Supporting Information for "Detectability of an AMOC decline in current and projected climate changes"

D. Lobelle[1,2,3], C. Beaulieu[4,1], V. Livina[2], F. Sévellec[5,1], E. Frajka-Williams[6]

[1]Ocean and Earth Science, University of Southampton, United Kingdom

[2]National Physical Laboratory, Hampton Road, Teddington, United Kingdom

[3]Institute for Marine and Atmospheric Research, Utrecht University, Utrecht, Netherlands

[4]Ocean Sciences Department, University of California, Santa Cruz, United States

[5]Laboratoire d'Océanographie Physique et Spatiale, Plouzané, France

[6]National Oceanography Centre, Southampton, United Kingdom

Contents of this file

1. Text S1 to S4
2. Figures S1 to S3
3. Tables S1 to S3

Ocean and Earth Science, University of Southampton, United Kingdom (d.m.a.lobelle@uu.nl)

Introduction

There are four main parts to the supporting information. (1) Details regarding the residual tests of the CMIP5 and RAPID data are in Sections S1 and S2, with the accompanying Table S1 showing the results of these tests and the values of the trends, variances, and AR(1) coefficients for each model and RAPID time series. (2) The changepoint analysis in Section S3 is described, with an accompanying Figure S1 and Table S2. (3) The parametric distribution equations used to fit the distributions of the CMIP5 trends, AR(1) coefficients and variances are explained in Section S4. (4) The effects of changing the start month to define the annual means of RAPID and CMIP5 data on the values of the parameters in Figures S2 and S3, respectively, and effects on n^* estimates in Table S3.

Durbin-Watson Test S1.

The Durbin-Watson test (Durbin & Watson, 1950) is used to determine whether the residuals of a time series exhibit AR(1) auto-correlation (i.e. short-term memory). If the residuals at time t are represented by: $N_t = \phi N_{t-1} + \epsilon_t$, where ϕ represents the AR(1) coefficient and ϵ_t is the white noise error process, then the hypotheses are:

$$H_0 : \phi = 0, \tag{1}$$

$$H_a : \phi \neq 0, \tag{2}$$

where the null hypothesis (H_0 in Equation 1) is that the autocorrelation coefficient is 0 and thus the residuals are independent and only consist of white noise. The alternative hypothesis (H_a in Equation 2) is therefore that the autocorrelation coefficient is not 0 and the time series exhibits short-term memory. The Durbin-Watson test is represented by:

$$dw = \frac{\sum_{t=2}^{n} (N_t - N_{t-1})^2}{\sum_{t=1}^{n} N_t^2}, \tag{3}$$

where n is the number of observations and dw can be a value between 0 and 4. The null hypothesis is accepted and the residuals are uncorrelated when $dw = 2$, whereas a value of 0 to 2 indicates positive autocorrelation and 2 to 4 indicates negative autocorrelation, and therefore the alternative hypothesis is accepted.

Engle-ARCH Test S2.

The AutoRegressive Conditionally Heteroscedastic (ARCH) model proposed by Engle (1982) is based on a Lagrange Multiplier test by examining the significance of fitting a linear regression model to squared residuals:

$$N_t^2 = a_0 + a_1 N_{t-1}^2 + ... + a_L N_{t-L}^2 + \epsilon_t, \tag{4}$$

where L is the number of lags. The hypotheses have the following form:

$$H_0 : a_0 = a_1 = ... = a_L = 0, \tag{5}$$

$$H_a : a_0 \neq a_1 \neq ... \neq a_L \neq 0, \tag{6}$$

where the null hypothesis (H_0 in Equation 5) is that the squared residuals exhibit a constant variance through time (i.e., homoscedastic) since all terms apart from the white noise are 0. The null hypothesis is rejected, and the alternative hypothesis (H_a in Equation 6) is subsequently accepted when the variance changes through time (i.e., heteroscedastic).

