Effects of strongly eddying oceans on multidecadal climate variability in the Community Earth System Model

André Jüling[1], Anna von der Heydt[1], and Henk A. Dijkstra[1]

[1]Institute for Marine and Atmospheric research Utrecht (IMAU), Utrecht University, Netherlands

Correspondence: André Jüling (a.juling@uu.nl)

Abstract. Climate variability on multidecadal time scales appears to be organized in pronounced patterns with clear expressions in sea surface temperature, such as the Atlantic Multidecadal Variability and the Pacific Decadal Oscillation. These patterns are now well studied both in observations and global climate models and are important in the attribution of climate change. Results from CMIP5 models have indicated large biases in these patterns with consequences for ocean heat storage

5   variability and eventually the global mean surface temperature. In this paper, we use two multi-century Community Earth System Model simulations at coarse (1°) and fine (0.1°) ocean model horizontal grid spacing to study the effects of the representation of mesoscale ocean flows on major patterns of multidecadal variability. We find that resolving mesoscale ocean flows both improves the characteristics of the modes of variability with respect to observations and increases the amplitude of the heat content variability in the individual ocean basins. The effect on the global mean surface temperature is relatively minor.

10   **Keywords.** Multidecadal Climate Variability

       Mesoscale ocean flows

       Ocean Heat Content

       Global Mean Surface Temperature

15   ## 1   Introduction

The ocean plays a key role in the climate system's heat budget, absorbing some 93% of the additional heat retained due to anthropogenic greenhouse gases (Stocker et al., 2013). Although the instrumental record of sea surface temperature (SST) is only about one and a half centuries, the observations indicate the existence of spatially correlated patterns of variability on multidecadal time scales, also referred to as (statistical) modes of variability (Deser et al., 2010). These modes are thought

20   to be part of the internal variability of the climate system and they affect the oceanic heat content by altering heat fluxes and consequently the global energy budget (Trenberth and Shea, 2006; Dijkstra, 2013; Zhang and Wang, 2013; Frajka-Williams et al., 2017). Disentangling these modes of internal variability from forced changes is hence important for detection and attribution studies of anthropogenic climate change (Zhang et al., 2019). Indeed, these SST patterns have played a major role

in the search for the origin of the most recent global mean surface temperature trend slowdown (Kosaka and Xie, 2013; England
25   et al., 2014). They may also reflect a memory component of the climate system, which can provide enhanced skill in long-term
climate predictions (Zhang et al., 2019). On a broader perspective, societies are impacted significantly by these SST patterns
through associated changes in ecosystem productivity (Mantua et al., 1997) and precipitation patterns (Sutton and Hodson,
2005).

To determine patterns of SST variability, scalar indices measuring average SST anomalies over one or more regions are
30   often regressed on the global SST field (Deser et al., 2010). Both the North Atlantic and North Pacific are known for their
low frequency SST variability and associated pattern. The Atlantic mode, indicated by the average North Atlantic SST, was
first named the Atlantic Multidecadal Oscillation by Kushnir (1994), but in the absence of a single spectral peak we use the
more modern naming convention of Atlantic Multidecadal Variability (AMV). Low frequency Pacific variability was revealed
through principle component analysis of North Pacific SSTs (Pacific Decadal Oscillation; Mantua et al. (1997)) and of the
35   global SSTs (Interdecadal Pacific Oscillation; Power et al. (1999)). The Pacific Decadal Oscillation is thought to arise as a
combination of low frequency tropical variability and local air-sea interaction (Newman et al., 2016). In the Southern Ocean,
SST observations are relatively sparse prior to satellite observations, but signatures of multidecadal variability are found, for
example, in the Antarctic sea ice extent (Simpkins et al., 2013). Southern Ocean variability is assessed here through an SST
index in the Atlantic sector of the Southern Ocean where eddy-mean flow interaction can result in multidecadal variability (Le
40   Bars et al., 2016; Hogg and Blundell, 2006; Jüling et al., 2018).

The importance of internal ocean variability in low-frequency climate variations has become more clear in recent years by
comparing ocean model results forced by either climatological or observed atmospheric forcing (Penduff et al., 2014). Mul-
tidecadal SST variability can arise from a range of processes (Dijkstra, 2013). Energy can be shifted to lower frequencies by
integration of high-frequency variability in one component by a slower component of a system, such as the coupled atmosphere-
45   ocean system (Hasselmann, 1976). It can also arise from internal variability in the atmosphere which is imprinted through heat
fluxes on the upper ocean (Eden and Jung, 2001). Internal multidecadal variability in the ocean can arise through specific insta-
bilities, such as the thermal Rossby wave mechanism (Te Raa and Dijkstra, 2002). Results of idealized models have indicated
that patterns of internal multidecadal SST variability may arise through noise excited (non)-normal modes (Frankcombe et al.,
2009; Weijer and van Sebille, 2014; Dijkstra, 2016) or due to a collective interaction of such modes (Berloff et al., 2007b;
50   Hogg and Blundell, 2006). Such collective interactions can change the vertical heat transport, affecting the stratification which
in turn influences the response of the mixed layer to heat fluxes and/or wind stress (Manucharyan et al., 2017) or impact deep
convection at multidecadal timescales (Dufour et al., 2017). Mesoscale eddies can also interact with the mean flow, for example
through rectification, which can lead to changes in momentum and vorticity input by wind stress (Hogg and Blundell, 2006;
Berloff et al., 2007a). Last, nonlinear advection of kinetic energy in the western boundary current separation region can lead to
55   an inverse temporal cascade in which spectral energy is shifted from high to low frequencies (Martin et al., 2019).

Models participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. (2012)) generally
underestimate multidecadal climate variability (Cheung et al., 2017; Mann et al., 2020). The amplitudes of Pacific and Atlantic
multidecadal variability, as measured by the AMV and PDO indices, are significantly lower than in observations. Further,

the regression patterns partially mismatch those in observations, in particular in the Pacific and the western boundary current

60 regions (Kajtar et al., 2019). Hence, CMIP5 models may miss crucial physical processes causing the patterns of multidecadal variability. The resolution of many, if not all, CMIP5 models is too coarse to capture all relevant internal variability. In at least one ocean model, multidecadal variability appears in a high resolution (strongly eddying) setup but is absent at a lower resolution (non-eddying) version (Le Bars et al., 2016).

Decreasing the horizontal ocean grid spacing from $1°$ in CMIP5 style GCMs to $0.1°$ enables new physics with the explicit

65 representation of mesoscale phenomena (Tulloch et al., 2011; Hallberg, 2013). Mesoscale variability arises due to barotropic and baroclinic instability and affects the ocean state in myriad ways (McWilliams, 2008), for example by changing the advection of tracers (in particular affecting the heat and salinity budgets), affecting (re-)stratification (Couvelard et al., 2015; Dufour et al., 2017), and by modifying the mean flow through rectification. Mesoscale eddies also form an integral part of the turbulent energy cascade connecting large scale potential and kinetic energy input to small-scale dissipation. Currents, such as narrow

70 western boundary currents, are also much better represented in high resolution simulations as are inter-ocean exchanges of water masses, most notably Agulhas rings (Biastoch et al., 2008).

In most CMIP5 model simulations, mesoscale eddy effects are parametrized with the isopycnal slope mixing parametrization of Gent and McWilliams (1990) (GM). While GM captures many effects of eddies on the circulation, it stabilizes the ocean state and suppresses low frequency variability (Hallberg and Gnanadesikan, 2006; Viebahn et al., 2018). Furthermore, GM does

75 not capture rectification effects of the eddies on the mean flow. In idealized non-eddying ocean models, modes of multidecadal variability exist that depend critically on the prescribed eddy diffusivity (Huck et al., 2014). Increasing the resolution to allow mesoscale eddies, Huck et al. (2014) find that the multidecadal variability persists despite changes in the mean circulation, suggesting that the eddy parametrization may suppress such low-frequency variability if tuned incorrectly. In a comparative study with a suite of coupled GFDL models at different resolutions, Griffies et al. (2015) find a stronger upward heat transport

80 by the mesoscale eddies which counteracts the general downward heat transport by the mean currents. This enhanced vertical heat transport leads to larger heat fluxes and faster adjustment to forcing (Delworth et al., 2012).

With the availability of multi-century simulations of global climate models with strongly eddying ocean components (Kirtman et al., 2012; van Westen and Dijkstra, 2017), with typical horizontal grid spacing of 10 km or less, it is timely to investigate how the mesoscale ocean flows affect the patterns of multidecadal variability. It is exactly what we do here in this paper, using

85 simulations with the Community Earth System Model. The paper is organized as follows: section 2 describes the simulations and the methods of analysis, section 3 presents the results, and section 4 discusses and summarizes the results and their implications.

