

RESEARCH ARTICLE

No critical slowing down in the Atlantic Overturning Circulation in historical CMIP6 simulations

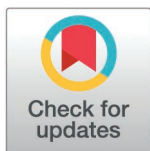
Maya Ben-Yami^{1,2*}, Lana Blaschke^{1,2}, Sebastian Bathiany^{1,2}, Niklas Boers^{1,2,3}

1 Complexity Science, Potsdam Institute for Climate Impact Research, Potsdam, Brandenburg, Germany,

2 Munich Climate Center and Earth System Modelling Group, Department of Aerospace and Geodesy, TUM School of Engineering and Design, Technical University of Munich, Ottobrunn, Bavaria, Germany

3 Department of Mathematics and Global Systems Institute, University of Exeter, Exeter, United Kingdom

* mayayami@pik-potsdam.de



OPEN ACCESS

Citation: Ben-Yami M, Blaschke L, Bathiany S, Boers N (2026) No critical slowing down in the Atlantic Overturning Circulation in historical CMIP6 simulations. PLOS Clim 5(10): e0000781. <https://doi.org/10.1371/journal.pclm.0000781>

Editor: Juan A. Añel, Universidade de Vigo, SPAIN

Received: August 12, 2025

Accepted: May 5, 2026

Published: October 7, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pclm.0000781>

Copyright: © 2026 Ben-Yami et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

Abstract

The Atlantic Meridional Overturning Circulation (AMOC) is a key component of the Earth's climate system, and has been suggested to have multiple stable states. Critical slowing down (CSD) can detect stability changes in Earth system components, and has been found in sea-surface temperature (SST) based fingerprints of the AMOC. Here, we look for CSD in historical simulations from 27 models from the sixth Climate Model Intercomparison Project (CMIP6). We calculate three different CSD indicators for the AMOC streamfunction strengths at 26.5°N and 35°N, as well as for a previously suggested SST-based AMOC index (ASSTI) based on averaging SSTs in the subpolar gyre region. No model shows CSD in the ASSTI, which is in marked disagreement with observations. This lack of CSD is reflected in the AMOC streamfunctions in most models, although individual ensemble members in some models do show signs of CSD even under a conservative significance calculation. We thus conclude that: 1) The historical AMOC in CMIP6 models is not losing stability, 2) studies of AMOC stability should consider an ensemble of realisations, 3) there is no non-AMOC-related physical process in CMIP6 models which could cause signs of CSD in North-Atlantic SSTs in the 1850–2014 period, and thus the CSD in the observed ASSTI is likely a sign of a change in the AMOC. This final result suggests that observed changes in the ASSTI could indicate a loss of stability in the real-world AMOC.

Introduction

The Atlantic Meridional Overturning Circulation (AMOC) is a vital part of the Earth's climate system. It consists of a system of currents that transport large amounts of heat and salt northward near the surface of the Atlantic Ocean. The AMOC is predicted to weaken under anthropogenic global warming [1], with substantial impacts

and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All CMIP6 data used in this study is available online on the ESGF database (<https://aims2.llnl.gov>). The HadCRUT5 dataset is available at <https://www.metoffice.gov.uk/hadobs/>. The ERSSTv5 operational dataset is available at <https://psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html>, and the ERSSTv5 uncertainties are available at <https://www.ncei.noaa.gov/pub/data/cmb/ersst/v5/tmp/>. No new data has been produced. The code used to analyse the data and generate figures is available at https://github.com/mayaby/PLOS_CSD.

Funding: This work was supported by the European Union's Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement No.956170 to MBY and NB, and grant agreement No.820970 to LB), the Volkswagen foundation (NB and SB), the European Union's Horizon Europe research and innovation programme (grant agreement No. 101137601 to NB and SB), the European Space Agency Climate Change Initiative (ESA-CCI) Tipping Elements SIRENE project (contract no. 4000146954/24/I-LR to LB), the Framework of international cooperation program by Deutsche Forschungsgemeinschaft (DFG) and National Research Foundation of Korea (NRF): Climate Resilience under Zero-Emission Commitment Scenarios (BO 4455/2-1; eBer-24-62950; RS-2024-00438471), and the Advanced Research + Invention Agency (ARIA) (funding reference SCOP-PR01-P003 - Advancing Tipping Point Early Warning AdvanTip). This is ClimTip contribution #165; the ClimTip project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101137601: Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: Niklas Boers is Editor-in-Chief at PLOS climate. There are no other competing interests.

on the global climate [2,3]. There is some observational and paleoclimate evidence that the AMOC has already weakened [4,5], although the decline is controversial and likely still consistent with natural variability [6–8]. The AMOC has also been identified as a potential climate tipping element [9,10], with support for bistability coming from paleoclimate proxy records [11], theoretical studies [12,13], as well as from experiments with some general circulation models [14–18]. However, no alternative stable states have been found in other models, and the question of such states in comprehensive climate models remains debated [19]. The AMOC's bistability is theorised to arise from a key positive feedback in the system, the salt advection feedback. A weaker AMOC leads to less northward salt transport and thus reduced surface density in the regions of North Atlantic deep water (NADW) formation. The reduction in NADW formation in turn leads to an even weaker AMOC, continuing until an alternative weak AMOC state is reached. A question of intense debate is whether or not the AMOC could undergo such an abrupt transition under future anthropogenic climate change [9,10]. The 6th assessment report (AR6) of the International Panel on Climate Change (IPCC) concludes that there is medium confidence that such a collapse will not occur before 2100 [20]. This conclusion is in large part based on results from complex climate models, and in particular on the experiments of the sixth Climate Model Intercomparison Projects (CMIP6). Some of these experiments have since been extended past 2100, and Drijfhout et al. 2025 [21] have shown that in high emission scenarios the AMOC consistently weakens to almost zero by 2300, although it should be noted that for most models there is no evidence that this weakening is a state transition and not a linear damping.

In this work we will call a weakening of the AMOC to almost zero strength an AMOC “collapse,” with the understanding that this collapse may be linear and does not necessarily mean the AMOC has tipped to an alternative state. The complexity and extent of the AMOC makes it challenging to simulate, and complex models can fail to capture key processes which could lead to a warming-induced collapse [16,22]. When the small number of older experiments that modelled a global-warming induced AMOC collapses are used to estimate the temperatures at which a collapse might happen, these range from 1.4 to 8 °C, in practice encompassing the full scope of future warming [9].

Another approach to understanding the stability of the AMOC is to use its characteristic statistical properties from the perspective of nonlinear dynamical systems. In particular, when a system approaches a bifurcation-induced transition to a different stable state, it undergoes a process called Critical Slowing Down (CSD), in which the rate of the system's return to equilibrium after perturbations decreases and eventually approaches zero. Statistical indicators of CSD, such as increasing variance or autocorrelation, are a measure of the stability of the system. In extreme cases, they can signal an approaching transition, and have thus sometimes been called early-warning signals [23]. Note that while non-linearity is a necessary condition for slowing down, as long as the linear stability weakens, CSD can still occur in systems even if they are not approaching a bifurcation point [24], although it is still uncertain how likely such a scenario is for the AMOC [25,26]. Boers 2021 [27] (hereafter B21) studied sea

surface temperature (SST)- and salinity-based observational proxies of the historical AMOC strength, and found a significant signal of CSD in these proxies. This result has often been interpreted as a sign that the real-world AMOC is losing stability and approaching a tipping point, and some studies have attempted to use this CSD signal to predict the AMOC's future tipping time [28], although concerns have been raised regarding the large uncertainties of such predictions [29].

In particular, the applicability of CSD to real-world systems is still under debate [30], as there are manifold uncertainties associated with using CSD signals to warn of real-world tipping points. Some of these stem from the theoretical assumption that the system is capable of tipping, but there are also many uncertainties that arise from our imperfect measurements of the Earth's climate [31]. Notably, continuous direct measurements of AMOC strength have only been available since 2004 [32], and thus any attempt to understand the AMOC on longer timescales is restricted to so-called fingerprints. Fingerprints are patterns derived from long-term observational records of variables that are in some way physically connected to the AMOC. While fingerprints in the literature use a range of variables, from sea-surface heat fluxes [8] to ocean heat content [33], the long time series length required for CSD detection constrains that analysis to fingerprints based on SST and salinity datasets. The most commonly used AMOC SST fingerprint was introduced by Caesar et al. 2018 [34], and averages SSTs from the subpolar gyre (SPG) area south of Greenland, normalized by subtracting the SST's global mean. Subsequent works have used different versions of this fingerprint, for example using a more general SPG area [35] or subtracting twice the global mean [28]. In this work, it will be called the AMOC SST Index (ASSTI) for clarity. This is the fingerprint that, among others, was used by Boers 2021 [27] to detect AMOC CSD, and by Ditlevsen & Ditlevsen 2023 [28] to predict an AMOC tipping time. The ASSTI has also been used as an emergent constraint for CMIP6 models, both to constrain AMOC strength and variability [36] and to constrain Arctic Ocean changes [37]. While Ben-Yami et al. 2023 [31] showed that the CSD in this fingerprint is still significant when considering the uncertainties stemming from the SST datasets, there has been increasing evidence in recent years that the ASSTI is only partially correlated with the AMOC streamfunction [5,8], and that this correlation is time and scenario dependent [38,39]. In particular for CSD, van Westen et al. 2024 [18] found no significant variance or autocorrelation increase preceding a hosing-induced AMOC tipping in CESM1. It should be noted, however, that both the forcing and the background climate in this experiment are different from the historical period. Although this body of evidence does indicate the need for the development of better AMOC fingerprints, it does not necessarily mean that we should disregard the CSD seen in the observed ASSTI. Indeed, there is a physical reason to suspect that losses in the AMOC's stability would mean slower recovery from perturbations in the SPG, as the SPG is connected to the AMOC through multiple different mechanisms [35,40,41]. However, the uncertain correlation between the ASSTI and the AMOC streamfunction leaves room for other physical processes influencing the changes in the ASSTI, for example changes in the mixed layer depth or atmospheric processes [42–45], and these changes may be causing the observed statistical changes in the ASSTI without those changes being mirrored in the AMOC itself. In this work, we will call such a scenario a “physical false positive” (as opposed to a statistical false positive caused by, e.g., dataset processing methods), as in this case the CSD signal is truly present in the SSTs, but does not indicate an AMOC slowdown. This is the main question we ask in this work — how certain are we that the changes we see in the observed ASSTI indicate changes in the AMOC?

