Geophysical Research Letters

Key Points:
- The estimated 14-year annually averaged RAPID AMOC trend of -0.1 Sv yr^{-1} is not statistically significant
- The detection window for a long-term AMOC decline is 14–42 years based on future scenario CMIP5 output statistical properties
- Autocorrelation of annually averaged AMOC data is weak and has a limited impact on detection of a long-term decline

Supporting Information:
- Supporting Information S1

Correspondence to:
D. Lobelle,
d.m.a.lobelle@uu.nl

Detectability of an AMOC Decline in Current and Projected Climate Changes

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

[1]Ocean and Earth Science, University of Southampton, Southampton, UK, [2]National Physical Laboratory, Teddington, UK, [3]Institute for Marine and Atmospheric Research, Utrecht University, Utrecht, Netherlands, [4]Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, CA, United States, [5]Laboratoire d'Océanographie Physique et Spatiale, Univ-Brest CNRS IRD Ifremer, Brest, France, [6]National Oceanography Centre, Southampton, UK

Abstract Determining whether the Atlantic Meridional Overturning Circulation (AMOC)'s transport is in decline is challenging due to the short duration of continuous observations. To estimate how many years are needed to detect a decline, we conduct a simulation study using synthetic data that mimics an AMOC time series. The time series' characteristics are reproduced using the trend, variance, and autocorrelation coefficient of the AMOC strength at 26.5°N from 20 Coupled Model Intercomparison Project Phase 5 (CMIP5) models under the RCP8.5 future scenario, and from RAPID observations (2004–2018). Our results suggest that the 14-year RAPID length has just entered the lower limits of the trend's "detection window" based on synthetic data generated using CMIP5 trends and variability (14–42 years; median = 24 years), but twice the length is required for detectability based on RAPID variability (29–67 years; median = 43 years). The annual RAPID trend is currently not statistically significant (-0.11 Sv yr^{-1}, $p > 0.05$).

Plain Language Summary There are ongoing discussions in the scientific community about whether the Atlantic Meridional Overturning Circulation transport is slowing down. This is of interest due to the importance of this circulation in transporting heat from the tropics to the northern latitudes. A consensus about its decline is hard to reach due to the limited direct observational data available; with the longest continuous data being 14 years long from 2004 to 2018. We therefore conduct a simulation experiment to examine how many years of data are required to detect a decline in the circulation. We create simulations of the North Atlantic transport based on statistical properties from 20 general circulation models with future climate change projections (until 2100) and from the RAPID array observations (since 2004). Our results demonstrate that the length of data we currently have from observations has just entered the "detection window" of 14–42 years (based on model simulations). However, the RAPID observations do not currently exhibit a statistically significant trend.

1. Introduction

Coupled ocean-atmosphere numerical models generally predict a decline of the Atlantic Meridional Overturning Circulation (AMOC) under the influence of anthropogenic warming in the 21st century (IPCC, 2019). The AMOC is responsible for ~25% of the globe's total meridional heat transport at 26°N, equivalent to 1.3 PW (1 PW = 10^{15} W) (Ganachaud & Wunsch, 2003; Hall & Bryden, 1982; Johns et al., 2011; Lavin et al., 1998). Heat is absorbed in surface waters of the tropical Atlantic and carried northward to be released to the atmosphere over the northeast Atlantic. This plays a role in maintaining milder northwestern European atmospheric temperatures, compared to the global average at the same latitudes (Pohlmann et al., 2006; Rhines et al., 2008).

At 26.5°N, the RAPID-MOCHA (Rapid Climate Change/Meridional Overturning Circulation and Heat-flux Array) program (hereafter, RAPID) has continuously monitored the AMOC since 2004 (Rayner & Kanzow, 2011). Initially, these observations suggested that a decline in AMOC transport could be occurring (Robson et al., 2014; Smeed et al., 2014), and later a sustained reduced mean transport post-2008 was detected (Smeed et al., 2018). Coupled Model Intercomparison Project Phase 5 (CMIP5) data and other studies using proxy indicators suggest that a long multidecadal declining trend could occur (Caesar et al., 2018; Cheng et al., 2013; IPCC, 2019; Rahmstorf et al., 2015; Robson et al., 2014; Sévellec et al., 2017; Stocker et al., 2013).

However, based on observations there are still ongoing discussions regarding whether the AMOC is in decline (McCarthy et al., 2012, 2015; Moat et al., 2020; Roberts et al., 2014; Smeed et al., 2018).

Detection of long-term (multidecadal) trends can be affected by numerous factors, such as the length of data available, the magnitude of the trend to be detected, the degree of variability (e.g., variance and autocorrelation, or short-term memory), and the trend detection technique used (Beaulieu et al., 2013; Henson et al., 2010; Tiao et al., 1990; Weatherhead et al., 1998). To explore these factors, we use the 14-year RAPID observations and 93-year CMIP5 future scenario model AMOC data which exhibit a range of trends and variability. We also compare four approaches for trend detection to assess sensitivity to the presence of autocorrelation and technique choice.

Previous studies have largely focused on time of emergence of an anthropogenically forced AMOC decline; that is, when a trend falls outside the range of its natural variability (Baehr et al., 2007, 2008; Jackson & Wood, 2020; Keller, Deutsch, et al., 2007; Keller, Kim, et al., 2007; Roberts & Palmer, 2012; Roberts et al., 2014; Santer et al., 1995; Vellinga & Wood, 2004; Williams et al., 2015). For example, Roberts et al. (2014) estimate that the -0.53 Sv yr^{-1} AMOC trend, assessed from the first 8 years of RAPID, takes 18 years to be significantly different from the internal variability found in the CMIP5 preindustrial control simulations. Here, we investigate the time needed to detect an AMOC decline from a monitoring perspective. As such, in contrast to the studies mentioned above, we do not use the variance from control runs to define the natural variability. We instead focus on detecting a trend signal, which is the step before investigating whether that trend falls outside of natural variability and is anthropogenically driven.