2 Model Simulations and Data Analysis

We analyze results from two multi-century present-day control simulations with the Community Earth System Model version

90 1.0.4 (CESM, Hurrell et al. (2013)), carried out at the Academic Computing Center in Amsterdam (SURFsara), see e.g., van Westen and Dijkstra (2017). Both control simulations use constant year 2000 atmospheric greenhouse gas concentrations forc-

ing, notably $[CO_2] = 367\,\mathrm{ppm}$ and $[CH_4] = 1760\,\mathrm{ppb}$. The CESM components are CAM5 (Community Atmosphere Model), POP2 (Parallel Ocean Program), CICE (sea ice model), and CLM (Community Land Model) which are coupled by the CESM1 coupler. The high resolution (HIGH) simulation employs a $0.1°$ ocean horizontal grid spacing on a tripolar grid, while the low

95 resolution (LOW) simulation has a $1°$ ocean horizontal grid spacing with a displaced dipolar grid. The effect of subgrid-scale processes on tracer and momentum transport is captured with a biharmonic diffusion operator in the HIGH simulation and the GM parameterization in the LOW simulation (Gent and McWilliams, 1990). The HIGH simulation was initialized from a simulation of several decades provided by the National Center for Atmospheric Research, while the LOW simulation was initialized from a decally averaged ocean state at year 1000 of a CESM 1.1.2 simulation performed with the same resolution.

100 CESM 1.1.2 and CESM 1.0.4 exhibit only minor differences in their ocean state with CESM 1.0.4 having slightly higher SSTs (not shown). Table 1 summarizes the important simulation characteristics.

Table 1. Overview of the two CESM simulations with characteristics and names of their ocean and atmosphere grids.

name	ocean grid	atmosphere grid	years analyzed
HIGH	$0.1°$ tripole, 42 levels to 6000 m (`tx0.1v2`)	$0.47° \times 0.63°$ (`f05`)	51-300
LOW	$1°$ dipole, 60 levels to 5500 m (`gx1v6`)	$0.9° \times 1.25°$ (`f09`)	11-260

From earlier simulations with the same model components (CCSM3.5: Kirtman et al. (2012) and CESM1: Small et al. (2014)), it is known that the climatology of the higher resolution simulation improves compared to the lower resolution simulation in many aspects. The overall SST biases reduce due to a better representation of boundary currents, ocean upwelling and

105 air-sea interactions (Small et al., 2014). Some biases remain, though, such as too warm high latitude SSTs which result in low sea ice extent and sea ice volume biases in the high resolution simulation (Kirtman et al., 2012).

Figure 1 shows the global mean surface heat flux into the ocean indicating the equilibration of the simulations. The HIGH simulation equilibrates significantly faster than the LOW simulation. The GFDL CM2 climate models show a similarly decreased drift in their higher resolution setups which is due to an enhanced upward heat transport by the explicitly resolved

110 eddies (Delworth et al., 2012; Griffies et al., 2015). In our CESM simulations, all basins are initiated too cool and hence the temperature increases at all depths, except for the deep Pacific and some layers at intermediate depths in the Indian Ocean which cool (not shown). On interannual time scales, the Pacific surface heat fluxes vary strongest in comparison with those in the other basins. In particular, in the LOW simulation a too strong El Niño-Southern Oscillation (ENSO) signal dominates the global interannual surface heat flux signal (Fig. 1b; Wieners et al. (2019)).

115 The results in Fig. 1 indicate that a trend must be removed despite the constant forcing as the model simulations are still adjusting towards their statistical equilibrium. We remove a quadratic trend at each grid point of the relevant model output. This approach allows for different proximities to the statistical equilibrium and different time scales of adjustment at each grid cell. A further choice in the analysis is the selection of model years which are used in the analysis. Of the HIGH simulation we ignore the first 50 spinup years as strong adjustment is evident in the global mean SST approximately until year 40. As

120 mentioned, the LOW simulation is initialized at year 1000 of a CESM 1.1.2 simulation with the same forcing and grid that
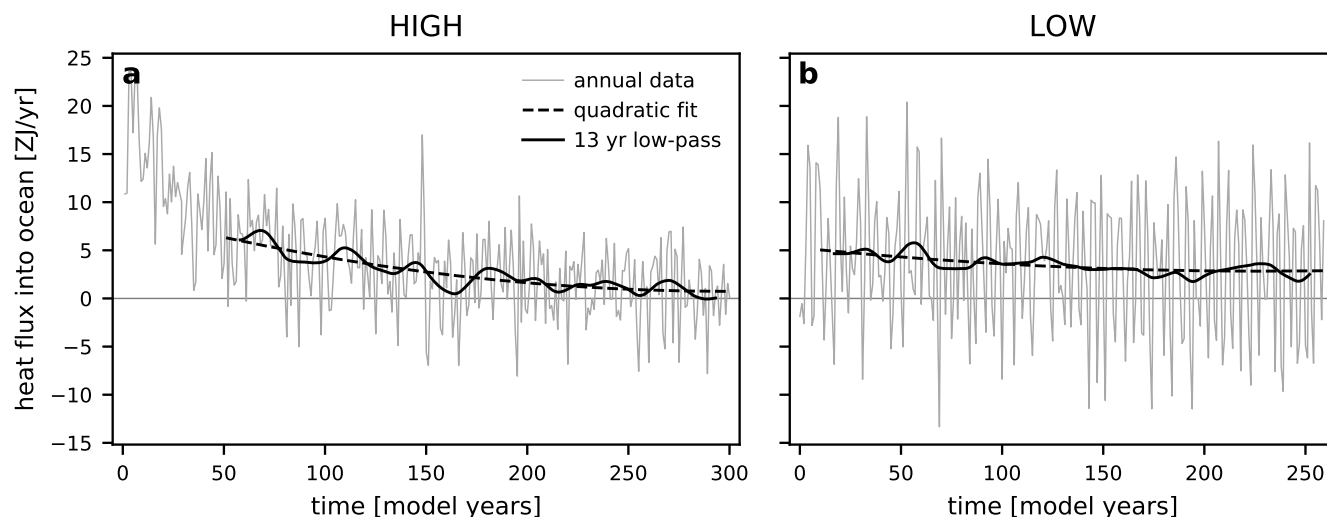
Figure 1. The globally averaged surface heat flux into the ocean of the HIGH (a) and LOW (b) simulations including the annual time series (thin), the quadratic fit to the 250 analyzed model years (dashed), and the 13 year low-pass filtered version (solid). For the low-pass filtered time series, 7 years are removed at each end to avoid filter edge effects.

maintains a near constant surface heat flux of $2.2 \, \mathrm{ZJ \, yr^{-1}}$ concurrent to the LOW simulation. We choose to discard the first 10 years of data to avoid fast adjustment processes and analyze years 11-260 (cf. Table 1). Monthly data is de-seasonalized by removing the mean annual cycle, i.e. the mean monthly difference to the annual mean.

To compare the SST model results with observations, we use the HadISST sea surface temperature dataset from 1870-2018 which is provided on a $1° \times 1°$ grid (Rayner et al., 2003). The observational dataset must be detrended with an appropriate estimate of the historical forced signal to allow for a fair comparison with the results of the model simulations, which are obtained under constant forcing. A simple linear detrending is unable to remove the non-linear historical forcing signal (Steinman et al., 2015). When an ensemble of model simulations is available, the forced signal can be approximated as the ensemble mean assuming that the internal variability of the individual ensemble members is uncorrelated. One can use either the mean of a single model ensemble (such as that of the Max-Planck Institute Grand Ensemble (Maher et al., 2019) or the CESM Large Ensemble (Kay et al., 2015)) or a multi-model ensemble (multi-model mean: MMM; such as that of the CMIP5 ensemble). The MMM is generally superior to the single model ensemble mean in the historical period (Frankcombe et al., 2018), so we use CMIP5 ensemble results. The forced signal can furthermore be estimated as a single time series (single factor detrending; used for example by Steinman et al. (2015)) or as a linear combination of different forcing signals (multi-factor detrending as described below).

