Antarctic continent to the ice edge (see Figures S4 and S5). The results are also found to be largely similar in equivalent experiments run at higher atmospheric resolution and in simulations where independent perturbations are applied to each parameter (not shown). It is likely that parameter specified asymmetries and atmospheric feedbacks make a contribution as well. The increase in the mean SIT is, for similar reasons, probably related to the fact that the SIT distribution is bounded below but not above, which means random perturbations are more likely to push the SIT higher.

An important caveat to this discussion is that Juricke and Jung (2014) showed that the response to perturbations can undergo distinct phases, with an initial multidecadal transient phase in which the system equilibrates with the perturbations, after which the response can then differ. The changes seen in our simulations thus only show the very initial part of the transient response, which is nevertheless relevant for weather as well as sub-seasonal and seasonal predictions. It is possible that our conceptual

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

model is only informative for this phase, with slower timescale and harder to predict physical feedbacks dominating the longer response.

6 Impact on seasonal forecast skill

6.1 Impact on sea ice forecast skill

We have seen that SPP changes both the ensemble mean and ensemble spread of the SIC forecast. We now examine how this impacts forecast skill, focusing on SIC. Figure 13 shows how SPP changes the RMSE, spread-error ratio and bias-corrected CRPS for seasonally averaged winter SIC. The change in RMSE is negative on average (i.e., the error goes down), but this is partially due to small improvements in the central Arctic, which are difficult to validate against observations. At the coastlines, where validation is more robust, there is a mixed impact. The tendency of SPP to enhance the amount of ice at the ice edge is sometimes increasing the model bias, and hence the RMSE. This is especially the case in the Greenland sea, where the CTRL forecast has its biggest bias. For spread-error ratios, there are large negative values in CTRL, confirming the underdispersion of SIC forecasts in the IFS. The exceptions to this are Hudson Bay and Chukchi Sea, where overdispersion occurs. Adding SPP generally enhances spread and hence brings the ratio closer to 1, though only by a small amount relative to the CTRL bias. The increased spread-error ratios manifest already within the first forecast month and are consistent thereafter (not shown). Interestingly, in Hudson Bay, and to a lesser extent the Chukchi Sea, where the CTRL forecast is overdispersive, SPP reduces the spread-error ratios. For CRPS, the improved spread results in broadly better (i.e., lower) scores, though some important exceptions exist at the ice edge, especially in the Greenland sea. While the SIC forecasts have been bias corrected to have no bias on average across the period 1993-2023, there are still many individual forecasts where both CTRL and SPP predict non-zero SIC values at gridpoints where no sea ice exists in observations. In such cases the forecast is heavily penalised. In other words, the binary nature of sea ice means that the typical bias correction approach cannot remove all the effects of model biases. This explains why the CRPS is worse with SPP in the Greeenland sea: SPP slightly increases the bias there resulting in greater penalties at gridpoints where no sea ice was present in observations. It is plausible that the use of calibration methods more tailored to sea ice, such as those discussed in Dirkson et al. (2022), would result in a more positive improvement due to SPP.

In summer, spread-error ratios are again almost uniformly improved. Changes to RMSE and CRPS are less clear cut, consistent with the more spatially varying changes to the mean SIC seen in Figure 2.

6.2 Impact on atmospheric forecast skill

We have already seen in Section 4.2 that the SPP-induced changes to the sea ice affect the atmosphere. We now examine how these changes affect forecast skill. We restrict attention here to winter only, given the limited impact seen in summer (Figure 7). We also restrict the discussion to T850 and Z500, thus avoiding T2M, where validation in the Arctic is highly uncertain (Tian et al., 2024).