There are multiple ways to approach this question. One way is to develop AMOC fingerprints with a higher correlation to the AMOC streamfunction, as is already being attempted in the literature [46,47]. Another is to look at scenarios where the AMOC tips due to global warming, and see if the ASSTI warns of this collapse. However, even then there is still the possibility of a physical false CSD warning coinciding with an AMOC collapse — to truly test the validity of ASSTI CSD, one would need additional control experiments in the same model with no collapse. And furthermore, a warming-induced AMOC collapse is rare in models, and only happens in CMIP6 models if the simulations are extended past 2100 [21]. Not only is it uncertain whether the AMOC collapse (weakening) seen in these models is a tipping or simply a linear damping, but the number of such extended simulations is low. Since we cannot be sure if the collapse is truly non-linear, a lack of CSD warning in these models could either mean that the collapse is linear, or that the fingerprint is not useful, and thus is

not a useful test of the fingerprint. In addition, the small number of models is an issue because the AMOC is a difficult system to model [22], and so its response and its connection to North Atlantic SSTs varies strongly across models (see range of correlations in Emirzade et al. 2025 [46]). Any model-based study aiming to make robust statements on the AMOC must therefore use a large number of models, each with multiple ensemble members.

These limitations lead to the approach we take in this work — instead of looking for an AMOC-collapse test for the ASSTI, we study what models can tell us about the period in which the real-world ASSTI CSD has been observed, namely 1850 to present. For this we use the CMIP6 historical simulations, and look for CSD both in the ASSTI and the AMOC streamfunction strength. This approach not only provides us with a large number of models and ensemble member runs, but also gives us three different angles of study: First, the results for the ASSTI can be directly compared to observations in a standard model evaluation approach. Second, the lack or presence of CSD in the AMOC streamfunction can tell us what is happening with the historical AMOC in these CMIP6 models. Finally, the results for the ASSTI and the streamfunction can be compared, to see if there is any relationship in their CSD indicator trends. To illustrate what this final comparison can tell us, we show the four possible combinations of ASSTI and streamfunction signals in Fig 1. Since for most models we do not know if the AMOC is truly approaching a critical transition, in this work we will call the ASSTI a reliable

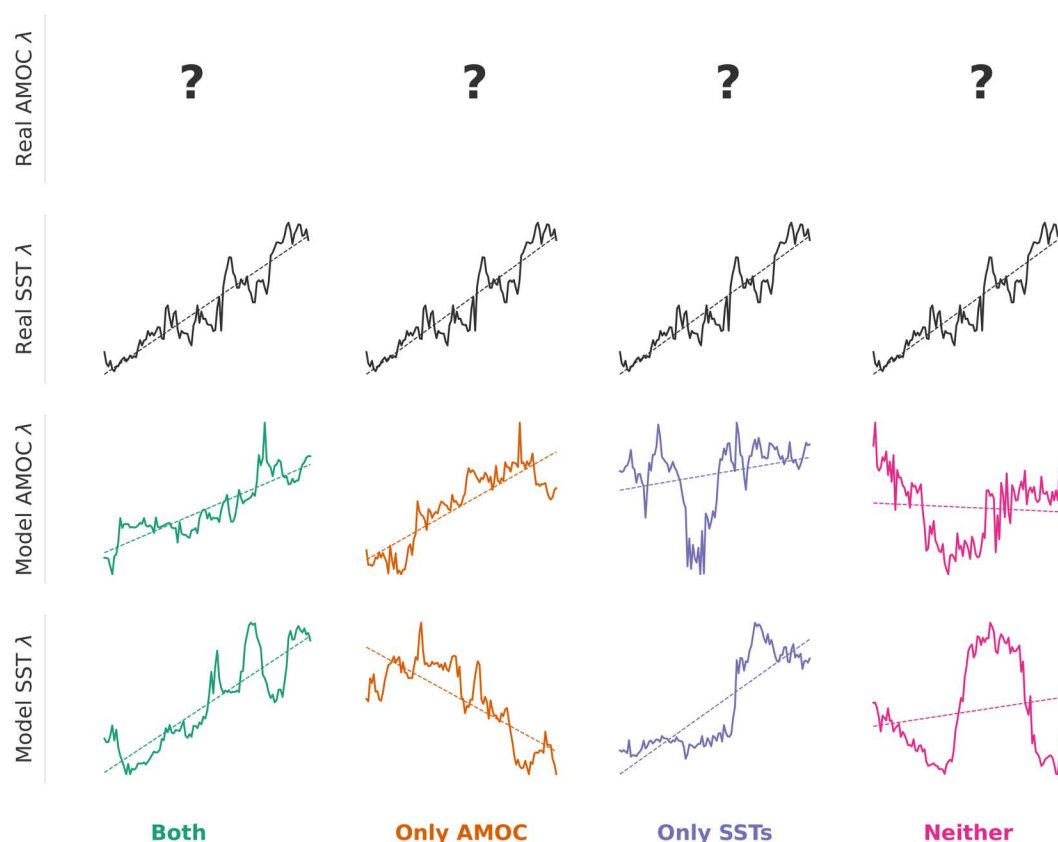


Fig 1. The four different possible results of our analysis. The rows, from top to bottom, represent the CSD indicator, restoring rate (λ) time series of: the real-world AMOC streamfunction (unknown), the real-world observed AMOC SST Index, the CMIP6 model AMOC streamfunction and the CMIP6 model AMOC SST Index. An increasing λ time series is a sign of CSD (destabilization). The columns represent four different scenarios for the model time series, namely, from left to right: both the streamfunction and the AMOC SST Index have a significant restoring rate increase (and so a sign of CSD) (turquoise), only the model streamfunction has a significant restoring rate increase (orange), only the model AMOC SST Index has a significant restoring rate increase (purple), and neither the model AMOC streamfunction nor the model AMOC SST Index have a significant restoring rate increase (pink).

<https://doi.org/10.1371/journal.pclm.0000781.g001>

fingerprint for detecting AMOC CSD if its CSD indicator changes follow those in the AMOC streamfunction strength — that is, if most of the runs are in either column 1 or column 4 of [Fig 1](#). Together with the other angles of study, the CMIP6 historical runs can thus help us better understand what is happening in the real-world Atlantic Ocean, even if we are not sure whether these runs end up in a tipped AMOC when projected into the future.

Methods

This work uses the historical simulations from the sixth Coupled Model Intercomparison Project (CMIP6). These experiments are initialised from different points in the model's pre-industrial control run (piControl), and are then forced with the natural and anthropogenic forcings of 1850–2014. For the AMOC streamfunction, we use the values calculated by Menary et al. 2020 [48], where the overturning streamfunction is calculated directly from the oceanic meridional velocities. For the streamfunction strength we take the maximum strength of overturning at latitudes 35°N and 26.5°N to represent the subpolar and subtropical AMOC. In the following, these time series will be abbreviated as S35 and S26.5, respectively. The AMOC SST Index is calculated as the mean SSTs in the area between 41° N and 60° N and 20° E and 55° E, minus the mean global SSTs, and will be abbreviated as ASSTI. The list of CMIP6 models used and their ensemble members can be found in Table A in [S4 Appendix](#) — there are 27 models and 133 ensemble members in total (the number of ensemble members per model varies). As we calculate both the ASSTI and the streamfunction strength at two latitudes, this totals to 399 AMOC time series.

Following B21, the main CSD indicator we will use is the restoring rate λ . Whilst it has been more common to use the variance and autocorrelation of a time series as CSD indicators, they can result in false positives or negatives when the statistics of the external conditions change. The restoring rate λ as introduced by B21 is estimated under the assumption of non-stationary correlated noise driving the system, and thus avoids biases induced by non-stationary noise (see [S1 Appendix](#)). When a system is in a stable equilibrium state, λ is negative, and it increases toward 0 from below as a critical transition is approached. In addition to λ , we also calculate the variance and lag-1 autocorrelation of the time series as additional CSD indicators. Thus, for each model ensemble member we have nine indicator time series — three AMOC time series times three CSD indicators. For 133 ensemble members, this totals to 1197 CSD indicator time series. In the main text, all indicators are calculated from the detrended time series in 60-year rolling windows, but the results of using other window lengths are shown in [S3 Appendix](#).

Since we are dealing with a large number of time series, even if these time series were completely random with no CSD, λ would still increase by chance sometimes. We thus need to look beyond a single λ increase in one time series, and find additional information that identifies what we see as a robust and significant sign of CSD. We approach this problem from two different angles. The first is that in the case of true AMOC CSD, it is likely that both streamfunction strengths at the two latitude bands, and possibly the AMOC SST Index, will have an increase in λ , as well as in the other two classical CSD indicators (variance and autocorrelation). Thus, to find a truly robust sign of CSD we need to see an increasing trend in all nine time series for a given ensemble member. This will be discussed further in the Results section. The second angle is statistical significance, which is tested individually for each time series: in particular, our null model is a completely random time series with the same variance and autocorrelation function as the AMOC time series. To calculate this significance we generate 1000 Fourier surrogates for each AMOC time series, and use those to calculate a p-value for the slope of the CSD indicator in each time series. Whilst we define $p \leq 0.05$ as a significant signal, we also use the p-values themselves to compare the models. Note that we calculate the surrogates from the original AMOC time series and not the CSD indicator time series, which is a more conservative approach than followed in B21 (see [31] for details). [S1 Appendix](#) contains additional information on the CSD indicators and the calculation of statistical significance, including our approach to the issue of multiple comparisons.

When comparing the CMIP6 results to the historical SST observations, we use the ERSSTv5 and HadCRUT5 SST datasets. The uncertainty ranges for the trend and restoring rate are calculated from the uncertainty ensembles of these datasets following [31].