The GLS fit requires the time series to exhibit non-spherical innovations, which means residuals that are autocorrelated and/or heteroscedastic (Brockwell & Davis, 2002). The stars in Table 1 of the main paper represent that the majority of the CMIP5 (11 models) reject the Durbin-Watson null hypothesis test and show autocorrelation. On the other hand, only one CMIP5 model rejects the Engle-ARCH null hypothesis test, therefore the majority of the numerical models and the observations have a constant variance through time. Following these results, when defining specifications for the AR simulations, the AR(1) coefficient is included and a constant variance value is used.

Changepoint Analysis S3.

It has been suggested that the RAPID time series exhibits trends and shifts, depending on the period and time resolution used for the analysis (Smeed et al., 2013, 2018; Moat et al., 2020). Detecting and distinguishing between trends or other signals (e.g., abrupt changes) in environmental data can be hindered by the presence of autocorrelation. To select the best model to describe the time series characteristics, we use the approach developed by Beaulieu and Killick (2018). Combinations of a trend or constant mean superimposed to white noise with or without autocorrelation are fitted and compared to time series with changepoints. We apply this technique to the RAPID and CMIP5 data to justify the use of a steady linear trend plus AR(1) model fit.

The changepoint analysis is run on `R` using the package `EnvCpt`. This analysis is developed such that a series of models are fit to the time series and the most appropriate model is selected according to an information criterion; here we use the Akaike information criterion (AIC). The information criterion are based on the log likelihood penalised by the number of parameters fitted. Further details on the algorithm and information criteria can be found in Beaulieu and Killick (2018).

Eight changepoint models are tested on the data which include:

(1) Mean: a constant mean and variance,

$$Y_t = \mu + \epsilon_t, \tag{7}$$

where μ (the mean), and ϵ_t are normally distributed residuals with a mean of zero and variance σ^2;

(2) Trend: a linear trend over time,

$$Y_t = \omega X_t + \lambda + \epsilon_t, \tag{8}$$

where λ is the intercept, or the constant term;

(3) Trend + AR(1): a linear trend over time with AR(1) errors,

$$Y_t = \omega X_t + \lambda + \phi Y_{t-1} + \epsilon_t, \tag{9}$$

(4) Mean cpt: multiple changepoints in the mean

$$Y_t = \begin{cases} \mu_1 + \epsilon_t, & t \leq c_1 \\ \mu_2 + \epsilon_t, & c_1 < t \leq c_2 \\ \vdots & \vdots \\ \mu_m + \epsilon_t, & c_{m-1} < t \leq n \end{cases}, \tag{10}$$

where the time series is split into m segments with means of $\mu_1, ..., \mu_m$, and variances $\sigma_1^2, ..., \sigma_m^2$. The timing of the changepoints between each segment is $c_1, ..., c_{m-1}$;

(5) Trend cpt: a trend with multiple changepoints in the regression parameters

$$Y_t = \begin{cases} \omega_1 X_t + \lambda_1 + \epsilon_t, & t \leq c_1 \\ \omega_2 X_t + \lambda_2 + \epsilon_t, & c_1 < t \leq c_2 \\ \vdots & \vdots \\ \omega_m X_t + \lambda_m + \epsilon_t, & c_{m-1} < t \leq n \end{cases}, \tag{11}$$

where $\omega_1, ..., \omega_m$ and $\lambda_1, ..., \lambda_m$ are the trend and intercept of each segment, respectively;

(6) Trend cpt + AR(1): a piecewise linear trend over time with AR(1) errors,

$$Y_t = \begin{cases} \omega_1 X_t + \lambda_1 + \phi_1 Y_{t-1} + \epsilon_t, & t \leq c_1 \\ \omega_2 X_t + \lambda_2 + \phi_2 Y_{t-1} + \epsilon_t, & c_1 < t \leq c_2 \\ \vdots & \vdots \\ \omega_m X_t + \lambda_m + \phi_m Y_{t-1} + \epsilon_t, & c_{m-1} < t \leq n \end{cases}. \tag{12}$$

