Linking Atmosphere-Only and Coupled Climate Simulations Through Perturbed Parameter Ensembles

B. S. Ferster[1†,2], J. Mignot[2], Guillaume Gastineau[2], F. Hourdin[3], and M. Coulon–Decorzens[3]

[1]CECI-CERFACS (Université de Toulouse, CNRS, IRD), Toulouse, France.

[2]LOCEAN-IPSL (Sorbonne Université, CNRS, IRD, MNHN), Paris, France.

[3]LMD-IPSL, Sorbonne University, CNRS, Paris, France.

Corresponding author: Brady S. Ferster (brady.ferster@cerfacs.fr)

ORCID: 0000-0001-9241-518X

†Current Affiliation.

Supporting Information:

https://docs.google.com/document/d/1RZsWm5lABhAgsNQ3F1UZ7-3UvYYKQDLx/edit?usp=sharing&ouid=109128369786086630786&rtpof=true&sd=true

Key Points:

- Twin perturbed parameter ensembles using atmosphere-only and coupled configurations are integrated to build tuning strategies.
- The upper-atmosphere warm anomaly found in the coupled configurations is linked to the radiative budget of atmosphere-only experiments.
- The anomalies in the upper troposphere found in the atmosphere-only runs are linked to polar climate anomalies of the coupled experiments.

Abstract

Calibration of climate models is essential for adjusting unresolved physical processes with large observational uncertainties through the adjustment of uncertain parameters, particularly those controlling clouds, the radiative budget, and ocean mixing properties. Perturbed parameter ensembles (PPEs) provide a systematic and objective alternative to manual tuning, but their use in coupled configurations is costly and remains therefore limited. Here, we analyze a pair of 120-member PPEs conducted with coupled and atmosphere stand-alone configurations of the IPSL model using identical parameter perturbations in both configurations. We use empirical orthogonal function analysis applied across the ensemble dimension to investigate the ensemble spread in both configurations. A first pattern is consistent with a thermal adjustment of the climate that is linked to the radiative budget of the atmospheric configuration. The second pattern is linked to the midlatitude higher-troposphere adjustments to upper level cloudiness, which is common in both configurations. This pattern is also related to sea ice changes in the coupled models. These results suggest that atmosphere-only configurations can be used to correct some of the biases in coupled models.

Plain Language Summary

1 Introduction

Uncertainty in climate projections remains a central challenge in climate science, largely due to variations in how models represent key physical processes and respond to external forcing (Hourdin et al., 2017; Knutti et al., 2017; Flynn and Mauritsen, 2020). Part of this uncertainty arises from parametric uncertainty. The latter is due to the fact that some processes cannot be explicitly resolved by the model equations, either due to their occurrence at spatial scales smaller than the model's resolution or because the governing for these processes remain insufficiently understood or unknown. The need to calibrate these parameters led to the development of frameworks for exploring the acceptable ranges of values for the parameters.

Previously, manual calibration and hand-chosen parameter combinations were the dominant approach in GCM calibration, with evidence that it may inadequately sample model diversity (Mauritsen et al., 2012; Sanderson et al., 2008, Hourdin et al., 2023). Manual calibration can be time consuming: the calibration of IPSL-CM6A-LR followed a long three-year sequence of improvements, bug fixes, and expert-driven manual tuning phases, largely focused on improving the sea ice extent and global mean surface temperature (Mignot et al., 2021). More recent studies used emulators, based on machine learning techniques, trained on large PPEs. A PPE is an ensemble of simulations using a modified set of parameters for each member. The investigation of such ensembles allows investigating the effect of the changes in each parameter and their interactions. The spread obtained with such an ensemble provides an evaluation of the so-called parametric uncertainty. The investigation of PPE usually involves having observational targets to reduce the dimensionality of the model outputs, and get physical or statistical insights on the adjustment in parameters needed to simulate correctly the target variables.

Several protocols have been previously proposed. For instance, Peatier et al. (2022; 2024) analyzed a PPE from the atmospheric component of CNRM-CM6-1, sampling 30 unconstrained parameters with a latin-hypercube strategy within their observational range. They propose selecting diverse but comparably performing candidates to better sample the parametric uncertainty and capture a range of possible climate responses. Eidhammer et al. (2024) similarly introduced an extensible PPE for CAM6 with the same framework to explore sensitivity to 45 parameters. Mikkelsen et al. (2025) used the same model and method but focused five aerosol–cloud parameters and compared the resulting simulations to surface observations.

In parallel, several methods have been developed using PPE results in order to infer climate characteristics from parameter combinations that were not directly sampled in the PPE. For instance, Williamson et al. (2013, 2017) built emulators based on gaussian processes to reconstruct and explore the climate in all the parameter space. On the other hand, Elsaesser et al. (2025) and Yang et al. (2025) used PPEs integrated with GISS ModelE atmosphere (version E3) to derive a calibrated physics ensemble (CPE) using emulators based on other machine learning methods.

These strategies offer alternatives to traditional manual tuning, emphasizing the need to build efficient targets with established observational uncertainties combined into statistical frameworks, in order to translate the expert driven tuning, done manually so far, into an objectified and automated expert driven tuning. As discussed in Williamson et al. (2013, 2017), Li et al., (2019), Couvreux et al. (2021), Hourdin et al. (2021, 2023), and Yamazaki et al. (2021), the use of emulator and process-based evaluation allows tuning to become more reproducible and potentially scalable across model resolutions. This approach reveals that distinct coupled climate states can emerge from multiple valid parameter combinations or unresolved structural errors (Hourdin et al., 2023; Peatier et al., 2024).

In summary, atmospheric stand-alone configurations provide means to test how atmospheric parameter perturbations influence radiation, circulation, and cloud processes (Vial et al., 2013; Webb et al., 2017; Duffy et al., 2023). These experiments have been widely used to calibrate atmospheric parameters, identify stable diagnostics, and explore emergent feedbacks (Mauritsen et al., 2012; Hourdin et al., 2017). This approach also allows separation of parametric uncertainty from structural uncertainty of climate models, which remains a critical challenge in interpreting models' ensembles diversity (Williamson et al., 2015; Sanderson et al., 2021; Peatier et al., 2024). However, the extent to which atmospheric parameter sensitivities derived from atmosphere-only configurations are transferable to coupled configurations remains mainly unknown.

With the emergence of single model perturbed parametric (PPE) protocols, recent efforts have aimed to quantify the impact of parametric uncertainty on simulated coupled climate responses (i.e., Sexton et al., 2021; Duffy et al., 2024; Bonnet et al., 2025; Yang et al., 2025; Hourdin et al., 2023, and several others). However, such an approach is costly CPU-wise to carry in coupled mode. More generally, a central challenge in climate model development is identifying which variables and diagnostics provide stable, transferable signals for tuning across experimental setups. Do atmospheric responses seen in atmosphere-only simulations carry over into coupled climate simulations, or do ocean–atmosphere interactions shift the parameter sensitivities once the ocean equilibrates?

This study addresses this question by directly comparing a pair of atmosphere-only and coupled PPEs built from the modification of 12 parameters. Specifically, we assess the opportunities of using atmospheric stand-alone simulations to precondition coupled model tuning, extending the approach and discussion of Hourdin et al. (2023). Our design allows us to isolate how ocean-atmosphere coupling alters the atmospheric diversity expressed in atmospheric ensembles and to test whether atmospheric stand-alone-based sensitivities can precondition coupled model tuning. Specifically, we (i) quantify the role of radiative adjustments and ocean-atmosphere coupling in masking, amplifying, or aligning ensemble diversity across configurations, (ii) compare ensemble spread across key atmospheric variables in atmosphere-only and coupled runs, and (iii) identify and understand the parameters controlling the spread in each configuration and assess their transferability between setups.

3. Experimental Protocol

2.1 Atmosphere-only and coupled configurations of IPSL-CM6

We use the latest version of IPSL-CM. It is derived from the CMIP6 (Coupled Model Intercomparison Project Phase 6) model version (Boucher et al., 2020), with small adjustments and bug corrections, and using newer versions of ocean and sea ice models. The coupled model consists of the LMDZ6 atmospheric component (Hourdin et al., 2020), the ORCHIDEE land surface model (Krinner et al., 2005), and the NEMO version 4 ocean model (Madec et al., 2019), coupled to the SI3 sea ice model (Vancoppenolle et al., 2023). The atmospheric resolution is 144 × 143 points in latitude and longitude (2.5° x 1.25°), with 79 vertical layers extending up to approximately 80 km. The ORCA1 ocean configuration is on a quasi-isotropic global tripolar grid with a 1° nominal resolution that increases latitudinal resolution to 1/3° in the equatorial region. Vertically, layer thickness varies from 1 m at the surface, 10 m at 100 m depth, and 200 m in the bottom layers (75 vertical levels).

The atmosphere-only simulations are run with imposed climatological sea surface temperatures (SSTs) and sea ice concentration (SIC) and last 2 years. The SST and SIC are derived from a climatology over 1979-2005 of the boundary conditions used in the AMIP (Atmospheric model intercomparison project) protocol. The first year of each simulation is discarded as spin-up.

The coupled simulations used fixed external forcings, such as greenhouse gases and aerosol concentrations, corresponding to the values of the year 2000 from the CMIP6 protocol. The initial conditions for the coupled experiments are taken from a control run with a spin-up of XX years obtained with the default values of the parameters. The simulations last at least 20 years. We are aware that this duration is too short with respect to the typical equilibrium time scale of the deep ocean, which could span decades to centuries. Yet, as explained next, we suggest that this duration is enough to investigate the drift toward an equilibration of the upper ocean, as we have chosen parameters that relate to the surface ocean rather than the deep ocean. To partly assess this important aspect, a subset of simulations is run over 60 years. Only the final 10 years of the simulations are considered, which helps to reduce the influence of drift.

