AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2021MS002565

Special Section:
The IPSL climate model used in CMIP6

Presentation and Evaluation of the IPSL-CM6A-LR Ensemble of Extended Historical Simulations

Rémy Bonnet[1], Olivier Boucher[1], Julie Deshayes[2], Guillaume Gastineau[2], Frédéric Hourdin[3], Juliette Mignot[2], Jérôme Servonnat[4], and Didier Swingedouw[5]

[1]Institut Pierre-Simon Laplace, Sorbonne Université/CNRS, Paris, France, [2]Laboratoire d'Océanographie et du Climat: Expérimentations et Approches Numériques, Institut Pierre-Simon Laplace, Sorbonne Université/CNRS/IRD MNHN, Paris, France, [3]Laboratoire de Météorologie Dynamique, Institut Pierre-Simon Laplace, Sorbonne Université/CNRS/Ecole Normale Supérieure/Ecole Polytechnique, Paris, France, [4]Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace, CEA/CNRS/UVSQ, Gif-sur-Yvette, France, [5]Environnements et Paléoenvironnements Océaniques et Continentaux, Université de Bordeaux/CNRS, Bordeaux, France

Abstract The Institut Pierre-Simon Laplace Climate Modeling Center has produced an ensemble of extended historical simulations using the IPSL-CM6A-LR climate model. This ensemble (referred to as IPSL-EHS) is composed of 32 members over the 1850–2059 period that share the same external forcings but differ in their initial conditions. In this study, we assess the simulated decadal to multidecadal climate variability in the IPSL-EHS. In particular, we examine the global temperature evolution and recent warming trends, and their consistency with ocean heat content and sea ice cover. The model exhibits a large low-frequency internal climate variability. In particular, a quasi-bicentennial mode of internal climate variability is present in the model and is associated with the Atlantic Meridional Overturning Circulation. Such variability modulates the global mean surface air temperature changes over the historical period by about ~ 0.1K. This modulation is found to be linked to the phase present in the initial condition state of each member. This variability appears to decrease during the 1850–2018 period in response to external forcings. The analysis of the ocean heat content reveals furthermore an overestimation of the ocean stratification, which likely leads to an overestimation of the recent warming rate on average.

Plain Language Summary The Institut Pierre-Simon Laplace (IPSL) developed an ensemble of 32 simulations over the 1850–2059 period using the IPSL-CM6A-LR climate model. Such a large ensemble allows a better sampling of the internally generated variability. Moreover, the ensemble averaging provides an estimation of the forced variability induced by the greenhouse gases and the aerosol concentration used as boundary conditions. In this study, we assess the simulated decadal to multidecadal climate variability in the IPSL ensemble. Relative to the large variability of the model, the evolution of observed surface temperature and sea ice cover is within the range of possibilities of the ensemble. The oceanic circulation and sea surface temperature over the North Atlantic are key players in the low-frequency internal variability of the model.

Key Points:

- A large part of the spread of temperature and sea ice trends in the IPSL ensemble is related to a large multicentennial internal variability
- Some members of the IPSL ensemble are consistent with the observed surface temperature, sea ice variations, and ocean heat content evolution
- The low-frequency internal climate variability of IPSL-CM6A-LR decreases since the 2000s in response to external forcing

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

R. Bonnet,
remy.bonnet@ipsl.fr

Citation:

Bonnet, R., Boucher, O., Deshayes, J., Gastineau, G., Hourdin, F., Mignot, J., et al. (2021). Presentation and evaluation of the IPSL-CM6A-LR ensemble of extended historical simulations. *Journal of Advances in Modeling Earth Systems*, *13*, e2021MS002565. https://doi.org/10.1029/2021MS002565

Received 12 APR 2021
Accepted 21 AUG 2021

1. Introduction

Uncertainties in climate projections can be decomposed into three contributions due to (a) internal variability—defined as the unforced intrinsic variability of the climate system—, (b) uncertainties in the model physics—resulting largely from the unresolved processes, notably linked to the model resolution—, and (c) the range of possible future external forcings which in turn depend on emission scenarios and natural forcings (Hawkins & Sutton, 2009). Over the coming decades, uncertainties related to decadal to multidecadal internal climate variability may modulate trends regionally and potentially also on the global scale. For instance, internal climate variability was found to contribute by up to more than 50% in projected precipitation change in 10 years over the British Islands (Lehner et al., 2020). In order to better understand the role and the uncertainties related to the internal climate variability, a simple method is to identify the recurring patterns in sea surface temperature at decadal to multidecadal timescale, also called "modes" of variability (Deser et al., 2010). However, this approach has limitations, as it is difficult to correctly separate the internal and forced components of climate variations given the too short observational (instrumental) record.

A leading pattern of observed sea surface temperature (SST) on these timescales is the Interdecadal Pacific Variability (IPV; Power et al.,1999). This mode is closely related to the Pacific Decadal Variability (PDV, also called Pacific Decadal Oscillation; Mantua et al., 1997), which can be regarded as the North Pacific component of the Pacific-wide IPV. Modeling studies (Dai et al., 2015; Meehl et al., 2013, 2016), as well as observational studies (Folland et al., 2018), indicate a potential role of the IPV in modulating the global surface air temperature (GSAT). Internal variability and its relation to ocean heat uptake have been put forward as an explanation for the recent slowdown in the rate of GSAT increase between 1998 and 2012, the so-called "hiatus" (England et al., 2014; Kosaka & Xie, 2013; Meehl et al., 2011; Swingedouw et al., 2017). The relative cooling trend observed in the Pacific Ocean around the 2000s, whose spatial pattern resembles the IPV, is due to the increase in equatorial upwelling associated with stronger westerlies over this period (England et al., 2014), trapping more heat in the deep ocean. The distribution and importance of this extra heat pumped into the Pacific Ocean are still under debate. Some studies suggest that a large fraction of this ocean heat uptake has been stored into the Indian Ocean via the Indonesian throughflow (Lee et al., 2015; Liu et al., 2016b; Nieves et al., 2015). Other studies point out that temperature increase at lower layers in the Atlantic and Southern Oceans could also account for the sequestration (Chen & Tung, 2014; Cheng et al., 2017; Gastineau et al., 2019; Llovel & Terray, 2016). Furthermore, the relative importance of internally driven and anthropogenically driven changes is still debated (Chen & Tung, 2016; Liu et al., 2016a).

At the decadal to multidecadal timescales, the variability of the Atlantic meridional overturning circulation (AMOC) is also believed to be an important driver of the climate variability, as it contributes to redistribute heat within the ocean across latitudes and across the two hemispheres (e.g., Delworth & Zeng, 2016; Dong & Sutton, 2005; Polyakov et al., 2010). In many climate models, one prominent feature related to the AMOC low-frequency variability is basin-wide low-frequency variations of SST in the North Atlantic (e.g., Danabasoglu et al., 2012; Delworth et al., 1993). The Atlantic Multidecadal Variability (AMV), also called Atlantic Multidecadal Oscillation (AMO), is the leading mode of SST variability on multidecadal timescales (Enfield et al., 2001; Kerr, 2000; Schlesinger & Ramankutty, 1994; Yeager & Robson, 2017). It has recently been suggested that at multidecadal timescales, the AMV is the main driver of the GSAT variability (Li et al., 2020). The AMV has been linked to many observed low-frequency global and regional climate variations, such as the Northern Hemisphere temperature (Zhang et al., 2007), Sahel rainfall (Zhang & Delworth, 2006), Arctic sea ice extent (Zhang, 2015), Atlantic hurricane activity (Zhang & Delworth, 2006), or European precipitation and temperature (Sutton & Dong, 2012). Although the AMV seems to be primarily linked to AMOC variations in climate models, some studies also suggest that AMV may result from a direct effect of atmospheric internal variability, without a role for ocean circulation (Clement et al., 2015). Additionally, there are many open questions and uncertainties about its nature and characterization. These uncertainties are mainly due to the short observational window (~150 years) in comparison to the AMV typical timescale, the difficulties to observe the AMOC, and the fact that the observational period is influenced by both external forcings and other modes of internal climate variability. The degree of influence of these two factors over the historical period is still under debate. Some studies suggest that anthropogenic aerosols have favored or even forced the negative phase of the AMV in the 1970s (Booth et al., 2012). Other studies suggest an influence of volcanic aerosols on the AMV over the second half of the 20th century (Otterå et al., 2010; Swingedouw et al., 2015), as well as anthropogenic aerosols (e.g., Booth et al., 2012). Using the Community Earth System Model (CESM) Large Ensemble and Last Millennium Ensemble, Bellomo et al. (2018) suggest that external forcings are the main drivers of the AMV. This was also suggested more recently by Haustein et al. (2019), using a two-box impulse response model. However, the pre-industrial control simulations in climate models generally show important multidecadal variations in the North Atlantic SST, which favors the hypothesis that the AMV is, at least partially, an internal mode of variability.

Multi-model ensembles such as those made available in the successive phases of the Coupled Model Intercomparison Project (Taylor et al., 2012, CMIP3 and CMIP5) can be used to estimate the combined effect of the three types of uncertainties described above, but it remains difficult to separate the internal variability from the forced response because of the insufficient number of simulations performed by every single model together with the large spread among models (Deser et al., 2012).

When only a single realization of the historical period is available, the forced climate response is classically approximated from a linear trend and any departure from the trend is interpreted as natural variability (e.g.,

Knight et al., 2005; Wyatt et al., 2012). This method, however, is inaccurate and may lead to misinterpretations (Frankcombe et al., 2018; Frankignoul et al., 2017; Mann et al., 2014). A better way to estimate the internal variability of a model is to use a multicentennial climate simulation with constant forcing such as the pre-industrial control simulations of the CMIP protocol (Taylor et al., 2012), as done for example, by Wittenberg et al. (2014). However, this method does not account for the influence of external forcings on the internal variability (Maher et al., 2015; Swingedouw et al., 2015). When enough realizations of the historical period are available for a given model, the effects of natural and anthropogenic external forcings can be isolated from the effects of internal variability by defining the forced signal as the ensemble average (Deser et al., 2012), assuming the effects of internal variability largely cancel out when the ensemble is large enough.

In this context, large ensembles of coupled climate models are necessary to isolate the internal variability from the forced climate response (Kay et al., 2015). The National Center for Atmospheric Research (NCAR) pioneered the field by performing a large ensemble of 62 members over the 1940–2080 period (Selten et al., 2004). This was followed by the Community Earth System Model Large Ensemble project (CESM-LE), composed of 42 members over the 1920–2100 period (Kay et al., 2015; Sanderson et al., 2018). Other modeling groups have also performed large ensembles: (a) the Geophysical Fluid Dynamics Laboratory (GFDL) developed a 30-member ensemble from 1850 to 2100 (Delworth et al., 2020) and (b) the Canadian Earth System Model Large Ensemble is composed of 50 members over the 1950–2005 period (Kirchmeier-Young et al., 2017); (c) a 30-member ensemble was developed for the 1940–2020 period (Frankignoul et al., 2017) based on the IPSL-CM5A-LR model (Dufresne et al., 2013); and (d) more recently the Max Planck Institute performed a grand ensemble (MPI-GE; Maher et al., 2019), with 100 members for the *historical* simulations and the four RCPs scenarios. Recently, Deser et al. (2020) reviewed the opportunities and perspectives given by these large ensemble simulations.

Over the observational period, the internal climate variability can be assessed by comparing each member of an ensemble of simulations to the observations, as each member represents a potential realization of the real world. The development of such large ensemble simulations has, therefore, led to some important insights regarding the uncertainties related to internal climate variability, in particular for temperature and precipitation (e.g., Dai & Bloecker, 2019; Deser et al., 2012, 2014; Thompson et al., 2015) or sea ice (e.g., Wettstein & Deser, 2014). Large ensembles have also been used to understand the respective roles of the forced and internal components, for example, in temperature changes at the global and regional scales (e.g., Deser et al., 2016), but also for extreme events, in terms of surface temperature (Suarez-Gutierrez et al., 2018; Trenberth et al., 2015), precipitation (Hagos et al., 2016; Wang et al., 2018), drought (Gu et al., 2019), or sea ice extent (Kirchmeier-Young et al., 2017). Multiple ensemble climate models can also be used to have a better estimation of the forced response (Kravtsov & Callicutt, 2017; Liguori et al., 2020).

While the number of modeling centers that have produced large ensembles is still very limited, it is important to develop new large ensemble simulations with additional models, to further constrain forced versus natural variability and to assess the robustness of the simulated internal climate variability and the forced response from different models. In this context, the Institut Pierre-Simon Laplace Climate Modeling Center (IPSL CMC, see https://cmc.ipsl.fr) took advantage of Phase 6 of the Coupled Model Intercomparison Project (CMIP6; Eyring et al., 2016) to develop an ensemble of extended *historical* simulations with the IPSL-CM6A-LR model (Boucher et al., 2020). The IPSL ensemble of extending historical simulations (IPSL-EHS) is composed of 32 members using initial conditions sampled along a pre-industrial control experiment. As a result, each member has a unique trajectory that is influenced by both the external forcings and internal climate variability. The IPSL-EHS allows an informative comparison between the model and the observations. As the IPSL-EHS is available on a longer period than the previous one (Frankignoul et al., 2017), the internal climate variability of the IPSL-EHS can be better assessed, especially at decadal to multidecadal timescales. Finally, the IPSL-EHS is based on a model that has significantly improved in terms of physics content (Hourdin, Rio, Grandpeix, et al., 2020) and climatology (Boucher et al., 2020) and that uses the more realistic external forcing reconstructions from CMIP6 (Lurton et al., 2020).

The first objective of this study are to evaluate the capacity of the IPSL-CM6A-LR model to encompass the observed variability and changes in surface temperature and ocean heat content despite its rather large climate sensitivity, with an equilibrium climate sensitivity of 4.5 K (though it reduces to 3.8 K if computed

from a $2 \times CO_2$ rather than $4 \times CO_2$ abrupt forcing, as discussed in Boucher et al., 2020). The implications on sea ice, which is a good indicator of global temperature change, will also be evaluated. The second objective is to evaluate the impact of the low-frequency internal variability of the IPSL-EHS on historical trajectory and to investigate the dynamics of the most coherent members against the observations regarding some known modes of variability.

This study is organized as follows. In Section 2, the model used is briefly described and the characteristics of the simulation are presented. The simulation of the surface temperature evolution is assessed in Section 3. The changes in the ocean heat content and sea ice are then assessed in Sections 4 and 5 respectively. The evolution of the Atlantic meridional overturning circulation and its influence on surface temperature trends are then studied in Section 6. In Section 7, the simulations of the Interdecadal Pacific Variability and the Atlantic Multidecadal Variability are evaluated. Finally, in Section 8, the main conclusions of these evaluations are drawn and possible use of the IPSL-EHS is discussed.

2. The IPSL-CM6A-LR Large Ensemble

2.1. The IPSL-CM6A-LR Model

IPSL-CM6A-LR is the standard configuration of the Institut Pierre-Simon Laplace (IPSL) coupled model developed for the 6th phase of the Coupled Model Intercomparison Project (CMIP6). It combines the LMDZ6 Atmospheric model (Hourdin, Rio, Jam, et al., 2020), the ORCHIDEE land surface model (Krinner et al., 2005), and the NEMO ocean model (Madec et al., 2017) using the LIM3 sea-ice model (Rousset et al., 2015). The atmospheric model resolution is now 144 × 143 points in latitude and longitude and 79 vertical layers (with a maximum height of about 80 km). This corresponds to the medium horizontal resolution of the previous model version (IPSL-CM5A-MR) but the number of vertical layers was roughly doubled. The horizontal resolution of the ocean model is increased to 1° with 75 layers in the vertical. The atmospheric physics of LMDZ6A is further described in Hourdin et al. (2019). The representation of clouds is described in Madeleine et al. (2020), while Hourdin, Rio, Jam, et al. (2020) linked the reduction in SST biases to improvements in atmospheric physics. Cheruy et al. (2019) discuss the atmosphere-surface coupling. Lurton et al. (2020) document the implementation of the CMIP6 climate forcings in the model while Boucher et al. (2020) present the coupled model and a preliminary evaluation of the *historical* simulations performed for CMIP6.

