Simulating AMOC tipping driven by internal climate variability with a rare event algorithm.

Matteo Cini (matteocini@hotmail.it)
Università degli Studi di Torino
Giuseppe Zappa
National Research Council of Italy, Institute of Atmospheric Sciences and Climate
Francesco Ragone
UC Louvain la Neuve
Susanna Corti
National Research Council of Italy, Institute of Atmospheric Sciences and Climate

Article

Keywords:

Posted Date: August 8th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3215995/v1

Additional Declarations: (Not answered)

Simulating AMOC tipping driven by internal climate variability with a rare event algorithm.

Matteo Cini[1,2] , Giuseppe Zappa [2] , Francesco Ragone[3,4] , Susanna Corti [2]

[1] *Department of Physics, Università degli studi di Torino, Turin, Italy*

[2] *National Research Council of Italy, Institute of Atmospheric Sciences and Climate (CNR-ISAC), Bologna, Italy*

[3] *Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium*

[4] *Royal Meteorological Institute of Belgium, Brussels, Belgium*

Abstract

This study investigates the possibility of Atlantic Meridional Overturning Circulation (AMOC) noise-induced tipping solely driven by internal climate variability without applying external forcing that alter the radiative forcing or the North Atlantic freshwater budget. We address this hypothesis with an innovative approach by applying a rare event algorithm to ensemble simulations of present-day climate with an intermediate complexity climate model. The algorithm successfully identifies trajectories leading to abrupt AMOC slowdowns, which are unprecedented in a 2000-year control run. Part of these AMOC weakened states lead to collapsed state without evidence of AMOC recovery on multi-centennial time scales. The temperature and Northern Hemisphere jet stream responses to these internally-induced AMOC slowdowns show strong similarities with those found in externally forced AMOC slowdowns in state-of-the-art climate models. The AMOC slowdown seems to be initially driven by Ekman transport due to westerly wind stress anomalies in the North Atlantic and subsequently amplified by a complete collapse of the oceanic convection in the Labrador Sea. These results demonstrate that transitions to a collapsed AMOC state purely due to internal variability in a model simulation of present-day climate are rare but physically possible. Additionally, these results show that rare event algorithms are a tool of valuable and general interest to study tipping points since they introduce the possibility of collecting a large number of tipping events that cannot be sampled using traditional approaches. This opens the possibility of identifying the mechanisms driving tipping events in complex systems in which little a-priori knowledge is available.

Introduction

The Atlantic Meridional Overturning Circulation (AMOC) is a large system of ocean currents within the Atlantic Ocean which is of key importance for climate since it regulates the meridional transport of heat and freshwater [1,2]. The main engine of these currents is the sinking of cold and salty dense water in the subpolar North Atlantic [2], mainly as a result of intense turbulent heat loss to the atmosphere in winter.

The AMOC is projected to weaken in response to anthropogenic climate change, due to a decrease in the surface density of the sea water in the subpolar North Atlantic [3,4]. Numerical experiments have shown that AMOC weakening can induce widespread regional and global climate impacts, such as significant decreases of North Hemisphere surface temperature, increases of Arctic sea ice and changes in the atmospheric circulation affecting global precipitation patterns [1,5–7]. In particular, in response to an AMOC weakening, the InterTropical Convergence Zone (ITCZ) tends to shift southward, while the North Hemisphere mid-latitude jet moves on average northward [1,5].

Of particular concern is the possibility that the AMOC, a key tipping element of the climate system, may pass a tipping point and reach a collapsed or substantially weakened stable state [2,8–12], which the sixth assessment report of the IPCC summarises as a low probability event but with a potentially very high impact on climate [3]. AMOC multistability is related to the salt-advection feedback, i.e. the fact that the AMOC transports salty surface water to the North Atlantic. The stronger the AMOC, the saltier the surface North Atlantic ocean, thus promoting a more sustained deep convection which in turn reinfroces the AMOC itself, and vice versa. Several studies have established that rapid changes in the AMOC triggered abrupt climate transitions between glacial and interglacial phases in the Pleistocene, such as Dansgaard-Oeschger (DO) events [2,13–19].

The studies on the weakening and/or stability of the AMOC are usually performed in simulations with external forcing elements like greenhouse gases [1], freshwater [5–7,20] or surface heat perturbations [21]. This allows the analysis of the climatic response to an AMOC slowdown and assess if the system has passed a bifurcation tipping point, so that it would remain in the new collapsed state, following a relaxation of the forcing elements [2]. The passing of a tipping point may also depend on the time scale of the forcing element, i.e. the critical threshold depends on the rate of change in the forcing [22].

However, understanding the possibility of a noise-induced tipping of the AMOC is equally important. By noise-induced it is meant a spontaneous transition of a system between two

distinct stable states via its own internal chaotic variability without any external forcing [3]. First of all, within long, paleo-climatic time scales, even rare events with very long return times have a non-negligible probability to occur. Dramatic spontaneous AMOC slowdowns appear in multi millennial simulations [13,23–26], which are relevant for the study of past climate transitions, e.g. DO events. Secondly, even in forced experiment an AMOC collapse is generally achieved through the combined action of external forcing and internal variability, so that the passing of the tipping point may be anticipated or delayed because of internal climate variability. Lastly, in simulations with forcing elements it may be difficult to isolate the impacts derived by the AMOC slowdown from the impacts derived directly by the forcing itself.

The importance of internal variability is evident also by the large variability in the AMOC projections found in a global warming scenario[1], and by the fact that when considering an ensemble simulation some members may collapse while others recover [27,28]. We need, consequently, to assess the potential of internal climate variability in triggering a noise-induced tipping of the AMOC.

To sample a spontaneous AMOC tipping event, very long ensemble simulations would be needed. This can become computationally untenable in particular when one wants to sample a large number of events to study the mechanism involved in the collapse on a robust physical basis. To tackle this problem, here we propose, for the first time in this context, the application of a rare event algorithm in order to internally push the coupled atmosphere-ocean system towards rare trajectories associated with a decline of the AMOC.

Rare events algorithms are computational techniques developed in statistical physics applied to a wide range of subjects (e.g. matter physics, chemistry, biology) [29,30] whose aim is to reduce the computational effort necessary to sample rare events in numerical simulations. The rare event algorithm used in this work has been previously applied to the computation of extreme heat waves in Europe [29–31]. It consists of a set of killing and cloning rules applied along the evolution of an ensemble simulation in order to populate the ensemble with model trajectories that feature rare values of a target metric of interest. In our case the metric is the value of the AMOC between 46-66°N, which the algorithm seeks to reduce. A different rare event algorithm has been recently applied to study ocean circulation noise-induced temporary transitions in a box model [32]. The rare event algorithm is applied to an intermediate complexity coupled climate model, composed by the Planet Simulator (PlaSim)[33,34] and the Large Scale Geostrophic Ocean (LSG)[35,36] . All simulations are performed at stationary greenhouse gases forcing. In this setup, by construction, any tipping is necessarily induced by the internal climate dynamics.

Overall, the goal of the paper is condensed in two main research questions:

- Is a rare event algorithm able to sample spontaneous AMOC collapses driven by the internal variability of the system?
- What is the climate response to an AMOC collapse as well as the physical processes driving the AMOC spontaneous collapse?

RESULTS

AMOC spontaneous slowdowns

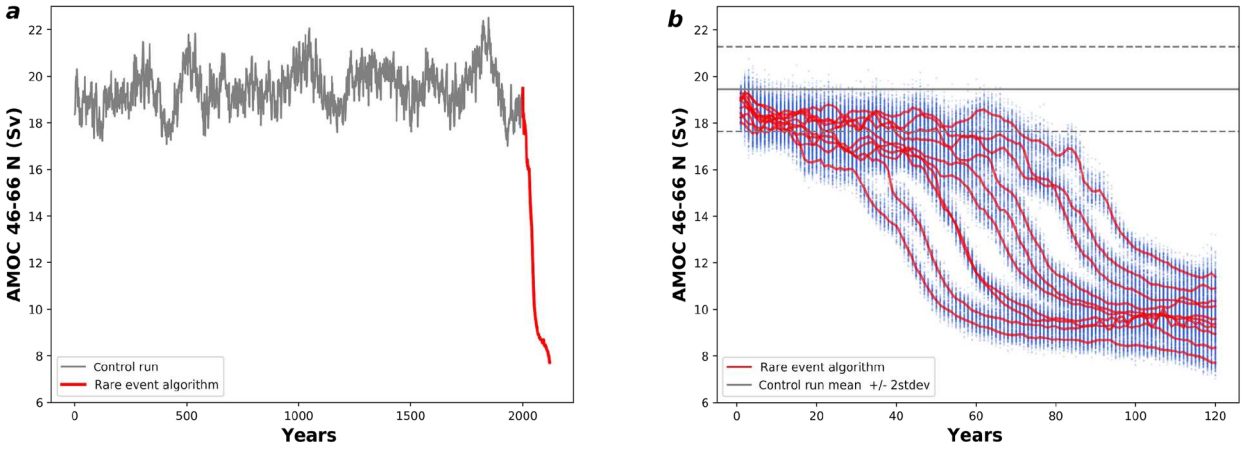

Fig. 1 *a) AMOC variability in a 2000ys control run (grey) and ensemble mean AMOC evolution in one 125-year long simulation performed with the rare event algorithm (red). b) 10 independent 125-year ensemble simulations performed with the rare event algorithm. The red lines represent the AMOC ensemble-mean evolution for each simulation. The blue points represent the values of the AMOC in all the 100 ensemble members of each simulation. The horizontal grey lines indicate the mean (full), plus/minus 2-standard-deviations (dashed), AMOC variability in the control run (Fig 1a in grey).*

The PlaSIM-LSG control run (Fig. 1a, grey line), performed as a single member simulation over 2000 years [37], shows a realistic natural variability of the AMOC ranging between 17-21 Sv with oscillations taking place over different time scales. The AMOC variability appears to be stable over the 2000 years of the simulation. However, a different behaviour quickly emerges as the rare event algorithm is switched on. A rare-event simulation initialised at the end of the control run (red line in Fig 1a) shows a dramatic decrease of the AMOC, which

reaches values clearly outside of the previous boundaries. The system is now exploring states with a significantly weakened AMOC. The weakening is also reached in a short time scale, and with a rate of change substantially exceeding that of the fluctuations characterising the control run. It is worth to stress that this is obtained without applying any external forcing. The algorithm samples rare trajectories where the AMOC weakens purely due to internal variability via atmosphere-ocean-ice interactions.

