https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Past, Present, and Future Variability of Atlantic Meridional Overturning Circulation in CMIP6 Ensembles

Arthur Coquereau[1], Florian Sévellec[1,2], Thierry Huck[1], Joël J.-M. Hirschi[3], and Quentin Jamet[4]

[1]Laboratoire d'Océanographie Physique et Spatiale, Univ Brest CNRS IRD Ifremer, Brest, France
[2]ODYSSEY Project-Team, INRIA CNRS, Brest, France
[3]Marine Systems Modelling, National Oceanography Centre, Southampton, SO14 3ZH, UK
[4]SHOM, Service Hydrologique et Océanographique de la Marine, Brest, France

Correspondence: Arthur Coquereau (arthur.coquereau@univ-brest.fr)

Abstract. The Atlantic Meridional Overturning Circulation (AMOC) is a key component of the climate system, exhibiting strong variability across daily to millennial timescales and significantly influencing global climate. Sensitive to external conditions such as freshwater input, greenhouse gas concentrations, and aerosol forcing, important variations of the AMOC can be triggered by anthropogenic emissions. This study presents a comprehensive analysis of sources of AMOC variance in state-of-the-art climate ensemble models. By decomposing the effects of scenario, model, ensemble, and time variability, along with their interactions, through an Analysis of Variance (ANOVA), we identify three distinct regimes of AMOC variability from 1850 to 2100. The first regime, spanning most of the historical period, is characterized by a relatively stable AMOC dominated by internal variability. The second regime, initiated by AMOC decline at the end of the 20^{th} century and lasting until mid-21^{st} century, is governed by a transient increase of time variability. Notably, the direct effect of forcing differences remains muted all along this regime, despite the start of emission-scenarios in 2015. The third regime, beginning around 2050, is marked by the emergence and rapid dominance of inter-scenario variability. Throughout the simulations, model variability remains the primary source of uncertainty, influenced by aerosol forcing response, AMOC decline magnitude, and the physical variability. A key finding of this work is the evidence that internal variability decreases simultaneously with AMOC intensity and seems proportional to emission-scenario intensity.

1 Introduction

The Earth's climate is a complex system made up of various intertwined components interacting with each other and varying over a wide range of temporal and spatial scales. Among the major components, the Atlantic Meridional Overturning Circulation (AMOC) plays a key role by controlling a substantial part of the poleward heat transport with major impacts on the surrounding regions (Srokosz et al., 2012). At 26°N, it transports about 1.2 PW (1 PW = 10^{15} W) representing 60% of the net poleward heat flux of the ocean and 30% considering both ocean and atmosphere (Ganachaud and Wunsch, 2000; Trenberth and Fasullo, 2017; Johns et al., 2023). Several remote influences have also been identified, with the North East Brazilian and Sahel rainfalls, or the Atlantic hurricanes activity to only name a few (Knight et al., 2006). Finally, the AMOC contributes significantly to carbon sequestration through deep-water formation at high latitudes (Zickfeld et al., 2008).

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Considering this leading role for the climate system, it appears critical to understand how the AMOC varies over time and identify the associated mechanisms. In this regard, a recent review provided a comprehensive description of the scales involved in AMOC variability (Hirschi et al., 2020). It shows that the variability unfolds over a large collection of scales, spanning from daily with synoptic weather events and near inertial gravity waves to interannual scales influenced by climate modes such as the North Atlantic Oscillation. In the middle, mesoscale eddies or large-scale baroclinic waves also play an important role. At the end of the spectrum, the AMOC has also been associated with abrupt climate events, such as Dansgaard-Oeschger and Heinrich events at millennial scale (Dansgaard et al., 1993; McManus et al., 2004; Böhm et al., 2015; Henry et al., 2016). These events have been associated with another important characteristic of the AMOC, namely its multi-stability (e.g., Sévellec and Fedorov, 2014). Indeed, a long literature has studied this aspect and shown that the AMOC is subject to collapses during which the system shifts from a first state with an intense circulation to another one where the circulation considerably weakens (Stommel, 1961; Lenton et al., 2008; Armstrong McKay et al., 2022).

Beyond these natural, internal, and chaotic variations of the AMOC, the system is also sensitive to external forcings, such as the anthropogenic emissions of aerosols and carbon in the atmosphere. These emissions can either reduce the surface temperature by increasing the reflection of incoming solar radiation (for the aerosols), or on the contrary, increase the temperature by intensifying the greenhouse effect (in the case of carbon dioxide or methane). In Subpolar North Atlantic regions, these change in surface temperatures can have a direct impact on the buoyancy of surface waters, and can therefore prevent or intensify the formation of deep-water that need sufficient density to sink. In addition, human activities can impact the freshwater input in the subpolar gyre, for instance by modifying precipitation regimes or by boosting the Greenland Ice Sheet melting rate, and thus further increase the surface buoyancy and decrease deep convection consequently (Gierz et al., 2015).

For our understanding of the future evolution of the AMOC and the potential impacts on the climate, it therefore appears crucial to study the different sources of variability and their role in the past, present, and future state of the AMOC. Our objective, in this work, is precisely to tackle this challenge and provide a comprehensive analysis of the AMOC variability in state-of-the-art climate models from the 6$^{\text{th}}$ Phase of the Coupled Model Intercomparison Project (CMIP6). In particular, we aim to separate the internal signal induced by internal modes of variability from the forced signal associated with anthropogenic fingerprint. The analysis takes advantage of a large body of work, over the last decades, dedicated to partitioning the variability (or the uncertainty) in climate simulations (Hawkins and Sutton, 2009, 2011; Yip et al., 2011; Lehner et al., 2020; Zhang et al., 2023).

In this study, we also benefit from an improvement over the last generation of CMIP models, namely the presence of relatively large ensemble simulations. Each ensemble consists of several simulations of a single model covering the same time period with the same forcing, but starting with different initial conditions. This initial difference together with the chaotic nature of the system, then, allows the simulations to sample the internal variability by covering the phase space of the system. The internal variability was previously computed as the time-variability over a given period. However, different works show that the phase-space of the climate can evolve over time and that the internal variability is everything but stationary when a forcing is applied (e.g., Cheng et al., 2016; MacMartin et al., 2016; Coquereau et al., 2024). It is therefore necessary to detach ourselves from this temporal dimension in order to accurately estimate this internal variability and its evolution. To evaluate the

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

role of forcing in climate variability, CMIP6 simulations provide different Shared Socio-Economic Pathways (SSP) allowing to investigate the response of the system under various forcing intensities. These scenarios extend after the 1850-2015 historical period up to 2100 (at least). The current analysis is based on 10 ensembles, each of them being produced with a different model, to draw a robust "model-independent" picture of the variability, and to investigate the sources of uncertainty associated with the representation of the AMOC in the different models.

To analyze this high-dimensional dataset and separate the different factors of variability, we used a proven and state-of-the-art method for climate datasets called Analysis of Variance (ANOVA, Zwiers, 1996; Wang and Zwiers, 1999; Hingray et al., 2007; Yip et al., 2011; Zhang et al., 2023). This method allows us to separate the total variability into the main effect of each dimension and the interactions among them. Zhang et al. (2023) computed an ANOVA decomposition involving three dimensions (3-way) and focusing on ensembles. They applied this decomposition to temperature and precipitation, with a separation between ensemble members, scenarios, and models. Here, we propose to incorporate the time dimension, crucial in our analysis, in a 4-way ANOVA. Indeed, time dimension is in the middle between internal and forced variability, and is influenced both by internal oscillations and by common trends. It is especially critical in the historical period where there is no forcing scenarios and where the forced variability is therefore not accounted for if the time dimension is not considered.

In the next section, we will present the ANOVA methodology and the data used for this study. Section 3 will be dedicated to the results. We will start with a general overview of the evolution of AMOC and the dominant sources of variance (3.1). Then, we will focus on the evolution of the different physical components of this variance during three periods of the time series representing three different regimes (3.2). The final part of the results section concerns model variability and associated uncertainty (3.3). Finally, in section 4, we will summarize and discuss the results in order to shed light on possible future evolutions of the AMOC, and to discuss their meaning from the point of view of uncertainty and predictability.

2 Materials and Methods

2.1 Material

The present work is based on state-of-the-art climate simulations from CMIP6. The AMOC intensity is derived from the maximum meridional overturning streamfunction in the Atlantic at 26°N. Here, we focus on ensemble simulations to investigate ensemble variance as a proxy of internal variability. Among the models, three large ensembles (with 25, 30, and 40 members) are available (Tab. 1), and we will mostly focus our analysis on these large ensembles. We, nonetheless, extend the analysis to seven smaller ensembles (3-6 members) to test the robustness of our results and improve the representation of model variability. The time series are separated in two parts: an historical period from 1850 to 2014 where the forcing are based on observations and a projection period from 2015 to 2100. For the 21st century, three major SSP are studied to estimate the forcing-scenario variability: SSP1-2.6 (“Sustainability”), SSP2-4.5 (“Middle of the road”) and SSP5-8.5 (“Fossil-fueled development”). Overall, four dimensions are investigated: model, scenarios, ensemble, and time.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Table 1. List of models and members used in the study.

Model	Ensemble size	Members	Reference
ACCESS-ESM1-5	40	r[1-40]i1p1f1	Ziehn et al. (2020)
MPI-ESM1-2-LR	30	r[1-30]i1p1f1	Olonscheck et al. (2023)
CanESM5	25	r[1-25]i1p2f1	Swart et al. (2019)
MRI-ESM2-0	5	r[1-5]i1p1f1	Yukimoto et al. (2019)
MIROC6	3	r[1-3]i1p1f1	Tatebe et al. (2019)
CESM2	3	r4i1p1f1, r10i1p1f1, r11i1p1f1	Danabasoglu et al. (2020)
UKESM-0-LL	5	r1i1p1f1, r2i1p1f1, r3i1p1f1 r4i1p1f1, r8i1p1f1	Sellar et al. (2019)
CNRM-CM6-1	6	r[1-6]i1p1f1	Voldoire et al. (2019)
IPSL-CM6A-LR	6	r1i1p1f1, r2i1p1f1, r3i1p1f1, r4i1p1f1, r6i1p1f1, r14i1p1f1	Boucher et al. (2020)
GISS-E2-2-G	5	r[1-5]i1p3f1	Rind et al. (2020)

2.2 Methods

2.2.1 Analysis of Variance

The Analysis of Variance (ANOVA) method allows to decompose and attribute the variance in a multidimensional data set. It enables to investigate the explanatory power of several qualitative factors (or dimensions) on a quantitative variable (here, the AMOC intensity). It returns the role of each individual factor effects associated with single dimensions and the role of interactions between them associated with two or more dimensions. As described in Berrington de González and Cox (2007), interaction occurs when the separate effects of the factors do not combine additively. The ANOVA is thus generally represented as a linear decomposition of the total variance, with main effects depending on single dimensions and interactions depending on multiple factors that cannot be separated. The ANOVA is particularly useful when investigating three or more dimensions, which is difficult with more classical analyzes as covariance or correlation. There are, however, connections between covariance and ANOVA 2^{nd} order interactions. Both methods assess relationships between variables. The covariance measures how two variables change together, while the interaction in the ANOVA captures how a combination of two dimensions explains the variability in a dataset that would not be captured by either factor alone. Therefore, both refer to a joint variability.

Originally, the one-dimension ANOVA (referred to as 1-way ANOVA) was first employed by Yates (1938) in research on agriculture. Since then, it has been extended to the 2-way ANOVA and extensively used in climate science to caracterize variability in simulations (e.g., Zwiers, 1996; Wang and Zwiers, 1999). In particular, different works used ANOVA to investigate the sources of uncertainty in climate projections, to separate model and scenario uncertainties in CMIP analyzes (Yip et al., 2011), but also in more regional contexts (Hingray et al., 2007). More recently, a study took advantage of the larger ensembles

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

available in CMIP6 to introduce the ensemble dimension directly in a 3-way ANOVA decomposition (Zhang et al., 2023). It therefore provides a separated dimension for internal variability, whereas this latter was previously represented as the residual of the decomposition.

In the present work, we build upon the previous 3-way ANOVA (Zhang et al., 2023) and extend it to a 4-way formulation by adding time dimension to investigate the importance of changes due to dynamical adjustment (i.e., temporal variability). This is done by applying a 30-year rolling window (a typical climate period, discussed below) for the decomposition.

