Anthropogenic Changes in Interannual-to-Decadal Climate Variability in CMIP6 Multiensemble Simulations

ARTHUR COQUEREAU,[a] FLORIAN SÉVELLEC,[a,b] THIERRY HUCK,[a] JOËL J.-M. HIRSCHI,[c] AND ANTOINE HOCHET[a]

[a] *Laboratoire d'Océanographie Physique et Spatiale, University Brest CNRS IRD Ifremer, Brest, France*
[b] *ODYSSEY Project-Team, INRIA CNRS, Brest, France*
[c] *Marine Systems Modelling, National Oceanography Centre, Southampton, United Kingdom*

(Manuscript received 4 October 2023, in final form 25 March 2024, accepted 2 April 2024)

ABSTRACT: As well as having an impact on the background state of the climate, global warming due to human activities could affect its natural oscillations and internal variability. In this study, we use four initial-condition ensembles from the CMIP6 framework to investigate the potential evolution of internal climate variability under different warming pathways for the twenty-first century. Our results suggest significant changes in natural climate variability and point to two distinct regimes driving these changes. The first is a decrease in internal variability of surface air temperature at high latitudes and all frequencies, associated with a poleward shift and the gradual disappearance of sea ice edges, which we show to be an important component of internal variability. The second is an intensification of the interannual variability of surface air temperature and precipitation at low latitudes, which appears to be associated with El Niño–Southern Oscillation (ENSO). This second regime is particularly alarming because it may contribute to making the climate more unstable and less predictable, with a significant impact on human societies and ecosystems.

KEYWORDS: ENSO; Anthropogenic effects/forcing; Climate; Ensembles; Interannual variability; Internal variability

1. Introduction

Earth's climate is a complex system that varies on a broad spectrum of time and space scales (Baede et al. 2001). This variability is often academically divided into two components. The first component is the forced variability, caused by external factors acting on the climate. Examples include volcanic eruptions that emit large amounts of aerosols that reflect solar radiation or anthropogenic CO_2 emissions that intensify the greenhouse effect. The second component is the internal variability related to Earth system itself and represents the spread of the range of possible system states for a given forcing or background state. A wide range of phenomena at very different scales contribute to this variability, from large climate oscillations such as El Niño–Southern Oscillation (ENSO) to atmospheric cyclones, eddies, or ocean waves. Despite its "internal" label, the latter component has been observed to evolve over time in response to changes in the climate background state and external forcings, as in the example of ENSO and other specific climate modes under global warming (Maher et al. 2018; Cai et al. 2021; Callahan et al. 2021; Coburn and Pryor 2023). These two components of Earth's variability are intertwined in the climate records, posing a challenging problem for scientists seeking to accurately disentangle the forced and internal contributions.

Since the late 1990s, the Coupled Model Intercomparison Project (CMIP) has enabled significant development of fully coupled global climate models (Bock et al. 2020) and facilitated their use to investigate the future of Earth's climate under various emission scenarios. More recently, the initial-condition large ensembles, which gather multiple coupled climate simulations from the same model but starting from different initial conditions, have experienced a rapid spread (Deser et al. 2012, 2020; Lehner et al. 2020; Maher et al. 2021). They allowed, among other things, to make important progress in estimating the intensity of internal variability in a given climate model using ensemble variance as a proxy. While this method has proved useful for extracting the forced signal from simulations, as well as for quantifying the importance of internal variability relative to other sources of uncertainty and its evolution over time (Lehner et al. 2020), less work has been done on the mechanisms linked to these changes in internal variability, and much remains to be learned about their evolution in the context of future warming scenarios.

Given that internal variability is a dominant factor of uncertainty in interannual-to-decadal projections (Hawkins and Sutton 2009; Lehner et al. 2020), it is a key component for realistic and reliable climate predictions. Assessing its evolution under a range of forcing scenarios—particularly in the next few decades—is a critical research objective as it has the potential to impact the predictability of the climate or alternatively temporarily mask or amplify long-term warming trends (e.g., Sévellec and Drijfhout 2018). In the case of masking, it could make potential arguments for minimizing the climate problem and make the case for action against climate change less compelling (as in the context of the global warming hiatus, e.g., Roberts et al. 2015). In the amplifying case, this could worsen the short-term consequences of climate change on societies and ecosystems in particular by causing more intense extreme events (such as more intense heatwaves, e.g., Perkins-Kirkpatrick and Gibson 2017). In this context, a previous study

Corresponding author: Arthur Coquereau, arthur.coquereau@univ-brest.fr

DOI: 10.1175/JCLI-D-23-0606.1

TABLE 1. CMIP6 models used in the study.

Model	Institution	Configuration	No. of members	Ocean/atmosphere (lon × lat)	Equilibrium climate sensitivity	Reference
ACCESS-ESM1-5	CSIRO (Australia)	p1 f1	40	ACCESS-OM2 (1° × 0.6°)/ HadGAM2 (1.87° × 1.25°)	3.9	Ziehn et al. (2020)
CanESM5	CCCma (Canada)	p2 f1	25	NEMO 3.4.1 (1° × 0.6°)/ CanAM5 (2.8° × 2.8°)	5.6	Swart et al. (2019)
MIROC6	MIROC (Japan)	p1 f1	50	COCO 4.9 (1° × 0.7°)/CCSR AGCM (1.4° × 1.4°)	2.6	Tatebe et al. (2019)
MPI-ESM1-2-LR	MPI-M (Germany)	p1 f1	30	MPIOM 1.63 (1.4° × 0.8°)/ ECHAM 6.3 (1.87° × 1.87°)	2.8	Olonscheck et al. (2023)

investigated the evolution of interannual internal variability of temperature using CMIP5 and CMIP6 ensembles over the historical period and under the RCP8.5/SSP5-8.5 scenario (Olonscheck et al. 2021). They detected a clear anthropogenic change emerging at the end of the twenty-first century from these high-emission scenarios and observed a latitudinal pattern of changes in internal variability with increasing variability in the tropics and decreasing at high latitudes. Other studies identified comparable latitudinal changes in variability without using ensemble simulations (Rehfeld et al. 2020; Shi et al. 2023). They estimated the climate variability by filtering or detrending the total signal. This implies making assumptions to remove the trend, corresponding to the forced signal, and makes the results potentially sensitive to these choices.

Building on these results, in the present work, we investigate the variability on interannual-to-decadal time scales, with precipitation in addition to surface temperature and using three different emission scenarios. We investigate how internal climate variability has evolved during this first past period of significant industrial human development (since 1850) and how it might evolve in the future (up to 2100). To this end, rather than studying the evolution of known modes of internal variability in our climate, we choose an objective approach by studying the internal variability signal from a general point of view, without pointing to specific climate modes. We use for this purpose a collection of four initial-condition large ensembles from phase 6 of CMIP to estimate the evolution of internal variability of surface air temperature and precipitation. We focus our analysis on the temporal scales of this variability, extending previous studies to decadal frequencies, and explore the spatial signature of the changes detected. In particular, we describe the underlying spatial patterns of these changes and study the physical mechanisms driving them. In the first section, the methods and datasets are described, whereas the results are presented in section 3. Discussion and conclusions follow in section 4.

2. Materials and methods

a. *Materials*

SURFACE AIR TEMPERATURE, PRECIPITATION, SEA ICE CONCENTRATION, AND SEA SURFACE TEMPERATURE DATA

The study is based on yearly averaged data from CMIP6 ensemble models. Two quantities were investigated: surface air temperature (SAT) and precipitation (PRECIP). Sea ice concentration (SIC) and sea surface temperature (SST) were used to contextualize and better understand changes in the SAT variance. The ensembles used have been chosen to satisfy three conditions: (i) more than 20 members in the same configuration (meaning the same physics and forcings), (ii) available from 1850 to 2100, and (iii) available for the three selected forcing scenarios. The scenarios investigated are "Sustainability" (SSP1-2.6) with a low radiative forcing of 2.6 W m^{-2}, "Middle of the Road" (SSP2-4.5) with a radiative forcing of 4.5 W m^{-2}, and "Fossil-fueled Development" (SSP5-8.5) with a high radiative forcing of 8.5 W m^{-2}. The radiative forcings mentioned above are the nominal values reached in 2100. Four models (Table 1) matched these conditions: ACCESS-ESM1-5 (Ziehn et al. 2020, CSIRO), CanESM5 (Swart et al. 2019, CCCma), MIROC6 (Tatebe et al. 2019, MIROC), and MPI-ESM1-2-LR (Olonscheck et al. 2023, MPI-M). They can be considered as quite independent since they are based on distinct atmosphere and ocean models. They also come from separate model families according to the evaluation performed by Brunner et al. (2020). This independence is important because it is the very motivation of our multiensemble strategy, assuming, to some extent, that model errors partially vanish if the models are independent. In practice, the models always share at least some assumptions and are therefore never completely independent. However, we believe that we can consider results derived from a set of models with more confidence than those derived from a single model, particularly when we are interested, as here, in the dispersion of each model around its own ensemble mean and not in the differences between the different models.

For all models, we decided to use a single configuration for physics and forcing per model to preserve the consistency among members concerning the physical processes simulated and their parameterization. We did not want to consider multiple configurations of the same model because we do not know exactly how much the configurations will differ from each other, and if they are too close, it would artificially increase the importance of the model in the multimodel average. For each model, we selected the most common configuration, which is generally the first configuration (referred to as p1 f1). The only exception is the CanESM5 model from CCCma, where we used the second physical configuration because the first configuration showed "cold spots in Antarctica,

which arise from a misspecified land fraction in p1 and were resolved in p2" according to Swart et al. (2019, p. 4860).

b. Methods

In this section, the methods are presented using SAT in the mathematical formulas but are also applied to PRECIP, SIC, and SST data throughout the work.

1) Ensemble variance as a proxy of internal variability

The variance computation in the present work is based on ensemble anomalies, which means that at each time step t, the ensemble mean $\hat{T}(t)$ is subtracted from each member j of the ensemble. Each model/ensemble is processed separately. The ensemble anomaly of a member (T'_j) corresponds to

$$T'_j(t) = T_j(t) - \hat{T}(t) \text{ or equivalently } T' = T - \hat{T}. \quad (1)$$

The use of ensemble variance and ensemble anomalies allows us to focus on the evolution of internal variability without being impacted by model differences/bias in absolute temperature or precipitation (Chen et al. 2021, IPCC, AR6 Chapter 1 section 1.4.1).