To test the robustness of this behaviour, we compare the results from 10 independent rare-event simulations in Fig. 1b. All the simulations follow the same qualitative time evolution: an initial slow weakening of the AMOC followed by an abrupt transition resembling the passing of a tipping point that leads to a final plateau-like weakened AMOC state. The simulations differ in the timing of the abrupt collapse and in the final value of the AMOC, which ranges between 7.7-11.4 Sv. An equivalent time evolution is also obtained when the AMOC maximum is evaluated within the entire North Atlantic ocean, rather than in the 46-66N band (see Fig S3). The main difference is that the basin wide AMOC maximum remains slightly higher (around 14 Sv), implying that the AMOC remains active in the subtropics. This difference is consistent with the fact that the algorithm seeks to minimise the AMOC at 46-66N.

Climate responses to the AMOC slowdown

By years 115-125, the abrupt weakening of the AMOC is associated with widespread changes in a number of key oceanic and atmospheric climate aspects. We examine these responses in terms of anomalies compared to the control run (Fig. 2).

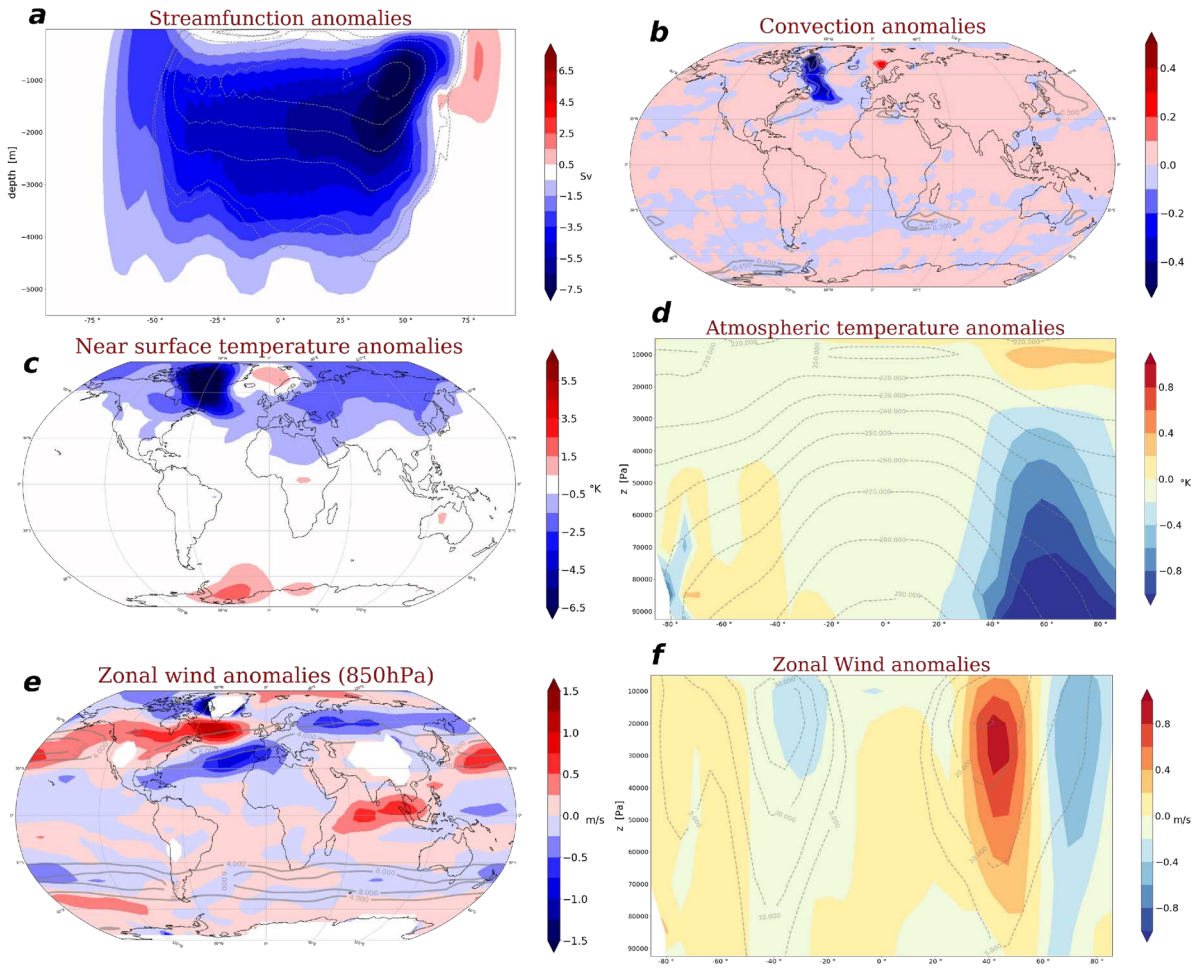

Fig. 2: *Climate responses associated with the AMOC slowdown, as obtained by the differences between the mean climate in the last 10 years of the rare-event-algorithm runs (average of years 115-125) relative to the control run (average of the last 10 years before the rare event simulations start). (a) AMOC stream function anomalies. Here the stream function is computed with Eq. 2a but starting from the annual mean meridional velocity. (b) Anomalies in the frequency of oceanic convection in the uppermost layer of the ocean. (c) Near-surface atmospheric temperature anomalies (2 meters) and d) atmoshperic temperature anomalies in the zonal-mean cross-section. (e) and (f) respectively represent zonal wind anomalies at 850 hPa and the respective zonal-mean cross-section. The climatologies of each variable in the control run are represented by the grey lines. All climate responses are evaluated as averages over the 10 rare-event-algorithm simulations shown in Fig 1b.*

The response in the meridional overturning stream function (Fig. 2a) reveals that even if the changes peak around 50°N, the anomaly is widely spread throughout the Atlantic.

Consistently with these circulation changes, anomalies in the oceanic convection frequency show the extinction of the branch of deep water formation over the Labrador Sea (ratio anomalies/climatology is almost -1 everywhere) in favour of a relative intensification of deep water formation in the Norwegian Sea (Fig 2b).

The slowdown of the AMOC and collapse of oceanic convection induces a general decrease of temperature in the North Hemisphere, especially in the Labrador Sea and Baffin Bay areas, which is consistent with the collapse of convection and a reduced meridional oceanic heat transport (Fig. 2c-d). The intense signal over the Labrador sea (up to 11°K anomaly) is also associated with sea ice growth (not shown). In contrast, there are warming areas over the Weddel sea and in the Greenland-Iceland-Norwegian (GIN) seas, which are associated with a spot of increasing convection (Fig 2a). Another study on spontaneous cooling events associated with a general decrease of the AMOC, documented a similar increased oceanic convection over Norwegian sea [26]. In our model, in that region, oceanic levels below the uppermost have a higher potential temperature, so that enhanced convection brings to the surface warmer waters. This can explain the warming anomalies over the Norwegian Sea. Overall, the abrupt AMOC weakening (up to 50% even considering the whole North Atlantic) results in a surface temperature drop of the order of 1.5°K over North Hemisphere continents, and of 0.5°K in the global mean.

The altered meridional temperature gradients induce changes in the mid-latitude jet streams (Fig 2e-f). The dominant effect is a strengthening and poleward shift of the North Hemisphere jet stream in the zonal-mean. However, changes in zonal wind at 850 hPa show different responses in the individual ocean basins. A predominant poleward shift characterises the North Atlantic jet, while the North Pacific jet tends to strengthen and slightly shift southward. This pattern is found even at higher altitude (300 hPa, not shown) suggesting that the contribution to the strengthening of the zonal-mean jet comes from the North Pacific region, while the contribution to the poleward shift comes from the North Atlantic region. These signals are relatively strong, up to 25% of the climatological amplitude. Positive westerlies anomalies at 850 hPa can be found over the equatorial Indian Ocean.

In general, for what concerns the climate response to the AMOC slowdown, there is a very good agreement of our simulations with other studies on AMOC slowdown under anthropogenic forcing [1,2,5,7,26]. This study shows large-scale patterns for near surface temperature and zonal wind anomalies similar to state-of-the-art climate models with external forcing elements. This clearly suggests that the main mechanisms of the climate

response to the AMOC weakening are already contained in this simplified setup composed by an intermediate complexity model and performed with internal-variability driven dynamics only.