The main idea behind ANOVA is to separate the variance in a multidimensional data set by computing the variance after averaging the data over some of the dimensions. At the origin, the multidimensional dataset is not averaged and varies over the four dimensions $X(s,m,r,t)$, where s represents scenario dimension, m represents various models, r the different realizations associated with the ensemble dimension, and t represents time. Then, the running window method is implemented by successively extracting 30-year subsets x_τ from the original dataset, each centered around a given time point τ, as follows:

$$x_\tau(s,m,r,t) = X(s,m,r,t) \quad \text{for} \quad t \in [\tau - 15, \tau + 15]. \tag{1}$$

In the analysis, it is important to acknowledge the two time axes: t, on which statistical analyses are computed, and τ, which is 30 years shorter and represents the successive windows of analysis. When averaging over a dimension, the data set cannot varies in that dimension, this is signaled by the bar over the x and by removing the dimension from the parentheses: for instance, $\overline{x_\tau}(s,r,t)$ represents an average among models, evaluated at time τ (See Section 2.2.2 for technical details). This central idea of ANOVA can be implemented for the four dimensions and lead to different remaining spread in the dataset after averaging some of the dimensions (Fig. 1).

We will now use a few specific examples to detail the calculation of the main effect and interactions. The formulations can be applied to all individual dimensions or combinations of dimensions. The ensemble main effect contribution (V_r^{main} or called R in the paper) is the variance of the dataset with respect to the ensemble of realizations, when averaging all dimensions but r. This corresponds to:

$$R(\tau) = V_r^{\text{main}}(\tau) = \frac{1}{N_r} \sum_{r=1}^{N_r} [\overline{x_\tau}(r) - \overline{x_\tau}]^2 \tag{2}$$

To compute the interactions, we use the total variance budget involving several dimensions. The total variance involving ensemble and time (r and t), for instance, corresponds to:

$$V_{rt}^{\text{total}}(\tau) = \frac{1}{N_r N_t} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} [\overline{x_\tau}(r,t) - \overline{x_\tau}]^2 \tag{3}$$

This total variance budget is also equal to the sum of the main effects of r and t and their interaction:

$$V_{rt}^{\text{total}}(\tau) = V_r^{\text{main}}(\tau) + V_t^{\text{main}}(\tau) + V_{rt}^{\text{interaction}}(\tau) \tag{4}$$

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

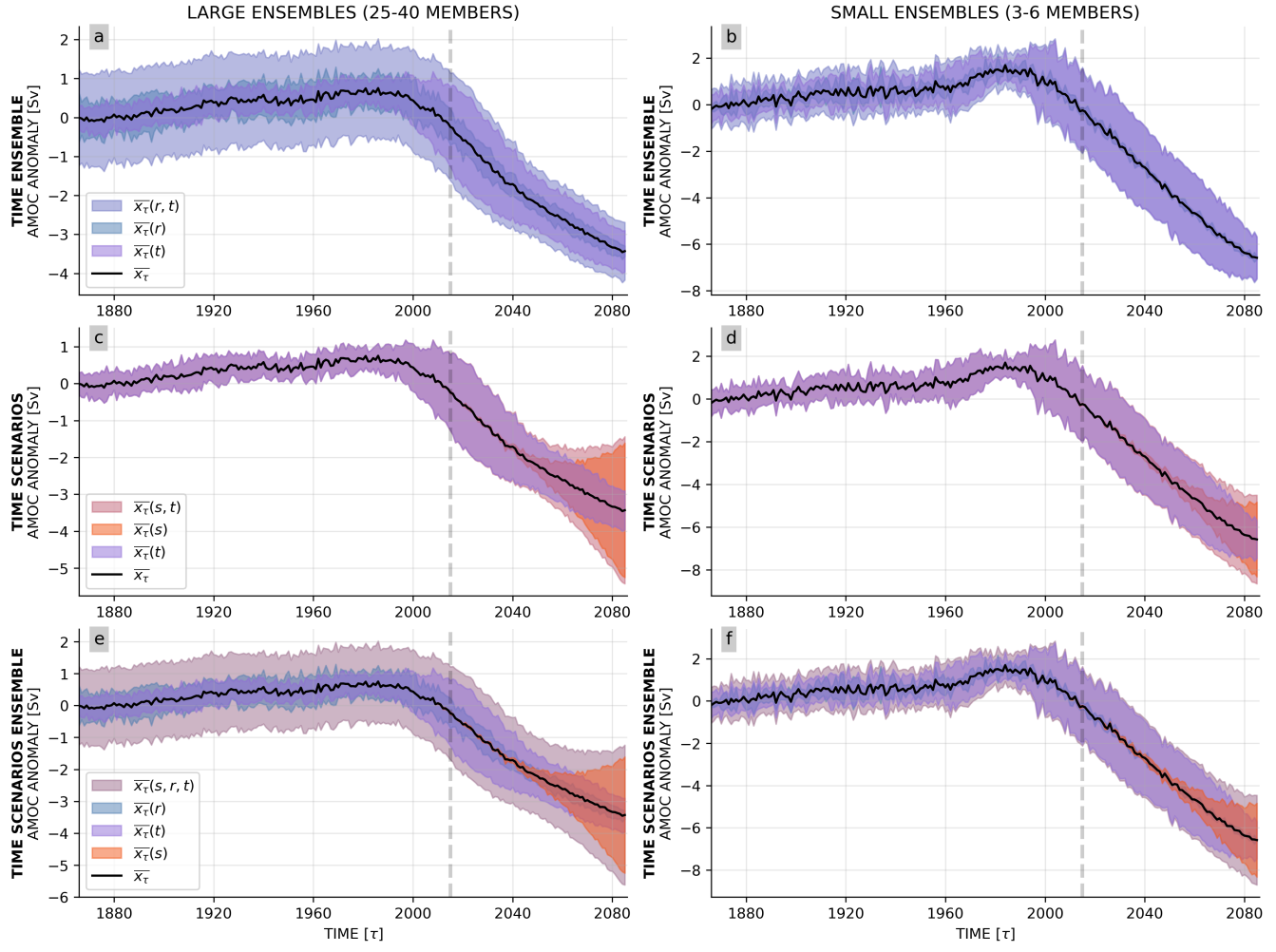

Figure 1. Example of multidimensional dataset used for the study and obtained from a single bootstrap with a 30-year running window. The dataset is then averaged over some dimensions inside the 30-year time window to show the remaining spread in the non-averaged dimensions. This running window allows us to study the temporal evolution of time-averaged quantities. The spread is characterized by a ± 1.96 standard deviation envelope around the average, representing approximately a 95% confidence interval. (a, b) Variability associated with time and ensemble dimensions. The three envelopes show: the remaining spread when averaging over scenarios and models leaving realizations and time ($\overline{x_\tau}(r,t)$); when averaging over scenarios, models and time leaving realizations ($\overline{x_\tau}(r)$); and when averaging over scenarios, model and realizations leaving time ($\overline{x_\tau}(t)$). (c, d) Variability associated with time and scenario dimensions. Envelopes show: the remaining spread when averaging over realizations and models leaving scenarios and time ($\overline{x_\tau}(s,t)$); when averaging over models, realizations and time leaving scenarios ($\overline{x_\tau}(s)$); and when averaging over scenarios, model and realizations leaving time ($\overline{x_\tau}(t)$). (e, f) Variability associated with time, scenario and ensemble dimensions. The four envelopes show: the remaining spread when averaging over models leaving scenarios, realizations and time ($\overline{x_\tau}(s,r,t)$); when averaging over scenarios, models and time leaving realizations ($\overline{x_\tau}(r)$); when averaging over models, realizations and time leaving scenarios ($\overline{x_\tau}(s)$); and when averaging over scenarios, model and realizations leaving time ($\overline{x_\tau}(t)$). Results are presented for large ensembles (a, c, e) with 25-40 members and small ensembles (b, d, f) with 3-6 members. The average among all dimensions is represented by the thin black line.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Thus, the 2nd order interaction between ensemble and time (called RT in the paper) can be obtain with:

$$RT(\tau) = V_{rt}^{\text{interaction}}(\tau) = V_{rt}^{\text{total}}(\tau) - V_{r}^{\text{main}}(\tau) - V_{t}^{\text{main}}(\tau)$$
$$= \frac{1}{N_r N_t} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} \left[\overline{x_\tau}(r,t)^2 - \overline{x_\tau}(r)^2 - \overline{x_\tau}(t)^2 - \overline{x_\tau}^2 - 2\overline{x_\tau}\left[\overline{x_\tau}(r,t) - \overline{x_\tau}(r) - \overline{x_\tau}(t)\right]\right] \quad (5)$$

We note that by averaging, the three terms in the last parenthesis will all become equal to $\overline{x_\tau}$. Thus, the expression can be simplified in:

$$RT(\tau) = \frac{1}{N_r N_t} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} \left[\overline{x_\tau}(r,t)^2 - \overline{x_\tau}(r)^2 - \overline{x_\tau}(t)^2 + \overline{x_\tau}^2 \right] \quad (6)$$

Similarly, the more classical 2nd order interaction formulation of ANOVA (presented in (7) or in Zhang et al., 2023) is equal to (6), as the cross terms disappear when averaging:

$$RT(\tau) = \frac{1}{N_r N_t} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} \left[\overline{x_\tau}(r,t) - \overline{x_\tau}(r) - \overline{x_\tau}(t) + \overline{x_\tau} \right]^2 \quad (7)$$

In the analysis, these interactions are described by the initial of the involved dimensions (e.g., RT corresponds to the interaction between time and ensemble dimensions). For higher-order interactions, we use the same principle by removing lower order terms from the total variance budget. For 3rd order interactions, SRT, for example, reads:

$$SRT = V_{srt}^{\text{interaction}} = V_{srt}^{\text{total}} - V_{sr}^{\text{interaction}} - V_{st}^{\text{interaction}} - V_{rt}^{\text{interaction}}$$
$$- V_{s}^{\text{main}} - V_{r}^{\text{main}} - V_{t}^{\text{main}} \quad (8)$$

and for 4th order, $SMRT$, for example, reads:

$$V_{smrt}^{\text{interaction}} = V_{smrt}^{\text{total}} - V_{smr}^{\text{interaction}} - V_{smt}^{\text{interaction}} - V_{srt}^{\text{interaction}}$$
$$- V_{mrt}^{\text{interaction}} - V_{sm}^{\text{interaction}} - V_{sr}^{\text{interaction}} - V_{mr}^{\text{interaction}}$$
$$- V_{st}^{\text{interaction}} - V_{mt}^{\text{interaction}} - V_{rt}^{\text{interaction}}$$
$$- V_{s}^{\text{main}} - V_{m}^{\text{main}} - V_{r}^{\text{main}} - V_{t}^{\text{main}} \quad (9)$$

To determine an overall picture of the sources of variability, Zhang et al. (2023) demonstrate that one should gather for each dimension its main effect and a part of its interactions. For instance, the 2nd order interaction between time and ensemble dimensions (i.e., RT) is divided by two and each half is allocated to one dimension: time or ensemble. As an example, here is the calculation for TIME using this attribution method:

$$\text{TIME} = T + \frac{1}{2}(RT + ST + MT) + \frac{1}{3}(SRT + SMT + RMT) + \frac{1}{4}SMRT \quad (10)$$

Although this method is very useful for obtaining an overview of the distribution, the sum can hide some of the results. An increase of a component can be compensated by, for example, a decline of another. The analysis of each individual component is therefore crucial, and the present work is dedicated to it.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Other methods have also been used to investigate the different sources of variability/uncertainty in climate simulations, such as the one used in Hawkins and Sutton (2009, 2011) or Lehner et al. (2020). While this method has the advantage of incorporating more models as it does not require ensembles, the internal variability is estimated as the residual of a polynomial fit. Here, given the importance of internal variability for our study, we decided to focus on ensemble simulations and use ANOVA. Another difference is the fact that this method does not evaluate the importance of interactions among sources of variance. While this simplifies the analysis, here we will see that interactions play an important role and must be considered to fully understand the evolution of AMOC variability.