The ensemble variance is also computed at each time step. Two ensemble variances are computed: the ensemble variance on yearly data at each location [Eq. (2)] and on globally averaged data [Eq. (3)]. Taking SAT anomalies as an example, ensemble variances are computed for each model or ensemble of size N as

$$\mathrm{EnVar}_{\mathrm{SAT}}(t, x, y) = \frac{1}{N-1}\sum_{j=1}^{N} T'^2_j(t, x, y) = \widehat{T'^2}, \quad (2)$$

$$\mathrm{EnVar}_{\mathrm{GSAT}}(t) = \frac{1}{N-1}\sum_{j=1}^{N} \langle T'_j\rangle^2(t) = \widehat{\langle T'\rangle^2}, \quad (3)$$

where $\langle . \rangle$ is the spatial average, such as $\langle T'_j\rangle = (1/S)\int\int_S (T_j - \hat{T})\,ds$, where ds is the surface area differential and S is Earth's surface. In the calculation of ensemble variance, $\widehat{\cdot}$ corresponds to $[1/(N-1)]\sum^N$ to take into account the degrees of freedom.

The computation of the multiensemble variance is detailed in the following section.

2) Construction of the multiensemble dataset

We based our work on a multiensemble approach for two reasons. First, as explained in the material section, the relative independence of the models is expected to partially remove model errors via averaging. Another way to think about it is that combining the internal variability of the different models makes it possible to study their common signal in terms of internal variability. Second, calculating the variance on a larger ensemble constructed from members of different models helps improve the estimation of internal variability by potentially encountering more possible states.

The multiensemble dataset was defined according to the "model democracy," which means that each model has the same weight in the dataset. The weighting of models is an active research question, and no general methodology is yet accepted (Knutti 2010). In addition, all the selected models have a large number of members (between 25 and 50) allowing us to assume a relatively good resolution of the internal variability in the ensembles given their relatively coarse spatial resolutions (around or greater than 1°). According to a recent paper, ensembles of 25 members could be sufficient to determine changes in temperature internal variability with about 10% error or less and with 30% when studying the internal response to forcing (Milinski et al. 2020). So here we decided to assign the same weight to all models. In addition, we also desired to use all the members available for the multiensemble computation. We therefore used two methodologies that return similar results but present different advantages:

- Averaging the individual results from the four models to construct an ensemble acknowledging all members from the four models (145 members). This is done by applying a weight on each individual member that corresponds to the inverse of the size of the ensemble it comes from (for example, MIROC6 has 50 members, and each member from MIROC6 receives a weight equal to 1/50).
- Use a resampling methodology where we select randomly 20 members for each model and compute the ensemble variance on the 80 (i.e., 20 × 4) members. Here, since we use ensemble anomalies (as explained in the section below), it is possible to compute the variance from a set of members coming from different models without being impacted by intermodel variance linked to their absolute climate state or their forced response. Repeating this many times (e.g., 100 times) allows us to assess the uncertainty.

While the first method has the advantage of being computationally efficient, the second one provides robust confidence intervals for the results. We therefore used the first method for the EOF decomposition discussed below and the other one for all the tasks that were less computationally expensive.

To compute the multiensemble variance maps, we first regridded the model outputs on a 1° × 1° grid with conservative algorithms (Jones 1999) using the xESMF python package (Zhuang et al. 2023).

3) Separation of time scales

The time scales have been separated using a rolling average on the ensemble anomalies time series. For a τ-yr filtering, one obtains

$$\overline{T'}^{\tau}(t) = \frac{1}{\tau}\int_{t-\tau/2}^{t+\tau/2} T'(t')dt', \quad (4)$$

where τ is the time period of the filter, $\overline{T'}^{\tau}$ is the filtered data, and dt' is the time unit. The τ-yr ensemble variance is thus computed by applying the EnVar, [Eq. (2)] or [Eq. (3)], on the τ-yr filtered temperature anomalies $\overline{T'}^{\tau}$. The τ-yr ensemble variance contains also the lower frequencies. We therefore define frequency bands (1–3, 3–5, 5–11, and >11 years) computed as the difference between two filterings. For instance, the 1–3 years is the difference between the 1-yr and the 3-yr

ensemble variance time series. Each τ-yr ensemble variance time series is finally filtered to have a final 11-yr resolution. Here, one must clearly distinguish the filtering on temperature anomalies before computing the variance that isolates a part of the variance "energy" and the filtering on variance time series that simply smooths the variance curves and homogenizes the resolution of variance time series to allow comparison between frequency bands.

4) Decomposition in spatial modes

The spatial patterns that control the (co)variance also known as empirical orthogonal functions (EOFs) and their temporal evolution called principal components (PCs) have been computed using a singular value decomposition (SVD). The entire ensemble anomaly dataset is constructed by concatenating all years, models, members, scenarios, and locations. Then, it has been area weighted (square root of the cosine of latitudes, to preserve variance), ensemble-size weighted (weight = inverse of ensemble size), and decomposed to obtain the first 1000 modes. Hence, here, the EOFs represent "typical" EOFs independent of time, ensemble members, and models. (Note that it is possible to define EOFs that are dependent on either time, ensemble members, or models, but this defeats the purpose of our computation that wishes to define a common basis.)

In the matrix form, the ensemble anomalies of SAT can be decomposed using the SVD into three matrices ($\mathbf{U}$, $\mathbf{\Sigma}$, and $\mathbf{V}^{\mathrm{T}}$, where superscript T represents the transpose operator) and reorganized into a sum of modes i with PCs (mode, time, and member) and EOFs (mode, lon, and lat):

$$T' = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}} = \sum_i \mathrm{PC}_i \mathrm{EOF}_i. \tag{5}$$

Globally averaged temperature anomalies can be obtained by spatially averaging the EOFs, enabling the contribution of each mode to the global internal variability signal to be studied:

$$\langle T' \rangle = \sum_i \mathrm{PC}_i \langle \mathrm{EOF}_i \rangle. \tag{6}$$

In this case, we are mainly interested in the signal associated with the first mode, which is significantly dominant (as explained below). The contribution of this first mode ($i = 1$) to the globally averaged anomalies is computed as

$$\langle T' \rangle_1 = \mathrm{PC}_1 \langle \mathrm{EOF}_1 \rangle. \tag{7}$$

And the ensemble variance of globally averaged surface air temperature explained by the first mode (only) is obtained by computing the squared anomalies and averaging over members ($\widehat{\cdot}$):

$$\widehat{\langle T' \rangle_1^2} = \widehat{\mathrm{PC}_1^2} \langle \mathrm{EOF}_1 \rangle^2 \tag{8}$$

Here, we have designed PC as a 2D matrix with a first dimension involving modes and a second dimension involving time and members. When we average over the members, we therefore only average over part of the second dimension, creating cross terms between modes, which are not investigated.

The decomposition is presented above for SAT data but can be applied to any ensemble model output. In this study, the same methodology is used to decompose the precipitation flux dataset in PCs and EOFs.

3. Results

We start the investigation with a broad scope using globally averaged data and progressively refine the analysis to move to local scope and finally to specific patterns.

a. *Opposite evolution of interannual and decadal variability at a global scale*

The ensemble variances computed for globally averaged temperature (Fig. 1) and precipitation (Fig. 2) strongly differ among models. For example, MIROC6, the ensemble with the highest resolution, has more than twice the variance of other models. However, these differences do not prevent the models to show similar a trend or behavior regarding the time evolution of this ensemble variance.

Looking specifically at the total variance, we see an increase over time for precipitation which seems also correlated to the intensity of the radiative forcing whereas no such trends are observed for temperature. For precipitation, the total ensemble variance reaches its maximum for all models in the 2065–95 segment under the SSP5-8.5 forcing scenario. For temperature, the discrepancies between models do not allow to identify the future evolution of the total ensemble variance. One model (CanESM5) shows an overall decrease over time and scenario, two models (MPI-ESM1-2-LR and ACCESS-ESM1-5) mainly exhibit stagnation, and the last one (MIROC6) shows an increase.

Despite this difference regarding the evolution of the total variance of precipitation and temperature, the two quantities share a strong common signal concerning the spectral change obtained with the filtering approach [see section 2b(3)]. Indeed, both show an intensification of interannual variability with a period between 1 and 3 years, in both absolute and relative contributions. At the same time, a weakening of the decadal variability of temperature with a period larger than 11 years is observed. This explains why despite the increase in interannual variability, temperature does not present a net increase in the total variance as observed in precipitation. The robust intermodel agreement of this shift allows us to analyze the multiensemble variance to describe and summarize this phenomenon.

Regarding the precipitation, the initial variances of each frequency band at the early historical period (1870–1900) are 6.1 $\mathrm{mm^2\ yr^{-2}}$ (52% of the total variance) for the 1–3-yr period band (i.e., interannual), 1.8 $\mathrm{mm^2\ yr^{-2}}$ (16%) for 3–5 years, 1.5 $\mathrm{mm^2\ yr^{-2}}$ (13%) for 5–11 years, and 2.3 $\mathrm{mm^2\ yr^{-2}}$ (20%) for periods longer than 11 years (i.e., decadal). The absolute (relative) contribution of the interannual variability increases to 8.4 $\mathrm{mm^2\ yr^{-2}}$ (55%) for the 1980–2010 time segment and, for the end of the twenty-first century, increases, respectively, to 8.8 $\mathrm{mm^2\ yr^{-2}}$ (58%), 9.8 $\mathrm{mm^2\ yr^{-2}}$(62%), and 12.3 $\mathrm{mm^2\ yr^{-2}}$