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

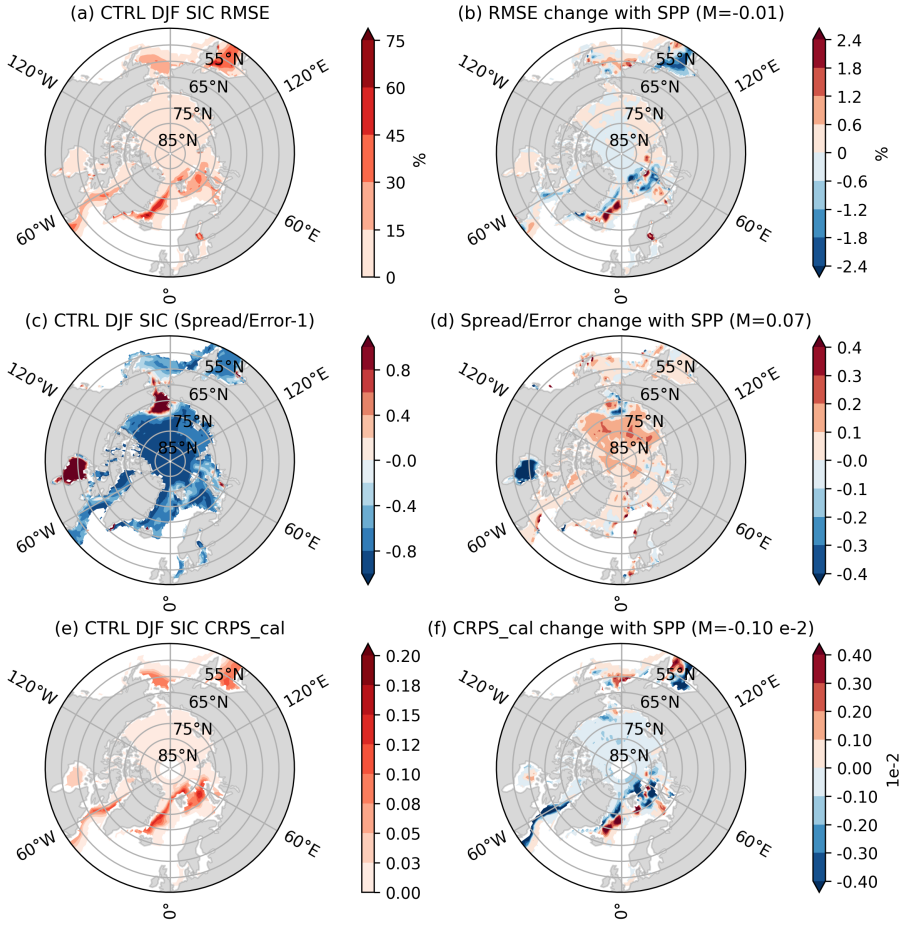

Figure 13. Impact of SPP on winter SIC forecast skill for months 2-4. In (a): the RMSE of DJF SIC for the CTRL forecast. In (b): the difference in RMSE with SPP. In (c) and (d): the same but for the spread-error ratio minus 1 (so a value of 0 means perfectly calibrated). In (e) and (f) the same but for the CRPS of the bias corrected forecasts. The values of M in the titles are the area-weighted mean across all gridpoints where the SIC is non-zero. Forecasts are validated against OSI-SAF.

Figure 15 shows the impact on T850 RMSE, spread-error ratios and bias-corrected CRPS. There is a decrease in the RMSE along the Greenland, Barents and Kara seas, which is consistent with a robust reduction in the mean bias here arising from the SPP induced mean change (see Figure 6(c) and Figure S9). Changes elsewhere are also consistent with changes to the mean bias (Figure S9) but these are less clearly related to robust shifts in the ice so may reflect noise. The spread-error ratios are also increased around Greenland and the Kara sea, but in the former case this is a small degradation. Given these mixed impacts on RMSE and spread-error ratios, it is not surprising the there is little change to the CRPS of bias-corrected fields.

Given the potential for stochastic sea ice perturbations to influence the midlatitude circulation via Arctic-midlatitude links (Strommen et al., 2022), it is of interest to assess changes in Z500 skill with SPP. Figure 16(a) shows the change in anomaly

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

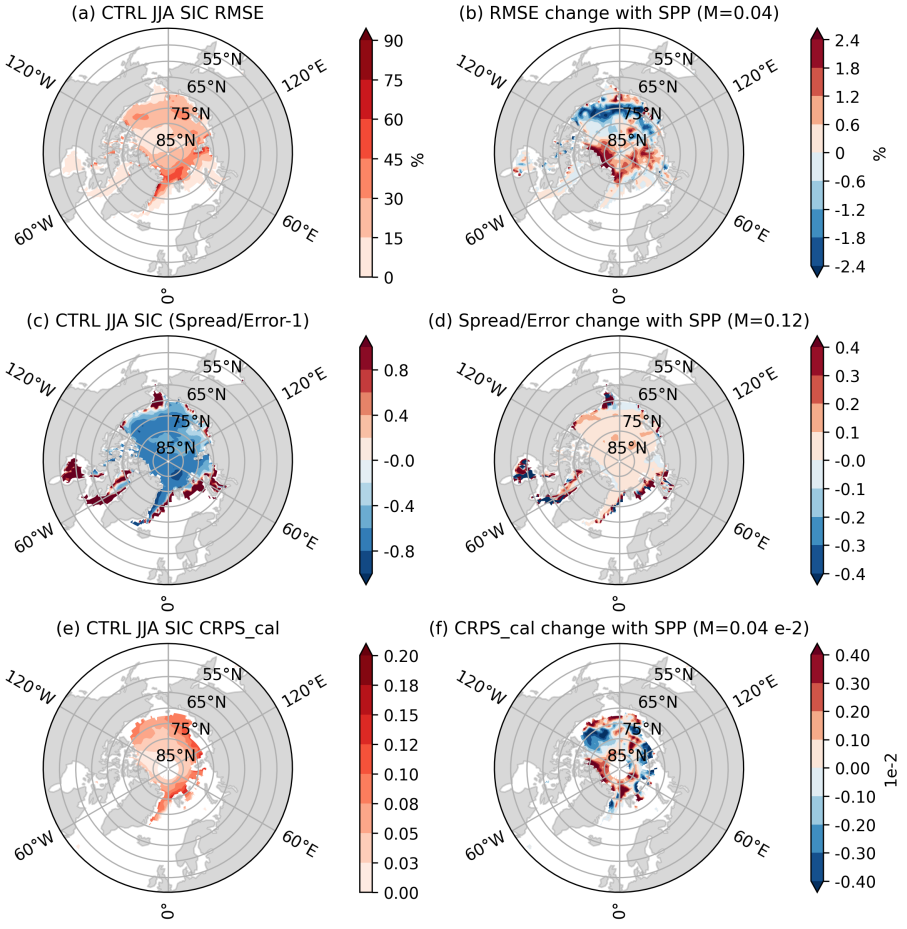

Figure 14. Impact of SPP on summer SIC forecast skill for months 2-4. As Figure 13 but for JJA.

correlation coefficient (ACC; the correlation between the ensemble mean and observations) with SPP in winter. There is a large and statistically significant increase in the sub-polar North Atlantic, extending into Scandinavia, with the difference peaking at around 0.5. If we define a broad sub-polar North Atlantic (SPNA) box by 55N-65N,60W-30W (black box in the figure), then the ACC of SPNA-averaged Z500 increases from around 0.3 in CTRL to 0.6 with SPP, a statistically significant difference ($p < 0.05$). This increase in skill is particularly intriguing because the CTRL skill is close to zero in this region (not shown). The spatial pattern of the skill increase only partially overlaps the NAO pattern, and the DJF NAO forecasts in SPP therefore exhibit only a somewhat higher ACC of 0.36 compared to 0.24 in CTRL. This increased NAO ACC is not statistically significant, but is consistent across each month of the winter season (Figure S10).