Mechanisms for AMOC slowdowns

The second part of the analysis is dedicated to the aspects of the internal climate variability that induce the AMOC weakening events in our simulations. In order to isolate the driving mechanisms of the AMOC slowdowns we examine the difference in the annual-mean climate between the ensemble members that are replicated versus the full set of ensemble members. In particular, we focus on the atmospheric driving mechanisms, i.e exchanges of heat, fresh water and momentum as discussed in Gregory et al (2016)[21], since it is the atmospheric state that drives climate variability in our rare events simulations (see Driving Mechanisms in Methods). The atmospheric internal variability is indeed a well-known key driver of interannual and multidecadal AMOC variability [11,38–44], via its impact on freshwater flux, heat exchanges and wind stress. However, its role in triggering abrupt changes or collapses of the AMOC is much less understood and atmosphere-ocean interactions that may prevent or favour the AMOC collapse have been both suggested [2,11,14,23].

The ensemble members that develop a weaker AMOC are characterised by stronger North Atlantic westerly winds in the core of the jet, increased net freshing of the northern North Atlantic, and a more complex pattern of turbulent heat flux anomalies (Fig 3). The strong ocean heat uptake in Labrador sea in Fig. 3c is probably due to the significant ice formation and latent heat release in the same area. However, the wind stress anomalies in the North Atlantic are by far the most intense signal, relative to the internal variability of the system.

Based on these results, we suggest that zonal wind stress is the main cause of the initial weakening of the AMOC via Ekman transport. Indeed, referring to Fig. 3a, the westerly wind anomalies at about 50N induce a net surface southward Ekman transport at the same latitudes (also see Fig S7), therefore opposing the AMOC. This is consistent with the important role of wind stress in driving AMOC anomalies in the interannual and seasonal time scale [38,41,43], and possibly even on a longer time scale [41].

The cause of the subsequent abrupt weakening of the AMOC is less clear, but we suggest it is linked to the gradual collapse of the convection over the Labrador Sea. The mentioned Ekman flow anomaly in the North Atlantic 35-55°N band is likely to contribute to this process,

by reducing the surface transport of hot-salty water directed toward the Labrador Sea (Fig S2a) and thus stabilising the water column. Indeed, the collapse of convection precedes the start of the AMOC abrupt weakening events in all the simulations (Fig S6). Note that simulations in which the convection collapse is slower are associated with a delayed AMOC abrupt transition, and with a higher value of the AMOC index at the end of the simulation. In this frame, the weakening of the convection in the Labrador Sea is thus the transmission belt between the wind stress signal and the AMOC abrupt weakening events.

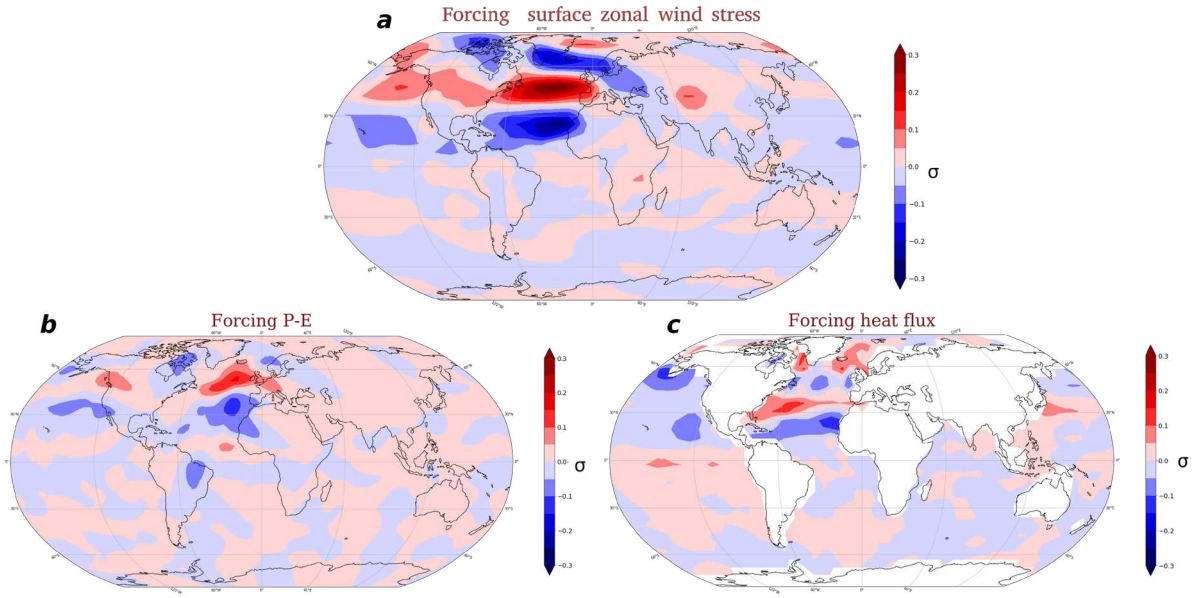

Fig. 3 *Possible atmospheric mechanisms driving the AMOC slowdown in the rare-event-algorithm simulations. This is obtained as the difference between the mean climate of the cloned members relative to the mean climate of all members (see methods). The panels show the anomalies in a) the zonal surface wind stress, b) the freshwater flux (difference between precipitation and evaporation) and , c) the net heat flux (turbulent, sensible and radiative) between the atmosphere and the ocean. Positive values reflect anomalies toward the ocean. All values are expressed in terms of sigma units, where sigma represents the amplitude of the inter-annual variability, as given by the averaged standard-deviation across all ensemble members for each resampling time step.*

Collapse, recovery, and critical thresholds

In the previous section it has been shown that internal climate variability is sufficient to

induce an abrupt slowdown of the AMOC. Now we ask if, after the abrupt transition, the system has tipped into a new stable state in which the AMOC is partially collapsed. This is tested by extending the previous simulations after year 125, but with the rare event algorithm turned off. In order to test the robustness of our finding we generate 20-member free-running ensemble simulations by randomly perturbing the sea level pressure. However, in order to reduce the computational costs, we only extend the first ensemble member from each of the ten rare-event-simulations. This setup leads to a 200-member free-running ensemble, featuring members starting from 10 different initial states.

The evolution of the AMOC in the free-running ensemble is given by the purple lines in Fig. 4a. Two opposite behaviours emerge, revealing a bimodal distribution in the evolution of the AMOC. For some ensemble members, the AMOC clearly bounces and overshoots, before relaxing to its starting value on a timescale of about 200 years. On the contrary, for other members, the AMOC keeps weakening even during the free run. This clearly shows that a noise-induced tipping has been achieved and the system evolves towards a new equilibrium state, at least on a time interval of 400 years. Only three ensemble members are able to pass from the collapsed to the active AMOC states in the course of the free-running simulations.

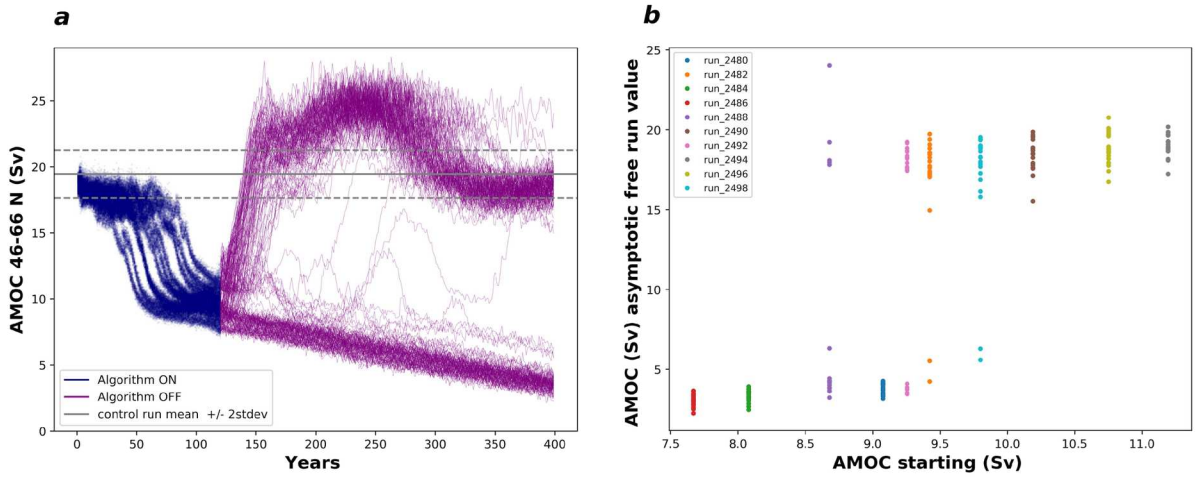

Fig. 4 *a) Evolution of the AMOC (purple lines) in a 200-member ensemble free-running extension of the simulations produced with the rare-event-algorithm (blue points, as in Fig 1b). The 200 members are composed of ten sub-ensemble of 20 members each. Each of the ten sub-ensembles branches at year 120 from one member of each of the ten simulations performed with the rare event algorithm. Each free-running simulation is run for an additional 280 years. b) Relationship between the final value of the AMOC (years 390-400) in the free-running simulations vs the value of the AMOC at the time it branched from the rare-event-algorithm simulations, for each of the 200 members*

In the light of the previous results one may wonder what are the climate aspects that determine whether the AMOC would eventually collapse or bounce back. A quantity that is considered fundamental for the behaviour of the AMOC is the AMOC strength itself, as the salt advection feedback implies that a very weak AMOC may not be able to recover. Indeed, other studies tried to determine a critical threshold or at least an indicator of the AMOC future qualitative evolution from the strength of the AMOC itself, though no robust conclusions were found [20,45] . Figure 4b explores this point by showing, for all 200 members, the value of the AMOC at the end of the 280-year free-running simulation as a function of the value at which the simulation started at year 120. All the ensemble members that start with an AMOC weaker than 8.5 Sv evolve towards an AMOC collapse. On the contrary, the AMOC bounces back and relaxes to the control run values in all members starting with an AMOC above 10 Sv. Between the two regimes, the evolution of the members is not unambiguously determined, and the same initial condition can either evolve into a collapse or a recovery of the AMOC. Since the different members, for any given starting AMOC, are initialised only via small random perturbations to the sea level pressure, the evolution of the system appears to be unpredictable. This implies the system is close to the edge of the two basins of attraction, and the chaotic atmospheric variability is sufficient to tip it in one state or the other.