All the coupled simulations were performed under present-day conditions. As explained in Mignot et al (2021), this protocol includes a small artificial increase of the oceanic surface albedo to mimic the effect of the present-day ocean heat uptake. This allows comparing the

results of the simulations against present-day observations, which are much more numerous and reliable than over the pre-industrial period.

Realm and model	Name		Min. Value	Max. Value	Current IPSL Model Value	Short Description
Atmosphere (LMDZ)	P1	CLC	1E-04	1,00E-03	6.5e-4	Autoconversion threshold for liquid cloud water.
	P2	FALLV	0.3	2.	0.8	Scaling parameter on the terminal fall velocity of ice crystals.
	P3	OMEPMX	0.0003	0.02	0.001	Ice-to-rain conversion efficiency of deep convection (1 — epmax)
	P4	DZ	0.04	0.12	0.07	Scaling factor on the thermal distance to boundary layer inversion.
	P5	EVAP	5E-05	5,00E-04	1,00E-04	Coefficient for re-evaporation of rain.
	P6	GKDRAG	0.2	2.	0.6	Subgrid-scale orographic drag coefficient.
Land Surface (ORCHIDEE)	P7	PCENT	0.3	1.	0.8	Maximum transpiration efficiency based on soil moisture.
	P8	ASNOW	5.	15.	10.	Snow albedo adjustment
Sea Ice (SI3)	P9	RNALB	0.	1.	0.50	Scaling factor for sea ice albedos.
	P10	RNCND	0.10	0.50	0.31	Thermal conductivity of snow over sea ice.
Ocean (NEMO)	P11	RNCE	0.06	0.08	0.06	Eddy diffusivity for mixed-layer processes.
	P12	RNLC	0.05	0.5	0.15	Langmuir cell mixing parameter.

Table 1. A descriptive table outlining the parameter names used within the PPE and the corresponding model component, value ranges, and short description.

2.2 Selection of parameters to be tested

Twelve unconstrained parameters are modified in the PPEs. These parameters are selected in atmospheric, oceanic and land surface components. The number of parameters was limited, since it scales the size of the ensemble: the number of simulations to adequately explore the parameter space using the iterative history matching method should be at least 10 times the number of parameters. These parameters, and their exploration ranges, are chosen based on expert judgment and are detailed Table 1. They are selected because they are known or suspected to influence the global climate and especially the coupling between atmosphere and ocean.

As processes controlling clouds are the most uncertain and affect global radiation budget the most, five of the twelve parameters varied in this experiment are involved in cloud-related parameterizations. They are chosen to most appropriately represent the diversity of cloud-related processes. Some of them control liquid cloud type (P1, P4, P5 in Table 1) while others affect mainly ice clouds (P2 and P3). P4 parameter is specific to the "thermal plume" parameterization, and affects mainly stratocumulus clouds, that are known to be key in regulating the climate and involved in atmosphere-ocean coupling (Hourdin et al 2015). P1 is one of the parameters that controls the liquid cloud reflectivity, impacting strongly global reflected shortwave radiation, and therefore global temperature, without affecting longwave global radiation. P2 is a shared parameter with other atmospheric GCMs that scales the terminal velocity fall of ice crystals. It affects both shortwave and longwave global TOA radiation (Hourdin et al 2017). It is known to be related with P3, a threshold parameter on condensed water-to-rain conversion in deep convection parameterization, that mainly controls the water remaining in the middle too high troposphere after deep convection event. P3 controls tropospheric humidity source term while P2 controls its sink. The sixth LMDZ parameter selected, P6, controls the influence of subgrid scale orography, which mainly affect Northern Hemisphere midlatitudes (Gastineau et al. 2020).

Parameters P7 and P8 respectively control land surface transpiration and land surface snow albedo in the land component, which affects land surface energy budget and land-atmospheric coupling. P9 and P10 are two parameters controlling respectively the sea ice albedo and the snow thermal conductivity of the SI3 model. Finally, two ocean parameters (P11 and P12) controlling upper ocean mixing processes are selected. Oceanic parameters were restricted to upper ocean processes due to the relatively short duration of the simulations for computing time constraints.

2.3 Preconditioning of the parameters range

We use an iterative history matching procedure to select the parameters of the twin PPEs (atmosphere-standalone and coupled simulation) of this study. The method is implemented with the High-Tune explorer software, as used in Hourdin et al (2021), Hourdin et al. (2023) or Raoult et al. (2024). It works iteratively : 1) generating a set of parameters via latin-hypercube sampling, 2) running a wave of PPE simulation, and 3) using emulators trained on these simulations to eliminate the combinations that produce climate metrics too far from target values, considering an expert given tolerance to error and the emulator uncertainty. At the end of each iteration, the space of acceptable parameters is reduced. The process can be repeated (resampling, running another wave of simulation, train emulators) to further narrow the acceptable parameter space. At the end, it provides a set of parameters values (or a range of) that allows simulating all climate metrics within their observational uncertainties. More details about the practical use and outcomes of the history matching can be found in the papers cited above. .

Following Hourdin et al 2021, in order to save computation time and ensure a good representation of some key phenomena at process scale, we start by constraining boundary layer processes, that control cumulus and stratocumulus cases. For this, we run the atmospheric model in single column mode, in which boundary conditions are chosen among a series of test cases, using target metrics computed on high resolution explicit simulations (LES) of the same

cloud scenes. For this initial step, only parameters P1 to P6 are constrained. Note however that all 12 parameters are sampled at each iteration. After 40 iterations, 71.7% of the volume of the initial parameter hypercube has been ruled out.

We then use waves of 2-year long atmosphere-only simulations, as presented in section 2.1. Parameters P1 to P8 are then constrained. The metrics used to characterize the climate are calculated from the top of the atmosphere (TOA) radiative fluxes and rainfall, which are compared to observations. The metrics consist of regional or global mean TOA radiative fluxes on the second year of the simulation, regional TOA radiative fluxes to target two major biases of the model, and three localized metrics related to precipitation. The targets for these metrics are computed from CERES-EBAF observations (Loeb et al., 2009) and the Global Precipitation Climatology Project (GPCP) monthly climatology (Huffman et al., 2001). We also use an a priori 7 W m^{-2} 1-sigma observational uncertainty, except for the global net TOA imbalance which follows a specific protocol explained below. The convergence of the initial parameter hypercube ruled out volume increases from 71.7% to 97.9% after the 41st iteration and 1st iteration of the atmosphere-only.

As in Hourdin et al. (2023), we target a global TOA radiative imbalance of 2.7 W m^{-2}, with an uncertainty of ± 1 W m^{-2}. This value was chosen empirically, as it has been noticed that it leads to a steady-state climate when using the same parameters in the coupled configuration. It is due to an imperfect energy conservation in the global model and to the different mean states of the coupled and stand-alone atmospheric simulations modulating the energy fluxes at TOA (Hourdin et al 2023). The analyses of the PPEs presented below will aim to verify this methodological choice and explore the related physical mechanisms.

GG : mention the links between parameters... especially the fact that *(P2) FALLV* and OMEPMAX are correlated. Maybe using the correlation matrix like Jeanne (M1 internship) did??

2.4 Final selection of the PPE

After 4 iterations of atmosphere-only simulations, the rules out volume of the initial parameter hypercube converges and reaches 98.8%. We finally select the 120 vectors of values for P1 to P12 of the 45th iteration. These vectors are used to construct the twin PPE of both atmosphere-only and coupled configurations. Note however that only the coupled simulations effectively use the P9 to P12 values for ocean and sea ice parameters. In practice, the 120 members of the atmosphere-only are first simulated for 2 years, after which the coupled model IPSL-CM is then run for 20 years using the same vectors of parameters. The ensemble of atmosphere-only simulations using each of these 120 vectors of parameters is hereafter called the atmosphere-only PPE . The ensemble of coupled simulations is designed as the coupled PPE in the rest of this manuscript. Additionally, 18 coupled configurations are run for 60 years in order to test the impacts of a longer adjustment in coupled mode. The extended configurations are chosen manually to achieve an even sampling of the net global radiative imbalance in the atmosphere-only simulations.

In comparison of the volume of the initial parameter hypercube, after the 45th iteration in the coupled represents 99.2% has been ruled out, as the coupled simulations includes the ocean and sea ice components (P9-P12). Moreover, among those top 100 vectors from the atmosphere-only simulations, there are 13, 16, 16, 18, and 37 respectively from wave 41, 42, 43, 44, and 45.

2.5 Reanalyses and Observational Datasets

For additional comparison, the NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration Version 4 (Fetterer et al., 2025) for years 1980 to 2014 are used to compare the coupled model sea ice extent. We additionally include ERSST v5 from 1980 to 2014 to compare the SST of the coupled model (Huang et al., 2017).