Results

AMOC SST index correlations

In the literature, using the ASSTI to represent the AMOC has been justified based on trend alignments in climate models. Caesar et al. 2018 [34] were interested in the trend of the real-world AMOC, and hence calculated the historical trend of the AMOC streamfunction and the ASSTI in multiple CMIP5 models. Plotting the model streamfunction trends against the model ASSTI trends, they found a strong correlation across models, concluding that every ASSTI historical trend corresponds to a streamfunction trend. Other studies focus on the annual or decadal correlations of the AMOC strength time series with the ASSTI time series, often noting that these are lower than the correlation of the trends [5,8]. While the Caesar et al. 2018 [34] results only state that the historical trend of the AMOC aligns with that of the ASSTI, a high correlation of the time series implies that every annual or decadal ASSTI value corresponds to an AMOC streamfunction value. As such correlations are a more stringent requirement than that shown by Caesar et al. 2018 [34], it is not surprising that studies focusing on the annual or decadal correlations find lower values.

Fig 2 shows both of these approaches in our 27 CMIP6 models. As previously shown by Menary et al. 2020 [48], the trends of the ASSTI time series are still highly correlated with the streamfunction strength trends both at 26.5°N and 35°N, although many of these trends fall outside the observed range of ASSTI trends (Fig 2b,c). This increase in AMOC strength in CMIP6 (as opposed to the observed negative trend) is mostly present in the time span of 1850–1985, and has in part been attributed to too-strong aerosol forcing (see [48–50] for discussion of this forcing and for full AMOC time series). In addition, other work has shown that the large scatter in slopes can also be caused by internal AMOC variability [36]. Indeed, although for each model the ensemble members have the same forcing, in some cases there is a large spread in ensemble member slopes (Fig A in S3 Appendix).

For this work, the more relevant correlation is the direct correlation of the streamfunction strength time series with the ASSTI time series (Fig 2a). Although this correlation is low in many models, for S26.5 56% of ensemble members show a decadal correlation above 0.5, and 27% have a correlation above 0.75 (49% and 26% for S35). One reason this correlation may differ from model to model is that the AMOC is difficult to model, and that models hence vary in how well they represent the connection between its components [22]. For example, the extent, location, and even existence of deep water formation in the SPG differs between models [22]. These differences in the processes connecting the SPG SSTs to the AMOC could influence the ASSTI's suitability to detect AMOC CSD, which may thus differ between models. However, note that similarly to the trend, the correlation between the ASSTI and AMOC strength time series also varies between different ensemble members from the same model, so it is difficult to identify specific models with higher correlation. Thus, while Fig 2a does show some of the low correlations discussed in previous works [5,8], the large percentage of correlations above 0.5 means that for our purposes it is still informative to investigate the ASSTI's reliability for detecting AMOC CSD.

AMOC CSD indicators

Out of the 399 time series for both streamfunction strengths and ASSTI for the 27 models, 36 have an increasing slope in the restoring rate λ with p-value below or equal to the 0.05 significance level (Fig 3). However, only 4 ensemble members have a significant λ increase in both streamfunction strengths (CanESM5 r7, MIROC6 r8 and CESM2 r3 and r8, see Fig 3). There is only one case in which both the ASSTI and a streamfunction strength have coinciding significant positive λ slopes (CanESM5 r10).

If the significant λ increases were due to the AMOC losing stability, we might expect to see them consistently across all ensemble members in a model. The result here is ambiguous: the significant increases are mainly found in certain models, while 11 models (96 ensemble members) show none at all. For example, at one end, IPSL-CM6A-LR has 10 ensemble members (two AMOC latitudes and the ASSTI for each ensemble member resulting in 30 time series for the

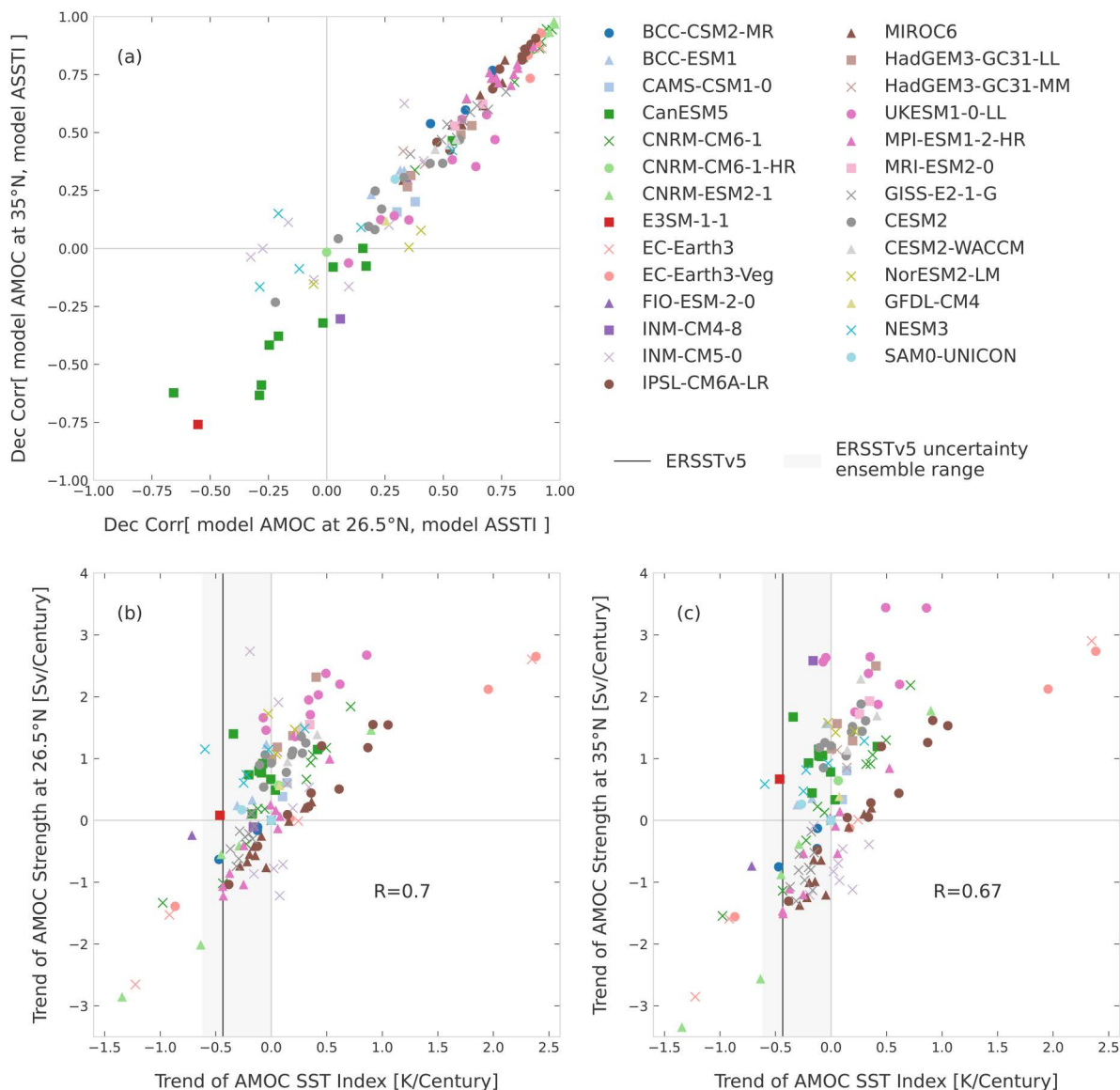


Fig 2. Correlation of the ASSTI fingerprint with the model AMOC strength. The top row (a) shows the correlation of the decadal rolling mean of the ASSTI with the AMOC streamfunction strengths at 35°N and 26.5°N for the individual ensemble members of 27 CMIP6 models. For each model, all ensemble member slopes are shown using the same colour-symbol combination. The bottom row (b,c) shows the 1850-2014 trend of the AMOC streamfunction strengths at 26.5°N (b) and 35°N (c) vs. the slope of the model ASSTI. The correlation of all data points is added as text in (b) and (c). The black line is the slope in the ERSSTv5-derived observational ASSTI, and the light grey ensemble range is calculated from the ERSSTv5 1000 member uncertainty ensemble.

<https://doi.org/10.1371/journal.pclm.0000781.g002>

model) but no significant λ increase, whilst at the other end, CanESM5 has 10 ensemble members and 8 significant λ increases (about a fourth of its 30 time series). The fraction of significant λ increases for each model are shown in Fig B in [S3 Appendix](#). Similarly, when we look beyond the 0.05 hard limit and look at the p value, we can see that some models have much smaller p-values than others (Fig C in [S3 Appendix](#)). The fact that no model has a significant increase in all of its ensemble members already indicates that there may be no robust signal of CSD in the CMIP6 historical simulations,