Our results do not appear sensitive, especially qualitatively, to the size of the time window (Fig. A1). However, it should be noted that increasing the window size reinforces the main effect of temporal variance and reduces the ensemble variance main effect. As 30 years is the typical period for studying the state of the climate (Arguez and Vose, 2011), we chose this duration to ensure a minimum coherence of the climate state and provide a good representation of temporal variance, while avoiding oversmoothing that would make transitions between regimes much more difficult to detect. Conversely, results concerning model variability are sensitive to the AMOC reference chosen, particularly depending on whether we use AMOC anomaly or absolute values. This sensitivity is due to the persistent AMOC bias in climate models and the difficulty of capturing the right AMOC intensity (Weijer et al., 2020). In our work, we consider the AMOC anomaly relative to the period 1850-1900 because we are interested in the model uncertainty/variability associated with the model response to historical and future emissions rather than the initial and overall AMOC differences between climate models. Results with absolute AMOC values are discussed in section 3.3 and shown in Fig. A2.

2.2.2 Bootstrapping

To obtain robust results and assess the model uncertainty it is important to mix the different models. As the ensemble models have different sizes, we used a Bootstrap methodology for the mixture. Only models of comparable size are mixed. Large ensembles (with 25 to 40 members) are mixed together and smaller ensembles (from 3 to 6 members) are mixed together. The bootstrapping procedure is relatively simple. For each large ensemble model, we select randomly 20 members (with replacement) which are assembled with 20 members of each other model. The ANOVA decomposition is thus applied on the 60 selected members from the three models. The selection of 20 random members and the ANOVA decomposition is then replicated a given number of time. In our case we use 100 resamplings and we did not observe substantial improvements by further increasing this number (e.g. to 1000). The results of the 100 resamplings are then averaged and analyzed. For the smaller members, the procedure is similar but by selecting 3 members per model instead of 20.

A sensitivity test was carried out to assess the impact of sub-sample size for large ensembles (Fig. A3a, c, e and A4a, c). The test indicates that reducing the subsample size from 20 to 3 members leads to a decrease of internal variability and an increase of time variability. Specifically, during the first regime, we observe a transfer from R and RT to T, and a transfer from RT and SRT to ST after the AMOC declines. The size of the ensemble has a direct impact on the level of internal variability. We also assessed the impact of a mixture of large and small ensembles versus the use of small ensembles only. In this case, we applied 3-member resampling for all ensembles. While the model-associated factors of variability are unaffected (Fig. A4),

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

the mixture leads to a decrease in all physical factors. The decrease is relatively homogeneous between components, as the variance distribution remains unchanged (Fig. A3f).

3 Results

Before delving into the various factors and contributions of the variance, we will start by drawing a general picture of the AMOC intensity evolution in the CMIP6 dataset to better understand the changes of variability.

The ensemble-averaged AMOC time series for different models and scenarios appear relatively stable during most of the historical period, from the mid-19$^{\text{th}}$ to the mid-20$^{\text{th}}$ (Fig. 2a and b). Some models then present a slow increase associated with the increasing aerosol concentration (Menary et al., 2013, 2020; Robson et al., 2022). In the last decades of the 20$^{\text{th}}$ century, the AMOC intensity initiates a substantial decrease, with no visible differences among emission pathways up to the middle of the 21$^{\text{st}}$. Afterward, the scenarios start to separate with the strongest forcing presenting a continuous and even intensified decline, while the weakest forcing scenario stabilizes or even slowly recovers depending on the models. At the end of the projection period, while the large ensemble models seem to converge under a given scenario in terms of relative decrease, the small ensembles present growing differences.

This quick overview depicts a strong variability of the AMOC over time, scenario, and models from the last decades of the 20$^{\text{th}}$ century. This behavior is expected to continue and even intensify along the 21$^{\text{st}}$ century.

3.1 A general picture of the variability

To analyze the evolution of the variability and its various contributions, we used the Analysis of Variance (ANOVA) method, described in section 2.2.1. The total variance from the multidimensional dataset is thus split into different contributions separated into main effects and associated interactions. Again, given the number of components, it is helpful to start the analysis simple and progressively increase complexity. For this purpose, we take advantage of the attribution method proposed by Zhang et al. (2023, and summarized in (10)). For each dimension of the dataset, we gather its main effect and redistribute the interaction terms (Fig. 2c-f).

The model variability (representing differences among models) appears to be the dominant factor of variance among most of the studied period. After a low-level of variability up to the beginning of the 20$^{\text{th}}$ century, due to the AMOC intensity reference taken as the 1850-1900 model average, the contribution rapidly increases to dominate both in large and small ensembles. Two consecutive periods of increase detach from this time series. The first one occurring in the second half of the 20$^{\text{th}}$ century, is likely associated with the differences of AMOC response to aerosol forcing among models. This increase is shifted in time between large and small ensembles. While it starts around 1940 in the larger ensembles, it does not appear until 1970 in the smaller ones. The second increase, starting at the very end of the 20$^{\text{th}}$ century, seems due to the important AMOC decline and to the differences of decline magnitude among models. After 2040, the end of the time series remains uncertain, as illustrated by the difference between small and large ensembles (grey dashed line). As discussed in section 2.2.1, this is due to the reference chosen for the AMOC anomaly. With the 1850-1900 reference period taken here, large ensembles eventually converge while

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

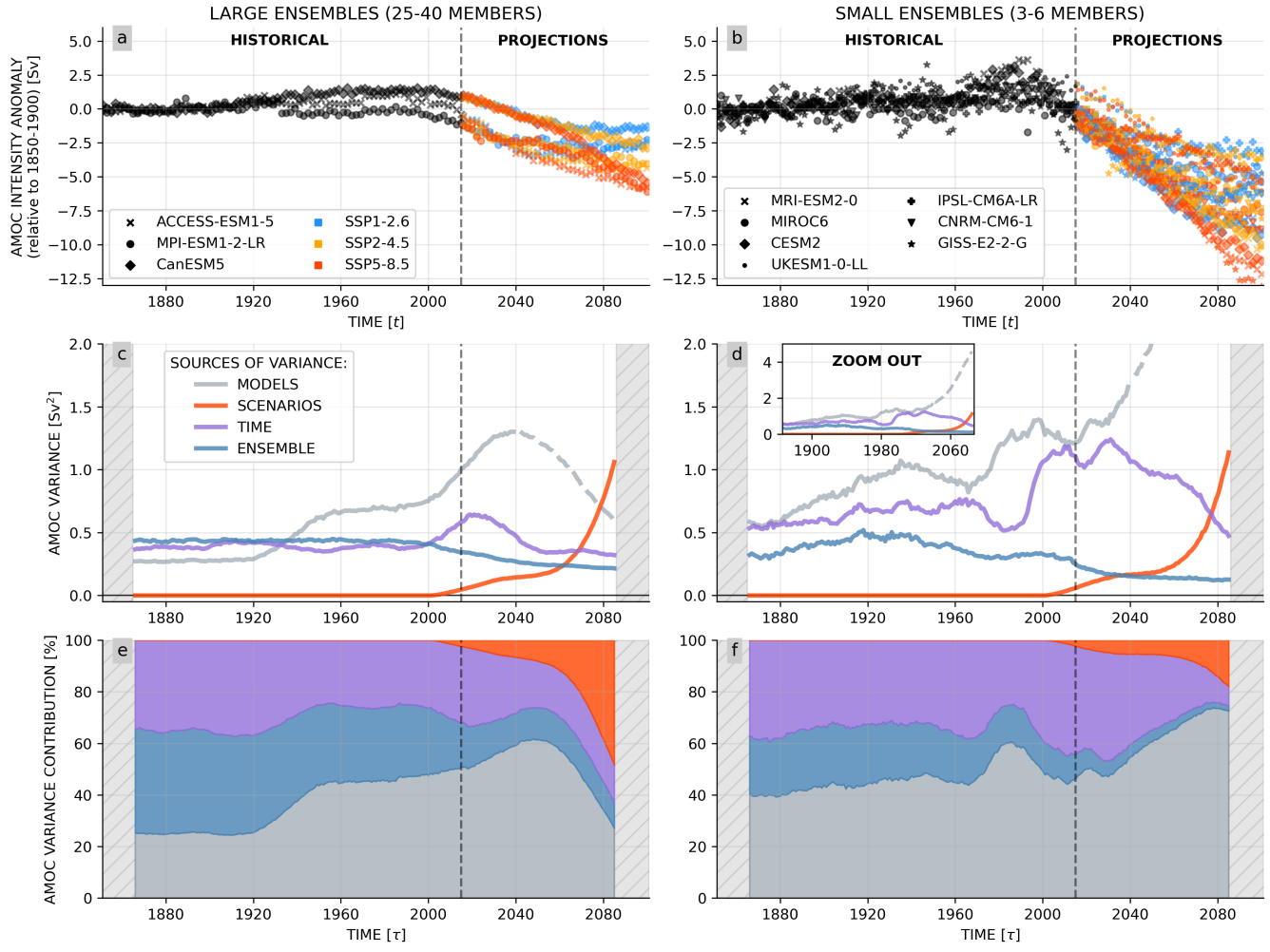

Figure 2. General picture of the AMOC evolution at 26°N from CMIP6 ensemble models. (a, b) AMOC intensity time series over historical (1850-2015, black) and projection (2015-2100, colors) periods. Reference value is the individual model average from 1850-1900. (c, d) AMOC variance associated with each of the four general factors based on the attribution methodology of Zhang et al. (2023). Model contribution has been represented with a dashed line after 2040 to highlight the lack of consistency between c and d regarding this result. Small and large ensembles have the same y-axis, and a zoom-out box is displayed for the variance of the small ensembles to show the full increase in model contribution (d). (e, f) Contribution of each factor to the total variance. Results are presented for large ensembles (a, c, e) with 25-40 members and small ensembles (b, d, f) with 3-6 members.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

small ones do not. Overall, the smaller ensembles show larger model variability. This is probably due both to the greater number of models considered, but also to the poorer separation of internal variability leading to differences between models.

Ensemble and time variability are the two other important factors of variance during the historical period, with comparable orders of magnitude. Nevertheless, they have very different evolutions. Time variability is relatively stable during most of the historical period. Then, at the end of the 20^{th} century, it presents a transient increase for a few decades as a signature of the AMOC decline. It is particularly striking in large ensembles, where the time variance bump is perfectly phased with the AMOC decline and the associated second increase of model variability. In small ensembles, the transient aspect of the increase is slightly less clear, but still visible. The ensemble variability is also very stable at the beginning of the historical period. Then, concomitantly to the AMOC decline, it initiates a slow and progressive decrease up to the end of the 21^{st} century.

Finally, the last component is the scenario variability, mainly associated with the divergence of future emission pathways. By definition, it is zero at the outset, since the scenarios only start in 2015. At the beginning of the 21^{st} century (due to the 30-year running window), it thus starts to increase but stabilizes at a low level until the mid-21^{st} century. After which, it rapidly increases to dominate over both internal and time variability in small ensembles, and even over the model variability in large ensembles.

To move a step further in understanding these changes of variability and the various phenomena involved, we will go beyond this overall attribution and analyze each individual contribution, including the various interactions. As we shall see, interaction terms can be responsible for a large part of the variability and provide crucial information that must be analyzed.

3.2 A focus on the three regimes of variability

In this section, we leave aside the model-associated factors of variance that do not refer to observable variability, and focus on analyzing the "physical factors" of variability. As it shows the robustness of the study, model-associated variability and interactions are discussed later in subsection 3.3. Also, analysis of the model-associated interactions shows that they are mainly governed by physical factors. This makes it more natural to analyze the physical factors first, even if their associated variance is smaller.

Three phases and regimes of variability can be isolated in the simulated time series from 1850 to 2100. Each of them will be analyzed in the following subsections.

Special attention is paid to the analysis of large ensembles as they provide a better cover of the phase-space and thus a more accurate picture of the variability, especially associated with internal factors. However, the small ensembles remain important both for assessing the significance of the results and for the analysis of model-associated variability (since there are more models with a small ensemble).

