




# Increased risk of near term global warming due to a recent AMOC weakening

Rémy Bonnet<sup>1</sup>, Didier Swingedouw<sup>2</sup>, Guillaume Gastineau<sup>3</sup>, Olivier Boucher<sup>1</sup>, Julie Deshayes<sup>3</sup>, Frédéric Hourdin<sup>4</sup>, Juliette Mignot<sup>3</sup>, Jérôme Servonnat<sup>5</sup> & Adriana Sima<sup>4</sup>

Some of the new generation CMIP6 models are characterised by a strong temperature increase in response to increasing greenhouse gases concentration<sup>1</sup>. At first glance, these models seem less consistent with the temperature warming observed over the last decades. Here, we investigate this issue through the prism of low-frequency internal variability by comparing with observations an ensemble of 32 historical simulations performed with the IPSL-CM6A-LR model, characterized by a rather large climate sensitivity. We show that members with the smallest rates of global warming over the past 6-7 decades are also those with a large internally-driven weakening of the Atlantic Meridional Overturning Circulation (AMOC). This subset of members also matches several AMOC observational fingerprints, which are in line with such a weakening. This suggests that internal variability from the Atlantic Ocean may have dampened the magnitude of global warming over the historical era. Taking into account this AMOC weakening over the past decades means that it will be harder to avoid crossing the 2 °C warming threshold.

<sup>1</sup>Institut Pierre-Simon Laplace, Sorbonne Université/CNRS, Paris, France. <sup>2</sup>Environnements et Paléoenvironnements Océaniques et Continentaux, Université de Bordeaux/CNRS, Bordeaux, France. <sup>3</sup>Laboratoire d'Océanographie et du Climat: Expérimentations et Approches Numériques, Institut Pierre-Simon Laplace, Sorbonne Université/CNRS/IRD/MNHN, Paris, France. <sup>4</sup>Laboratoire de Météorologie Dynamique, Institut Pierre-Simon Laplace, Sorbonne Université/CNRS/Ecole Normale Supérieure/Ecole Polytechnique, Paris, France. <sup>5</sup>Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace, CEA/CNRS/UVSQ, Gif-sur-Yvette, France. ✉email: [remy.bonnet@ipsl.fr](mailto:remy.bonnet@ipsl.fr)

Projections of future climate from Earth System Models (ESM) represent a crucial source of information for adaptation and mitigation planning. The Equilibrium Climate Sensitivity (ECS), defined as the warming resulting from a CO<sub>2</sub> doubling once the system has equilibrated, as well as the Transient Climate Response (TCR), defined as the warming produced after 70 years by a 1% per year increase in CO<sub>2</sub> up to a doubling of CO<sub>2</sub>, are often used to estimate the sensitivity of ESMs to anthropogenic CO<sub>2</sub> emissions. Some models from the Coupled Model Intercomparison Project phase 6 (CMIP6) exhibit relatively large climate sensitivities. Indeed, one-third of the models<sup>1</sup> have an effective ECS greater than the IPCC AR5 “likely” range (1.5–4.5 K) and one-fifth of the models<sup>2</sup> have a TCR greater than the IPCC AR5 “likely” range (1–2.5 K). Consistently, these models with higher sensitivity project greater warming over the 21st century than previously reported in CMIP5, although a direct comparison is challenging because emission scenarios used to drive the models have also evolved. These larger sensitivities from some of the CMIP6 models are also difficult to reconcile with estimates of ECS and TCR based on observations over the historical period<sup>3–5</sup> (renamed S\_hist and TCR\_hist hereafter), which are around 2.5 K for S\_hist according to a recent study<sup>6</sup> and around 1.6 K for TCR\_hist<sup>5</sup>. The associated uncertainties to S\_hist and TCR\_hist, however, are large. Alternative estimates using the observed variability have also been proposed<sup>7</sup>, but do not allow us to rule out high ECS values.

This new generation of models with large effective ECS raises important questions: (i) are these models realistic in comparison to the last few decades of climate observations? (ii) how to interpret these highly sensitive models in relation to climate change over the historical era? If these models are not falsifiable, it would imply higher risks and costs induced by future climate change impacts and the need for greater and faster mitigation efforts to achieve climate targets than previously thought. To address these issues, recent studies<sup>2,8,9</sup> tried to constrain CMIP6 climate model projections with observed warming trends over the last decades. Although it makes sense to focus on the last few decades to estimate the impact of greenhouse gases (GHGs), because aerosol forcing has not changed much, there is a need to understand the model spread and to explore and attribute the recent observed changes. In order to take into account all the information available over the historical period, a recent study<sup>10</sup> developed a new statistical method, leading to a reduction of the uncertainty on the projected future warming by about 50% and a slightly higher future warming in CMIP6 relative to CMIP5. However, the authors showed that this method poorly accounted for the internal variability in some CMIP6 models<sup>10</sup> that exhibit stronger multi-decadal to centennial internal climate variability than CMIP5 models<sup>11</sup>, which might have strong implications. Indeed, the low-frequency internal variability at decadal to multi-centennial time scales can temporarily enhance or reduce the long-term imprints of externally forced climate change. In particular, such variability might have a strong impact on the way climate models should be compared to observations.

In this study, we address these issues highlighting new perspectives, which allows us notably to evaluate the risk of larger warming over the next decades, using the IPSL-CM6A-LR model. This model is one of those new generation models characterized by a rather large sensitivity, close to the upper bound of the constrained projections, with an effective ECS of 4.5 K—computed from a 4xCO<sub>2</sub> abrupt forcing experiment and a 150-year regression—and a TCR of 2.4 K. It is also characterized by a relatively high low-frequency internal climate variability in comparison to the CMIP6 models<sup>11</sup>. Taking advantage of CMIP6, the Institut Pierre-Simon Laplace Climate Modelling Centre has produced an ensemble of extended historical simulations<sup>12</sup> (referred to as

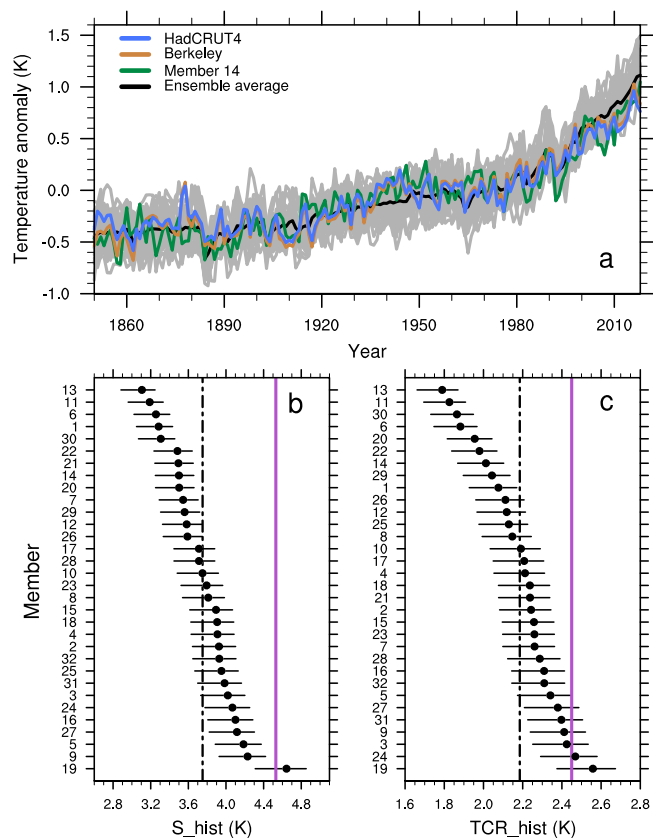
IPSL-EHS, see Methods). By using this ensemble, we show the influence of multi-centennial internal climate variability on the Global near-Surface Air Temperature (GSAT) warming since the middle of the 20<sup>th</sup> century, as well as the related uncertainties on the estimation of S\_hist and TCR\_hist. The spread in the GSAT trends is related to trends in AMOC. Specifically, we show that members matching best the observed record of GSAT are also those with a large internally-driven weakening of the AMOC in the past decades. Several AMOC fingerprints confirm the consistency of the AMOC variations in these members with observations. The mechanisms involved are also confirmed by the analysis of other climate model ensembles. Over the next decades, the low-frequency internal variability in these members leads to slightly more warming than the rest of the ensemble.