Our aim is to provide a timely estimate of the number of years (hereafter, n^*) required to detect a long-term decline in the AMOC at 26.5°N. We compare results using linear regressions that account for autocorrelation versus those that do not and CMIP5 variability versus RAPID variability. This latter comparison follows one of key findings from Roberts et al. (2014) that the CMIP5 variability could be underestimated. We present these results as probability density functions which provide uncertainties ("detection windows") around our n^* estimates.

2. Data

We use the numerical CMIP5 ensemble model output under the RCP8.5 ("business-as-usual") future scenario. Monthly means from 2006 to 2099 are used from the 20 models we select (Table S1 in the supporting information; the data are available from the portal Earth System Grid-Center for Enabling Technologies, on https://esgf-node.llnl.gov/search/cmip5/). To reduce model-associated biases, only one model per institute has been selected. We use the first ensemble member for each CMIP5 model used (r1i1p1, which specifies the realization, initial conditions, and physical parameter constants). For comparison to the observations, the zonally and vertically integrated CMIP5 AMOC streamfunction data are extracted at the closest latitude to 26.5°N and the depth at which the maximum streamfunction is found.

The observations from the RAPID transbasin array at 26.5°N are estimated as a twice-daily time series (McCarthy et al., 2015) from April 2004 to September 2018 (14.5 years; available from https://rapid.ac.uk). The RAPID AMOC time series is averaged to monthly means, then both CMIP5 and RAPID data are averaged to annual means from March to February, leading to a singular value for each year (Figure 1). The explanation for using a March start month can be found in section 4.1. The CMIP5 annual time series are therefore 93 years long (from March 2006 to February 2099).

3. Trend Detection Methods

Throughout this study, the term "linear trend" refers to the slope of a straight line. Four linear regression methods are used to evaluate the impact of autocorrelation on the n^* estimate. These include (1) ordinary least squares (OLS), (2) prewhitening (PW), (3) OLS with the effective number of degrees of freedom (OLSneff), and (4) generalized least squares (GLS), outlined from sections 3.1 to 3.4 below. All methods start with a simple linear trend model equation as a function of time, stated as:

$$Y_t = \omega X_t + \lambda + N_t, \qquad t = 1, \dots, n, \qquad (1)$$

where Y_t are the data (annual means of AMOC transport [1 Sv = 10^6 m^3 s^{-1}]), ω is the trend magnitude, λ is the constant term (or intercept), N_t are the residuals, t is time, and n is the sample size (i.e., corresponding
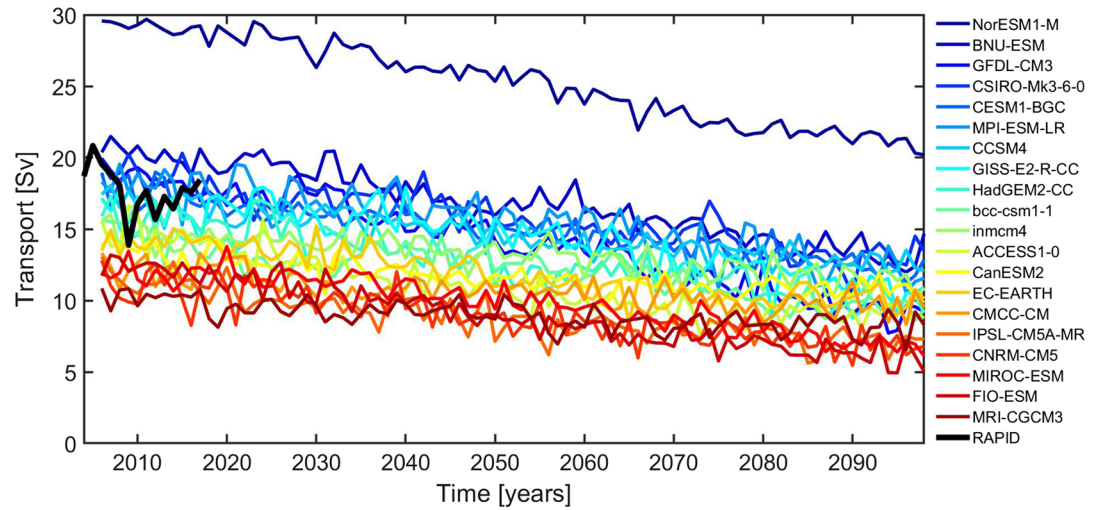
Figure 1. Time series of the annually averaged maximum AMOC streamfunction at 26.5°N represented by observations from the RAPID time series (black line) between 2004 and 2018 and 20 CMIP5 models (colored lines) between 2006 and 2099 following the RCP8.5 scenario.

to the total number of years of the time series Y_t). Classic linear regression models, such as OLS, assume the errors or residuals to be random and independent (white noise) and can be represented as ϵ_t. This assumption can be violated in climate time series data, where successive data points are often correlated due to short-term memory in the system (Harris et al., 2019; Tiao et al., 1990; Weatherhead et al., 1998). It is standard practice to take into account temporal dependence in climate change detection studies (Hartmann et al., 2013). The other three methods used in this study (PW, OLSneff, and GLS) account for autocorrelation by assuming a first-order autoregressive process [AR(1)]. Lastly, following Smeed et al.'s (2018) detection of a shift in the RAPID data mean strength in 2008, we also compare long-term trends to changepoint models to verify which models best describe the CMIP5 and RAPID AMOC time series (Text S3, using the R package EnvCpt described in Beaulieu & Killick, 2018).