Although the findings in Smeed et al. (2018) and Moat et al. (2020), show that for the RAPID observations the best model selection is a change in the mean occurring between 2008 and 2009, here, the AIC results select model (1) Mean as the best fit. The difference being that we use annual data as opposed to monthly data used in those previous studies, suggesting that a changepoint is not significant in annual RAPID data. In Table S1 half of the CMIP5 models show model (3) Trend + AR(1) as the best fit for the multi-decadal time series. This justifies the use of a linear trend and autocorrelation for the 100 years of simulated AMOC transport. We stress that for 14 years of RAPID data, the best model fit does not include a trend when using the AIC. Furthermore, as mentioned in the manuscript, the -0.11 Sv yr^{-1} is not significant ($p>0.05$); these two points support why the RAPID trend is not used to generate any simulations.

Empirical Distribution Equations S4.

In order to generate simulations of the AMOC time series, we randomly select 1,000 values within the range of the empirical cumulative distribution functions (eCDFs) from the 20 CMIP5 trend, variance and AR(1) coefficient values. To do this, we first choose the parametric cumulative distribution functions (pCDFs) with the best fit to each eCDF of the three CMIP5 parameters. After testing several distributions, the Johnson distribution best fits the eCDF of the CMIP5 trend and AR(1) distributions, and the inverse Gaussian distribution best fits the CDF of the CMIP5 variance distribution (shown in Figure 2). The Johnson system (Johnson, 1949) is a flexible parametric family of distributions including a wide variety of shapes. There are four distribution types that correspond to the following transformations: exponential or lognormal (S_L), logistic or bounded (S_B), hyperbolic sine or unbounded (S_U) and identity or normal (S_N). The scale and location parameters to describe the distributions are γ, η, τ and θ, and are used to produce a random variate, r, as follows:

$$r = \theta h(\frac{z-\gamma}{\eta}) + \tau, \tag{13}$$

where z is the standard normal random variable $N(0,1)$ and h is one of the four transformations described above. In this study, the CMIP5 AR(1) distribution type is S_B (therefore, h is a logistic transformation) and the CMIP5 trend distribution type is S_U (therefore, h is a hyperbolic sine transformation). The four parameters that describe the CMIP5 trend distribution are: γ = 0.38, η = 1.19, τ = –0.05, and θ = 0.02 (see Figure 2a). The four parameters that describe the CMIP5 AR(1) coefficient distribution are: γ = –1.46, η = 1.57, τ = –0.84, and θ = 1.42 (see Figure 2b).

The inverse Gaussian distribution (also known as the Wald distribution (Wald, 1944)) is used to model nonnegative positively skewed data, as is the case for the CMIP5 variance distribution. Two parameters are used to describe the distribution, the mean, where $\mu > 0$ and the shape parameter, where $\delta > 0$, as follows:

$$r = \sqrt{\frac{\delta}{2\pi x^3}} exp\left[-\frac{\delta}{2\mu^2 x}(x-\mu)^2\right], \tag{14}$$

where $\mu = 5.16$ and $\delta = 0.64$ in this study (see Figure 2c).

References

Beaulieu, C., & Killick, R. (2018). Distinguishing trends and shifts from memory in climate data. *Journal of Climate*, *31*(23), 9519–9543. doi: 10.1175/JCLI-D-17-0863.1

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting* (2nd ed.). New York: Springer. doi: 10.1007/978-3-319-29854-2

Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika*, *37*(3-4), 409–428. doi: 10.1093/biomet/37.3-4.409

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987–1007.

Johnson, N. L. (1949). Systems of Frequency Curves Generated by Methods of Translation. *Biometrika*, *36*(1), 149–176.