## Results and discussion

**Climate sensitivity of IPSL-CM6A-LR over the historical period.** The IPSL-CM6A-LR ensemble mean shows a larger than observed warming over the historical period (Fig. 1a). This global warming bias can be due to (i) an overestimated climate sensitivity in the model, (ii) an overestimation of the net anthropogenic radiative forcing, or (iii) multi-decadal to multi-centennial internal variability which offsets some of the warming in the observations. While the radiative forcing by aerosols in IPSL-CM6A-LR is relatively weak at  $-0.6 \text{ Wm}^{-2}$ , which might be in favor of hypothesis (ii), we argue that this is compensated by an underestimation of the (positive) radiative forcing by non-CO<sub>2</sub> greenhouse gases<sup>13</sup>. Thus, although we cannot discard hypothesis ii) in the light of the large uncertainties on the aerosol forcing, we focus this study on the exploration of hypotheses (i) and (iii).

Despite the warmer than the observed ensemble mean, some members have a consistent representation of the historical and more recent warming trends (Fig. 1a and Supplementary Fig. 3a). Member #14 for example, has one of the best representations of the GSAT in comparison to observations, with the lowest root mean square error (RMSE) of 0.14 K for the annual GSAT anomaly over the 1900–2018 period (Supplementary Fig. 3b).

We estimate S\_hist and TCR\_hist for each individual member of the IPSL-EHS using a similar method as former studies<sup>3,4</sup> based on the changes of surface air temperature in 1999–2018 compared to the preindustrial period (see Methods, Fig. 1b, c). The ECS value of 4.5 K diagnosed from the abrupt-4xCO<sub>2</sub> experiment is at the top of the S\_hist range inferred in the IPSL-EHS, which is consistent with the previous studies<sup>6,14</sup>. This is likely due to the non-linear relationship between radiative forcing and the change in GSAT as a result of stronger feedbacks in a warmer climate<sup>15</sup>. The pattern effect might also contribute. Indeed, the short-term warming of models has a pattern of surface air temperature warming different from the stabilized warming pattern leading to weaker low cloud feedback. The range of S\_hist values is indeed more consistent with the value diagnosed from an abrupt-2xCO<sub>2</sub> experiment at 3.8 K. The same is true for the TCR value of 2.4 K diagnosed from the 1pctCO<sub>2</sub> experiment which is larger than the mean of TCR\_hist values. More importantly, the spread in modelled S\_hist and TCR\_hist values is substantial ( $\pm 2\sigma = 1.25 \text{ K}$ ), suggesting such values are largely influenced by internal climate variability (Fig. 1b, c). If the low-frequency internal variability simulated by the model is realistic then we should expect a large uncertainty in observational estimates of S\_hist, which question its definition and its utility for assessing correctly future greenhouse emissions admissible to remain below-given thresholds in GSAT. Furthermore, this low-frequency internal variability presumably strongly interacts with the transient response to external forcing and thereby affects TCR\_hist.



**Fig. 1 Sensitivity of IPSL-CM6A-LR over the historical period.** **a** Global mean near-surface air temperature (GSAT) anomaly (K) relative to the 1880–2018 period for the ensemble average (black), individual members (gray), and member #14 (green) of the IPSL ensemble of extended historical simulations (IPSL-EHS), the infilled HadCRUT4-CW<sup>41,42</sup> (blue) and the Berkeley<sup>45</sup> observational datasets (brown). **b** The climate sensitivity calculated over the historical period ( $S_{\text{hist}}$ , in K) for each individual member of the IPSL-EHS ranked from the lowest to the highest value (see Methods). The black dot-dashed line indicates the ensemble mean of  $S_{\text{hist}}$  values and the purple line indicates the Equilibrium Climate Sensitivity (ECS) value<sup>13</sup> of the IPSL-CM6A-LR model. **c** Same as **b** but for the transient climate response calculated over the historical period ( $TCR_{\text{hist}}$ , in K), the purple line indicating the Transient Climate Response (TCR) value<sup>13</sup> of IPSL-CM6A-LR. Error bars in **b** and **c** are calculated using the lowest and largest values of the forcing due to a doubling in atmospheric  $\text{CO}_2$  concentration (see Methods for more details).

**Low-frequency internal climate variability of the global surface temperature.** IPSL-CM6A-LR is characterized by a multi-centennial variability of AMOC and GSAT in the piControl simulation<sup>13</sup>. A recent study<sup>16</sup> on the associated mechanisms suggests that a positive AMOC anomaly is associated with warm temperature anomalies in the Northern Hemisphere, a reduced Arctic sea-ice extent, and salinity-driven positive density anomalies in the Nordic Seas and in the North Atlantic subpolar gyre. During a positive AMOC, both ocean and sea ice freshwater export from the Arctic Ocean decreases, which contributes to build a large freshwater anomaly in the top 100 m of the Arctic Ocean, while the Atlantic inflow of salty water into the Arctic increases the salinity below 100 m. After several decades, the freshwater anomaly covers most of the western Arctic which eventually increases the oceanic freshwater export from the Arctic into the Nordic Seas through the Fram Strait. This slows down the AMOC. The Northern Hemisphere climate then cools and Arctic sea ice grows, reversing the phase of the cycle. The 32

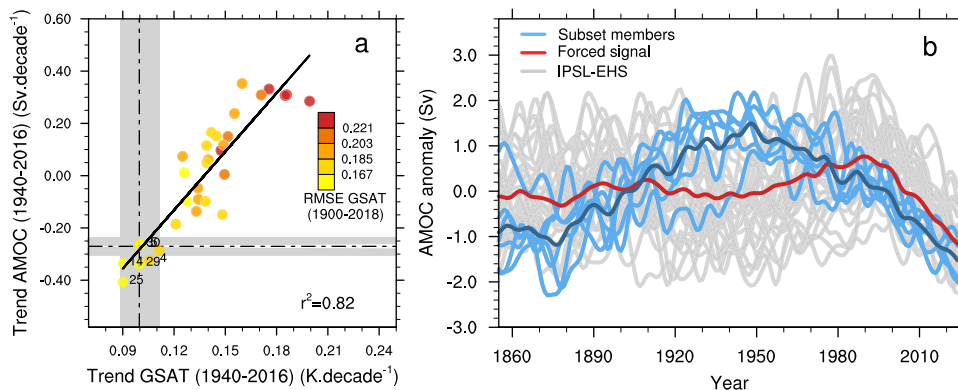
members of the ensemble were run with the same external forcings but initialised from different years in the pre-industrial control simulation in order to sample this multi-centennial variability. Due to the link between the AMOC variability and the initial state used in the historical simulations (Supplementary Fig. 4), a clear positive significant relationship exists between the AMOC trend in 1850–2010 and the  $S_{\text{hist}}$  and  $TCR_{\text{hist}}$ .

A recent AMOC reconstruction<sup>17</sup>, using an index (named “Caesar index” hereafter) based on the relation between the North Atlantic Subpolar Gyre Sea Surface Temperature (SST) relative to the global SST from November to May and the AMOC, suggests a weakening of the AMOC since the 1940s. Although some debate exists concerning the robustness of this index, the implied AMOC weakening trend is supported by other reconstructions and indices<sup>18</sup>. There are still substantial uncertainties though, and other AMOC fingerprints will be considered later on in this study.