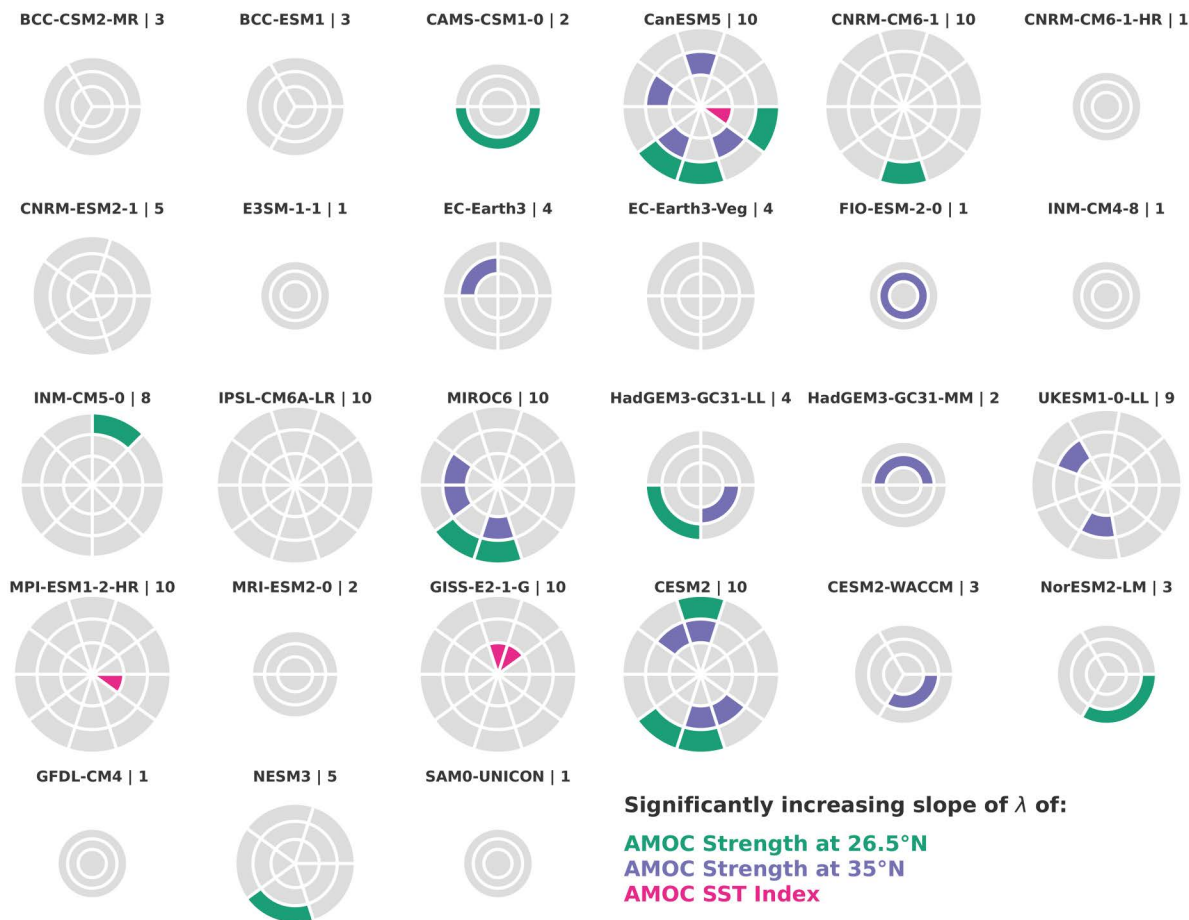


Fig 3. Overview of ensemble members and the occurrence of significant increases in λ for AMOC strength and the SST fingerprint. Each model is represented by a circle which is cut into a number of sectors (slices) corresponding to the ensemble members analysed in this study. These slices are also cut into three ring segments (layers), each representing one of either the AMOC streamfunction strength time series or the AMOC SST Index. So a model with 10 ensemble members will have $10 \times 3 = 30$ slice layers in the circle corresponding to AMOC time series. For each time series, if its λ increases significantly at 0.05 confidence level, the corresponding slice layer is coloured, otherwise it is left grey. The order from outside in is: S26.5 (turquoise), S35 (purple) and AMOC SST index (pink). The name of the model and the number of ensemble members are printed above each circle. The first ensemble member (r1) is always the one above the $+90^\circ$ line from the top (3 o'clock), and the ensemble members then proceed counterclockwise. For example, for CanESM5 the SST Index (pink) has a significant slope in λ in member r10, and the S26.5 index in members r7, r8 and r10. The corresponding list can be found in Table D in [S4 Appendix](#).

<https://doi.org/10.1371/journal.pclm.0000781.g003>

and in particular the lack of CSD indicator increases in the ASSTI time series means that there is no model that completely matches the observations.

The next step in testing the validity of a CSD signal is to check the agreement between different CSD indicators. As described in the Methods, we also calculate the variance and autocorrelation of the time series, and check for their significant increases following the same methods. Out of the 399 AMOC time series, there are 73 time series in total with any sort of significant CSD indication (Fig 4). Of these, 47 time series show significant change in only one indicator, 20 time series in two indicators, and six time series show significant trends in all three indicators (Table A in [S2 Appendix](#)). These last six time series are all from different models, and five of them are the time series of the AMOC streamfunction at 35°N. Again, there is an uneven distribution of these indicators among the models (see Fig D in [S3 Appendix](#) for the fraction of significance per model). While this may indicate that the streamfunction at 35°N is better at reflecting the AMOC's stability

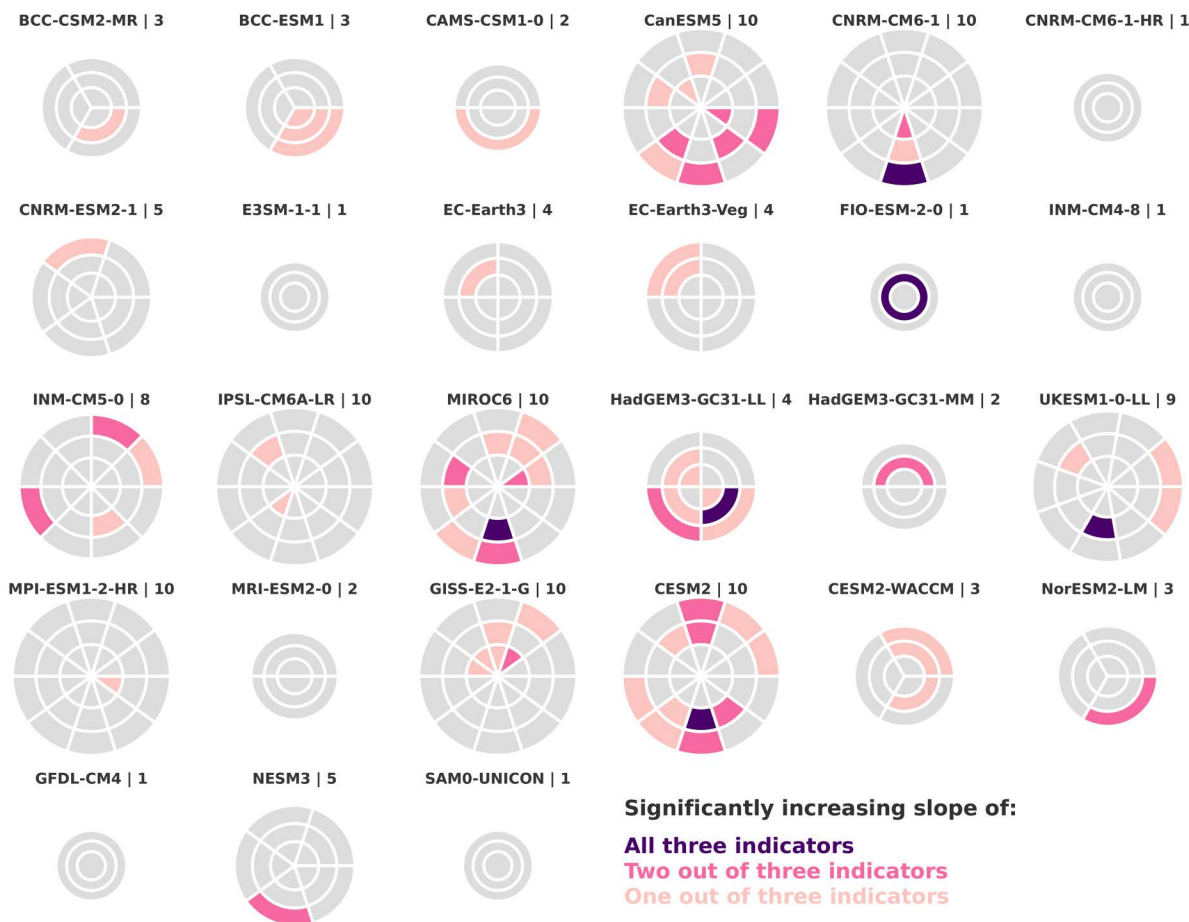


Fig 4. Number of indicators with a significantly increasing slope in the ensemble members' AMOC strengths and SST fingerprints. The structure of the Figure is the same as Fig 3, but instead of representing a significant increase in λ , the colours now represent different amounts of indicators with a significant increase. These are all three (dark purple), two out of three (pink) and just one (light pink).

<https://doi.org/10.1371/journal.pclm.0000781.g004>

than the other latitude, the number of AMOC time series with significant trends in all indicators is too low to make any definitive statement about AMOC stability, especially since these time series all belong to different models.

Up until this point, we have called an individual time series' CSD indicator slope "significant" when it had a p values lower than or equal to 0.05, calculated from a CSD indicator trend distribution of the time series' Fourier surrogates. However, this calculation only truly indicates statistical significance when we investigate only one time series, as we do with the observations. Because we have a large number of time series in our analysis, even if all time series were completely random we would still expect about 5% that fit the criteria of $p \leq 0.05$. Thus, from now on we will not call individual time series significant when the trend of their CSD indicator has $p \leq 0.05$, but instead discuss them simply as time series with increases with $p \leq 0.05$. This then allows us to analyse the probability of seeing certain amounts of these increases with $p \leq 0.05$, and to discuss the significance of these amounts for individual models and the ensemble as a whole.

First, we calculate these probabilities for the distribution of CSD indicator trends in our 133 ensemble members as a whole. As the different CSD indicators and the AMOC strengths and index are not independent (Fig 6d and Figs. E, F and G in S3 Appendix), we thus consider each indicator-strength combination separately when calculating probabilities. For each of these combinations we can then use the binomial distribution to calculate the probability that a given number of

significant outcomes out of 133 occurs randomly, where the probability of a random time series to have a $p \leq 0.05$ is 5%. An in depth explanation of our methods can be found in [S1 Appendix](#). The specific resulting probabilities of this calculation can be found in Table C in [S2 Appendix](#), but the overall results can be well understood from [Fig 5](#), which compares the number of $p \leq 0.05$ indicator increases to the number we would expect from a random group of 133 time series. The number of increases for the two streamfunction strengths is clearly higher than for random time series for all three CSD indicators, and indeed for the two strengths the probability of the total number of increases with $p \leq 0.05$ occurring randomly is less than 5%. However, the number for the ASSTI lies either at or below the expected number ([Fig 5](#)), and the probability calculation shows that the number of $p \leq 0.05$ increases seen in its 133 time series is likely to occur randomly. In fact, the probability of having this number of increases is only as low as 10–15% because the number of $p \leq 0.05$ ASSTI restoring rate trends is less than the expected number, not more. In other words, compared to random time series the only unlikely thing about the number of increases in the ASSTI CSD indicators is that there are so few of them.