Moat, B., Smeed, D., Frajka-Williams, E., Desbruyères, D., Beaulieu, C., Johns, W., … Bryden, H. (2020). Pending recovery in the strength of the meridional overturning circulation at 26° N. *Ocean Science Discussions*(January), 1–22. doi: 10.5194/os-2019-134

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., Moat, B. I., Frajka-Williams, E., … McCarthy, G. D. (2018). The North Atlantic Ocean is in a state of reduced overturning. *Geophysical Research Letters*, *45*(3), 1527–1533. doi: 10.1002/2017GL076350

Smeed, D. A., McCarthy, G., Cunningham, S., Frajka-Williams, E., Rayner, D., Johns, W., … Bryden, H. (2013). Observed decline of the Atlantic Meridional Overturning Circulation 2004 to 2012. *Ocean Science Discussions*, *10*(5), 1619–1645. doi: 10

.5194/osd-10-1619-2013

Wald, A. (1944). On cumulative sums of random variables. *The Annals of Mathematical Statistics*, *15*(3), 283–296.

Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X.-L., Choi, D., Cheang, W.-K., … Frederick, J. E. (1998). Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *Journal of Geophysical Research*, *103*(D14), 17149–17161. doi: 10.1029/98JD00995

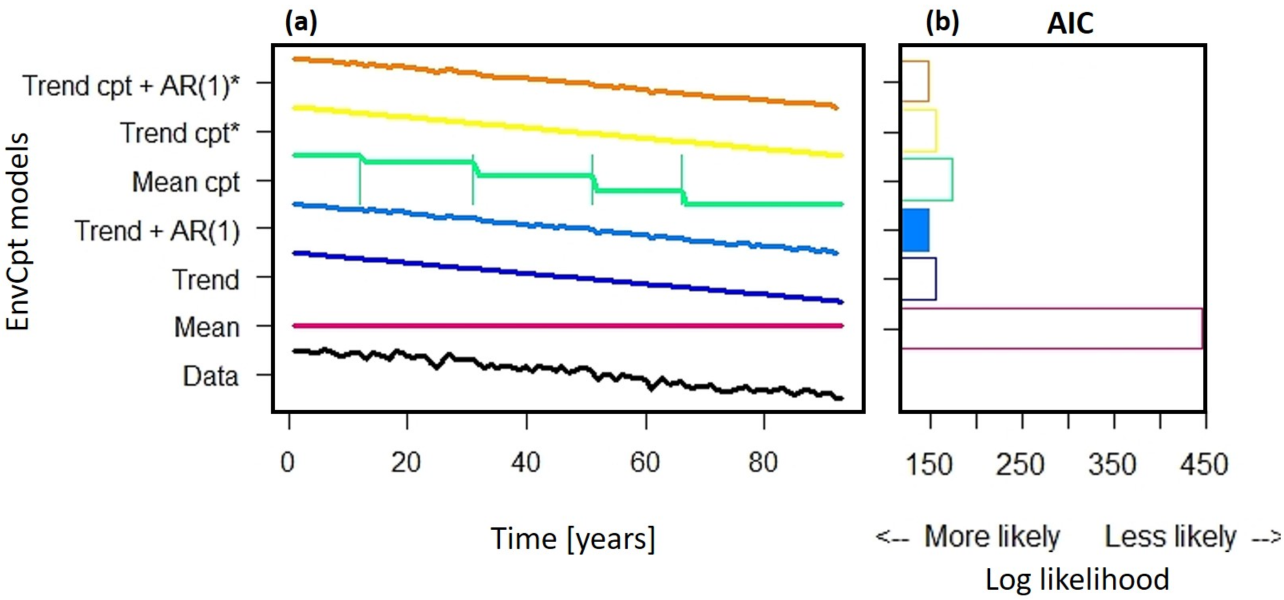

Figure S1. An example of the changepoint analysis output. (a) The models fit to the 93 years of NorESM1-M CMIP5 time series (black line). The stars next to 'Trend cpt + AR(1)' and 'Trend cpt' indicate that there is no changepoint in this fit. (b) The AIC log likelihood of each model, showing that the best-performing model (with the smallest AIC value) is the 'Trend + AR(1)'.