These results seem to support the idea that the AMOC strength, consistent with the salt advection feedback, is a key parameter in order to determine the following behaviour of the system. However, when additional experiments are performed, it is found that these thresholds are not robust, rather they seem to be linked to the specific state reached at the end of the rare-event simulations. In particular, for the two rare-event simulations in which the AMOC is fully collapsed at year 120 (see Fig. 4b for runs 2480 and 2486), we performed nine additional "early-start" free-running ensemble simulations initialised before year 120, i.e at years 40, 50, 60,..., 110. The outcomes of the new experiments, though generally consistent with the previous results, reveal two initial conditions for which an AMOC collapse is achieved despite an AMOC starting value greater than 10 Sv. In particular members starting at year 40 of run 2486 (Fig. 5a, starting AMOC strength of 14.00 Sv) and members starting at year 60 of run 2480 (Fig. 5b, starting AMOC strength of 12.12 Sv) having already passed the tipping, shows an AMOC abrupt decrease as in the runs with the algorithm on.

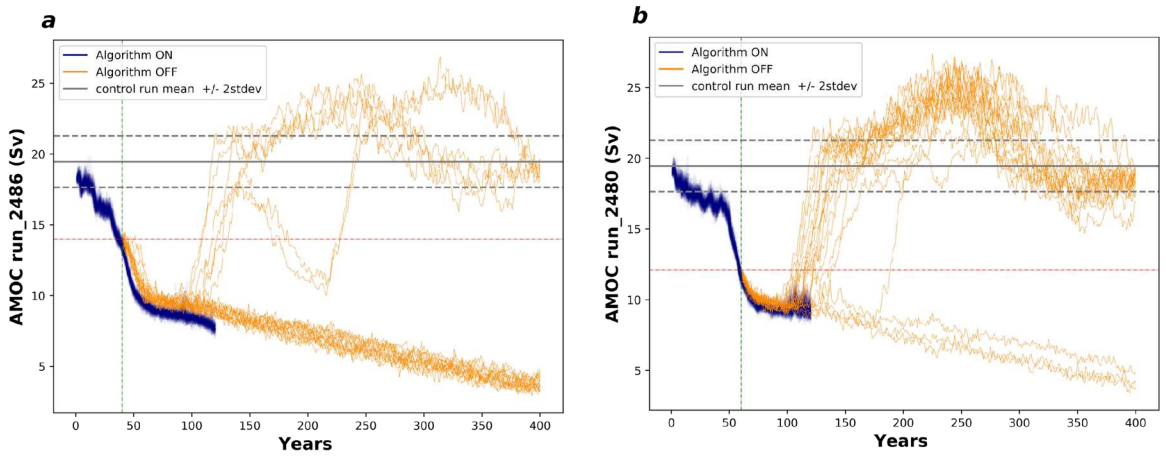

Fig. 5 *a) Example of early start experiment. The 20 orange lines represent AMOC ensemble members spontaneous evolution starting at years 40 of the run_2486 performed with the algorithm (blue points representing ensemble members evolution). Green vertical line is fixed at year 40, while the red horizontal line is plotted at 14.00 Sv. b) As in a) but for the 20 free members (orange lines) starting at year 60 of the run_2486 performed with the algorithm (blue points representing ensemble members evolution). Green vertical line is fixed at year 60, while the red horizontal line is plotted at 12.12 Sv*

The results from all the free-running ensemble simulations (Fig 6a) reveal that much care is needed before inferring the evolution of the AMOC based on the AMOC value itself. For example members starting near 14 Sv tend to collapse while the ones starting between 10 Sv and 12 Sv all recover. This is enough to reject the idea of a hard AMOC strength threshold that controls the behaviour of the system. Instead, we suggest that these results imply an increasingly small, but non zero, risk of AMOC collapse even for high starting values of the AMOC. This is consistent with the ability of the rare-event simulations to find internal trajectories leading to a collapsed AMOC starting from the control run at around 20 Sv. On the other hand, we show that there is no evidence of the recovery of the AMOC in trajectories in which the value of the AMOC reaches a value below 7.5 Sv (Fig 6b). This may be seen as a sufficient, but not necessary, condition to have an AMOC collapse in PLASIM-LSG since paths to collapse exist well above that value.

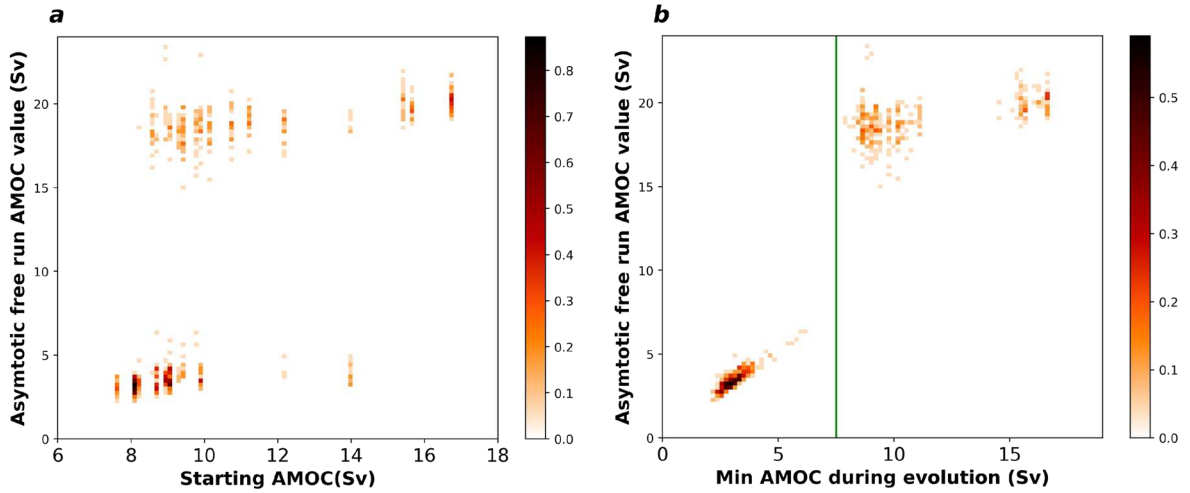

Fig 6 *a) As in Fig 4b, but also including the 18 free-running ensemble simulations initialised before year 120 of the rare-event simulations. The colours of the dot shows the density function. b) as in a) but against the weakest AMOC value crossed during the evolution of that member in the free run.*

DISCUSSION

This work has assessed for the first time the possibility of applying a Rare Event Algorithm to the study of noise-induced tipping events in a coupled climate model. We show that the rare event algorithm is able to identify AMOC collapses induced by internal atmospheric variability in the intermediate complexity PlaSim-LSG climate model, thus showing the potential of this novel approach to explore tipping events in the climate system.

With the approach proposed in this study, we are able to isolate the atmospheric processes driving the AMOC slowdown from the climate response to the weakened AMOC. We suggest that an intensified jet stream over the North Atlantic is the initial driver of the AMOC slowdown at the beginning of the rare-event simulation, via the induced oceanic Ekman circulation. Subsequently, after a few decades of atmospheric forcing, deep convection collapses in the Labrador Sea, leading to an abrupt weakening of the AMOC. Therefore, in this climate model, the repeated occurrence of an intensified jet stream for at least a few decades seems to be sufficient to lead to an AMOC collapse without possibility of recovery on multi-centennial time scales. The process driving the AMOC slowdown is likely to be dependent upon the model used, but at the same time, the applicability of the rare event algorithm to identify such processes in the different models is of general interest.

The weakened AMOC found at the end of the 125-year rare-event simulation is itself driving a cooling of North Hemisphere temperatures and a strengthening and poleward shift of the North Hemisphere jet streams. These responses are consistent with those found in experiments adopting an external forcing to weaken the AMOC in state-of-the-art climate models. Therefore, this suggests that the key elements of the response to an AMOC slowdown are already embodied in this simplified unforced context, thus supporting the use of EMICs to analyse these problems.

We conclude that:

- Rare event algorithms are very promising tools to test the multistability of a system and in particular to sample noise-induced tipping events.
- Atmospheric internal variability can lead to an abrupt weakening of the AMOC even without external forcing elements, but with an analogous climate response as in externally forced experiments.
- After 120 years of simulation performed with the rare event algorithm, some of these simulations have reached AMOC collapsed states from which the system does not recover in the following 280 years. This implies that the rare event algorithm is able to sample the AMOC collapsed states.

The proposed approach can be easily applied to a large variety of phenomena once one has defined the quantity that needs to be minimised or maximised to induce a tipping event. In particular, the algorithm offers the opportunity of generating a large number of tipping events without introducing any external forcing. This can be very useful in studies on tipping points in which it is unknown a priori what may trigger the events. On the contrary, as we did, we are able to use the trajectories generated with the algorithm in order to infer the mechanisms driving the transitions.