2.2. The Simulations and Data Availability

The IPSL ensemble of extended *historical* simulations' initial conditions is taken from different years in a long preindustrial *picontrol* simulation after it has reached a quasi-stationary state. The simulations are started from the atmospheric, oceanic and land surface initial conditions of the 1st January from different years. A large multicentennial variability in the global mean surface air temperature (GSAT) is present in the *piControl* simulation (Figure 1a). The start dates of the *historical* simulations are therefore spaced every 20–40 years (see Table S1) in order to sample this variability. Members #1 and #32 start on the same date in the *piControl* simulation, but member #32 was run on a different supercomputer than the other members. As in other climate models (Hobbs et al., 2016), although the *piControl* simulation inherits from a long spin-up, the model is not completely equilibrated given the long timescale associated with the deep ocean. As a result, there is a negative drift of −37 ZJ per century in the 0–2,000 m OHC, which can be decomposed into drifts of −4.4 and −33 ZJ per century in the 0–700 m and 700–2,000 m layers, respectively. A negative drift of −0.011 K per century is also present in the GSAT of the *piControl* simulation. Unless otherwise stated, we neglect this surface temperature drift in the rest of the analysis as it is small compared to the forced signal in the *historical* simulations.

The IPSL-EHS follows the CMIP6 protocol for historical simulations (Eyring et al., 2016) for the period 1850–2014. Simulations were extended until 2060 using radiative forcings from the SSP245 scenario (Gidden et al., 2019), except for the ozone field which has been kept constant to its 2014 climatology (as this particular forcing was not available at the time of performing the extensions). This implies that these simulations do not simulate the ozone hole recovery and changes in tropospheric ozone, in contrast to the official CMIP6 projections. Therefore the comparison with other CMIP6 large ensembles is less straightforward

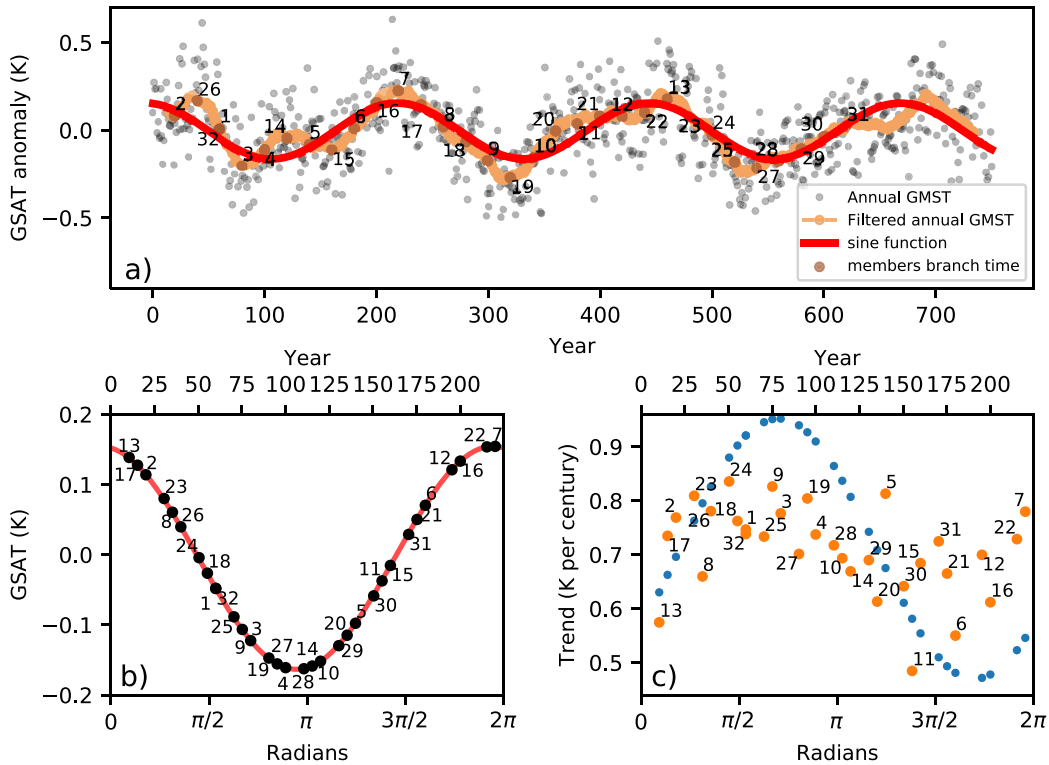

Figure 1. (a) Time evolution of the annual global mean near-surface air temperature anomaly (GSAT, in K) from the *piControl* simulation (gray circles), the low-pass filtered GSAT (with a 31 years window, brown curve), the sine function fit (red curve) and the start dates of the IPSL-EHS members (dark brown circles). (b) Locations of the member start dates (black dots) repositioned over one period of the *piControl* bicentennial mode of temperature variability (red curve, in K) of the IPSL-CM6A-LR model. The locations are estimated from Equation 1. (c) Scatter plot of the GSAT trend calculated over the 1850–2018 period (K per century) relative to the locations of the member start dates from the 32 members of the IPSL-EHS (orange dots) and from GSAT reconstructions based on the 168-years of the sine function of each member relative to their starting dates in the piControl simulation, to which is added the ensemble mean (blue dots).

after 2014. However, this also provides an opportunity to isolate and study the influence of the tropospheric ozone trend after 2014, by comparing the IPSL-EHS with other CMIP6 large ensembles.

The SSP245 scenario was chosen because it is a middle-of-the-road scenario. An additional motivation for that choice is that SSP245 is the scenario used of DCCP-A and DCPP-B protocols (Boer et al., 2016). This will allow relevant comparisons between the IPSL-EHS and initialized decadal predictions under the same external forcing.

2.3. Influence of the Initial Conditions on the Historical Simulations

The IPSL-CM6A-LR model is characterized by a multicentennial temperature variability in the control simulation (Boucher et al., 2020; Figure 1a), which was related to the AMOC (Jiang et al., 2021). The analysis of the mechanisms behind this variability suggests that it is related to a progressive accumulation of surface freshwater in the Arctic Ocean when the AMOC is in an ascending phase. After several decades, this accumulation eventually leaks into the Nordic Seas through the Fram Strait, which then slows down the AMOC and reverses the phase of this cycle. These mechanisms lead to an almost oscillatory behavior at multicentennial timescales. A strong AMOC phase in IPSL-CM6-LR is therefore associated with a warm North Atlantic Ocean and a large Arctic sea-ice loss, associated with warming over the whole Northern Hemisphere. This long-term variability is likely to affect the GSAT evolution of the *historical* simulations. In order to look at the impact of this variability on the historical simulations, we have fitted a sine function (in the least squares sense) to the detrended GSAT evolution of the *piControl* simulation over the 1850–2599 period:

$$y(t) = A\sin(\omega t + \phi) + C \quad (1)$$

where A is the amplitude, $\omega = 2\pi/T$ is the angular frequency corresponding to the period T, ϕ the phase where oscillation starts and C a constant. An amplitude $A = 0.16$ K and a period $T = 225$ years are found using this method. This approximation allows to summarize the starting point of each *historical* simulation by a single number (the phase between 0 and 2π) in relation to this bicentennial variability. The locations of these starting points are summarized in Figure 1b when projected along one period of this bicentennial variability.

As the historical period (168 years) represents 0.75 period of the sine function, the members initialized with a phase between $\pi/2$ and π modulo 2π (i.e., years 56 and 112 of the period) are expected to result in the largest warming if the bicentennial variability remains unchanged (Figure 1b). Indeed these members start during a negative phase of the bicentennial variability and end after 3/4 of a period during a positive phase. This turns out to be the case, as these members have an average trend of 0.75 K per century (range 0.70–0.83 K) over the 1850–2018 period in comparison to the average trend of 0.71 K per century for the 32 members (Figure 1c). Conversely, the members initialized with a phase between $4\pi/3$ and 2π modulo 2π (i.e., years 150 and 225 of the period) are expected to simulate the smallest warming, as these members start during the beginning of a positive phase, and end during a negative phase. This is also the case, with an average trend of 0.66 K per century (range 0.48–0.78 K) for these members (Figure 1c). The quasi-bicentennial variability of IPSL-CM6A-LR therefore seems to modulate the temperature changes over the historical period by about 0.1 K per century. Note that this estimate strongly depends on the period chosen for computing the trends. Over the 110 year period, for example, it can affect the trend by as much as 0.25 K per century.

In order to further test the influence of the quasi-bicentennial variability over the historical period, we reconstructed the GSAT evolution for each member by adding the sine function (from the start date and for the 168 years of the historical period) to the ensemble mean which estimates the temperature response to external forcings. This procedure assumes the quasi-bicentennial variability of these GSAT reconstructions remains unchanged over the historical period as compared to the *piControl* simulations, as well as that the effect of external forcing adds without interactions onto internal variability. The reconstructions with a phase between $\pi/2$ and π modulo 2π show a larger warming than the *historical* simulations, with an average trend of 0.92 K per century (range 0.87–0.95 K) (Figure 1c). Similarly, the reconstructions with a phase between $4\pi/3$ and 2π modulo 2π show a smaller warming than the corresponding *historical* simulations, with an average trend of 0.52 K per century (range 0.47–0.61 K). This figure also shows that a simple superposition of the *piControl* variability on the *historical* ensemble mean would have modulated the temperature changes over the historical period by about 0.4 K per century. The influence of the quasi-bicentennial variability on temperature changes is therefore diminished in the *historical* simulations, suggesting a decrease of the variability over the historical period, which could be due to external forcings. The standard deviation of the GSAT in the detrended *historical* simulations (i.e., after the ensemble mean has been removed) is 0.14 K on average (range 0.12–0.19 K), which is less than the standard deviation of GSAT in the detrended *piControl* simulation (0.2 K). This confirms the decrease of the bicentennial variability over the historical period.

As a result of this variability, there is a relation between the GSAT state at the beginning of the *historical* simulations and the GSAT change over the historical period, with a coefficient of determination r^2 of 0.40, significant with p-value < 5% using a two-tailed probability Student's t-test (Figure 2). The members with the lowest GSAT at the beginning of the historical period tend to be the members with the largest GSAT changes over the historical period and conversely. However, this relationship is limited, on the one hand, by the noise of this mode of variability and by the other modes of internal climate variability that occur on a range of different timescales, and, on the other hand, by the external forcings, which as discussed above seems to damp this mode variability. Indeed, this relationship is stronger when considering the historical GSAT evolution from the conceptual model based on the sine function (described in the previous paragraph) with a r^2 of 0.83 and significant with p-value < 5%, consistently with Figure 1c.

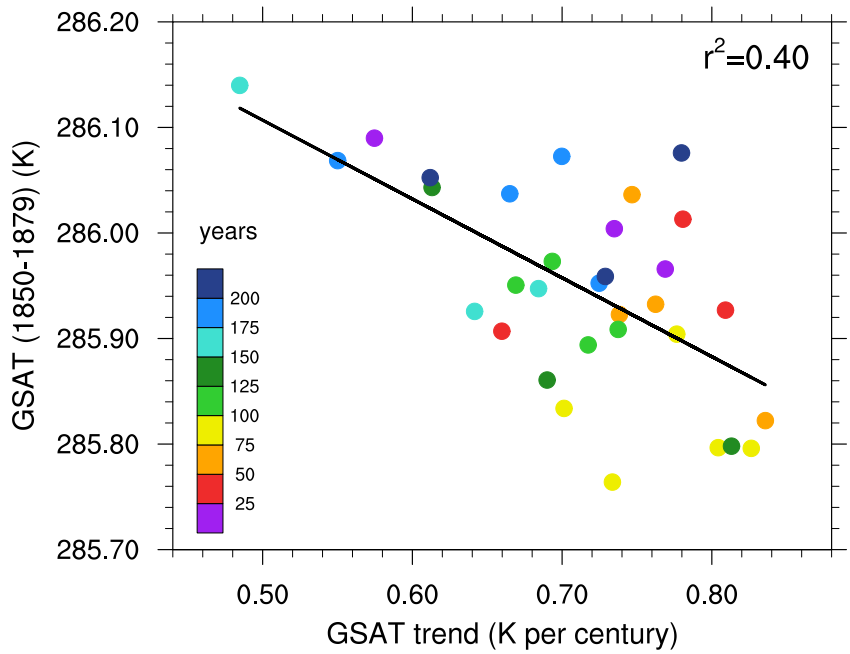

Figure 2. Scatter plot between the average GSAT (K) over the first 30 decades and the GSAT trend (K per century) calculated over the 1850–2018 period from the 32 members of the IPSL-EHS. The dot colors indicate the position (or phase) of the member in the bicentennial GSAT variability period (Figure 1b, upper *x* axis). The black line represents the linear relationship between these two variables, with a determination coefficient r^2 of 0.40.

3. Comparison of the Simulated Surface Temperature With Observations

3.1. Global Mean Surface Air Temperature

The IPSL-EHS covers a large range of GSAT variations and trends over the historical period (Figure 3a). The standard deviation calculated among members over the historical period is of 0.14 K on average. As explained above, the GSAT ensemble mean can be interpreted as the climate response to external forcings (both natural and anthropogenic) while the variations around it are associated with internal variability.

The infilled HadCRUT4-CW (Cowtan & Way, 2014; Jones et al., 2012) observational data set is used to evaluate the historical simulations, allowing us to avoid problems with missing values. It reports blended Sea Surface Temperature (SST) and temperature at the surface, a metric that reports significantly lower temperature than the GSAT metric (Cowtan et al., 2015). A factor of 1.06 is applied to annual temperature to take this discrepancy into account (Gillett et al., 2021). Because of the large uncertainties in the observations before the 1880s (Morice et al., 2012), the evaluation is limited here to the 1880–2018 period. Note that HadCRUT4-CW is very consistent with the Berkeley data set (Rohde, Muller, Jacobsen, Muller, et al., 2013; Rohde, Muller, Jacobsen, Perlmutter, et al., 2013), with a correlation coefficient of 0.99 when using global annual mean (Figure 3a). In this data set, the temperature at the surface is considered at the sea ice level rather than the SST, making it closer to the GSAT metric.

The observed evolution of the GSAT anomaly is within the spread of the historical simulations for both data set (Figure 3a). The modeled ensemble mean departs from the observations over the period 1935–1945, with smaller than observed warming, and after the year 2000, with larger than observed warming. In general, such departures can be due to (a) decadal to multidecadal internal variability in the observations, (b) modeling issues (incorrect representation of external forcings such as aerosol concentrations changes or climate response to the forcings) or (c) observational uncertainties as hypothesized by Folland et al. (2018) for the World War II period. The response to the Krakatoa (1883) and the Pinatubo (1991) volcanic eruptions are represented by the ensemble mean, with a decrease of the GSAT consistent with the observations (Figure 3a).

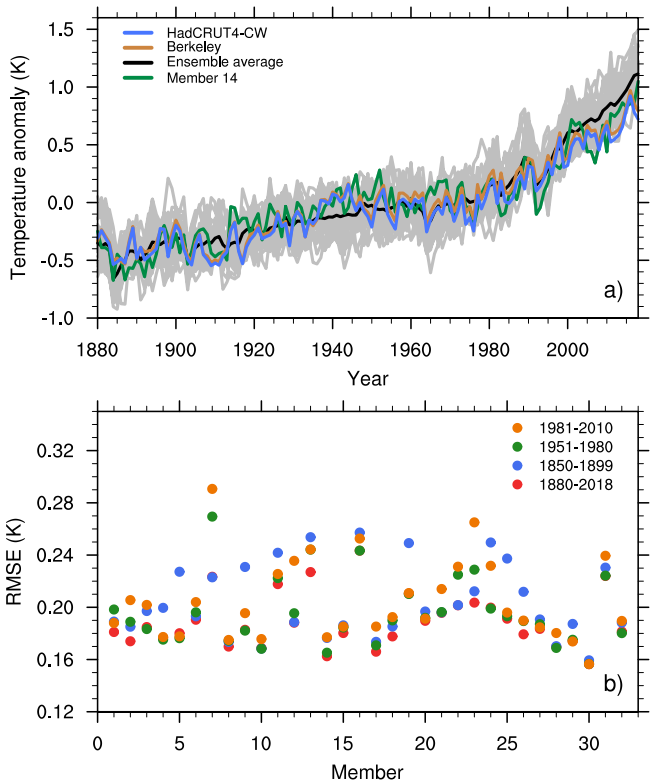

Figure 3. (a) Time evolution of the annual global mean near-surface air temperature (GSAT, in K) anomaly from the IPSL-EHS, with the ensemble average (in black) and each individual member (in gray), the HadCRUT4-CW (Cowtan & Way, 2014; Jones et al., 2012, in blue) and the Berkeley (Rohde, Muller, Jacobsen, Muller, et al., 2013; Rohde, Muller, Jacobsen, Perlmutter, et al., 2013, in brown) data sets, and the closest simulation to the observations based on the root mean square error (RMSE) (member #14, in green). The GSAT anomalies are computed relative to the 1880–2018 average. (b) The RMSE was calculated for each member against the HadCRUT4-CW data set over the 1880–2018 period based on different reference periods: 1880–2018 (red), 1850–1899 (blue), 1951–1980 (green), 1981–2010 (orange).