For the HadISST dataset, we choose to separate the natural and anthropogenic forcing signals which was found to be superior to single-factor detrending (Frankcombe et al., 2018). We use the CMIP5 MMM forced with observed aerosol and greenhouse gas concentrations until 2005 and those from the RCP8.5 scenario afterwards until 2018 (like, e.g. Kajtar et al. (2019)). The

natural forcing is assumed to equal the CMIP5 MMM of the historical (natural forcing only) simulations. The anthropogenic

140 contribution ($\mathrm{MMM_{ant}}$) due to greenhouse gases and aerosols is calculated as the difference between the fully forced MMM ($\mathrm{MMM_{all}}$) and the MMM with only natural forcing ($\mathrm{MMM_{nat}}$), i.e.,

$$\mathrm{MMM_{ant} = MMM_{all} - MMM_{nat}} \tag{1}$$

At each grid point, scaled versions of these two time series are subtracted from the historical reconstructed SST, such that the detrended SST is

145 $$\mathrm{SST_{detr}}(x,y,t) = \mathrm{SST}(x,y,t) - \alpha_1(x,y)\mathrm{MMM_{nat}}(t) - \alpha_2(x,y)\mathrm{MMM_{ant}}(t) - \alpha_3(x,y) \tag{2}$$

where the coefficients α_i at each location are determined through multiple linear regression. Stolpe et al. (2017) show for the AMV index and the index of Pacific multidecadal variability by Henley et al. (2015) (which does not equal the PDO index, but shares characteristics) that the various detrending methods result in similar behavior.

To assess the spatial expression of the modes of multidecadal variability, we use regression plots (Deser et al., 2010). The

150 plotted regression values R are defined as the covariance of the SST field and the normalized index $X(t)$, where $X = $ AMV, PDO, or SOM, according to

$$\mathrm{R}(x,y) = \frac{\mathrm{cov}(\mathrm{SST}(x,y,t), X(t))}{\mathrm{std}(X(t))}, \tag{3}$$

such that R has units of temperature. The significance against a no-correlation null hypothesis is tested with a two-tailed Student's t-test. Because the time series are autocorrelated and filtered, the effective number of data points n' is lower than

155 the original sample size n and can be estimated as the maximum of the reduced sample size due to filtering and due to autocorrelation (Trenberth, 1984):

$$n' = n \times \max\left(f\Delta t, \frac{1 - r_{1,X}r_{1,Y}}{1 + r_{1,X}r_{1,Y}}\right), \tag{4}$$

where $r_{1,X}$ ($r_{1,Y}$) is the lag-1 autocorrelation of time series X (Y), f the filtering frequency, and Δt the time step.

We analyze time series in the spectral domain with multi-taper spectra (Ghil et al., 2002). This estimator of spectral density

160 is superior to the classic periodogram in that it reduces spectral leakage and is statistically robust, i.e. the estimated noise reduces with more data points. However, as a trade-off the effective spectral resolution is reduced which becomes problematic at periodicities near the length of the time series. We therefore limit our analysis to periodicities below 50 years when focussing on model-observation comparisons but extend the range to 100 year periods for comparisons between the simulations. We use a bandwidth parameter of 2 and the number of tapers is 3. To test significance against a red noise null hypothesis (Hasselmann,

165 1976), we generate 10,000 Monte-Carlo first order autoregressive (AR(1)) processes. The autocorrelation coefficient and noise amplitude are estimated with the maximum likelihood estimator via Kalman filtering from the deseasonalized and detrended but unfiltered monthly time series Durbin and Koopman (2012).

3 Results

The results are divided into four subsections: the first describes the variability in SST by means of the chosen indices, the
170 second focuses on the surface heat fluxes (SHF), the third explores the spatial structure of multidecadal variability of the Ocean
Heat Content (OHC), and the fourth shows the consequences for the Global Mean Surface Temperature (GMST).

3.1 Sea Surface Temperature
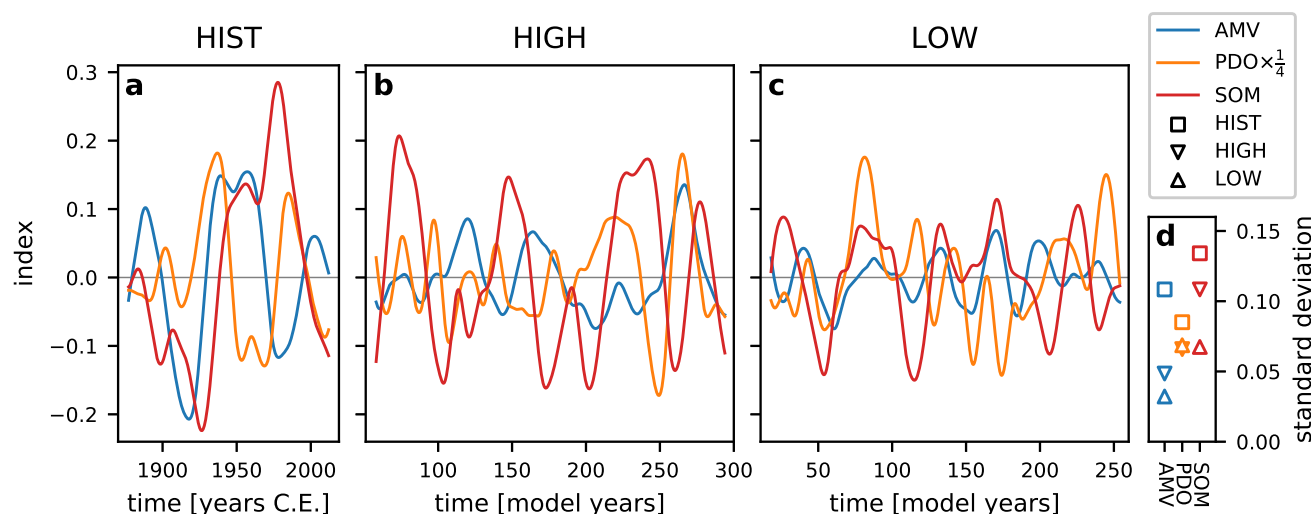
Figure 2. The three 13 year low-pass filtered indices of multidecadal variability for the two-factor detrended HadISST data (a) and the
quadratically detrended HIGH (b) and LOW (c) simulations. Atlantic Multidecadal Variability (blue) and Southern Ocean Mode (red) indices
are in units of Kelvin. The monthly time series of the Pacific Decadal Oscillation index is dimensionless with unit standard deviation is also
low-pass filtered and has been divided here by 4 for a comparable amplitude to the other indices (orange). At each end of the time series 7
years are removed to avoid filter edge effects. Panel (d) shows the standard deviation of the low-pass filtered time series.

From the deseasonalized and detrended model and observational SST fields, we determine area averages and principle
components as indices of modes of multidecadal variability. The AMV index is calculated here over the domain $[0°\mathrm{N}, 60°\mathrm{N}] \times$
175 $[80°\mathrm{W}, 0°\mathrm{E}]$, very similar to that in Stolpe et al. (2017) who use the zonal extent $[75°\mathrm{W}, 5°\mathrm{W}]$. The PDO is captured by the
first principal component of the Pacific SSTs north of $20°\mathrm{N}$ as originally proposed by Mantua et al. (1997). The Southern
Ocean Mode (SOM) index, proposed by Le Bars et al. (2016) as a signature of the multidecadal variability in the Atlantic
sector of the Southern Ocean, is computed over the region $[50°\mathrm{S}, 35°\mathrm{S}] \times [50°\mathrm{W}, 0°\mathrm{E}]$. The index regions are outlined in the
respective regression maps of Fig. 3. The indices are defined as the 13 year Butterworth low-pass filtered monthly time series
180 with the first and last 7 years removed to avoid filter edge effects. Figure 2 shows the AMV, PDO, and SOM indices which all
display variability at multidecadal time scales (Fig. 2a-c). The AMV and SOM indices (in units of Kelvin) exhibit a smaller
amplitude in the simulations than in the historical data (Fig. 2d). Both amplitudes are higher in the HIGH simulation than in