The hypothesis that the AMOC value itself could control the following evolution of the system towards the active or collapsed AMOC state is not fully sustained by our data. Indeed, we have not been able to define universal thresholds on AMOC strength that characterise which trajectories are bound to collapse or recover. Consistently with the fact the algorithm finds paths to a collapsed AMOC from initial states at 20 Sv, spontaneous AMOC collapses in free runs can already be found from 14 Sv. On the contrary when the AMOC crosses a value of 7,5 Sv it is not able to recover anymore. Therefore, despite the relative simplicity of the model, the mechanism seems more complex than simply regulated by a single control parameter.

If on the one side there is great potential to apply this method to new tipping elements, there are nevertheless several aspects related to internal AMOC tipping events that still need to be better assessed. First of all, the hypothesis on the identified atmospheric driving mechanism should be tested in a state-of-the-art climate model. Furthermore what is presented in this paper could be integrated in a system in which, for example, the GHGs concentration are perturbed. This would allow to explore the effects of GHGs increase on the risk of a spontaneous collapse of the AMOC at different levels of global warming.

METHODS

The Climate Model: PlaSim coupled to LSG

We perform numerical experiments using the Planet Simulator [33,34] General Circulation Model (PlaSim) coupled to the Large-Scale Geostrophic Ocean model (LSG) [36,36]. PlaSim-LSG, in different configurations, has been already used in several studies ranging from aquaplanets, to paleoclimatic reconstructions, future climate projections and AMOC variability [11,39,46]. PlaSim is based on the wet primitive equations representing the conservation of momentum and mass, the first law of thermodynamics and the equation of state, simplified by the hydrostatic approximation [34]. PlaSim includes parameterizations of sea ice and land processes, as well as a slab ocean model. The slab ocean acts as a thermal bath and water reservoir for absorption and release of heat and freshwater but has no representation of oceanic currents.

LSG is a simple ocean model that accounts for the geostrophic dynamics of the ocean on spatial scales larger compared to the internal Rossby radius as well as on time scales larger compared to the typical period of gravity modes and barotropic Rossby wave modes. The model implicitly solves the primitive equations neglecting the nonlinear terms in the Navier-Stokes equation. The interaction between PlaSim and LSG occurs through the slab ocean layer of PlaSim (50 m) that determines average 10-day temperature to be used as the uppermost layer for LSG, while average 10-day fluxes of momentum and freshwater are given by the sea ice and atmospheric modules. In turn, the slab ocean of PlaSim is forced to relax towards the uppermost layer of LSG in the following 10 days of PlaSim evolution [11,37].

Following Mehling (2022)[39] and Angeloni (2022)[37], the coupled PlaSim-LSG model is run at T21 horizontal spectral resolution and 10 vertical levels in the atmosphere, 3.5 deg and 22 vertical levels in the ocean. The state of the AMOC in this model has been shown to be sensitive to the parameterisation of the vertical profile of the ocean vertical (diapycnal) diffusivity, in particular to the value of the diffusivity at the upper layers [37]. In this work we use the values of diffusivity proposed and tested in Angeloni (2022)[37] (see Supplementary Informations for further details). The AMOC has a stable behaviour on the millennial time scale, with values between 18-22Sv, and multi centennial oscillations up to 4 Sv (see Fig. 1a) . Overall, the model shows a satisfactory large-scale climatology given its low resolution, with model biases concerning surface temperature and sea ice cover being mostly confined to the Southern Hemisphere [37]. However, a relevant bias for this study is that North Atlantic oceanic currents don't follow the usual counterclockwise circulation around the North Atlantic subpolar gyre, but hot surface water tends to go more directly from the Gulf of Mexico to the strong convective area in the Labrador Sea (see Fig. S2 and Fig. 2e grey contours).

Greenhouse gases are taken into account by a single value of CO2-equivalent concentrations, which is kept fixed at 354 ppm. With this CO2 level and diffusivity parameter, Angeloni (2022)[37] demonstrated that the model has two AMOC stable states: a state with the already mentioned values of the AMOC and a weak AMOC state of about 5 Sv without oscillations. This makes PlaSim-LSG particularly suitable to explore whether it may be possible to find internally driven transitions to the collapsed state from the strong circulation state with the rare event algorithm. In Angeloni (2022)[37] no transition from one state to the other during millennial scale simulations have been found.

The use of an intermediate complexity model (EMIC) allows us to test our experimental setup, exploring the parameter phase space of the algorithm with low computational cost. It is important to stress that dealing with an EMIC can also be a way to isolate fundamental mechanisms in a given phenomenon. In other words, if we can obtain results comparable to General Circulation Models (GCMs) in a simplified climate model, then the key processes of that phenomenon must already be contained in the simplified model.

AMOC Metric

The North Atlantic meridional overturning streamfunction Ψ at latitude θ and depth z, is defined as :

$$\Psi_\theta(z) = -\int_{\varphi_{west}}^{\varphi_{est}} \int_{z_{bottom}}^{z} v_v r_T |\cos\theta| \, d\varphi \, dz' \; ; \qquad \text{[eq.1a]}$$

where Z_{bottom} is the oceanic depth, v_v the meridional velocity, φ the longitude angle and r_T the Earth radius, φ_{est} and φ_{west} represent respectively the east and west margin of the Atlantic Ocean basin at latitude θ. The AMOC strength is obtained as the maximum of the $\Psi_\theta(z)$ between 46° and 66°N and below 700m,

$$AMOC = max_{z,theta}\{\Psi_\theta(z)\}_{z>700\,m, \theta \in 46°-66°N} \quad . \qquad \text{[eq.1b]}$$

The AMOC strength is evaluated at every ocean model timestep (10 days) and then averaged to obtain an annual mean AMOC index.

Application of the Rare Event Algorithm

The general characteristics of the algorithm are described in several works [27-29]. Here we summarise the key aspects, and describe the implementation adopted for this study. In this work, an experiment with the rare event algorithm consists of an ensemble of 100 members initialised from slightly different initial conditions, that are obtained by applying small random perturbations to the sea level pressure of the same starting configuration. Each simulation is run for 125 years. At constant intervals of 1 year, which is the resampling time scale of the experiment, each trajectory i is associated to a weight defined as

$$W_i = \frac{e^{k \cdot AMOC_i}}{\Sigma_n e^{k \cdot AMOC_n}} \; ; \qquad \text{[eq.2]}$$

where *k* is a negative parameter (see below) and AMOC is the 1-year average AMOC index during the last resampling time. Before simulating the following resampling time interval, each ensemble member generates a number of clones proportional to its weight. The number of replication of the single trajectory *i* at step *n* for the next step *n+1*, is

$$c_{n,i} = integer\left[W_{n,i} \cdot N + r_{n,i}\right], \qquad \text{[eq.3]}$$

where *integer[ ]* defines the integer part, N is the number of members in each resampling step, and $r_{n,i}$ a random number between 0 and 1. Trajectories with low value of the AMOC index will generate a large number of replicas in the following resampling time (since k is negative), while trajectories with large values of the AMOC index will generate 0 replicas, being effectively killed. The selection and cloning is done in such a way that the number of ensemble members remains constant throughout the integration ($\Sigma_i c_{n,i} = N$). If this condition

is not met, the $c_{n,i}$ are adjusted by selecting random members to be killed or cloned. All the replicas are perturbed with a small random perturbation applied to the sea level pressure. This perturbation, of the relative order of $10^{-4} - 10^{-5}$, allows the system to explore the neighbourhood of the phase space. The generated trajectories will diverge from each other during the model simulation, allowing to explore possible alternative evolutions of climate variability starting from the parent trajectory. It should be clear that the small perturbation added is not the driving mechanism for the AMOC decrease rather it is just a way to explore the phase space.

The application of the algorithm will populate the ensemble of trajectories characterised by extreme low values of the AMOC averaged along the entire simulation length. The absolute value of the parameter k, that we remind being negative, determines how stringent the selection is, where the larger |k| the larger the number of trajectories discarded at every resampling step. We explored the sensitivity to the parameter k. The results shown in the paper refer to the parameter k set to -3 Sv^{-1}, but simulations were also performed for k= -1 Sv^{-1} and k= -10 Sv^{-1}. The results are not discussed in this paper, however we found that for k= -1 Sv^{-1} the selection is so loose that a significant slowdown can't emerge in 125 years, while for k= -10 Sv^{-1} the selection is so tight that we can have problems with local minima of the AMOC.

In order to test the robustness in our findings we performed 10 independent ensemble simulations, each with 100 members and 125-year long. The 10 ensemble simulations are initialised from 10 different years of the control run in the period 2480-2500 every two years, i.e. 2480, 2482, 2484, 2486,…, 2498.

Driving Mechanisms

The algorithm selects and clones trajectories at the end of every 1-year resampling time interval. The driving mechanisms able to induce a weakening of the AMOC in a time scale of one year are statistically preferred to those that develop in a longer time scale. Considering that, in order to generate the ensemble members, the algorithm adds stochastic perturbations to the atmosphere only, and since the dynamical evolution of the atmosphere is much faster than the one of the ocean, we conclude that the initial driving mechanism of the AMOC slowdown resides in atmospheric signals which are subsequently transferred to the ocean. Each ensemble member, due to the chaotic nature of the atmosphere, will evolve in a different trajectory, leading to different impacts on the AMOC.