3.2.1 First century of the historical period controlled by internal variability

Among the different physical factors (Fig. 3), the interaction between time and ensemble (RT) dominates the variability during most of the historical period: from the mid-19^{th} to the last decades of the 20^{th} century. It is particularly striking in large ensembles where it represents around 80% of the physical factor variability (Fig. 3e). In small ensembles, it is lower but still

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

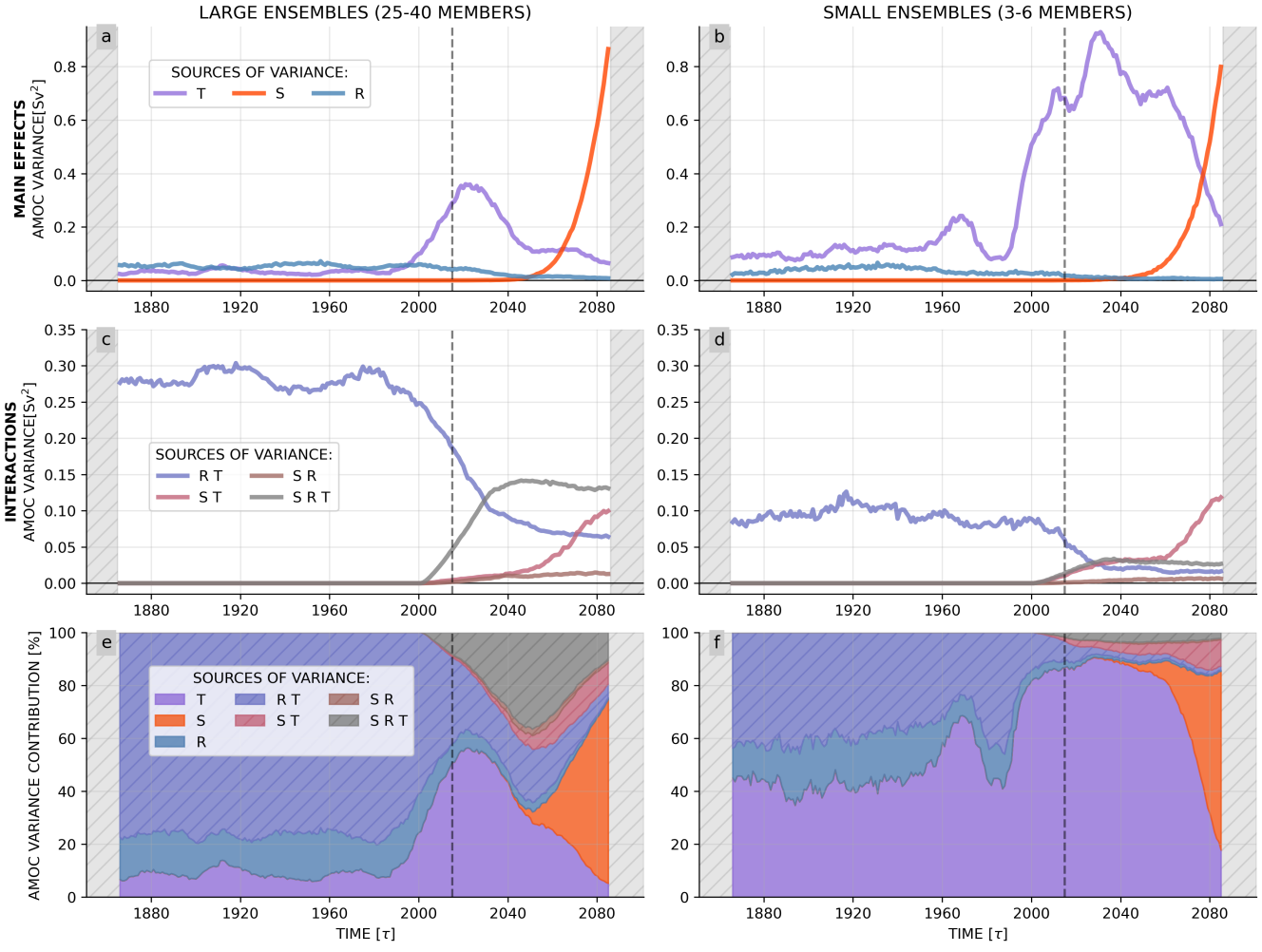

Figure 3. Evolution of the variability of the physical factors including ensemble, time, and scenarios dimensions. (a-d) Variance associated with each main effect (a, b) and their interaction (c, d). (e, f) Relative variance contribution of each main effect and their interactions to the total physical factor variability. Interactions are highlighted by hatches. Results are presented for large ensembles (a, c, e) with 25-40 members and small ensembles (b, d, f) with 3-6 members.

represents approximately 40% of the total physical factor variability (Fig. 3f). The remainder is divided into the main effects of ensemble (R) and time (T).

In a relatively stable context, such as the beginning of the historical period, quasi-ergodicity can be assumed (Hingray and Saïd, 2014). Ergodicity defines a situation where the ensemble and time statistics tend toward the same values. Here, we cannot speak of ergodicity because of the impact of the trend on time variance. However, as the trend is small during this period so is its impact on time variance allowing to use the term quasi-ergodicity. This can be highlighted by computing the ratio of time and ensemble main effects (R/T ratio, Fig. 4a). During this regime, the ratio appears close to one, indicating relatively similar

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

ensemble and time variance. Time and ensemble dimensions can, therefore, be both associated with internal variability. This explains the remarkable importance of RT interaction as the system varies both in time and among ensemble members in the historical period. Thus, the main effects of time and ensemble are smaller than their interaction since the variability strongly decreases when averaging over one of the two dimensions, as illustrated in Fig. 1a and b.

Going into detail, the R/T ratio appears slightly greater than one for large ensembles and slightly lower for smaller ones. This indicates that ensemble dimension dominates in large ensembles while small ensembles are dominated by the time dimension. The study of the number of independent variables can shed light on this result. Large ensembles provide 20 independent variables at each time step, through the bootstrapping of members. For the computation of time variability, we used a 30-year time window to ensure a relatively consistent climate state among the considered periods. By analyzing AMOC time series, we highlight a decorrelation timescale (calculated as the e-folding timescale of the auto-correlation) between 2 and 40 years (Fig. A5). This means that in 30-year windows, only a few data points are truly independent over time. The covering of the phase-space of internal variability is thus much better in the ensemble dimension than in the time dimension, explaining why the first one dominates. The effect of information quantity is strengthened by the difference between large and small ensembles, with RT variability more than twice as low in small ensembles, based on 3-member resampling. This impact of degree-of-freedom is also highlighted by the sensitivity tests where increasing bootstrapping subsample size results in an increase of R and a decrease of T (Fig. A3a and c), and conversely increasing the running time window leads to an increase of T and a decrease of R (Fig. A1).

The decomposition of the total variance budget involving time and ensemble dimensions is particularly insightful to understand the role of these two dimensions (Fig. 5a and b). It corresponds to the sum of the main effects and their interaction, as shown in (4). Following (3), it is also equal to the variance of the initial dataset after averaging models and scenarios (as illustrated in Fig. 1a and b). The decomposition of the variance clearly highlights the dominant role of the interaction especially in large ensembles, and the relatively similar ensemble and time main effects highlight the quasi-ergodicity of the regime (Fig. 5a). Figure 1a supports this result with the spread of the combined dimensions ($\overline{x_\tau}(r,t)$) being much greater than the spread of individual dimensions ($\overline{x_\tau}(r)$ or $\overline{x_\tau}(t)$). In small ensembles, the interaction is much lower consistently with previous results on subsample size. The time main effect is larger than the ensemble main effect showing that quasi-ergodicity is not reached because of the too small ensemble size.

During this period, scenarios variability is zero, by definition, since scenarios only start in 2015.

3.2.2 AMOC decline insensitive to forcing-scenarios up to mid-21st century

After the first regime of relatively stable AMOC intensity and variance components, the AMOC enters a phase of intense decline reflected in the change of the variance attribution (Fig. 2). This declining trend is directly responsible for the substantial and transient increase of T main effect that reaches almost 60% of the total physical variability at its peak during the first half of the 21st century for large ensembles and almost 90% for small ensembles (Fig. 3e and f). This regime is thus associated with a clear loss of ergodicity (Fig. 4a), where the ensemble over time main effects ratio decrease by a factor $\sim$30 or $\sim$50 for large and small ensembles, respectively.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

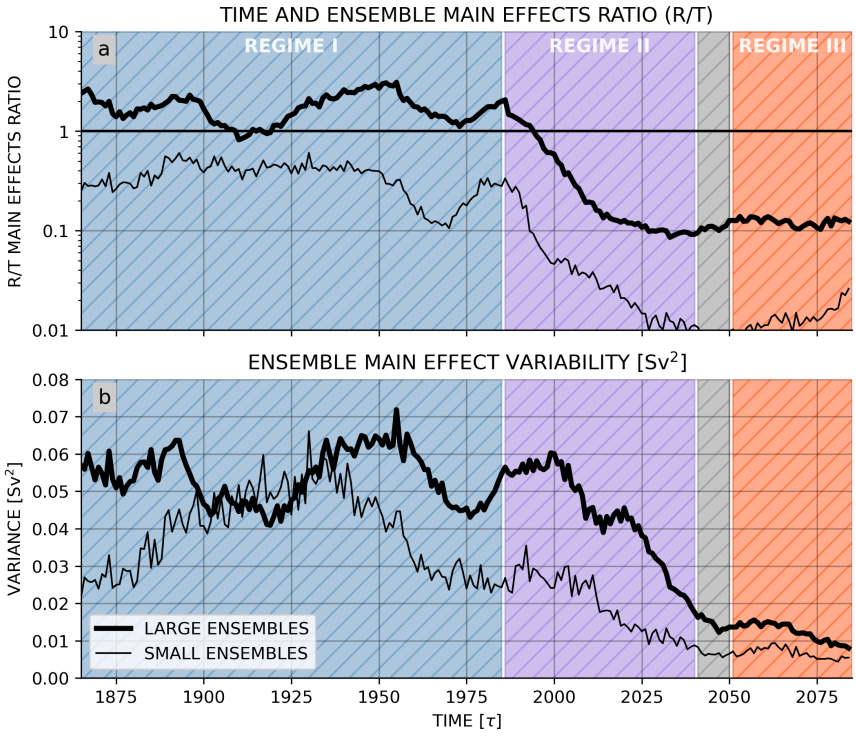

Figure 4. Time and ensemble main effects. (a) Comparison of their variability with the ratio of ensemble main effect over time main effect. We use this ratio as a proxy of ergodicity. (b) Zoom on the evolution of ensemble main effect variability over time. Thick lines represent large ensembles and thin lines small ensembles.

The signature of this loss of ergodicity can be found in the RT interaction that reflects the intricacy of time and ensemble dimensions under quasi-ergodicity. RT, that was previously the dominant physical factor, presents a substantial decrease together with the AMOC decline around 1990 in large ensembles and slightly later in smaller ones (Fig. 3c and d).