Understanding the increased Z500 skill with SPP is challenging, not just because of the small signal-to-noise ratios in the midlatitudes, but also because it could reflect contributions from several distinct events (e.g., improved forecasts over the Labrador sea in some years and improved forecasts around Scandinavia in others), and the causal mechanisms for each could

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

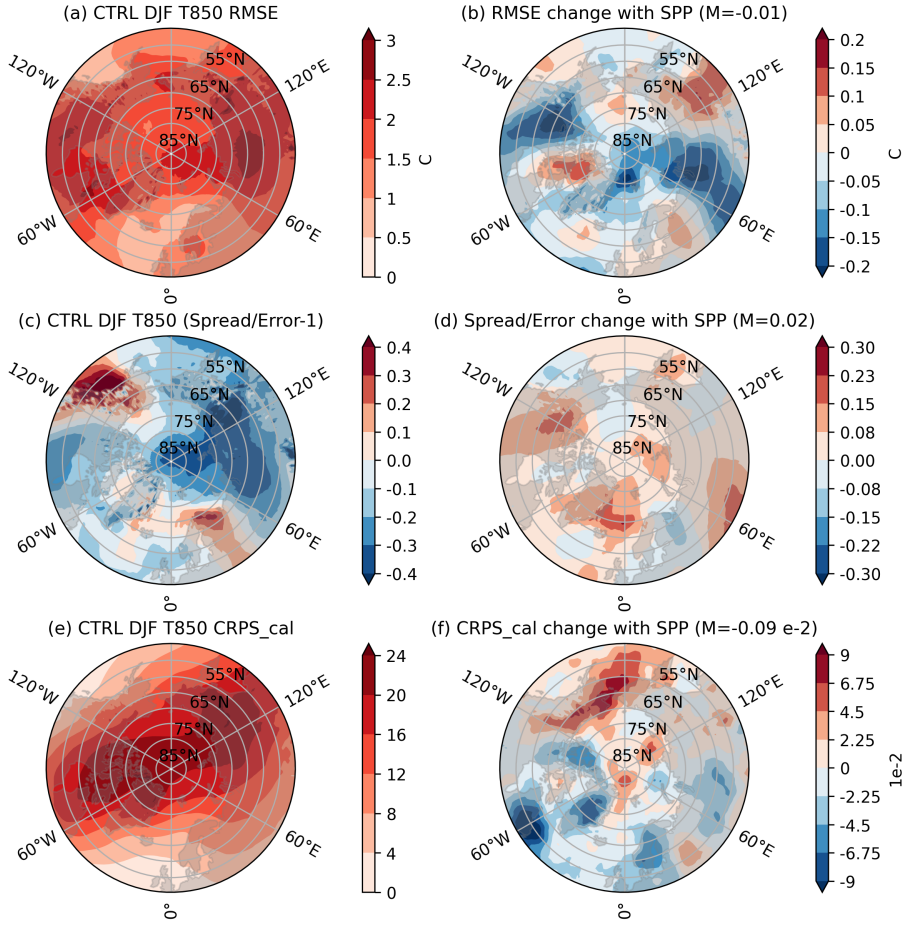

Figure 15. Impact of SPP on winter T850 forecast skill for months 2-4. In (a): the RMSE of DJF T850 for the CTRL forecast. In (b): the change in RMSE with SPP. In (c) and (d): the same but for the spread-error ratio minus 1 (so a value of 0 means perfectly calibrated). In (e) and (f) the same but for the CRPS of the bias corrected forecasts. The values of M in the titles are the area-weighted mean across all gridpoints. Forecasts are validated against ERA5.

be unrelated. Furthermore, despite the statistical significance of the differences, it is hard to rule out a contribution from internal variability, and that the true magnitude of the impact of SPP is smaller. However, we identify at least one potential pathway from sea ice changes to improved Z500 forecasts. Figure 16(b) shows the CTRL bias in T850, with the change in T850 mean with SPP overlain. The CTRL model is uniformly biased warm across the Arctic. A consequence of this is that the meridional temperature gradient in the North Atlantic domain is biased weak in CTRL. As previously noted, SPP increases the SIC mean along the edge of the Greenland, Barents and Kara seas, resulting in a robust decrease in T850, thereby reducing the warm bias. This bias reduction could impact forecasts of the North Atlantic jetstream by strengthening the meridional temperature gradient, thereby modulating jetstream variability (Breul et al., 2025). To assess this hypothesis, we define an IceEdge box by

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

67N-77N, 60W-90W (mauve box in Figure 16(b)). We then resample SPP ensemble members with replacement to generate 5000 synthetic values of the SPNA ACC and IceEdge T850 means. A scatter plot of these is shown in Figure 16(c), showing a statistically significant correlation of -0.28 between the two. Thus, SPP members that show greater 850hPa cooling along the ice edge show greater Z500 skill in the SPNA region, consistent with the hypothesis that changes to the meridional temperature gradient play a role in driving the increased skill. Additional evidence towards this hypothesis comes from the fact that the area-averaged Z500 in the SPNA box correlates with the NAO index in ERA5 ($R^2 = 0.54, p \ll 0.05$), and increased SIC in the IceEdge box has previously been identified as a potential causal driver of the NAO (Strong et al., 2009). Changes to SIC in this region could plausibly therefore also drive variations in SPNA Z500.