In addition to calculating these probabilities for the whole CMIP6 ensemble, we can also calculate a similar probability for each individual model using binomial distributions. Here the calculation is more complicated, as we have to account not only for the number of ensemble members in a model, but also for the large number of models we analyse. A detailed explanation can be found in [S2 Appendix](#). Table B in [S2 Appendix](#) shows the probabilities for the five model-indicator-strength combinations that are below the 5% probability level: CanESM5 λ S35, CESM2 λ S35, CESM2 AR1 S26.5, HadGEM3-GC31-LL AR1 S35 and MIROC6 AR1 S35 (Table B in [S2 Appendix](#)). Even for these, however, only either the lag-1 autocorrelation or the restoring rate λ , and either S35 or S26.5 are significant, where ideally all would need to show significance to state that a model AMOC is losing stability. Note, however, that these probabilities are highly conservative (see [S2 Appendix](#)).

However, probabilities should always be examined contextually [\[51,52\]](#). And even in the models where the number of $p \leq 0.05$ increases is statistically unlikely, these $p \leq 0.05$ increases are unevenly spread across ensemble members, with some members having no $p \leq 0.05$ increases at all ([Fig 4](#)). The statistical unlikelihood of the number of $p \leq 0.05$ increases is thus not enough by itself to conclude that these models have a robust sign of CSD. If the results were more consistent

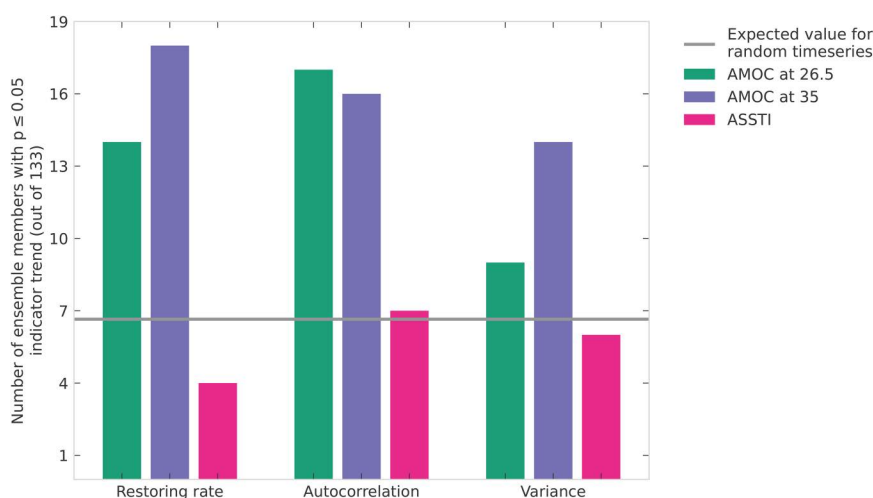


Fig 5. The number of CSD indicator increases with $p \leq 0.05$ compared to the expected number of random increases with $p \leq 0.05$. The height of the bars shows the number of ensemble members for the AMOC strength at 26.5°N (turquoise) and at 35°N (purple), as well as the corresponding number for the ASSTI. The bars are grouped along the horizontal axis according to the CSD indicator the number of increases are calculated for. The black line shows the number of increases we would expect if the time series were completely random, namely $133 * 0.05 = 6.65$. Note that the number of increases for the ASSTI is either at or lower than the expected number.

<https://doi.org/10.1371/journal.pclm.0000781.g005>

within models, we could try to formulate necessary criteria for CSD. As it is, there is no model for which we can say for certain that the historical AMOC is losing stability. In contrast, the one thing that is true across models is that the changes in the model ASSTI do not match the observations: not only could the number of CSD indicator increases with $p \leq 0.05$ easily occur by chance, but in the observations there is a significant increase in all three of the CSD indicators in the ASSTI. This is not present in any of the models' ensemble members. This means that the CMIP6 historical runs lie firmly in the fourth column of [Fig 1](#) — overall no significant increase either in the AMOC streamfunction strength or the ASSTI. The five model-indicator-strength combinations listed in Table B in [S2 Appendix](#) are the exception, lying in the second column of [Fig 1](#), but as they are only 5 combinations out of $27 \cdot 3 \cdot 3 = 243$ the overall result still holds.

Whilst the list of time series with significant CSD indicator slopes changes slightly when the CSD analysis parameters are changed (see Figs. H, I, J and K in [S3 Appendix](#) and Tables B, C in [S4 Appendix](#)), the statistics and conclusions from the results remain the same.

Correlation of CSD significance with other parameters

Finally, we look for reasons why some models have lower p-value for λ slopes than others (even if these are overall not a conclusive sign of CSD). We do this by looking for correlations between the p-values of the different CSD indicator slopes and other properties and observables. This could, for example, explain why only a few ensemble members in each model show CSD, since we have seen the AMOC slope also varies between ensemble members. We compare the p-values of the different CSD indicator slopes with the p-values of the other AMOC time series, with the slope of the historical AMOC, and with the mean AMOC strength ([Fig 6](#)).

We find almost no correlation between p-values of λ -slopes and these properties ([Fig 6](#)). The same is also true for the AR1 and variance (Figs. F and G in [S3 Appendix](#)). The only exception is the correlation between the p-values of the two streamfunction strengths, which is strongest for the variance but also present for the AR1 and λ slopes ([Fig 6d](#), and Figs. Fd and Gd in [S3 Appendix](#)). This correlation is also present for the ensemble mean values (Fig L in [S3 Appendix](#)). This may be expected, as the streamfunction strengths at different latitudes are correlated to some extent. For the ensemble means we also find a negative correlation between each streamfunction strength p-value and the corresponding mean strength (Fig Lg-i in [S3 Appendix](#)). However, this correlation disappears when two outliers with very low mean streamfunction strength are removed.

Discussion and conclusions

The main question we address in this study is: how certain are we that the changes we see in the observation-based ASSTI indicate changes in the AMOC? And more generally, what is happening to the real-world historical AMOC? In the introduction we listed three ways that analysing the higher-order statistics of AMOC time series in CMIP6 historical runs can help us move towards an answer to this question: First, by allowing us to compare the changes in the model ASSTI to those observed; Second, by showing us how the stability of the CMIP6 historical AMOC has changed; Third, by giving us a case study for the connection between the changes in the ASSTI and in the AMOC streamfunction strength. In the interest of clarity, we will divide the discussion into two parts: the overall conclusions from the results for all 27 models, and the exception of the four models listed in Table B in [S2 Appendix](#), where there is some significant sign of CSD but where this sign is not robust enough for concluding AMOC destabilization.

Overall results

First, let us discuss the overall results of the three points posed in the introduction:

1. **No CMIP6 model shows the same significant increase in the ASSTI CSD indicators that the observed ASSTI has** ([Figs 3–5](#)). There are only four $p \leq 0.05$ increases in λ in the SST Index (out of 133 time series), lower than the

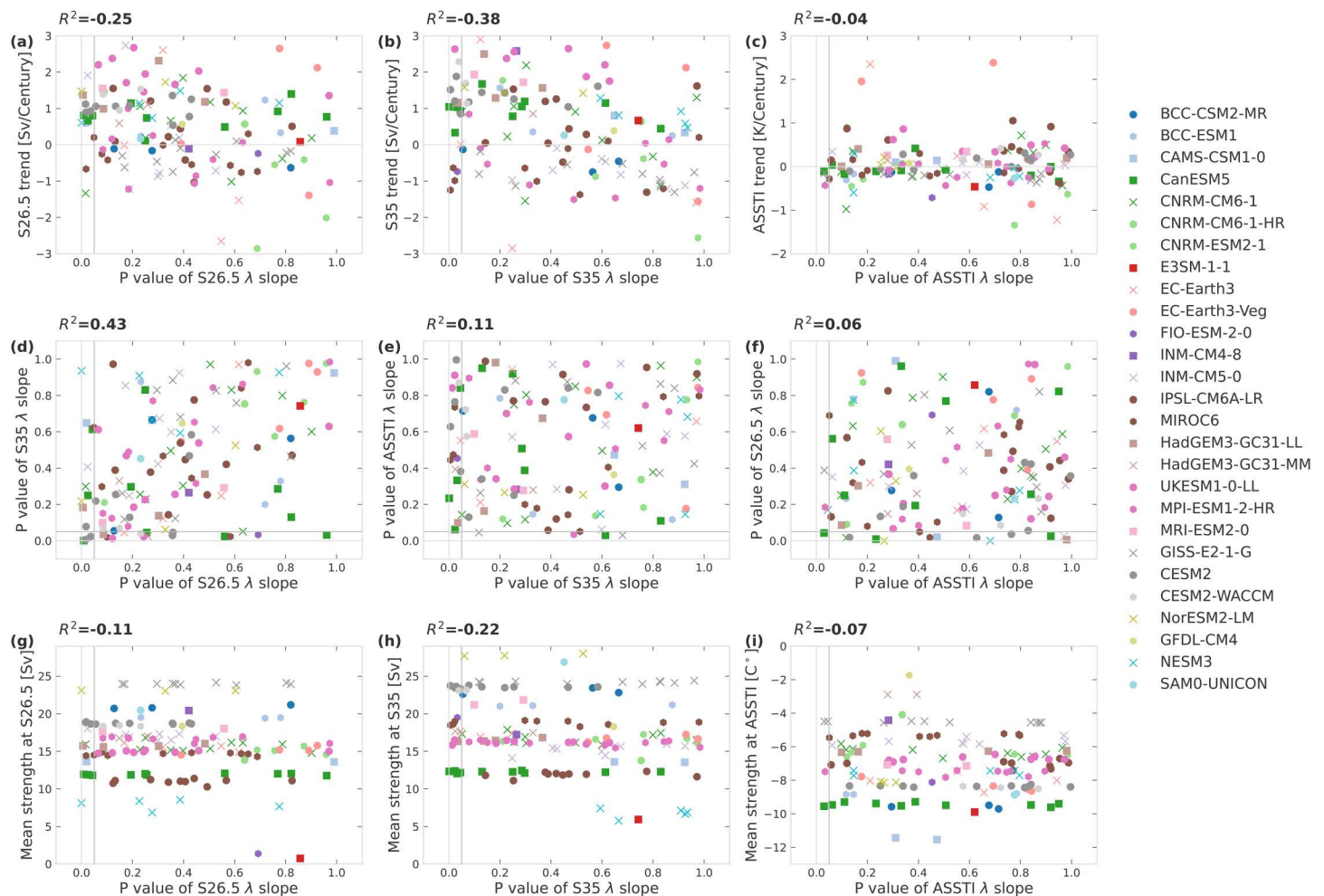


Fig 6. Correlations between the p-values of λ slopes and other quantities. S26.5, S35 and ASSTI stand for the AMOC streamfunction strengths at 26.5°N and 35°N and the AMOC SST Index, respectively. The horizontal axis always shows the p-value of the slope of λ calculated from S26.5 (a,d,g), S35 (b,e,h) or ASSTI (c,f,i). This is plotted against the slope of the AMOC time series itself (a-c), the p-value of the slope of λ calculated from another of the time series (d-f) and the mean strength of the time series (g-i). All panels show the individual ensemble members of 27 CMIP6 models. For each model all ensemble member slopes are shown using the same coloured symbol. Where appropriate darker grey lines mark the 0.05 significance region.