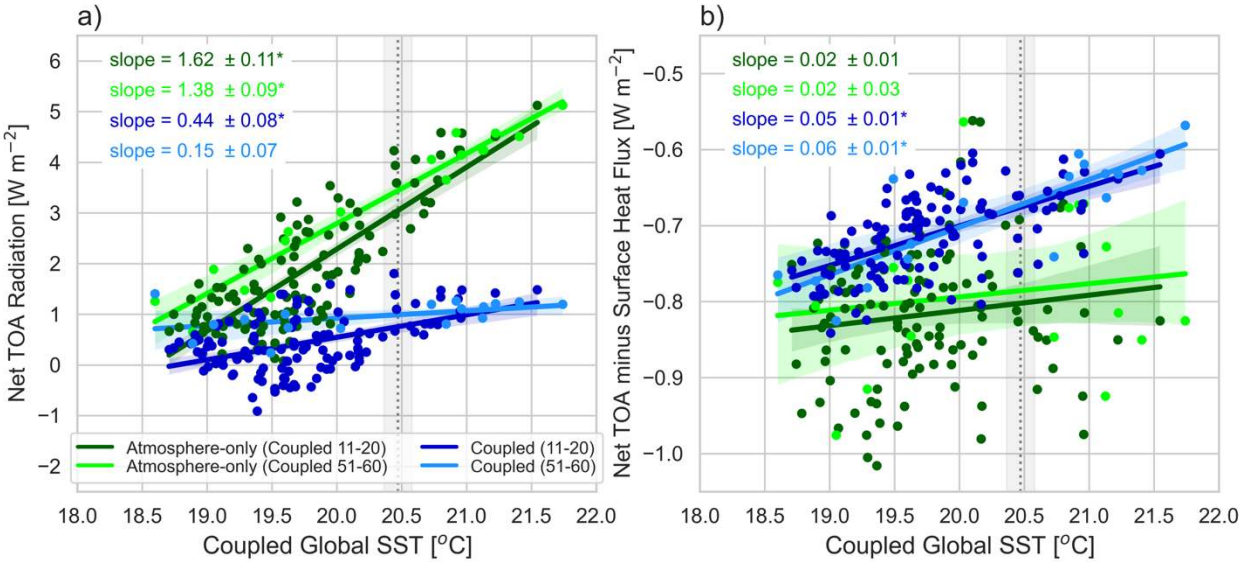


Figure 1: Regressions of globally averaged annual mean a) net top of the atmosphere (TOA) radiation and b) net TOA minus net surface radiation plotted against the Coupled globally averaged sea surface temperature (SST). Atmosphere-only values represent year 2 of simulation plotted against the corresponding Coupled simulations for years 11-20 and years 51-60. The vertical dashed gray line represents EN4 between 1980 and 2014 and its uncertainty. Global SST is defined as 60ºS-60ºN.

3 Results & Discussion

3.1. Global energetic and temperature adjustment

We begin by examining the adjustment of the top-of-atmosphere (TOA) radiative imbalance (net radiative flux) and global SST (60ºS-60ºN) across the 120-member ensembles. The TOA imbalance is defined as the globally averaged net radiative flux at the TOA. In the atmosphere-only PPE, which uses prescribed SSTs, the TOA imbalance ranges from 0 to 5 W m^{-2} (Fig. 1a). This is not by chance. As explained above, the center of the distribution is around 2.7 W/m^2, which corresponds to the target of the TOA imbalance in the tuning (see section 2.3). As several parameters target cloud physics, vertical displacement, and longevity of poleward distribution (P1-6), this TOA imbalance adjustment occurs mainly through changes in the cloud radiative effect (CRE).

In Fig. 1a the radiative imbalance in atmosphere-only PPE is plotted against the global SST of the coupled PPE for years 11-20 (dark green circles) and 51-60 (light green circle). The global SST of the coupled-PPE is linearly related with the TOA imbalance found in the corresponding atmosphere-only simulations. Although this finding was expected and has already been used in previous studies (e.g. Mignot et al. 2021, Hourdin et al. 2023), it is the first time it is formally verified in a twin-PPE protocol. It confirms that the parametric spread in

atmospheric stand-alone global radiative imbalance can be used to anticipate the global SST responses in the equivalent coupled simulations. The black vertical line in Fig. 1a represents the mean SST from observations. Using a regression, we find that in this version of the IPSL-CM model, a radiative imbalance of 3.1 ± 0.1 W m^{-2} (3.4 ± 0.1 W m^{-2}) is optimal to simulate the observed mean SST for year 11-20 (51-60). This verifies a posteriori the choice for the value of 2.7 ± 1 W m-2 used previously for the tuning (Hourdin et al. 2023). The difference obtained when using year 11-20 and 51-60 can be explained by the smaller sampling for the year 51-60, and by a non converged temperature adjustment in years 11-20 when compared to years 51-60.

The TOA imbalance of the coupled experiments (Fig. 1a blue circles) can be used to evaluate the adjustments of the simulations. As expected, it is much weaker than in the corresponding stand-alone configurations as the SST adjusts so as to reduce the TOA imbalance. Assuming no changes in the boundary conditions (i.e., external forcings) and ignoring the effect of internal climate variability, the TOA imbalance in coupled mode should approach zero once the ocean is equilibrated. Here, the TOA imbalance does not fully cancel in the coupled simulation, for two possible reasons. The first reason may be the insufficient length of the coupled experiments so that the ocean is not yet adjusted. The regression slope for years 51–60 is indeed approximately one-third of that for years 11–20. The regression is not significant at the 5% level for years 51-60, meaning that all the coupled PPE members have converged toward the same net TOA radiation imbalance. Yet, they have not converged towards zero but the coupled TOA radiation is on average ~ 0.9 W m^{-2} in the PPE. This residual offset leads into the second point: the atmospheric model does not fully conserve energy, with an energy numerical sink of about 0.7 to 1 W/m2 as found by Irving et al. (2021) or Mignot et al. (2021). This is confirmed by the difference between the TOA and surface (including land and glaciers) energy fluxes in the Atmosphere-only PPE (Fig. 1b, green dots): The difference is spread between 0.7 and1 W/m2, with no significant correlation with the associated SST in coupled mode. It is thus a purely atmospheric internal leakage.

3.2 Vertical thermal structure of the atmosphere

To investigate in more detail the vertical thermal adjustment of the atmosphere in the coupled PPEs, we first show the regression of the zonally averaged air temperature anomalies on the global mean SST (Fig. 2a). It shows that members with an anomalously warm SST show relatively warm temperatures throughout the whole troposphere, and anomalies of opposite sign in the stratosphere. The anomalous warming is strongest in the tropical upper troposphere, with another maximum in the lower-troposphere north of 70°N, as seen in global warming projections (Lee et al. 2021). The stratospheric cooling has a tripolar shape, with cooling above 100-hPa in the tropics, and at 200-hPa poleward of 40° in both hemispheres. This pattern is consistent with past analyses of stratosphere–troposphere coupling in response to radiative forcing changes (e.g. Thompson and Solomon, 2005). The occurrence of stratospheric cooling associated to tropospheric warming in the tropics is consistent with the quasi-steady, quasi-balanced response of the stratosphere to tropospheric heating (Lin and Emanuel 2024a, 2024b). The stratospheric cooling at midlatitudes might reflect the changes induced in the Brewer-Dobson circulation, as well as the modification of the radiative cooling linked to clouds at the tropopause. Modifications of the upward propagating planetary waves can also explain these anomalies through the downward control theory (Haynes, et al., 1991). Quantitatively, the

majority of the air temperature variance across the PPE in the troposphere is explained by SST dispersion, with R^2 values exceeding 0.7 across most vertical levels, and approaching 1.0 in some layers (Fig. 2b). This result confirms that the tropospheric spread in the coupled configuration is mainly attributable to ocean-mediated adjustments. Comparatively, the lower stratosphere and the tropopause have notably weaker R^2 values up to 0.5 in the mid-latitudes near 150 hPa. This suggests that the upper-atmosphere PPE spread is influenced by other processes than direct SST-adjustments.

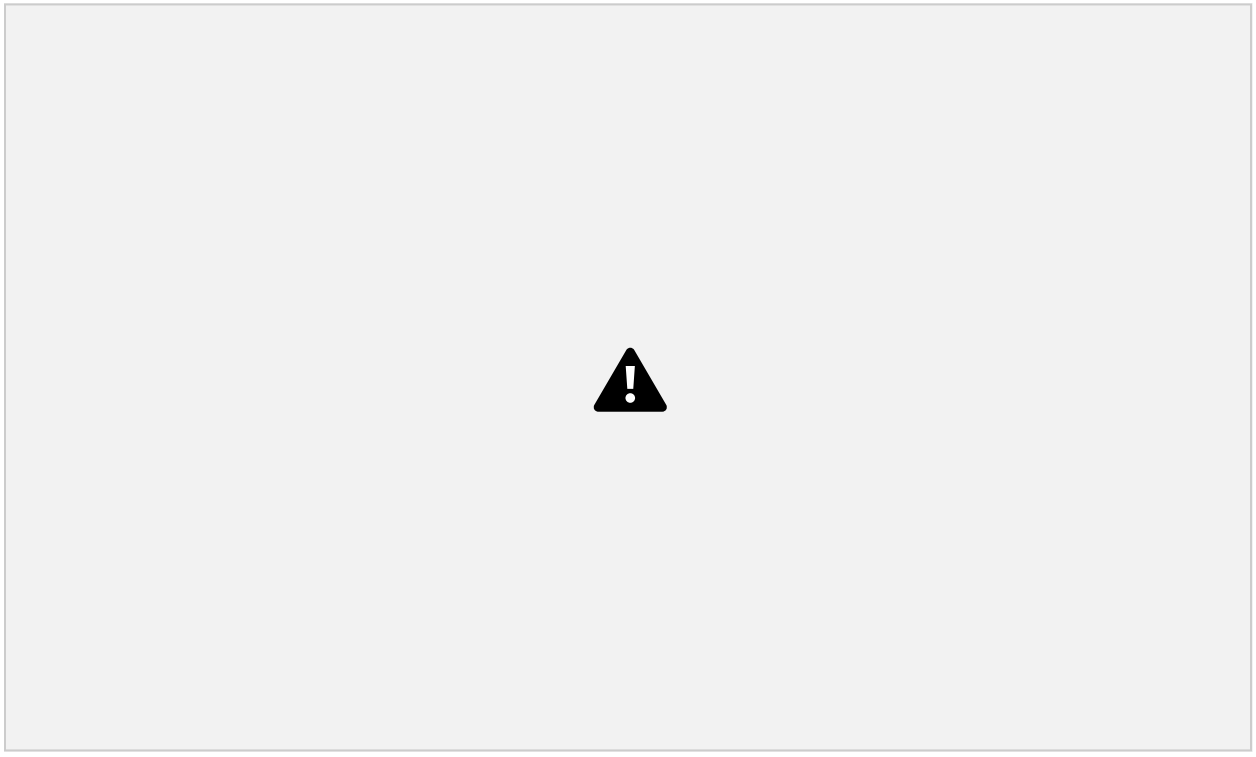


Figure 2: a) Slope and b) explained variance of the regression of zonal mean temperature on the global mean SST. c) The EOF1, shown as covariance of the annual mean zonally averaged temperature across the ensemble of Coupled simulations. d) Same as c) for the residue of the regression coupled PPE on the global mean SST. e) Same as c) for the atmosphere-only PPE. Similar to (c), EOF1 shown as covariance in the (d) Coupled $SST_{removed}$ and (e) Atmosphere-only. Atmosphere-only simulations represent year 2 and the coupled years 11-20. f) Scatter plot of the PC1 of the zonal mean temperature from the coupled PPE (blue) and the coupled $SST_{removed}$ PPE (orange) versus the same PC1 from the atmosphere-only PPE, demonstrating the similarity between each member of the PPE. Regression slopes that are significant at the 95%-confidence are noted with a *. Plots c, d, and h) all include the variance explained through the EOF1 in the top right.