Around 2000, the increase of time main effect, leading it to exceed the ensemble main effect in large ensembles (Fig. 5a), is also detectable in Fig. 1a and b where the spread of $\overline{x_\tau}(t)$ becomes much larger than the spread of $\overline{x_\tau}(r)$. In addition to the increase of T this change of dominant factor is also due to the decline of R that follows the same decreasing pathway as RT. During this regime R losses around 3/4 of its variance in both ensembles (Fig. 3 and 4b). The reason for such decline of ensemble main effect is not straightforward. We thus further investigate the evolution of ensemble variability over time and scenarios by computing for each individual model and scenario the evolution of ensemble variance over time (Fig. 6). This analysis shows an overall decrease of the ensemble variance over time, illustrating a contraction of the phase-space. To assess this visual decline, we compute the average ensemble variance over 50-year time window in the historical and projection periods. In the historical period, we select the 1900-1950 window, located a few decades after the beginning of the historical simulations (allowing for possible AMOC adjustment from the control preindustrial forcing to the historical one) and before the

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

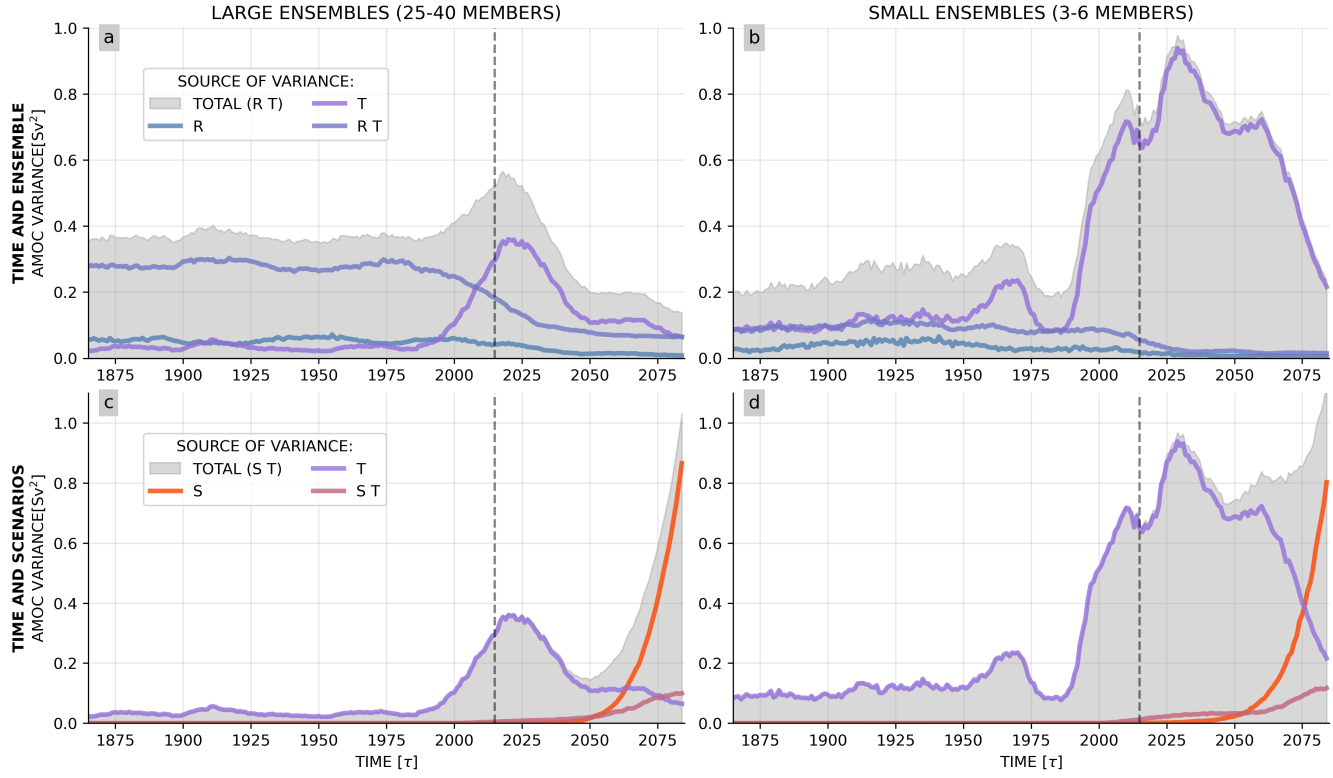

Figure 5. Evolution of variance associated with time and ensemble (a, b) and time and scenarios (c, d). Grey shaded areas represent the total variance associated with the considered dimensions. It corresponds to the sum of the colored lines representing the different variance contributions including the two main effects and their interaction. Results are presented for large ensembles (a and c) with 25-40 members and small ensembles (b and d) with 3-6 members.

increase of aerosol forcing. For the projection, we select the last 50 years to target the potential maximum impact of scenarios. In all models, the 1900-1950 period presents a larger ensemble variance than the 2050-2100 one, with an intensifying decrease over time. If we compare now the different scenarios, the majority of models (7/10), including the three large ensembles, presents weaker ensemble variance under SSP5-8.5 than under SSP1-2.6; 6/10 models present weaker variability under SSP5-8.5 than SSP2-4.5; and 8/10 a weaker variability under SSP2-4.5 than under SSP1-2.6. For the three large ensembles, the ensemble variability is sorted according to scenario intensity, except for CanESM5 where SSP2-4.5 shows a slightly weaker variance than SSP5-8.5. This results therefore suggests a link between the anthropogenic forcing intensity and the AMOC ensemble variance decrease.

From 2015, scenarios start and the associated variability emerges (Fig. 2). This emergence is not generated by the scenario main effect (S) but by the interaction between time, scenarios, and ensemble (SRT, Fig. 3). This increase of scenario-associated variability before stabilizing, is therefore not the direct effect of the forcing. It is, rather, a sort of chaotic internal variability triggered by slight differences among scenarios allowing the simulations to spread. Indeed, during the historical

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

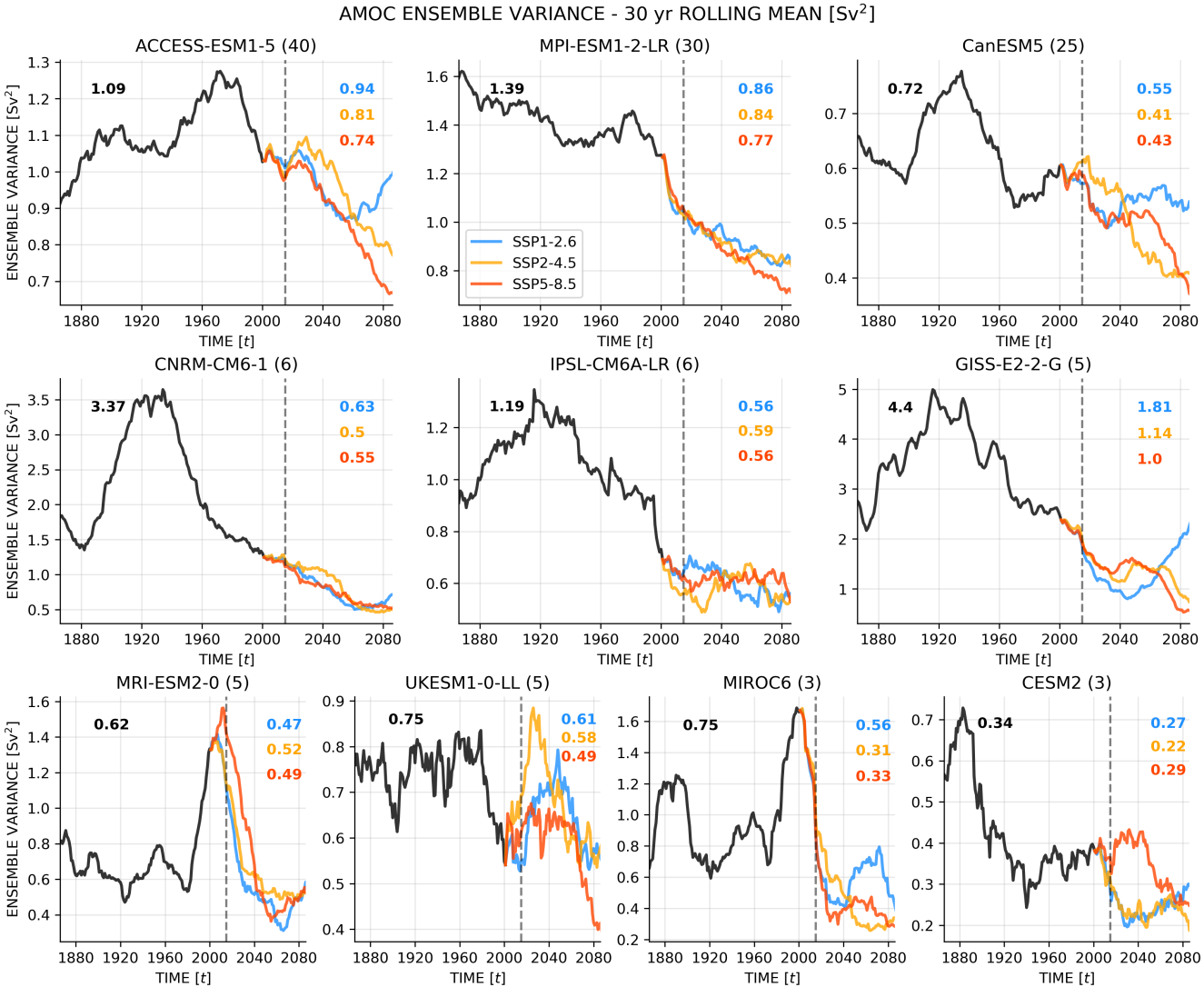

Figure 6. Evolution of internal variability, measured by the ensemble variance, over time and scenarios. Each subplot presents historical time series (black) followed by the three scenarios SSP1-2.6 (blue), SSP2-4.5 (orange), or SSP5-8.5 (red) for a given model. The values written in each subplots correspond to 50-year averaged period of internal variability: 1900-1950 for historical (black) and 2050-2100 for scenarios (colors). A 30-year rolling averaged is applied to isolate long term trends.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

period, the three time series corresponding to scenarios for each member of models are perfect replicas of the historical simulation. However, when scenarios really start in 2015, slight differences or perturbations in forcing appear among them and allow the three time series to spread as pseudo ensemble members. This internal and chaotic nature of SRT variability is underlined by the magnitude of this variability, which appears to be three times greater in large ensembles as in small ones. The absence of direct scenario main effect variability is underline by the spread of $\overline{x_\tau}(s)$ in Fig. 1e and f, that does not increase before 2040.

Finally, the last component is the interaction between time and scenarios (ST). This one weakly increases when scenarios appear but remains at a very low level up to mid-21st century. The analysis of terms included in the total variance involving time and scenario dimensions shows that time is clearly the leading factor of the associated variability during this phase, with S and ST negligible compared to T (Fig. 5c and d). This result is strengthened by the strong dominance of the spread of $\overline{x_\tau}(t)$ in Fig. 1c and d. During this regime, while the scenarios have started, the AMOC variability remain driven by the declining trend associated with a dynamical-adjustment without direct impact of forcing scenarios.

Beyond the decline associated with the transient increase of time variability, the T main effect decreases and stabilizes with SRT around 2050. RT interaction decrease rate also flattens, while ST only exhibits a very weak increase. Before the start of the third regime, we therefore observe an intermediate phase where all components of variability appear stable. This short phase around 2040-2050 is dominated by the SRT interactions associated with a "scenario-perturbed" internal variability. This is therefore in a way another phase were internal variability is important despite the significant trend. While most of the components in small ensembles present a clear signature of this intermediate regime, it is less clear for T main effect. We suggest this is due to the less accurate diagnostic of internal variability in small ensembles, hence impacting the separation of internal and dynamical-adjustment variability.

3.2.3 Late separation of anthropogenic emissions scenarios after the mid-21st century

The third regime starts with the emergence of the scenario main effect S. It presents a steep increase from the middle of the 21st century and becomes, in a few decades, the dominant factor of physical variability. In parallel, the ST interaction presents a weaker but still substantial increase, since the scenarios lead to an AMOC dynamical adjustment. The variance involving time and scenario dimensions (Fig. 5), shows that after being mostly driven by T, S takes the lead around 2060 and 2075 in large and small ensembles, respectively. This is also shown by the strong increase of the spread of $\overline{x_\tau}(s)$ in Fig. 1c-f. The scenario variability becomes the first order factor of variability even above model variability in large ensembles after 2080. The increasing and prominent role of scenarios/forced variability demonstrates that the system leaves a regime driven by past, historical forcings to enter in a phase of forced evolution driven at first order by the anthropogenic emissions scenarios.

After focusing on the physical factors of variability, we can investigate model-associated variability and the drivers of model differences along these multisecular simulations. This will measure the confidence and robustness of our physical factor results.