<https://doi.org/10.1371/journal.pclm.0000781.g006>

expected value, and similarly the number of $p \leq 0.05$ increases in the AR1 and variance are also not unlikely (Fig 5 and Table C in S2 Appendix).

- Overall, the historical AMOC in CMIP6 is not losing stability, as no model has a robust sign of CSD.
- The CMIP6 historical runs lie firmly in the fourth column of Fig 1 — that is, there is no significant sign of CSD in either the ASSTI or the AMOC streamfunction strengths. Although this is not a complete test of its reliability, we can conclude from this that the ASSTI is not prone to a “physical false positive” CSD signal (Fig 1 column three).

Point (3) has natural implications for the observed CSD in the real-worlds ASSTI. If the ASSTI is not prone to “physical false positive” CSD, then the observed ASSTI CSD is likely caused by changes in the AMOC. However, both point (3) and its implications require us to assume that most other large scale changes influencing the ASSTI, other than the AMOC stability, are well represented in the CMIP6 models. In theory, there could be some other physical process in the real-world

which is unrelated to the AMOC, induces signs of CSD, and is not well-represented in any CMIP6 model. Still, we believe the existence of such a process is unlikely, and that the above assumption is well-founded, since any such alternative physical process would not only have to react strongly enough to global warming to cause CSD in the ASSTI, but also have a persistent stability bias across all CMIP6 models analysed. And while the AMOC and its stability are known to be especially hard to model [19,22,53], with consistent stability biases across CMIP6 models [16,54,55], the representation of other oceanic and radiative processes associated with the ASSTI [45] varies among models [56–59]. It is thus doubtful that not one out of the 27 models used in this study would include this other physical process, and so show CSD in the model ASSTI. However, to fully support this assumption, further work should investigate all physical processes that might influence the stability of North Atlantic SSTs.

It is also interesting to note that one of the alternative physical processes which may have been able to cause signs of CSD in the ASSTI is the destabilization and tipping of the SPG [60,61]. But in fact, in our results there is no indication of CSD in the ASSTI in the runs identified by Swingedouw et al. 2021 [44] as having an abrupt collapse of the SPG in the 21st century (r1i1p1f1 from CESM2-WACCM, MRI-ESM2–0, NORESM2-LM and CESM2). This is an indication that using the ASSTI to detect CSD in the AMOC would not produce a false warning in case of a collapse of only the SPG, and is in line with recent work disentangling the fingerprints for the AMOC and SPG [46].

A simple summary of our overall results would be the statement “something different is happening in the historical model Atlantic Ocean than in the real-world.” Taking all three above results together, there are a few possible reasons for this difference:

- The model AMOC is stable and will stay stable, whilst the real-world AMOC is losing stability. This difference could be caused by the bias of climate models towards stability [53].
- The model AMOC could tip due to warming in future projections (e.g., the collapse we see in high emission scenarios after 2100), but this tipping point is in the more distant future in the models than it is in the real-world, and thus the CSD that has emerged in the observations has yet to become significant in the models.
- Both the modelled and the real-world AMOC are stable. The observed signs of CSD in the North Atlantic would then be caused by some other phenomenon unrelated to the AMOC. As the third part of our results in the list above shows, this option is unlikely — signs of CSD in the ASSTI do not occur easily.
- The model AMOC could tip due to warming in future projections, but the concept of CSD is not useful for climate change AMOC tipping. This could be due to anthropogenic forcing being too fast for reliable detection of CSD, which assumes the system to be at least close to equilibrium, or because a complex high-dimensional system like the AMOC cannot be captured by a one-dimensional fingerprint signal. This explanation would again require the observed North Atlantic CSD to be caused by something other than the AMOC, which is unlikely according to our results.

The last two points are possible, but unlikely, as it is difficult to imagine a situation in which at the same time CSD is not useful for the AMOC and a false signal is present in the ASSTI. In addition, and as we argue above, this work shows that the ASSTI is not prone to physical false positives. We are thus left with the first two points: the model AMOC is either stable, or further away from a tipping point than the real-world AMOC [62]. The true explanation could vary between models, and in fact is even likely to do so — there is already strong evidence that the AMOC in the GISS-E2-1-G model tips to an alternative weak state in high emission scenarios past 2100 [63], and also evidence that some of the other CMIP6 models (EC-Earth3, MPI-ESM1–2-HR) have a very stable AMOC [19,64]. In addition, the theory that the model AMOC may have the potential for tipping but is farther away from this tipping point than the real-world AMOC has been introduced and substantiated by Shin et al. 2025 [62] for CESM2. The reason for this delay would be the bias in the freshwater convergence into the Atlantic basin (called M_{OV} or F_{OV} in some studies), which is positive in most models and negative in observations [54]. As it is theorized that a certain freshwater convergence threshold needs to be passed for the AMOC to tip, “starting”

further away from this threshold at higher value would explain why the historical model AMOC would be further away from a tipping point [16,26,62].

Our overall conclusions for the CMIP6 historical runs are thus clear: models do not give any indication that the ASSTI is prone to physical false positives, and the AMOC in these models is either stable or further away from tipping than the real-world AMOC.

Significant but inconclusive CSD signs

We thus come to the four models listed in Table B in [S2 Appendix](#): CanESM5, HadGEM3-GC31-LL, CESM2 and MIROC6. For some combination of streamfunction strength latitude and CSD indicator, the number of $p \leq 0.05$ increases in the ensemble members of these models is unlikely to occur by chance. This can be seen clearly in [Figs 3](#) and [4](#), where these models have a much higher concentration of colour than other models. The latter two of these also have one ensemble member with what we would classically expect for an indication of CSD: significant increases of all three CSD indicators in the streamfunction strength at 35°N, and an increase in two out of three CSD indicators in the streamfunction strength at 26.5°N. Indeed, out of the six time series with significant increases in all three CSD indicators, five are for the streamfunction strength at 35°N. If this is due to the AMOC truly destabilizing in these models, it would make sense that the signal would be stronger in the more northern latitude, as the destabilization usually originates in the North Atlantic.

However, it is still difficult to call the result for these models a robust sign of the AMOC losing stability, since the $p \leq 0.05$ increases in the CSD indicators are still only in some ensemble members, and even in those mostly only in one or two CSD indicators. The overall conclusion for CMIP6 historical runs thus still remains that discussed above, namely that there is no robust sign of CSD. However, what we see in these four models fits well with the theory described above that in some models the AMOC is approaching a tipping point, but is too far away from this tipping point for a robust sign of CSD to emerge. Indeed, measuring CSD too far from a tipping point, we would expect to see what we do in these models: the trend emerging above the noise in some members and indicators, but not yet in others. In particular, for CESM2, CanESM5 and HadGEM3-GC31-LL there is evidence supporting the existence of an alternative weak AMOC stable state [22,65], and for CanESM5 there is a future projection simulation where the AMOC collapses beyond 2100 [21].

There is of course also the possibility that the signs of CSD in these models are either being masked by other processes in the ensemble members without a signal, or caused by some other underlying ocean process in the members that do have a signal. Both of these explanations are possible because, as discussed in the previous section, the AMOC is a complex high-dimensional system, and it could be that the one-dimensional signals considered here cannot properly capture the changes in its stability. However, unlike the above reasoning, these explanations do not immediately fit with our conclusions for the overall results, and would require additional evidence to be substantiated.

Final comments

Our results have multiple implications for our understanding of the AMOC. First, the marked difference between the changes in higher-order statistics of North Atlantic SSTs in the models and observations highlights the persistent biases in the ocean representation of complex climate models, and points towards directions for improvement. In addition, we find that no model ASSTI truly matches the way the observed ASSTI changes over time, meaning that studies using the ASSTI as an emergent constraint should be careful to investigate the effect different higher order statistics may have on their results [36,37]. In addition, our results show that signs of CSD in the ASSTI are not a common occurrence. There must thus exist some real-world physical process that is causing these signs of CSD, and that is not represented in any of the models. If one assumes that any non-AMOC-related process with the potential to cause ASSTI CSD would be well represented in at least a few of the analysed models, our results indicate that the CSD we detect in the observed ASSTI suggests something is changing in the AMOC. These results cannot prove that this AMOC change is a loss of stability, but simply ruling out other processes greatly increases the likelihood that the observed ASSTI CSD is caused by AMOC

destabilization. Future work on AMOC CSD in complex climate models should focus on scenarios where the AMOC undergoes a warming-induced abrupt transition, and such work can answer whether the ASSTI is not only not prone to false positives, but also reliably shows true positives. Finally, our results also highlight the importance of using multiple ensemble members when investigating changes in the ocean of a complex climate model. As Romanou et al. 2023 [63] have shown for the GISS-E2-1-G model, the AMOC can have substantially different responses to anthropogenic warming in different ensemble members, and robust conclusions can only be made about the ensemble as a whole.