Table S1. The GLS trend magnitude [Sv yr^{-1}], variance [Sv^2] and AR(1) coefficients of 20 CMIP5 model simulations following the RCP8.5 scenario and the RAPID observations. [See S1 and S2 for descriptions on the residual analyses mentioned as footnotes; the Durbin-Watson test (Durbin & Watson, 1950) and the Engle-ARCH test (Engle, 1982)].

CMIP5 models	**Institute**	**Trend (ω) [Sv yr^{-1}]**	**Variance (σ_ϵ^2) [Sv^2]**	**AR(1) coeff (ϕ) [-]**
FIO-ESM	FIO (China)	−0.08 *	0.45	−0.13
NorESM1-M	NCC (Norway)	−0.10 *	0.32	0.30*
GFDL-CM3	NOAA GFDL (USA)	−0.12 *	0.53	0.30*
CESM1-BGC	NSF-DOE-NCAR (USA)	−0.06 *	0.47	−0.02
BNU-ESM	GCESS (China)	−0.09 *	0.52	0.28*
CCSM4	NCAR (USA)	−0.05 *	0.44	0.00
MIROC-ESM	MIROC (Japan)	−0.06 *	0.44	0.23*
ACCESS1-0	CSIRO-BOM (Australia)	−0.07 *	0.88*	0.21*
CNRM-CM5	CNRM-CERFACS (France)	−0.05 *	0.79	−0.17
GISS-E2-R-CC	NASA GISS (USA)	−0.07 *	1.06	0.14
HadGEM2-CC	MOHC (UK)	−0.07 *	0.96	0.15
MPI-ESM-LR	MPI-M (Germany)	−0.07 *	1.29	0.09
EC-EARTH	EC-EARTH (Europe)	−0.06 *	0.65	0.20*
inmcm4	INM (Russia)	−0.04 *	0.60	−0.02
CSIRO-Mk3-6-0	CSIRO-QCCCE (Australia)	−0.07 *	0.76	0.38*
IPSL-CM5A-MR	IPSL (France)	−0.05 *	0.64	0.26*
bcc-csm1-1	BCC (China)	−0.06 *	0.59	0.48*
CMCC-CM	CMCC (Italy)	−0.03 *	0.59	0.07
CanESM2	CCCMA (Canada)	−0.03 *	0.42	0.29*
MRI-CGCM3	MRI (Japan)	−0.02 *	0.48	0.23*
RAPID		−0.11	2.33	0.29

* in 'trend' column: the trend is significantly different from zero (5% significance)
* in 'variance' column: Engle's ARCH null hypothesis rejected (5% significance): variance changes through time
* in 'AR(1) coeff.' column: Durbin-Watson null hypothesis rejected (5% significance): residuals are autocorrelated

Table S2. The changepoint statistical analysis model selections for the time series of each CMIP5 model using the Akaike information criterion (AIC). For 10 of the 20 models Trend + AR(1) is selected as the best model.

CMIP5 models	AIC
FIO-ESM	Trend + AR(1)
NorESM1-M	Trend + AR(1)
GFDL-CM3	Mean cpt
CESM1-BGC	Trend + AR(1)
BNU-ESM	Trend + AR(1)
CCSM4	Trend
MIROC-ESM	Trend cpt
ACCESS1-0	Trend cpt AR(1)
CNRM-CM5	Trend + AR(1)
GISS-E2-R-CC	Trend + AR(1)
HadGEM2-CC	Trend + AR(1)
MPI-ESM-LR	Trend + AR(1)
EC-EARTH	Mean cpt
inmcm4	Trend
CSIRO-Mk3-6-0	Mean cpt
IPSL-CM5A-MR	Trend + AR(1)
bcc-csm1-1	Trend cpt AR(1)
CMCC-CM	Mean cpt
CanESM2	Trend + AR(1)
MRI-CGCM3	Mean cpt