3.3 Evolution of inter-model variability and uncertainty

Overall, the dominant factor in model-associated variability lies in the model main effect (M, Fig. 7). This follows the same two-step increase, described at the beginning with the attribution method (Fig. 2). As a reminder the two steps are associated

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

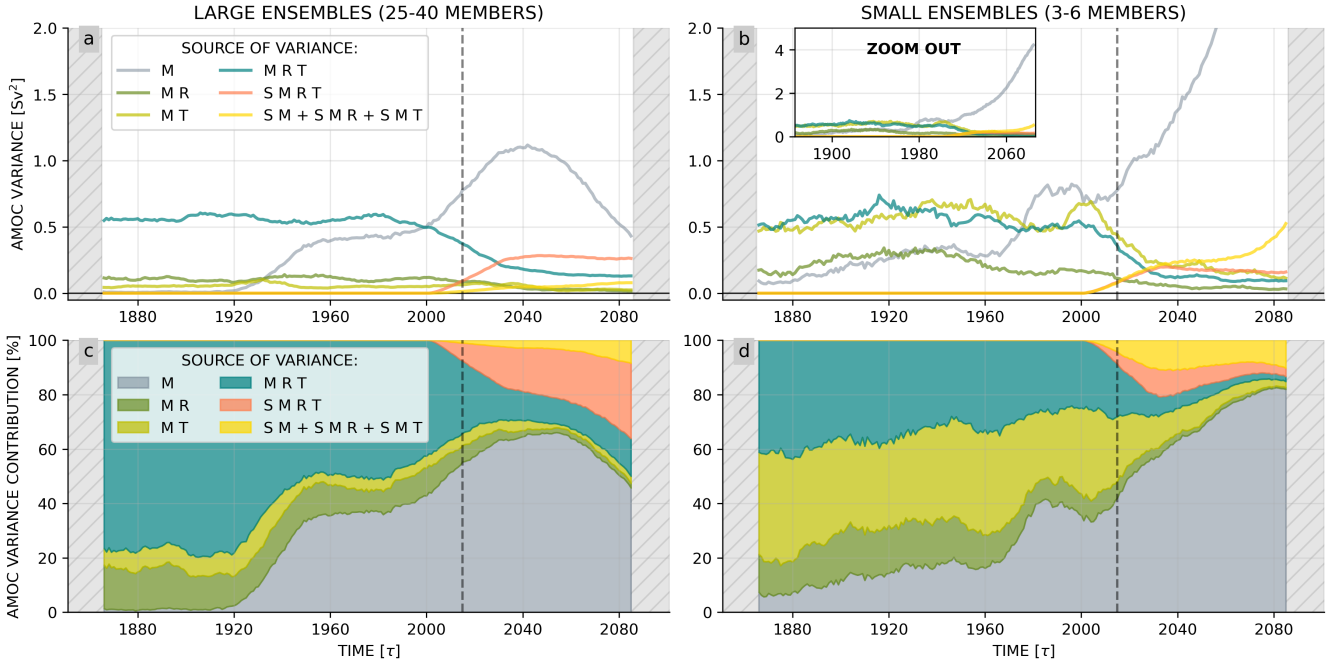

Figure 7. Evolution of the model-associated factors of variability. (a, b) Variance associated to model main effect and interactions. Small and large ensembles have the same y-axis, and a zoom-out box is displayed for the variance of the small ensembles to show the full increase in model contribution (b). (c, d) Variance contribution of model main effect and interactions to the total model-associated variability. Results are presented for large ensembles (a, c) with 25-40 members and small ensembles (b, d) with 3-6 members.

with two uncertainties of AMOC response predictions among models: (i) the response to the increase of aerosol concentration, and (ii) the magnitude of the AMOC decline. The low level of model variability at the beginning of the simulations is artificially caused by the AMOC intensity reference taken as the individual model average over the 1850-1900 time period. It is difficult to analyze the M main effect at the end of the simulations considering the differences between small and large ensembles. Indeed, while large ensembles present a decreasing model variability due to the convergence of relative AMOC intensity decline, small ensembles present an exponential increase of model variability, linked to a divergence in terms of relative AMOC intensity decline. However, this is sensitive to the reference chosen, as discussed in the method section. Indeed, if absolute AMOC intensity is considered, after 2040, we would observe a convergence of small ensembles and a divergence of large ones (Fig. A2).

The second main feature of the inter-model variability is the fact that it closely follows the evolution of physical components. Most of the interactions involving models follow the same path as the corresponding interaction without the model dimension (Fig. 3 and 7). For instance, the interaction between models, time, and ensemble (MRT) evolves like the interaction between time and ensemble (RT). This is the same for MR and R or $SMRT$ and SRT. This seems to indicate that when physical factors generate a substantial variability, this translates in differences and variability among models. MT appears to be an

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

exception to these results, as the transient increase in T associated with the decline in AMOC is not clearly visible. We suspect that this absent signal is actually hidden in the second increase of M starting from 2000 onwards.

4 Summary and discussion

This 4-way analysis of variance of the AMOC at 26°N in CMIP6 ensembles depicts three phases (Fig. 4) associated with three variance regimes, each dominated by a particular factor.

1. The first phase from 1850 to around 1990 presents a relatively stable AMOC, whose variability is driven by internal variability, associated with time and ensemble dimension given the quasi-ergodicity of this regime;
2. The second phase from 1990 to 2050 is characterized by a decline of the AMOC associated with an adjustment to past historical emissions. It is thus a regime where time changes dominate, the forcing-scenario effect remaining weak;
3. Finally the third phase starting around 2050 is a regime forced by emission scenarios, where the impact of anthropogenic emissions pathways takes control on the divergence of the simulated trajectories.

An important finding of this work is the existence of two increasing phases of scenario variability. The first one, after 2015, rapidly stabilizes in a few decades. The second one, emerging after 2050, presents an exponential increase of the variability. During the first phase, scenarios simply act as little perturbations creating, in a way, “larger ensembles”. This is particularly evident as the main signal associated with scenarios in this period lies in the interaction between scenarios, ensemble, and time dimensions (SRT). This also leads us to detect an intermediate regime around 2040-2050, where most of the variability components stabilize and where the system is driven by this "scenario-perturbed" internal variability. In the second phase, the increase of variability is purely forced by the divergence of scenarios and scenarios main effect takes the lead.

The AMOC intensity decline associated with the transient increase of time variability, because of the absence of clear scenario variability before 2050, suggests that this evolution is not forced by scenarios nor synchronous to anthropogenic emissions. Here the AMOC behavior differs from CO_2 concentration, anthropogenic radiative forcing, and temperature change that already start diverging one or two decades earlier (O’Neill et al., 2016). In our analysis, this AMOC decline is therefore a response to previous historical forcings, rather than a scenario-driven response.

This study also provides robust evidence that the AMOC internal variability in CMIP6 decreases in the projection period compared to historical reference. We also detected a potential link between this decline and emission pathways intensity. This is consistent with MacMartin et al. (2016) results with a single Earth System Model or Cheng et al. (2016) on interdecadal variability of the previous CMIP exercise. This also aligns with previous results on large-scale climate showing that internal variability is highly sensitive to forced variability and mean state (Coquereau et al., 2024). The decline of internal variability observed in the present work corresponds to a contraction of the phase-space and therefore an increasing predictability of the AMOC in the future from one time step to another. However, it does not substantially increase our ability to predict the future of the AMOC, since internal variability is a small component of the AMOC uncertainty, smaller than model or

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

scenario-uncertainty after mid-century. Yet, internal variability seems to decrease concomitantly to the AMOC magnitude, thus it could be a potential proxy or warning signal of an AMOC decline hardly detectable because of the large AMOC interannual variability (Lobelle et al., 2020).

Finally, the level of model variability demonstrates that the magnitude of AMOC decline remains relatively uncertain. This is consistent with the important model biases on AMOC intensity in CMIP6 (Weijer et al., 2020). The disagreement between small and large ensembles at the end of the 21st century raises interesting questions about ensemble strategies and (i) the use of a small number of large ensembles, versus (ii) the use of a large number of small ensembles. There is thus a choice to be made between relying on few models (i) or not having a good separation of internal variability (ii). Here, all results are displayed separately between small and large ensembles to give readers all the information. The fact that interactions involving model dimension follow the same evolution as the corresponding interactions without this dimension is interesting from an uncertainty perspective. Indeed, it shows that when the physical aspects vary greatly, it is more difficult for the models to accurately find the correct state. A good illustration of this is the evolution of internal variability, which represents the range of possible states the system can occupy under a given forcing. As said previously, we observed a decline of ensemble variance when the AMOC intensity decreases. If the space of possible states contracts, one expects the absolute difference between models concerning this phase-space location to contract consequently.

Code and data availability. The program code (in Python) for computing the 4-way ANOVA is available at https://github.com/coquereau/ANOVA_4way. CMIP6 ensemble model outputs can be downloaded from the various Earth System Grid Federation (ESGF) nodes. The article references for each model can be found in Tab. 1.

Author contributions. Conceptualization: A.C., F.S., and Q.J. Methodology: A.C., F.S., T.H., J.J.H, and Q.J. Investigation: A.C., F.S., and Q.J. Visualization: A.C. Supervision: F.S., T.H., and J.J.H. Writing—original draft: A.C., F.S., and Q.J. Writing—review and editing: A.C., F.S., T.H., J.J.H, and Q.J.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the ARVOR project funded by the LEFE IMAGO program, by the OceaniX project funded by the French ANR program, and by the ISblue project (Interdisciplinary graduate school for the blue planet, ANR-17-EURE-0015) funded by a grant from the French government under the program “Investissements d’Avenir”. This work was also supported by the EERIE project (Grant Agreement No 101081383) funded by the European Union. Views and opinions expressed are however those of the authors 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.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

References

Arguez, A. and Vose, R. S.: The Definition of the Standard WMO Climate Normal: The Key to Deriving Alternative Climate Normals, Bulletin of the American Meteorological Society, 92, 699–704, 2011.

Armstrong McKay, D. I., Staal, A., Abrams, J. F., Winkelmann, R., Sakschewski, B., Loriani, S., Fetzer, I., Cornell, S. E., Rockström, J., and Lenton, T. M.: Exceeding 1.5°C global warming could trigger multiple climate tipping points, Science, 377, eabn7950, https://doi.org/10.1126/science.abn7950, 2022.

Berrington de González, A. and Cox, D. R.: Interpretation of interaction: A review, The Annals of Applied Statistics, 1, 371–385, https://doi.org/10.1214/07-AOAS124, 2007.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al.: Presentation and evaluation of the IPSL-CM6A-LR climate model, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002010, 2020.

Böhm, E., Lippold, J., Gutjahr, M., et al.: Strong and deep Atlantic meridional overturning circulation during the last glacial cycle, Nature, 517, 73–76, https://doi.org/10.1038/nature14059, 2015.

Cheng, J., Liu, Z., Zhang, S., Liu, W., Dong, L., Liu, P., and Li, H.: Reduced interdecadal variability of Atlantic Meridional Overturning Circulation under global warming, Proceedings of the National Academy of Sciences, 113, 3175–3178, https://doi.org/10.1073/pnas.1519827113, 2016.

Coquereau, A., Sévellec, F., Huck, T., Hirschi, J. J., and Hochet, A.: Anthropogenic Changes in Interannual-to-Decadal Climate Variability in CMIP6 Multiensemble Simulations, Journal of Climate, 37, 3723–3739, https://doi.org/10.1175/JCLI-D-23-0606.1, 2024.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001916, 2020.

Dansgaard, W., Johnsen, S. J., clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., Hvldberg, C. S., Steffensen, J. P., Sveinbjörnsdóttir, A. E., Jouzel, J., and Bond, G.: Evidence for general instability of past climate from a 250-kyr ice-core record, Nature, 364, 218–220, 1993.

Ganachaud, A. and Wunsch, C.: Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data, Nature, 408, 453–457, https://doi.org/10.1038/35044048, 2000.

Gierz, P., Lohmann, G., and Wei, W.: Response of Atlantic overturning to future warming in a coupled atmosphere-ocean-ice sheet model, Geophysical Research Letters, 42, 6811–6818, https://doi.org/10.1002/2015GL065276, 2015.

Hawkins, E. and Sutton, R.: The Potential to Narrow Uncertainty in Regional Climate Predictions, Bulletin of the American Meteorological Society, 90, 1095–1108, 2009.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation change, Clim Dyn, 37, 407–418, https://doi.org/10.1007/s00382-010-0810-6, 2011.

Henry, L. G. et al.: North Atlantic ocean circulation and abrupt climate change during the last glaciation, Science, 353, 470–474, https://doi.org/10.1126/science.aaf5529, 2016.

Hingray, B. and Saïd, M.: Partitioning Internal Variability and Model Uncertainty Components in a Multimember Multimodel Ensemble of Climate Projections, Journal of Climate, 27, 6779–6798, https://doi.org/10.1175/JCLI-D-13-00629.1, 2014.

Hingray, B., Mezghani, A., and Buishand, T.: Development of probability distributions for regional climate change from uncertain global mean warming and an uncertain scaling relationship, Hydrology and Earth System Sciences, 11, 1097–1114, 2007.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Hirschi, J. J.-M., Barnier, B., Böning, C., Biastoch, A., Blaker, A. T., Coward, A., et al.: The Atlantic meridional overturning circulation in high-resolution models, Journal of Geophysical Research: Oceans, 125, e2019JC015 522, https://doi.org/10.1029/2019JC015522, 2020.