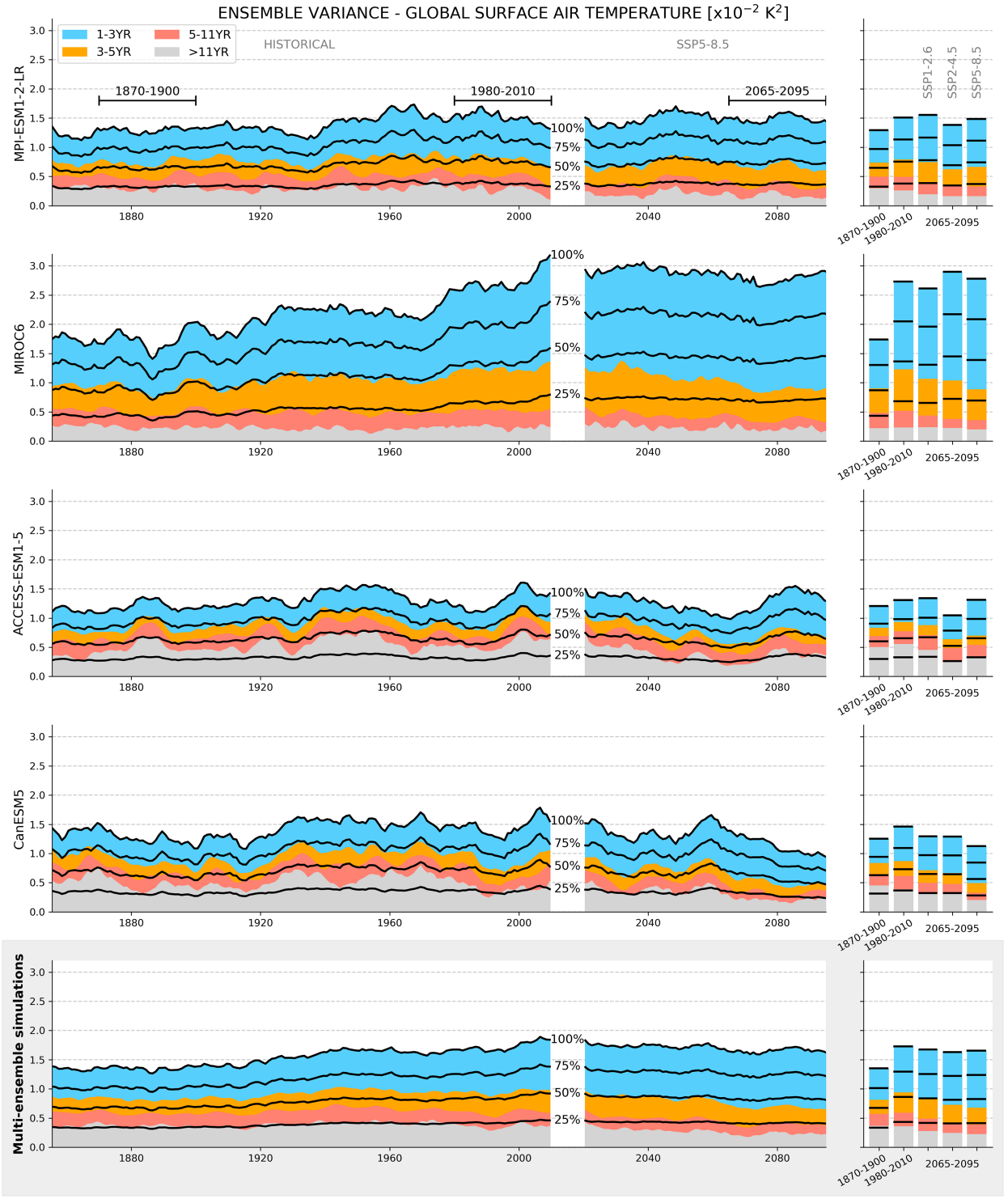

FIG. 1. Ensemble variance of globally averaged SAT. The four first rows show the ensemble variance for four different models: MPI-ESM1-2-LR, MIROC6, ACCESS-ESM1-5, and CanESM5. The last row presents the variance computed from a multiensemble constructed from the concatenation of the four previous models. (left) The yearly results for the historical and projected period with the Fossil-fueled Development scenario (SSP5-8.5). The colors represent the contribution of the various frequency bands to the total variance: blue for the variability with a period band of 1–3 years, orange for 3–5 years, red for 5–11 years, and gray for periods longer than 11 years. The bands are obtained from a temporal filtering of temperature at the given period, and all the bands of variance are finally filtered to be consistent with the 11-yr period. (right) The results for a given climatic period (averaged over 30 years): early historical (1870–1900), end of historical (1980–2010), and end of projections (2065–95) for three shared socioeconomic pathways: Sustainability (SSP1-2.6), Middle of the Road (SSP2-4.5), and Fossil-fueled Development (SSP5-8.5). The black lines represent the quarters of the total variance.

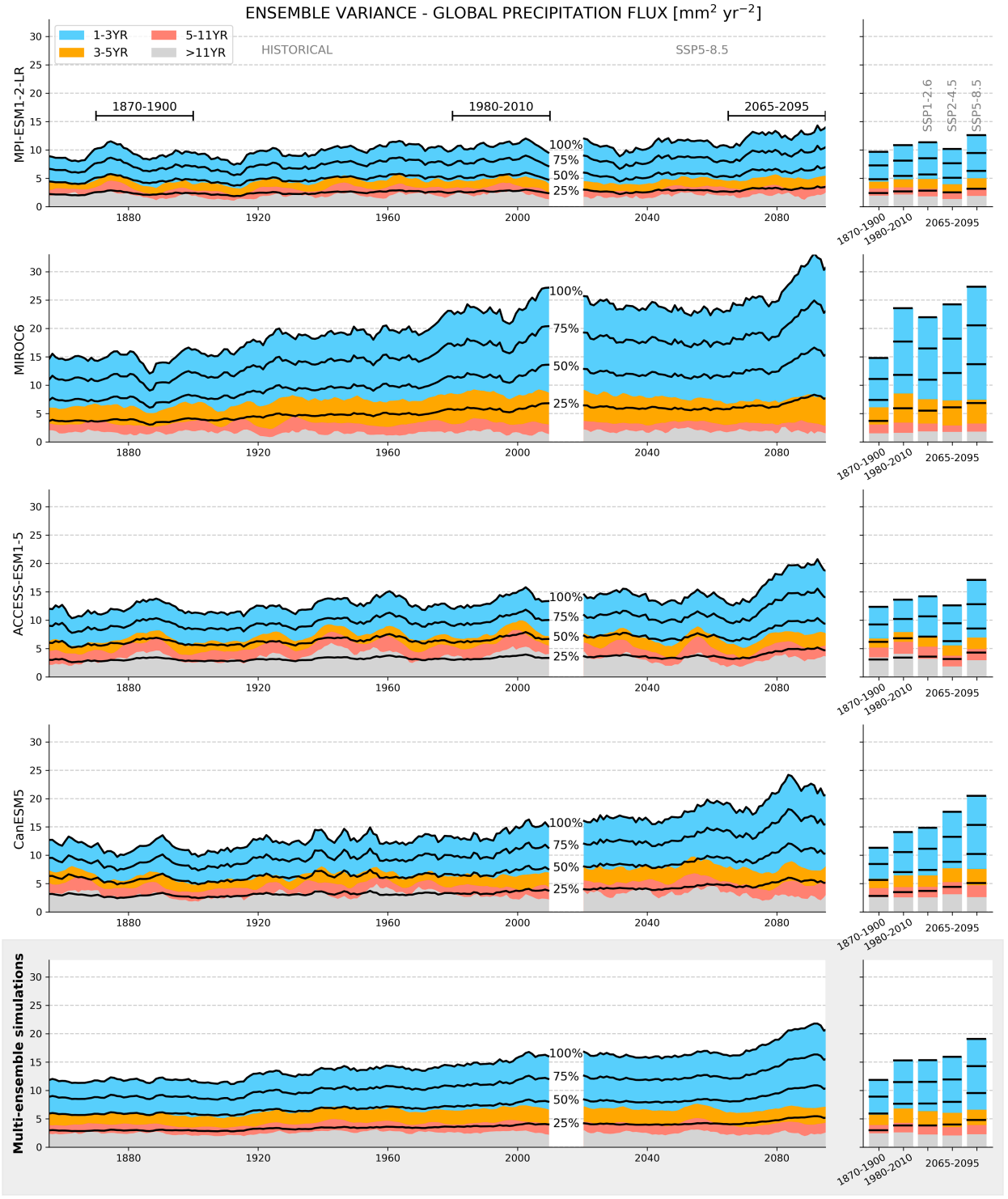

FIG. 2. Ensemble variance of globally averaged precipitation. As in Fig. 1, but for precipitation.

(65%) for SSP1-2.6, SSP2-4.5, and SSP5-8.5. The absolute variance of the other bands does not change much over time or scenario, except the 3–5-yr period band that increases from 1870–1900 to 1980–2010 and then stabilizes.

Concerning the SAT, the initial variances are 5.3×10^{-3} K^2 (40% of the total variance) for the interannual component, 2.4×10^{-3} K^2 (18%) for a period band of 3–5 years, 2.1×10^{-3} K^2 (15%) for 5–11 years, and 3.6×10^{-3} K^2 (27%)

for decadal variability. The absolute (relative) contribution of the interannual variability increases to 7.7×10^{-3} K^2 (45%) for the 1980–2010 time segment. For the end of the current century, it increases to 8.3×10^{-3} K^2 (50%), 8.9×10^{-3} K^2 (55%), and 9.6×10^{-3} K^2 (59%) for SSP1-2.6, SSP2-4.5, and SSP5-8.5, respectively. After an increase of their variance from 1870–1900 to 1980–2010, frequency bands above 1–3 years show significant declines. The maximum decrease is observed for the decadal variability that diminishes to 2.2×10^{-3} K^2 for SSP5-8.5 in 2065–95, which represents a 40% loss relative to the early historical period.

These frequency changes in temperature and precipitation variance are very intense. Indeed, the ensemble variance from the 1–3-yr band almost doubles from 1870–1900 to 2065–95 for the SSP5-8.5 scenario, when at the same time, temperature variance in the other bands decreases significantly, up to 40%. The interannual time scale dominates the internal variability of temperature and precipitation, in both the past and the present, and even more so in the future. In contrast, the contribution of low frequencies to the total variability of these two quantities at a global scale appears to be of second-order importance and is projected to decrease even further over the twenty-first century.

To better understand this phenomenon, we refined the analysis on the five identified 30-yr segments (1870–1900, 1980–2010, and 2065–95 for the three scenarios) with spatial maps of local ensemble variance (Figs. 3 and 4). The objective is to determine, among other things, if the signal is spatially homogeneous or located in specific regions. We also seek to determine whether the opposite changes in temperature variance are driven by the same mechanism or not. For that, we use the multiensemble dataset built from the 100-times resampling method [see section 2b(2)] to investigate a robust signal. The area where the signal-to-noise ratio of the change in ensemble variance is less than one (i.e., where the 95% confidence interval associated with the change in variance is greater than the mean magnitude of that change) indicates low significance and covers most of the area where the changes are weak.

The maps of temperature variance present, first of all, a strong initial (1870–1900) ensemble variance located over the ocean at high latitudes (Fig. 3, first column). This signal shows significant intensity at all frequencies and reaches its peak for the interannual variability, in the 1–3-yr band. On the other hand, an important ensemble variance signal is observed at lower latitudes on land and in the equatorial central Pacific, for the interannual variability. The changes computed for the four following segments (Fig. 3, second to fifth column) appear as a progressive evolution of two regimes. The first one is associated with a decrease in the ensemble variance located in the ocean at high latitudes and for all time scales, in agreement with the initial variance signal. The second one is associated with an increase solely visible at a band frequency of 1–3 years and located at lower latitudes. Only this second regime appears in the precipitation maps (Fig. 4). Its signature is particularly visible around the equatorial band of the Pacific. According to the assessment of significance, represented by the gray hatches, these two patterns of change clearly stand out as being robust at the 95% confidence level. In the next two subsections, we will discuss these two changes more specifically.

b. Decline of internal variability at high latitudes

The decrease in temperature variance located at high latitudes and acting on all time scales has an amplitude varying with the intensity of the scenarios (Fig. 3). In the two less emissive scenarios (SSP1-2.6 and SSP2-4.5), this decrease seems also linked, in the Northern Hemisphere, to a local increase of the variance further poleward of the decrease patch. Given the location of the decrease and the associated increase poleward, we have tested if this decrease is associated with sea ice–atmosphere interactions.