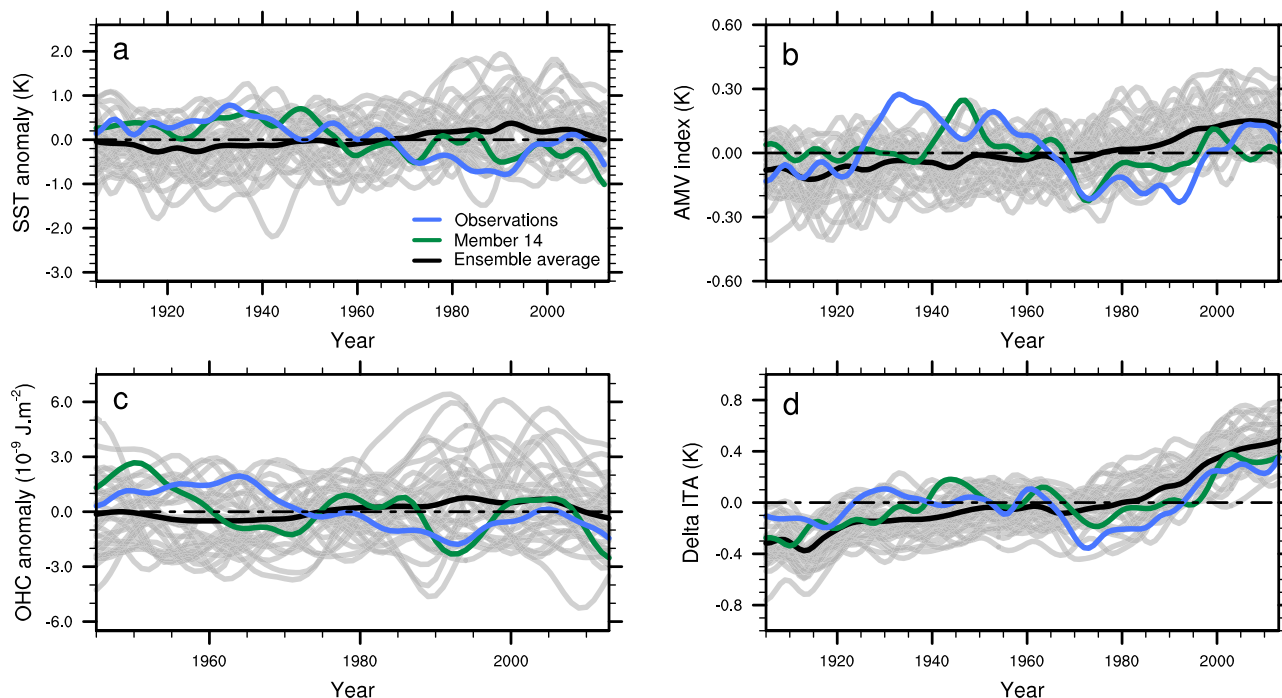
A significant positive relationship is found in IPSL-EHS between GSAT and AMOC trends over the 1940–2016 period (when AMOC weakening is suggested to start in the reconstruction), with a determination coefficient  $r^2=0.82$  (Fig. 2a), in line with previous studies<sup>19–21</sup>. Unlike the historical period, no clear relationship is found between the AMOC trends over the 1940–2016 period and the calculated values of  $S_{\text{hist}}$  or  $TCR_{\text{hist}}$ , as these quantities rely on the temperature difference between the pre-industrial (1850–1879) and present-day (1999–2018) periods. Figure 2a shows that ensemble members with the weakest GSAT warming are those with the strongest AMOC weakening and the members with the largest GSAT warming are those with the strongest AMOC strengthening. The members whose warming levels match best the GSAT observations both in terms of the recent trends and RMSE over the 1900–2018 period are also characterized by a negative AMOC trend (Fig. 2a, Supplementary Fig. 3) suggesting a possible link. Member #14 of IPSL-EHS in particular is again one of the closest members to observations in terms of both the trends in reconstructed AMOC index<sup>17</sup> and GSAT. We pre-select members that are consistent with the observed GSAT and AMOC trends, and then select the six members with the lowest GSAT RMSE among those. The subset is composed of members #14, #4, #5, #25, #29, and #30, with RMSE ranked 1<sup>st</sup>, 8<sup>th</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 5<sup>th</sup>, and 7<sup>th</sup>, respectively.

These members are characterized by a strengthening of the AMOC until the 1940s followed by a decline, although member #30 shows a different evolution from the other members until the 1920s (Fig. 2b). Conversely, the forced AMOC estimated by the ensemble mean shows only weak changes, with a small strengthening up to the 1990s followed by a weakening, as found in other CMIP6 models<sup>22</sup>. The evaluation of the relative influence of internal variability and forced response in the AMOC weakening of these members indicates that internal climate variability has been the main driver of the AMOC variability in IPSL-EHS since the 1940s (Supplementary Section 1). From the 21st century onwards, however, external forcings have an increasing influence on internal variability, and dominate the decline in AMOC (Fig. 2b). Consequently, the AMOC weakening suggested in the Caesar reconstruction<sup>17</sup> might be mainly internally-driven rather than externally forced. Conversely, an internally generated enhancement of the AMOC in the early 20th century might have warmed GSAT at that time, in agreement with other studies<sup>23</sup>, suggesting an important role of internal variability for this early century warming. This is consistent with the increase of the Caesar index over the beginning of the 20<sup>th</sup> century (Fig. 3a).

We now check whether the members identified previously present realistic trends in North Atlantic SST or are characterized by excessive cooling in comparison to the observations and the



**Fig. 2 Relationship between GSAT and AMOC over the historical period.** **a** Scatter plot of Global near-Surface Air Temperature (GSAT) ( $\text{K decade}^{-1}$ ) versus Atlantic Meridional Overturning Circulation (AMOC) ( $\text{Sv decade}^{-1}$ ) trends calculated over the 1940-2016 period from the IPSL ensemble of extended historical simulations (IPSL-EHS; filled circles) and the observations (dot-dashed lines), with HadCRUT4-CW<sup>41,42</sup> for the temperature and the Caesar index<sup>17</sup> as a proxy for the AMOC evolution, with the related uncertainties (gray). The black solid line represents the least square regression between these two variables in IPSL-EHS, with a determination coefficient  $r^2 = 0.82$  ( $p < 0.1$ , see Method). The color scale represents the RMSE between simulated and observed (HadCRUT4-CW<sup>41,42</sup>) annual GSAT anomalies over the period 1900-2018. **b** Time evolution of the low-pass filtered AMOC strength anomaly from IPSL-EHS (gray), the AMOC forced signal (or ensemble mean, red), and the subset of 6 members labelled in **a** (light blue), with the subset average (dark blue). The anomaly is calculated for each member with respect to its 1850-2018 average. A Lanczos low-pass filter with a cutoff period of 11 years is used.