The mechanisms driving the AMOC slowdowns are evaluated by examining the difference in the annual-mean climate between the ensemble members that are replicated versus the full set of ensemble members, where the signal of the replicated members is weighted over the number of replications generated for the following resampling time step. This difference is computed for every year, and averaged over the 125 years of simulation with the rare event algorithm. Finally, the mean difference is scaled over the internal variability of the system (sigma units), computed as the average annual-mean variability of the full set of ensemble-members. Thus we are able to compare the relative size of the anomalies across different observables. In summary, if A is an annual-mean quantity,

$$A_{driving} = \frac{1}{years} \cdot \Sigma_n^{years} \frac{\left[A\right]^{(n)} - \langle A^{(n)} \rangle}{\sigma^{(n)}}; \qquad \text{[eq.4]}$$

where

$$\left[A\right]^{(n)} = \Sigma_i^{members} A_i^{(n)} \cdot c_i^{(n)} \quad ; \quad \langle \psi \rangle = \Sigma_i^{members} \frac{\psi_i}{members} \; ; \; \sigma^{(n)} = \sqrt{\left\langle \left(A^{(n)} - \left\langle A^{(n)} \right\rangle \right)^2 \right\rangle}$$

$members = 100$; $years = 125$.

DATA AVAILABILITY

The datasets generated and analysed during the current study will be available in a open repository with a persistent link.

ACKNOWLEDGEMENTS

This is TiPES contribution #203; the TiPES (Tipping Points in the Earth System) project has received funding from the European Union's Horizon 2020 research and innovation program

under grant agreement No. 820970.
We thank Michela Angeloni assistance with PlaSIM-LSG and for providing the PlaSIM-LSG pre-industrial control simulations.

590

AUTHORS CONTRIBUTION

GZ and SC conceived the study; MC, GZ and FR designed the experiments; MC run the experiments and performed the data analysis; All authors contributed to interpret the results; MC drafted the paper, with contributions from SC, FR and GZ.

595

COMPETING INTERESTS

The authors declare no competing interests.

600 ## SUPPLEMENTARY INFORMATION

Oceanic vertical diffusion parametrization

The state of the AMOC in the PlaSim-LSG model has been shown to be sensitive to the
605 parameterisation of the vertical profile of the ocean vertical (diapycnal) diffusivity, in particular to the value of the diffusivity at the upper layers.

$$A_V(z) = a_* + a_{range} \arctan ,$$

610 Where *z* represents the vertical oceanic coordinate. In our setup vertical diffusion parameterization values are:

$$z_* \; 2500\,m, \lambda = 4.5 \cdot 10^{-3} m^{-1}, a_* \; 0.8714 \cdot 10^{-4}, a_{range} = 0.2843 \cdot 10^{-4}$$

615 ***Supplementary Plots***

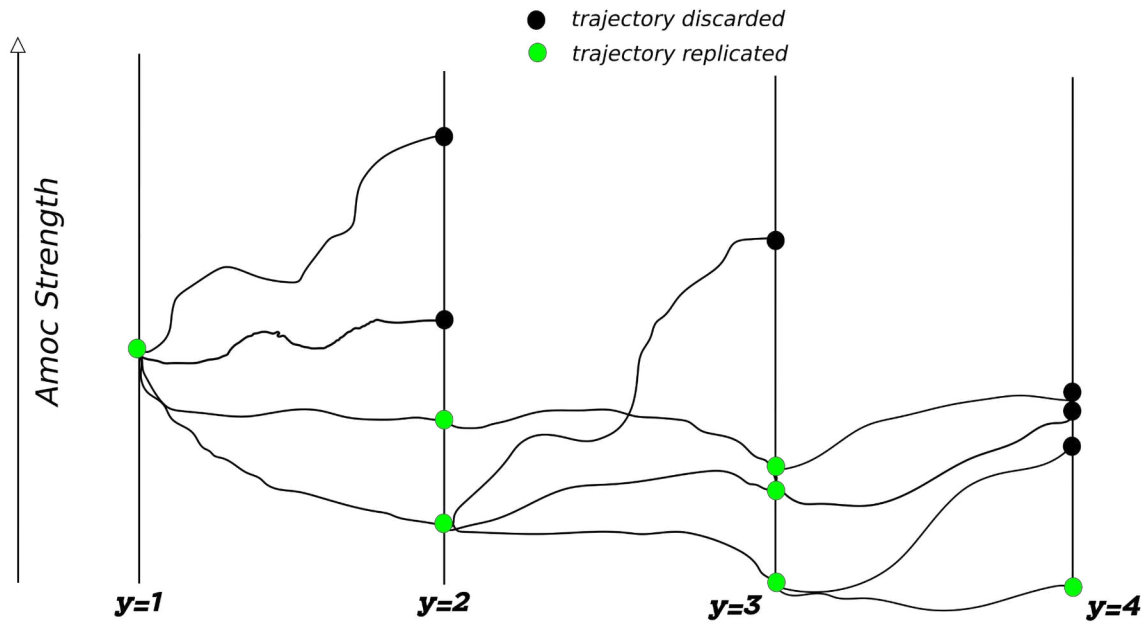

Fig. S1 *Example of how the algorithm works assuming, for the sake of presentation, an ensemble of only 4 members. Starting from the initial condition (y=1) the algorithm initialises 4 members via 4 different small random perturbations of the atmospheric part of the system. After a resampling time of 1 yr, the simulation is stopped and the algorithm assigns to each trajectory a weight that is a function of the mean AMOC in the previous year. Then, according to these weight, the algorithm initalitses a variable number of trajectories for the next resampling step, in a way that the weaker is the AMOC the greater is the number of members that are replicated. The trajectories with a higher value of the AMOC are not replicated and hence are discarded*

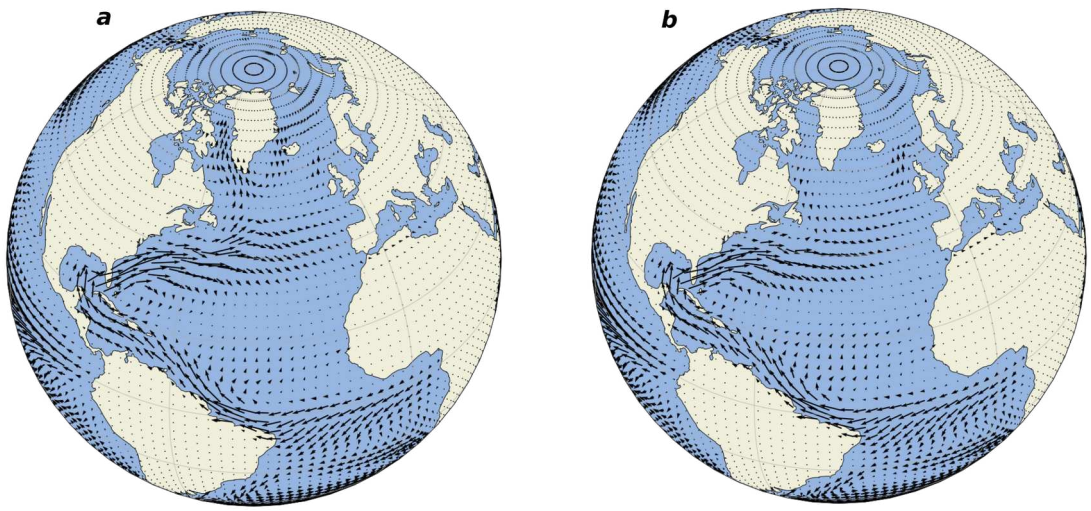

Fig. S2 *Surface oceanic currents in control run (left) and at the end of 125 years run with the rare event algorithm on (right). Data represents mean year value for currents and it was interpolated from LSG 76x72 grid into a 64x128 grid. Left figure represents the mean of the last 30 years of the control run while the right figure represents the mean of the 10 simulations at year 125 (i.e. the end of the simulations with the algorithm on). Deep changes are limited to the region above 40°. We can see in the control run the pattern is characterised by a stream that goes directly from Florida to Labrador sea. At the end of the simulation the northern branch of this stream and the stream from Iceland are turned off. It seems that the previous surface current going from the Gulf of Mexico to the Labrador sea now stops at 40°N*

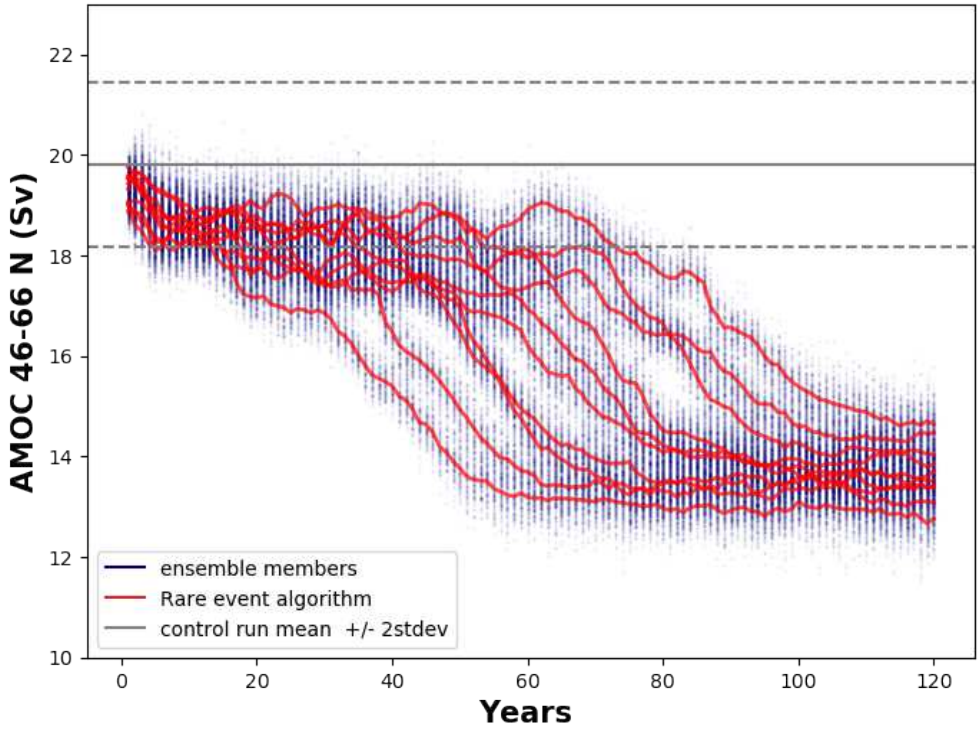

Fig. S3 *Equivalent to Fig 2b but the AMOC maximum is evaluated over the entire North Atlantic. The qualitative behaviour is the same, but the weakened AMOC states remain 4-5 Sv more intense that in the case of Fig 2b.*