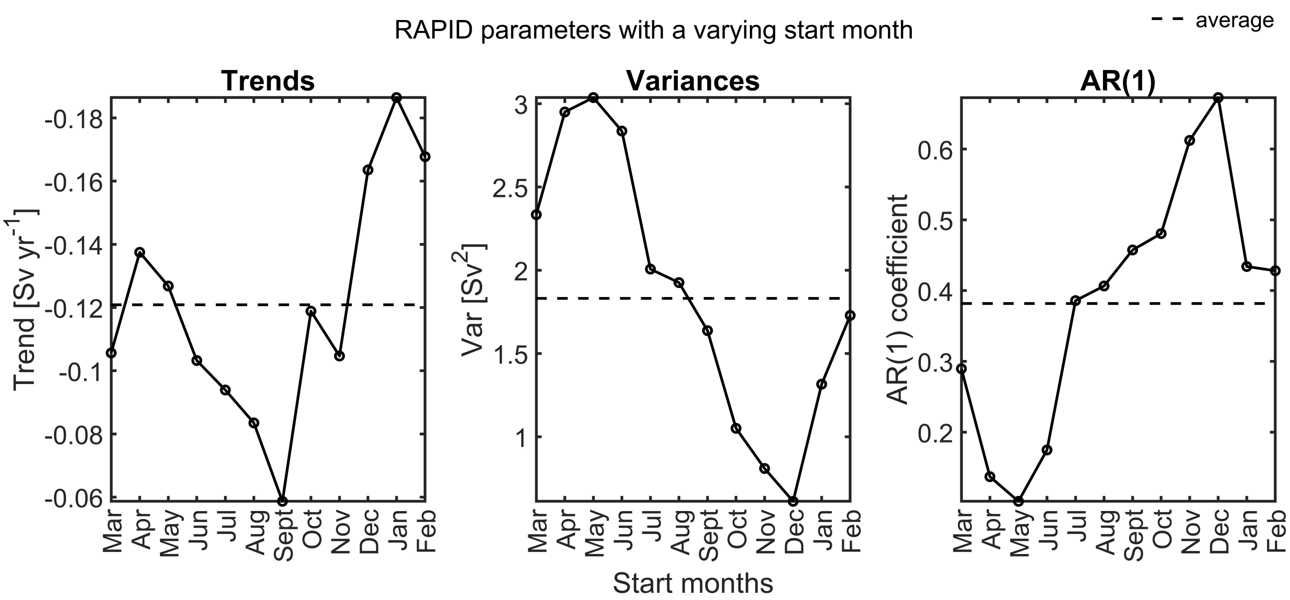

Figure S2. The sensitivity of the RAPID AMOC parameters (trend; ω, variance; σ_ϵ^2, AR(1) coefficient; ϕ) to the start month, when obtaining annual averages of the RAPID time series from monthly means. All three parameters are highly sensitive to the start month. We use a March start month, partially due to the how close they each are to their respective means. These means are displayed as dashed lines; (ω = -0.12 Sv yr^{-1}, σ_ϵ^2 = 1.83 Sv^2, ϕ = 0.38). Regardless of the start month, the trends are not significantly different from zero (p>0.05). The RAPID AMOC trend is therefore not used in the study while generating the simulations (i.e., in simRAPIDvar).