Another approach to characterizing the air temperature variation across a PPE is to apply an empirical orthogonal function (EOF) analysis. The first EOF mode, EOF1, explains 89% of the variance of the coupled ensemble (Fig. 2c). Its pattern closely matches the regression shown in Fig. 2a, showing that the dominant mode of the PPE corresponds to the SST driven adjustments. Members showing a heavy load in this pattern (high values of the corresponding principal component) correspond to the anomalously warm members in terms of SST and thus of global mean SST in Fig. 1. The coupled PC1 is indeed significantly correlated (r=0.99) with the global mean SST, with a slope of 0.6/°C. This linear correlation emphasizes the strong linearity of the

global mean SST adjustment in shaping the anomalous vertical temperature structure along the EOF1 pattern in the coupled PPE.

To identify sources of dispersion beyond the response to the global SST, we compute residuals by subtracting from any analyzed field its regression against the global SST in the coupled PPE. The residual fields obtained are hereafter called $SST_{removed}$ fields. Fig. 2d shows EOF1 of the $SST_{removed}$ zonal mean air temperature in the coupled PPE. The resulting pattern shows a modest warming throughout the lower-troposphere at 40° and polar regions and is associated with two large cooling patterns in the tropopause and lower stratosphere poleward of 40° in both hemispheres.

An EOF analysis is also conducted on the atmosphere-only PPE (Fig. 2e). EOF1 patterns of the atmosphere only and $SST_{removed}$ coupled mean air temperature are very similar to each other, both in vertical structure and spatial amplitude (Fig. 2e). The associated first principal components (PC1) from the coupled-$SST_{removed}$ ensemble are also highly correlated with those from atmospheric-stand-alone ensemble (Fig. 2f, r = 0.87; orange). On the contrary, there is only a weak correlation between the PC1 from the raw coupled PPE outputs and that of the atmosphere-only PPE (r = 0.36; dark blue). This indicates that, beyond the direct thermal adjustment of the coupled model linked to the response of the global SST to the global radiative imbalance that accounts for most of the inter-member spread, the vertical atmospheric thermal structure in the coupled PPE reflects a parametric adjustment that closely resembles the one found in the atmosphere-only runs. EOF1 of the coupled-$SST_{removed}$ explains ~51% of the PPE variance while EOF1 of the atmosphere-only PPE explains ~23% of the variance.

3.3. Linking ensemble spread to cloudiness

To identify the parameters responsible for driving the spread among the configurations, we compute the correlation between the PC1 of the zonal mean air temperature and each model parameter values across the PPEs (Fig. 3). In the raw coupled PPE, *CLC* (P1; auto-conversion threshold for liquid cloud water) and *RNLC* (P12; non-shear-driven Langmuir vertical mixing in the ocean) emerge as dominant drivers (r > 0.6 and 0.5, respectively) of the intermember spread. *(P1) CLC* is associated with tropical, low-level clouds forming at the top of the atmospheric boundary-layer. Increasing *(P1) CLC* value increases the low cloud cover and thereby cools the climate through an increasing planetary albedo. This result retrospectively justifies the choice, based on expert judgment, to use CLC as the final adjustment parameter for the global temperature or SST of the IPSL-CM. (Mignot el al., 2021). On the other hand, *(P12) RNLC* governs upper ocean mixing processes, with increasing *(P12) RNLC* causing increasing mixing in the upper-ocean and cooling the SST. Both of these parameters strongly influence SST, and it is thus coherent that they primarily set the anomalous pattern discussed in Fig. 2c.

In both atmosphere-only and coupled-$SST_{removed}$ PPE, the parameters *FALLV* (P2; fudge factor on the falling velocity of ice crystals) and *OMEPMX* (P3; threshold on the conversion of deep convection condensates to rainfall) exhibit the strongest positive correlations with PC1 (r > 0.8 and 0.6, respectively). These two parameters control the upper-level cloud structure and thus associated radiative responses. Notably, the coupled PPE also retains a significant correlation with *(P2) FALLV* (r > 0.3), highlighting the impact, yet secondary, of this parameter in the adjustment of this ensemble. These findings are robust across alternative ranking methods, including SHAP, Sobol, and Spearman analyses.

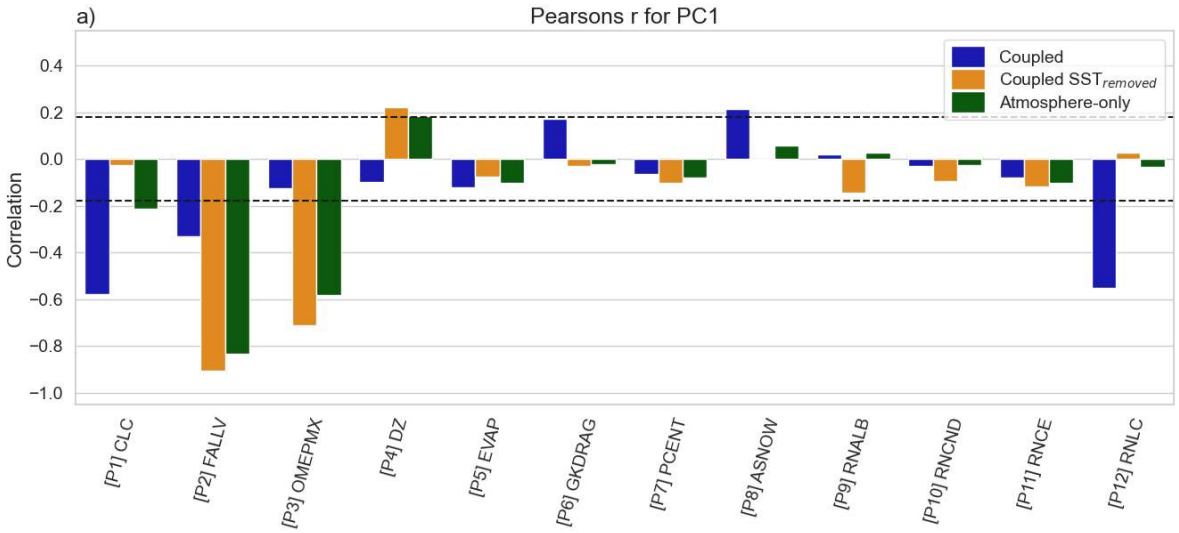


Figure 3: a) Correlation of the values of parameters with the PC1 of zonal mean temperature from the coupled and atmosphere-only PPE. The horizontal dashed lines represent the threshold correlation for a significance level of 0.05.

To further confirm the role of the different types of clouds for the air temperature, we show scatter plots of the globally-averaged high level (Fig. 4b) and low level (Fig. 4a) cloud covers versus the PC1 of the zonal mean temperature across the PPEs. PC1 from the raw coupled PPE shows a positive relationship with globally averaged low-level cloudiness, while the atmosphere-only and coupled-SST$_{removed}$ configurations exhibit weaker, negative regressions (Fig. 4a). This is consistent with the role of low clouds in increasing the planetary albedo, which cools the climate, primarily visible in the coupled ensemble. In contrast, PC1 in both atmospheric and coupled-SST$_{removed}$ PPE shows a strong, negative correlation with high-level cloud fraction (Fig. 4b). The raw coupled ensemble exhibits a weaker negative correlation with high clouds, confirming reduced sensitivity to upper-level cloud changes. This is indeed consistent with the role of *(P2) FALLV* and *(P3) OMEPMX* in modulating the high cloud cover (Fig. 3). This is also consistent with the associated patterns (Fig. 2) from the atmosphere-only and coupled-SST$_{removed}$ PPE, where less high clouds are associated with negative radiative warming in the upper troposphere at the cloud top. Note that the dominant impact of *(P2) FALLV* and *(P3) OMEPMX* in modulating upper-level cloud structure is consistent with prior findings, which discussed how cloud-ice microphysical parameters (i.e., upper level clouds) can drive climate or PPE diversity (Zelinka et al., 2013; Gettelman et al., 2015; Hourdin et al., 2017).