Johns, W. E., Elipot, S., Smeed, D. A., Moat, B., King, B., Volkov, D. L., and Smith, R. H.: Towards two decades of Atlantic Ocean mass and heat transports at 26.5° N, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 381, 20220 188, https://doi.org/10.1098/rsta.2022.0188, 2023.

Knight, J. R., Folland, C. K., and Scaife, A. A.: Climate impacts of the Atlantic Multidecadal Oscillation, Geophysical Research Letters, 33, L17 706, https://doi.org/10.1029/2006GL026242, 2006.

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

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, Proceedings of the National Academy of Sciences, 105, 1786–1793, https://doi.org/10.1073/pnas.0705414105, 2008.

Lobelle, D., Beaulieu, C., Livina, V., Sévellec, F., and Frajka-Williams, E.: Detectability of an AMOC decline in current and projected climate changes, Geophys. Res. Lett., 47, e2020GL089 974, 2020.

MacMartin, D. G., Zanna, L., and Tziperman, E.: Suppression of Atlantic Meridional Overturning Circulation Variability at Increased CO_2, Journal of Climate, 29, 4155–4164, https://doi.org/10.1175/JCLI-D-15-0533.1, 2016.

McManus, J., Francois, R., Gherardi, J. M., et al.: Collapse and rapid resumption of Atlantic meridional circulation linked to deglacial climate changes, Nature, 428, 834–837, https://doi.org/10.1038/nature02494, 2004.

Menary, M. B., Roberts, C. D., Palmer, M. D., Halloran, P. R., Jackson, L., Wood, R. A., Müller, W. A., Matei, D., and Lee, S.-K.: Mechanisms of aerosol-forced AMOC variability in a state of the art climate model, Journal of Geophysical Research: Oceans, 118, 2087–2096, https://doi.org/10.1002/jgrc.20178, 2013.

Menary, M. B., Robson, J., Allan, R. P., Booth, B. B. B., Cassou, C., and Gastineau, G.: Aerosol-forced AMOC changes in CMIP6 historical simulations, Geophysical Research Letters, 47, e2020GL088 166, https://doi.org/10.1029/2020GL088166, 2020.

Olonscheck, D., Suarez-Gutierrez, L., Milinski, S., Beobide-Arsuaga, G., Baehr, J., Fröb, F., et al.: The New Max Planck Institute Grand Ensemble with CMIP6 forcing and high-frequency model output, Journal of Advances in Modeling Earth Systems, 15, e2023MS003 790, https://doi.org/10.1029/2023MS003790, 2023.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

Rind, D., Orbe, C., Jonas, J., Nazarenko, L., Zhou, T., Kelley, M., et al.: GISS Model E2.2: A climate model optimized for the middle atmosphere—Model structure, climatology, variability, and climate sensitivity, Journal of Geophysical Research: Atmospheres, 125, https://doi.org/10.1029/2019JD032204, 2020.

Robson, J., Menary, M. B., Sutton, R. T., Mecking, J., Gregory, J. M., Jones, C., Sinha, B., Stevens, D. P., and Wilcox, L. J.: The Role of Anthropogenic Aerosol Forcing in the 1850–1985 Strengthening of the AMOC in CMIP6 Historical Simulations, Journal of Climate, 35, 6843–6863, https://doi.org/10.1175/JCLI-D-22-0124.1, 2022.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., et al.: UKESM1: Description and evaluation of the U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Srokosz, M., Baringer, M., Bryden, H., Cunningham, S., Delworth, T., Lozier, S., Marotzke, J., and Sutton, R.: Past, Present, and Future Changes in the Atlantic Meridional Overturning Circulation, Bulletin of the American Meteorological Society, 93, 1663–1676, https://doi.org/10.1175/BAMS-D-11-00151.1, 2012.

Stommel, H.: Thermohaline Convection with Two Stable Regimes of Flow, Tellus, 13, 224–230, https://doi.org/10.1111/j.2153-3490.1961.tb00079.x, 1961.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geoscientific Model Development, 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.

Sévellec, F. and Fedorov, A. V.: Millennial Variability in an Idealized Ocean Model: Predicting the AMOC Regime Shifts, Journal of Climate, 27, 3551–3564, https://doi.org/10.1175/JCLI-D-13-00450.1, 2014.

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., Sudo, K., Sekiguchi, M., Abe, M., Saito, F., Chikira, M., Watanabe, S., Mori, M., Hirota, N., Kawatani, Y., Mochizuki, T., Yoshimura, K., Takata, K., O'ishi, R., Yamazaki, D., Suzuki, T., Kurogi, M., Kataoka, T., Watanabe, M., and Kimoto, M.: Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6, Geoscientific Model Development, 12, 2727–2765, https://doi.org/10.5194/gmd-12-2727-2019, 2019.

Trenberth, K. E. and Fasullo, J. T.: Atlantic meridional heat transports computed from balancing Earth's energy locally, Geophysical Research Letters, 44, 1919–1927, https://doi.org/10.1002/2016GL072475, 2017.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., et al.: Evaluation of CMIP6 DECK experiments with CNRM-CM6-1, Journal of Advances in Modeling Earth Systems, 11, 2177–2213, https://doi.org/10.1029/2019MS001683, 2019.

Wang, X. L. and Zwiers, F. W.: Interannual Variability of Precipitation in an Ensemble of AMIP Climate Simulations Conducted with the CCC GCM2, Journal of Climate, 12, 1322–1335, 1999.

Weijer, W., Cheng, W., Garuba, O. A., Hu, A., and Nadiga, B. T.: CMIP6 Models Predict Significant 21st Century Decline of the Atlantic Meridional Overturning Circulation, Geophysical Research Letters, 47, e2019GL086 075, https://doi.org/10.1029/2019GL086075, 2020.

Yates, F.: Orthogonal Functions and Tests of Significance in the Analysis of Variance, Supplement to the Journal of the Royal Statistical Society, 5, 177–180, https://doi.org/10.2307/2983655, 1938.

Yip, S., Ferro, C. A. T., Stephenson, D. B., and Hawkins, E.: A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions, Journal of Climate, 24, 4634–4643, https://doi.org/10.1175/2011JCLI4085.1, 2011.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M.: The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component, Journal of the Meteorological Society of Japan. Ser. II, 97, 931–965, https://doi.org/10.2151/jmsj.2019-051, 2019.

Zhang, S., Zhou, Z., Peng, P., and Xu, C.: A New Framework for Estimating and Decomposing the Uncertainty of Climate Projections, Journal of Climate, 37, 365–384, https://doi.org/10.1175/JCLI-D-23-0064.1, 2023.

Zickfeld, K., Eby, M., and Weaver, A. J.: Carbon-cycle feedbacks of changes in the Atlantic meridional overturning circulation under future atmospheric CO2, Global Biogeochemical Cycles, 22, GB3024, https://doi.org/10.1029/2007GB003118, 2008.

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., and Srbinovsky, J.: The Australian Earth System Model: ACCESS-ESM1.5, Journal of Southern Hemisphere Earth Systems Science, 70, 193–214, https://doi.org/10.1071/ES19035, 2020.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Zwiers, F.: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2, Climate Dynamics, 12, 825–847, https://doi.org/10.1007/s003820050146, 1996.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

Appendix A: Additional figures

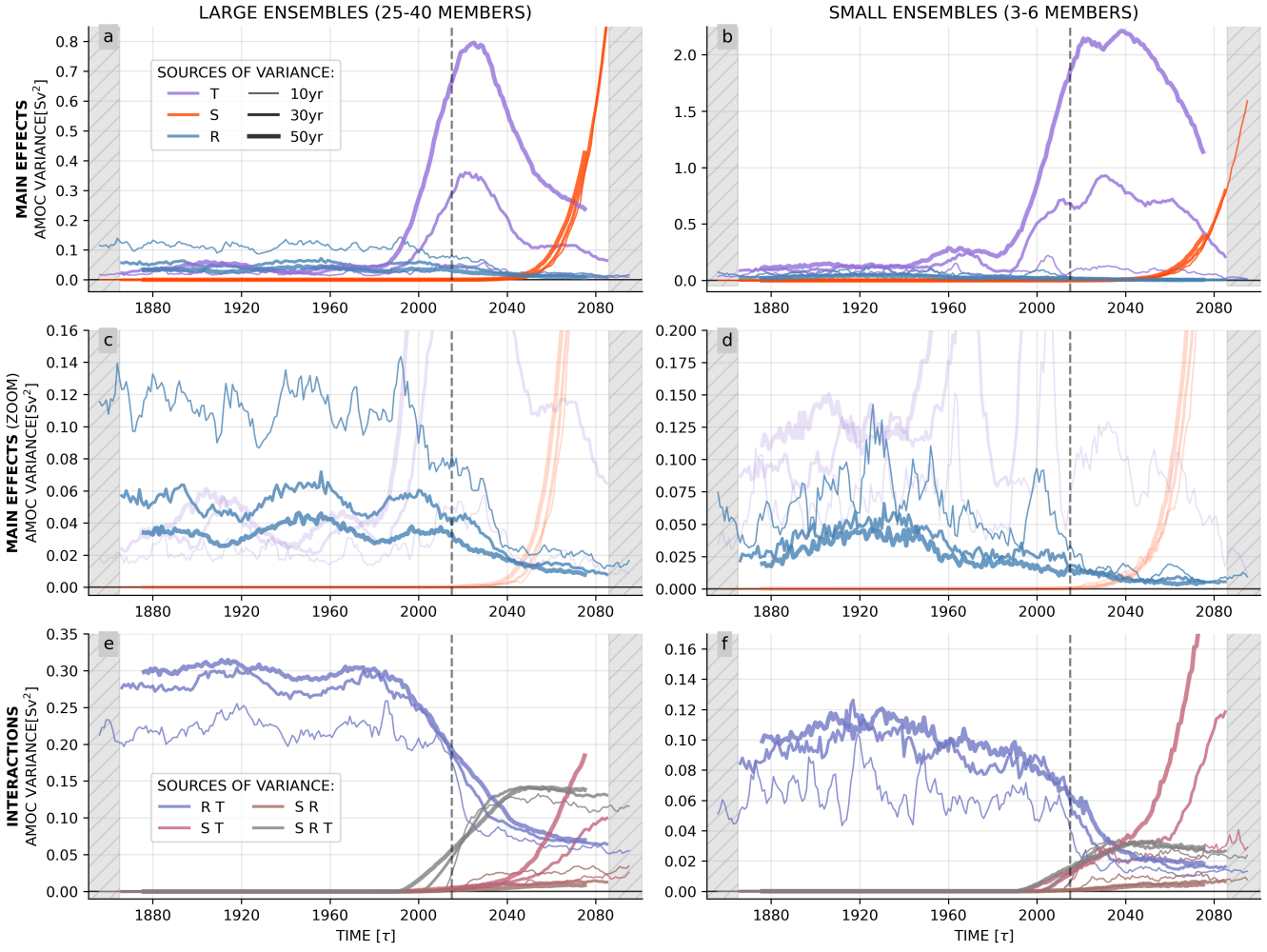

Figure A1. Sensitivity test on time window size. Variance associated to each main effect (a-d) and interaction (e,f). (c,d) Zoom on ensemble main effect. Line thickness represents the size of the time window from 10 (thin) to 50 years (thick). Results are presented for large ensembles (a, c, e) with 25-40 members and small ensembles (b, d, f) with 3-6 members.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