the LOW simulation. Larger PDO amplitudes after lowpass filtering of the underlying monthly, unit standard deviation time series indicate larger spectral power at frequencies below the filter cutoff frequency.
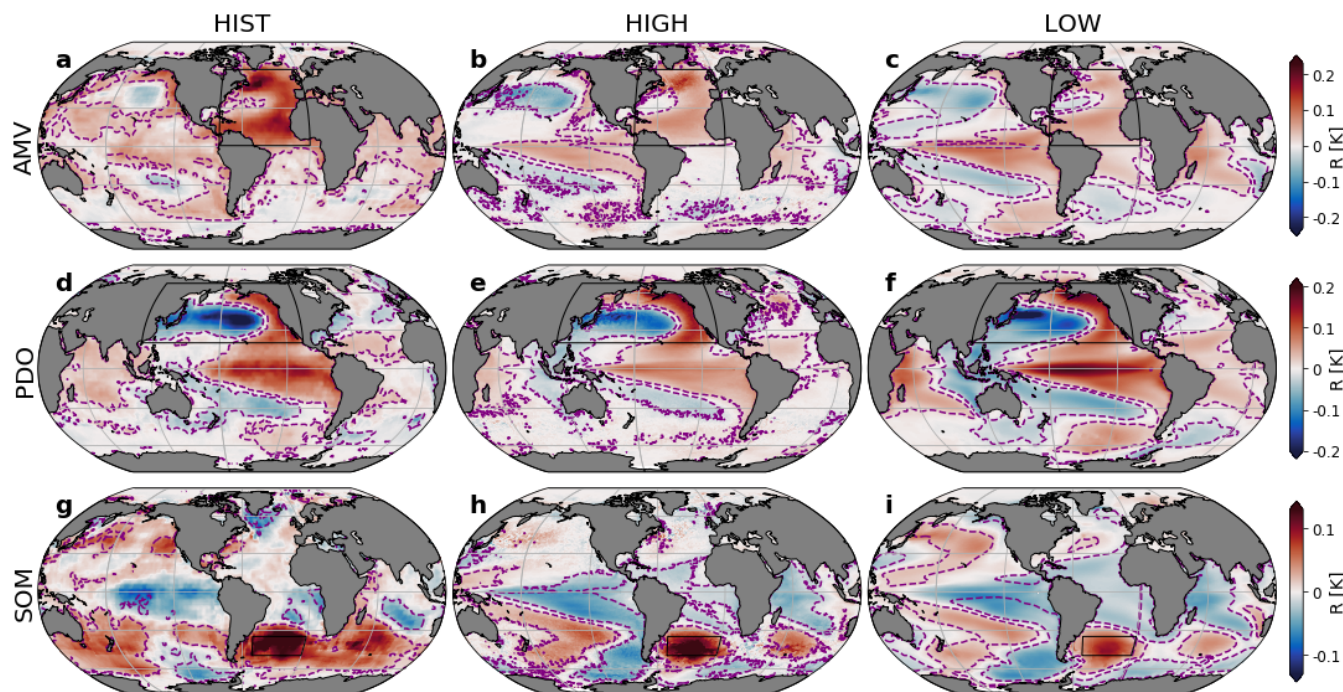
Figure 3. Regression maps (values of $R(x, y)$ as in eqn. 3 in Kelvin) of the detrended SST fields onto the AMV (a-c), PDO (d-f), and SOM (g-i) indices for the 149 year detrended HadISST historical dataset (left) and the 250 year data sets of HIGH (center) and LOW (right) simulations. Purple dashed lines demarcate the areas of significant correlations at the 98% level. Black boxes outline the index areas. Parallels and meridians are drawn every 30° and 60°, respectively. Note the different colorbar ranges.

185   Figure 3 shows the regression patterns of the detrended SST on the AMV, PDO, and SOM indices for the detrended historical data (HIST) and the two simulations (HIGH and LOW). For the AMV index, the regression pattern of the detrended historical SST data shows a horse-shoe shape with significant positive regression values throughout the North Atlantic, with a maximum in the subpolar gyre and a secondary maximum in the subtropical gyre. The regression patterns of both HIGH and LOW simulations show a comparable pattern to the historical data, albeit with lower regression values. Like the observations,

190   the HIGH simulation exhibits the strongest regression values in the subpolar gyre, while the LOW simulation shows more pronounced regression values in the subtropical gyre. The LOW simulation exhibits significant negative correlations inside the horseshoe pattern that are not present in the historical observations or the HIGH simulation. In the historical data and the HIGH simulation regression values are smaller in magnitude outside of the North Atlantic, but the LOW simulation's tropical Pacific exhibits positive correlation values exceeding those in the North Atlantic. Both simulations show a significant ENSO/PDO-

195   like regression pattern in the Pacific and dipole structure in the Indian Ocean, whereas the historical data shows only largely non-significant positive correlations in the tropical Pacific and a positive correlations throughout the northern Indian Ocean.

Overall, the areas of significant correlations are larger in the LOW simulation than in both the HIGH simulation and the historical data. This suggests possible teleconnections between the Indian and Pacific basins and the Atlantic basin at multidecadal time scales especially in the LOW simulation, but such correlations are not significant in observations (Steinman et al., 2015).

200   The PDO regression pattern is characterized by negative values in the Kuroshio extension and positive values in an enclosing horseshoe pattern to the East. In the observations, the minimum is located east of the dateline and significant positive values exist in the Eastern tropical Pacific. Weaker negative (positive) regression values exists in the southwest Pacific (the Bellingshausen Sea). The simulations show largely similar patterns, although the minima in the Kuroshio extension area are shifted west of the dateline, especially so in the LOW simulation. The historical data shows uniformly positive correlations in the

205   Indian Ocean where both simulations display a dipole pattern which is more pronounced in the LOW simulation. All PDO patterns exhibit positive values in the tropical Pacific, but the regression values of the LOW simulation are closer to those of the historical data. In the LOW simulation a large area of significant correlations in the tropical Atlantic exists that is absent in both the observations and HIGH simulation.

The SOM index regression pattern shows a coherent pattern spanning much of the Southern Ocean at significant levels in

210   the historical data with maximum values in the SOM region and a secondary maximum north of the Kerguelen Plateau in the West Indian Ocean sector of the Southern Ocean. Both simulations exhibit these two positive value areas and whereas the HIGH simulation's correlation values are closer to those of the historical SST data, both feature a more azonal pattern with negative values in the Eastern Weddell Gyre. All patterns exhibit negative regression values in the tropical Pacific, but only the simulations feature this structure at significant levels and then as part of a larger South Pacific horseshoe shape pattern

215   that extends south to the Bellingshausen Sea; there, slightly negative values occur in the observations as well. The LOW simulation's area of significant correlations extends further into the Northern Hemisphere, especially in the Pacific.

Figure 4 shows the spectral power of the unfiltered monthly index time series of AMV, PDO, and SOM. Again, the historical indices cover 149 years while the simulated ones cover 250 years, resulting in different spectral resolutions. To allow a comparison between the results, also the period of variability of the historical data has been extended to 50 years, although the higher

220   period results are less reliable than in the models, because of the difference in data series length. In the detrended observations, significant (99%) spectral power for all three indices occurs at large periodicities above 40 years. The PDO signal also extends beyond the red noise (95%) at periods of about 20 years. Most previously published spectral analysis of the North Atlantic observed temperatures find significant periodicities around 50-70 years, but they overwhelmingly remove a linear trend only which distorts the signal (Steinman et al., 2015). Longer proxy data records reveal significant multidecadal variability around

225   the North Atlantic from sources including Greenland ice cores (Chylek et al., 2011), tree rings (Gray et al., 2004), sediments (Knudsen et al., 2011), or a combination of proxies (Delworth and Mann, 2000; Wang et al., 2017). In the HIGH simulation, significant spectral power peaks only exist at 40-50 years for the AMV index (>95%), around 13-15 and 20-25 years for the PDO index (>99%), while the SOM index shows significant variability around 15 years (>95%) and above 40 years (>95%). The indices of the LOW simulation exhibit no significant spectral power for either the AMV or SOM indices and the red noise