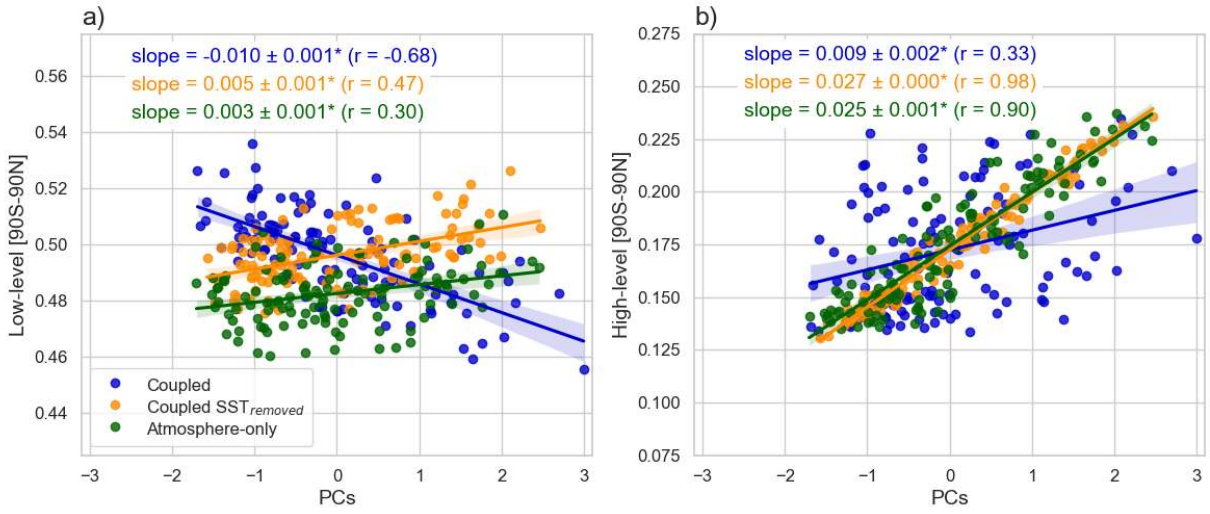


Figure 4: Regressions of annual mean a) low-level and b) high-level cloudiness plotted against the PC 1. Regression slopes that are significant at the 95%-confidence are noted with a *.

These connections can also be explored spatially through the annual mean regressionof the zonal-mean cloud fraction against PC1 (Fig. 5). show in all three PPEs a significant increase in high clouds above 200-hPa in the deep tropics in 15°N-15°S.

An increase is also seen in the descending branches of the residual mean atmospheric circulation connecting the tropics to mid-latitude and polar regions between 600-hPa and 200-hPa poleward for 30° in both hemispheres. Again, this is consistent with Fig. 4b and the influence of *(P2) FALLV* and *(P3) OMEPMX* that set the quantity of high clouds formed in the deep tropics convective regions, and the residence time of the ice cloud formed. In the coupled configuration, though, an additional feature can be seen: PC1 also relates to a decrease in low-level cloudiness below 800 hPa, most notably in the tropics and mid-latitudes. This pattern extends upward to 200 hPa in the tropical regions. This is consistent with the strong response of the PPE to SST adjustment (Fig. 2) itself largely driven by low-level clouds (Fig. 4) and the parameter *(P1) CLC* (Fig. 3). The same analysis was repeated using the seasonal mean cloudiness averaged in January-February-March (JFM) and June-July-August (JJA) (see Fig. SX1). The changes of mid and high clouds are systematically larger in the winter hemisphere, as the detrainment of the clouds from the tropics by the Hadley circulation is dominant through the winter hemisphere. The low clouds modification also shows some seasonal changes, with an increase larger in the summer mid-latitudes than in the winter mid-latitude.

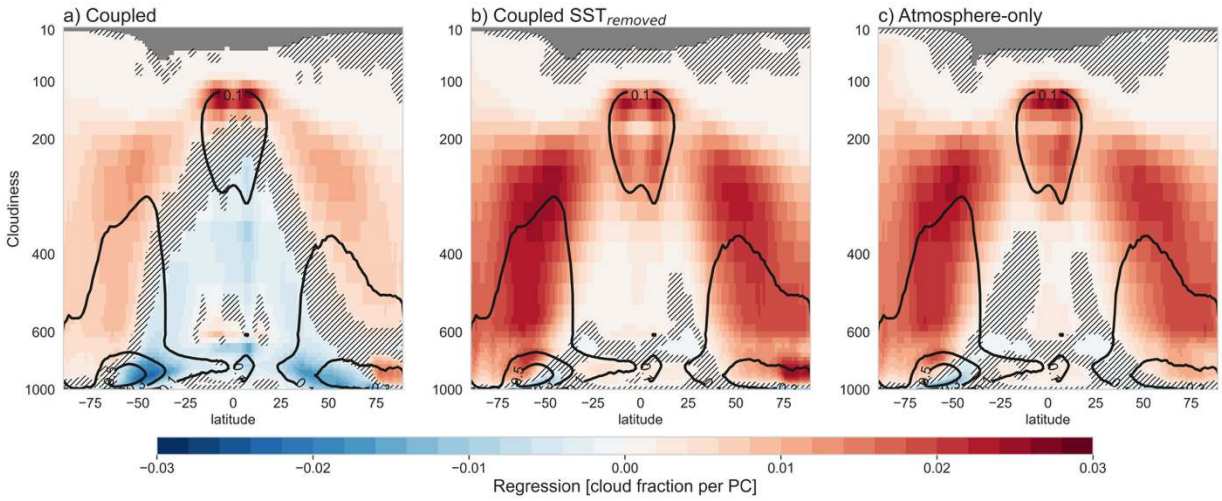


Figure 5: Regressions of zonally averaged (a-c) annual mean cloud fraction and the PC 1 of the Coupled, Coupled SST$_{removed}$, and Atmosphere-only. Seasonal regressions are shown in Figure SX1. Hatching represents values not significant at the 95% confidence level.

Taken together, these results show that the anomalous atmospheric thermal vertical structure in the atmospheric stand-alone PPE is linked to its anomalous high cloud cover. This adjustment is transferable to the coupled model but appears as second order after the adjustment to parameters controlling low level clouds covers and the global mean SST adjustment.

3.4 Climate impacts of parametric modulations

To understand the effect of clouds on the climate, we evaluate the TOA cloud radiative effect (CRE), defined by the difference in net TOA radiation between the clear-sky minus all-sky calculations. The CRE thus corresponds to how the atmosphere reflects the balance between cloud longwave warming and shortwave cooling. Regressions of PC1 onto TOA net CRE (Fig. 6a-c) for the coupled PPE reveals large positive net CRE over mid-latitude oceans as well as in the tropical western Atlantic and eastern Pacific basins. The tropical regions are strongly affected by the changes in low cloud cover, with a positive net SW CRE seen from space when low clouds are more frequent, the low clouds having almost no direct effect on the TOA LW radiation. The increase in high clouds reduces LW radiation to space strongly, because of the strong contrast in temperature between the high clouds altitude and the mid troposphere where clear sky LW radiation comes from. It also creates a strong greenhouse effect at the surface.

In the SST$_{removed}$ and atmosphere-only PPE, the CRE is positive in the mid-latitude to polar regions and negative in the tropics. The positive relationship with mid-latitude corresponds to the effect of increasing high clouds, which warms the mid-latitude regions, as their effect on trapping long-wave radiation increases. It is weaker in the deep tropics, in particular over the maritime continent or over the equatorial Indo-Pacific and Atlantic and poleward of 70°N in both hemispheres.

The analysis was repeated using seasonal means (Fig. S2). It reveals that the positive CRE poleward of 70°N in the atmosphere-only and SST$_{removed}$ PPE is dominant during winter and absent in summer. This is consistent with the larger cloud changes in the winter hemisphere.

More generally, these signals of high latitude CRE aligns with previous studies emphasizing the critical role of upper-level clouds in regulating the polar energy budget and surface temperature (Crook et al., 2011; Taylor et al., 2013; Payne et al., 2015; Stuecker et al., 2018). The coupled $SST_{removed}$ and atmosphere-only PPE SST signature is indeed dominated by the significant warming signals in the Southern Ocean and Arctic regions. This warming is amplified in the winter hemisphere (Fig. S2). To conclude, the atmosphere-only and coupled $SST_{removed}$ PPE indeed agree with a dominant influence of high-clouds in the winter hemisphere that has a large influence for the polar surface air temperature. The divergence of net CRE and SAT between the coupled and coupled $SST_{removed}$ PPE within the mid-latitudes and tropics highlights that SST pattern feedbacks significantly influence the mean state and spatial structure of climate sensitivity (Proistosescu et al., 2018; Dong et al., 2020).

Warming in the polar region can reach values approaching 4ºC per unit of PC in the coupled PPE (Fig. 6a). This result emphasizes that the mid-latitude to polar regions of the coupled PPE reflect the atmospheric response, independent of the tropics and SST-feedbacks. Moreover, we speculate that these strong changes are in part controlled by the net CRE and global climate adjustment to global warming, with the characteristic polar amplification in coupled models, based on the relationships of the PPEs' PC and the relationships with the parameters and cloud responses. For example, results from Fig. 3 suggest that the link between PC1 and CRE is likely controlled by *(P2) FALLV*. In all coupled, coupled-$SST_{removed}$, and atmosphere-only PPEs, increasing *(P2) FALLV* indeed reduces the annual mean net CRE and cools the surface throughout the tropics, mid-latitudes, and polar regions (Fig. S3). There are not significant areas in both the coupled and coupled $SST_{removed}$ in the polar regions compared to the atmosphere-only net CRE results, likely related to non-linearities in the PPE diversity of sea ice extent. Taken globally, these findings confirm the physically consistent pathway that occurs across coupled, coupled-$SST_{removed}$, and atmosphere-only simulations by which a single parameter drives upper-level cloud changes that affect net global cloud radiative effect and the temperature surface. This underscores how an atmosphere-only calibrated parameter like *(P2) FALLV* can help precondition the polar regions net CRE and surface temperature behavior in coupled simulations through upper-level cloud cover.