To evaluate the representation of the GSAT evolution by the 32 historical members, root-mean-square error (RMSE) is calculated for each member over the 1880–2018 period (Figure 3b). Member #14 has one of the best representations of the GSAT in comparison to both data sets, with the lowest RMSE of annual GSAT anomaly over the 1880–2018 period (Figures 3a and 3b). Large differences are present among the historical members, with some members having a good consistency with the observed GSAT, with RMSE between 0.15 and 0.17 K (e.g., members #5, 14, 30). On the contrary, some members have a poor representation of the GSAT over the 1880–2018 period, with RMSE larger than 0.23 K (e.g., members #7, 13, 31). The reference period used to calculate the anomaly generally does not have a large impact on the score, members with a good score remain generally better than members with a lower score irrespectively of the reference periods.

3.2. Recent Trends in Surface Temperature

We focus here on temperature trends over the 1978–2018 period, which is very well observed thanks to the advent of remote sensing and is characterized by quite linear warming warming. The recent warming trend of the GSAT over the 1978–2018 period exhibits a large spread in the IPSL-CM6A-LR ensemble of historical simulations, ranging from 0.20 to 0.47 K per decade for the surface temperature over land, with an average of 0.32 K per decade, and from 0.11 to 0.23 K per decade for the sea surface temperature, with an average of 0.18 K per decade (Figure 4). In comparison, the observed trends of surface temperature are 0.28 K per decade over land and 0.15 K per decade over the ocean. Note that prior to the analysis, the model data, sea surface temperature (SST) over the ocean, and temperature at the surface (TAS) over land, are regridded onto the observations and the missing grid points in the observations are masked temporarily. Consistent with the relative larger than observed warming in the model ensemble mean many of the historical members show a warming trend that is, larger than observations, particularly over the ocean. The trend of the ensemble mean is 0.027 K per decade larger than the trend in the observations over the ocean, and 0.043 K per decade larger over land. Some members have global trends close to those observed both over the ocean and land. Members #5 and 14, which have a good fit with the GSAT evolution, also have warming trends that are consistent with the observations, with respective biases of $-3.2\,10^{-4}$ and $7.3\,10^{-3}$ K per decade for ocean, and of $-2.3\,10^{-2}$ and $-4.1\,10^{-3}$ K per decade for land. Members #16 and 31, which have a high RMSE with the observed GSAT (Figure 3b), show the strongest warming trends over ocean and land. These members have respective biases of 0.080 and 0.076 K per decade for the ocean and of 0.19 and 0.16 K per decade for land. In contrast, member #22, which also has a relatively poor representation of the GSAT, has a much lower warming trend, both over the ocean and land. The consistency of some members with the observed warming, however, is not proof that the model forcings and feedbacks are correct. It could also be due to bias compensations, for example, between greenhouse gases, aerosol forcing, and associated climate feedbacks. It can also be explained by a large internal variability aliasing the high climate sensitivity of the model. Thus the fact that the ensemble encompasses the observations indicates that the model behavior and notably its high climate sensitivity cannot be refuted based only on these observations.

Figure 5 shows the patterns of the recent observed and simulated warming trends for the IPSL-EHS mean and for the four members with the lowest and the highest spatial RMSE respectively. Overall, the model reproduces the land/sea warming contrast, with a warming ratio of 1.61 for the ensemble mean, ranging in the IPSL-EHS from 1.52 to 1.79 between land and ocean over the 1978–2018 period compared to the observed ratio of 1.67 from the HadCRUT4-CW data set (Cowtan & Way, 2014; Jones et al., 2012). The Arctic

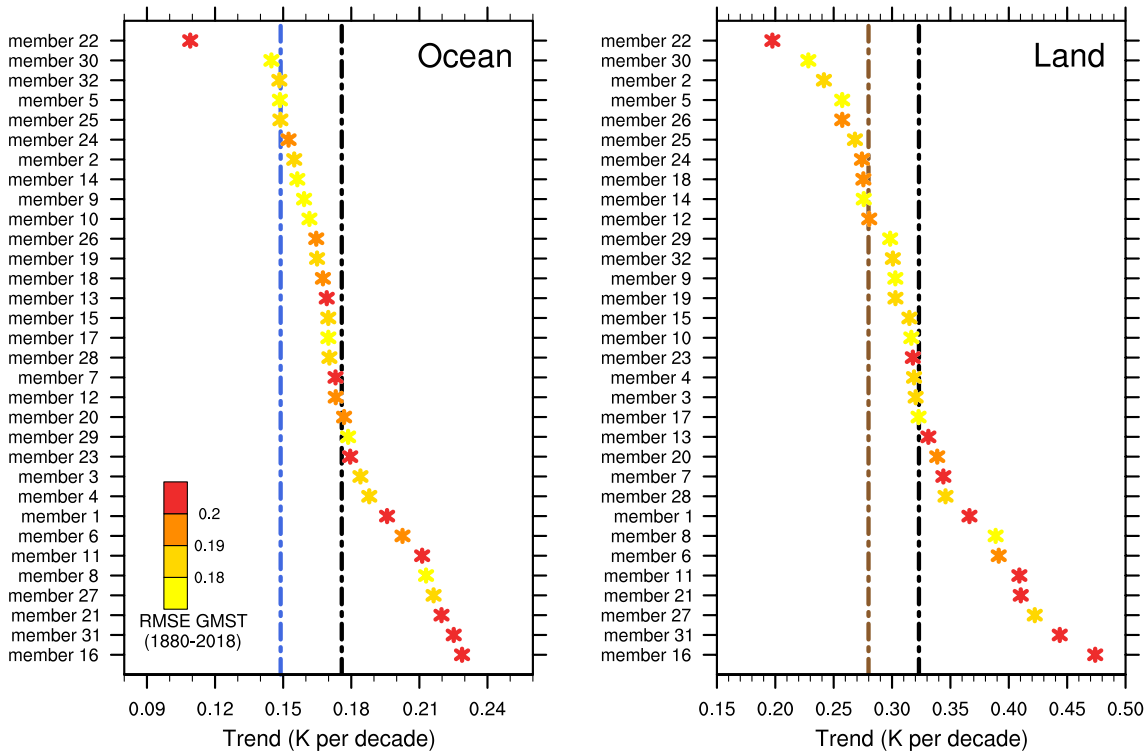

Figure 4. (Left) Trends of average sea surface temperature over the oceans for the 1978–2018 period from the HadSST3 data set of observations (blue dashed line; Kennedy et al., 2011a, 2011b), the ensemble mean of the historical simulations (black dotted line) and the 32 members (stars) ranked from the weakest to strongest trend (from top to bottom). The stars colors are relative to the RMSE of the GSAT (1880–2018). (Right) same as left panel, but for the average land surface temperature using the CruTEM4 observational data set (brown dashed line; Osborn & Jones, 2014).

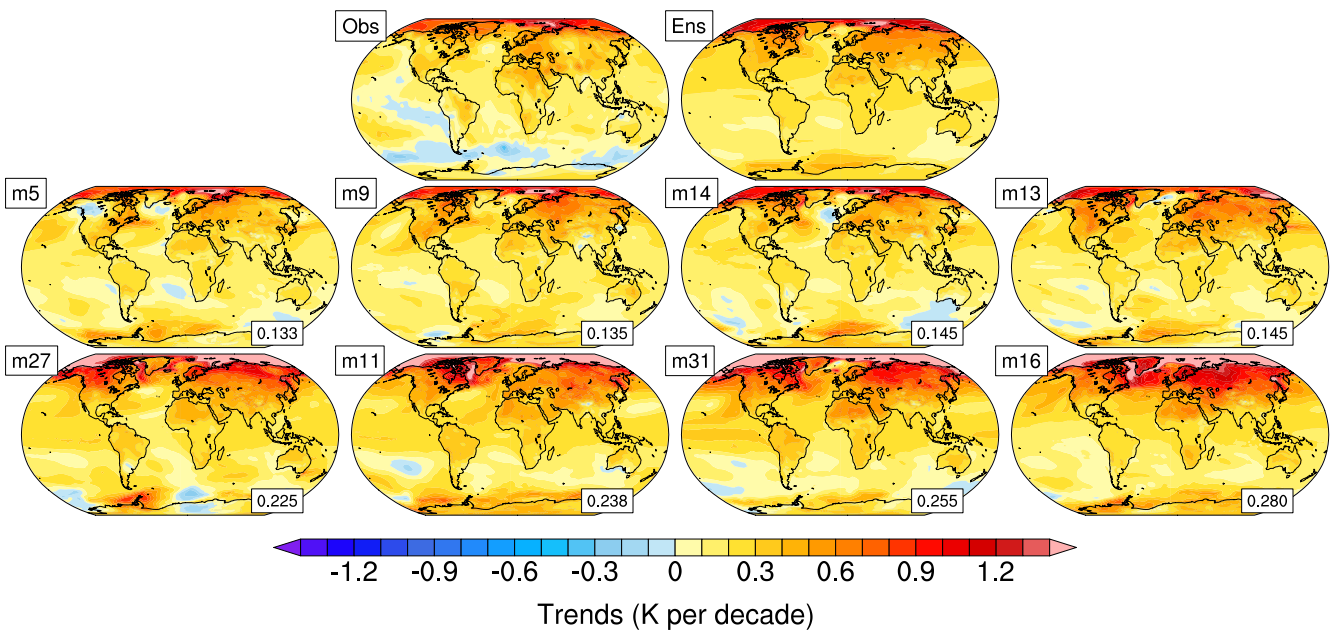

Figure 5. Trends in near surface air temperature (K decade^{-1}) over the 1978–2018 period from the HadCRUT4-CW data set (Cowtan & Way, 2014; Jones et al., 2012) (top left panel) and from the *historical* members of the IPSL-CM6A-LR model, with the ensemble average (top right), the four members with the lowest RMSE relative to the observations (middle panel) and the four members with the highest RMSE (bottom panel). The member number and the RMSE values are indicated on the top-left and bottom-right corners of each panel, respectively. To calculate the RMSE, the model data is regridded onto the grid of the observations with a bilinear interpolation.

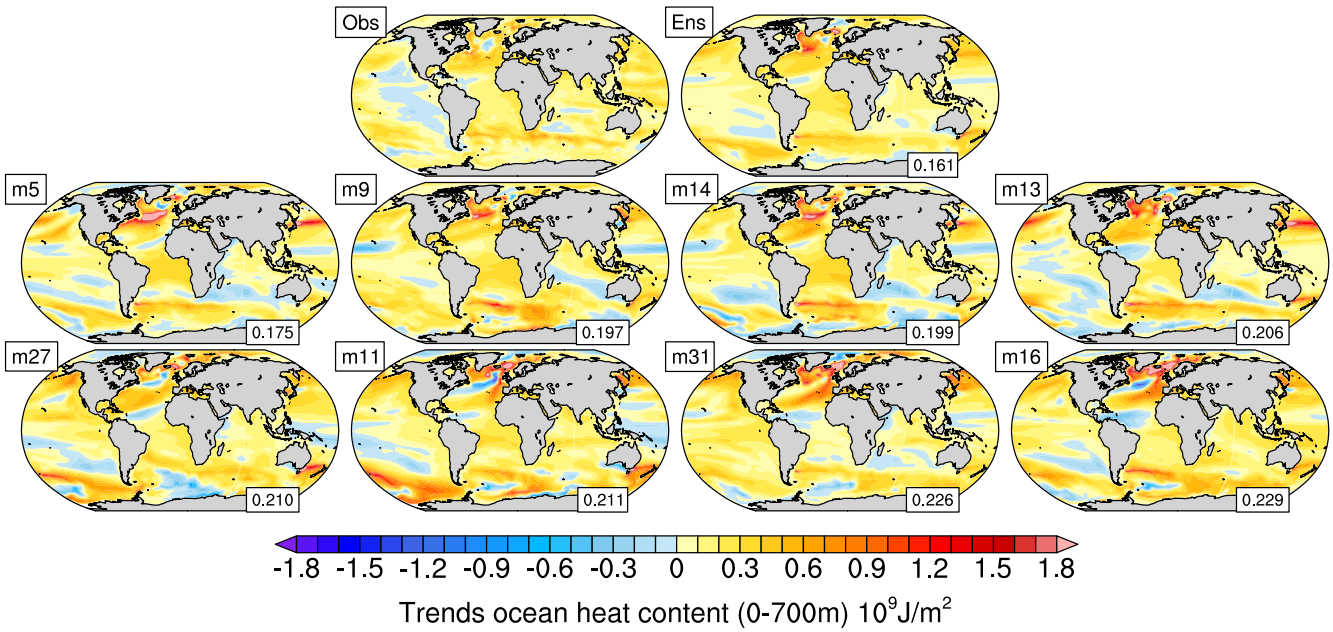

Figure 6. Spatial trends in ocean heat content between 0 and 700 m (10^9 J m^{-2} per decade) over the 1978–2018 period from the observational data set of Cheng and Zhu (2016) (top left) and from the *historical* members of the IPSL-CM6A-LR model, with the ensemble average (top right), the four members with the lowest spatial RMSE (middle panel) and the four members with the highest spatial RMSE (bottom panel) for the surface temperature trends (Figure 5). To calculate the spatial RMSE of the upper ocean heat content trends (indicated at the bottom right), the model data is regridded onto the grid of the observations with a bilinear interpolation.

amplification is also simulated by the IPSL-EHS, but tends to be overestimated by the ensemble mean. Over the 70 − 90°N region, IPSL-EHS is 0.88 K per decade, (from 0.22 to 1.58 K per decade) relative to the trend of 0.26 K per decade over the whole Earth (from 0.16 to 0.36 K per decade), whereas the trend is about 0.79 K per decade over the 70 − 90°N region for the HadCRUT4-CW data set and about 0.19 K per decade over the whole Earth. The members with the greatest Arctic amplifications are those with the greatest overestimation of the temperature trends over land and ocean, highlighting the important contribution of this process in the members that warm most (not shown). Members #31 and 16 exhibits much larger global warming trends than observed (Figure 4), mainly because of the contribution of the Northern Hemisphere. Some members are closer to the observations, with, for example, a representation of the so-called "warming hole" in the Atlantic Ocean. Members with a global warming trend close to the observations are also those with a low spatial RMSE (Figure 4). No member, however, reproduces the cooling trend observed in the Southeastern Pacific or in the Southern Ocean. The ensemble mean shows warming over every region of the Earth.

4. Simulation of the Ocean Heat Content

Ocean heat content is the main energy reservoir of the climate system. Since the 1970s, it has been estimated that approximately 93% of the excess heat energy of the Earth, with respect to climate change, is stored in the ocean (Levitus et al., 2012). A good representation of its evolution is therefore essential to have a plausible estimate of the GSAT. In this section, we assess the OHC of the IPSL-EHS as compared to two data sets of observations (Cheng & Zhu, 2016; Levitus et al., 2012), available over the 0–2,000 m depth. Large uncertainties are present in the OHC observations, mainly related to the lack of data and the data coverage evolution, as well as the different measurement techniques used over the historical period. The OHC drift calculated in the *piControl* simulation is removed from the IPSL-EHS prior to the analysis.