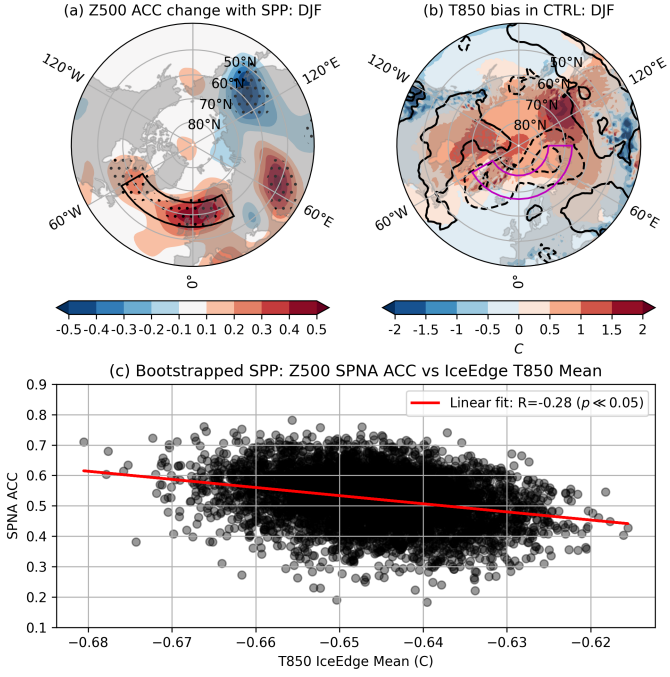

Figure 16. Impact of SPP on winter Z500 forecast skill for months 2-4. In (a): difference in DJF Z500 ACC between SPP and CTRL at every gridpoint. The black box highlights the SPNA region. Stipling denotes significance ($p < 0.05$). In (b): the DJF T850 bias in CTRL (filled contour) and the change in the T850 ensemble mean with SPP (line contours; marked at $\pm 0.05, \pm 0.15, \pm 0.25$ C). The mauve box highlights the IceEdge region. In (c): scatter plot of SPNA ACC's versus the mean T850 in the IceEdge region, obtained by bootstrap resampling the SPP ensemble with replacement 5000 times. The red line is the linear fit. Forecasts are validated against ERA5.

An important caveat to the above discussion is that the causality could go the other way. If the Z500 circulation in the SPNA region changes due to random internal variability, then this would impact the sea ice in the IceEdge box, which could result in the relationship between SPNA ACC and T850 mean we found. However, we point out that if the change in skill is not just random variability, then the causality necessarily comes from the sea ice, not the atmosphere, since the inclusion of SPP to the sea ice is the only difference between the two ensembles. Finally, we note that SPP also results in a significant increase in

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

T850 ensemble spread over Greenland (Figure 6(d)), but while this could also play a role, assessing this is more complicated and goes beyond the scope of the present work.

To summarise, SPP leads to an overall small reduction in T850 RMSE in winter, a small increase in spread-error ratios, but has a near-neutral impact on bias-corrected CRPS. SPP also results in a large increase in winter Z500 ACC over the sub-polar North Atlantic extending into Scandinavia, which may be related to how changes to the ice edge induced by SPP alter the mean meridional temperature gradient.

7 Discussion

The SPP scheme achieves its main goal of rapidly enhancing ensemble spread and thereby producing better calibrated sea ice forecasts. However, there are arguably two potential limitations, which we now discuss.

Firstly, despite applying relatively large perturbations to 9 different parameters, SPP only increases ensemble spread in coupled forecasts by around 10% relative to the ensemble without sea ice perturbations, and the perturbed ensemble is therefore still notably underspread. This reflects in parts the lack of spread in the ice and ocean initial conditions. In particular the sea ice initial conditions from the ORAS6 reanalysis ensemble are known to have insufficient spread. The stochastic perturbations are therefore being applied to an ensemble whose sea ice evolution is highly constrained by the initial conditions. Importantly, the modest impact of SPP should therefore probably not be seen as a failure, since attempting to compensate for lack of initial condition spread with parameter perturbations could quickly risk pushing physical parameters well beyond any reasonable range of their estimated true value. Work is ongoing by the third author to apply a variant of the SPP scheme described here in ocean-ice data assimilation, in order to enhance initial condition spread. Given the large relative increase in spread seen in uncoupled simulations (Figure 9), this is a promising avenue. Future work aims to additionally explore the effect of perturbing liquid-ocean parameters, also available via STOPACK (Storto and Andriopoulos, 2021). Since SSTs exert a strong influence on ice formation, this could enhance ensemble spread further. All this being said, it must be kept in mind that the ensemble spread may already be close to saturated in the IFS, and that further improvements to spread-error ratios require a reduction of the error rather than an increase of the spread.

A second limitation is the change in the mean state, since this can sometimes increase the bias. Indeed, one of the key appeals for using parameter perturbations is to enhance spread while limiting mean state changes by ensuring local conservation of mass and energy (Lang et al., 2021). We find that for sea ice, the strong asymmetries in the system mean that changes to spread are not independent of changes to the mean, and that some mean state changes are probably inevitable. It might be possible to remedy this by retuning the default value of the perturbed parameters to compensate for the induced mean state change. Since the changes to the mean are fairly modest, on the order of $\approx 4\%$ for SIC and ≈ 10cm for SIT, the retuning required is likely minimal. However, this could come at the cost of reducing the spread increase. We encourage future exploration along these lines to obtain a version of sea ice SPP which maximally enhances spread while minimising mean state changes.