**Fig. 3 Evaluation of Atlantic Meridional Overturning Circulation fingerprints.** **a** Time evolution of the Caesar index computed as the low-pass filtered sea surface temperature anomaly (K) averaged over the North Atlantic Subpolar Gyre from November to May from the ERSSTv5 observational dataset<sup>54</sup> (blue), the 32 historical members (gray) with the ensemble mean (black) and member #14 (green). **b** Same as **a** but for the Atlantic Multidecadal Variability index (see Methods). **c** Same as **a** but for the Ocean Heat Content (OHC) anomaly ( $10^9 \text{ J m}^{-2}$ ) between 0 and 700 m averaged over the Newfoundland region minus the OHC averaged over the North Atlantic Subpolar Gyre region from the IAP observational dataset<sup>55,56</sup> (blue). **d** Same as **a** but for the interhemispheric near-surface air temperature difference (K), with the HadCRUT4-CW observational dataset<sup>41,42</sup> (blue). A Lanczos low-pass filter with a cutoff period of 11 years is used in **c** and **d** and the anomalies are computed over the displayed period.

rest of the ensemble (Supplementary Fig. 5). Over the 1940-2018 period, member #14 has an averaged SST trend of  $0.06 \text{ K decade}^{-1}$  over the North Atlantic, slightly weaker than the observed trend of  $0.065 \text{ K decade}^{-1}$  (Supplementary Fig. 5). The pattern is overall close to the observations, with a spatial correlation coefficient of 0.73. The subset of members previously identified are also among the members with average trends over the North Atlantic closest

to the observations. Due to its strong AMOC weakening (Fig. 2a), member #25 shows the smallest SST trend in the North Atlantic with a value of  $0.042 \text{ K decade}^{-1}$ .

**Further lines of evidence from AMOC fingerprints.** In order to strengthen the hypothesis that a weakening of the AMOC in the

real world since the middle of the 20th century may have masked a fraction of the anthropogenic global warming, we evaluate four different observable AMOC fingerprints against our ensemble members. A first SST fingerprint is calculated following the Caesar index<sup>17</sup>. Then, we use the Atlantic Multidecadal Variability<sup>24–26</sup> index (AMV index, see Methods), as basin-wide low-frequency variations in the North Atlantic SST are related to the AMOC low-frequency variability in many climate models<sup>27</sup>. Whether low-frequency variations in AMOC and AMV are the results of naturally-occurring internal variability, external forcings, or a mix of both is still debated<sup>27–29</sup>. This is also the case in the IPSL-EHS, with significant Pearson correlations ranging from 0.56–0.97 between the AMOC and the AMV index over the 1850–2018 period in the various members (see Methods). We calculate a third index based on the upper (0–700 m) Ocean Heat Content (OHC) averaged offshore Newfoundland minus the average over the North Atlantic Subpolar Gyre ( $\Delta$ OHC). The upper OHC is indeed significantly correlated with the AMOC in IPSL-CM6A-LR over these two regions, with a negative correlation offshore Newfoundland and a positive correlation over the North Atlantic Subpolar Gyre (Supplementary Fig. 6a). This correlation dipole is consistent with observations and other climate model simulations<sup>30</sup>. As no significant correlations are found between the upper OHC and the AMOC forced signal, these correlations are mainly due to the AMOC internal variability (Supplementary Fig. 6b). Finally, the fourth index is defined as the difference between the Northern and the Southern Hemisphere anomalies of near-surface air temperature ( $\Delta$ ITA), which might also be a good indicator of the AMOC low-frequency variability, given its impact on interhemispheric temperature<sup>31</sup>. The time series of these four observation-based indices are included within the range of IPSL-EHS (Fig. 3), which gives confidence in the capacity of the model to represent climate dynamics accurately.

Furthermore, a good consistency is found between the six members identified above as best fitting GSAT over the observed period and the observations for the four AMOC fingerprints. In member #14 for example, Pearson correlations amount to 0.48 for the Caesar index (Fig. 3a), 0.49 for the AMV index (Fig. 3b), 0.46 for the  $\Delta$ OHC index (Fig. 3c) and 0.74 for the  $\Delta$ ITA index (Fig. 3d) and are significant for the Caesar index, the  $\Delta$ OHC index and  $\Delta$ ITA index (see Method). Notwithstanding, the decrease of  $-0.13$  K decade<sup>-1</sup> of the Caesar index in member #14 since the 1940s, consistent with a weakening AMOC, is close to the observed trend of  $-0.11$  K decade<sup>-1</sup>. The decrease in the observed  $\Delta$ OHC index of  $-0.45$   $10^9$  J m<sup>-2</sup> decade<sup>-1</sup> is also close to observations in member #14, with a decrease of about  $-0.42 \times 10^9$  J m<sup>-2</sup> decade<sup>-1</sup>. The pattern of the observed North Atlantic upper OHC trend calculated over the 1940–2020 period is consistent with a weakening AMOC, with significant positive trends over Newfoundland and significant negative trends over the Subpolar Gyre (Supplementary Fig. 6c). This pattern is no longer visible considering only the last decades (1980–2020), a period more influenced by the external forcing and characterized by strong warming of the upper OHC over the polar region (Supplementary Fig. 6d). The fact that these six members whose AMV is overall in phase with observations are those with a GSAT evolution most consistent with observations underscores the important role of the AMV in the GSAT variability at least in this model. This is consistent with a recent study<sup>32</sup> showing the main role of the AMV in the internal variability of GSAT on multi-decadal time scales. This good consistency is reinforced by the fact that member #14 has low RMSEs for these four AMOC fingerprints.

Member #14 is not the only one with overall low RMSEs for these four AMOC fingerprints. Members #29, #4, #5, #25, and #30, identified previously, also have overall the lowest RMSEs in

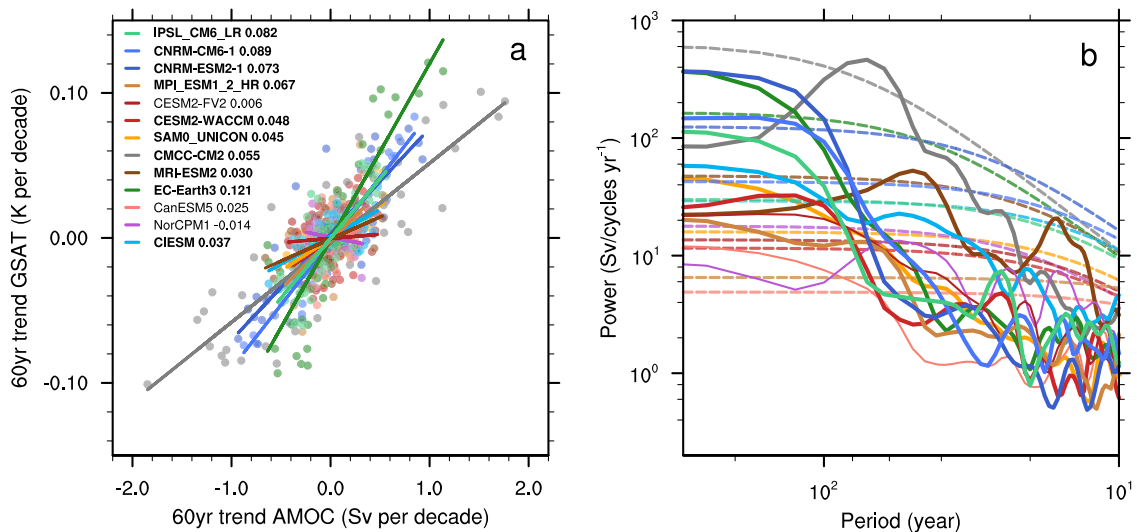
comparison to the rest of the ensemble (Supplementary Fig. 6). Notwithstanding, the consistency between AMOC fingerprints in the subset of members identified in Fig. 2a and in observations reinforces the hypothesis that the AMOC may have been weakening since the second half of the 20th century, thereby masking a fraction of the anthropogenic global warming as seen in the IPSL-EHS.