Figure 6 shows the pattern of the recent warming trend (1978–2018) of the upper OHC (0–700 m) in the observations and in the one simulated by the IPSL-EHS ensemble mean, as well as for the four members with the lowest and the highest spatial RMSE of surface temperature trends (Figure 5). Overall, the model tends to have larger trends than the observations over the 1978–2018 period, mainly at high and mid latitudes. The

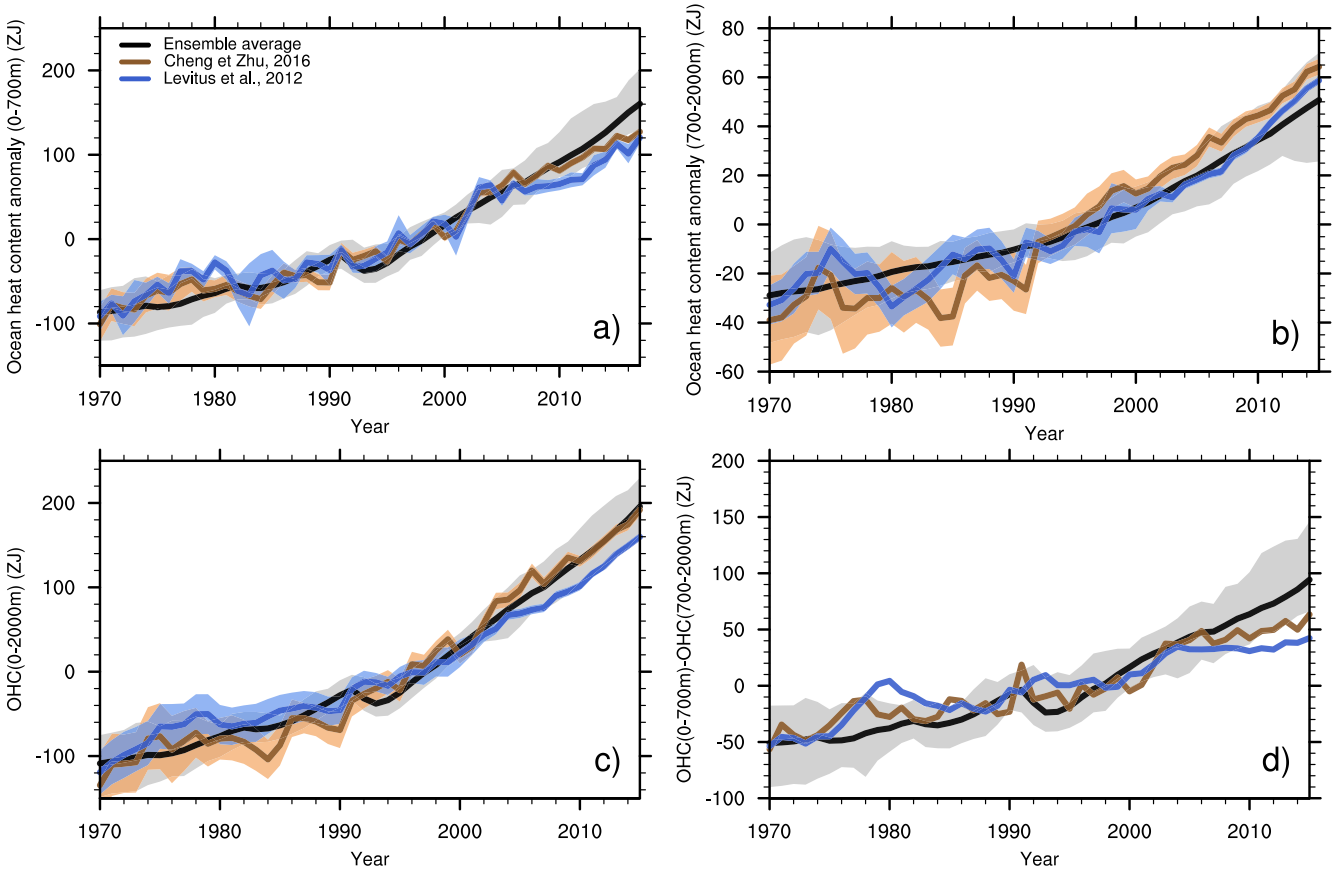

Figure 7. (a) Time evolution of the annual global mean upper (0–700 m) Ocean Heat Content (OHC) anomaly in ZJ (1 ZJ = 10^{21} J) for two observationally based estimates (Levitus et al., 2012, in blue) and (Cheng & Zhu, 2016; Cheng et al., 2017, in brown), with their two standard deviation uncertainties (shading in blue and brown), and for the 32 *historical* members of the IPSL-CM6A-LR model (shading in gray) with the model ensemble mean (black line). The reference period is 1970–2017. (b) Same as panel (a) for the mid-depth (700–2,000 m) OHC, except for the Levitus et al. (2012) data set which reports a 5-years running mean and for the reference period (1970–2015). (c) Same as panel (b) for the 0–2,000 m OHC. Panel (d) represents the time evolution of the 0–700 m minus the 700–2,000 m OHC (panel a minus panel b).

IPSL-EHS mean warms up more than the observations, with a larger warming trend in the North Atlantic and North Pacific Oceans. The ensemble mean also differs from observations in the Eastern Pacific Ocean, probably linked to the large Interdecadal Pacific Variability (England et al., 2014). Despite these larger than observed variations in the model, some members exhibit patterns relatively close to the observations, with, for example, a pattern resembling the warming hole in the North Atlantic Ocean (e.g., members #5 or #14).

Figure 7a shows the time evolution of the global OHC of the first 700 m of the ocean relative to the 1970–2018 period. Despite some discrepancies between the two observationally based estimates used, such as the weaker trend around 2006 or the larger OHC in 1980 in Levitus et al. (2012), the two data sets are broadly consistent with a correlation coefficient of 0.99 for annual values. These differences are mainly due to the different methods used to overcome the insufficient and irregular data coverage.

The observations of the global 0–700 m OHC are broadly in the range of the IPSL-EHS. Some discrepancies, however, are present between the IPSL-EHS and the Levitus et al. (2012) data set, with an underestimation of the OHC in 1980 and an overestimation of the OHC from 2010 on. Consistently with the GSAT analysis, the ensemble mean departs from the observations at the end of the period, with a larger than observed upper ocean warming. This deviation starts around 2008 for the Levitus et al. (2012) data set or around 2014 for the Cheng and Zhu (2016) data set.

Unlike the upper OHC, the IPSL model ensemble mean has a lower than observed warming of the mid-depth (700–2,000 m) OHC at the end of the period, which occurs around 2002 for the Cheng and Zhu (2016)

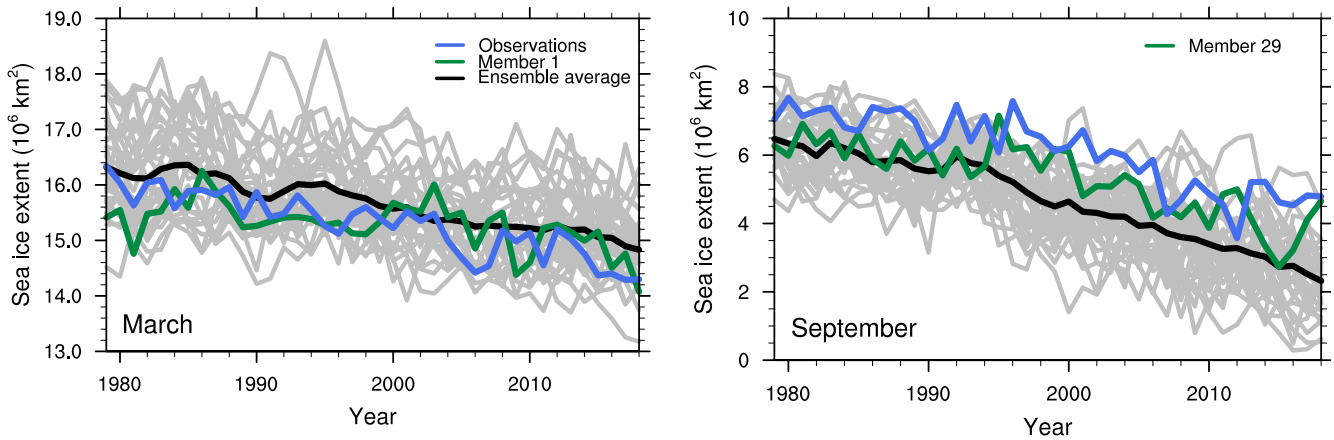

Figure 8. Time evolution of the sea ice extent (10^6 km^2, top panels) in the Northern Hemisphere in March (left) and September (right) from the 32 *historical* members (gray lines) and the ensemble mean (black line) and from the (Fetterer et al., 2017) data set (blue line), with the closest simulation to the observations based on the RMSE (green line).

data set and around 2012 for the Levitus et al. (2012) data set (Figure 7b). The IPSL-CM6A-LR model has an overestimated stratification of the ocean (Boucher et al., 2020), which could explain a part of the overestimation of the upper OHC warming trend and the underestimation of the mid-depth OHC warming trend by limiting the propagation of heat from the surface to the deeper ocean. The origin of this excessive stratification has not yet been determined. Further investigation is required to determine if it is the consequence of the tuning procedure of the IPSL-CM6A-LR model or a more robust feature of the IPSL model. In order to further quantify this potential model bias, we calculated the difference between the upper OHC and the mid-depth OHC (Figure 7d). The observations are clearly outside the range from the IPSL-EHS from 2008 onwards, which confirms that the model transfers too little heat downwards.

Considering the entire 0–2,000 m depth column of water, these discrepancies in surface and mid-depth OHC partly compensate for each other. Both sets of observations are within the range of the IPSL-EHS when taking into account the range of uncertainties (Figure 7c). Therefore, the right amount of energy seems to be captured by the ocean in the model.

5. Simulation of the Sea Ice

We now turn to the assessment of the simulated sea ice, looking at the months of March and September, which correspond to the maximum and the minimum of sea ice in the Arctic region, respectively. The evaluation is focused on sea ice extent in the Northern Hemisphere, as the sensitivity of the Arctic sea ice to global warming is greater than in the Antarctic (Cavalieri et al., 1997). The sea ice extent is defined as the total area enclosed within the 15% sea ice fraction. We choose the sea ice extent variable rather than the sea ice area because of the smaller observational uncertainty, especially during the melting season (Andersen et al., 2007).

In the Northern Hemisphere, the variability of the observed sea ice extent is in the range of the model variability (Figure 8). Large differences in the intensity of interannual variability are present between the different historical members, especially in March. Some members have a good representation of the sea ice variability over the 1979–2019 period, like member #2, which has the lowest RMSE in sea ice extent in March.

The *historical* simulations show nevertheless a large diversity in trends of sea ice extent averaged over the Northern Hemisphere for the 1979–2019 period both in March and in September (Figure 9). The observed trends are within the range of the model, but in the tail of the distribution for the sea ice extent in summer (September), the majority of members having stronger negative trends than observed. In winter (March), the observed trend is similar to that of the ensemble mean. Except for members #22 and 26, all the historical simulations show negative trends of sea ice extent in March. The slight positive trend in sea ice for these members in March is consistent with the lower than an observed trend in surface temperature in this region

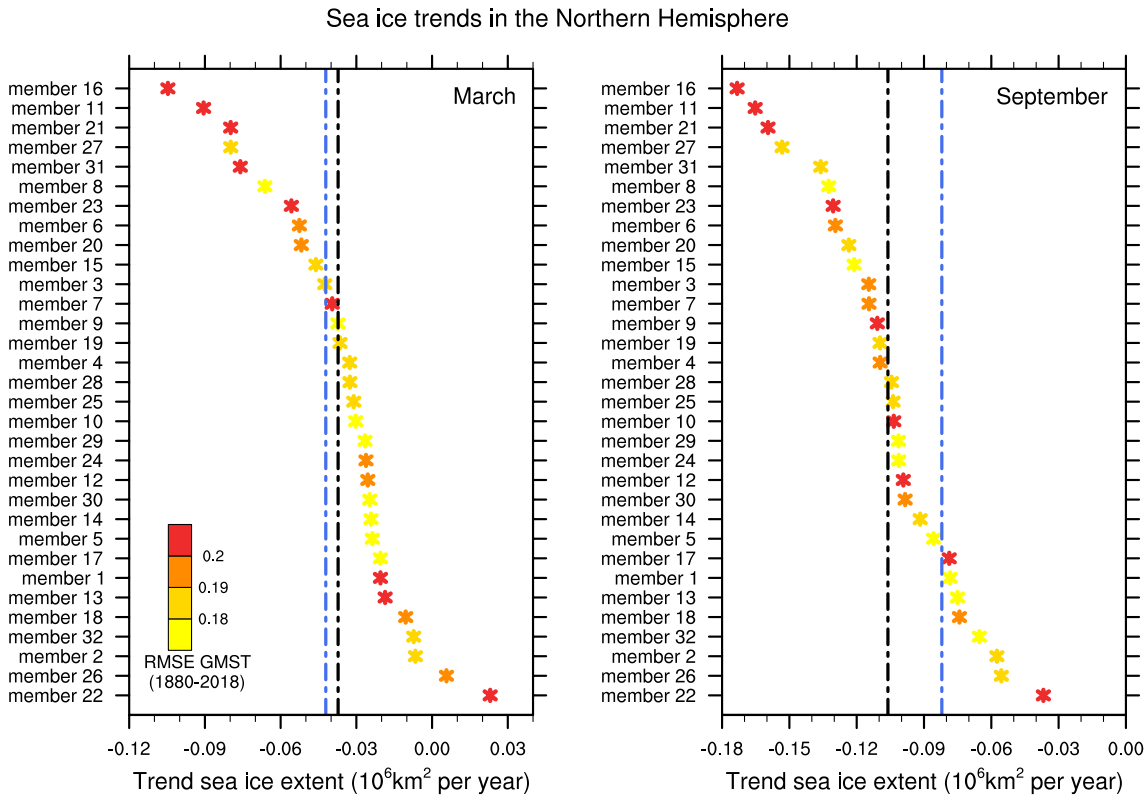

Figure 9. Trends (10^6 km^2 yr^{-1}) in annual sea ice extent over the Northern Hemisphere in March (left) and September (right) calculated over the 1979–2018 period from the (Fetterer et al., 2017) observational data set (blue dashed line), the IPSL-EHS mean (black dashed line) and the 32 historical members of the IPSL-CM6A-LR model (stars) ranked from weakest to highest trend (top to bottom). The stars colors are relative to the RMSE of the GSAT (1880–2018).

(Figure 5). In September, all the members show a reduction in sea ice extent since 1978. This consistency between the members highlights the influence of external forces on sea ice melting. Members #16, 11, 21 and 27 have the strongest decline in sea ice area, up to 0.95 10^6 km^{2}yr^{-1} in March and 0.15 10^6 km^{2}yr^{-1} in September for member #16.

Trends in sea ice extent in most CMIP5 models were lower than observed in September and this was the case of the majority of members of the IPSL-CM5A-LR ensemble (Stroeve et al., 2012). On the opposite, the sea ice trends in the IPSL-EHS tend to be larger than observed. This is consistent with the results of Notz et al. (2020), where the IPSL-CM6A-LR model shows a larger sensitivity of sea ice area to GSAT and CO_2 concentration changes in comparison to IPSL-CM5A-LR, in better agreement with observations. It is nevertheless still unclear to what degree the better performance of CMIP6 models in comparison to CMIP5 in capturing the observed sensitivity of sea ice area to change in GSAT and CO_2 concentration is related to a change in the forcing, a stronger climate sensitivity or an improvement of the model physics (Notz et al., 2020).

6. Simulation of the Atlantic Meridional Overturning Circulation

The AMOC plays an important role in regulating the Earth's climate system, as it participates in redistributing heat within the ocean across latitudes and across the two hemispheres.

In the IPSL-CM6A-LR model, the AMOC maximum intensity at 26°N is underestimated by about 25% (Boucher et al., 2020), which is a common bias in low-resolution climate models (Hirschi et al., 2020), especially for those without overflow parametrization (Danabasoglu et al., 2014). The meridional heat transport of the Atlantic Ocean is also underestimated in IPSL-CM6A-LR at all latitudes, and particularly at tropical latitudes (Boucher et al., 2020). This bias is nevertheless weaker than in previous model versions.