To investigate this hypothesis of sea ice impact on SAT variance, we computed, following the same methodologies, the multiensemble variance of the yearly SIC and its evolution (Fig. 5). The SAT variance changes precisely match the SIC variance changes. In the Northern Hemisphere, we observe a decrease in the ensemble variances in the Barents Sea in 1980–2010 that intensifies and extends progressively in the future scenarios to the Greenland Sea and to the Irminger and Labrador Seas for the higher scenarios. In the Southern Hemisphere, the decrease is a bit smaller but still exists. It is mainly located at both sides of the Antarctic Peninsula and propagates and intensifies all around Antarctica proportionally to the radiative forcing. At the same time, we observe an increase in the ensemble variance of temperature at higher latitudes that corresponds to an increase in SIC variance change in the Northern Hemisphere but that cannot be matched to SIC variance changes in the south because of the presence of the land.

We explain this almost perfect match between the variance signals with the impact of sea ice on surface air temperature as explained by the schematic diagram in Fig. 6. In the presence of sea ice ("ice-covered" area), the atmosphere is isolated from the warmer ocean and its temperature can decrease to very low levels. Conversely, in the absence of sea ice ("ice-free" area), the ocean warms the atmosphere and prevents its temperature from decreasing to very low levels. The area of variance changes is therefore simply linked with the presence/absence of sea ice among members. The region where some members present sea ice and some others do not (referred to as "sea ice edge") will have very high SAT ensemble variance. In contrast, the regions that are covered by sea ice either in all members or in none of them have much lower SAT ensemble variance. The signature showing a decrease in variance is therefore the footprint of the intermember sea ice edge during the early historical period changing from being ice-covered in some members and ice-free in others to being ice-free in all or most of the members. Following the same line of reasoning, the signature showing an increase in variance is the footprint of the intermember sea ice edge in a given scenario changing from being ice-covered in all or most of the members to being ice-covered in some members and ice-free in others. The combined increasing and decreasing patches indicate the displacement of the sea ice edge. The fact that SSP5-8.5 presents almost no increase in ensemble variance reflects the almost

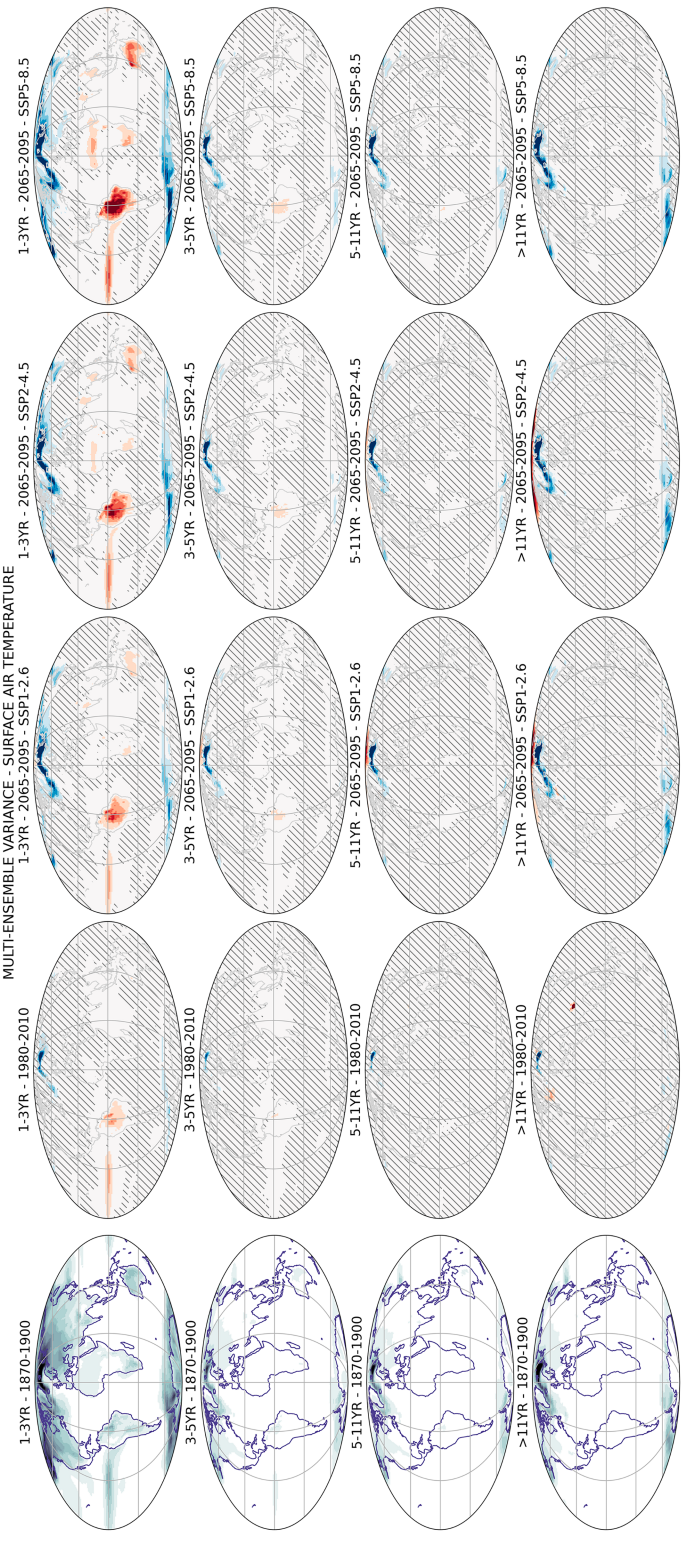

FIG. 3. Multiensemble variance of local SAT. Each column corresponds to a given 30-yr segment from the collection of four models assembled together. The first column is the ensemble variance for the early historical reference taken between 1870 and 1900. The four following columns show the deviation from this reference. The second column is the end of the historical period averaged between 1980 and 2010. The three last columns present projections of the same time period at the end of the projection runs (2065–95) for three shared socioeconomic pathways: Sustainability (SSP1-2.6), Middle of the Road (SSP2-4.5), and Fossil-fueled Development (SSP5-8.5). Each row corresponds to a given frequency band of the variability extracted from the time filtering of the variance signal: (row 1) 1–3 years, (row 2) 3–5 years, (row 3) 5–11 years, and (row 4) variability with period greater than 11 years. The gray hatched area corresponds to the area where the change in variance is not significant, i.e., where the signal-to-noise ratio of the variance is less than 1, where the 95% confidence interval associated with the change in variance is greater than the magnitude of that change.

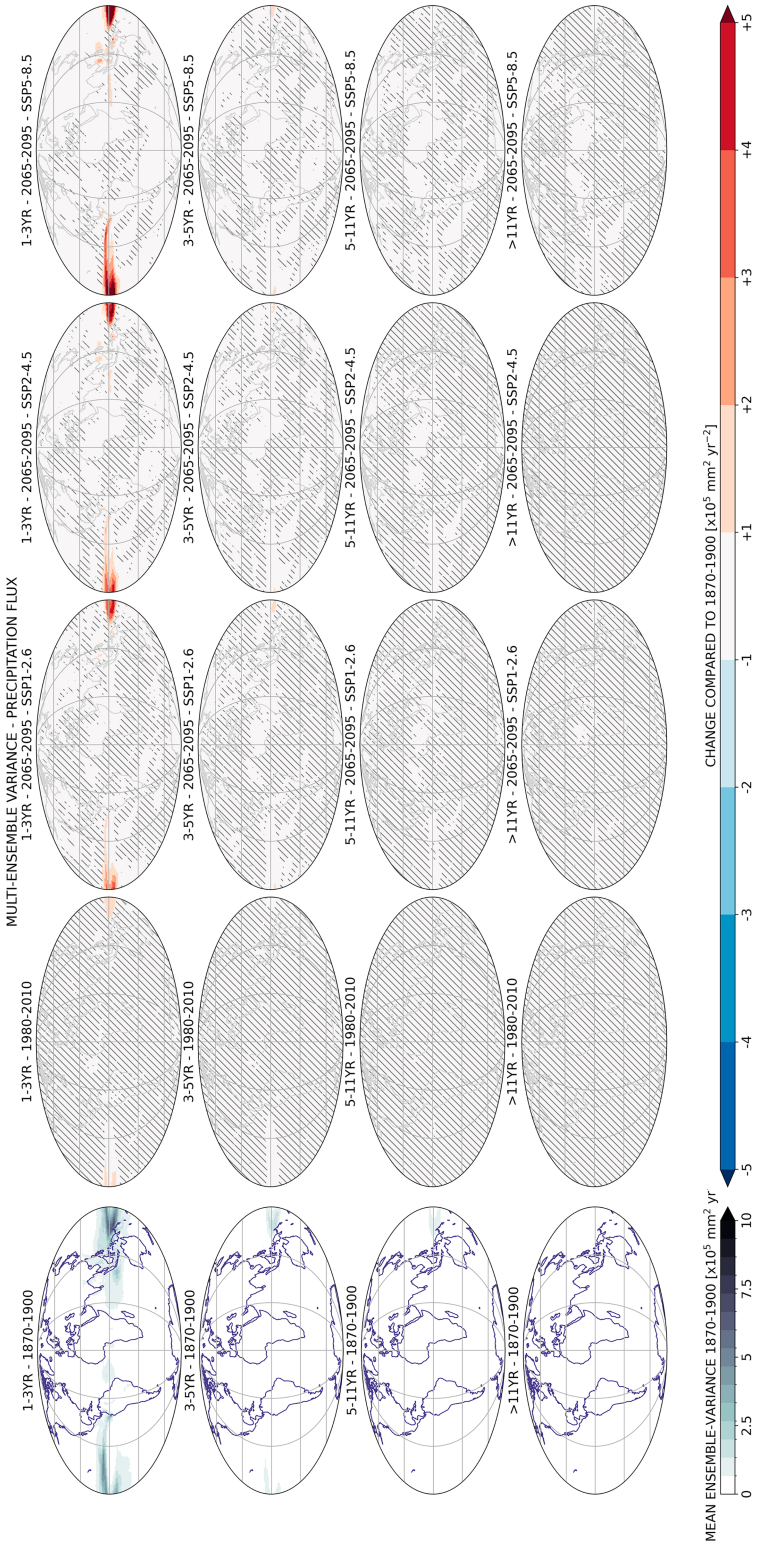

FIG. 4. Multiensemble variance of precipitation flux (see detailed explanations in Fig. 3).