The impacts we reported on how SPP affects seasonal forecasts of the winter atmosphere need to be treated cautiously. Some of the largest changes to winter air temperature occur around the Greenland, Barents and Kara seas, which is consistent with

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

these regions being where the maximum amplitude ice-driven heatfluxes occur (Koenigk et al., 2009). SPP tends to increase the SIC along the edge of these regions, which exacerbates the SIC bias in the CTRL model. However, the additional ice cools the overlying atmosphere by suppressing heat fluxes, which reduces the atmospheric warm bias in the CTRL model. Thus, some of the improvements to atmospheric forecasts arising from SPP arise from a compensation of biases rather than a genuine improvement to the coupled system. More generally, whether or not SPP improves or degrades atmospheric forecasts in a particular region will likely depend sensitively on the baseline CTRL biases in both the ice and atmosphere. As a result, while we expect the impact SPP has on the SIC to be fairly robust across different model versions (at least on the seasonal timescales we are concerned with), we do not necessarily expect the same degree of consistency for the impact on atmospheric forecasts.

The apparent increase in Z500 skill in the Euro-Atlantic, while particularly intriguing, should be interpreted with the same level of caution. Our hypothesis for the causal pathway goes via the reduced warm bias around Greenland, and to a lesser extent, Barents and Kara, which, if correct, implies the Z500 skill increase again arises from a compensation of biases. The increased Z500 skill therefore might not translate to other forecast models. A counterpart ensemble forecast run at higher atmospheric resolution with and without a modified version of SPP still reveals an increase in Z500 skill in the Euro-Atlantic domain with SPP (not shown), but one which is smaller in magnitude and spatially displaced somewhat southwards compared to what is shown in Figure 16(a). This difference could be related to a general reduction in the Arctic warm bias in this higher resolution counterpart. Thus, while there is some evidence that the increased skill is a robust effect, more work would be required to substantiate this. We therefore suggest that these results should be viewed as consistent with the body of literature establishing that Arctic sea ice can exert a small but robust influence on the jetstream in winter (Smith et al., 2022), and that stochastic sea ice perturbations can thereby modulate the jetstream. We add to this by finding evidence that the impact seems to be large enough to affect seasonal forecast skill, although the precise magnitude and spatial pattern of the signal is unclear.

Relatedly, we find that the result from Strommen et al. (2022) is not reproduced in these forecasts. Strommen et al. (2022) found that including stochastic perturbations to the P^* parameter in the ice and various ocean parameters results in clearly enhanced correlations between Barents-Kara sea ice and the NAO in winter. In the seasonal forecasts considered here we see no meaningful difference between the correlations (Figure S11). There are several possible reasons for this failure to reproduce. Firstly, it could indicate that the results of Strommen et al. (2022) were inflated by sampling variability, and the null-result seen in the larger ensemble used here is correct. Secondly, a crucial difference between the forecasts considered here and the simulations used in Strommen et al. (2022) is that the latter are free-running climate simulations using a model with different biases. Past work has shown that the atmospheric response to sea ice anomalies in a particular region can depend crucially on the exact change in the ice (Petoukhov and Semenov, 2010); that the Arctic regions that are able to exert an influence on the NAO in models can depend on the basic mean state of the ice (Strommen and Cooper, 2024); and that the response to Arctic sea ice anomalies exhibits mean state dependence more broadly (Smith et al., 2017). Thus, the different behaviour could reflect the different mean states of the free-running EC-Earth3 simulations in Strommen et al. (2022) and the initialised IFS forecasts used here, as well as the somewhat differing responses to stochastic perturbations in both. Indeed, in the EC-Earth3 simulations, the largest changes to the mean and variance occur in the Barents-Kara region, while in the seasonal forecasts considered

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

here the net change in these regions is small, and the larger, more consistent change occurs in the Greenland sea (Figure 1). Thus, the pathway from the ice to the NAO in these seasonal forecasts may be more related to Greenland SIC as opposed to Barents-Kara SIC. The potential importance of Greenland SIC was also highlighted in Strommen and Cooper (2024). It would be interesting to see how the impact of SPP evolves when allowed to run freely for longer, including how it compares to the impact in the free-running EC-Earth3 simulations. Thirdly, the simulations in Strommen et al. (2022) also included stochastic perturbations to various ocean parameters, and these may have contributed to the change in Arctic teleconnections by altering teleconnections from the tropics (see ibid). The absence of ocean perturbations in our seasonal forecasts could therefore play a role. Exploring all of these possibilities in depth goes beyond the scope of this paper.

8 Conclusions

The inclusion of stochastic perturbations to sea ice parameters, via an SPP scheme, results in a robust increase in ensemble spread for SIC and SIT in both summer and winter, and, consequently, better spread-error ratios in the sea ice forecasts. The relative increase in spread is much greater in uncoupled forecasts than in coupled forecasts. Systematic but small changes to the mean are also found, with a general decrease in SIC and increase in SIT across all seasons. In winter, there is also a net redistribution of ice from the ice pack to the ice edge, while in summer the perturbed ice undergoes seemingly random motions. We presented a simple conceptual model for understanding these mean state changes via asymmetries in the system. The boundedness of both SIC and SIT likely explains the net decrease of SIC and increase of SIT across both seasons. In winter, the barrier placed by the high strength of the ice pack likely explains why random perturbations systematically move ice outwards, while in summer, the ice can move more freely, and hence the random perturbations move the ice around more randomly. An important consequence of this is that enhanced spread is not independent of changes to the mean in our forecasts, meaning a tradeoff between maximising the increased ensemble spread and minimising the mean state change is likely necessary.