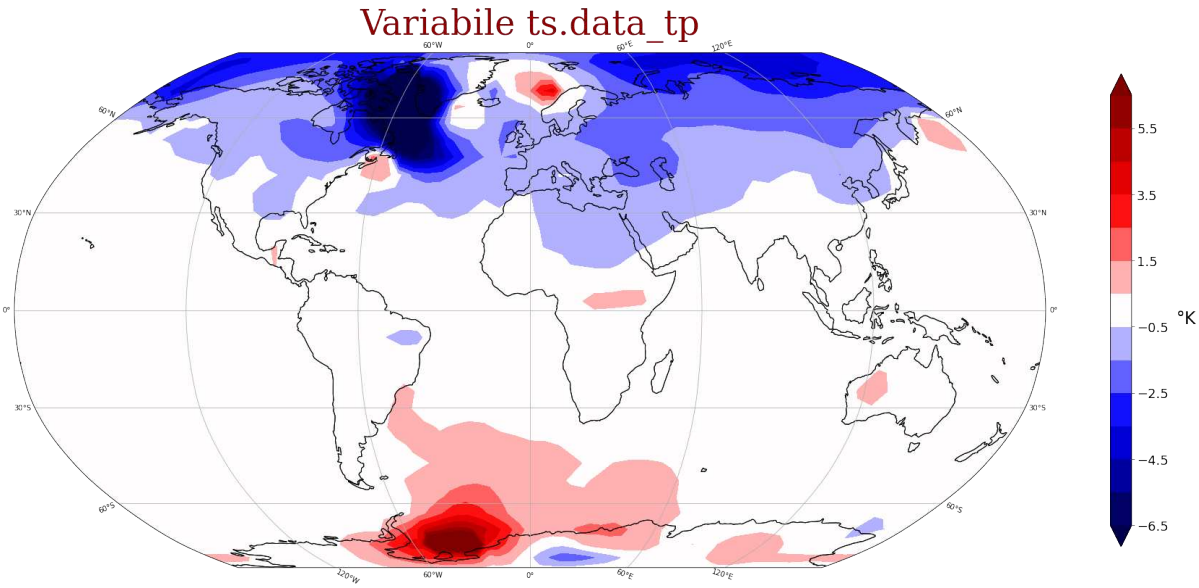

Fig. S4 *Surface temperature differences between the last 10 years average around year 400 of the 3 simulations with collapsed AMOC in the free run (Fig 4 run_2480, run_2484, run_2486) and the last 10 ys average of the control run*

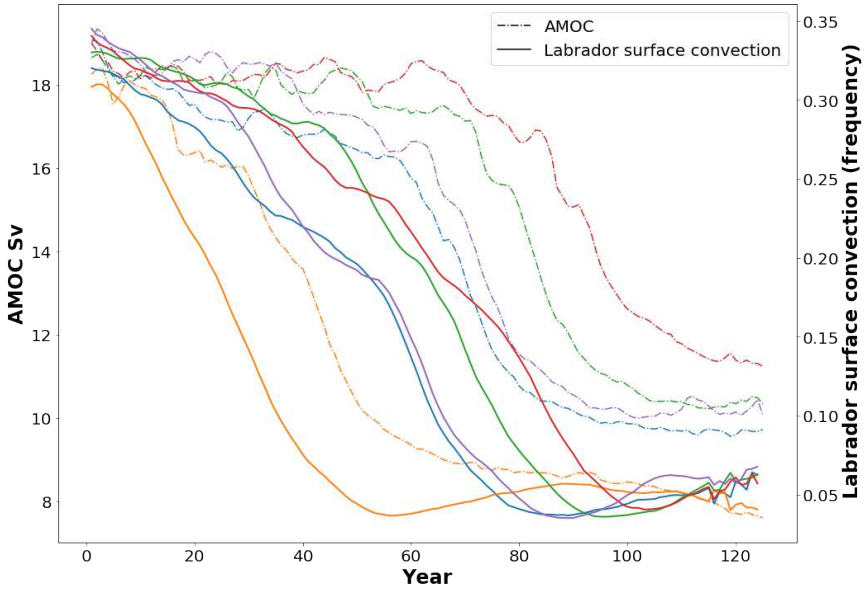

Fig. S5 *Time series of the surface convection frequency in the Labrador sea (full lines) and of the AMOC index (dashed lines) for 5 simulations. Each simulation is marked with a*

different colour.

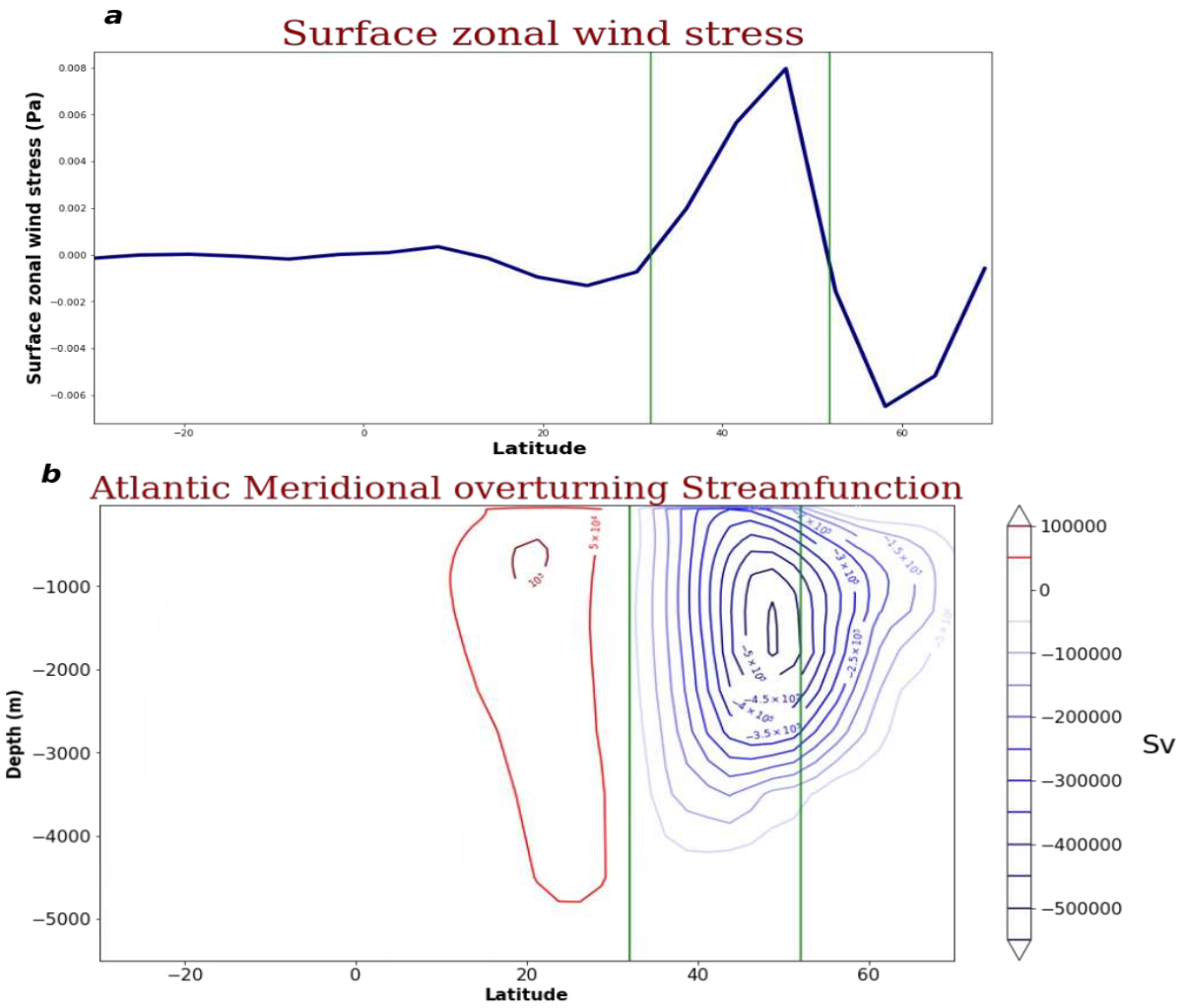

Fig. S6 *a) As in Fig 3, but for a) the meridional profile of surface zonal wind stress averaged over the North Atlantic ocean, and b) the Atlantic meridional overturning streamfunction. Vertical green lines are plotted at 32° and 52°N respectively. Within this band, the overturning stream function anomaly is qualitatively consistent with the effect induced by the Ekman transport induced by the zonal wind. However, other processes are also likely to contribute, leading to an extension of the overturning streamfunction anomaly poleward of 52N.*

Bibliography

1. 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.* **12**, 1–10 (2021).
2. Weijer, W. *et al.* Stability of the Atlantic Meridional Overturning Circulation: A Review and Synthesis. *J. Geophys. Res. Oceans* **124**, 5336–5375 (2019).
3. Masson-Delmotte, V., P. Zhai et al. *IPCC, 2021. Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. (Cambridge University Press, 2023). doi:10.1017/9781009157896.
4. Smeed, D. A. *et al.* The North Atlantic Ocean Is in a State of Reduced Overturning. *Geophys. Res. Lett.* **45**, 1527–1533 (2018).
5. Liu, W., Fedorov, A. V., Xie, S. P. & Hu, S. Climate impacts of a weakened Atlantic meridional overturning circulation in a warming climate. *Sci. Adv.* **6**, 1–9 (2020).
6. Bellomo, K. *et al.* Impacts of a weakened AMOC on precipitation over the Euro-Atlantic

region in the EC-Earth3 climate model. *Clim. Dyn.* (2023) doi:10.1007/s00382-023-06754-2.