230   null hypothesis cannot be rejected at 95% level, and only the PDO index has a significant spectral peak (95%) at 16 years.
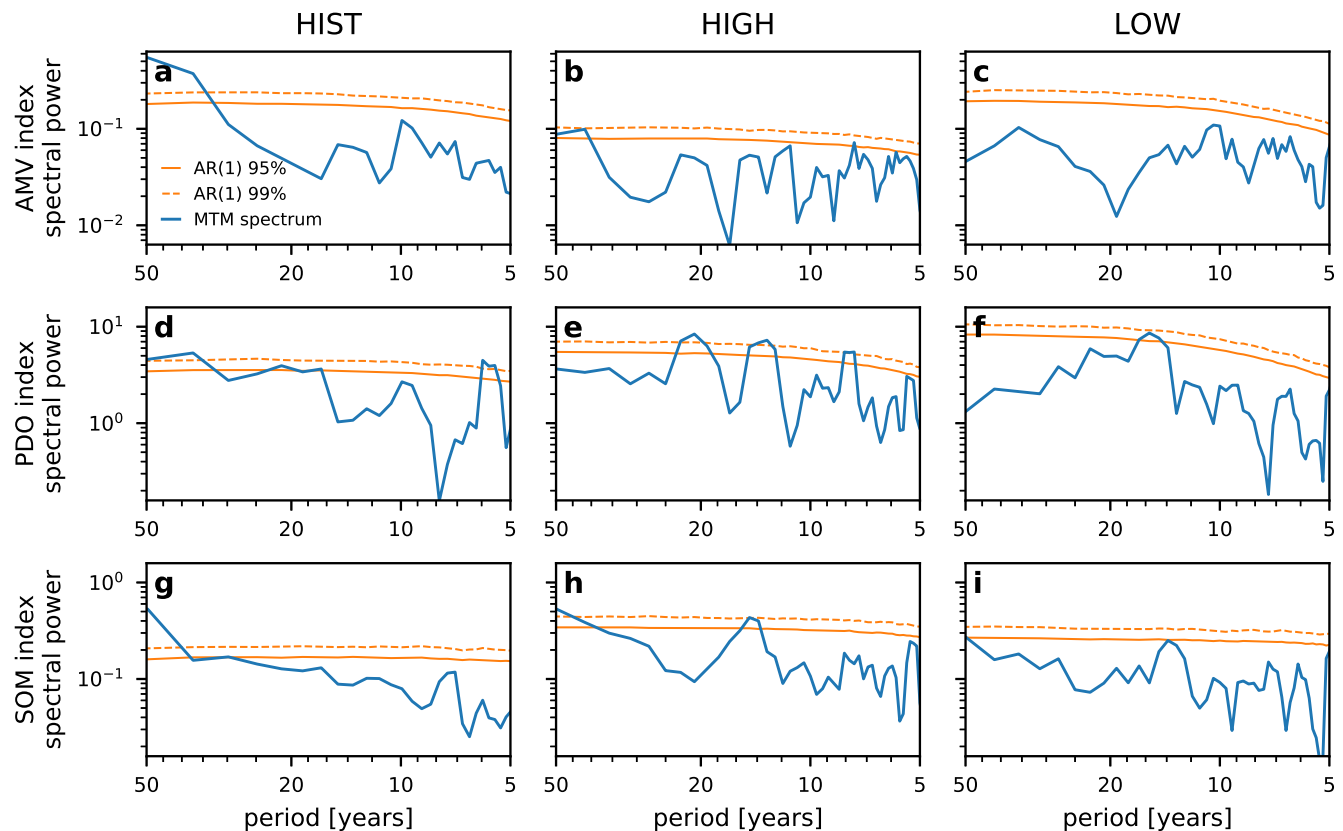
Figure 4. Multi-taper spectral estimates of the unfiltered monthly SST indices (blue) underlying the filtered time series of Fig. 2. Panels (a-c) show the AMV, (d-f) the PDO, and (g-i) the SOM index spectra for the two-factor detrended 149 year HadISST dataset (left) and the 250 year HIGH and LOW simulations (center and right, respectively). The units of the spectral power are $[\mathrm{K}^2\,\mathrm{yr}]$ for the SST average AMV and SOM and $[\mathrm{yr}]$ for PDO index which is dimensionless as a principal component. As a null hypothesis 10,000 AR(1) processes were simulated to estimate the 95% and 99% red noise confidence interval (solid and dashed orange lines).

3.2 Surface Heat Fluxes

Figure 5 shows the time series and spectra of the surface heat flux (SHF) into the global ocean (black) as well as into the major ocean basins (Atlantic: blue, Pacific: orange, Southern Ocean: red). We define the Southern Ocean as south of the parallel at Cape Agulhas at $34°$S, the Atlantic Ocean is bounded in the north by the Labrador Sea and extends to Iceland, while the Pacific

235   extends up to the Bering Strait (inset map in Fig. 5b). For global as well as Atlantic and Pacific heat fluxes, the LOW spectra show enhanced power around 4 and 8 years which is due to a strong and regular ENSO in this simulation. At multidecadal time scales, the HIGH simulation clearly exhibits larger surface heat flux variability than the LOW simulation in particular in the Southern Ocean and, as a result, globally. The Atlantic and Pacific SHF spectral comparison is more subtle, as there are periods between 25-35 years where the LOW simulation shows higher variability than the HIGH simulation. However, the
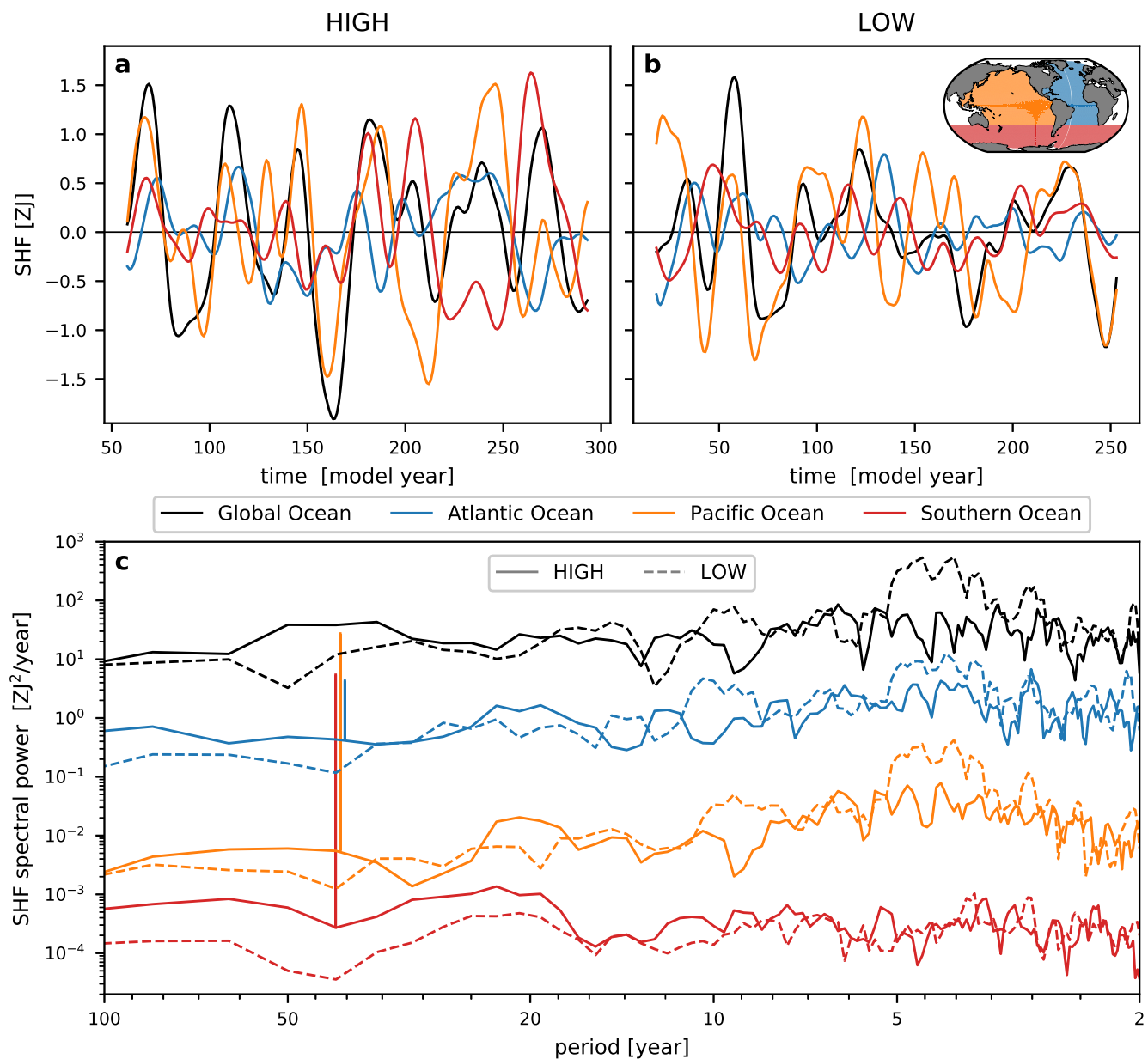
Figure 5. Time series for the 13 year low-pass filtered surface heat fluxes (SHF) into the global (black; equals thick solid line of Fig. 1), Atlantic (blue), Pacific (orange), and Southern (red) oceans for the (a) HIGH and (b) LOW simulations. The inset global map defines the three major ocean basins. Panel (c) shows the multi-taper spectral estimates of the unfiltered, but quadratially detrended time series. For visual clarity, the spectra were separated vertically by multiplying the basin flux spectra by constant factors, these shifts are indicated by vertical lines.

240   integrated Atlantic and Pacific multidecadal spectral power is higher in the HIGH simulation compared to the LOW simulation.
      Remarkable is further that the Southern Ocean SHF spectrum is nearly white with similar power at all frequencies, while both
      the Atlantic and Pacific show decreasing power at lower frequencies.

3.3   Ocean Heat Content

      The basin-scale ocean heat content (OHC) changes are largely determined by surface heat fluxes while horizontal heat di-
245   vergences play a minor role. The strong multidecadal surface heat flux variability motivates the investigation of the structure
      of OHC anomalies in this section. Long reconstructions of the OHC exist for the industrial period (1870-2015; Zanna et al.
      (2019)) and even the Common Era (15-2015; Gebbie and Huybers (2019)), but they lack the detail that we aim to investigate
      here so that we compare model data only.