3.1. Ordinary Least Squares (OLS)

Following Equation 1 above, the simplest method to estimate the trend, ω, is OLS:

$$\hat{\omega} = \frac{\sum_{t=1}^{n}(X_t - \overline{X})(Y_t - \overline{Y})}{\sum_{t=1}^{n}(X_t - \overline{X})^2}, \tag{2}$$

where the $\hat{}$ represents an estimate and the $\overline{}$ represents the mean. The white noise residuals are then:

$$\epsilon_t = Y_t - (\hat{\lambda} + \hat{\omega}X_t). \tag{3}$$

The unbiased estimate of the white noise variance (σ_ϵ^2) is derived from Equation 3:

$$\hat{\sigma}_\epsilon^2 = \frac{1}{n-2}\sum_{t=1}^{n}\epsilon_t^2, \tag{4}$$

where the denominator is $n-2$ since 2 degrees of freedom out of the original n are spent on fitting the two parameters (ω and λ).

3.2. Prewhitening (PW)

Prewhitening is a transformation to filter out the autocorrelation without removing the trend signal (Cochrane & Orcutt, 1949). The OLS regression is first fitted and then the AR(1) coefficient, ϕ, is estimated:

$$\hat{\phi} = \frac{\sum_{t=1}^{n-1}(e_t)(e_{t+1})}{\sum_{t=1}^{n}e_t^2}. \tag{5}$$

The prewhitening phase then occurs:

$$Y'_t = Y_{t+1} - \hat{\phi}Y_t, \qquad t = 1, \dots, n-1. \tag{6}$$

OLS is then once again applied to the transformed time series (Y'_t).

3.3. OLS With the Effective Number of Degrees of Freedom (OLSneff)

In the presence of autocorrelation, the number of independent observations is reduced. As such, one can account for autocorrelation in the residuals using OLS, by adjusting for the effective number of degrees of freedom (Von Storch & Zwiers, 2003):

$$n_{\text{eff}} = n\left(\frac{1-\hat{\phi}}{1+\hat{\phi}}\right), \tag{7}$$

where n_{eff} is the effective sample size. The estimate of the variance in Equation 4 is therefore adjusted using $n_{\text{eff}} - 2$ instead of $n - 2$ in the denominator.

3.4. Generalized Least Squares (GLS)

Once again, the starting point is the same linear regression as Equation 1, written in matrix notation as:

$$\mathbf{Y} = \omega\mathbf{X} + e. \tag{8}$$

Now $\mathbf{e}$ is treated as a random vector from the multivariate normal distribution $N(0, \boldsymbol{V})$, where $\boldsymbol{V}$ is a covariance matrix. In order to estimate the linear regression coefficient, the covariance matrix must be incorporated into the previous OLS estimate in Equation 2, as follows:

$$\hat{\omega}_{GLS} = (\mathbf{X}'\boldsymbol{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{V}^{-1}\mathbf{Y}, \tag{9}$$

where $\mathbf{X}'$ reflects the transpose of $\mathbf{X}$ (defined through the classical Euclidian scalar product). The covariance matrix has the effect of scaling the errors and hence "decorrelating" them (a detailed description can be found in Draper & Smith, 1998 or Brockwell & Davis, 2002). Here, it is assumed that $\boldsymbol{V}$ is a covariance matrix of an AR(1) process: $\boldsymbol{V} = (v_{tj}), v_{tj} = \sigma_\epsilon^2 \phi^{|i-j|}$. Since the covariance matrix is unknown, the σ_ϵ^2 and ϕ from the initial OLS residuals in Equations 4 and 5, respectively, are used as approximations. The R package `nlme` is used to estimate the GLS trend (Pinheiro & Bates, 2000).

4. Simulation Experiment Setup

4.1. Synthetic Series

We define a "year" of data as an average of monthly means from March to February. The RAPID AMOC time series parameters (the trend, variance and AR(1) coefficient) are sensitive to the start month chosen to generate annual means (Figure S2). A March start month is chosen for three reasons: (1) the three RAPID parameter values are close to their respective averages (computed using the 12 start months; Figure S2), (2) it allows for conservation of data, producing 14 years of RAPID data, and (3) CMIP5 model parameters are not very sensitive to start months (Figure S3).

The three parameters are then extracted from the 93-year-long numerical model simulations or 14-year-long RAPID observations. As described in section 3, the parameters are (1) the trend magnitude (ω), (2) the white noise variance (σ_ϵ^2), and (3) AR(1) coefficient (ϕ). Two sets of simulations are then generated from these parameters. The first set uses the three parameters from the CMIP5 models only (hereafter, simCMIP5var)
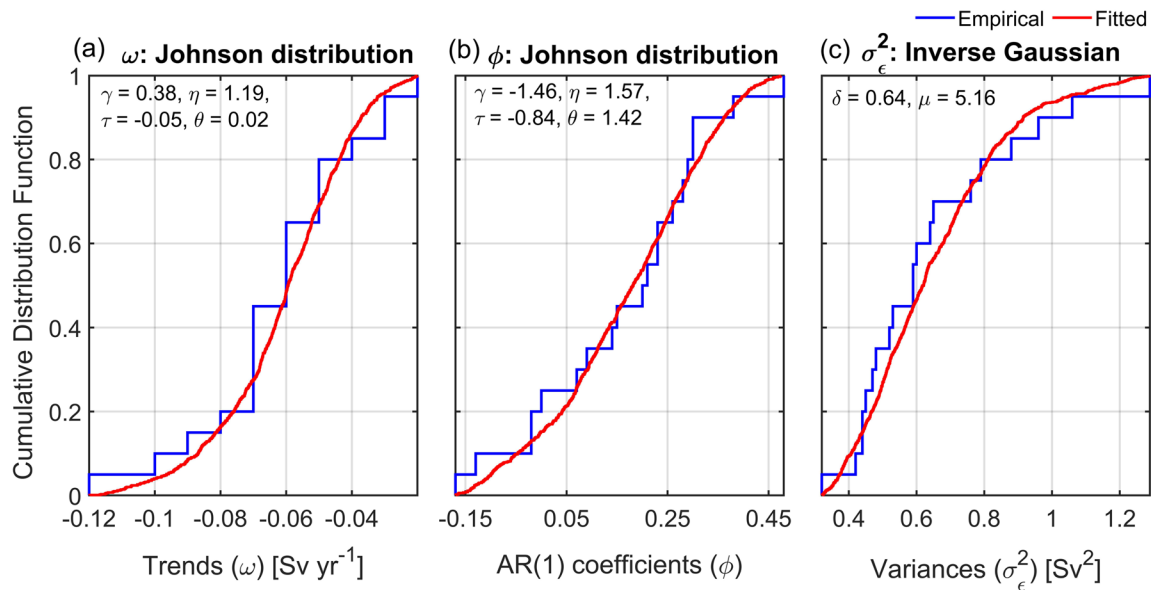
Figure 2. Empirical cumulative distribution functions (eCDFs; blue) and fitted parametric cumulative distribution functions (pCDFs; red) representing the three parameters extracted from the 20 CMIP5 models (in Table S1). Panels (a) and (b) represent the trend (ω) and AR(1) coefficient (ϕ) distributions, respectively, which follow a truncated Johnson distribution and panel (c) represents the variance (σ_ϵ^2) distribution which follows a truncated inverse Gaussian distribution. The coefficients describing the shape of each curve are also displayed, top-left of each subplot.