7. Jackson, L. C. *et al.* Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM. *Clim. Dyn.* **45**, 3299–3316 (2015).
8. Lenton, T. M. *et al.* Tipping elements in the Earth's climate system. *Proc. Natl. Acad. Sci.* **105**, 1786–1793 (2008).
9. Boers, N. Observation-based early-warning signals for a collapse of the Atlantic Meridional Overturning Circulation. *Nat. Clim. Change* **11**, 680–688 (2021).
10. Liu, W., Xie, S.-P., Liu, Z. & Zhu, J. Overlooked possibility of a collapsed Atlantic Meridional Overturning Circulation in warming climate. *Sci. Adv.* **3**, e1601666 (2017).
11. Andres, H. J. & Tarasov, L. Towards understanding potential atmospheric contributions to abrupt climate changes: characterizing changes to the North Atlantic eddy-driven jet over the last deglaciation. *Clim. Past* **15**, 1621–1646 (2019).
12. Boulton, C. A., Allison, L. C. & Lenton, T. M. Early warning signals of atlantic meridional overturning circulation collapse in a fully coupled climate model. *Nat. Commun.* **5**, 1–9 (2014).
13. Vettoretti, G., Ditlevsen, P., Jochum, M. & Rasmussen, S. O. Atmospheric CO2 control of spontaneous millennial-scale ice age climate oscillations. *Nat. Geosci.* **15**, 300–306 (2022).
14. Klockmann, M., Mikolajewicz, U., Kleppin, H. & Marotzke, J. Coupling of the Subpolar Gyre and the Overturning Circulation During Abrupt Glacial Climate Transitions. *Geophys. Res. Lett.* **47**, e2020GL090361 (2020).
15. Henry, L. G. *et al.* North Atlantic ocean circulation and abrupt climate change during the last glaciation. *Science* **353**, 470–474 (2016).
16. Lippold, J., Dokken, T. & Thil, F. Events During the Last Glacial. 1–8 (2015) doi:10.1002/2014GL062512.Received.
17. Rahmstorf, S. Ocean circulation and climate during the past 120,000 years. *Nature* **419**, 207–214 (2002).

18. Broecker, W. S., Bond, G., Klas, M., Bonani, G. & Wolfli, W. A salt oscillator in the glacial Atlantic? 1. The concept. *Paleoceanography* **5**, 469–477 (1990).

19. Dansgaard, W. *et al.* Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* **364**, 218–220 (1993).

20. Jackson, L. C. *et al. Understanding AMOC stability: the North Atlantic Hosing Model Intercomparison Project*. https://gmd.copernicus.org/preprints/gmd-2022-277/ (2022) doi:10.5194/gmd-2022-277.

21. Gregory, J. M. *et al.* The Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP) contribution to CMIP6: Investigation of sea-level and ocean climate change in response to CO2 forcing. *Geosci. Model Dev.* **9**, 3993–4017 (2016).

22. Lohmann, J. & Ditlevsen, P. D. Risk of tipping the overturning circulation due to increasing rates of ice melt. (2021) doi:10.1073/pnas.2017989118/-/DCSupplemental.y.

23. Li, C. & Born, A. Coupled atmosphere-ice-ocean dynamics in Dansgaard-Oeschger events. *Quat. Sci. Rev.* **203**, 1–20 (2019).

24. Rheology, P. & Rheology, K. M. An abrupt climate event in a coupled ocean±atmosphere simulation without external forcing. *Nature* **4997**, 2–5 (2001).

25. Drijfhout, S., Gleeson, E., Dijkstra, H. A. & Livina, V. Spontaneous abrupt climate change due to an atmospheric blocking-Sea-Ice-Ocean feedback in an unforced climate model simulation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 19713–19718 (2013).

26. Goosse, H., Renssen, H., Selten, F. M., Haarsma, R. J. & Opsteegh, J. D. Potential causes of abrupt climate events: A numerical study with a three-dimensional climate model. *Geophys. Res. Lett.* **29**, 7-1-7–4 (2002).

27. Orbe, C. *et al.* Atmospheric Response to a Collapse of the North Atlantic Circulation Under A Mid-Range Future Climate Scenario: A Regime Shift in Northern Hemisphere Dynamics. *J. Clim.* **1**, 1–52 (2023).

28. Romanou, A. *et al.* Stochastic Bifurcation of the North Atlantic Circulation Under A Mid-Range Future Climate Scenario With The NASA-GISS ModelE. *J. Clim.* **1**, 1–49 (2023).

29. Ragone, F., Wouters, J. & Bouchet, F. Computation of extreme heat waves in climate

models using a large deviation algorithm. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 24–29 (2018).

30. Ragone, F. & Bouchet, F. Computation of Extreme Values of Time Averaged Observables in Climate Models with Large Deviation Techniques. *J. Stat. Phys.* **179**, 1637–1665 (2020).
31. Ragone, F. & Bouchet, F. Rare Event Algorithm Study of Extreme Warm Summers and Heatwaves Over Europe. *Geophys. Res. Lett.* **48**, e2020GL091197 (2021).
32. Castellana, D., Baars, S., Wubs, F. W. & Dijkstra, H. A. Transition Probabilities of Noise-induced Transitions of the Atlantic Ocean Circulation. *Sci. Rep.* **9**, 20284 (2019).
33. Fraedrich, K., Jansen, H., Kirk, E., Luksch, U. & Lunkeit, F. The Planet Simulator: Towards a user friendly model. *Meteorol. Z.* **14**, 299–304 (2005).
34. Lunkeit, F., Borth, H., Bottinger, M. & , Fraedrich, K., Jansen, H., Kirk, E., Kleidon, A., Luksch, U., Paiewonsky, P., Schubert, S., Sielmann, S., and Wan, H. Planet Simulator-reference manual, version 16. https://www.mi.uni-hamburg.de/en/arbeitsgruppen/theoretische-meteorologie/modelle/sources/psreferencemanual-1.pdf (2011).
35. Maier-Reimer, E. & Mikolajewicz, U. *The Hamburg large scale geostrophic ocean general circulation model Cycle 1*. 38 (1992).
36. Maier-Reimer, E., Mikolajewicz, U. & Hasselmann, K. Mean Circulation of the Hamburg LSG OGCM and Its Sensitivity to the Thermohaline Surface Forcing. *J. Phys. Oceanogr.* **23**, 731–757 (1993).
37. Angeloni, M. Climate Variability in an Earth System Model of intermediate complexity: from interannual to centennial time scale. (2022).
38. Polo, I., Robson, J., Sutton, R. & Balmaseda, M. A. The importance of wind and buoyancy forcing for the boundary density variations and the geostrophic component of the AMOC at 26°N. *J. Phys. Oceanogr.* **44**, 2387–2408 (2014).
39. Mehling, O., Bellomo, K., Angeloni, M., Pasquero, C. & von Hardenberg, J. High-latitude precipitation as a driver of multicentennial variability of the AMOC in a climate model of

intermediate complexity. *Clim. Dyn.* (2022) doi:10.1007/s00382-022-06640-3.

40. Delworth, T. L. & Zeng, F. The Impact of the North Atlantic Oscillation on Climate through Its Influence on the Atlantic Meridional Overturning Circulation. *J. Clim.* **29**, 941–962 (2016).
41. Kostov, Y. *et al.* Distinct sources of interannual subtropical and subpolar Atlantic overturning variability. *Nat. Geosci.* **14**, 491–495 (2021).
42. Leroux, S. *et al.* Intrinsic and Atmospherically Forced Variability of the AMOC: Insights from a Large-Ensemble Ocean Hindcast. *J. Clim.* **31**, 1183–1203 (2018).
43. Hirschi, J. & Marotzke, J. Reconstructing the meridional overturning circulation from boundary densities and the zonal wind stress. *J. Phys. Oceanogr.* **37**, 743–763 (2007).
44. Roberts, C. D., Garry, F. K. & Jackson, L. C. A multimodel study of sea surface temperature and subsurface density fingerprints of the Atlantic meridional overturning circulation. *J. Clim.* **26**, 9155–9174 (2013).
45. Jackson, L. C. & Wood, R. A. Hysteresis and Resilience of the AMOC in an Eddy-Permitting GCM. *Geophys. Res. Lett.* **45**, 8547–8556 (2018).
46. Hertwig, E., Lunkeit, F. & Fraedrich, K. Low-frequency climate variability of an aquaplanet. *Theor. Appl. Climatol.* **121**, 459–478 (2015).

820

825