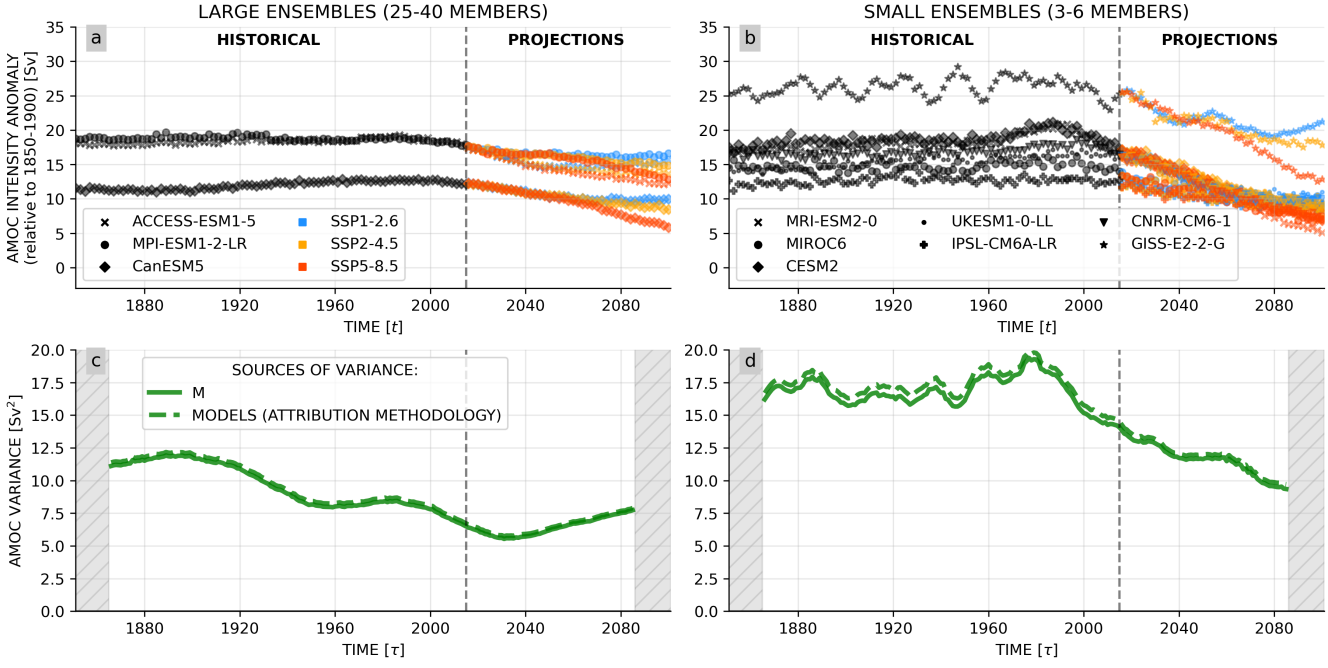

Figure A2. Sensitivity test with AMOC absolute intensity. (a, b) AMOC absolute intensity time series over historical (1850-2015, black) and projection (2015-2100, colors) periods. (c, d) Variance associated with model main effect only (solid line) and with all model-associated factors combined with the attribution methodology by Zhang et al. (2023) (dashed line). Values correspond to absolute AMOC intensity. Results are presented for large ensembles (a, c) with 25-40 members and small ensembles (b, d) with 3-6 members.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

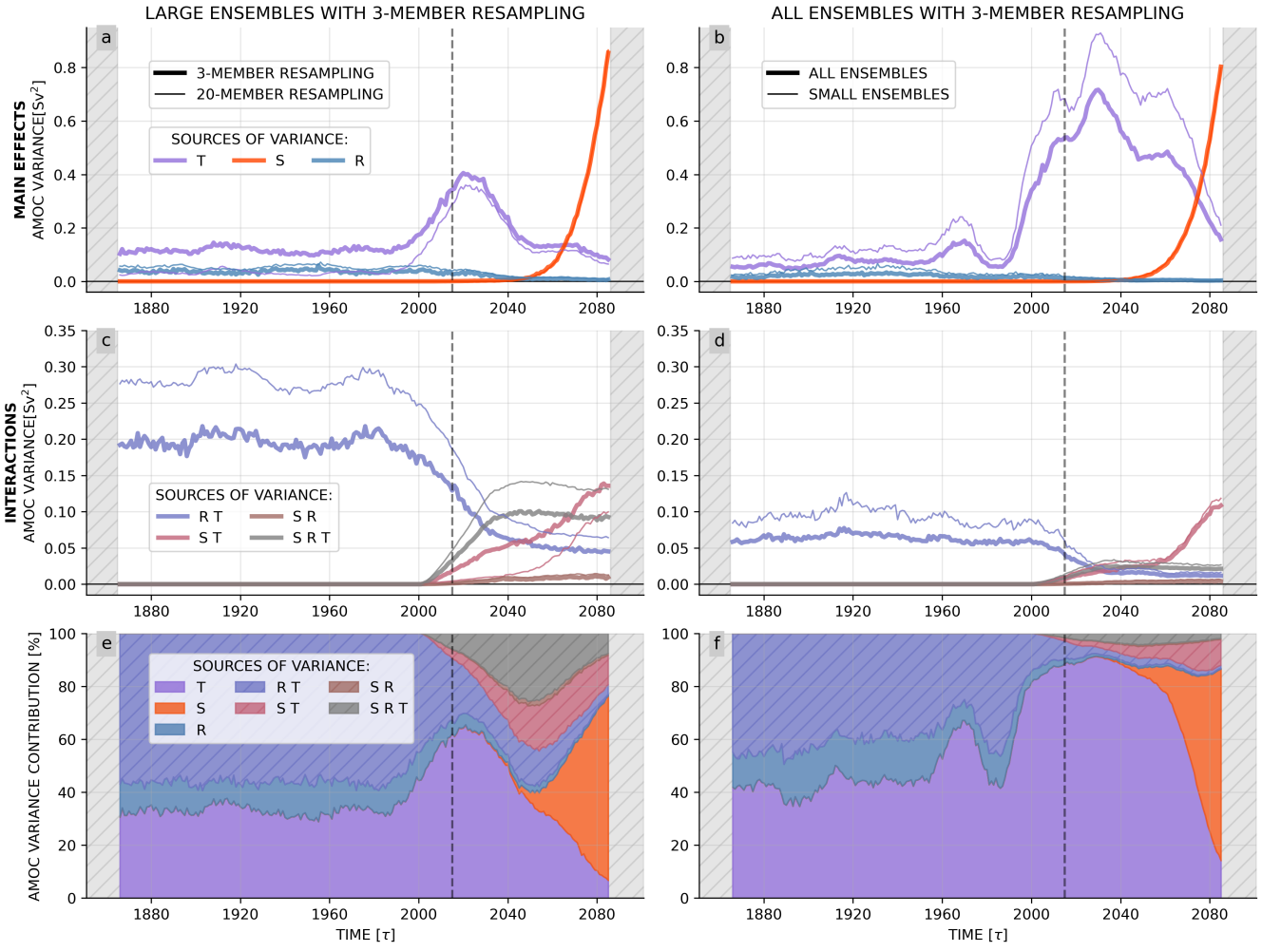

Figure A3. Sensitivity test on bootstrapping methodology: application to the variability of physical factors. (a-d) Variance associated to each main effect (a, b) and their interaction (c, d). (e, f) Relative variance contribution of each main effect and their interactions to the total physical factor variability. Left-hand panels show the sensitivity of the large ensemble results to the number of members in each resampling. Thin lines present the control experiment with 20 members of each model (identical to Fig. 3 a, c, e) and thick lines present the results obtained using 3 members of each model per resampling. Right-hand panels show the sensitivity to mixing large and small ensembles. Thin lines present the control experiment with only small ensembles (identical to Fig. 3 b, d, f) and thick lines present the results obtained with all ensembles. 3-member resampling are used here.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

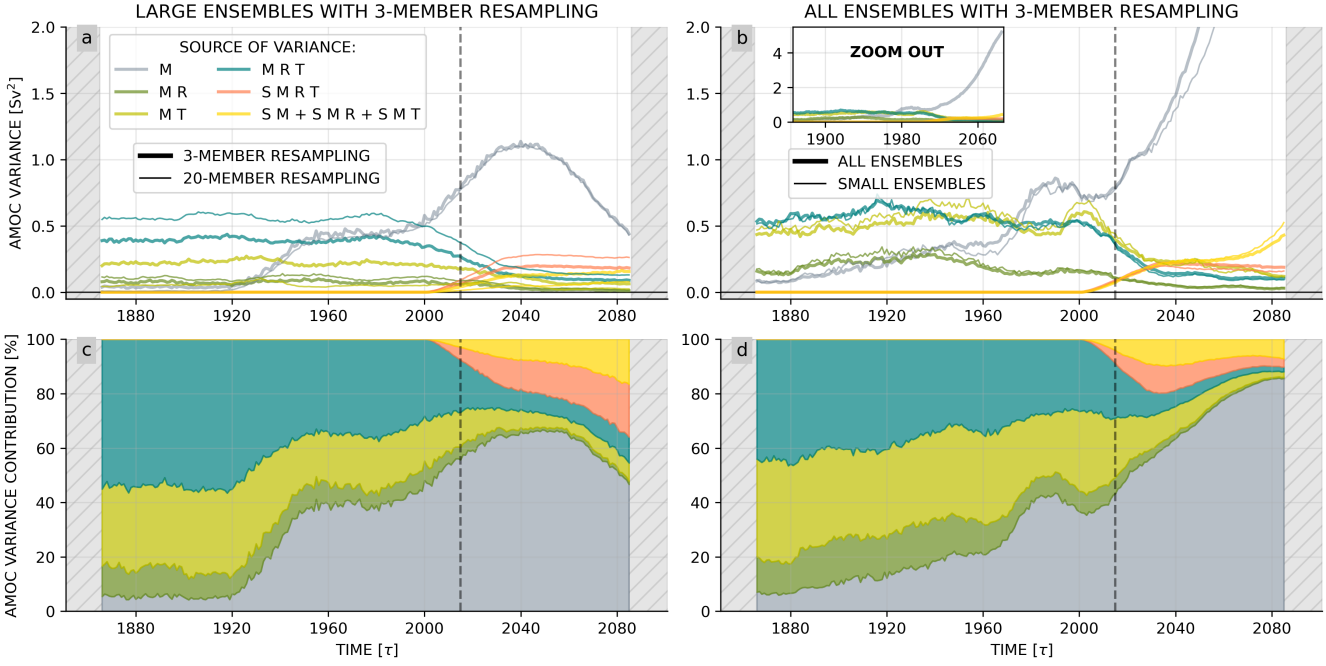

Figure A4. Sensitivity test on bootstrapping methodology: application to the variability of model-associated factors (a, b) Variance associated to model main effect and interactions. (c, d) Variance contribution of model main effect and interactions to the total model-associated variability. Left-hand panels show the sensitivity of the large ensemble results to the number of members in each resampling. Thin lines present the control experiment with 20 members of each model (identical to Fig. 3 a, c, e) and thick lines present the results obtained using 3 members of each model per resampling. Right-hand panels show the sensitivity to mixing large and small ensembles. Thin lines present the control experiment with only small ensembles (identical to Fig. 3 b, d, f) and thick lines present the results obtained with all ensembles. 3-member resampling are used here.

https://doi.org/10.5194/egusphere-2025-17
Preprint. Discussion started: 27 January 2025

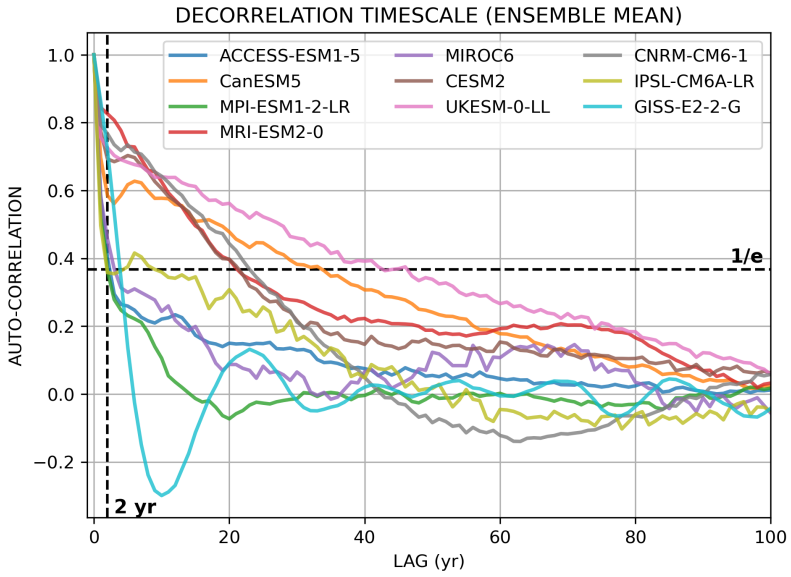

Figure A5. Decorrelation timescale of AMOC intensity time series for each model. Evolution of the normalized auto-correlation with respect to the considered lag. Decorrelation timescale corresponds to the the e-folding timescale, i.e. the lag when the normalized auto-correlation falls below $1/e$.