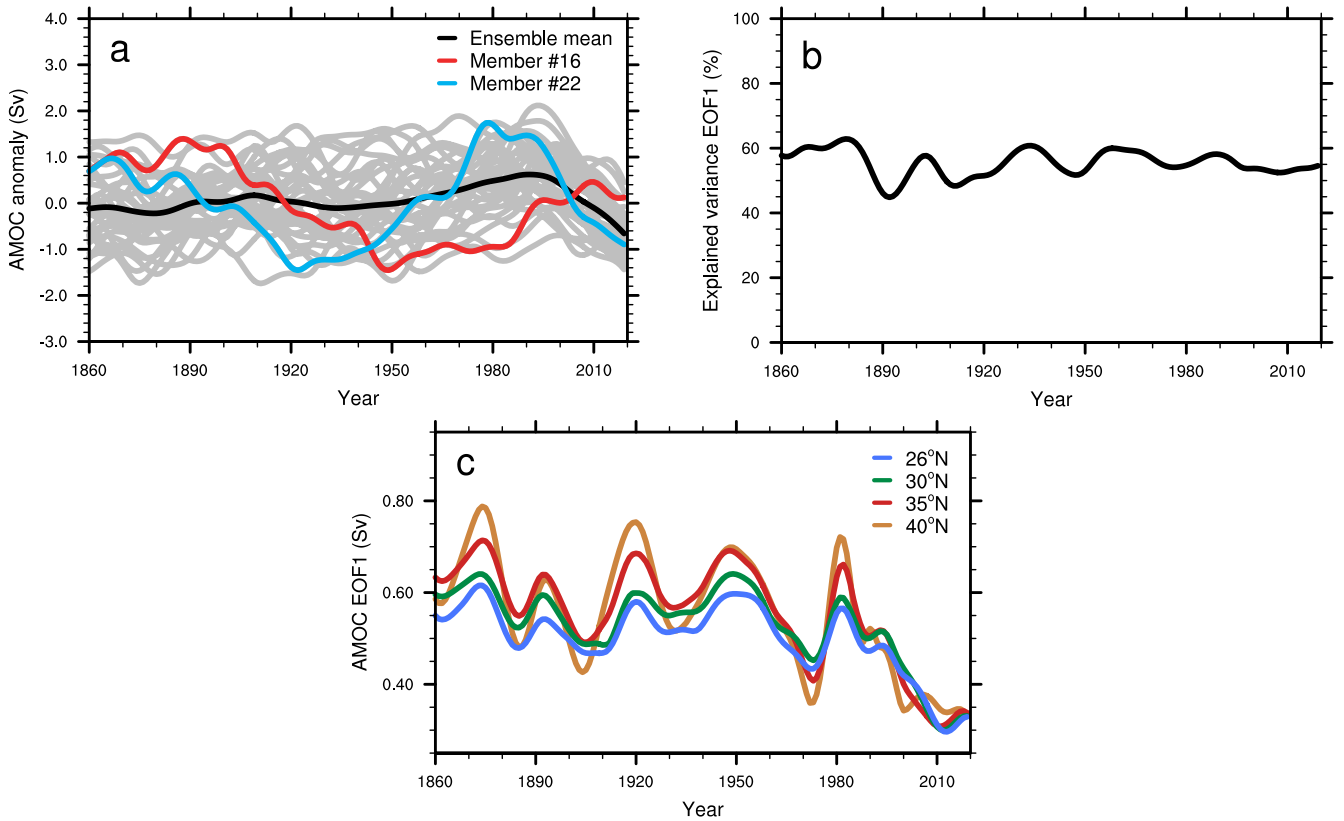

Figure 10. (a) Low-pass filtered annual anomaly of the AMOC at 26°N (in Sv) relative to the 1880–2018 period from the 32 *historical* members (in gray), the model ensemble average (in black), member #16 (in red), member #22 (in blue). A Lanczos low-pass filter is used with a cutoff period of 11 years (b) Variance of first intra-member AMOC at 26°N empirical orthogonal function (EOF-PC1) (%) calculated from the IPSL-EHS. (c) Intensity of intra-member AMOC EOF corresponding to one standard deviation of the PC1 (in Sv) at (green) 26°N, (red) 30°N, (black) 35°N and (blue) 40°N, calculated from the IPSL-EHS.

Unlike the previous CMIP3 and CMIP5 simulations, the IPSL-EHS exhibits strong multidecadal variations of the AMOC over the 1880–2018 period (Figure 10a). The IPSL-EHS mean shows a slight increase of the AMOC between 1940 and 1990, which was related in CMIP6 models to the large aerosol emissions during that period, although the associated mechanism remains largely unknown (Menary et al., 2020). The ensemble mean then shows a weakening of the AMOC since the 1990s that continues in future scenario simulations (not shown), which is consistent with CMIP3 (Gregory et al., 2005), CMIP5 (Cheng et al., 2013), and CMIP6 (Weijer et al., 2020) results. Member #22, characterized by the smallest warming trend in the IPSL-EHS over the 1978–2018 period for land and ocean (Figure 4) as well as with little reduction in sea ice (Figure 9), has one of the largest negative trends of the AMOC since the 1980s (Figure 10b). Conversely, member #16, which is characterized by a strong warming since the end of the 20th century (Figure 4), shows a positive trend in the AMOC since the 1980s. This link between the AMOC and the GSAT variability results in a strong positive relationship between the 1978–2018 GSAT and AMOC trends, with a $r^2 = 0.59$ (not shown). The ensemble members with the strongest decrease of the AMOC are those with the weakest GSAT warming and *vice versa*. The modulation of the GSAT and AMOC by the dominant bicentennial variability mode previously illustrated in the preindustrial control run is consistent with these relationships.

As suggested by the previous analyses, a decrease in this variability is present over the historical period (Section 2.3). To better describe it, the first component empirical orthogonal function (EOF) of the AMOC is calculated among the members of the IPSL ensemble (Figures 11b and 11c). This is done using low-pass filtered yearly Atlantic meridional streamfunction field, using a fifth order Butterworth filter with 20-years as cutoff period. The standardized principal components (PCs) are then calculated from the covariance matrix using the ensemble dimension instead of the time dimension. The EOFs are defined as the regressions onto the Principal Components (PCs). Fairly consistently over the historical

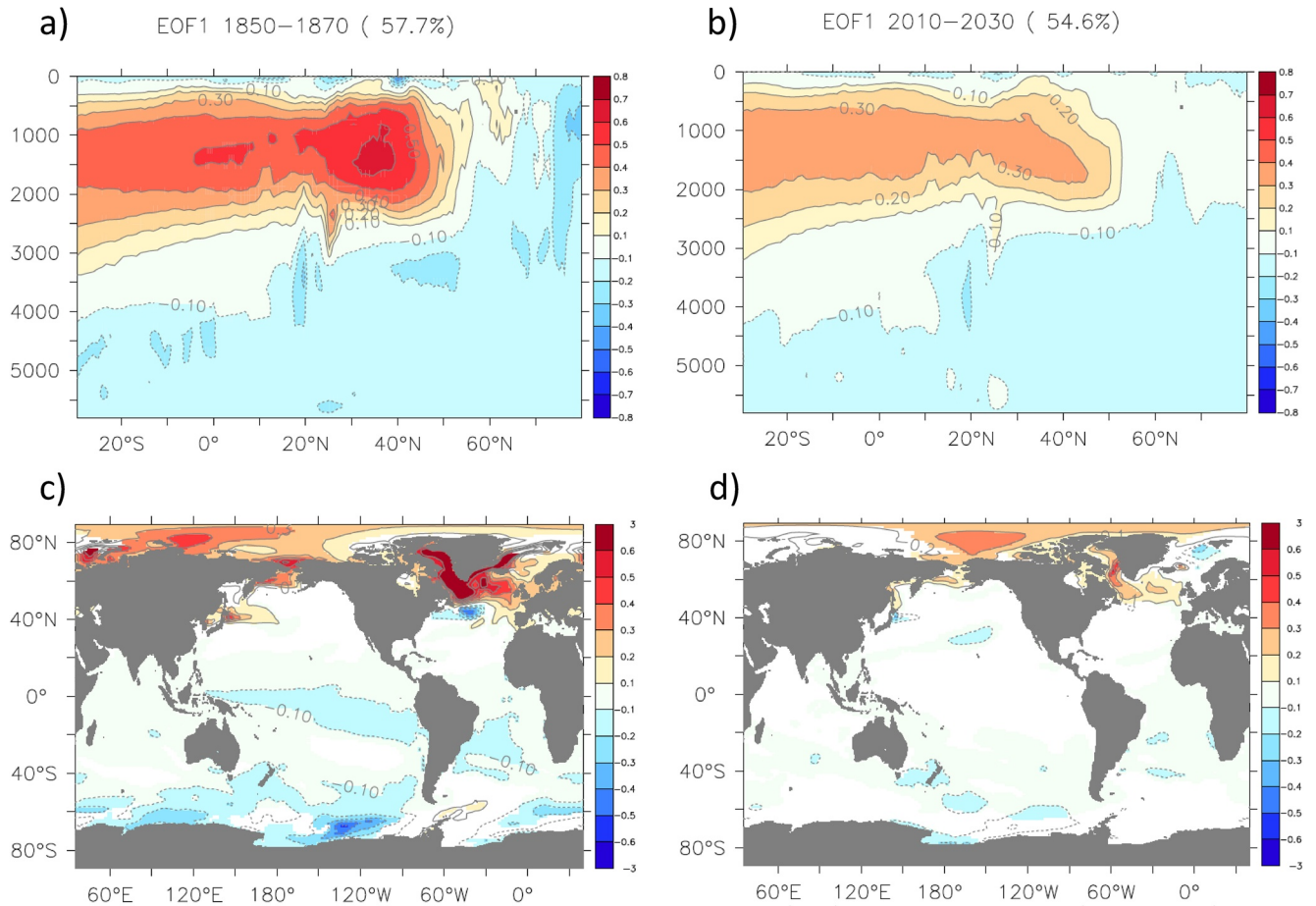

Figure 11. (a) First empirical orthogonal function (EOF) of the AMOC (Sv per standard deviation) calculated over the 1850–1870 period among the members of the IPSL ensemble. (b) Same as panel (a) for the 2010–2030 period. (c) Regression of the SST on the AMOC EOF (K per standard deviation) calculated over the 1850–1870 period among the members of the IPSL ensemble. (d) Same as panel (c) for the 2010–2030 period. In panels (c) and (d), the colors are shown if the significance level is below 10%.

period, about 50% of the intra-members AMOC variance is explained by the EOF-PC1 (Figure 11b). However, a decrease of the EOF intensity is visible from a mean value of 0.55 Sv before the 1980s to 0.35 Sv after the 2000s, regardless of the latitude (Figure 11c). Therefore, the decrease of the AMOC variability seems mainly due to a decrease of the internally driven AMOC variability, probably due related to the interaction with the external forcing, rather than an externally forced phasing between members.

This decrease of the intra-member AMOC variability for the 2010–2030 period calculated from the EOF in 2020, in comparison to the 1850–1870 period, calculated from the EOF in 1860 is very visible between 500 and 2,000 m depth and at all latitudes, although the main pattern remains similar (Figures 11a and 11b). In order to investigate the influence on SST of the decrease in the AMOC variability, we calculated the regression of the low-pass filtered temperature (same filter as the AMOC) onto the AMOC-PC1 through the ensemble dimension. A large decrease of the SST variability related to the AMOC variability is detected, especially in the North Atlantic (Figures 11c and 11d). While the AMOC impacts in 1850–1870 are similar to that found in the pre-industrial control simulation (see Jiang et al., 2021, their Figure 3), with a large increase of the Arctic SST associated with sea ice loss. The AMOC impacts are largely reduced in 2010–2030. Such reduction remains to be understood and could be linked to the Arctic sea-ice reduction after the 2000s, or to the increase of oceanic stratification as the North Atlantic subpolar gyre warms.

7. Modes of Variability

This section has two main objectives, the first is to assess whether the IPSL-EHS has a good representation of internal decadal climate variability over the historical period and the second is to look at whether the members are consistent with the observed evolution of the global temperature are those that are most consistent with two of the main modes of decadal climate variability, located over the Atlantic and Pacific Oceans.

7.1. Interdecadal Pacific Variability

The Interdecadal Pacific Variability (IPV) is characterized by a horseshoe shape over the North Pacific, with for a positive phase, an eastern warming and a cooling in the Kuroshio-Oyashio Extension, similar to the Pacific Decadal Variability pattern, a warming over the tropical Pacific region, and a cooling over the southwestern Pacific Ocean (Newman et al., 2016). To evaluate the IPV in the IPSL-EHS, we used the tripole index (TPI) developed by Henley et al. (2015). The TPI is defined as the weighted difference between averaged monthly SST anomalies (with the seasonal cycle removed) (SSTA) over the central equatorial Pacific ($SSTA_2$, 10°S–10°N, 170°E–90°W), and the Northwest ($SSTA_1$, 25°N–45°N, 140°E–145°W) and the Southwest Pacific ($SSTA_3$, 50°S–15°S, 150°E–160°W):

$$TPI = SSTA_2 - (SSTA_1 + SSTA_3)/2 \quad (2)$$

A Lanczos low-pass filter with a cutoff period of 13 years is then used. The interest of this index is to be expressed in Kelvin unit, which makes it easier to interpret than EOF-based definitions.

To evaluate the simulated IPV pattern, the TPI index is regressed on SST in each grid point for the observations (Huang et al., 2017) and for each member of the IPSL-EHS. The observations are regridded onto the IPSL-CM6A-LR grid prior to the evaluation. The results are summarized in Figure 12 using a Taylor diagram (Taylor, 2001) and a panel of the IPV pattern maps is in (Figure S1). Overall, the IPSL-EHS has a good spatial representation of the IPV, with pattern correlations ranging from 0.54 to 0.78 (Figure 12). Although some members have pattern variability close to the observations, the model tends to underestimate it. The IPV pattern from the *piControl* simulation is in line with the IPSL-EHS, suggesting that the externally forced component is rather small in shaping the pattern. The representation of the IPV pattern seems to be better represented in IPSL-CM6A-LR than in IPSL-CM5A-LR, the CMIP5 version of the IPSL model, with a higher spatial correlation in the *piControl* simulation and in most of the historical simulations, although the spatial pattern variability of the *piControl* run seems slightly better in CMIP5 version than in CMIP6 one.

In terms of temporal variations, the observed TPI is overall within the range simulated by the IPSL-EHS over the historical period, except around the 1900–1910 period (Figure 13). The observed temporal variability of the TPI has a standard deviation of 0.26 K, which is within the range of the IPSL-EHS whose standard deviations range from 0.15 to 0.29 K. When considering the previous version of the model (IPSL-CM5A-LR), the standard deviations range from 0.20 to 0.29 K over the 1880–2005 period. Therefore, both models seem consistent with the observed variability in terms of temporal standard deviations of the TPI index. No clear relationship seems to exist between the TPI and the GSAT in the model. Member #16, with the lowest RMSE (0.25 K) over the 1880–2018 period and the best correlation (0.4), has a poor representation of the GSAT over the 1880–2018 period (Figure 3b).

7.2. Atlantic Multidecadal Variability

The spatial pattern of the AMV is characterized by SST anomalies of the same sign over the North Atlantic, with a maximum on the subpolar gyre and a second maximum off the Iberian Peninsula and in the tropical Atlantic. The AMV index is generally defined as the average SST over the North Atlantic after the removal of the forced signal. A low-pass filter is then used to retain only the low-frequency variations.

We use the Trenberth and Shea (2006) method, which estimates the effect of external forcings as the averaged SST between 60°S and 60°N. We also use the ensemble mean to estimate the forced variability in the model. A comparison between these two methods is then made to see the impact of the method used.

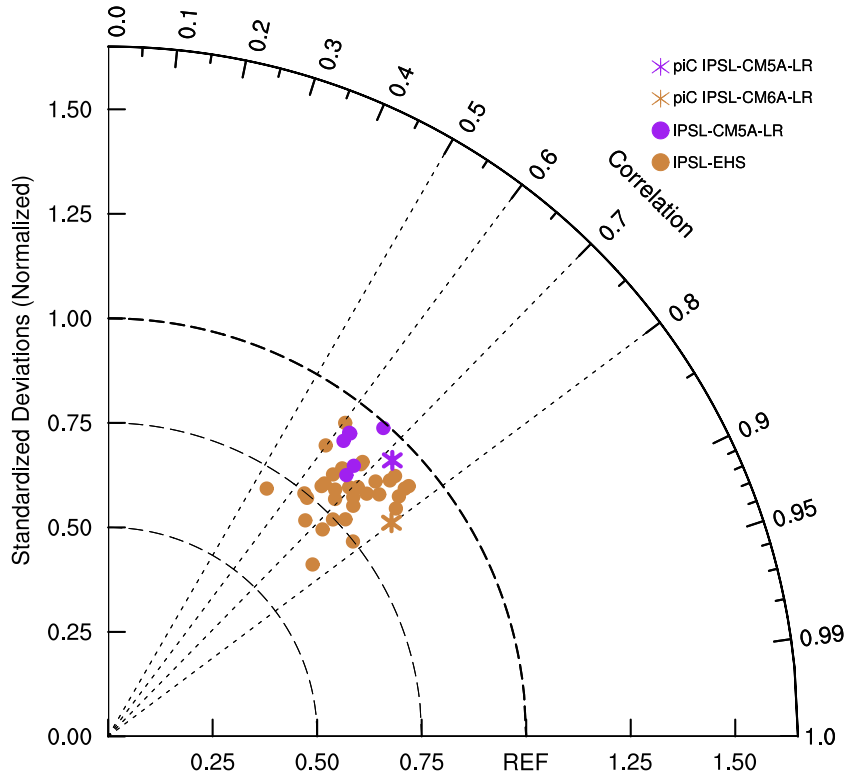

Figure 12. Taylor diagram of the TPI index regressed on SST over the 1880–2018 period, with the observed IPV pattern from the ERSSTv5 data set of SST (Huang et al., 2017) as reference (REF on the *X* axis). The ratio of standard deviation (radial distance), spatial correlation (radial angle) is indicated. The brown dots are the results from the IPSL-EHS, the brown star from the *piControl* simulation of IPSL-CM6A-LR, the purple dots from the CMIP5 historical simulations performed with IPSL-CM5A-LR (limited to the 1880–2005 period), and the purple star from the CMIP5 *piControl* simulation of IPSL-CM5A-LR.