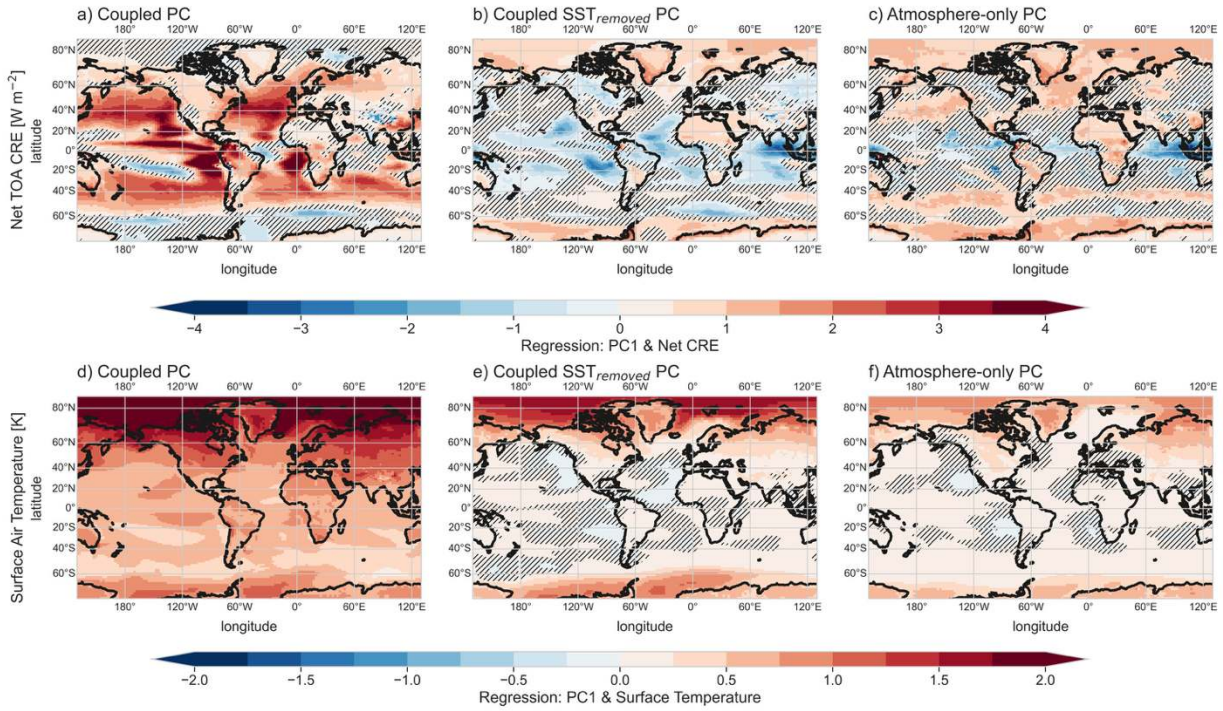


Figure 6: Regressions of (a-c) annual mean net CRE and the PC 1 of the coupled, coupled $SST_{removed}$, and atmosphere-only. Regressions of (d-f) annual mean surface temperature and the PC 1 of the coupled, coupled $SST_{removed}$, and atmosphere-only. The atmosphere-only comparisons represent year 2 and the coupled years 11-20. Hatching represents values not significant at the 95% confidence level.

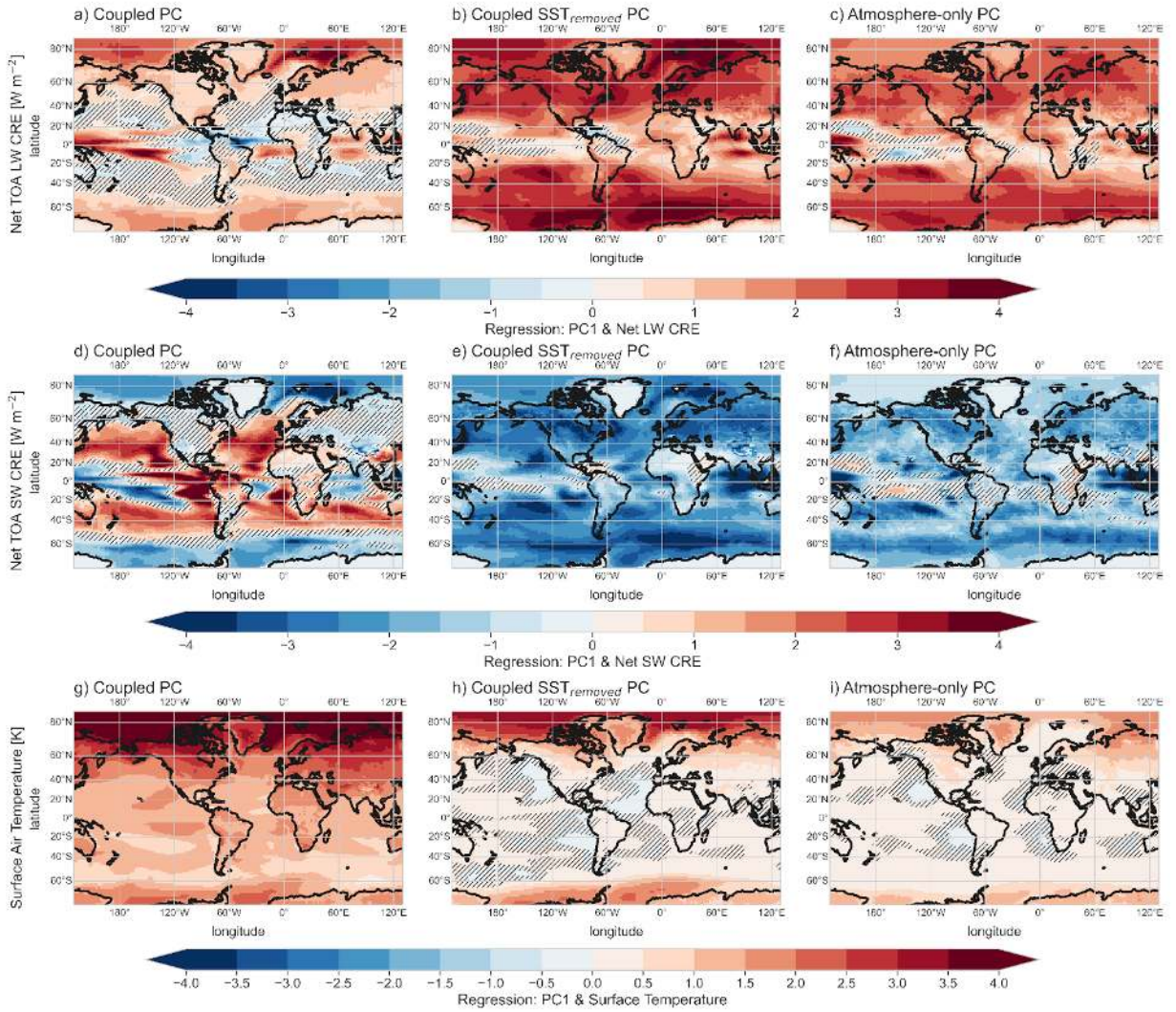


Figure 6 Alternate: LW -> clear signal in polar regions of all, SSTremoved and Atmosphere-only similar spatial pattern. SW has features of sea ice in coupled, otherwise SSTremoved and Atmosphere-only similar. Similar pattern and agrees to cloudiness? SW CRE by low clouds and ice, LW CRE by high clouds?

To further explore the climate responses in the PPEs, regressions of zonally-averaged zonal wind velocity and PC1 (Fig. S4) in both coupled and coupled-SST$_{removed}$ are performed Here, the atmosphere-only PPE is not investigated as we suspect that the duration of one year is insufficient to capture the wind stress change, the internal climate variability having a large influence. The regressions show a poleward shift of the subtropical jets and of the eddy-driven jet in the Southern Hemisphere associated to positive loads of PC1. The zonal wind is also strongly modified in the stratosphere in thermal wind balance with the temperature changes. The polar jet is intensified especially over its poleward flank which further suggest associated changes in the Brewer-Dobson circulation. In the Northern Hemisphere, the raw coupled ensemble shows an enhanced magnitude compared to the coupled SST$_{removed}$, likely due to broader meridional SST and troposphere temperature gradients in the coupled EOF1 pattern. Still, the vertical structure and position of the jet anomalies the coupled and SST$_{removed}$

configurations share some common points, suggesting that zonal winds, both surface and upper-level, can in part be preconditioned using atmosphere-only simulations. In both regressions, the Southern Hemisphere jet has a larger magnitude and reaches the surface in JFM months compared to JAS months. The Southern Hemisphere relationship is weaker and constrained to the lower-stratosphere during JAS months. Moreover, the relationship with seasonality of both hemispheres in the regressions further demonstrates the similarities in PPEs, mostly in the spatial patterns above 200 hPa. This pattern is true in the extra-tropics, particularly well in the Southern Hemisphere. Again, the magnitude between coupled and $SST_{removed}$ do not match. These results suggest a positive PC based on the EOF of a "cooler" stratosphere of the coupled $SST_{removed}$ (and thus atmosphere-only) is related to a more poleward position of the mid-latitude jets and the seasonality across the PPEs, with emphasis at vertical heights near the tropopause and lower-stratosphere.

Fig. S5 extends this analysis of PC1 to zonal surface wind stress, in the coupled and coupled $SST_{removed}$. The relationships between the PC1 wind stress regressions show strong (moderate) agreement between the raw coupled and coupled-$SST_{removed}$, particularly in the Southern (Northern) Hemisphere mid-latitudes in JFM. However, the magnitude of regression is weaker in the coupled-$SST_{removed}$ compared to the coupled. This hemispheric difference is in agreement with the zonally averaged zonal wind relationship with PC1 (Fig. S4) in that there is a stronger spatial similarity in the Southern Hemisphere than the Northern Hemisphere and that in JFM months the relationship of the jet reaches the surface compared to JAS months.

Additionally, the changes agree with the zonal mean temperature anomalous pattern that showed warming over the whole atmospheric column at the edges of the tropical band, showing an enhanced tropical zone, with atmospheric eddy-driven jets shifted poleward. In the coupled PPE, this enhanced tropical band is reinforced by the global warming effect (Lu et al. 2007; Gastineau et al. 2008).