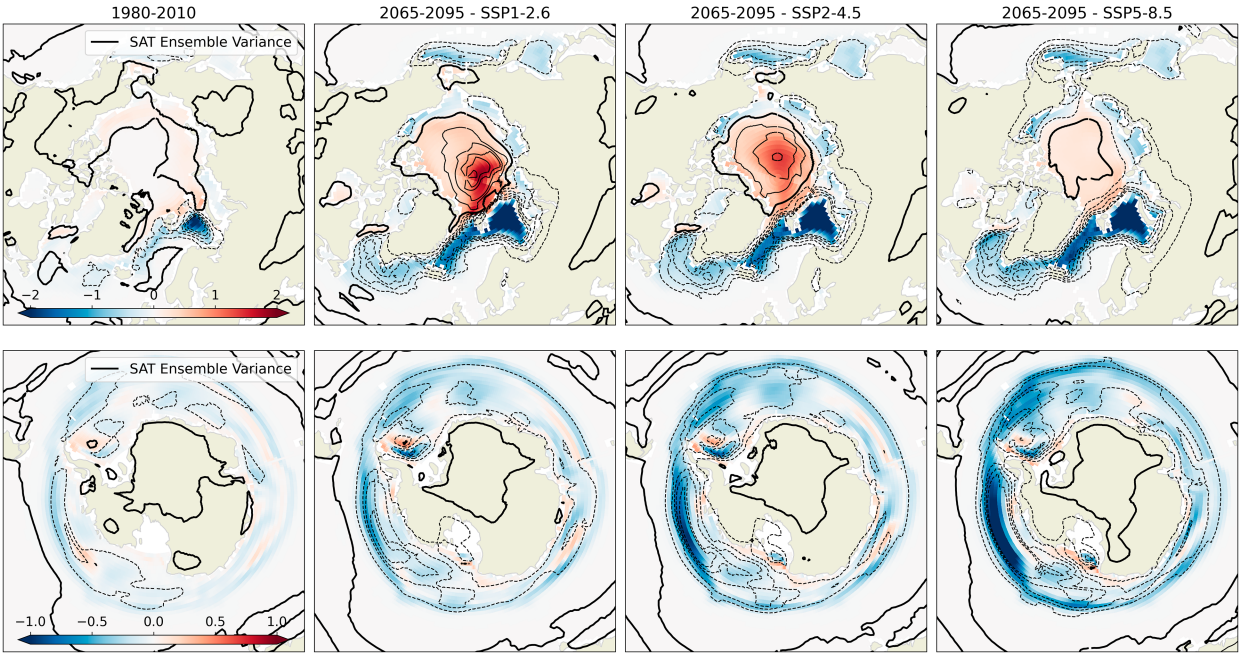

FIG. 5. Comparison between SIC and SAT ensemble variance changes. The colors represent the changes in SIC ensemble variance (%) compared to the early historical reference (1870–1900). Areas in red represent local increases of ensemble variance, and the blue areas represent local decreases. Contours represent the changes in SAT ensemble variance (K^2) compared to the early historical reference. The thick lines represent the zero contour, and the thin solid and dashed lines represent, respectively, the positive (increase) and negative (decrease) contours. (first column) The results for the end of the historical period (1980–2010). The three following columns present the results for the end of the twenty-first century for the three scenarios studied: (second column) SSP1-2.6, (third column) SSP2-4.5, and (fourth column) SSP5-8.5. (top) The results around the North Pole. The contours are plotted from -2 to 2 K^2 every 0.4 K^2. (bottom) The variances are smaller around the South Pole; therefore, the colormap and contour values are reduced by half.

disappearance of sea ice in yearly averaged sea ice concentration, linked to global warming. Indeed, in SSP5-8.5, the multiensemble mean sea ice concentration does not have any cell in the Northern Hemisphere with more than 50% of sea ice after 2090 and no cell with more than 75% of sea ice after 2062. The system described here therefore involves three components with mutual impact on their variances: the ocean, the atmosphere, and the sea ice. In this relationship, sea ice is the main driver of variance, enabling or preventing heat exchange between the ocean and the atmosphere.

To test the validity of the proposed system, we extended the investigations to sea surface temperature. In Fig. 7, the evolution of the ensemble variance of SST with respect to the early historical period is compared to the ensemble variance of sea ice. As for the SAT case, the SST variance changes match the changes in the sea ice variance. This results from the fact that under sea ice, SST is locked near the freezing point, so when sea ice concentration changes, temperature changes. Their variances are directly related. Although obvious, this result reinforces our confidence in this three-component (ocean, sea ice, and atmosphere) system.

c. *Reinforcement of interannual variability at low latitudes*

As discussed above, the second regime observed on local ensemble variance maps is characterized by an increase in variance at low latitudes (Figs. 3 and 4). For temperature, the changes are located both above the ocean, in the eastern and central equatorial Pacific region, and above the land, over the Amazon rain forest or in the tropical regions in Africa, Asia, and Australia. The evolution appears proportional to the radiative forcing with a maximum reached for SSP5-8.5 at the end of the twenty-first century. For this time segment, the maximum change is located in the Amazonian region where the initial (1870–1900) variance is 0.34 K^2 and increases to 1.59 K^2. In terms of standard deviation, it corresponds to an increase by 0.67 K of the interannual variability, which is more than a doubling.

Local ensemble variances for precipitation are consistent with the results from globally averaged data and on the local variance of temperature (Fig. 4). Indeed, we find the low-latitude signature associated with the increasing variance at the interannual scale. The increase seems again proportional to the radiative forcing and closely looks like a typical ENSO signature. For the end of the twenty-first century time segment, the maximum change is located in the western Pacific where the initial variance was 4×10^5 mm^2 yr^{-2} and increased to 13×10^5 mm^2 yr^{-2}. In terms of standard deviation, it corresponds to an increase by 520 mm yr^{-1} of the interannual variability (i.e., close to a doubling). The absence of signal at lower frequency is also an argument to associate this precipitation variance signal to the signal of SAT variance and consider they are driven by the same mechanism.

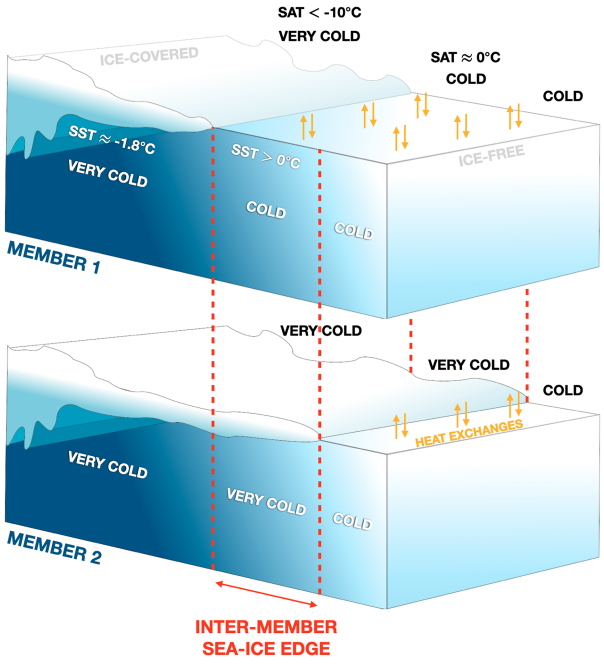

FIG. 6. Schematic link between ensemble variance of sea ice and ensemble variance of surface air and ocean temperature. The inter-member sea ice edge is the area where some members are covered by sea ice and others are not. Sea ice controls and enables heat exchange between the ocean and the atmosphere. Consequently, in these transition zones, variance will be important not only for SIC but also for atmospheric and ocean surface temperature.

To determine whether this pattern of increasing variance is driven by a specific mechanism or is the result of the superposition of different mechanisms, we decomposed the SAT and precipitation into a collection of EOFs (referred to as spatial modes) and their associated PCs (referred to as time series of the modes), using a singular value decomposition [see section 2b(4)]. This is done by using the entire dataset (i.e., all time periods, scenarios, models, and ensembles). It allows to build a typical variability pattern valid across and representative of all models, scenarios, and time periods.

For both SAT and precipitation, the leading modes obtained (Fig. 8) are largely dominant and explain a significant part of the variance. The first mode (EOF1) of SAT alone explains 15% of the variance of the entire dataset, while the following mode explains only 6%. Regarding the precipitation, EOF1 explains 24% of the total variance, whereas EOF2 is representative of only 5%. These EOFs are particularly intensified at low latitudes. Within the 20°S–20°N region, the first mode of SAT and of precipitation explains 46% and 28% of the variance, respectively. Focusing on the Pacific sector only (20°S–20°N and 130°–290°E), they are even more important. Indeed, there, EOF1 of SAT and of precipitation explains 55% and 35% of the variance, respectively.

These leading EOFs for temperature and precipitation with their strong signature in the equatorial Pacific and teleconnections transporting their influence around the world mainly in the tropical region are very close to the typical ENSO pattern that dominates the variability of climate at a global scale (Cassou et al. 2021).

To determine whether these ENSO-like patterns are responsible for the increase in ensemble variance frequency, we computed the ensemble variance of globally averaged data explained by the first mode alone (Figs. 9 and 10 for SAT and precipitation, respectively). We used for each quantity its PC1 and its globally averaged EOF1. In this work, we do not investigate the impact of cross terms involving the first mode and the other ones.