      The HIGH simulation temperature data has been interpolated to a $0.4°$ rectangular grid and the OHC calculations were
250   performed on that grid. The OHC of the LOW simulation is calculated on the original displaced dipole grid and zonal integrals
      are performed along the grid x-direction (so not along parallels in the high northern latitudes). The first two columns of Figure 6
      are Hovmöller diagrams of 13 year low-pass filtered zonally and vertically integrated OHC anomalies, indicated by OHC_v,
      i.e.,

$$\mathrm{OHC}_v(y,t) = c_p \int \int \rho\theta(x,y,z,t) \ \mathrm{d}x\,\mathrm{d}z \ , \tag{5}$$

255   where θ is the potential temperature, ρ the density, and $c_p = 3996\,\mathrm{J\,kg^{-1}\,K^{-1}}$ the sea water heat capacity used in the CESM.
      The last column in Figure 6 shows the standard deviation of the 13 year low-pass filtered anomalies as a function of latitude.
      Almost everywhere the variability as measured by this standard deviation is higher in HIGH compared to LOW. The first row
      shows the global signal which also captures the Southern Ocean signal south of $34°$S (marked by the green line). Here signs of
      the SOM in the HIGH simulation can be seen with north- and southward propagating anomalies; these are absent in the LOW
260   simulation. The Atlantic behavior is qualitatively similar between the HIGH and LOW simulations with southward propagating
      anomalies between $40°$N and $10°$N, although the anomaly amplitude is larger in the HIGH simulation. Both simulations further
      show meridionally coherent anomalies south of $10°$N in the Atlantic. On the other hand, in the Pacific remarkable differences
      exist: only in the HIGH simulation OHC anomaly signals propagate equatorward around $30°$N, imprinting on the global pattern.
      This is also visible to a lesser extent in the South Pacific just north of $30°$S. In the equatorial Pacific, the unrealistically strong
265   ENSO signal of the LOW simulation is filtered out by the 13 year low-pass filter. However, compared to the HIGH OHC_v,
      the difference is reduced between the tropical and extratropical low frequency spectral power as revealed by the standard
      deviations. The Atlantic and Pacific peaks at northern midlatitudes are shifted equatorward in the HIGH simulation compared
      to the LOW simulation due to a better representation of the western boundary current separation. In the Indian Ocean, the
      HIGH simulation shows a less meridionally coherent pattern compared to the LOW simulation.

270   Figure 7 shows 13 year low-pass filtered Hovmöller diagrams of the horizontally integrated OHC anomalies, given by

$$\mathrm{OHC}_h(z,t) = c_p \int \int \rho\theta(x,y,z,t) \ \mathrm{d}x\,\mathrm{d}y, \tag{6}$$
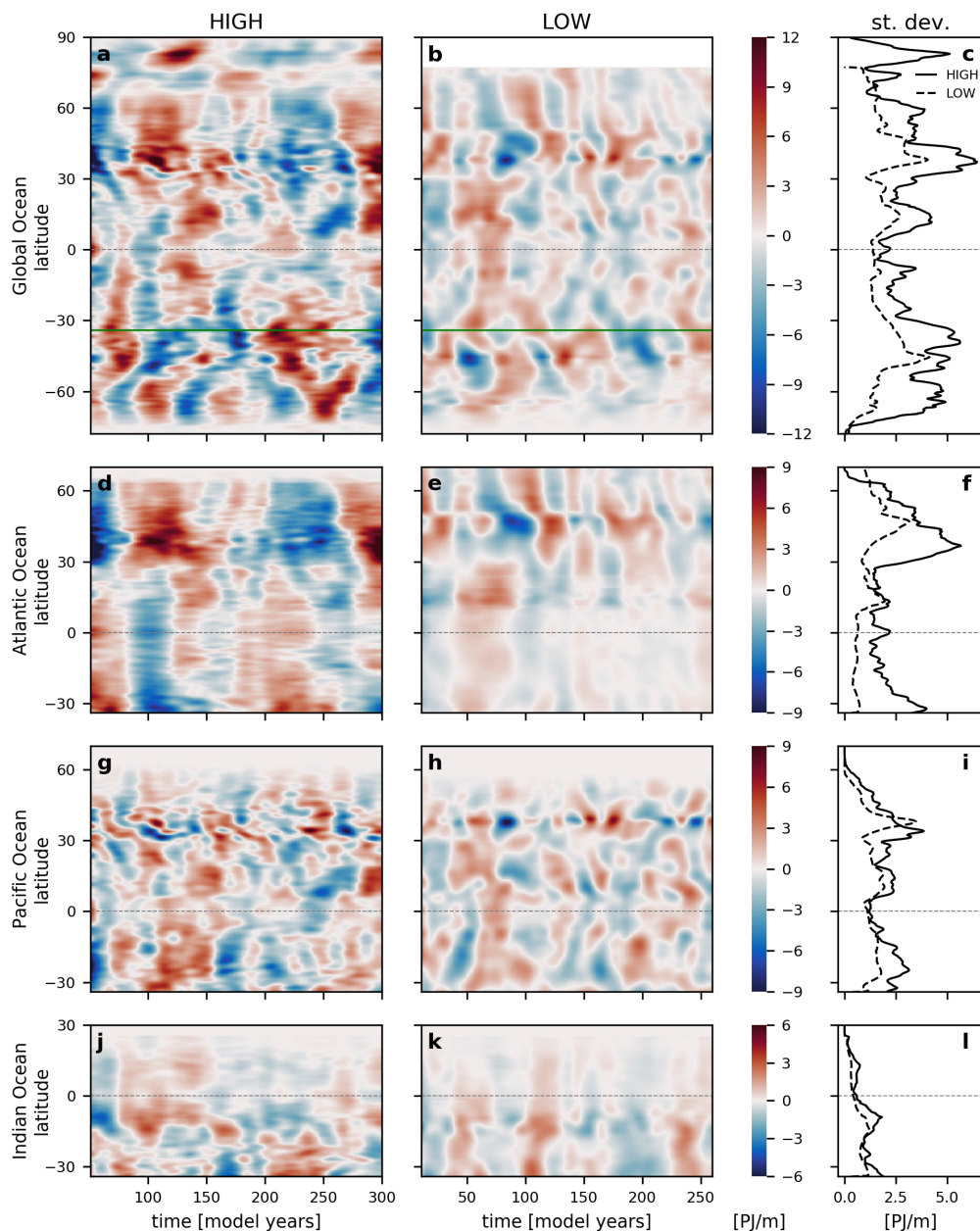
Figure 6. The 13 year low-pass filtered zonally and depth integrated OHC anomalies ($\mathrm{OHC}_v(t,y)$ of Eq. (5)) as Hovmöller diagrams of the HIGH (left) and LOW simulations (center), respectively. The top row shows the globally integrated OHC anomaly (which includes the Southern Ocean signal south of $34°\mathrm{S}$, demarcated with the green line in the first row), while the lower rows show the individual ocean basins. The equator is marked with a thin dashed line. The right column shows the standard deviation in time of the zonally and depth integrated OHC (solid: HIGH, dashed: LOW). The LOW latitudes values are averaged along the grid x-coordinate, such that north of $60°\mathrm{N}$ they are not exactly representing the true latitudes. Note the different color scale ranges of each row with units of $\mathrm{PJ/m} = 10^{15}\,\mathrm{J/m}$.
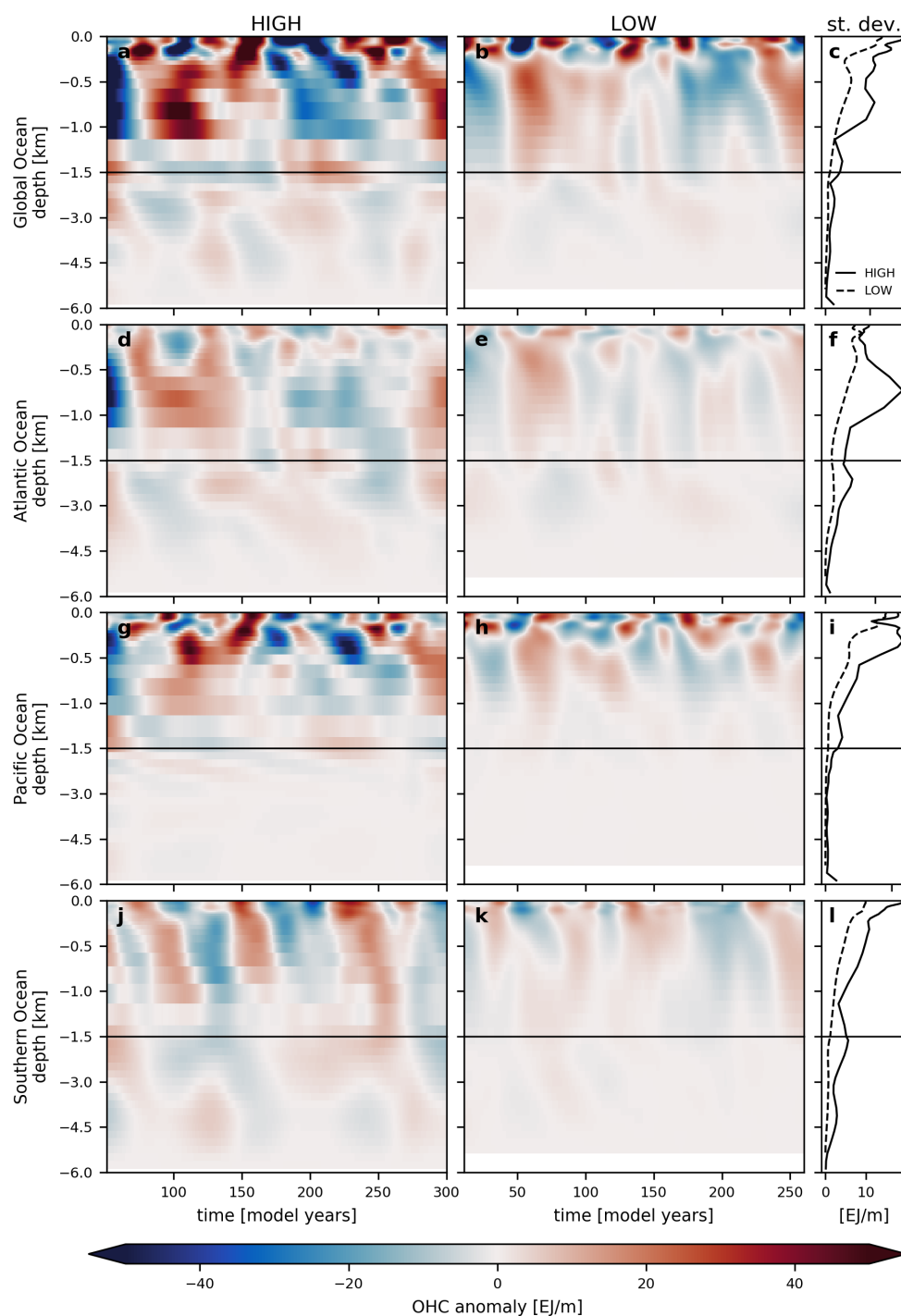
Figure 7. The 13 year low-pass filtered horizontally integrated OHC anomalies ($\mathrm{OHC}_h(t,y)$ of Eq. (6)) as Hovmöller diagrams of the HIGH (left) and LOW (center) simulations, respectively. The different rows show the global, Atlantic, Pacific, and Southern Ocean integrals over time. Note the change of depth scale at 1500 m. The right column shows the standard deviation in time of horizontally integrated OHC (solid: HIGH, dashed: LOW), calculated similarly as in Fig. 6. Units are in $\mathrm{EJ/m} = 10^{18}\,\mathrm{J/m}$.