3.5 Calibrating the sea ice extent in climate models

Recent studies have shown that surface winds can modulate the sea ice extent in the Southern Hemisphere (Blanchard-Wrigglesworth et al., 2021; Roach et al., 2023; Cesana et al., 2025), and can also influence the CRE (Cesana et al., 2025). Sea ice extent is thus a critical variable to adjust in climate models. We now compare the sea ice cover adjustment in the various ensembles. Regressions of PC1 onto sea ice concentration in the Northern (Fig. 7a–c) and Southern (Fig. 7e–g) Hemispheres reveal robust seasonal and regional signals. In the Northern Hemisphere, the coupled and $SST_{removed}$ coupled PPEs both show a significant negative regression in the Nordic and Barents Seas, as well as the Bering Sea during the boreal winter (Fig. 7a,e). They differ in the western subpolar and Labrador Sea, the $SST_{removed}$ PPE showing a significant negative relationship where the coupled is not significant. In the boreal summer (Fig. 7b,f) both PPEs indicate a negative regression with PC1 with the summer Arctic sea ice concertation.

The Southern Hemisphere austral summer (Fig. 7c,g) shows similar negative patterns spatially in both the coupled and $SST_{removed}$ coupled regressions, with strong negative similarities in the Ross and Weddell Seas and the sea ice extent. The austral winter (Fig. 7d,h) indicates stronger differences. The $SST_{removed}$ sea ice regression with the corresponding PC1 demonstrates a strong relationship with the sea ice extent across the Atlantic sector, but spatially more local in

the fully coupled ensemble. In the austral winter, the spatial pattern in the eastern Pacific sector of the Southern Hemisphere is different between the two PPEs, which shows a weaker relationship in the coupled PPE to that of the negative regressions in the Eastern Pacific of the $SST_{removed}$. Interesting to note, both the boreal (Fig. 7b,f) and austral (Fig. 7c,g) summer regressions indicate that the coupled relationships are twice as strong as the $SST_{removed}$ and have similar spatial patterns. However, both boreal (Fig. 7a,e) and austral (Fig. 7d,h) show less coherence between the coupled and coupled $SST_{removed}$, with the $SST_{removed}$ regressions being significantly stronger in magnitude and covering a larger area.

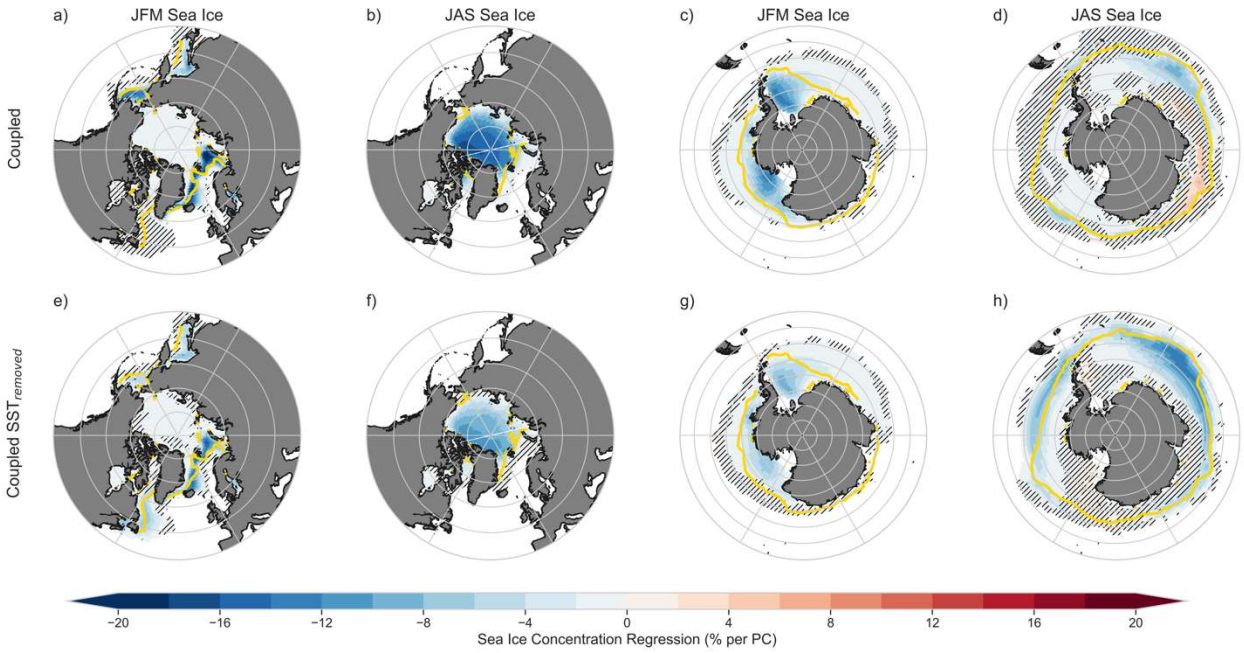
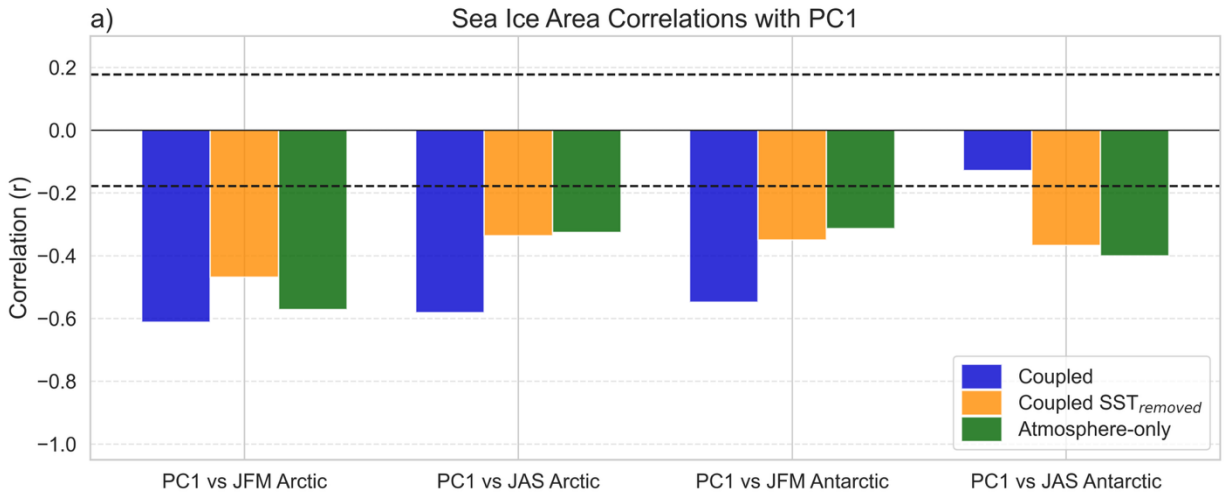


Figure 7: Regressions of PCs onto Coupled JFM (a-c) sea ice concentration and d) Arctic sea ice area. (e-h) Similar, but for Coupled JAS sea ice concentration and Antarctic sea ice area. The Coupled PCs years 11-20 and the Coupled $SST_{removed}$ are regressed on the sea ice concentration field with the SST-effect removed. The yellow contour represents the 1980-2000 NOAA NSIDC v4 0.15 sea ice extent. Hatching represents values not significant at the 95% confidence level.

Fig. 8 further demonstrates this relationship with sea ice extent by correlating the PC1 with the Arctic and Antarctic winter and summer total sea ice extents. To include the atmosphere-only the correlation compares the atmosphere-only PC1 with the sea ice concentration of the $SST_{removed}$ analysis. Thereby, the atmosphere-only PC1 is included to show the sea ice extent that could have been obtained with the atmosphere-only configuration and as a benchmark to compare with the coupled and $SST_{removed}$ PPEs. In the Arctic, the raw coupled PC1 indicates moderate negative correlations ($r \sim -0.6$) in both boreal winter and summer, while the coupled $SST_{removed}$ and atmosphere-only being weaker in winter ($r < -0.45$) and summer ($r < -0.3$), but still significant. In the Antarctic, the raw coupled reveals moderate negative correlations ($r \sim -0.5$) in austral summer, but is not significant in austral winter. Both the coupled $SST_{removed}$ and atmosphere-only PC1 are significant in austral summer and winter ($r < -0.3$).

While the general negative anomalies in sea ice extent are consistent with the global troposphere warming pattern found in the coupled EOF1, the specific anomalies in the coupled

585 $SST_{removed}$ and atmosphere-only PPE can be attributed to the enhanced high-cloud cover. In the
586 coupled $SST_{removed}$, more high clouds increase the longwave CRE in winter (Fig. S2), which
587 promotes sea ice decline. Negative anomalies in sea ice volume and extent can persist through
588 the summer and manifest as reduced sea ice extent climate states in the PPE. The wind
589 anomalies might play an additional role in the Southern Hemisphere, with enhanced westerlies
590 in JAS increasing the sea ice Ekman drift toward the equator, and increasing the upwelling of
591 relatively warm circumpolar deep waters (ref).



592

593 Figure 8: a) Pearson's r of the seasonal Arctic and Antarctic sea ice area with PC 1. The coupled $SST_{removed}$ are
594 correlated with the sea ice concentration field with the SST-effect removed. To compare the
595 atmosphere-only, the PC1 are correlated with the coupled $SST_{removed}$ sea ice detrended field.

596 These results demonstrate the importance of the zonally averaged atmospheric
597 temperature pattern, shown as EOF1 in Fig. 2, on Arctic sea ice area and Antarctic austral
598 summer sea ice area in the coupled PPE. In the context of relating the PC1 with the EOF1
599 pattern, the regression with sea ice extent for coupled $SST_{removed}$ suggests a significant
600 relationship to upper troposphere ensemble diversity. The similarities between the coupled and
601 $SST_{removed}$ ensembles suggests that upper-level cloud and CRE processes originating in
602 atmosphere-only can influence the cryospheric state in the coupled model (see also Cesena et
603 al., 2025, on Antarctic cloud-ice interactions), independent of SST-feedbacks and ocean and sea
604 ice parameters used in a coupled model. This result stresses how the upper-atmosphere and
605 AMIP configurations can significantly contribute to preconditioning the polar climates through
606 the spatial temperature structure and cloud feedbacks.