For both SAT and precipitation, the variability is dominated mainly by the 1–3-yr band and by the 3–5-yr band to a lesser extent. This is typical of ENSO, generally assumed to have a period between 2 and 7 years (Fredriksen et al. 2020). For SAT (Fig. 9), the relative importance of the 1–3-yr band is very close between models, starting around 50% at the beginning of the historical period and reaching 65%–75% at the end of the SSP5-8.5 scenario. Regarding precipitation (Fig. 10), the differences are much larger among models. MPI shows a weak increase in interannual variability, starting around 60% and finishing at 63%. MIROC6 presents the largest increase from 64% to 77%. Interannual variability also increases in ACCESS but is limited by its already high level at the start of the simulation. The relative importance of its 1–3-yr band increases from 73% to 80%. Finally, CanESM5 presents also a significant increase from 68% to 76%. The differences regarding the increase of interannual variability seem linked to the concomitant increase of the 3–5-year frequency band that can dilute a part of the 1–3-yr band increase in relative importance. In general, the total ensemble variance increases with time and radiative forcing, in line with recent literature based on SST (Cai et al. 2022; Maher et al. 2023) and as indicated in the IPCC AR6 report (Lee et al. 2021, Chapter 4, section 4.3.3.2) for precipitation. According to our results, this is achieved through the absolute and relative increase of the interannual variability (i.e., 1–3-yr period band). For the precipitation and using the multiensemble, the absolute (relative) ensemble variance at the interannual scale is 0.27 $mm^2\ yr^{-2}$ (65%) during the early historical (1870–1900), increases to 0.37 $mm^2\ yr^{-2}$ (66%) between 1980 and 2100, and reaches at the end of the century 0.47 (69%), 0.52 (72%), and 0.59 $mm^2\ yr^{-2}$ (74%) for SSP1-2.6, SSP2-4.5, and SSP5-8.5, respectively. For the SAT, the interannual multiensemble variance is $2.5 \times 10^{-3}\ K^2$ (53%) during the early historical, increases to $3.7 \times 10^{-3}\ K^2$ (56%) between 1980 and 2100, and reaches at the end of the century 4.1 (60%), 4.5 (63%), and $4.6 \times 10^{-3}\ K^2$ (67%) for SSP1-2.6, SSP2-4.5, and SSP5-8.5, respectively. The interannual variance component almost doubles between the early historical and the end of the twenty-first century for SSP5-8.5.

MIROC6 presents a small decrease of ensemble variance at the end of the SSP5-8.5 scenario for both SAT and precipitation. This is due to a change in the low-latitude variability pattern, moving from canonical-type/eastern ENSO to Modoki-type/central ENSO (Shin et al. 2022). In our framework, this implies a transfer of the variability to another EOF. CanESM5 also presents such a decrease, but its spatial resolution is much lower and we were not able to recognize a clear Modoki pattern emerging. However, we clearly identify alternative patterns emerging in the equatorial Pacific

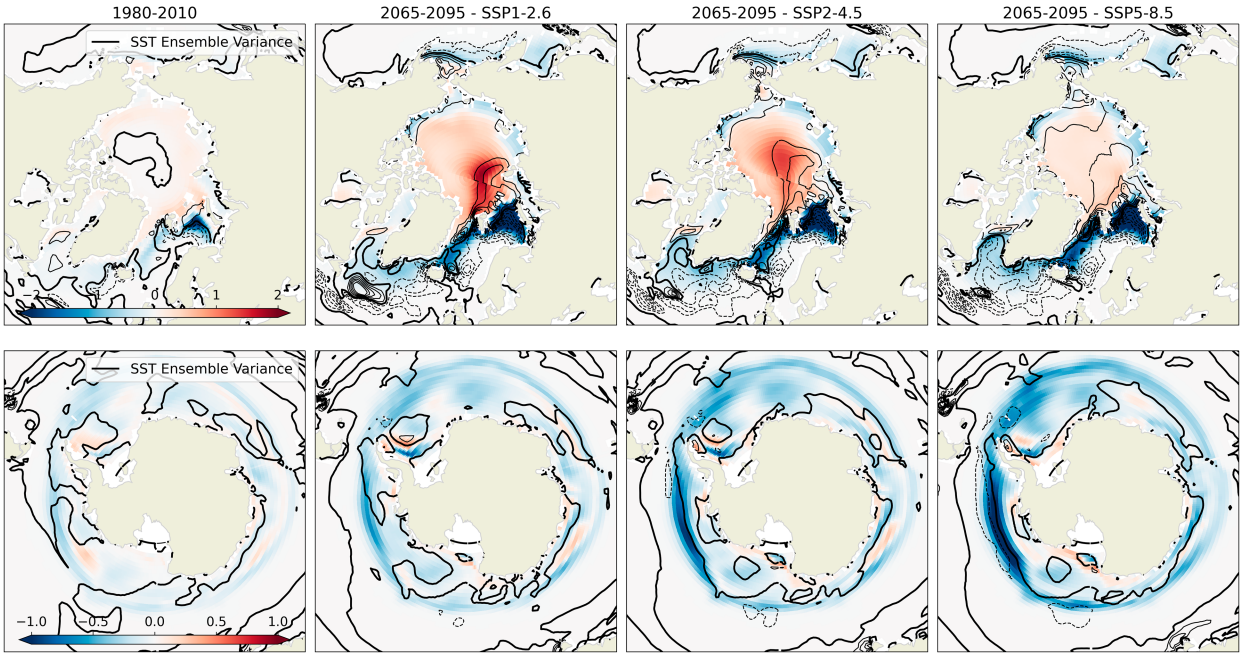

FIG. 7. Comparison between SIC and SST ensemble variance changes. As in Fig. 5, the colors represent the changes in SIC ensemble variance (%) compared to the early historical reference (1870–1900). Contours represent the changes in SST ensemble variance (K^2) compared to the early historical reference. The thick lines represent the zero contour, and the thin solid and dashed lines represent, respectively, the positive (increase) and negative (decrease) contours. (top) The results around the North Pole, and the contours are plotted from -1 to 1 K^2 every 0.2 K^2. (bottom) The variances are smaller around the South Pole; therefore, the colormap and contour values are reduced by half.

and therefore expect leakage of the variance of its first mode to these other EOFs.

4. Discussion

Using four ensemble models from CMIP6, we studied the evolution of ensemble variance as an indicator of internal variability. Our results suggest that this internal variability has changed since the beginning of the historical period and is likely to change in the future depending on the emission scenario. We paid particular attention to separating the different temporal scales of variability, by highlighting and associating frequencies with the simulated changes. This aspect has not been much studied in the literature, particularly using different realistic forcing scenarios. For instance, Rehfeld et al. (2020) studied variability frequencies,

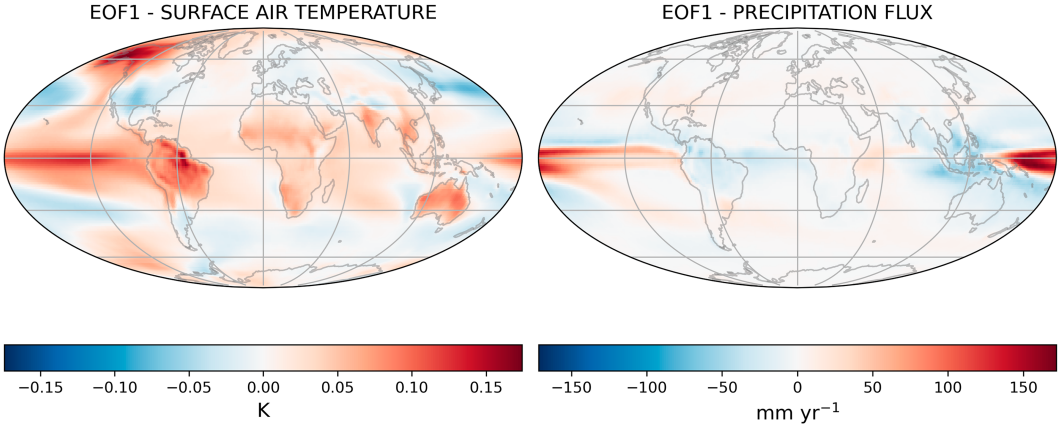

FIG. 8. First EOFs from a SVD of SAT and precipitation. The first EOF modes for (left) SAT and (right) precipitation multiplied by the standard deviation of the first principal component. The first mode of SAT explains 15% of the variance of the entire dataset, and the second mode explains only 6%. For precipitation, the first mode explains 24% of the variance and the second mode explains only 5%.

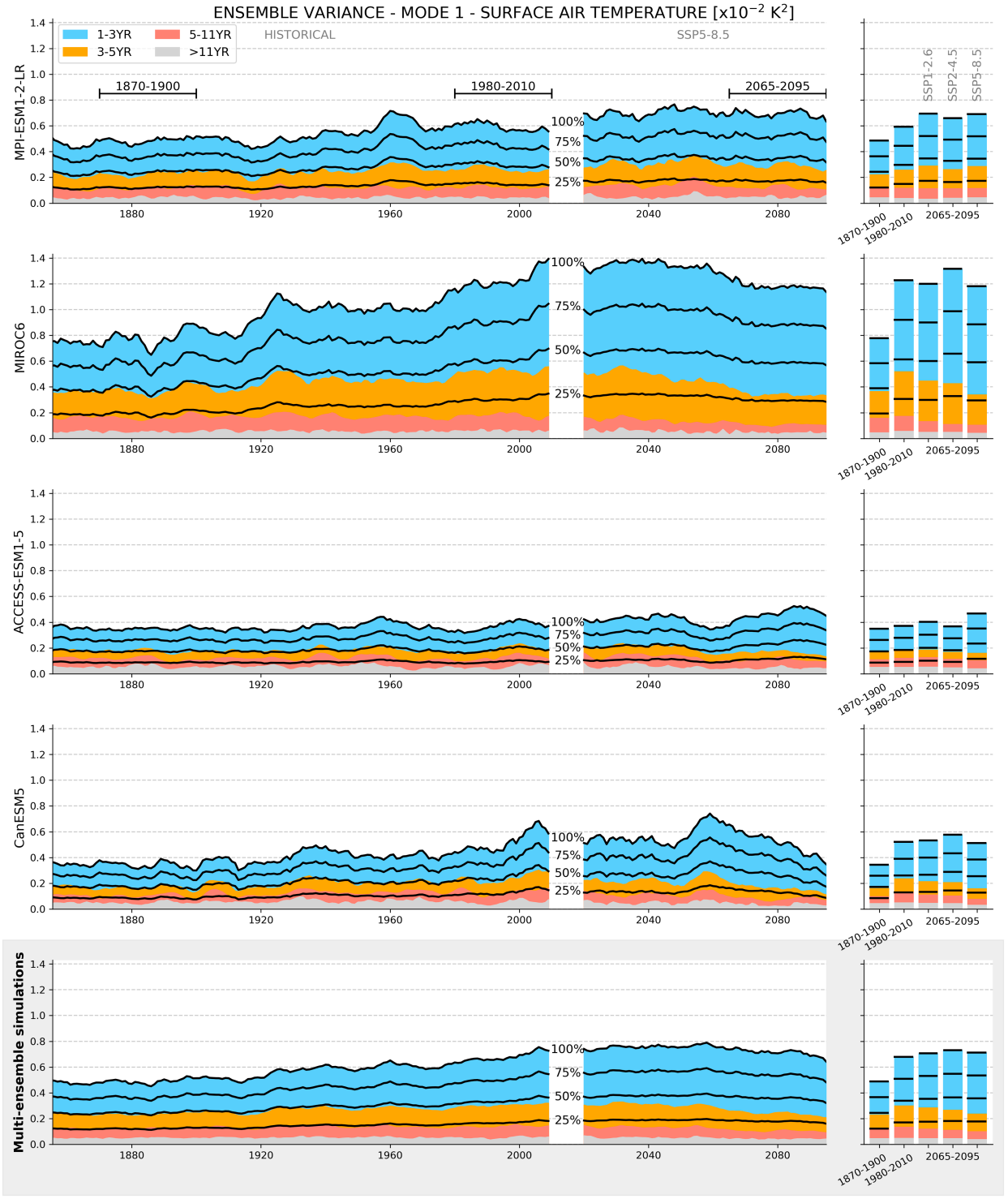

FIG. 9. Ensemble variance of globally averaged SAT explained by the first EOF mode. Each row corresponds to a given model, except the last one which is the concatenation of all of them. The rows present the ensemble variance over the historical and projection (SSP5-8.5) periods. As in Fig. 1, the colors represent the contribution of the various frequency bands of variability to the variance, and the right panels gather information for three climatic periods (1870–1900, 1980–2010, and 2065–95) and for the three investigated scenarios for the last period (2065–95). The black lines represent the quarters of the total variance.