**Multi-model analysis.** We now examine how other climate models support our analysis of IPSL-EHS. A similar significant positive relationship between the AMOC and the GSAT 60-year trends is found in the piControl simulation of IPSL-CM6A-LR, as well as in ten out of thirteen other CMIP6 models (IPSL-CM6A-LR, both CNRM models, MPI-ESM1-2, CESM2-WACCM, SAM0-UNICON, CMCC-CM2, MRI-ESM2, EC-Earth3, and CIESM) (Fig. 4a). Those models also generally show significant multi-centennial AMOC variability compared to red noise at periods longer than 100 years (Fig. 4b). The models with the strongest AMOC-GSAT trends relationships are also those with the largest GSAT variance in their piControl simulations. Taken together these results highlight the role of internal variability in these models and suggest that the relationship between the AMOC and GSAT trends found in IPSL-CM6A-LR is rather widespread in CMIP6 models albeit with varying strength and range.

Only a few models with large historical ensembles are available to repeat the analysis performed with the IPSL-EHS: MPI-ESM1.1 with a low ECS, CNRM-CM6A-1, and CanESM5 with a high ECS. We investigate the consistency of our previous results with these models in Supplementary Section 2, which is summarized here. First, one can see that some members in the CNRM and the MPI ensembles reproduce the observed warming (similar GSAT trend as observations), unlike the CanESM5 ensemble. This latter model is therefore omitted. This may be due to its large climate sensitivity and/or its weak internal climate variability. Furthermore, a similar positive relationship between AMOC and GSAT trends is found in the CNRM historical simulations as in the IPSL-EHS, with, nevertheless, larger regression coefficients due to the larger internal climate variability of CNRM-CM6A-1<sup>11</sup>. A positive relationship is also found in the MPI ensemble, with a surprisingly large fraction of the historical members showing larger GSAT trends than observed despite a low climate sensitivity for this model. This might be due to its weak aerosol forcing. In addition, the members which fit best the GSAT and the AMOC weakening also match best the AMOC fingerprints in the CNRM-CM6 ensemble (Supplementary Fig. 9). This is not the case for the MPI ensemble where these members are scattered throughout the other members of the ensemble for the three fingerprints, which means that the members of the MPI ensemble with a weakening AMOC do not project well onto the observed AMOC fingerprints (Supplementary Fig. 8).

In conclusion, the analysis of three other large ensembles of historical simulations does not invalidate our results that suggest that the weakening in AMOC reconstructions since the second half of the 20th century might be due to internal variability. Indeed, similar behaviour can be seen in one model (CNRM-CM6 with a high ECS), the CanESM5 model does not reproduce the IPSL-CM6A-LR features as it appears to have warming not compatible with the observations, and the MPI-ESM1.1 model, in which the members with an AMOC trend consistent with a recent reconstruction<sup>17</sup> are not those with the best representation of the four observed AMOC fingerprints analyzed in this study.

**Implications for future global warming and discussion.** Our analysis suggests that a fraction of anthropogenic warming might



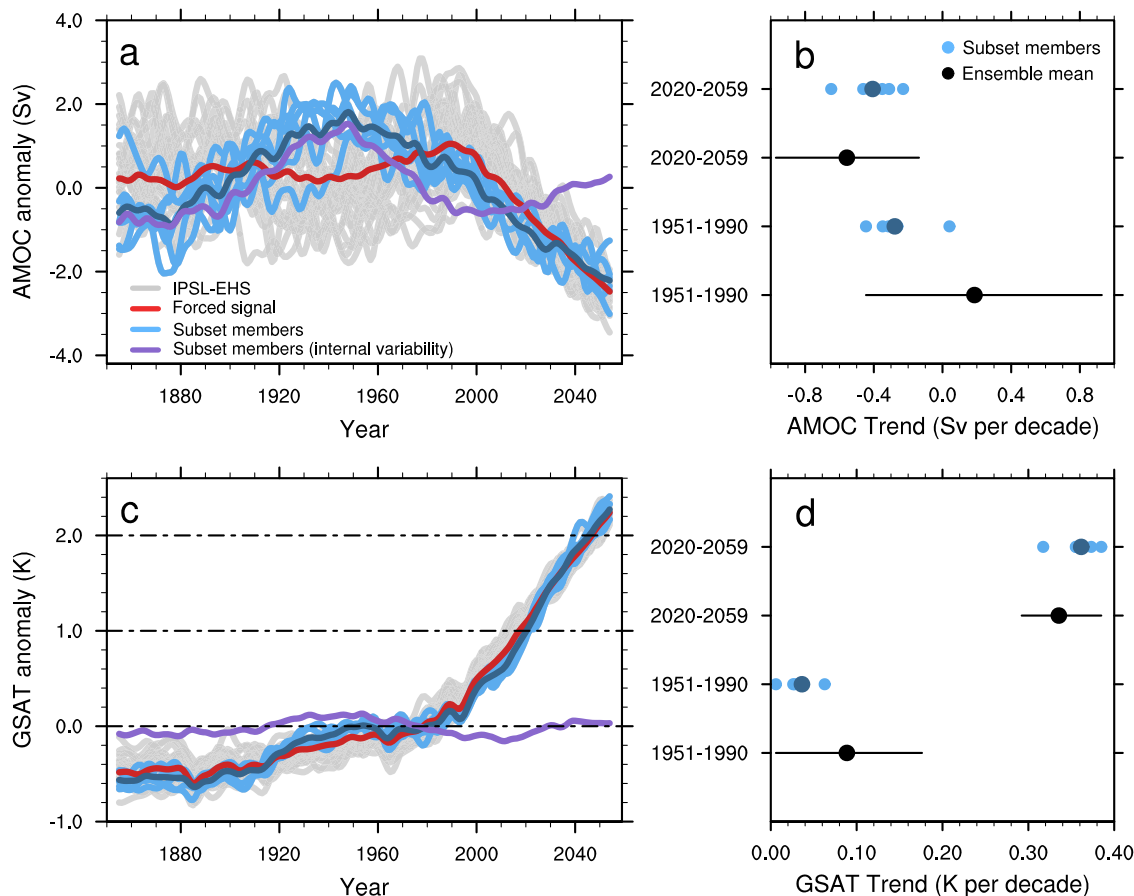
**Fig. 4 Relationship between GSAT and AMOC in CMIP6 control simulations.** **a** Scatter plot of Atlantic Meridional Overturning Circulation (AMOC) ( $\text{Sv decade}^{-1}$ ) versus Global near-Surface Air Temperature (GSAT) ( $\text{K decade}^{-1}$ ) trends calculated overall 60 year windows increasing by 10 years from the first 500 years of the pre-industrial control (piControl) simulations of thirteen CMIP6 models. The lines represent the least square regression between these two variables for each model. The regression coefficients ( $\text{K Sv}^{-1}$ ) are indicated in bold beside the model's name when the regression is significant ( $p < 0.1$ , see Methods). **b** Smoothed power spectra of the AMOC time series from the first 500 years of the piControl simulations of the same thirteen CMIP6 models (solid lines, thick when the regression is significant), with the 95% confidence limit estimated from a red noise (dashed lines). A Lanczos low-pass filter with a cutoff period of 5 years is used.

have been hidden by an AMOC weakening that started in the middle of the 20th century and is mainly related to internal climate variability. Indeed, the subset of members identified previously, which is characterized by a strong AMOC weakening over the 1951–1990 period in comparison to the IPSL-EHS (Fig. 5a, b) shows a lower warming over the same period relative to the full IPSL-EHS (Fig. 5c, d). The same subset of members experiences an internally-generated AMOC strengthening over the next few decades. This is associated with a larger warming rate of about 0.36 K per decade relative to the IPSL-EHS mean of 0.34 K per decade. This influence of a phase change in the AMOC is however limited, as both the AMOC and GSAT low-frequency internal variability are projected to decrease in response to external forcings<sup>33</sup>. Taken together, these results reinforce the risk of crossing the 2 °C warming objective. Indeed, our subset of members with a low-frequency variability consistent with the observations shows larger warming than the ensemble means over the next decades. Moreover, taking into account this AMOC weakening over the historical era reinforces the risk of faster warming, as a fraction of anthropogenic global warming, which could have been hidden by this low-frequency internal variability, is expected to materialize in the coming decades. These results thus seem to be in line with a recent study<sup>34</sup> suggesting that the pattern effect could be in part related to internal climate variability. This internally-driven pattern effect could have masked part of human-induced global warming in recent decades. Nevertheless, the time period investigated here is longer than the last four decades analyzed in this other study, and our mechanism is related to the North Atlantic variability, rather than the Pacific variability, as found in the previous studies<sup>35</sup>. Future work is required to quantify how accounting for North Atlantic low-frequency variability, both in the Atlantic and Pacific oceans, modifies observational constraints on future warming levels.