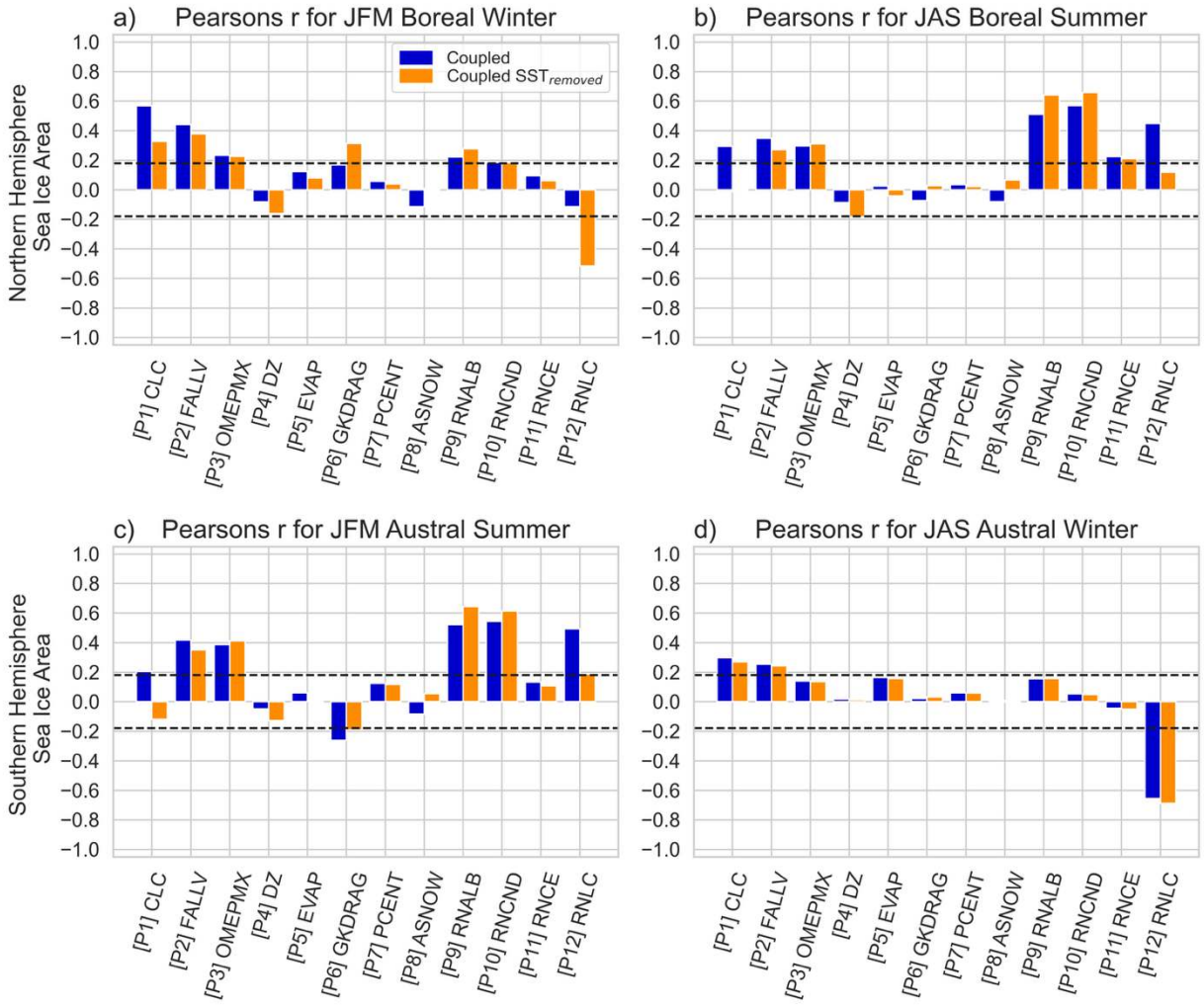


Figure 9: Similar to Figure 3, but Pearson's r correlation for Arctic and Antarctic JFM and JAS sea ice area and the 12 parameters used in the PPE.

4. Implications for Climate Model Tuning & Conclusion

This study evaluates whether atmosphere-only simulations can capture and precondition the diversity expressed in coupled climate simulations, using a PPE constructed with identical atmospheric parameter perturbations. By comparing atmosphere-only, coupled, and coupled $SST_{removed}$ ensembles across TOA radiative imbalance, cloud structure, parameter sensitivity, and surface climate responses, we show that a large fraction of the coupled model's atmospheric diversity is already expressed in atmosphere-only, particularly in the upper troposphere and lower stratosphere and in the mid and high latitudes of both hemispheres. The additional spread in coupled runs is largely due to global thermal adjustment and surface-driven feedbacks.

An EOF analysis reveals that the leading mode of variability (PC1) is consistently reproduced across atmosphere-only and coupled-$SST_{removed}$ ensembles, with clear upper troposphere and lower stratosphere signal across the PPE. This EOF1 signal further relates to the extra-tropical jet responses in the PPEs, reflecting stratosphere–troposphere coupling

processes (Baldwin et al., 2024). These localized relationships suggest that the upper troposphere could be a good region to identify additional metrics for the atmosphere-only to precondition the coupled model. Several atmospheric parameters, especially those related to upper-level cloud microphysics (e.g., ice auto-conversion, cloud lifetime), exert a persistent influence on ensemble spread in both atmosphere-only and coupled frameworks. This strengthens the case for targeting upper-tropospheric diagnostics in calibration, while also underscoring that tuning efforts must be grounded in physical understanding to avoid compensating errors (Sherwood et al., 2020).

Among these diagnostics (see Fig. S6), the Northern Hemisphere midlatitude temperature at 150 hPa, which is identified through the EOF spatial pattern that shows a near 1:1 relationship between atmosphere-only and coupled simulations PC1. The temperature metric shows a positive correlation in all three coupled, $SST_{removed}$, and atmosphere-only. This would suggest that when the North Hemisphere midlatitude stratosphere is "warmer", there is more Arctic sea ice seasonally. This same relationship is true for the Southern Hemisphere and the Antarctic sea ice extent. The correlations range from 0.35 to 0.50 in the North Hemisphere and Southern Hemisphere JFM months. The correlations are weaker in the Southern Hemisphere for JAS months (~0.25).

Additionally, LW CRE estimated globally or in the polar regions further provide significant correlations. The hemispheric polar values provide a higher correlation than the globally-averaged LW CRE. The coupled, $SST_{removed}$, and atmosphere-only all suggest a significant negative relationship, so if the LW CRE increases, the sea ice extent decreases. Correlations range from -0.35 to -0.70, except in the Southern Hemisphere for JAS months (~0.3). Finally, the net global radiative imbalance, which is already a metric within our approach and Hourdin et al., (2023), indicates strong correlations with winter sea ice extent in both hemispheres. Correlations in the coupled and $SST_{removed}$ exceed -0.75, suggesting in the coupled model there is a strong dependence of winter sea ice extent to the net global radiative imbalance. The correlations are weaker in the atmosphere-only, but still significant.

These metrics could further provide added value as to constrain the model to be within the given error tolerance. Moreover, these metrics are comparable in fidelity to TOA-based metrics and largely insensitive to ocean feedbacks (Hourdin et al., 2023). In contrast, precipitation-based diagnostics have not been explored due to the high frequency variability relative to the simulation length and have been shown to exhibit poor atmosphere-only to coupled alignment (Ceppi & Gregory, 2017; Mikkelsen et al., 2025).

The implications extend to polar regions. Sea ice regressions demonstrate that coupled $SST_{removed}$ ensemble PC1 structure (and therefore linked to the AMIP) projects onto coupled sea ice patterns in the Weddell and Ross Seas of the Antarctic and Nordic Seas of Arctic, validating the ability of AMIP diagnostics to precondition polar climate states (i.e., sea ice extent) even in the presence of ocean feedbacks. This is consistent with recent findings showing that high-cloud processes and upper-tropospheric radiative feedbacks are stronger predictors of high-latitude diversity and sea ice extent than precipitation-based metrics (Cesena et al., 2025). Related work further links lower-stratospheric temperature biases to sea ice extent and air–sea coupling strength (Studholme et al., 2025), highlighting the importance of upper-tropospheric and lower-stratospheric layers as effective tuning targets. Here, applying both atmosphere, ocean,

667 and sea ice parameters in the PPE, we demonstrate how seasonally the atmospheric state can
668 be vital in preconditioning the sea ice extent.
669 Together, these findings extend the PPE framework proposed by Hourdin et al. (2023) by
670 incorporating vertically and spatially resolved diagnostics. AMIP-based PPEs emerge as a
671 practical tool for isolating robust atmospheric parameter impacts, preconditioning coupled
672 simulations, and reducing reliance on costly manual tuning (Murphy et al., 2004; Sanderson et
673 al., 2008). The demonstrated AMIP–coupled alignment in key diagnostics, combined with tools
674 such as hightune-explore, provides a pathway toward reproducible, systematic model
675 calibration and more reliable projections of polar climate change ahead of CMIP7.
676

677

(All figures and tables should be cited in order. For initial submission, please embed figures, tables, and their captions within the main text near where they are cited. At revision, figures should be uploaded separately, as we need separate files for production. Tables and all captions should be moved to the end of the file.)

References should use a name-date format, not numbers. Enclose citations in parentheses with authors in upright text (non italics) as in: (Smith et al., 2009) or Smith et al. (2009). More information on in-text citations can be found in our Brief Style Guide,"Reference Formatting."

Acknowledgments

B.S.F. acknowledges support from the European Union as part of the EPOC project (Explaining and Predicting the Ocean Conveyor; grant number: 101059547). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Open Research

AGU requires an Availability Statement for the underlying data and code needed to understand, evaluate, and build upon the reported research at the time of peer review and publication. Additionally, authors should include an Availability Statement for the software and other research products that have a significant impact on the research. Details and templates are in the Availability Statement section of the Data & Software for Authors Guidance. For physical samples, use the IGSN persistent identifier, see the International Geo Sample Numbers section.

References