and the second combines the CMIP5 trends with the RAPID variance and AR(1) coefficient (hereafter, simRAPIDvar).

For simCMIP5var, the distribution of the 20 CMIP5 values for each parameter is used to calculate empirical cumulative distribution functions (eCDF) of ω, σ_ϵ^2, and ϕ (Figure 2). In order to randomly generate plausible values from CMIP5 distributions, we select parametric cumulative distribution functions (pCDFs) that best fit the eCDFs. A Johnson distribution (Johnson, 1949) best fits the ω and ϕ eCDFs and an inverse Gaussian distribution (Wald, 1944) best fits the σ_ϵ^2 eCDF (the coefficients to describe each pCDF are found in Figure 2 and Text S4). The pCDFs are truncated such that their tails do not extend beyond the minimum and maximum value within the CMIP5 ranges. In order to generate synthetic time series simulations, random values are selected 1,000 times from each parameters' truncated pCDF. These drawn values are therefore used to produce 1,000 synthetic simulations of the AMOC transport.

For simRAPIDvar, the same randomly picked CMIP5-based trend values as in simCMIP5var are used (from the pCDF in Figure 2a). The difference here is that all 1,000 simulations use RAPID's white noise variance (2.33 Sv^2) and AR(1) coefficient (0.29). This simRAPIDvar setup follows suggestions from previous studies that AMOC variability is underestimated in CMIP5 models on interannual timescales (McCarthy et al., 2012; Roberts et al., 2014).

4.2. n^* Approximation

Trends are estimated from the simulated time series as a function of sample size. For each of the 1,000 randomly selected combination of parameters, 100 time series are generated for different sample sizes (8 through to 100 years), in order to ensure independence throughout the trend detection method. For each sample of simulated data, we use the four regression models, described in section 3 and identify whether the trend is significantly different from zero. Note that since model projections agree on a decreasing trend, the hypothesis tested here is only a negative trend (i.e., one-tailed test). A statistically significant declining trend with 95% confidence is identified when $\omega < 0$ and $p < 0.05$. The proportion of the 100 time series generated for each sample size with significant declining trends is counted. Three different powers of detection are used in this study, meaning that the minimum sample size with significant trends in 80%, 90%, or 95% of the 100 time series is identified as the length of data required to detect a declining trend (n^*). This is repeated for each of the 1,000 simulations, and the results are represented as a probability density function (PDF); and summarized as a median n^* and the 5th–95th percentile interval representing the uncertainty, which we also call a "detection window."

To assess the impact of changes of parameters more generally, we also include the Weatherhead et al. (1998) approximation which estimates n^* in monthly time series as a function of the noise standard deviation, the trend magnitude, and the first-order autocorrelation:

$$n^* = \zeta \left(\frac{\sigma_N}{|\omega_0|} \sqrt{\frac{1+\phi}{1-\phi}} \right)^{2/3},$$ (10)