Our analysis shows that perturbations to the ice have a measurable impact also on seasonal forecasts of the atmosphere, at least in winter. By enhancing the spread and altering the mean SIC at the ice edge, SPP modulates the strong heat fluxes taking place here in winter, resulting in small but robust changes of order 0.5C to mean air temperature up to at least 850hPa. Air temperature ensemble spread is also enhanced at both surface and 850hPa by around 10% relative to the baseline spread, both over the Arctic and to some extent at lower latitudes. These changes appear to be sufficient in magnitude to affect seasonal forecast skill of Z500 in winter at lower latitudes: we find a robust improvement of DJF Z500 forecast skill over the sub-polar North Atlantic and extending into Scandinavia. The change projects to some extent onto the NAO, and so the inclusion of SPP somewhat enhances NAO skill. This appears to be mostly related to changes in the mean SIC around Greenland, and we find no evidence that SPP changes the teleconnection between Barents-Kara SIC and the NAO in these forecasts, unlike what was reported in Strommen et al. (2022). Thus, while adding stochastic perturbations to the sea ice does appear to be able to positively affect the midlatitude circulation, including the NAO, the exact nature of the change and its origin are likely sensitive to the details of the perturbations, the simulations analysed (free-running versus initialized) and the model used.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

Moving forward, ECMWF will be exploring combining SPP for sea ice with SPP for the ocean, both for use in data assimilation and coupled forecasts. Further assessment of the impact of SPP for sea ice in the IFS is ongoing. In particular, forthcoming work will evaluate the impact on medium-range and sub-seasonal forecast skill.

Data availability. Data statement.

ERA5 data is publicly available via the Copernicus Data Store Hersbach et al. (2023). IFS forecast data is publicly available via MARS [IN PROGRESS]. OSI-SAF data is publicly available courtesy of EUMETSAT: https://osi-saf.eumetsat.int/products/sea-ice-products.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

Author contributions. KS led the analysis, manuscript writing and produced the figures. MM led the initial implementation and testing of SPP, produced the main forecast experiments, conducted preliminary analysis and assisted with the manuscript. AS led the development of SPP, produced some of the experiments used, and provided expert input when implementing SPP. JS provided supplementary diagnostics and assisted with analysis and the manuscript. ST provided project supervision and funding, assisted with implementation of SPP and creation of forecast experiments, provided feedback on the manuscript, and provided expert input on the analysis conducted.

Competing interests. We declare there are no competing interests.

Acknowledgements. This research has been supported by the European Union's Horizon Europe project ACCIBERG (grant no. 101081568). MM was additionally supported by the Horizon Europe project EXPECT (grant no. 101137656). We thank Stephan Juricke for helpful discussions and feedback.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

References

Berner, J., Achatz, U., Batté, L., Bengtsson, L., De La Cámara, A., Weisheimer, A., Weniger, M., Williams, P. D., and Yano, J.-I.: Stochastic parameterizations: Toward a New View of Weather and Climate Models, Bulletin of the American Meteorological Society, 98, 565–588, https://doi.org/10.1175/BAMS-D-15-00268.1, 2017.

Breul, P., Ceppi, P., Simpson, I. R., and Woollings, T.: Seasonal and regional jet stream changes and drivers, Nature Reviews Earth & Environment, pp. 1–19, https://doi.org/https://doi.org/10.1038/s43017-025-00749-9, 2025.

Buizza, R., Milleer, M., and Palmer, T. N.: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, Quarterly Journal of the Royal Meteorological Society, 125, 2887–2908, https://doi.org/https://doi.org/10.1002/qj.49712556006, 1999.

Christensen, H. M., Berner, J., Coleman, D. R., and Palmer, T. N.: Stochastic parameterization and El Niño-southern oscillation, Journal of Climate, 30, 17–38, https://doi.org/10.1175/JCLI-D-16-0122.1, 2017.

Davini, P., Von Hardenberg, J., Corti, S., Christensen, H. M., Juricke, S., Subramanian, A., Watson, P. A., Weisheimer, A., and Palmer, T. N.: Climate SPHINX: Evaluating the impact of resolution and stochastic physics parameterisations in the EC-Earth global climate model, Geoscientific Model Development, 10, 1383–1402, https://doi.org/10.5194/gmd-10-1383-2017, 2017.

Dawson, A. and Palmer, T. N.: Simulating weather regimes: impact of model resolution and stochastic parameterization, Climate Dynamics, 44, 2177–2193, https://doi.org/10.1007/s00382-014-2238-x, 2015.

Dirkson, A., Denis, B., and Merryfield, W.: A multimodel approach for improving seasonal probabilistic forecasts of regional Arctic sea ice, Geophysical Research Letters, 46, 10 844–10 853, https://doi.org/https://doi.org/10.1029/2019GL083831, 2019.

Dirkson, A., Denis, B., Merryfield, W. J., Peterson, K. A., and Tietsche, S.: Calibration of subseasonal sea-ice forecasts using ensemble model output statistics and observational uncertainty, Quarterly Journal of the Royal Meteorological Society, 148, 2717–2741, https://doi.org/https://doi.org/10.1002/qj.4332, 2022.

Fortin, V., Abaza, M., Anctil, F., and Turcotte, R.: Why should ensemble spread match the RMSE of the ensemble mean?, Journal of Hydrometeorology, 15, 1708–1713, https://doi.org/https://doi.org/10.1175/JHM-D-14-0008.1, 2014.

Gurvan, M., Bourdallé-Badie, R., Chanut, J., Clementi, E., Coward, A., Ethé, C., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Rousset, C., Storkey, D., Vancoppenolle, M., Müeller, S., Nurser, G., Bell, M., and Samson, G.: NEMO ocean engine, https://doi.org/10.5281/zenodo.3878122, 2019.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather and Forecasting, 15, 559–570,

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Sabater, M., J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 monthly averaged data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.f17050d7, 2023.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., et al.: SEAS5: the new ECMWF seasonal forecast system, Geoscientific Model Development, 12, 1087–1117, https://doi.org/https://doi.org/10.5194/gmd-12-1087-2019, 2019.