In conclusion, we have found that there are no robust and significant signs of CSD in the CMIP6 historical runs of 27 models, neither in the AMOC SST Index or the AMOC streamfunction strength. In some models, this may be caused by the modelled historical AMOC being biased to start further away from the tipping point than the real-world AMOC, and thus the CSD signal taking longer to emerge from the noise. In addition, the lack of CSD in the model AMOC SST Index means that it is not prone to show CSD for changes unrelated to the AMOC, and that the observed CSD in the real-world AMOC SST Index likely indicates that the AMOC is losing stability.

Supporting information

S1 Appendix. Critical slowing down indicators and significance calculation.

(PDF)

S2 Appendix. Conservative probability calculation.

(PDF)

S3 Appendix. Extended figures. Fig A: 1850–2014 trends in AMOC streamfunction strength ensemble members for each model at 26.5°N and 35°N. Fig B: Fraction of timeseries per model with a significant increase in λ for both streamfunction strength and SST Index. Fig C: As Fig 3 in the main text, with slices coloured by p-value of λ . Fig D: Fraction of timeseries per model with a significant increase in any of the three CSD indicators. Fig E: As Fig 6 in the main text, for correlations between p-values of different CSD indicators. Fig F: As Fig 6 in the main text, for AR1 instead of λ . Fig G: As Fig 6 in the main text, for variance instead of λ . Fig H: As Fig 3 in the main text, with a window size of 50 years. Fig I: As Fig 3 in the main text, with a window size of 70 years. Fig J: As Fig 4 in the main text, with a window size of 50 years. Fig K: As Fig 4 in the main text, with a window size of 70 years. Fig L: As Fig 6 in the main text, for ensemble mean values of each model. Fig M: Distribution of ensemble member p-values for the λ trend of AMOC streamfunction strengths and SST Index, by ocean component. Fig N: Mean p-value for all λ timeseries per model. Fig O: Mean p-value for all CSD indicator timeseries per model.

(PDF)

S4 Appendix. Extended tables. Table A: List of models and ensemble members used in this study. Table B: Number of timeseries with a significant increase in λ , AR1, or variance for a window size of 50 years. Table C: As Table B for a window size of 70 years. Table D: Full list of ensemble members and AMOC timeseries for which λ shows a significant increase. Table E: Full list of ensemble members and AMOC timeseries for which the autocorrelation shows a significant increase. Table F: Full list of ensemble members and AMOC timeseries for which the variance shows a significant increase.

(PDF)

Acknowledgments

We would like to thank Dr Matthew Menary for providing the CMIP6 AMOC data.

Author contributions

Conceptualization: Maya Ben-Yami, Lana Blaschke, Niklas Boers.

Formal analysis: Maya Ben-Yami.

Methodology: Maya Ben-Yami, Lana Blaschke.

Project administration: Niklas Boers.

Supervision: Sebastian Bathiany, Niklas Boers.

Visualization: Maya Ben-Yami.

Writing – original draft: Maya Ben-Yami.

Writing – review & editing: Maya Ben-Yami, Lana Blaschke, Sebastian Bathiany, Niklas Boers.

References

1. Lee J, Marotzke J, Bala G, Cao L, Corti S, Dunne J. Future global climate: Scenario-based projections and near-term information. In: Masson-Delmotte V, Zhai P, Pirani A, Connors S, Péan C, Berger S. *Climate Change 2021: The Physical Science Basis*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 2021. p. 553–672.
2. Bellomo K, Angeloni M, Corti S, von Hardenberg J. Future climate change shaped by inter-model differences in Atlantic meridional overturning circulation response. *Nat Commun*. 2021;12(1):3659. <https://doi.org/10.1038/s41467-021-24015-w> PMID: [34135324](#)
3. Liu W, Fedorov AV, Xie S-P, Hu S. Climate impacts of a weakened Atlantic Meridional Overturning Circulation in a warming climate. *Sci Adv*. 2020;6(26):eaz4876. <https://doi.org/10.1126/sciadv.aaz4876> PMID: [32637596](#)
4. Caesar L, McCarthy G, Thornalley D, Cahill N, Rahmstorf S. Current Atlantic Meridional Overturning Circulation Weakest in Last Millennium. *Nature Geoscience*. 2021;14(3):118–20. <https://doi.org/10.1038/s41561-021-00699-z>
5. Zhu C, Liu Z, Zhang S, Wu L. Likely accelerated weakening of Atlantic overturning circulation emerges in optimal salinity fingerprint. *Nat Commun*. 2023;14(1):1245. <https://doi.org/10.1038/s41467-023-36288-4> PMID: [36871075](#)
6. Kilbourne KH, Wanamaker AD, Moffa-Sanchez P, Reynolds DJ, Amrhein DE, Butler PG, et al. Atlantic circulation change still uncertain. *Nat Geosci*. 2022;15(3):165–7. <https://doi.org/10.1038/s41561-022-00896-4>
7. Latif M, Sun J, Visbeck M, Hadi Bordbar M. Natural variability has dominated Atlantic Meridional Overturning Circulation since 1900. *Nat Clim Chang*. 2022;12(5):455–60. <https://doi.org/10.1038/s41558-022-01342-4>
8. Terhaar J, Vogt L, Foukal NP. Atlantic overturning inferred from air-sea heat fluxes indicates no decline since the 1960s. *Nat Commun*. 2025;16(1):222. <https://doi.org/10.1038/s41467-024-55297-5> PMID: [39814728](#)
9. McKay D, Staal A, Abrams J, Winkelmann R, Sakschewski B, Loriani S. Exceeding 1.5 $^{\circ}\text{C}$ global warming could trigger multiple climate tipping points. *Science*. 2022;377(6611). <https://doi.org/10.1126/science.abn7950>
10. Boers N, Ghil M, Stocker TF. Theoretical and paleoclimatic evidence for abrupt transitions in the Earth system. *Environ Res Lett*. 2022;17(9):093006. <https://doi.org/10.1088/1748-9326/ac8944>
11. Henry LG, McManus JF, Curry WB, Roberts NL, Piotrowski AM, Keigwin LD. North Atlantic ocean circulation and abrupt climate change during the last glaciation. *Science*. 2016;353(6298):470–4. <https://doi.org/10.1126/science.aaf5529> PMID: [27365315](#)
12. Stommel H. Thermohaline convection with two stable regimes of flow. *Tellus*. 1961;13(2):224–30. <https://doi.org/10.3402/tellusa.v13i2.9491>
13. Cessi P. A Simple Box Model of Stochastically Forced Thermohaline Flow. *J Phys Oceanogr*. 1994;24(9):1911–20. [https://doi.org/10.1175/1520-0485\(1994\)024<1911:asbmos>2.0.co;2](https://doi.org/10.1175/1520-0485(1994)024<1911:asbmos>2.0.co;2)
14. Rahmstorf S. Ocean circulation and climate during the past 120,000 years. *Nature*. 2002;419(6903):207–14. <https://doi.org/10.1038/nature01090> PMID: [12226675](#)
15. Hawkins E, Smith RS, Allison LC, Gregory JM, Woollings TJ, Pohlmann H, et al. Bistability of the Atlantic overturning circulation in a global climate model and links to ocean freshwater transport. *Geophys Res Lett*. 2011;38(10):n/a–n/a. <https://doi.org/10.1029/2011gl047208>
16. Liu W, Xie S-P, Liu Z, Zhu J. Overlooked possibility of a collapsed Atlantic Meridional Overturning Circulation in warming climate. *Sci Adv*. 2017;3(1):e1601666. <https://doi.org/10.1126/sciadv.1601666> PMID: [28070560](#)
17. Jackson L, Wood R. Hysteresis and resilience of the AMOC in an eddy-permitting GCM. *Geophysical Research Letters*. 2018;45(16):8547–56. <https://doi.org/10.1029/2018GL078104>
18. van Westen RM, Kliphuis M, Dijkstra HA. Physics-based early warning signal shows that AMOC is on tipping course. *Sci Adv*. 2024;10(6):eadk1189. <https://doi.org/10.1126/sciadv.adk1189> PMID: [38335283](#)
19. Jackson L, Alastrué De Asenjo E, Bellomo K, Danabasoglu G, Haak H, Hu A. Understanding AMOC stability: the North Atlantic Hosing Model Inter-comparison Project. *Geoscientific Model Development*. 2023;16(7):1975–95. <https://doi.org/10.5194/gmd-16-1975-2023>
20. Arias P, Bellouin N, Coppola E, Jones R, Krinner G, Marotzke J. Technical Summary. *Climate Change 2021: The Physical Science Basis*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 2021. p. 33–144.
21. Drijfhout S, Angevaere JR, Mecking J, van Westen RM, Rahmstorf S. Shutdown of northern Atlantic overturning after 2100 following deep mixing collapse in CMIP6 projections. *Environ Res Lett*. 2025;20(9):094062. <https://doi.org/10.1088/1748-9326/adfa3b>