where ζ is a unitless constant that depends on the critical level and the power of detection.
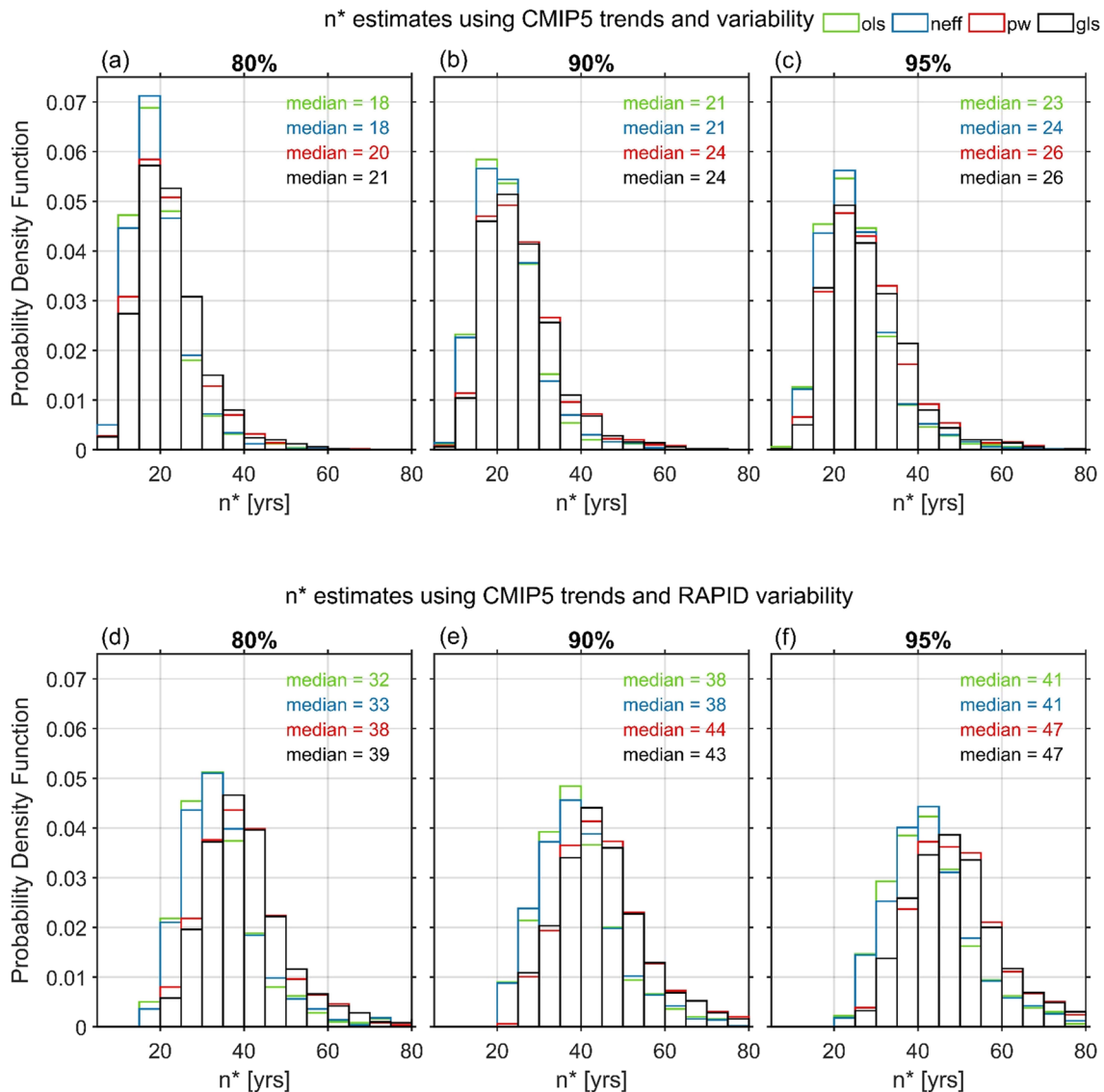
5. Results

The 20 CMIP5 parameter values range from -0.02 to -0.12 Sv yr^{-1} for ω, 0.32 to 1.29 Sv^2 for σ_e^2, and -0.17 to 0.48 for ϕ (see Table S1 for each model's values and results from the Durbin and Watson (1950) and Engle (1982) residual analyses; Texts S1 and S2). A model comparison between linear trend and changepoint models (Beaulieu & Killick, 2018; Killick et al., 2012) suggests that the best model fit for 10 of the 20 model time series is a trend with first-order autocorrelation (see Table S2). Though this is the most commonly selected model, there lacks a unanimous agreement from the information criterion results of the other 10 CMIP5 models. This serves as a motivation to verify the impact of accounting for autocorrelation on trend detection of an AMOC decline.

The n^* PDFs (Figure 3) indicate the minimum number of years required to detect a statistically significant long-term decline in the AMOC (using a 5% critical level). From simCMIP5var, the largest n^* values result from the GLS regression method, with median values of 21 [12–36], 24 [14–42], and 26 [16–45] years for increasing powers of detection; 80%, 90%, and 95%, respectively (Figures 3a–3c). The ranges presented in brackets after the median values are the 5th–95th percentiles (or "detection window"). Results from the OLS analysis produce the lowest n^* estimate; 18 [10–32], 21 [12–36], and 23 [14–41] years, respectively. These results are expected since autocorrelation in a time series can be confounded with a linear trend, hence n^* increases when using a regression method that accounts for short-term memory (Beaulieu et al., 2013; Harris et al., 2019; Weatherhead et al., 1998). The other two regression methods, OLSneff and PW, produce median n^* values above the OLS median and below the GLS median results. When considering the uncertainty of these results, the four regression methods produce median n^* estimates that are within the detection window of all GLS results, suggesting that accounting for autocorrelation does not have a large impact on annual AMOC trend detection.

For simRAPIDvar, the 14-year annual trend from RAPID of -0.11 Sv yr^{-1} is not used, first, because it is not significantly different from zero ($p > 0.05$, Table S1), and second, because this study's aim is to estimate the n^* for a long-term (multidecadal) decline scenario. Here, compared to simCMIP5var, the median GLS n^* values from the PDFs increase by 18–20 years (Figures 3d–3f; to 39 [26–60] years, 43 [29–67] years and 47 [32–72] years for an 80%, 90%, and 95% power of detection, respectively). Again, the OLS, OLSneff, and PW median n^* results fall within the GLS detection window. Comparing results from simCMIP5var and simRAPIDvar indicates that, as expected, n^* increases with an increase in variability or noise; RAPID's variance is almost twofold larger than the largest CMIP5 value (MPI-ESM-LR; Table S1).