where the last column shows the temporal standard deviation at each depth. Globally and in the individual ocean basins, the HIGH simulation exhibits stronger and deeper OHC variability than the LOW simulation. The global OHC_h anomalies on multidecadal timescales are dominated by those in the upper 1200 m (Fig. 7c). The HIGH global multidecadal signal below
275 the mixed layer is made up in approximately equal parts of the Atlantic, Pacific, and Southern Ocean components, as they exhibit similar magnitudes in their OHC_h standard deviations. In the LOW simulation, the global anomalies are dominated by those in the Pacific. The Atlantic HIGH OHC_h anomaly standard deviation shows a pronounced peak between 400 and 1500 m depth. In the Pacific, the variability is larger between 200 m and 1500 m in the HIGH simulation. In the upper 1000 m, the anomalies propagate quickly downwards in the Atlantic but slower in the Pacific. At Pacific depths below 1000 m, a very slow
280 downward propagation of anomalies is visible in the HIGH simulation that is absent in the LOW simulation. The faster vertical propagation of heat anomaly signals in the Atlantic compared to the Pacific can be explained by the presence of the North Atlantic downward branch of the meridional overturning circulation (Buckley and Marshall, 2016). In the Southern Ocean large differences are evident: the HIGH multidecadal variability is much stronger and extends to the full depth as opposed to the LOW variability that is weaker at all depths, in correspondence with results in Le Bars et al. (2016). The SOM index
285 (Fig. 2) correlates very well with Southern Ocean surface OHC_h signal.

3.4 Global Mean Surface Temperature

The multidecadal global mean surface temperature (GMST) evolution is a consequence of the heat flux convergence in the atmosphere as energy is exchanged through the sea surface with the oceans and through the top of the atmosphere with outer space. Figure 8 shows the time series and spectral estimates of the GMST. At periods around 4 years, the too strong ENSO of the
290 LOW simulation (as extensively analysed in Wieners et al. (2019)), leads to a larger interannual variability in the GMST time series and manifests iteself as a peak of the GMST spectral signal. At multidecadal time scales (beyond periods of 13 years) the integrated spectral power in the HIGH simulation is higher than in the LOW simulation, but for periodicities at 35-60 years the LOW GMST spectral power is higher than that of the HIGH simulation. The spectral power of the two-factor detrended historical GMST data lies between both simulations at sub-decadal time scales, agrees well with both simulations at (inter-)
295 decadal time scales, and exceeds both simulations at periods above 30 years.

4 Summary and Discussion

We investigated the effect of ocean model resolution on multidecadal variability by contrasting two multi-century simulations with the Community Earth System Model (CESM), one with a non-eddying ocean typical of the CMIP5 models (LOW) and one with a strongly eddying ocean (HIGH). The enhanced horizontal ocean model resolution allows for more detailed features of
300 the circulation with many documented improvements, such as better representation of boundary currents, reduced SST biases, air-sea exchange, and internal variability of both high and low frequency relative to the life times of mesoscale ocean eddies. In particular, we focused on three modes of multidecadal SST variability: the Atlantic Multidecadal Variability (AMV), the Pacific Decadal Oscillation (PDO), and the Southern Ocean Mode (SOM). We find that these modes of multidecadal variability are
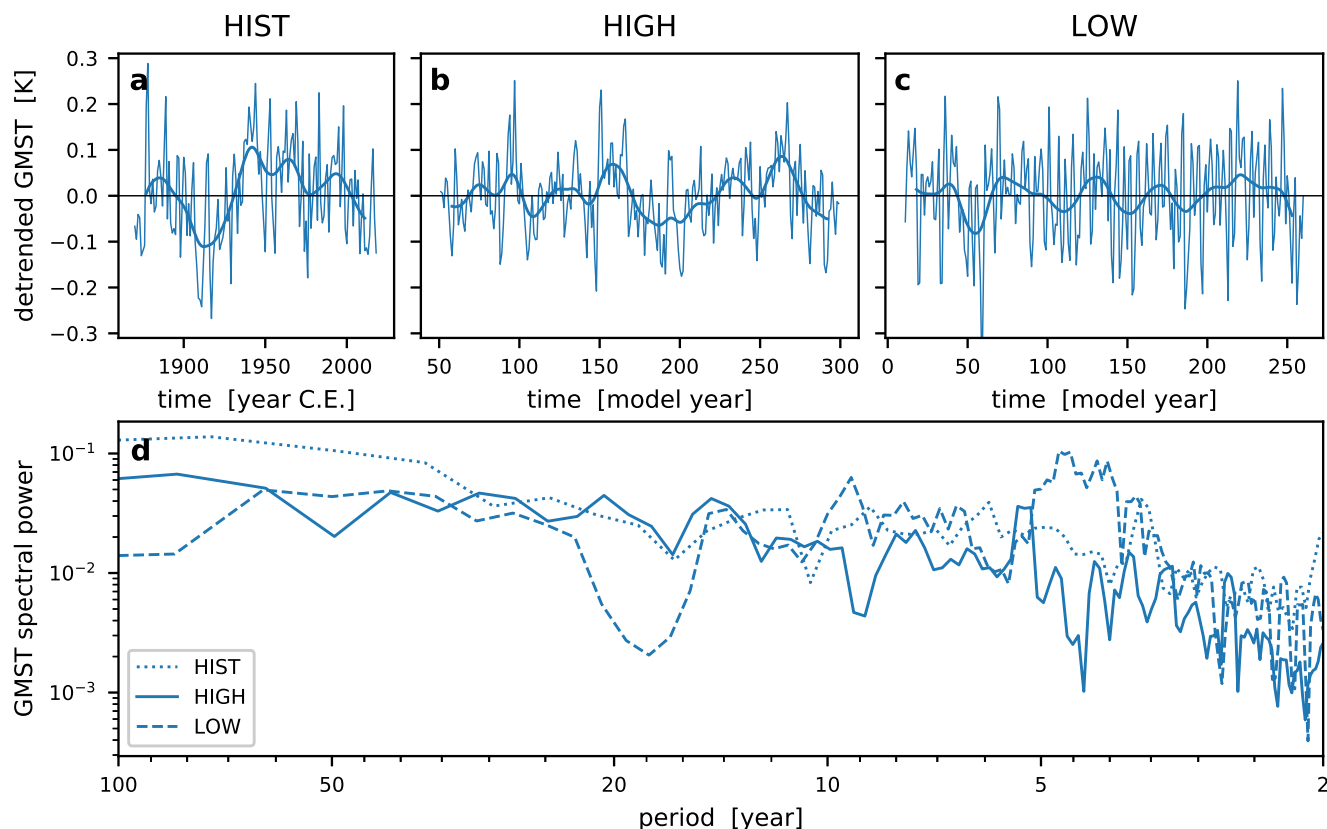
Figure 8. Time series and multi-taper spectra of the annual global mean surface temperature in units of $[K^2\,yr]$. The historical GMST estimate is derived from the two-factor detrended time series as discussed in section 2, while model (HIGH and LOW) spectral estimates are based on the quadratically detrended time series.