The realism of the multi-centennial low-frequency internal variability found in some of the CMIP6 models<sup>10,11</sup> is a crucial element of our results. Although the instrumental period is short,

previous studies suggested that the Atlantic Multidecadal Variability is underestimated in CMIP5 models, with a persistence lower than the one deduced from observations<sup>36,37</sup>. To gain insight on this realism, paleodata provides additional information, but it usually suffers from large uncertainties concerning quantified estimations of the variability, especially at large spatial scales. While it is usually believed that model simulations might have too low multi-centennial variability as compared to proxy records<sup>38</sup>, a recent study<sup>11</sup> suggests that GSAT interdecadal variability of some CMIP6 models might be overestimated over the period 1450–1840 in comparison to the pre-industrial control simulations. Therefore, we cannot exclude that some CMIP6 models, such as IPSL-CM6A-LR have too much internal variability. On the longer time scale of the last two thousand years, the different methods to reconstruct global mean surface temperature do show considerable uncertainty in terms of the magnitude of the multi-centennial variability<sup>39</sup>, while the external forcings and their impact remain poorly estimated in models. Therefore, it seems difficult at the moment to properly assess the realism of model simulations with those reconstructions. Such an evaluation of model simulations might deserve a dedicated analysis using the PAGES2K database and last millennium simulations from CMIP6, using advanced techniques like pseudo-proxy approaches<sup>40</sup> in order to compare model simulations and reconstructions in a coherent framework.

To conclude, we have shown here that the different indications of an AMOC weakening since the mid-20th century might be mainly of internal origin coming from multi-centennial variability of the ocean. If true, this might mean that transient climate sensitivity estimated from the observational records, especially over the last 6–7 decades may be underestimated. Thus, emergent constraint approaches that try to constrain future warming using the recent decades should fully embrace the issue of low-frequency internal variability and take into account individual ensemble members rather than ensemble means, as this might have crucial implications in terms of how different models are weighted in such studies.



**Fig. 5 Implication for near-future AMOC and GSAT change.** **a** Time evolution of the low-pass filtered Atlantic Meridional Overturning Circulation (AMOC) anomaly (Sv) relative to the 1900–2018 period for the IPSL ensemble of extended historical simulations (IPSL-EHS) (gray), the ensemble mean (red), the subset of members identified in Fig. 2 (member #14, 5, 25, 29, 4 and 30 in light blue), the subset mean (dark blue) and the mean internal variability (calculated by removing the ensemble mean for each member) of the subset of members (purple). **c** Same as **a** but for the Global near-Surface Air Temperature (GSAT) anomaly (K). **b** AMOC trends (Sv per decade) in IPSL-EHS (black), with the ensemble mean (black dot) and the full minimum–maximum range (black line) and the subset of members (light blue), with the subset mean (dark blue). **d** Same as **c** but for GSAT.

**Methods**

**Observational data.** We used the infilled Cowtan and Way<sup>41,42</sup>, named HadCRUT4-CW in this study, observational surface temperature, which avoids the problem of missing values. This dataset is composed of blending sea surface temperature over the ocean and near-surface air temperature over the land. This metric of global mean temperature, however, warms significantly less than the GSAT. To take into account this discrepancy, we applied a factor of 1.06 to the HadCRUT4-CW surface temperature, as used in a recent study<sup>10</sup>. This factor is similar to those of 1.05<sup>43</sup>, 1.06<sup>44</sup>, and 1.08<sup>5</sup> used in several other studies. One of the interests of this dataset is to provide an ensemble of realizations, allowing to consider the associated uncertainties. Here, the median of the dataset is considered the best estimate, and the maximum and the minimum is used for the uncertainty. Note that the HadCRUT4-CW dataset is very consistent with the Berkeley<sup>45</sup> observational surface air temperature, with a Pearson correlation coefficient between the annual GSAT values of 0.99. Very small differences are found with the infilled HadCRUT5 dataset when accounting for the missing values of this dataset. Therefore, we keep the HadCRUT4-CW in order to have a global coverage without missing values.

**Simulations used.** The atmospheric resolution of IPSL-CM6A-LR is 1.26° × 1.25° with 79 levels (model top at 1 Pa). The model uses a nominal resolution of 1° in the ocean and 75 levels. The IPSL-CM6A-LR ensemble<sup>12</sup> used in this study follows the CMIP6 protocol<sup>46</sup> for historical simulations for the period 1850–2014. The historical simulations initial conditions are taken from different years in a long pre-industrial simulation after it has reached a quasi-stationary state. Specifically, the simulations are started from the atmospheric, oceanic, and land surface initial conditions of the 1<sup>st</sup> January from different years. The simulations were extended until 2060 using all forcings from the SSP245 scenario<sup>47</sup>, 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 “see” the ozone hole recovery and changes in tropospheric

ozone, in contrast to the official CMIP6 projections. However, this is seen as a minor shortcoming for the purpose of this study. All the simulations used are summarized in Supplementary Table 1.

**ECS and TCR estimates.** The Sensitivity and the Transient climate response calculated over the historical period, named as S\_hist and TCR\_hist in this study, are defined as:

$$S_{hist} = \frac{F_{2*CO_2} * \Delta T}{\Delta F - \Delta Q} \tag{1}$$

$$TCR_{hist} = \frac{F_{2*CO_2} * \Delta T}{\Delta F} \tag{2}$$

with  $F_{2*CO_2} = 3.77 \text{ Wm}^{-2}$  being the radiative forcing due to a doubling of the atmospheric CO<sub>2</sub> concentration. This latter quantity is calculated by averaging the Effective Radiative Forcing (ERF) values from three different experiments performed with IPSL-CM6A-LR, a fixed SST experiment<sup>48</sup> and two experiments based on this method<sup>49</sup>, using linear regressions of the top-of-atmosphere net radiative flux against surface temperature carried out over the first 20 years of the abrupt experiment, with a subtraction of the corresponding piControl quantities for both, but based on a year-to-year basis for the first case or using 20-year climatologies for the second case. We estimate the uncertainty related to using the lowest (3.50 W m<sup>-2</sup>) and largest (3.94 W m<sup>-2</sup>) values of the three experiments. The net ERF is calculated between the pre-industrial (1850–1879) and present-day (1999–2018) periods by averaging the ERFs of the three members of piClim-histall RFMIP experiments and is estimated at  $\Delta F = 2.09 \text{ Wm}^{-2}$ .  $\Delta T$  is the change in GSAT between the same periods, and  $\Delta Q$  the rate of the total increase in Earth system heat content as diagnosed in the model, calculated over the 1999–2018 period. Here, we approximate the Earth System Heat Content by calculating  $\Delta Q$  from the Ocean Heat Content increase over the historical period divided by 0.89, as the oceans store about 89% of the excess heat<sup>50</sup>. This allows us to take into account an

estimation of the continental and atmospheric heat uptake, as well as the heat uptake owing to ice melt.  $\Delta F$  does not change much across the members so we use a single value. However both  $\Delta T$  and  $\Delta Q$  vary more substantially and are estimated separately for each member.