To evaluate the simulated AMV pattern, the index is regressed on the unforced annual SST over the 1880–2018 period for the ERSSTv5 (Huang et al., 2017) data set and for each member of the IPSL-EHS. The observations are regridded onto the IPSL-CM6A-LR grid prior to the evaluation. The spatial correlations between the AMV pattern simulated in the IPSL-EHS and the observations range from 0.5 to 0.7 (Figure 14). The IPSL-CM6A-LR model underestimates the spatial pattern variability of the AMV, by around 25% for a large number of members of the IPSL-EHS. Overall, the positive phase of the AMV is characterized by a positive response over the Northern Hemisphere, and negative response over the Southern Hemisphere (Figure S2). The IPSL-EHS exhibits a strong subpolar response to the AMV, larger than observed, as well as in the tropical Atlantic. The model also simulates a link between the AMV and the North Pacific, with a pattern resembling a negative phase of the Pacific Decadal Oscillation associated with a positive AMV phase as noted by Jiang et al. (2021) in the *piControl* run. As for the IPV, the AMV pattern seems to be better represented in IPSL-CM6A-LR than in IPSL-CM5A-LR (higher correlations in Figure 14 as well as standard deviations closer to observations). The AMV pattern of the *piControl* simulation of the CMIP5 IPSL model has, in particular, a much lower spatial correlation than IPSL-CM6A-LR. Although there are fewer biases in the spatial variability of the historical simulations from the IPSL CMIP5 model, the spatial correlations are overall lower than in IPSL-CM6A-LR.

A large uncertainty is associated with the way to remove the effect of external forces. Using the IPSL-EHS mean to remove response from the external forcing in the observations, a pattern correlation of 0.79 is found with the observed AMV pattern based on the Trenberth and Shea (2006) method (Figure 14). For the IPSL-EHS, an underestimation of the spatial magnitude is present with the Trenberth and Shea (2006) method, which is no longer visible when using the ensemble mean. The AMV pattern from the *piControl* simulation is relatively close to the observations, with a pattern correlation of 0.68. The pattern magnitude

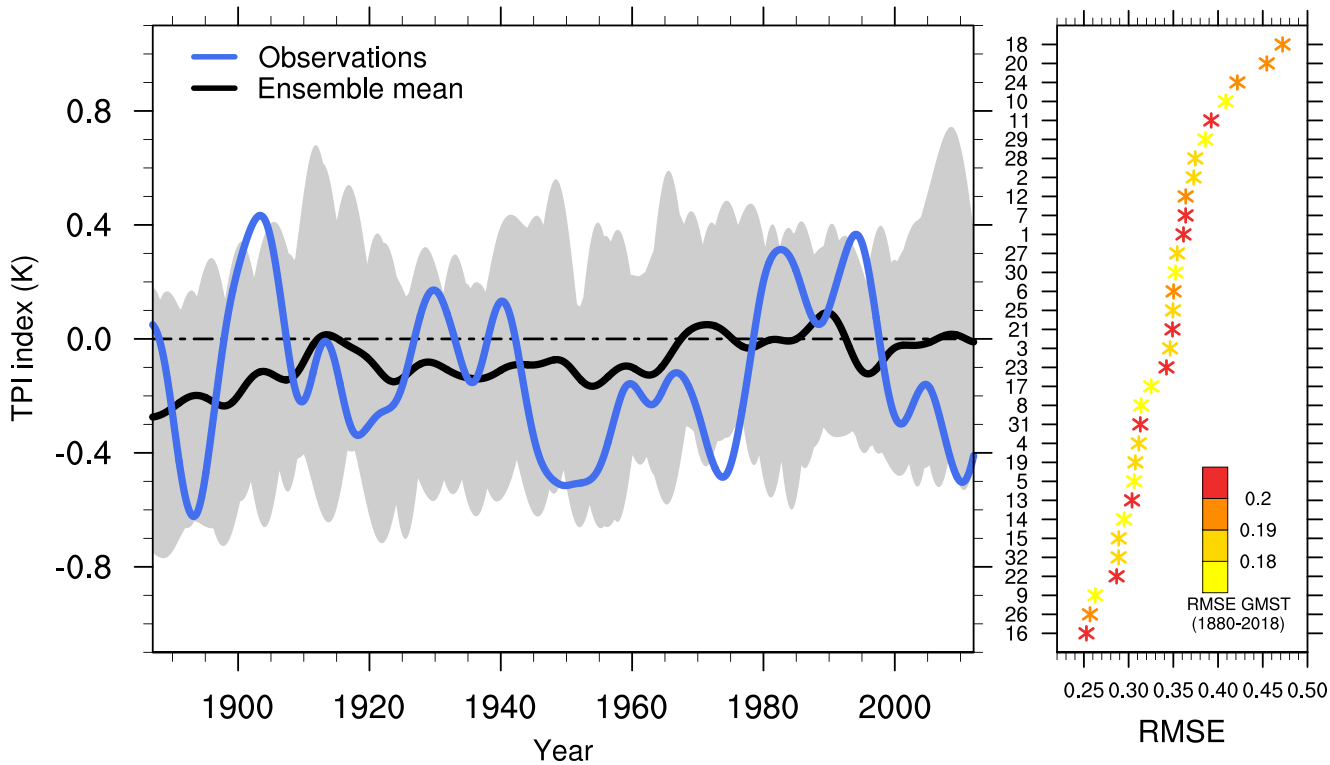

Figure 13. (Left) Time evolution of the TPI index calculated from the ERSSTv5 sea surface temperature (SST) observational data set (Huang et al., 2017, in blue) and from the 32 *historical* members (spread in gray). (Right) RMSE calculated for each member ranked from the highest to the lowest (top to bottom), the color of markers indicating the RMSE of the GSAT calculated over the 1880–2018 period.

is slightly larger than in the observations. The pattern magnitude of the *piControl* simulation is much larger than that from the IPSL-EHS using the Trenberth and Shea (2006) method.

The time variations of the observed AMV index are broadly in the range of the IPSL-EHS when using the Trenberth and Shea (2006) method to remove external forcing signal, albeit with an underestimation by the model ensemble around 1935 and an overestimation around 1998 (Figure 15). The temporal standard deviation of the observed AMV index is 0.13 K, inside the range of the IPSL-EHS, which goes from 0.07 to 0.2 K. As for the IPV, the observed temporal standard deviation of the AMV index is also within the range of the historical simulations of the CMIP5 version of the IPSL model. The negative anomalies of the ensemble mean during the first part of the 20th century, as well as the ensemble mean increase since 1990, suggests nevertheless that the Trenberth and Shea (2006) method does not fully estimate the external forcings in the IPSL-EHS, consistent with the previous study (Qasmi et al., 2017). Using the ensemble mean to estimate the external forcing, the observations remain included in the range of the IPSL-EHS (Figure 15).

As for the TPI index, the RMSE scores are calculated between simulated and observed AMV index (Figure 15, right panel). In contrast to the TPI index, the members with the lowest RMSE scores for the AMV index tend to have a low RMSE score for the GSAT evolution. Therefore, the low-frequency internal variability of the North Atlantic SST seems to have a larger influence on GSAT than those of the Pacific in the IPSL model. This result is consistent with the large bicentennial variability of the AMOC present in this model and the associated oceanic variables in the North Atlantic.

8. Discussion and Conclusion

A large ensemble of *historical* simulations has been performed with the IPSL-CM6A-LR model (Boucher et al., 2020), sampling start dates every 20 years from a preindustrial control simulation. This method allowed to sample adequately the dominant mode of the model internal climate variability of the GSAT and to evaluate the uncertainty related to the IPSL-EHS. This bicentennial variability present in the IPSL-CM6A-LR

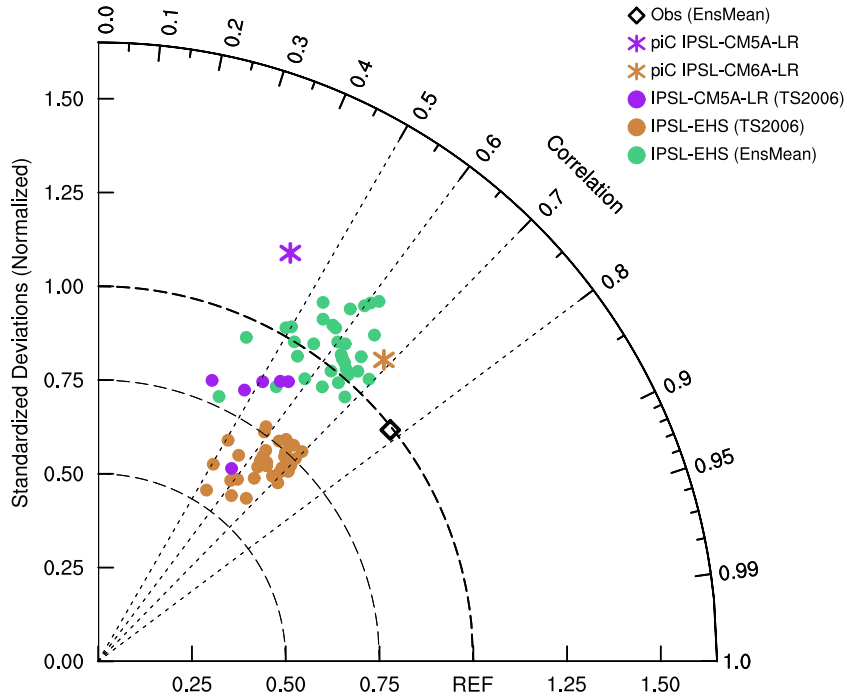

Figure 14. Taylor diagram of the AMV index regressed on SST over the 1880–2018 period, with the observed AMV pattern from the ERSSTv5 data set of SST (Huang et al., 2017) as reference (REF on the *X* axis). The ratio of standard deviation (radial distance), spatial correlation (radial angle) is indicated. The dots are the results from the IPSL-EHS using the Trenberth and Shea (2006) method to estimate the external forcings, the green dots using the ensemble mean and the purple dots are from the CMIP5 historical simulations using the IPSL-CM5A-LR model (limited to the 1880–2005 period). The stars indicate the results from the (brown) CMIP6 and (purple) CMIP5 *piControl* simulations. The black rhombus is the result of the observed AMV pattern regressed on the unforced SST using the IPSL-EHS ensemble mean. The external forcings are removed from the observed SST prior to the analysis using the −60 to 60°N average or the IPSL-EHS ensemble mean relative to the method used for the calculation of the AMV index.

model indeed modulates the temperature changes over the historical period. We show that this variability is damped in the model along the historical periods, probably due to the effect of external forcing on this mode (e.g., Ma et al., 2021). If it had remained unchanged, this variability would have modulated the temperature changes over the historical period by about 0.4 K per century. This very long memory of initial conditions in the IPSL model, conditioning the evolution of temperature over the historical period, could have important implications in terms of predictability and evolution of the AMOC.

As a result of this large low-frequency internal variability, a strong link is found between the AMOC internal variability and the climate variability of the Northern Hemisphere. Members whose AMOC internal variability has a strong strengthening over the period 1978–2018 tend to have greater warming and stronger ice melt in the Northern Hemisphere than members whose AMOC internal variability has a slowing down, which will have much more limited warming. This slowdown tends to be amplified by the forced response of the AMOC to temperature increase. As a consequence of this strong AMOC variability, the variability of the North Atlantic seems to modulate GSAT multidecadal trends in the model, whereas the Interdecadal Pacific Variability does not seem to have a strong influence. Further investigation is needed to better understand the processes involved in this low-frequency internal variability, which is also present in other CMIP6 models (Parsons et al., 2020), whether it is an artifact of the model, and how it is modulated by external forcings.

Members #1 and #32 start from the same initial condition but were performed using different supercomputers. Due to slight differences in the precision of the supercomputers and due to the chaotic nature of the climate system (e.g., Lorenz, 1963), their exact trajectories have important differences on multidecadal and shorter timescales, consistent with the study of Guarino et al. (2020). However, the GSAT and AMOC multicentennial variability of these two members are close, thanks to the memory of the multicentennial

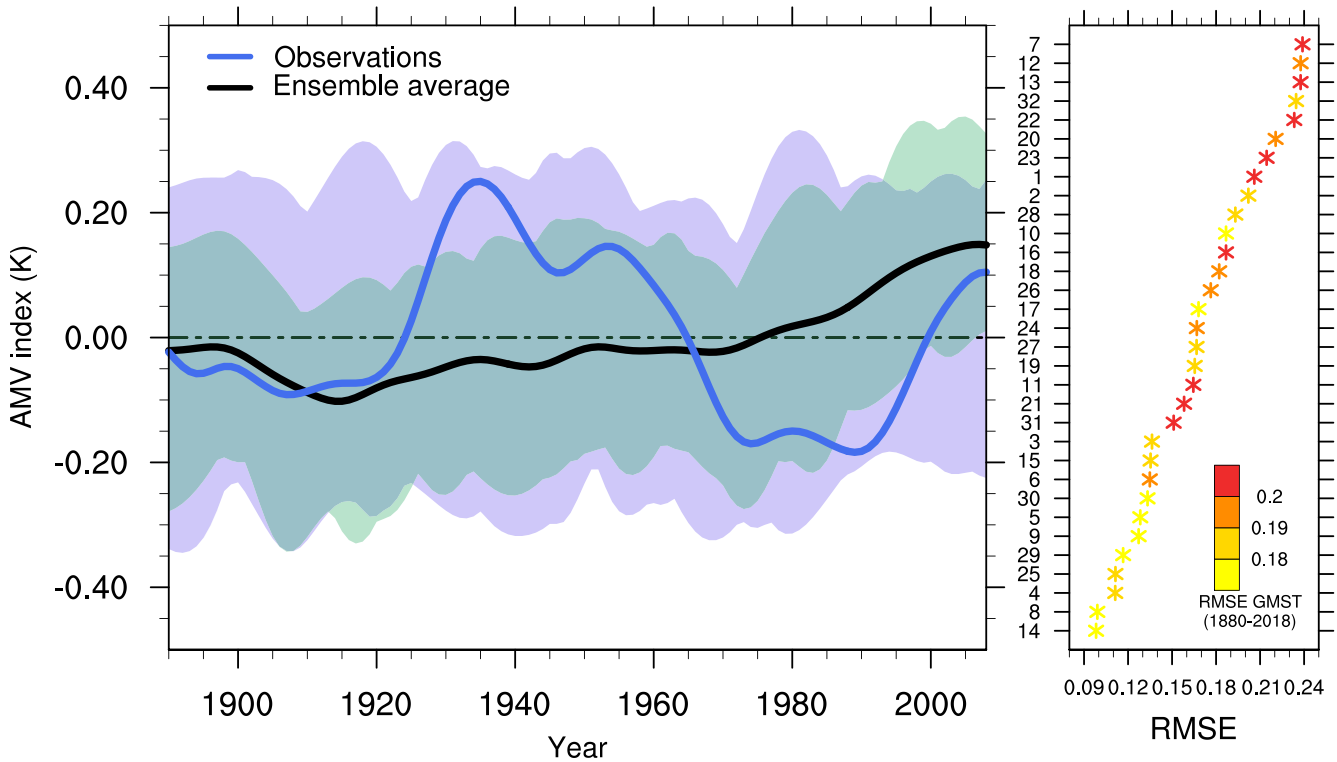

Figure 15. Time evolution of the Atlantic multidecadal variability (AMV) index, calculated from the ERSSTv5 sea surface temperature (SST) observational data set (Huang et al., 2017, in blue) and from the 32 *historical* members using the Trenberth and Shea (2006) method to remove the response from the external forcings (spread in green) with the ensemble mean (in black). The spread in purple represents the same as the spread in green but using the ensemble mean as an estimation of the external forcing. A 10-years window low-pass filter is applied to the AMV index. (Right) RMSE calculated for each member ranked from the highest to the lowest (top to bottom), the color of markers indicating the RMSE of the GSAT calculated over the 1880–2018 period.

variability in IPSL-CM6A-LR. As a result, the trends of these variables over the historical period are very close in these two members (Figure 1c for the GSAT).