The Weatherhead et al. (1998) n^* approximation in Equation 10, also supports that the number of years necessary to detect a trend is influenced by the three terms: the trend magnitude, the standard deviation, and the strength of autocorrelation. When using a critical level of 5% and a power of detection of 90%, the constant, ζ, is on the order of 4.1 (or 3.6 for a 95% detection and 4.5 for an 80% detection). With the same ranges of trend magnitude and autocorrelation coefficient, a change in the noise standard deviation impacts the detection timescale. Taking, for example, the median CMIP5 trend of -0.06 Sv yr^{-1}, the median CMIP5 variance of 0.6 Sv^2, and the median CMIP5 autocorrelation coefficient of 0.2, we obtain an n^* of 26 years. If we now replace the median variance by a variance of 2.3 Sv^2, as observed in the RAPID AMOC time series, n^* increases to 40 years.

Based on this approximation, a positive autocorrelation increases n^* by a factor of $\left(\frac{1+\phi}{1-\phi} \right)^{1/3}$. For example, when comparing two time series with the same trend magnitude (say -0.1 Sv yr^{-1}) and a noise standard deviation of 1 Sv, the n^* from a time series with no autocorrelation versus a positive autocorrelation coefficient of 0.5 increases from 19 to 27 years. These results from the approximation in Equation 10 demonstrate that n^* increases with a larger variance, stronger autocorrelation, and weaker trend magnitude.

Figure 3. PDFs of the minimum number of years required to detect a declining AMOC trend (n^*). Panels (a)–(c) are results from simCMIP5var (generated using 1,000 values from the pCDFs of the CMIP5 trends, AR(1) coefficients, and variances in Figure 2). Panels (d)–(f) are results from simRAPIDvar (generated with the same trends as in simCMIP5var but RAPID's AR(1) coefficient, 0.29, and variance, 2.33 Sv^2). The four linear regressions used are displayed: ordinary least squares ("ols") in green, OLS with effective number of degrees of freedom ("neff") in blue, prewhitening ("pw") in red, and generalized least squares ("gls") in black. All methods produce similar results, when comparing the medians of the PDFs. The power of detection increases from 80% in (a) and (d), to 90% in (b) and (e), and to 95% in (c) and (f). See section 4 for further details on the simulation experiment setup.

We also quantify the effect of different start months to compute a year of data on the n^* using the approximation from Weatherhead et al. (1998). We set the trend value to the median CMIP5 trend of −0.06 Sv yr^{-1} (since neither start months produce a significant RAPID AMOC trend, $p > 0.05$), while the 12 different RAPID AMOC variance and AR(1) coefficients from the different start months are used. The n^* estimates only vary between 39 and 44 years, with a median of 43 years, which coincides with the March start month n^* (Table S3).

6. Discussion

We have estimated the number of years of AMOC transport monitoring required to detect a statistically significant long-term decline at 26.5°N. We use synthetic simulations based on parameters from CMIP5 numerical model output and RAPID in situ observations. For a 90% power of detection, synthetic time series based on CMIP5 trends and variability (simCMIP5var) produce a median n^* of 24 [14–42] years, with

brackets representing the 5th–95th percentiles or "detection window". The RAPID time series length is therefore at the edge of the simCMIP5var detection window. When including RAPID's larger variability and autocorrelation (simRAPIDvar), n^* increases to 43 [29–67] years. This result is to be expected since a lower signal-to-noise ratio increases the trend detection time. This is corroborated with the n^* approximation from Weatherhead et al. (1998); the combination of a strong trend, low white noise standard deviation and weak short-term memory produces the fastest detection time.

Previous work has accounted for the AMOC's autocorrelation by using different techniques, such as generation of ARMA simulations (Roberts et al., 2014), adjusting for the effective number of degrees of freedom (Smeed et al., 2014), and bootstrapping techniques (Baehr et al., 2007). Here, a number of regression methods are used to compare results from trend detection approaches that do not account for autocorrelation (OLS) with a few that do (PW, OLSneff, and GLS). While n^* values are smaller than GLS when using OLS (by 3–7 years, depending on the power of detection and simulation experiment setup), all n^* median values are within the detection window of each other. This demonstrates that autocorrelation in the AMOC time series on an annual timescale has a limited impact on detecting its decline. Such a finding is beneficial as a baseline for future AMOC trend detection studies since our results suggest that a simple least squares regression technique is sufficient to capture the signal. We do not rule out, however, that the decorrelation timescale of the AMOC could be less than a year and hence using a daily or monthly mean transport may change these results.

The n^* estimates are dependent on the simulation experiment setup, and therefore careful choices have been made to best replicate the AMOC's statistical properties (i.e., trend, variance and AR(1) coefficient), especially for simRAPIDvar. The 14-year RAPID AMOC trend ($-0.11 \, \text{Sv yr}^{-1}$) is not used since it is not significantly different from zero ($p > 0.05$). This is contrary to the fivefold larger trend magnitude estimated from 2004 to 2012 ($-0.53 \, \text{Sv yr}^{-1}$, $p < 0.01$) by Roberts et al. (2014). This difference can potentially be explained by the fact that the minimum AMOC transport strength recorded in 2009–2010 (McCarthy et al., 2012) induces a stronger 8-year trend, and the slight increased mean transport post-2012 (Smeed et al., 2018) weakens the 14-year trend. RAPID AMOC variability is used in simRAPIDvar to account for the potential underestimated variability in CMIP5 time series on interannual timescales (McCarthy et al., 2012; Roberts et al., 2014). Finally, we choose to characterize the overall noise in the system (the variance and autocorrelation) from the same time series (RAPID) as opposed to the Roberts et al. (2014) approach which uses the CMIP5 AR(1) coefficients and the white noise variability from the RAPID data.

Although the RAPID AMOC parameters (trend, variance and AR(1) coefficient) are sensitive to the start month used for annual averaging (Figure S2), which could be due to large subannual variability (Hirschi et al., 2007), n^* is not affected. This is shown quantitatively by the Weatherhead et al. (1998) n^* approximation (Table S3) that produces a range of only 5 years (from 39 to 44 years) using different start months. This could be due to the variance and AR(1) coefficient values showing opposing behaviors as a function of start month (Figure S2), and hence having a minor impact on the overall noise term and n^*. The March start month has been used here which coincides with the median Weatherhead et al. (1998) n^* of 43 years (Table S3).