Although the piControl simulation inherits from a long millennial spin-up during the development process, the model is not completely equilibrated given the long timescale associated with the deep ocean. As a result, a very small drift is present in the piControl simulations, as it is often the case with climate models<sup>51</sup>. We diagnose the drifts in the piControl for OHC ( $-1.46 \times 10^{21} \text{ J yr}^{-1}$ ) and GSAT ( $-0.00011 \text{ K yr}^{-1}$ ) and remove it in the historical simulations prior to the  $\Delta Q$  and  $\Delta T$  calculation.

**Calculation and analysis of the AMOC index.** The AMOC is evaluated using the maximum of the annual Atlantic meridional stream function at  $20^\circ\text{N}$ – $50^\circ\text{N}$ . We consider CMIP6 models with Atlantic meridional stream function data available in the piControl simulation for at least 500 years. Thirteen models, including IPSL-CM6A-LR, were available at the time of our analysis.

**Scaling of the internal versus externally-forced AMOC variation.** From the IPSL ensemble, we define the forced (or externally-driven) response of the AMOC as the ensemble mean of AMOC variations. This method is appropriate because the initial conditions of the 32 ensemble members started every 20 years from the piControl simulation to sample various phases of the multi-centennial variability. Then, for each member, the internal (or internally-driven) variability is obtained as the difference between that member and the forced response. To estimate the forced and unforced components in the AMOC strength variations for each member of the IPSL ensemble, we used a linear model as follows:

$$y_{AMOC \text{ unforced}} = \lambda_{\text{unforced}} x_{AMOC} + \epsilon_{\text{unforced}} \quad (3)$$

$$y_{AMOC \text{ forced}} = \lambda_{\text{forced}} x_{AMOC} + \epsilon_{\text{forced}} \quad (4)$$

where  $x_{AMOC}$  is the same in both equations and designate the AMOC time series over the 1940–2016 period,  $y_{AMOC \text{ unforced}}$  the AMOC internal variability and  $y_{AMOC \text{ forced}}$  the AMOC forced signal while the  $\lambda$  are the regression coefficients and  $\epsilon$  the error terms. When both  $\lambda$  values are significantly positive,  $\lambda$  estimates the respective roles of internal variability and forced changes in driving the simulated AMOC variations (e.g., Supplementary Fig. 1).

**Calculation of the AMV index.** The AMV index is calculated as the mean SST in the North Atlantic Ocean between  $0^\circ\text{N}$  and  $60^\circ\text{N}$ . How to remove the effect of external forcings in order to study the AMV due to internal processes from the observations and the historical simulations is a thorny issue. A good way to do so in ensemble simulations is to remove the ensemble mean. However, this cannot be done in observations for which we only have one realisation. Therefore, when comparing the simulated AMV to the observations, as in Fig. 3b, the forced signal is estimated from the average SST between  $60^\circ\text{S}$  and  $60^\circ\text{N}$ <sup>52</sup>. In both observation and model<sup>12</sup>, the AMV spatial pattern 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. A Lanczos low-pass filter with a cutoff period of 11 years is then used to retain only the low-frequency variations.

**Significance testing.** The significance of the relationship between the AMOC trends and the TCR\_hist and TCR\_hist and the regression calculated in Fig. 2a are estimated with a two-tailed Student t-test considering a  $p$ -value  $< 0.1$  and 30 degrees of freedom. For the significance of the regression between GSAT and AMOC trends in the piControl simulations (Fig. 4a), we used a nonparametric method<sup>53</sup> based on a random phase of resampling, using 1000 surrogates and  $p$ -value  $< 0.1$ , in order to take into account the serial correlation, as the trends overlap. The same method is used to evaluate the significance of the correlations between AMV and AMOC time series in IPSL-EHS and between the four observed AMOC fingerprints and the member #14 in the third section.

## Data availability

All of the observational datasets used are publicly available online. The CMIP6 model outputs from the piControl and historical simulations are available through the Earth System Grid Federation (ESGF) portal (<https://esgf-node.llnl.gov/projects/cmip6/>). The data based on the extended (beyond 2014) historical simulations from the IPSL-CM6A-LR model used in the main figures of this study are available in this repository: <https://doi.org/10.5281/zenodo.5159426>.

## Code availability

The codes used for the analyses are available from the authors upon request.

Received: 30 March 2021; Accepted: 28 September 2021;

Published online: 20 October 2021

## References

- Forster, P. M., Maycock, A. C., McKenna, C. M. & Smith, C. J. Latest climate models confirm need for urgent mitigation. *Nat. Clim. Change* **10**, 7–10 (2020).
- Nijse, F. J. M. M., Cox, P. M. & Williamson, M. S. Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS) from historical warming in CMIP5 and CMIP6 models. *Earth Syst. Dyn.* **11**, 737–750 (2020).
- Otto, A. et al. Energy budget constraints on climate response. *Nat. Geosci.* **6**, 415–416 (2013).
- Lewis, N. & Curry, J. The impact of recent forcing and ocean heat uptake data on estimates of climate sensitivity. *J. Clim.* **31**, 6051–6071 (2018).
- Richardson, M., Cowtan, K., Hawkins, E. & Stolpe, M. B. Reconciled climate response estimates from climate models and the energy budget of Earth. *Nat. Clim. Change* **6**, 931–935 (2016).
- Sherwood, S. et al. An assessment of Earth's climate sensitivity using multiple lines of evidence. *Rev. Geophys.* <https://doi.org/10.1029/2019RG000678> (2020).
- Annan, J. D., Hargreaves, J. C., Mauritsen, T. & Stevens, B. What could we learn about climate sensitivity from variability in the surface temperature record? *Earth Syst. Dyn.* **11**, 709–719 (2020).
- Tokarska, K. B. et al. Past warming trend constrains future warming in CMIP6 models. *Sci. Adv.* **6**, eaaz9549 (2020).
- Liang, Y., Gillett, N. P. & Monahan, A. H. Climate model projections of 21st century global warming constrained using the observed warming trend. *Geophys. Res. Lett.* **47**, e2019GL086757 (2020).
- Ribes, A., Qasmi, S. & Gillett, N. P. Making climate projections conditional on historical observations. *Sci. Adv.* **7**, eabc0671 (2021).
- Parsons, L. A., Brennan, M. K., Wills, R. C. J. & Proistosescu, C. Magnitudes and spatial patterns of interdecadal temperature variability in CMIP6. *Geophys. Res. Lett.* **47**, e2019GL086588 (2020).
- Bonnet, R. et al. Presentation and evaluation of the IPSL-CM6A-LR ensemble of extended historical simulations. *J. Adv. Model. Earth Syst.* **13**, e2021MS002565 (2021).
- Boucher, O. et al. Presentation and evaluation of the IPSL-CM6A-LR climate model. *J. Adv. Model. Earth Syst.* **12**, e2019MS002010 (2020).
- Rugenstein, M. et al. Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophys. Res. Lett.* **47**, e2019GL083898 (2020).
- Knutti, R., Rugenstein, M. A. A. & Hegerl, G. C. Beyond equilibrium climate sensitivity. *Nat. Geosci.* **10**, 727–736 (2017).
- Jiang, W., Gastineau, G. & Codron, F. Multicentennial variability driven by salinity exchanges between the Atlantic and the Arctic ocean in a coupled climate model. *J. Adv. Model. Earth Syst.* **13**, e2020MS002366 (2021).
- Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G. & Saba, V. Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature* **556**, 191–196 (2018).
- Caesar, L., McCarthy, G. D., Thornalley, D. J. R., Cahill, N. & Rahmstorf, S. Current Atlantic meridional overturning circulation weakest in last millennium. *Nat. Geosci.* **14**, 118–120 (2021).
- Maroon, E. A., Kay, J. E. & Karnauskas, K. B. Influence of the Atlantic meridional overturning circulation on the Northern Hemisphere surface temperature response to radiative forcing. *J. Clim.* **31**, 9207–9224 (2018).
- Stolpe, M. B., Medhaug, I., Sedláček, J. & Knutti, R. Multidecadal variability in global surface temperatures related to the Atlantic meridional overturning circulation. *J. Clim.* **31**, 2889–2906 (2018).
- Garuba, O. A., Lu, J., Liu, F. & Singh, H. A. The active role of the ocean in the temporal evolution of climate sensitivity. *Geophys. Res. Lett.* **45**, 306–315 (2018).
- Menary, M. B. et al. Aerosol-forced AMOC changes in CMIP6 historical simulations. *Geophys. Res. Lett.* **47**, e2020GL088166 (2020).
- Hegerl, G. C. et al. Causes of climate change over the historical record. *Environ. Res. Lett.* **14**, 123006 (2019).
- Schlesinger, M. E. & Ramankutty, N. An oscillation in the global climate system of period 65–70 years. *Nature* **367**, 723–726 (1994).
- Kerr, R. A. A North Atlantic climate pacemaker for the centuries. *Science* **288**, 1984–1985 (2000).
- Enfield, D. B., Mestas-Núñez, A. M. & Trimble, P. J. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.* **28**, 2077–2080 (2001).
- Delworth, T., Manabe, S. & Stouffer, R. J. Interdecadal variations of the thermohaline circulation in a coupled Ocean-Atmosphere model. *J. Clim.* **6**, 1993–2011 (1993).
- Steinman, B. A., Mann, M. E. & Miller, S. K. Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures. *Science* **347**, 988–991 (2015).
- Qin, M., Dai, A. & Hua, W. Quantifying contributions of internal variability and external forcing to Atlantic multidecadal variability since 1870. *Geophys. Res. Lett.* **47**, e2020GL089504 (2020).