The IPSL-CM6A-LR model is characterized by a rather large effective ECS of 4.5 K (though it reduces to 3.8 K if computed from a $2 \times CO_2$ rather than a $4 \times CO_2$ abrupt forcing, as discussed in Boucher et al., 2020), a TCR of 2.4 K, and a relatively low aerosol forcing of -0.6 W m^2, in comparison to the others CMIP6 models (Tokarska et al., 2020), as well as to the estimation based on observations (Sherwood et al., 2020). In relation to this rather large climate sensitivity, the ensemble mean of the *historical* simulations warms more than the observations. Despite the relative greater than observed warming in the IPSL-EHS means, some members are consistent with the global surface temperature variability and the recent warming trends, probably due to the large influence of decadal to multicentennial internal climate variability in IPSL-CM6A-LR. Consistent with this average warming bias, the IPSL-EHS mean shows a stronger than observed decrease for sea ice extent in September. As the quasi-bicentennial variability is linked to interactions between the Arctic and the Atlantic Ocean, the Arctic sea ice also shows a large internal variability so that the observed decrease remains within the range of possibilities of the IPSL-EHS. The ensemble mean of the trend in sea ice extent in March in the IPSL-EHS is close to observations. We show that the upper OHC (above 700 m) shows a larger than observed increase, while the OHC at depth (below 700 m) shows a weaker than observed increase. We relate this bias to the fact that the model is generally too stratified in the ocean, which limits heat transfer from the surface ocean to the mid and deep ocean, which might then play a role and possibly explain the larger than observed warming found in the ensemble mean. This comparison of the modeled OHC to observations points to a potential positive bias for the transient climate response (TCR) but not for the equilibrium climate sensitivity (ECS).

A possible interesting perspective from this ensemble would be to identify the member or a subset of members with a better fit to the observations is, for example, to apply a regional downscaling. These regional

experiments are often bias corrected against a short observational record (e.g., Cannon et al., 2015). It could be interesting to evaluate if performing such a bias correction over a simulated period that is, inconsistent with the observed period in terms of multidecadal variability results in under or over correction of the biases when applied to future projected climate conditions. A perfect model experiment, using opposite phase members of the AMV for example, could be an opportunity to test this. The development of large ensemble simulations such as the IPSL-EHS offers great opportunities in the context of quantifying climate impacts.

This article provides a description of this large ensemble with a range of evaluation metrics as well as the first overview of its capabilities. It is important, however, to keep in mind that the best member(s) to be used may depend on the region and/or climate property of interest, and, therefore, have to be evaluated according to the study of interest. Additionally, the consistency of some members with some climate features (in this study, the GSAT evolution for example), does not mean that the forcings and feedbacks are correct.

Some other precautions must be taken when using IPSL-EHS, as the memory of the initial conditions is indeed relatively persistent along the whole historical period because of the strong internal variability at multicentennial timescales. This could have implications, when considering a limited number of members for specific studies and in particular over regions largely influenced by this low-frequency internal variability, such as Europe for example, Indeed, the analysis of two members in an opposite phase in terms of this low-frequency climate variability as compared to the observations could lead to biased results. Nevertheless, this marked variability can also be seen as an opportunity to study conditional occurrences of extremes, for example, depending on the specific phase of internal large-scale modes of variability, through an ad hoc sampling of the members. To conclude, the use of this large ensemble requires some care so as to properly account for the long and strongly marked internal variability in particular for bias correction methods for example, Yet, it offers also a unique opportunity to address important questions linking specific interannual events to decadal to multicentennial internal variability.

Overall, the IPSL-EHS is an interesting tool that can be used to address uncertainties relative to the internal climate variability over the historical period and until the mid-21st century. Additionally, the IPSL-EHS can be used in conjunction with other large ensembles in order to assess the robustness of internal modes of climate variability or of the climate response to external forcings.

Data Availability Statement

The IPSL ensemble of historical simulations are publicly available on the Earth System Grid Federation website (https://esgf-node.ipsl.upmc.fr/projects/cmip6-ipsl/). Extensions for the period 2015–2059 are not official CMIP6 simulations and therefore are only available upon request.

Acknowledgments
The authors are grateful to the developers of the IPSL-CM6A-LR climate model. This work was undertaken in the framework of the IPSL Climate Graduate School EUR. As such it benefited from the French state aid managed by the ANR under the "Investissements d'avenir" programme with the reference ANR-11-IDEX-0004-17-EURE-0006. The CMIP6 project at IPSL used the HPC resources of TGCC under the allocations 2016-A0030107732, 2017-R0040110492 and 2018-R0040110492 (project gencmip6) provided by GENCI (Grand Équipement National de Calcul Intensif). This study benefited from the ESPRI (Ensemble de Services Pour la Recherche à l'IPSL) computing and data center (https://mesocentre.ipsl.fr) which is supported by CNRS, Sorbonne Université, École Polytechnique and CNES and through national and international grants. Support from the European Commission's Horizon 2020 Framework Programme is acknowledged, under Grant Agreement number 820829 for the "Constraining uncertainty of multidecadal climate projections (CONSTRAIN)" project. D. Swingedouw, J. Mignot, and G. Gastineau received support from the Blue-Action (European Union's Horizon 2020 research and innovation program, Grant Number: 727852) and EUCP (European Union's Horizon 2020 research and innovation programme under grant agreement no 776613) projects.

References

Andersen, S., Tonboe, R., Kaleschke, L., Heygster, G., & Pedersen, L. T. (2007). Intercomparison of passive microwave sea ice concentration retrievals over the high-concentration Arctic sea ice. *Journal of Geophysical Research*, *112*(C8), C08004. https://doi.org/10.1029/2006jc003543

Bellomo, K., Murphy, L. N., Cane, M. A., Clement, A. C., & Polvani, L. M. (2018). Historical forcings as main drivers of the Atlantic multidecadal variability in the CESM large ensemble. *Climate Dynamics*, *50*(9–10), 3687–3698. https://doi.org/10.1007/s00382-017-3834-3

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., et al. (2016). The decadal climate prediction project (DCPP) contribution to CMIP6. *Geoscientific Model Development*, *9*(10), 3751–3777. https://doi.org/10.5194/gmd-9-3751-2016

Booth, B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., & Bellouin, N. (2012). Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, *484*(7393), 228–232. https://doi.org/10.1038/nature10946

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., & Vuichard, N. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advanced Modeling Earth Systems*, *12*(7), 2019MS002010. https://doi.org/10.1029/2019MS002010

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *Journal of Climate*, *28*(17), 6938–6959. https://doi.org/10.1175/jcli-d-14-00754.1

Cavalieri, D. J., Gloersen, P., Parkinson, C. L., Comiso, J. C., & Zwally, H. J. (1997). Observed hemispheric asymmetry in global sea ice changes. *Science*, *278*(5340), 1104–1106. https://doi.org/10.1126/science.278.5340.1104

Chen, X., & Tung, K.-K. (2014). Varying planetary heat sink led to global-warming slowdown and acceleration. *Science*, *345*(6199), 897–903. https://doi.org/10.1126/science.1254937

Chen, X., & Tung, K.-K. (2016). Variations in ocean heat uptake during the surface warming hiatus. *Nature Communications*, *7*(1), 1–3. https://doi.org/10.1038/ncomms12541

Cheng, L., Trenberth, K. E., Fasullo, J., Boyer, T., Abraham, J., & Zhu, J. (2017). Improved estimates of ocean heat content from 1960 to 2015. *Science Advances*, *3*(3), e1601545. https://doi.org/10.1126/sciadv.1601545

Cheng, L., & Zhu, J. (2016). Benefits of CMIP5 multimodel ensemble in reconstructing historical ocean subsurface temperature variations. *Journal of Climate*, *29*(15), 5393–5416. https://doi.org/10.1175/jcli-d-15-0730.1

Cheng, W., Chiang, J. C., & Zhang, D. (2013). Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations. *Journal of Climate*, *26*(18), 7187–7197. https://doi.org/10.1175/jcli-d-12-00496.1

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, E., Gastineau, G., & Zhao, Y. (2019). Improved near surface continental climate in IPSL-CM6 by combined evolutions of atmospheric and land surface physics. *Journal of Advanced Modeling Earth Systems*, *12*(10)2019MS002005. https://doi.org/10.1029/2019MS002005

Clement, A., Bellomo, K., Murphy, L. N., Cane, M. A., Mauritsen, T., Rädel, G., & Stevens, B. (2015). The Atlantic multidecadal oscillation without a role for ocean circulation. *Science*, *350*(6258), 320–324. https://doi.org/10.1126/science.aab3980

Cowtan, K., Hausfather, Z., Hawkins, E., Jacobs, P., Mann, M. E., Miller, S. K., et al. (2015). Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophysical Research Letters*, *42*(15), 6526–6534. https://doi.org/10.1002/2015gl064888

Cowtan, K., & Way, R. G. (2014). Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society*, *140*(683), 1935–1944. https://doi.org/10.1002/qj.2297

Dai, A., & Bloecker, C. E. (2019). Impacts of internal variability on temperature and precipitation trends in large ensemble simulations by two climate models. *Climate Dynamics*, *52*(1–2), 289–306. https://doi.org/10.1007/s00382-018-4132-4

Dai, A., Fyfe, J. C., Xie, S.-P., & Dai, X. (2015). Decadal modulation of global surface temperature by internal climate variability. *Nature Climate Change*, *5*(6), 555–559. https://doi.org/10.1038/nclimate2605

Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., et al. (2014). North Atlantic simulations in coordinated ocean-ice reference experiments phase II (CORE-II). Part I: Mean states. *Ocean Modelling*, *73*, 76–107. https://doi.org/10.1016/j.ocemod.2013.10.005

Danabasoglu, G., Yeager, S. G., Kwon, Y.-O., Tribbia, J. J., Phillips, A. S., & Hurrell, J. W. (2012). Variability of the Atlantic meridional overturning circulation in CCSM4. *Journal of Climate*, *25*(15), 5153–5172. https://doi.org/10.1175/jcli-d-11-00463.1

Delworth, T., Manabe, S., & Stouffer, R. J. (1993). Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *Journal of Climate*, *6*(11), 1993–2011. https://doi.org/10.1175/1520-0442(1993)006<1993:ivottc>2.0.co;2

Delworth, T., & Zeng, F. (2016). The impact of the North Atlantic Oscillation on climate through its influence on the Atlantic meridional overturning circulation. *Journal of Climate*, *29*(3), 941–962. https://doi.org/10.1175/jcli-d-15-0396.1

Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J.-H., & Dunne, K. A. (2020). SPEAR: The next generation GFDL modeling system for seasonal to multidecadal Prediction and projection. *Journal of Advances in Modeling Earth Systems*, *12*(3), e2019MS001895. https://doi.org/10.1029/2019ms001895

Deser, C., Alexander, M. A., Xie, S.-P., & Phillips, A. S. (2010). Sea surface temperature variability: Patterns and mechanisms. *Annual Review of Marine Science*, *2*, 115–143. https://doi.org/10.1146/annurev-marine-120408-151453

Deser, C., Lehner, F., Rodgers, K., Ault, T., Delworth, T., DiNezio, P., & Ting, M. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286. https://doi.org/10.1038/s41558-020-0731-2

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, *38*(3–4), 527–546. https://doi.org/10.1007/s00382-010-0977-x

Deser, C., Phillips, A. S., Alexander, M. A., & Smoliak, B. V. (2014). Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *Journal of Climate*, *27*(6), 2271–2296. https://doi.org/10.1175/jcli-d-13-00451.1

Deser, C., Terray, L., & Phillips, A. S. (2016). Forced and internal components of winter air temperature trends over North America during the past 50 years: Mechanisms and implications. *Journal of Climate*, *29*(6), 2237–2258. https://doi.org/10.1175/jcli-d-15-0304.1

Dong, B., & Sutton, R. T. (2005). Mechanism of interdecadal thermohaline circulation variability in a coupled ocean-atmosphere GCM. *Journal of Climate*, *18*(8), 1117–1135. https://doi.org/10.1175/jcli3328.1

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., & Aumont, O. (2013). Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5. *Climate Dynamics*, *40*(9–10), 2123–2165. https://doi.org/10.1007/s00382-012-1636-1

Enfield, D. B., Mestas-Nuñez, A. M., & Trimble, P. J. (2001). The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophysical Research Letters*, *28*(10), 2077–2080. https://doi.org/10.1029/2000gl012745

England, M. H., McGregor, S., Spence, P., Meehl, G. A., Timmermann, A., Cai, W., et al. (2014). Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nature Climate Change*, *4*(3), 222–227. https://doi.org/10.1038/nclimate2106

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geophysical Model Development*, *9*, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fetterer, F., Knowles, K., Meier, W., Savoie, M., & Windnagel, A. (2017). *Updated monthly. Sea ice index*. NSIDC: National Snow and Ice Data Center.

Folland, C. K., Boucher, O., Colman, A., & Parker, D. E. (2018). Causes of irregularities in trends of global mean surface temperature since the late 19th century. *Science Advances*, *4*(6), eaao5297. https://doi.org/10.1126/sciadv.aao5297

Frankcombe, L. M., England, M. H., Kajtar, J. B., Mann, M. E., & Steinman, B. A. (2018). On the choice of ensemble mean for estimating the forced signal in the presence of internal variability. *Journal of Climate*, *31*(14), 5681–5693. https://doi.org/10.1175/jcli-d-17-0662.1

Frankignoul, C., Gastineau, G., & Kwon, Y.-O. (2017). Estimation of the SST response to anthropogenic and external forcing and its impact on the Atlantic multidecadal oscillation and the Pacific decadal oscillation. *Journal of Climate*, *30*(24), 9871–9895. https://doi.org/10.1175/jcli-d-17-0009.1

Gastineau, G., Friedman, A. R., Khodri, M., & Vialard, J. (2019). Global ocean heat content redistribution during the 1998–2012 Interdecadal Pacific Oscillation negative phase. *Climate Dynamics*, *53*(1–2), 1187–1208. https://doi.org/10.1007/s00382-018-4387-9

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., et al. (2019). Global emissions pathways under different socioeconomic scenarios for use in CMIP6: A dataset of harmonized emissions trajectories through the end of the century. *Geophysical Model Development*, *12*(4), 1443–1475. https://doi.org/10.5194/gmd-12-1443-2019

Gillett, N. P., Kirchmeier-Young, M., Ribes, A., Shiogama, H., Hegerl, G. C., Knutti, R., et al. (2021). Constraining human contributions to observed warming since the pre-industrial period. *Nature Climate Change*, *11*(3), 207–212. https://doi.org/10.1038/s41558-020-00965-9

Gregory, J., Dixon, K., Stouffer, R., Weaver, A., Driesschaert, E., & Eby, M. (2005). A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric concentration. *Geophysical Research Letters*, *32*(12), L12703. https://doi.org/10.1029/2005gl023209

Gu, L., Chen, J., Xu, C.-Y., Kim, J.-S., Chen, H., Xia, J., & Zhang, L. (2019). The contribution of internal climate variability to climate change impacts on droughts. *The Science of the Total Environment*, *684*, 229–246. https://doi.org/10.1016/j.scitotenv.2019.05.345

Guarino, M.-V., Sime, L. C., Schroeder, D., Lister, G., & Hatcher, R. (2020). Machine dependence and reproducibility for coupled climate simulations: The hadGEM3-GC3. 1 CMIP preindustrial simulation. *Geoscientific Model Development*, *13*(1), 139–154. https://doi.org/10.5194/gmd-13-139-2020

Hagos, S. M., Leung, L. R., Yoon, J.-H., Lu, J., & Gao, Y. (2016). A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western North America from the large ensemble CESM simulations. *Geophysical Research Letters*, *43*(3), 1357–1363. https://doi.org/10.1002/2015gl067392

Haustein, K., Otto, F. E., Venema, V., Jacobs, P., Cowtan, K., Hausfather, Z., et al. (2019). A limited role for unforced internal variability in twentieth-century warming. *Journal of Climate*, *32*(16), 4893–4917. https://doi.org/10.1175/jcli-d-18-0555.1

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108. https://doi.org/10.1175/2009bams2607.1

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., & Folland, C. K. (2015). A tripole index for the interdecadal Pacific oscillation. *Climate Dynamics*, *45*(11–12), 3077–3090. https://doi.org/10.1007/s00382-015-2525-1

Hirschi, J. J.-M., Barnier, B., Böning, C., Biastoch, A., Blaker, A. T., Coward, A., et al. (2020). The Atlantic meridional overturning circulation in high-resolution models. *Journal of Geophysical Research: Ocean*, *125*(4), e2019JC015522. https://doi.org/10.1029/2019jc015522

Hobbs, W., Palmer, M. D., & Monselesan, D. (2016). An energy conservation analysis of ocean drift in the CMIP5 global coupled models. *Journal of Climate*, *29*(5), 1639–1653. https://doi.org/10.1175/jcli-d-15-0477.1

Hourdin, F., Jam, A., Rio, C., Couvreux, F., Sandu, I., Lefebvre, M.-P., et al. (2019). Unified parameterization of convective boundary layer transport and clouds with the thermal plume model. *Journal of Advanced Modeling Earth Systems*, *11*, 2910–2933. https://doi.org/10.1029/2019MS001666