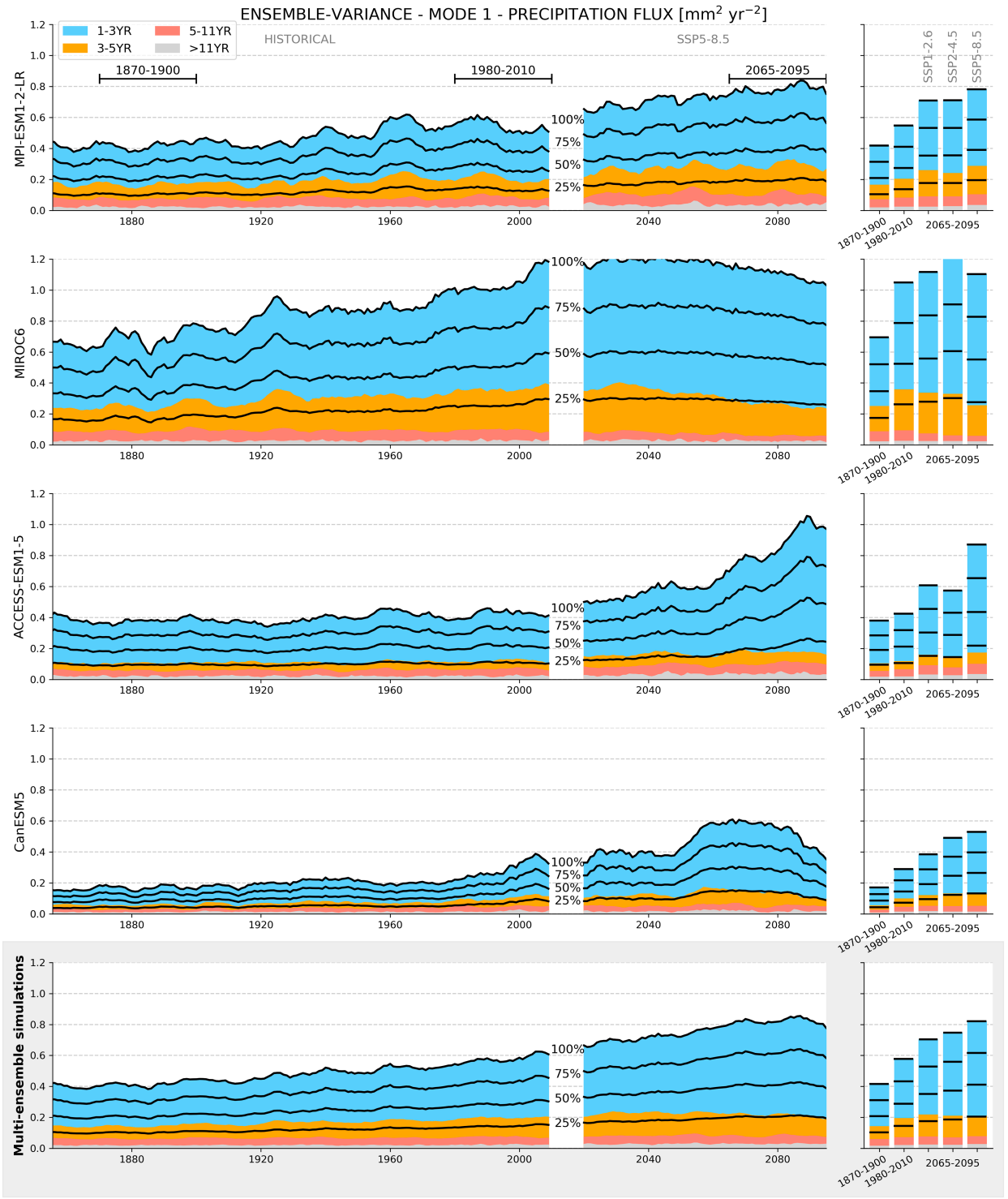

FIG. 10. Ensemble variance of globally averaged precipitation flux explained by the first EOF mode. As in Fig. 9, but for precipitation.

but only using idealized experiments (e.g., 1pctCO_2 or abrupt-4xCO_2). Our results highlighted two distinct patterns of changes consistent with previous works, using ensemble simulations (Olonscheck et al. 2021), or single-member datasets (Rehfeld et al. 2020; Shi et al. 2023). However, the two mechanisms driving the changes identified in the present study are to some extent different from those proposed in these previous studies.

The first one, inducing a net loss of internal variability at high latitudes, is linked to the progressive disappearance of sea ice, as also suggested by Olonscheck et al. (2021) or Shi et al. (2023). Here, we show that the changes are shared from interannual-to-decadal time scales and show that the dominant signal is related to the sea ice edge shifting poleward. In addition, we highlight the importance of the coupled ocean–atmosphere–sea ice system reinforced by the similar changes observed in sea surface temperature variance. Olonscheck et al. (2021) have proposed that the meridional temperature gradient may also play an important role in reducing temperature variability from high to midlatitudes. Although this mechanism may indeed play a role in the decrease in variance, we believe that this role is second order, since the decrease at midlatitudes is not visible compared to the one at high latitudes. Shi et al. (2023) suggest that snow-cover changes on land contribute to the decrease in temperature variability at high latitudes. This is a potential mechanism explaining the small decrease in variability we observe at high latitudes over land. However, we find that the signal is mainly oceanic, and therefore, the sea ice edge displacement is likely the dominant mechanism to explain the decreasing signal detected at a global scale.

The second mechanism acts at low latitudes where we suggest an increase in the internal variability frequency in the future. Although Olonscheck et al. (2021) also observed an increase in variability at low latitudes, they linked these observations to changes in vegetation cover, with a transition to drier surface types. In particular, they suggest a shift from tropical forest to semiarid landscapes. Another study based on single-member datasets suggests that changes at low latitudes are linked to a decrease in soil moisture and an increase in clouds and longwave radiations associated (Shi et al. 2023). In the present work, thanks to the study of precipitation in addition to temperature, an ENSO-like pattern emerges naturally from the singular value decomposition of the data. Moreover, the separation of frequencies enables us to investigate this aspect of changes concerning the entire low-latitude pattern. We suggest that the increasing frequency of internal variability observed at low latitudes may reflect an increase in the frequency of ENSO, with an interannual component becoming increasingly dominant reflecting stronger associated events. These observations can be linked to the work of Fredriksen et al. (2020) describing a shift of the Niño-3.4 index toward the higher frequency in its power spectrum. Here, we obtain the same results using a completely different and independent method making a supplementary argument in the direction of this hypothesis. In addition, we bring additional evidence to answer the question opened by Fredriksen et al. (2020) about the proportionality of the frequency increasing with the radiative forcing. Our results with EOF decomposition suggest that the increase in frequency is greater if the radiative forcing is stronger. Our description of the increasing frequency of ENSO is reinforced by the robust agreement among models regarding this evolution despite their large differences in other aspects. Indeed, they all used different resolutions and model components. They also exhibit various ensemble variance time series and equilibrium climate sensitivity (Meehl et al. 2020, Table 1) and are relatively independent inside the CMIP6 framework (Brunner et al. 2020). Previous studies (Cai et al. 2014, 2015) have shown an increase in the frequency of extreme El Niño and La Niña events which seems consistent with the general increase in the frequency we identified. Finally, this increase appears compatible with the mechanisms explaining the changes in variability at low latitudes proposed in Olonscheck et al. (2021) and Shi et al. (2023) (i.e., changes in vegetation, clouds, soil moisture, etc.) which could even be consequences of the change of ENSO, since its teleconnections are known to have a profound impact on the water cycle in America, Asia, and Africa, closely linked to these mechanisms.

The results obtained with the time filtering to separate the time scales of variability have been compared to various low, high, and bandpass filterings. The relative importance of each time scale appears sensitive to the filtering method with, for instance, a more important 3–5-yr band and a lower 1–3-yr band using a fourth-order Butterworth filter. However, regardless of the method used, we found the increasing frequency of the variability over time and an intensification of the interannual variability. The dominant period can be either 1–3 or 3–5 years depending on the methodology used, but it remains an interannual variability strengthening and a shift toward higher frequencies. In the present work, we have not studied shorter time scales such as seasonal or intra-annual variability. Some high-impact events (e.g., repeated heatwaves) occur on these shorter time scales, and changes in their variability could be different from those studied. Shi et al. (2023) investigated such seasonal changes and showed that the decreasing high-latitude signal is reinforced in winter.

The computational cost of ensemble simulations remains a limiting factor for the study of internal variability. On the one hand, the quite coarse spatial resolution of CMIP6 models means that some of them may have difficulty representing specific climate modes such as the different types of El Niño (canonical or Modoki). However, although some of the selected models have shown such difficulties, they all show a relatively good ability to represent ENSO in general, and more particularly to predict its periodicity (Hou and Tang 2022), which is the focus of our attention here. On the other hand, ensemble size is also a limitation. Large ensembles with a hundred members have already been run, but not on such long time scales with such a wide variety of scenarios. A larger number of members could help improve the resolution of internal variability by visiting more possible states. However, our main results are consistent between the four ensemble models, which have different sizes, and the confidence intervals built by selecting members randomly are very narrow. We can therefore expect the size of the ensembles to be sufficient to resolve the processes studied here.

The internal variability of the climate is known to have a potentially strong impact on human societies and ecosystems. ENSO is a perfect example, with warm El Niño events causing forest fires and droughts in the western Pacific, Australia, and Asia and marine heat waves in the eastern Pacific, while cold La Niña events cause droughts in South America and floods in Asia and Australia (Cai et al. 2021). In the present

study, we found dramatic changes over the Amazon rain forest with interannual SAT standard deviation increasing by 0.67 K, which corresponds to more than a doubling.