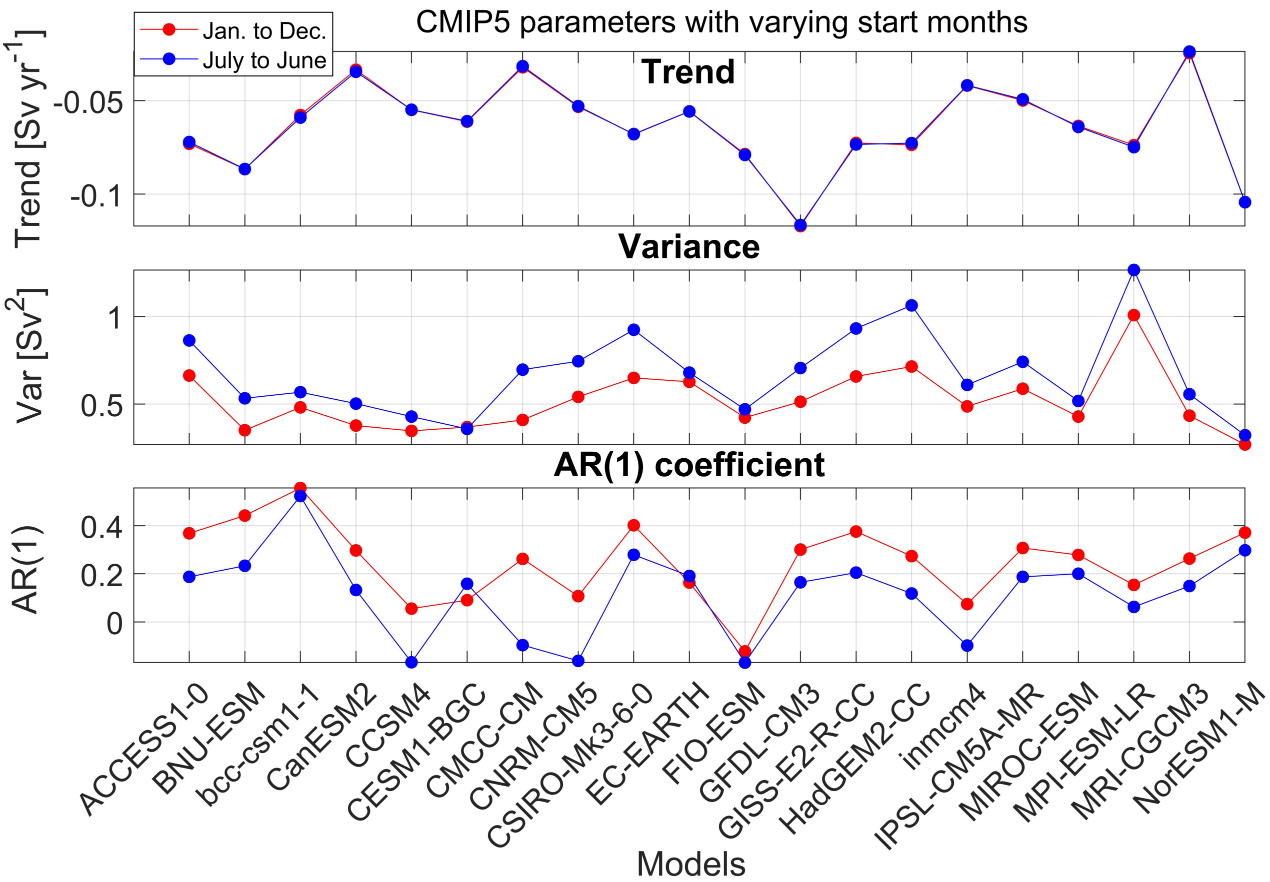

Figure S3. The sensitivity of the 20 CMIP5 AMOC model parameters (trend, variance, AR(1) coefficient) to the start month when, obtaining annual averages from monthly means. Here, two examples of start months are used; January to December (red) and July to June (blue). The intramodel comparison of the parameters shows that the trend is not sensitive to the start month for each model, whereas all models have a slightly higher variance using a July start month and almost all models have a slightly higher AR(1) coefficient using a January start month. Since the sensitivity to the start month is a lot larger in the RAPID data in Figure S2 relative to CMIP5, it is used to determine which start month to use for annually averaging the data in this study.

Table S3. The effect of changing the start month to compute the annually averaged time series on n^*. The Weatherhead et al. (1998) n^* approximation described in equation (10) in the main text is used with a 90% power of detection (hence $\zeta = 4.1$). The trend values remain constant (-0.6 Sv yr^{-1}; the median of the CMIP5 values) since the RAPID trends show p>0.05 (Figure S2). The AR(1) coefficient and variance are selected from the twelve RAPID AMOC values as a function of start month, in Figure S2. The median n^* is 43 years, which is also the value for the start month used in this study (March).

Start month	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.
n^* **[years]**	43	42	41	43	44	44	44	39	40	39	40	44