Several assumptions were made in this study that may impact the results. First, the parameters estimated from the models and observations are assumed to be stationary and therefore constant through time. However, this assumption seems reasonable since the Engle-ARCH test (Text S2 and Table S1) shows that 19 of the 20 models have a constant variability through time. As for the trend itself, even though we do not rule out that it may change, we do not currently have evidence for a changing trend to include it in our simulation experiment. The presence of long-term memory in the AMOC time series could affect our results since long-term memory will tend to increase trend detection time (e.g., Ludescher et al., 2017). However, our study suggests that the annual AMOC time series (from CMIP5 models and RAPID) displays short-term memory, if any.

7. Conclusion

Due to the numerous potential impacts of a decline in the AMOC transport, detecting whether its trend is significant is timely. Here, our main findings are threefold. First, the longest continuous observations available (14 years) from the annual RAPID time series do not exhibit a significant trend. Second, statistically significant AMOC trend detection timescales using the variability from CMIP5 under the RCP8.5 scenario result in the 14–year length just entering the n^* detection window (14–42 years) and the inclusion of the

larger RAPID variability means that at least 29 years are required. And finally, autocorrelation does not significantly affect trend detection of a multidecadal AMOC decline.

Data Availability Statement

The CMIP5 data can be found from the portal Earth System Grid-Center for Enabling Technologies (ESG-CET) at this site (https://esgf-node.llnl.gov/search/cmip5/). Data from the RAPID AMOC monitoring project are funded by the Natural Environment Research Council and are freely available at this site (https://rapid.ac.uk).

References

Baehr, J., Haak, H., Alderson, S., Cunningham, S., Jungclaus, J., & Marotzke, J. (2007). Timely detection of changes in the meridional overturning circulation at 26°N in the Atlantic. *Journal of Climate*, *20*(23), 5827–5841. https://doi.org/10.1175/2007JCLI1686.1

Baehr, J., Keller, K., & Marotzke, J. (2008). Detecting potential changes in the meridional overturning circulation at 26°N in the Atlantic. *Climatic Change*, *91*(1–2), 11–27. https://doi.org/10.1007/s10584-006-9153-z

Beaulieu, C., Henson, S. A., Sarmiento, J. L., Dunne, J. P., Doney, S. C., Rykaczewski, R. R., & Bopp, L. (2013). Factors challenging our ability to detect long-term trends in ocean chlorophyll. *Biogeosciences*, *10*(4), 2711–2724. https://doi.org/10.5194/bg-10-2711-2013

Beaulieu, C., & Killick, R. (2018). Distinguishing trends and shifts from memory in climate data. *Journal of Climate*, *31*(23), 9519–9543. https://doi.org/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. https://doi.org/10.1007/978-3-319-29854-2

Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., & Saba, V. (2018). Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature*, *556*(7700), 191–196.

Cheng, W., Chiang, J. C. H., & Zhang, D. (2013). Atlantic Meridional Overturning Circulation (AMOC) in CMIP5 models: RCP and historical simulations. *Journal of Climate*, *26*(18), 7187–7197. https://doi.org/10.1175/JCLI-D-12-00496.1

Cochrane, D., & Orcutt, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, *44*(245), 32–61.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York, NY: John Wiley & Sons.

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

Ganachaud, A., & Wunsch, C. (2003). Large-scale ocean heat and freshwater transports during the world ocean circulation experiment. *Journal of Climate*, *16*, 696–705. https://doi.org/10.1175/1520-0442(2003)016<0696:LSOHAF>2.0.CO;2

Hall, M. M., & Bryden, H. L. (1982). Direct estimates and mechanisms of ocean heat transport. *Deep Sea Research Part A, Oceanographic Research Papers*, *29*(3), 339–359. https://doi.org/10.1016/0198-0149(82)90099-1

Harris, P., Sotirakopoulos, K., Robinson, S., Wang, L., Livina, V., Harris, P., et al. (2019). A statistical method for the evaluation of long term trends in underwater noise measurements. *The Journal of the Acoustical Society of America*, *228*, 228–242. https://doi.org/10.1121/1.5084040

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., et al. (2013). Observations: Atmosphere and surface. In T. F. Stocker, et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of working group i to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Henson, S. A., Sarmiento, J. L., Dunne, J. P., Bopp, L., Lima, I., Doney, S. C., et al. (2010). Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences*, *7*, 621–640. https://doi.org/10.5194/bg-7-621-2010

Hirschi, J. J.-M., Killworth, P. D., & Blundell, J. R. (2007). Subannual, seasonal, and interannual variability of the North Atlantic meridional. *Journal of Physical Oceanography*, *37*, 1246–1265. https://doi.org/10.1175/JPO3049.1

IPCC (2019). Summary for policymakers. In H.-O. Pörtner, et al. (Eds.), *IPCC special report on the ocean and cryosphere in a changing climate*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Jackson, L. C., & Wood, R. A. (2020). Fingerprints for early detection of changes in the AMOC. *Journal of Climate*, *33*, 7027–7044. https://doi.org/10.1175/JCLI-D-20-0034.1

Johns, W. E., Baringer, M. O., Beal, L. M., Cunningham, S. A., Kanzow, T., Bryden, H. L., et al. (2011). Continuous, array-based estimates of Atlantic Ocean heat transport at 26.5°N. *Journal of Climate*, *24*(10), 2429–2449. https://doi.org/10.1175/2010JCLI3997.1

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, *36*(1), 149–176.