Juricke, S. and Jung, T.: Influence of stochastic sea ice parametrization on climate and the role of atmosphere-sea ice-ocean interaction, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372, https://doi.org/10.1098/rsta.2013.0283, 2014.

Juricke, S., Lemke, P., Timmermann, R., and Rackow, T.: Effects of Stochastic Ice Strength Perturbation on Arctic Finite Element Sea Ice Modeling, Journal of Climate, 26, 3785 – 3802, https://doi.org/10.1175/JCLI-D-12-00388.1, 2013.

Juricke, S., Goessling, H. F., and Jung, T.: Potential sea ice predictability and the role of stochastic sea ice strength perturbations, Geophysical Research Letters, 41, 8396–8403, https://doi.org/https://doi.org/10.1002/2014GL062081, 2014.

Juricke, S., Palmer, T. N., and Zanna, L.: Stochastic sub-grid scale ocean mixing: Impacts on low frequency variability, Journal of Climate, pp. JCLI–D–16–0539.1, https://doi.org/10.1175/JCLI-D-16-0539.1, 2017.

Keeley, S., Mogensen, K., Bidlot, J., Balmaseda, M. A., and Hatfield, S.: ECMWF Newsletter on NEMO-SI3, https://www.ecmwf.int/en/newsletter/180/earth-system-science/introduction-new-ocean-and-sea-ice-model-based-nemo4-si3, accessed: 2025-10-03, 2024.

Koenigk, T., Mikolajewicz, U., Jungclaus, J. H., and Kroll, A.: Sea ice in the Barents Sea: Seasonal to interannual variability and climate feedbacks in a global coupled model, Climate Dynamics, https://doi.org/10.1007/s00382-008-0450-2, 2009.

Kretschmer, M., Coumou, D., Donges, J. F., and Runge, J.: Using causal effect networks to analyze different arctic drivers of midlatitude winter circulation, Journal of Climate, https://doi.org/10.1175/JCLI-D-15-0654.1, 2016.

Lang, S. T. K., Lock, S.-J., Leutbecher, M., Bechtold, P., and Forbes, R. M.: Revision of the Stochastically Perturbed Parametrisations model uncertainty scheme in the Integrated Forecasting System, Quarterly Journal of the Royal Meteorological Society, 147, 1364–1381, https://doi.org/https://doi.org/10.1002/qj.3978, 2021.

MacLeod, D. A., Cloke, H. L., Pappenberger, F., and Weisheimer, A.: Improved seasonal prediction of the hot summer of 2003 over Europe through better representation of uncertainty in the land surface, Quarterly Journal of the Royal Meteorological Society, 142, 79–90, 2016.

Notz, D. and the SIMIP Community: Arctic Sea Ice in CMIP6, Geophysical Research Letters, 47, e2019GL086 749, https://doi.org/https://doi.org/10.1029/2019GL086749, e2019GL086749 10.1029/2019GL086749, 2020.

Ollinaho, P., Lock, S.-J., Leutbecher, M., Bechtold, P., Beljaars, A., Bozzo, A., Forbes, R. M., Haiden, T., Hogan, R. J., and Sandu, I.: Towards process-level representation of model uncertainties: stochastically perturbed parametrizations in the ECMWF ensemble, Quarterly Journal of the Royal Meteorological Society, 143, 408–422, https://doi.org/https://doi.org/10.1002/qj.2931, 2017.

Palmer, T., Buizza, R., Jung, T., Leutbecher, M., Shutts, G. J., Steinheimer, M., and Weisheimer, A.: Stochastic Parameterization and Model Uncertainty, vol. 598, http://www.ecmwf.int/publications/, 2009.

Palmer, T. N.: Towards the probabilistic Earth-system simulator: A vision for the future of climate and weather prediction, Quarterly Journal of the Royal Meteorological Society, 138, 841–861, https://doi.org/10.1002/qj.1923, 2012.

Perlin, K.: Improving noise, in: Proceedings of the 29th annual conference on Computer graphics and interactive techniques, pp. 681–682, 2002.

Peterson, K. A., Smith, G. C., Lemieux, J.-F., Roy, F., Buehner, M., Caya, A., Houtekamer, P. L., Lin, H., Muncaster, R., Deng, X., et al.: Understanding sources of Northern Hemisphere uncertainty and forecast error in a medium-range coupled ensemble sea-ice prediction system, Quarterly Journal of the Royal Meteorological Society, 148, 2877–2902, https://doi.org/https://doi.org/10.1002/qj.4340, 2022.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

Petoukhov, V. and Semenov, V. A.: A link between reduced Barents-Kara sea ice and cold winter extremes over northern continents, Journal of Geophysical Research: Atmospheres, 115, https://doi.org/https://doi.org/10.1029/2009JD013568, 2010.

Roberts, C. D. and Leutbecher, M.: Unbiased calculation, evaluation, and calibration of ensemble forecast anomalies, Quarterly Journal of the Royal Meteorological Society, p. e4993, https://doi.org/https://doi.org/10.1002/qj.4993, 2025.

Sanchez, C., Williams, K. D., and Collins, M.: Improved stochastic physics schemes for global weather and climate models, Quart. J. Roy. Meteor. Soc., 2015.

Shapiro, R.: Smoothing, filtering, and boundary effects, Reviews of geophysics, 8, 359–387, https://doi.org/https://doi.org/10.1029/RG008i002p00359, 1970.