It is usual to interpret climate changes as a modification of the forced variability or similarly of the background state. However, here we have shown that we may also be entering an era of changing "internal" climate variability. For some regions, this means a more variable, volatile, and unstable climate. This has the potential to strongly impact ecosystems (Dee et al. 2020) and societies (Bathiany et al. 2018), in particular around the Pacific where the changes are stronger, but also on the whole planet given the variance changes observed over tropical Africa and the known teleconnections of ENSO. It is key to integrate these variance changes in future adaptation plans.

Acknowledgments. We thank the World Climate Research Program, especially the Coupled Modeling Working Group for coordinating the CMIP program and the Earth System Grid Federation (ESGF) for distributing the datasets, and thank the various modeling groups for making their model outputs available. We are grateful to all contributors to the xarray Python package (Hoyer et al. 2023, v2023.09.0) that was used for the analyses in this work. We would also like to thank the editor, Isla Ruth Simpson, and the two anonymous reviewers who helped us improve our paper. This research was supported by the ARVOR project funded through the French CNRS/INSU/LEFE program, by the OceaniX project funded through the French ANR program, and by the ISblue project, Interdisciplinary graduate school for the blue planet (ANR-17-EURE-0015) and co-funded by a grant from the French government under the program "Investissements d'Avenir." A. H. is supported by an ESA Living Planet Fellowship (PACIFIC project). The authors have no relevant financial or nonfinancial interests to disclose.

Data availability statement. CMIP6 ensemble model outputs can be downloaded from the different Earth System Grid Federation (ESGF) nodes. The references of the four CMIP6 ensemble simulations used are Ziehn et al. (2020) for ACCESS-ESM1-5, Swart et al. (2019) for CanESM5, Tatebe et al. (2019) for MIROC6, and Olonscheck et al. (2023) for MPI-ESM1-2-LR.

REFERENCES

Baede, A. P. M., E. Ahlonsou, Y. Ding, and D. S. Schimel, 2001: The climate system: An overview. *Climate Change 2001: Impacts, Adaptation and Vulnerability*, B. Bolin and S. Pollonais, Eds., Cambridge University Press, 87–98.

Bathiany, S., V. Dakos, M. Scheffer, and T. M. Lenton, 2018: Climate models predict increasing temperature variability in poor countries. *Sci. Adv.*, **4**, eaar5809, https://doi.org/10.1126/sciadv.aar5809.

Bock, L., and Coauthors, 2020: Quantifying progress across different CMIP phases with the ESMValTool. *J. Geophys. Res. Atmos.*, **125**, e2019JD032321, https://doi.org/10.1029/2019JD032321.

Brunner, L., A. G. Pendergrass, F. Lehner, A. L. Merrifield, R. Lorenz, and R. Knutti, 2020: Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth Syst. Dyn.*, **11**, 995–1012, https://doi.org/10.5194/esd-11-995-2020.

Cai, W., and Coauthors, 2014: Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat. Climate Change*, **4**, 111–116, https://doi.org/10.1038/nclimate2100.

——, and Coauthors, 2015: Increased frequency of extreme La Niña events under greenhouse warming. *Nat. Climate Change*, **5**, 132–137, https://doi.org/10.1038/nclimate2492.

——, and Coauthors, 2021: Changing El Niño–Southern Oscillation in a warming climate. *Nat. Rev. Earth Environ.*, **2**, 628–644, https://doi.org/10.1038/s43017-021-00199-z.

——, B. Ng, G. Wang, A. Santoso, L. Wu, and K. Yang, 2022: Increased ENSO sea surface temperature variability under four IPCC emission scenarios. *Nat. Climate Change*, **12**, 228–231, https://doi.org/10.1038/s41558-022-01282-z.

Callahan, C. W., C. Chen, M. Rugenstein, J. Bloch-Johnson, S. Yang, and E. J. Moyer, 2021: Robust decrease in El Niño/Southern Oscillation amplitude under long-term warming. *Nat. Climate Change*, **11**, 752–757, https://doi.org/10.1038/s41558-021-01099-2.

Cassou, C., A. Cherchi, and Y. Kosaka, Eds., 2021: Annex IV: Modes of variability. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 2153–2192.

Chen, D., and Coauthors, 2021: Framing, context, and methods. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 147–286.

Coburn, J., and S. C. Pryor, 2023: Evolution of the internal climate modes under future warming. *J. Climate*, **36**, 511–529, https://doi.org/10.1175/JCLI-D-22-0200.1.

Dee, L. E., D. Okamtoto, A. Gårdmark, J. M. Montoya, and S. J. Miller, 2020: Temperature variability alters the stability and thresholds for collapse of interacting species. *Philos. Trans. Roy. Soc.*, **B375**, 20190457, https://doi.org/10.1098/rstb.2019.0457.

Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, https://doi.org/10.1007/s00382-010-0977-x.

——, and Coauthors, 2020: Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Climate Change*, **10**, 277–286, https://doi.org/10.1038/s41558-020-0731-2.

Fredriksen, H.-B., J. Berner, A. C. Subramanian, and A. Capotondi, 2020: How does El Niño–Southern Oscillation change under global warming—A first look at CMIP6. *Geophys. Res. Lett.*, **47**, e2020GL090640, https://doi.org/10.1029/2020GL090640.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, **90**, 1095–1108, https://doi.org/10.1175/2009BAMS2607.1.

Hou, M., and Y. Tang, 2022: Recent progress in simulating two types of ENSO – from CMIP5 to CMIP6. *Front. Mar. Sci.*, **9**, 986780, https://doi.org/10.3389/fmars.2022.986780.

Hoyer, S., and Coauthors, 2023: xarray. Zenodo, accessed 1 October 2023, https://doi.org/10.5281/zenodo.8379187.

Jones, P. W., 1999: First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon. Wea. Rev.*, **127**, 2204–2210, https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2.

Knutti, R., 2010: The end of model democracy? *Climatic Change*, **102**, 395–404, https://doi.org/10.1007/s10584-010-9800-2.

Lee, J.-Y., and Coauthors, 2021: Future global climate: Scenario-based projections and near-term information. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 553–672.

Lehner, F., C. Deser, N. Maher, J. Marotzke, E. M. Fischer, L. Brunner, R. Knutti, and E. Hawkins, 2020: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth Syst. Dyn.*, **11**, 491–508, https://doi.org/10.5194/esd-11-491-2020.

Maher, N., D. Matei, S. Milinski, and J. Marotzke, 2018: ENSO change in climate projections: Forced response or internal variability? *Geophys. Res. Lett.*, **45**, 11 390–11 398, https://doi.org/10.1029/2018GL079764.

——, S. Milinski, and R. Ludwig, 2021: Large ensemble climate model simulations: Introduction, overview, and future prospects for utilising multiple types of large ensemble. *Earth Syst. Dyn.*, **12**, 401–418, https://doi.org/10.5194/esd-12-401-2021.

——, and Coauthors, 2023: The future of the El Niño–Southern Oscillation: Using large ensembles to illuminate time-varying responses and inter-model differences. *Earth Syst. Dyn.*, **14**, 413–431, https://doi.org/10.5194/esd-14-413-2023.

Meehl, G. A., C. A. Senior, V. Eyring, G. Flato, J.-F. Lamarque, R. J. Stouffer, K. E. Taylor, and M. Schlund, 2020: Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Sci. Adv.*, **6**, eaba1981, https://doi.org/10.1126/sciadv.aba1981.

Milinski, S., N. Maher, and D. Olonscheck, 2020: How large does a large ensemble need to be? *Earth Syst. Dyn.*, **11**, 885–901, https://doi.org/10.5194/esd-11-885-2020.

Olonscheck, D., A. P. Schurer, L. Lücke, and G. C. Hegerl, 2021: Large-scale emergence of regional changes in year-to-year temperature variability by the end of the 21st century. *Nat. Commun.*, **12**, 7237, https://doi.org/10.1038/s41467-021-27515-x.

——, and Coauthors, 2023: The new Max Planck institute grand ensemble with CMIP6 forcing and high-frequency model output. *J. Adv. Model. Earth Syst.*, **15**, e2023MS003790, https://doi.org/10.1029/2023MS003790.

Perkins-Kirkpatrick, S. E., and P. B. Gibson, 2017: Changes in regional heatwave characteristics as a function of increasing global temperature. *Sci. Rep.*, **7**, 12256, https://doi.org/10.1038/s41598-017-12520-2.

Rehfeld, K., R. Hébert, J. M. Lora, M. Lofverstrom, and C. M. Brierley, 2020: Variability of surface climate in simulations of past and future. *Earth Syst. Dyn.*, **11**, 447–468, https://doi.org/10.5194/esd-11-447-2020.

Roberts, C. D., M. D. Palmer, D. McNeall, and M. Collins, 2015: Quantifying the likelihood of a continued hiatus in global warming. *Nat. Climate Change*, **5**, 337–342, https://doi.org/10.1038/nclimate2531.

Sévellec, F., and S. S. Drijfhout, 2018: A novel probabilistic forecast system predicting anomalously warm 2018-2022 reinforcing the long-term global warming trend. *Nat. Commun.*, **9**, 3024, https://doi.org/10.1038/s41467-018-05442-8.

Shi, J., Z. Tian, X. Lang, and D. Jiang, 2023: Projected changes in the interannual variability of surface air temperature using CMIP6 simulations. *Climate Dyn.*, **62**, 431–446, https://doi.org/10.1007/s00382-023-06923-3.

Shin, N.-Y., J.-S. Kug, M. F. Stuecker, F.-F. Jin, A. Timmermann, and G.-I. Kim, 2022: More frequent central Pacific El Niño and stronger eastern Pacific El Niño in a warmer climate. *npj Climate Atmos. Sci.*, **5**, 101, https://doi.org/10.1038/s41612-022-00324-9.

Swart, N. C., and Coauthors, 2019: The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci. Model Dev.*, **12**, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019.

Tatebe, H., and Coauthors, 2019: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geosci. Model Dev.*, **12**, 2727–2765, https://doi.org/10.5194/gmd-12-2727-2019.

Zhuang, J., and Coauthors, 2023: pangeo-data/xESMF: v0.8.2. Zenodo, accessed 1 October 2023, https://doi.org/10.5281/zenodo.8356796.

Ziehn, T., and Coauthors, 2020: The Australian Earth System Model: ACCESS-ESM1.5. *J. South. Hemisphere Earth Syst. Sci.*, **70**, 193–214, https://doi.org/10.1071/ES19035.