30. Zhang, R. Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophys. Res. Lett.* **35**, 035463 (2008).
31. Marshall, J., Donohoe, A., Ferreira, D. & McGee, D. The ocean's role in setting the mean position of the Inter-Tropical Convergence Zone. *Clim. Dyn.* **42**, 1967–1979 (2014).
32. Li, Z. et al. A robust relationship between multidecadal global warming rate variations and the Atlantic Multidecadal Variability. *Clim. Dyn.* **55**, 1945–1959 (2020).
33. Cheng, J. et al. Reduced interdecadal variability of Atlantic meridional overturning circulation under global warming. *Proc. Natl Acad. Sci.* **113**, 3175–3178 (2016).
34. Zhou, C., Zelinka, M. D., Dessler, A. E. & Wang, M. Greater committed warming after accounting for the pattern effect. *Nat. Clim. Change* **11**, 132–136 (2021).
35. Dong, Y., Proistosescu, C., Armour, K. C. & Battisti, D. S. Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a green's function approach: the preeminence of the western pacific. *J. Clim.* **32**, 5471–5491 (2019).
36. Cheung, A. H. et al. Comparison of low-frequency internal climate variability in CMIP5 models and observations. *J. Clim.* **30**, 4763–4776 (2017).
37. Qasmi, S., Cassou, C. & Boé, J. Teleconnection between atlantic multidecadal variability and european temperature: diversity and evaluation of the coupled model intercomparison project phase 5 models. *Geophys. Res. Lett.* **44**, 11,140–11,149 (2017).
38. Laepple, T. & Huybers, P. Ocean surface temperature variability: Large model–data differences at decadal and longer periods. *Proc. Natl Acad. Sci.* **111**, 16682–16687 (2014).
39. Neukom, R. et al. Consistent multi-decadal variability in global temperature reconstructions and simulations over the Common Era. *Nat. Geosci.* **12**, 643–649 (2019).
40. Ortega, P. et al. A model-tested North Atlantic Oscillation reconstruction for the past millennium. *Nature* **523**, 71–74 (2015).
41. Jones, P. D. et al. Hemispheric and large-scale land-surface air temperature variations: an extensive revision and an update to 2010. *J. Geophys. Res. Atmospheres* **117** (2012).
42. Cowtan, K. & Way, R. G. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Q. J. R. Meteorol. Soc.* **140**, 1935–1944 (2014).
43. Cowtan, K. et al. Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophys. Res. Lett.* **42**, 6526–6534 (2015).
44. Gillett, N. P. et al. Constraining human contributions to observed warming since the pre-industrial period. *Nat. Clim. Change* **11**, 207–212 (2021).
45. Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S. & Mosher, S. Berkeley Earth temperature averaging process. *Geoinformatics Geostat. Overv.* **01**, 1000103 (2013).
46. Eyring, V. et al. Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
47. Gidden, M. J. et al. Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. *Geosci. Model Dev.* **12**, 1443–1475 (2019).
48. Forster, P. M. et al. Recommendations for diagnosing effective radiative forcing from climate models for CMIP6. *J. Geophys. Res. Atmospheres* **121**, 12,460–12,475 (2016).
49. Gregory, J. M. et al. A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.* **31**, 01847 (2004).
50. von Schuckmann, K. et al. Heat stored in the Earth system: where does the energy go? *Earth Syst. Sci. Data* **12**, 2013–2041 (2020).
51. Hobbs, W., Palmer, M. D. & Monselesan, D. An energy conservation analysis of ocean drift in the CMIP5 global coupled models. *J. Clim.* **29**, 1639–1653 (2016).
52. Trenberth, K. E. & Shea, D. J. Atlantic hurricanes and natural variability in 2005. *Geophys. Res. Lett.* **33**, 026894 (2006).
53. Ebisuzaki, W. A method to estimate the statistical significance of a correlation when the data are serially correlated. *J. Clim.* **10**, 2147–2153 (1997).
54. Huang, B. et al. Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J. Clim.* **30**, 8179–8205 (2017).
55. Cheng, L. & Zhu, J. Benefits of CMIP5 multimodel ensemble in reconstructing historical ocean subsurface temperature Variations. *J. Clim.* **29**, 5393–5416 (2016).
56. Cheng, L. et al. Improved estimates of ocean heat content from 1960 to 2015. *Sci. Adv.* **3**, e1601545 (2017).

## Acknowledgements

The IPSL-CM6 experiments were performed using the HPC resources of TGCC under the allocations 2017-R0040110492, 2018-R0040110492, and 2019-A0060107732 (project gencmip6) provided by GENCI (Grand Equipement National de Calcul Intensif). The IPSL-CM6 team is acknowledged for data curation of the CMIP6 model output. RB and OB were supported by the European Union's Horizon 2020 research and innovation programme under grant agreement number 820829 for the "Constraining uncertainty of multi-decadal climate projections (CONSTRAIN)" project (07/2019-06/2024). JM, GG, and DS were supported by Blue-Action (European Union's Horizon 2020 research and innovation program under grant agreement no 727852) and EUCP (European Union's Horizon 2020 research and innovation programme under grant agreement no 776613) projects. This work also 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.

## Author contributions

R.B., D.S., O.B., and J.M. designed the study. R.B. performed all of the IPSL-EHS and the multi-model analysis against observations. G.G. computed the A.M.O.C. index based on the O.H.C. AS provided the radiative forcing estimates. RB wrote the paper, and O.B., D.S., J.M., J.D., G.G., F.H., and J.S. contributed to the text.

## Competing interests

The authors declare no competing interest.

## Additional information


**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-26370-0>.

**Correspondence** and requests for materials should be addressed to Rémy Bonnet.

**Peer review information** *Nature Communications* thanks Nathan Gillett, Aixue Hu and the other anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021