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., et al. (2020). LMDZ6: The improved atmospheric component of the IPSL coupled model. *Journal of Advanced Modeling Earth Systems*, *12*. https://doi.org/10.1029/2019MS001892

Hourdin, F., Rio, C., Jam, A., Traore, A.-K., & Musat, I. (2020). Convective boundary layer control of the sea surface temperature in the tropics. *Journal of Advanced Modeling Earth Systems*, *12*(6), e2019MS001988. https://doi.org/10.1029/2019ms001988

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., et al. (2017). Extended reconstructed sea surface temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *Journal of Climate*, *30*(20), 8179–8205. https://doi.org/10.1175/jcli-d-16-0836.1

Jiang, W., Gastineau, G., & Codron, F. (2021). Multi-centennial variability driven by salinity exchanges between the Atlantic and the arctic ocean in a coupled climate model. *Journal of Advanced Modeling Earth Systems*, *13*, e2020MS002366. https://doi.org/10.1029/2020ms002366

Jones, P., Lister, D., Osborn, T., Harpham, C., Salmon, M., & Morice, C. (2012). Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *Journal of Geophysical Research: Atmospheres*, *117*(D5), D05127. https://doi.org/10.1029/2011jd017139

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., et al. (2015). The community earth system model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333–1349. https://doi.org/10.1175/BAMS-D-13-00255.1

Kennedy, J., Rayner, N., Smith, R., Parker, D., & Saunby, M. (2011a). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *Journal of Geophysical Research*, *116*(D14), D14104. https://doi.org/10.1029/2010jd015218

Kennedy, J., Rayner, N., Smith, R., Parker, D., & Saunby, M. (2011b). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization. *Journal of Geophysical Research*, *116*(D14), D14104. https://doi.org/10.1029/2010jd015220

Kerr, R. A. (2000). A North Atlantic climate pacemaker for the centuries. *Science*, *288*(5473), 1984–1985. https://doi.org/10.1126/science.288.5473.1984

Kirchmeier-Young, M. C., Zwiers, F. W., & Gillett, N. P. (2017). Attribution of extreme events in Arctic sea ice extent. *Journal of Climate*, *30*(2), 553–571. https://doi.org/10.1175/jcli-d-16-0412.1

Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., & Mann, M. E. (2005). A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophysical Research Letters*, *32*(20). https://doi.org/10.1029/2005gl024233

Kosaka, Y., & Xie, S.-P. (2013). Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, *501*(7467), 403–407. https://doi.org/10.1038/nature12534

Kravtsov, S., & Callicutt, D. (2017). On semi-empirical decomposition of multidecadal climate variability into forced and internally generated components. *International Journal of Climatology*, *37*(12), 4417–4433. https://doi.org/10.1002/joc.5096

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., et al. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeological Cycles*, *19*(1), GB1015. https://doi.org/10.1029/2003GB002199

Lee, S.-K., Park, W., Baringer, M. O., Gordon, A. L., Huber, B., & Liu, Y. (2015). Pacific origin of the abrupt increase in Indian Ocean heat content during the warming hiatus. *Nature Geoscience*, *8*(6), 445–449. https://doi.org/10.1038/ngeo2438

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E., Brunner, L., et al. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth Systems Science Data*, *11*, 491–508. https://doi.org/10.5194/esd-2019-93

Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., & Locarnini, R. A. (2012). World ocean heat content and thermosteric sea level change (0–2000 m). *Geophysical Research Letters*, *39*(10), L10603. https://doi.org/10.1029/2012gl051106

Li, Z., Zhang, W., Jin, F.-F., Stuecker, M. F., Sun, C., Levine, A. F., & Liu, C. (2020). A robust relationship between multidecadal global warming rate variations and the Atlantic multidecadal variability. *Climate Dynamics*, *55*(7), 1945–1959. https://doi.org/10.1007/s00382-020-05362-8

Liguori, G., McGregor, S., Arblaster, J. M., Singh, M. S., & Meehl, G. A. (2020). A joint role for forced and internally-driven variability in the decadal modulation of global warming. *Nature Communications*, *11*(1), 1–7. https://doi.org/10.1038/s41467-020-17683-7

Liu, W., Xie, S.-P., & Lu, J. (2016a). Reply to: "Correspondence: Variations in ocean heat uptake during the surface warming hiatus". *Nature Communications*, *7*, 12542. https://doi.org/10.1038/ncomms12542

Liu, W., Xie, S.-P., & Lu, J. (2016b). Tracking ocean heat uptake during the surface warming hiatus. *Nature Communications*, *7*(1), 1–9. https://doi.org/10.1038/ncomms10926

Llovel, W., & Terray, L. (2016). Observed southern upper-ocean warming over 2005–2014 and associated mechanisms. *Environmental Research Letters*, *11*, 124023. https://doi.org/10.1088/1748-9326/11/12/124023

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, *20*(2), 130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:dnf>2.0.co;2

Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Brockmann, P., et al. (2020). Implementation of the CMIP6 forcing data in the IPSL-CM6A-LR model. *Journal of Advanced Modeling Earth Systems*, *12*. https://doi.org/10.1029/2019MS001940

Ma, X., Liu, W., Burls, N. J., Chen, C., Cheng, J., Huang, G., & Li, X. (2021). Evolving AMOC multidecadal variability under different CO_2 forcings. *Climate Dynamics*, *1–18*. https://doi.org/10.1007/s00382-021-05730-y

Madec, G., Bourdallé-Badie, R., Bouttier, P., Bricaud, C., Bruciaferri, D., Calvert, D., & Vancoppenolle, M. (2017). *NEMO ocean engine (version v3.6)*. Notes du Pôle de modélisation de l'Institut Pierre-Simon Laplace (IPSL). https://doi.org/10.5281/zenodo.1472492

Madeleine, J.-b., Hourdin, F., Grandpeix, J.-y., Rio, C., Dufresne, J.-L., & Vignon, E. (2020). Improved representation of clouds in the atmospheric component LMDZ6A of the IPSL-CM6A earth system model. *Journal of Advances in Modeling Earth Systems*, *12*(10), e2020MS002046. https://doi.org/10.1029/2020ms002046

Maher, N., McGregor, S., England, M. H., & Gupta, A. S. (2015). Effects of volcanism on tropical variability. *Geophysical Research Letters*, *42*(14), 6024–6033. https://doi.org/10.1002/2015gl064751

Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., et al. (2019). The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *Journal of Advanced Modeling Earth Systems*, *11*(7), 2050–2069. https://doi.org/10.1029/2019MS001639

Mann, M. E., Steinman, B. A., & Miller, S. K. (2014). On forced temperature changes, internal variability, and the AMO. *Geophysical Research Letters*, *41*(9), 3211–3219. https://doi.org/10.1002/2014gl059233

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., & Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society*, *78*(6), 1069–1079. https://doi.org/10.1175/1520-0477(1997)078<1069:apicow>2.0.co;2

Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A., & Trenberth, K. E. (2011). Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change*, *1*(7), 360–364. https://doi.org/10.1038/nclimate1229

Meehl, G. A., Hu, A., Arblaster, J. M., Fasullo, J., & Trenberth, K. E. (2013). Externally forced and internally generated decadal climate variability associated with the Interdecadal Pacific Oscillation. *Journal of Climate*, *26*(18), 7298–7310. https://doi.org/10.1175/jcli-d-12-00548.1

Meehl, G. A., Hu, A., Santer, B. D., & Xie, S.-P. (2016). Contribution of the Interdecadal Pacific Oscillation to twentieth-century global surface temperature trends. *Nature Climate Change*, *6*(11), 1005–1008. https://doi.org/10.1038/nclimate3107

Menary, M. B., Robson, J., Allan, R. P., Booth, B. B., Cassou, C., & Gastineau, G. (2020). Aerosol-forced AMOC changes in CMIP6 historical simulations. *Geophysical Research Letters*, *47*(14), e2020GL088166. https://doi.org/10.1029/2020gl088166

Morice, C. P., Kennedy, J. J., Rayner, N. A., & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research*, *117*(D8), D08101. https://doi.org/10.1029/2011jd017187

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., et al. (2016). The Pacific decadal oscillation, revisited. *Journal of Climate*, *29*(12), 4399–4427. https://doi.org/10.1175/jcli-d-15-0508.1

Nieves, V., Willis, J. K., & Patzert, W. C. (2015). Recent hiatus caused by decadal shift in Indo-Pacific heating. *Science*, *349*(6247), 532–535. https://doi.org/10.1126/science.aaa4521

Notz, D., Dörr, J., Bailey, D. A., Blockley, E., Bushuk, M., & Debernard, J. B. (2020). Arctic sea ice in CMIP6. *Geophysical Research Letters*, *47*(10), e2019GL086749. https://doi.org/10.1029/2019GL086749

Osborn, T. J., & Jones, P. (2014). The CRUTEM4 land-surface air temperature data set: Construction, previous versions and dissemination via Google Earth. *Earth System Science Data*, *6*(1), 61–68. https://doi.org/10.5194/essd-6-61-2014

Otterå, O. H., Bentsen, M., Drange, H., & Suo, L. (2010). External forcing as a metronome for Atlantic multidecadal variability. *Nature Geoscience*, *3*(10), 688–694. https://doi.org/10.1038/ngeo955

Parsons, L. A., Brennan, M. K., Wills, R. C., & Proistosescu, C. (2020). Magnitudes and spatial patterns of interdecadal temperature variability in CMIP6. *Geophysical Research Letters*, *47*(7), e2019GL086588. https://doi.org/10.1029/2019gl086588

Polyakov, I. V., Alexeev, V. A., Bhatt, U. S., Polyakova, E. I., & Zhang, X. (2010). North Atlantic warming: Patterns of long-term trend and multidecadal variability. *Climate Dynamics*, *34*(2–3), 439–457. https://doi.org/10.1007/s00382-008-0522-3

Power, S., Casey, T., Folland, C., Colman, A., & Mehta, V. (1999). Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics*, *15*(5), 319–324. https://doi.org/10.1007/s003820050284

Qasmi, S., Cassou, C., & Boé, J. (2017). Teleconnection between Atlantic multidecadal variability and European temperature: Diversity and evaluation of the Coupled Model Intercomparison Project phase 5 models. *Geophysical Research Letters*, *44*(21), 11–140. https://doi.org/10.1002/2017gl074886

Rohde, R., Muller, R., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., & Wickham, C. (2013). A new estimate of the average Earth surface land temperature spanning 1753 to 2011. *Geoinformatics & Geostatistics: An Overview*, *7*, 2.

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., Wurtele, J., & Mosher, S. (2013). Berkeley Earth temperature averaging process. *Geoinformatics & Geostatistics: An Overview*, *1*(2), 20–100. https://doi.org/10.4172/2327-4581.1000103

Rousset, C., Vancoppenolle, M., Madec, G., Fichefet, T., Flavoni, S., Barthélemy, A., et al. (2015). The Louvain-La-Neuve sea ice model LIM3. 6: Global and regional capabilities. *Geoscientific Model Development*, *8*(10), 2991–3005. https://doi.org/10.5194/gmd-8-2991-2015

Sanderson, B. M., Oleson, K. W., Strand, W. G., Lehner, F., & O'Neill, B. C. (2018). A new ensemble of GCM simulations to assess avoided impacts in a climate mitigation scenario. *Climatic Change*, *146*(3–4), 303–318. https://doi.org/10.1007/s10584-015-1567-z

Schlesinger, M. E., & Ramankutty, N. (1994). An oscillation in the global climate system of period 65–70 years. *Nature*, *367*(6465), 723–726. https://doi.org/10.1038/367723a0

Selten, F. M., Branstator, G. W., Dijkstra, H. A., & Kliphuis, M. (2004). Tropical origins for recent and future Northern Hemisphere climate change. *Geophysical Research Letters*, *31*(21). https://doi.org/10.1029/2004GL020739

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K., Forster, P. M., & Hargreaves, J. C. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *58*(4), e2019RG000678. https://doi.org/10.1029/2019rg000678

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., & Meier, W. N. (2012). Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophysical Research Letters*, *39*(16), L16502. https://doi.org/10.1029/2012gl052676

Suarez-Gutierrez, L., Li, C., Müller, W. A., & Marotzke, J. (2018). Internal variability in European summer temperatures at 1.5°C and 2°C of global warming. *Environmental Research Letters*, *13*(6), 064026. https://doi.org/10.1088/1748-9326/aaba58

Sutton, R. T., & Dong, B. (2012). Atlantic Ocean influence on a shift in European climate in the 1990s. *Nature Geoscience*, *5*(11), 788–792. https://doi.org/10.1038/ngeo1595

Swingedouw, D., Mignot, J., Guilyardi, E., Nguyen, S., & Ormières, L. (2017). Tentative reconstruction of the 1998–2012 hiatus in global temperature warming using the IPSL-CM5A-LR climate model. *Comptes Rendus Geoscience*, *349*(8), 369–379. https://doi.org/10.1016/j.crte.2017.09.014

Swingedouw, D., Ortega, P., Mignot, J., Guilyardi, E., Masson-Delmotte, V., Butler, P. G., & Séférian, R. (2015). Bidecadal North Atlantic Ocean circulation variability controlled by timing of volcanic eruptions. *Nature Communication*, *6*(1), 1–12. https://doi.org/10.1038/ncomms7545

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000jd900719

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/bams-d-11-00094.1

Thompson, D. W., Barnes, E. A., Deser, C., Foust, W. E., & Phillips, A. S. (2015). Quantifying the role of internal climate variability in future climate trends. *Journal of Climate*, *28*(16), 6443–6456. https://doi.org/10.1175/jcli-d-14-00830.1

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., & Knutti, R. (2020). Past warming trend constrains future warming in CMIP6 models. *Science Advances*, *6*(12), eaaz9549. https://doi.org/10.1126/sciadv.aaz9549

Trenberth, K. E., Fasullo, J. T., & Shepherd, T. G. (2015). Attribution of climate extreme events. *Nature Climate Change*, *5*(8), 725–730. https://doi.org/10.1038/nclimate2657

Trenberth, K. E., & Shea, D. J. (2006). Atlantic hurricanes and natural variability in 2005. *Geophysical Research Letters*, *33*(12), L12704. https://doi.org/10.1029/2006gl026894

Wang, S. S., Zhao, L., Yoon, J.-H., Klotzbach, P., & Gillies, R. R. (2018). Quantitative attribution of climate effects on Hurricane Harvey's extreme rainfall in Texas. *Environmental Research Letters*, *13*(5), 054014. https://doi.org/10.1088/1748-9326/aabb85

Weijer, W., Cheng, W., Garuba, O. A., Hu, A., & Nadiga, B. (2020). CMIP6 models predict significant 21st century decline of the Atlantic meridional overturning circulation. *Geophysical Research Letters*, *47*(12), e2019GL086075. https://doi.org/10.1029/2019gl086075

Wettstein, J. J., & Deser, C. (2014). Internal variability in projections of twenty-first-century Arctic sea ice loss: Role of the large-scale atmospheric circulation. *Journal of Climate*, *27*(2), 527–550. https://doi.org/10.1175/jcli-d-12-00839.1

Wittenberg, A. T., Rosati, A., Delworth, T. L., Vecchi, G. A., & Zeng, F. (2014). ENSO modulation: Is it decadally predictable? *Journal of Climate*, *27*(7), 2667–2681. https://doi.org/10.1175/jcli-d-13-00577.1

Wyatt, M. G., Kravtsov, S., & Tsonis, A. A. (2012). Atlantic multidecadal oscillation and Northern Hemisphere's climate variability. *Climate Dynamics*, *38*(5–6), 929–949. https://doi.org/10.1007/s00382-011-1071-8

Yeager, S., & Robson, J. (2017). Recent progress in understanding and predicting Atlantic decadal climate variability. *Current Climate Change Reports*, *3*(2), 112–127. https://doi.org/10.1007/s40641-017-0064-z

Zhang, R. (2015). Mechanisms for low-frequency variability of summer Arctic sea ice extent. *Proceedings of the National Academy of Sciences USA*, *112*(15), 4570–4575. https://doi.org/10.1073/pnas.1422296112

Zhang, R., & Delworth, T. L. (2006). Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophysical Research Letters*, *33*(17), L17712. https://doi.org/10.1029/2006gl026267

Zhang, R., Delworth, T. L., & Held, I. M. (2007). Can the Atlantic Ocean drive the observed multidecadal variability in Northern Hemisphere mean temperature? *Geophysical Research Letters*, *34*(2), L02709. https://doi.org/10.1029/2006gl028683