Siegert, S., Bellprat, O., Ménégoz, M., Stephenson, D. B., and Doblas-Reyes, F. J.: Detecting improvements in forecast correlation skill: Statistical testing and power analysis, Monthly Weather Review, 145, 437–450, https://doi.org/https://doi.org/10.1175/MWR-D-16-0037.1, 2017.

Smith, D. M., Dunstone, N. J., Scaife, A. A., Fiedler, E. K., Copsey, D., and Hardiman, S. C.: Atmospheric response to Arctic and Antarctic sea ice: The importance of ocean–atmosphere coupling and the background state, Journal of Climate, 30, 4547–4565, https://doi.org/https://doi.org/10.1029/2009JD013568, 2017.

Smith, D. M., Eade, R., Andrews, M. B., Ayres, H., Clark, A., Chripko, S., Deser, C., Dunstone, N. J., García-Serrano, J., Gastineau, G., Graff, L. S., Hardiman, S. C., He, B., Hermanson, L., Jung, T., Knight, J., Levine, X., Magnusdottir, G., Manzini, E., Matei, D., Mori, M., Msadek, R., Ortega, P., Peings, Y., Scaife, A. A., Screen, J. A., Seabrook, M., Semmler, T., Sigmond, M., Streffing, J., Sun, L., and Walsh, A.: Robust but weak winter atmospheric circulation response to future Arctic sea ice loss, Nature Communications, 13, 727, https://doi.org/10.1038/s41467-022-28283-y, 2022.

Storto, A. and Andriopoulos, P.: A new stochastic ocean physics package and its application to hybrid-covariance data assimilation, Quarterly Journal of the Royal Meteorological Society, 147, 1691–1725, https://doi.org/https://doi.org/10.1002/qj.3990, 2021.

Strommen, K. and Cooper, F. C.: Physical and unphysical causes of nonstationarity in the relationship between Barents-Kara Sea ice and the North Atlantic Oscillation, Geophysical Research Letters, 51, e2023GL107 609, https://doi.org/https://doi.org/10.1029/2023GL107609, 2024.

Strommen, K., Christensen, H. M., Berner, J., and Palmer, T. N.: The impact of stochastic parametrisations on the representation of the Asian summer monsoon, Climate Dynamics, 0, 1–14, https://doi.org/10.1007/s00382-017-3749-z, 2017.

Strommen, K., Christensen, H. M., Macleod, D., Juricke, S., and Palmer, T. N.: Progress towards a probabilistic Earth system model: Examining the impact of stochasticity in the atmosphere and land component of EC-Earth v3.2, Geoscientific Model Development, https://doi.org/10.5194/gmd-12-3099-2019, 2019.

Strommen, K., Juricke, S., and Cooper, F.: Improved teleconnection between Arctic sea ice and the North Atlantic Oscillation through stochastic process representation, Weather and Climate Dynamics, 3, 951–975, https://doi.org/https://doi.org/10.5194/wcd-3-951-2022, 2022.

Strong, C., Magnusdottir, G., and Stern, H.: Observed Feedback between Winter Sea Ice and the North Atlantic Oscillation, Journal of Climate, 22, 6021 – 6032, https://doi.org/10.1175/2009JCLI3100.1, 2009.

Subramanian, A., Juricke, S., Dueben, P., and Palmer, T.: A stochastic representation of subgrid uncertainty for dynamical core development, Bulletin of the American Meteorological Society, 100, 1091–1101, 2019.

https://doi.org/10.5194/egusphere-2025-6402
Preprint. Discussion started: 2 January 2026

Tian, T., Yang, S., Høyer, J. L., Nielsen-Englyst, P., and Singha, S.: Cooler Arctic surface temperatures simulated by climate models are closer to satellite-based data than the ERA5 reanalysis, Communications Earth & Environment, 5, 111, https://doi.org/https://doi.org/10.1038/s43247-024-01276-z, 2024.

Vancoppenolle, M., Rousset, C., Blockley, E., Aksenov, Y., Feltham, D., Fichefet, T., Garric, G., Guémas, V., Iovino, D., Keeley, S., Madec, G., Massonnet, F., Ridley, J., Schroeder, D., and Tietsche, S.: SI3, the NEMO Sea Ice Engine, https://doi.org/10.5281/zenodo.7534900, 2023.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wang, L., Ting, M., and Kushner, P. J.: A robust empirical seasonal prediction of winter NAO and surface climate, Scientific Reports, 7, 279, https://doi.org/10.1038/s41598-017-00353-y, 2017.

Watson, P. A. G., Berner, J., Corti, S., Davini, P., von Hardenberg, J., Sanchez, C., Weisheimer, A., and Palmer, T. N.: The impact of stochastic physics on tropical rainfall variability in global climate models on daily to weekly timescales, Journal of Geophysical Research: Atmospheres, 122, 5738–5762, https://doi.org/10.1002/2016JD026386, 2017.

Xu, Z., Chen, J., Mu, M., Tao, L., Dai, G., Wang, J., and Ma, Y.: A stochastic and non-linear representation of model uncertainty in a convective-scale ensemble prediction system, Quarterly Journal of the Royal Meteorological Society, 148, 2507–2531, https://doi.org/https://doi.org/10.1002/qj.4322, 2022.

Zuo, H., Balmaseda, M. A., de Boisseson, E., Browne, P., Chrust, M., Keeley, S., Mogensen, K., Pelletier, C., de Rosnay, P., and Takakura, T.: ECMWF's next ensemble reanalysis system for ocean and sea ice: ORAS6, ECMWF Newsletter, 180, https://doi.org/10.21957/hzd5y821lk, 2025.