22. Jackson LC, Hewitt HT, Bruciaferri D, Calvert D, Graham T, Guivarc'h C, et al. Challenges simulating the AMOC in climate models. *Philos Trans A Math Phys Eng Sci*. 2023;381(2262):20220187. <https://doi.org/10.1098/rsta.2022.0187> PMID: 37866390
23. Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53–9. <https://doi.org/10.1038/nature08227> PMID: 19727193
24. Kéfi S, Dakos V, Scheffer M, Van Nes EH, Rietkerk M. Early warning signals also precede non-catastrophic transitions. *Oikos*. 2012;122(5):641–8. <https://doi.org/10.1111/j.1600-0706.2012.20838.x>
25. Zimmerman CC, Wagner TJW, Maroon EA, McNamara DE. Slowed Response of Atlantic Meridional Overturning Circulation Not a Robust Signal of Collapse. *Geophysical Research Letters*. 2025;52(2). <https://doi.org/10.1029/2024gl112415>
26. van Westen R, Vanderborcht E, Dijkstra H. A saddle-node bifurcation is causing the AMOC collapse in the community Earth system model. 2025. <https://egusphere.copernicus.org/preprints/2025/egusphere-2025-14/>
27. Boers N. Observation-based early-warning signals for a collapse of the Atlantic Meridional Overturning Circulation. *Nat Clim Chang*. 2021;11(8):680–8. <https://doi.org/10.1038/s41558-021-01097-4>
28. Ditlevsen P, Ditlevsen S. Warning of a forthcoming collapse of the Atlantic meridional overturning circulation. *Nat Commun*. 2023;14(1):4254. <https://doi.org/10.1038/s41467-023-39810-w> PMID: 37491344
29. Ben-Yami M, Morr A, Bathiany S, Boers N. Uncertainties too large to predict tipping times of major Earth system components from historical data. *Sci Adv*. 2024;10(31):eadl4841. <https://doi.org/10.1126/sciadv.adl4841> PMID: 39093979
30. Rietkerk M, Skiba V, Weinans E, Hébert R, Laepple T. Ambiguity of early warning signals for climate tipping points. *Nat Clim Chang*. 2025;15(5):479–88. <https://doi.org/10.1038/s41558-025-02328-8>
31. Ben-Yami M, Skiba V, Bathiany S, Boers N. Uncertainties in critical slowing down indicators of observation-based fingerprints of the Atlantic Overturning Circulation. *Nat Commun*. 2023;14(1):8344. <https://doi.org/10.1038/s41467-023-44046-9> PMID: 38102135
32. Frajka-Williams E, Moat B, Smeed D, Rayner D, Johns W, Baringer M. Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux Array-Western Boundary Time Series) array at 26N from 2004 to 2020 (v2020.1). 2021.
33. Ren Q, Li Y, Hu S, Xie S, Lyu Y, Wang F. Heat Storage Pattern Linked to the Atlantic Meridional Overturning Circulation Slowdown. *Geophysical Research Letters*. 2025;52(13). <https://doi.org/10.1029/2025gl116801>
34. Caesar L, Rahmstorf S, Robinson A, Feulner G, Saba V. Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature*. 2018;556(7700):191–6. <https://doi.org/10.1038/s41586-018-0006-5> PMID: 29643485
35. Menary MB, Hodson DLR, Robson JI, Sutton RT, Wood RA. A Mechanism of Internal Decadal Atlantic Ocean Variability in a High-Resolution Coupled Climate Model. *Journal of Climate*. 2015;28(19):7764–85. <https://doi.org/10.1175/jcli-d-15-0106.1>
36. Bonnet R, Swingedouw D, Gastineau G, Boucher O, Deshayes J, Hourdin F, et al. Increased risk of near term global warming due to a recent AMOC weakening. *Nat Commun*. 2021;12(1):6108. <https://doi.org/10.1038/s41467-021-26370-0> PMID: 34671020
37. Devilliers M, Olsen S, Langehaug H, Guo C, Mahmood R, Tian T. Constraining CMIP6 simulations for Atlantic Water in the Arctic using an AMOC-SST index. *Front Clim*. 2025. <https://doi.org/10.3389/fclim.2025.1550772>
38. Little C, Zhao M, Buckley M. Do surface temperature indices reflect centennial-timescale trends in Atlantic meridional overturning circulation strength?. *Geophysical Research Letters*. 2020;47(22):1–10. <https://doi.org/10.1029/2020GL090888>
39. Mackay Q, Fan Y, Karnauskas KB, Li L. Nonstationarity of the Atlantic Meridional Overturning Circulation's Fingerprint on Sea Surface Temperature. *Geophysical Research Letters*. 2024;51(19). <https://doi.org/10.1029/2024gl109789>
40. Swingedouw D, Houssais M, Herbaut C, Blaizot A, Devilliers M, Deshayes J. AMOC recent and future trends: a crucial role for oceanic resolution and Greenland melting?. *Frontiers in Climate*. 2022;4(April). <https://doi.org/10.3389/fclim.2022.838310>
41. Sun J, Latif M, Park W. Subpolar Gyre – AMOC – Atmosphere Interactions on Multidecadal Timescales in a Version of the Kiel Climate Model. *Journal of Climate*. 2021;:1–56. <https://doi.org/10.1175/jcli-d-20-0725.1>
42. Shi H, Jin F-F, Wills RCJ, Jacox MG, Amaya DJ, Black BA, et al. Global decline in ocean memory over the 21st century. *Sci Adv*. 2022;8(18):eabm3468. <https://doi.org/10.1126/sciadv.abm3468> PMID: 35522743
43. Keil P, Mauritsen T, Jungclauss J, Hedemann C, Olonscheck D, Ghosh R. Multiple drivers of the North Atlantic warming hole. *Nat Clim Chang*. 2020;10(7):667–71. <https://doi.org/10.1038/s41558-020-0819-8>
44. Swingedouw D, Bily A, Esquerdo C, Borchert LF, Sgubin G, Mignot J, et al. On the risk of abrupt changes in the North Atlantic subpolar gyre in CMIP6 models. *Ann N Y Acad Sci*. 2021;1504(1):187–201. <https://doi.org/10.1111/nyas.14659> PMID: 34212391
45. Fan Y, Chan D, Clothiaux E, Zhang P, Li L. Subpolar North Atlantic cooling reinforced by colder, drier atmosphere with a weakening Atlantic meridional overturning circulation; 2025. Available from: <https://www.science.org>
46. Emirzade B, Ajagun-Brauns J, Ben-Yami M, Bathiany S, Shin Y, Boers N. Optimal Fingerprints for Decoupling Atlantic Overturning and Subpolar Gyre. *Research Square*. 2025. <https://doi.org/10.21203/rs.3.rs-7073175/v1>
47. Michel SLL, Dijkstra HA, Guardamagna F, Jacques-Dumas V, van Westen RM, von der Heydt AS. Deep learning based reconstructions of the Atlantic meridional overturning circulation confirm twenty-first century decline. *Environ Res Lett*. 2025;20(6):064036. <https://doi.org/10.1088/1748-9326/add7f0>

48. Menary MB, Robson J, Allan RP, Booth BBB, Cassou C, Gastineau G, et al. Aerosol-Forced AMOC Changes in CMIP6 Historical Simulations. *Geophysical Research Letters*. 2020;47(14). <https://doi.org/10.1029/2020gl088166>
49. Robson J, Menary MB, Sutton RT, Mecking J, Gregory JM, Jones C, et al. The Role of Anthropogenic Aerosol Forcing in the 1850–1985 Strengthening of the AMOC in CMIP6 Historical Simulations. *Journal of Climate*. 2022;35(20):3243–63. <https://doi.org/10.1175/jcli-d-22-0124.1>
50. Weijer W, Cheng W, Garuba OA, Hu A, Nadiga BT. CMIP6 Models Predict Significant 21st Century Decline of the Atlantic Meridional Overturning Circulation. *Geophysical Research Letters*. 2020;47(12). <https://doi.org/10.1029/2019gl086075>
51. Vidgen B, Yasseri T. P-values: Misunderstood and misused. 2016.
52. Colquhoun D. An investigation of the false discovery rate and the misinterpretation of p-values. *R Soc Open Sci*. 2014;1(3):140216. <https://doi.org/10.1098/rsos.140216> PMID: 26064558
53. Valdes P. Built for stability. *Nature Geosci*. 2011;4(7):414–6. <https://doi.org/10.1038/ngeo1200>
54. van Westen RM, Dijkstra HA. Persistent climate model biases in the Atlantic Ocean's freshwater transport. *Ocean Sci*. 2024;20(2):549–67. <https://doi.org/10.5194/os-20-549-2024>
55. Dijkstra HA. Characterization of the multiple equilibria regime in a global ocean model. *Tellus A: Dynamic Meteorology and Oceanography*. 2007;59(5):695. <https://doi.org/10.1111/j.1600-0870.2007.00267.x>
56. Liu H, Song Z, Wang X, Misra V. An ocean perspective on CMIP6 climate model evaluations. *Deep Sea Research Part II: Topical Studies in Oceanography*. 2022;201:105120. <https://doi.org/10.1016/j.dsr2.2022.105120>
57. Song F, Mao Y, Liu S, Wu L, Dong L, Su H, et al. A long-term decline in downward surface solar radiation. *Natl Sci Rev*. 2025;12(3):nwaf007. <https://doi.org/10.1093/nsr/nwaf007> PMID: 39989914
58. Wild M. The global energy balance as represented in CMIP6 climate models. *Clim Dyn*. 2020;55(3):553–77. <https://doi.org/10.1007/s00382-020-05282-7> PMID: 32704207
59. Xu J, Zhang X, Feng C, Yang S, Guan S, Jia K, et al. Evaluation of Surface Upward Longwave Radiation in the CMIP6 Models with Ground and Satellite Observations. *Remote Sensing*. 2021;13(21):4464. <https://doi.org/10.3390/rs13214464>
60. Born A, Stocker T, Raible C, Levermann A. Is the Atlantic subpolar gyre bistable in comprehensive coupled climate models?. *Climate Dynamics*. 2013;40(11–12):2993–3007. <https://doi.org/10.1007/s00382-012-1525-7>
61. Born A, Stocker TF. Two Stable Equilibria of the Atlantic Subpolar Gyre. *Journal of Physical Oceanography*. 2014;44(1):246–64. <https://doi.org/10.1175/jpo-d-13-073.1>
62. Shin Y, Oh J, Bathiany S, Ben-Yami M, Årthun M, Lee H. Reconciled warning signals in observations and models imply approaching AMOC tipping point. 2025. <https://doi.org/10.48550/arXiv.2503.22111>
63. Romanou A, Rind D, Jonas J, Miller R, Kelley M, Russell G, et al. Stochastic Bifurcation of the North Atlantic Circulation under a Midrange Future Climate Scenario with the NASA-GISS ModelE. *Journal of Climate*. 2023;36(18):6141–61. <https://doi.org/10.1175/jcli-d-22-0536.1>
64. Guo C, Yang S, Schiller-Weiss I, Bernales J, Olsen S, Koenigk T. Reversible Atlantic overturning despite continued Greenland Ice Sheet melt. 2025.
65. Chapman R, Sinet S, Ritchie PDL. Tipping mechanisms in a conceptual model of the Atlantic Meridional Overturning Circulation. *Weather*. 2024;79(10):316–23. <https://doi.org/10.1002/wea.7609>