Keller, K., Deutsch, C., Hall, M. G., & Bradford, D. F. (2007). Early detection of changes in the North Atlantic meridional overturning circulation: Implications for the design of ocean observation systems. *Journal of Climate*, *20*(2), 145–157. https://doi.org/10.1175/JCLI3993.1

Keller, K., Kim, S. R., Baehr, J., Bradford, D. F., & Oppenheimer, M. (2007). What is the economic value of information about uncertain climate thresholds? *Human-induced climate change: An interdisciplinary assessment* (pp. 343–354). Cambridge: Cambridge University Press.

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Lavin, A., Bryden, H. L., & Parrilla, G. (1998). Meridional transport and heat flux variations in the subtropical North Atlantic. *The Global Atmosphere and Ocean System*, *6*, 269–293.

Ludescher, J., Bunde, A., & Schellnhuber, H. J. (2017). Statistical significance of seasonal warming/cooling trends. *Proceedings of the National Academy of Sciences*, *114*(15), 1–6. https://doi.org/10.1073/pnas.1700838114

McCarthy, G. D., Frajka-Williams, E., Johns, W. E., Baringer, M. O., Meinen, C. S., Bryden, H. L., et al. (2012). Observed interannual variability of the Atlantic meridional overturning circulation at 26.5°N. *Geophysical Research Letters*, *39*, L19609. https://doi.org/10.1029/2012GL052933

McCarthy, G. D., Smeed, D. A., Johns, W. E., Frajka-Williams, E., Moat, B. I., Rayner, D., et al. (2015). Measuring the Atlantic meridional overturning circulation at 26°N. *Progress in Oceanography*, *130*, 91–111. https://doi.org/10.1016/j.pocean.2014.10.006

Moat, B., Smeed, D., Frajka-Williams, E., Desbruyères, D., Beaulieu, C., Johns, W., et al. (2020). Pending recovery in the strength of the meridional overturning circulation at 26°N. *Ocean Science Discussions*, *16*, 863–874. https://doi.org/10.5194/os-2019-134

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. New York: Springer-Verlag.

Pohlmann, H., Sienz, F., & Latif, M. (2006). Influence of the multidecadal Atlantic meridional overturning circulation variability on European climate. *Journal of Climate*, *19*(23), 6062–6067. https://doi.org/10.1175/JCLI3941.1

Rahmstorf, S., Box, J. E., Feulner, G., Mann, M. E., Robinson, A., Rutherford, S., & Schaffernicht, E. J. (2015). Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. *Nature Climate Change*, *5*(5), 475–480. https://doi.org/10.1038/nclimate2554

Rayner, D., & Kanzow, T. (2011). The design strategy and methodology of the RAPID-MOC project moorings. *Underwater Technology*, *29*(4), 159–171. https://doi.org/10.3723/ut.29.159

Rhines, P., Häkkinen, S., & Josey, S. A. (2008). Is oceanic heat transport significant in the climate system? In R. R. Dickson, J. Meincke, P. Rhines (Eds.), *Arctic-subarctic ocean fluxes* (pp. 87–109). Dordrecht: Springer.

Roberts, C. D., Jackson, L., & McNeall, D. (2014). Is the 2004–2012 reduction of the Atlantic meridional overturning circulation significant? *Geophysical Research Letters*, *41*, 3204–3210. https://doi.org/10.1002/2014GL059473.1

Roberts, C. D., & Palmer, M. D. (2012). Detectability of changes to the Atlantic meridional overturning circulation in the Hadley Centre Climate Models. *Climate Dynamics*, *39*(9–10), 2533–2546. https://doi.org/10.1007/s00382-012-1306-3

Robson, J., Hodson, D., Hawkins, E., & Sutton, R. (2014). Atlantic overturning in decline? *Nature*, *7*(1), 2–3. https://doi.org/10.1038/ngeo2050

Sévellec, F., Federov, A. V., & Liu, W. (2017). Arctic sea ice decline weakens the Atlantic Meridional Overturning Circulation. *Nature Climate Change*, *7*, 604–610. https://doi.org/10.1038/nclimate3353

Santer, B. D., Mikolajewicz, U., Brüggemann, W., Cubasch, U., Hasselmann, K., Höck, H., et al. (1995). Ocean variability and its influence on the detectability of greenhouse warming signals. *Journal of Geophysical Research*, *100*(C6), 10,693–10,725. https://doi.org/10.1029/95JC00683

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

Smeed, D. A., McCarthy, G. D., Cunningham, S. A., Frajka-Williams, E., Rayner, D., Johns, W. E., et al. (2014). Observed decline of the Atlantic meridional overturning circulation 2004–2012. *Ocean Science*, *10*, 29–38. https://doi.org/10.5194/os-10-29-2014

Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., et al. (2013). Summary for policymakers. *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 3–29). Cambridge, UK, and New York, USA: Cambridge University Press.

Tiao, G. C., Reinsel, G. C., Xu, D., Pedrick, J. H., Zhu, X., Miller, A. J., et al. (1990). Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation. *Journal of Geophysical Research*, *95*(D12), 20,507–20,517. https://doi.org/10.1029/JD095iD12p20507

Vellinga, M., & Wood, R. A. (2004). Timely detection of anthropogenic change in the Atlantic meridional overturning circulation. *Geophysical Research Letters*, *31*, L14203. https://doi.org/10.1029/2004GL020306

Von Storch, H., & Zwiers, F. W. (2003). *Statistical analysis in climate research*. Cambridge: Cambridge University Press.

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., et al. (1998). Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *Journal of Geophysical Research*, *103*(D14), 17,149–17,161. https://doi.org/10.1029/98JD00995

Williams, J., Hughes, C. W., & Tamisiea, M. E. (2015). Detecting trends in bottom pressure measured using a tall mooring and altimetry. *Journal of Geophysical Research: Oceans*, *120*, 5216–5232. https://doi.org/10.1002/2015JC010955