more pronounced in the HIGH simulation than in the LOW simulation (Fig. 2) and compare more favourably to observations
305 both in regression patterns (Fig. 3) and spectral properties (Fig. 4). At multidecadal time scales, stronger surface heat fluxes (globally and in all major ocean basins) are associated with this larger SST variability (Fig. 5). The integrated ocean heat content (OHC) signal varies more strongly in the HIGH simulation than in the LOW simulations and anomalies extend to greater depths (Figs. 6, 7). However, while the integrated spectral power of the global mean surface temperature (GMST) at multidecadal tie scales is larger in the HIGH simulation than the LOW simulation, there are frequencies at which the LOW
310 GMST variability is larger (Fig. 8). From the analyzed LOW and HIGH CESM simulations, we conclude that representing mesoscale eddies in the CESM leads to an enhanced intensity of multidecadal variability in the ocean.

In the Atlantic, an improvement in the simulation of the AMV pattern is evident in the HIGH simulation compared to the LOW simulation. While the overall AMV pattern is similar in observations and model simulations, the subpolar maximum in the regression patterns in the HIGH simulation matches the detrended observations in contrast to the subtropical maximum in

315   the LOW resolution. Due to the specifics of our detrending approach, the historical regression pattern looks somewhat different

from Deser et al. (2010), especially outside the North Atlantic. In particular, the pattern shows a statistically significant positive

correlation in the northwestern tropical Pacific that is not evident in the regression pattern of Deser et al. (2010). The AMV

spectra also show that the HIGH simulation performs better compared with the historical SST data by showing significant

power at multidecadal periodicities against the red noise null hypothesis, while the spectral estimate of the LOW AMV signal

320   fails to reject this null hypothesis (Fig. 4). While the magnitude of OHC variability is larger in the HIGH simulation, the

meridional and depth structures of the Atlantic OHC variability are similar between the HIGH and LOW simulations (Figs. 6

7). This suggests that the AMOC effects on the OHC variability are captured with the coarse resolution in agreement with earlier

studies (Delworth et al., 1993; Delworth and Mann, 2000) providing support for physical mechanisms of the AMV deduced

from idealized models which are independent of mesoscale variability (Te Raa and Dijkstra, 2002). Increasing the resolution

325   can still improve the AMV relative to observations by reducing ocean mean state biases, in particular the representation of the

Gulf Stream and of deep water formation.

      We use the PDO index to capture Pacific low frequency variability and the resulting regression patterns are similar between

observations and simulations. The spectral estimates of both the historical PDO and simulated ones extend beyond the 95%

confidence interval interdecadal time scales around 20 year periodicity. The LOW simulation exhibits a strong and regular

330   ENSO signal (Wieners et al. (2019); visible e.g. in Fig. 5c) and the low frequency tropical Pacific SHF spectral power is rela-

tively larger than the extratropical spectral power compared to the HIGH simulation (Fig. 5i). That the PDO partly comprises

a low-frequency ENSO signal is visible in the strong tropical Pacific regression maximum on the LOW simulation. With the

better representation of the western boundary currents, in particular the Kuroshio, in the HIGH simulation, there are qualitative

differences in the meridional structure of the Pacific OHC anomaly propagation: around $30°$N, and to a lesser degree around

335    $30°$S, equatorward propagation is evident in the HIGH simulation while meridionally coherent anomalies appear in the LOW

simulation (Fig. 6).

      The representation of the Antarcic Circumpolar Current in the Southern Ocean changes dramatically between the HIGH

and LOW simulations due to the presence of mesoscale eddies. The multidecadal SHF variability is much enhanced in the

strongly eddying HIGH simulation compared to the LOW simulation. Ocean heat anomalies also penetrate much deeper into

340   the Southern Ocean, suggesting a crucial difference between explicitly simulated mesoscale eddies and their parametrization

in advecting heat downward (Fig. 7). In the HIGH simulation, the SOM is visible in the meridional structure with north- and

southward moving anomalies, while anomalies simply converge on $45°$S in the LOW simulation (Fig. 6). The presence of the

SOM is expected in the HIGH simulation as it is thought to be caused by the interaction of mesoscale eddies and the mean

flow (Hogg and Blundell, 2006; Le Bars et al., 2016; Jüling et al., 2018). However, the SOM regression patterns of the HIGH

345   and LOW simulations are more similar to one another than to the observations with a more azonal, circumpolar wavenumber-3

pattern and stronger correlations extending into the South Pacific. However, the pattern in the HadISST dataset is likely to be

biased due to the limited number of observations in this region, in particular prior to satellite observations.

      Multiple mechanisms through which mesoscale eddies can generate multidecadal variability have been alluded to in the

introduction, such as eddy-mean flow interactions (Hogg and Blundell, 2006; Berloff et al., 2007a), the changes to the stratifi-

350 cation introducing memory (Manucharyan et al., 2017) or the modification of conditions for convection (Dufour et al., 2017). Furthermore, only eddying models are capable of simulating observed high-frequency variability correctly (Penduff et al., 2011), where an 'inverse temporal cascade' is important for the shift of variance from high to low frequencies. Any mechanism, involving eddies or not, leading to such a cascade thus requires a source of spectral energy at high frequencies which can be provided by the eddies. It is beyond the scope of this study to analyze the exact mechanisms at play in the HIGH simulation,

355 but it will be the subject of further study.

Naturally, the comparison between observations and model results is fraught with challenges, not least of which is the choice of an appropriate detrending procedure. The inadequate removal of the forced signal in the observations leads to biased regression patterns (Brown et al., 2015). In the model simulations the forcing is held constant and there is only a modest model drift to detrend (Fig. 1) for which we chose to a second order polynomial. The observations are based on the climate system's

360 response to non-stationary forcing and we chose the scaled multi-model mean approach (Frankcombe et al., 2015). Here the detrending signal is derived from the multi-model mean of the CMIP5 models which are biased in their own ways and have different relative climate sensitivities to different forcings. Furthermore, the external forcing prescribed in the CMIP5 models is in itself uncertain and hence may alter the relative contribution due to external forcing and internal variability. Furthermore, the atmospheric grid spacing decreases from $1°$ in the LOW to $0.5°$ in the HIGH simulation. However, no new essential

365 atmospheric processes are resolved, so no significant changes are expected apart from coupling to different ocean boundary conditions.

These caveats do not detract from our main point which is to emphasize that low frequency variability may be underestimated in climate models with low-resolution ocean model components and that representing mesoscale eddies in climate models can improve the simulation of this variability. Of course, we showed this here only for the CESM but expect the improvement of

370 multidecadal variability to be generalizable to other coupled climate models, as some of the features like enhanced vertical heat transport are consistent with results obtained with other models (Griffies et al., 2015). The introduction of high frequency variability through mesoscale ocean features and the reduced ocean state bias can further improve the simulations' multidecadal variability skill. Certainly, better eddy parametrizations may improve shortcomings of the traditional Gent-McWilliams approach (Zanna et al., 2017). A recent study claimed an absence of evidence for multidecadal variability in CMIP5 models

375 (Mann et al., 2020), but all these models use non-eddying ocean components. As most CMIP6 models outside the High Resolution Model Intercomparison Project (Haarsma et al., 2016) still use non-eddying ocean components, not much improvement is expected on the representation of multidecacal variability in these models compared to CMIP5 models.

Internal variability obscures any forced GMST signal and periods of accelerated and decelerated warming are observed, such as the recent warming trend slowdown (Medhaug et al., 2017). Many studies use global circulation models with coarse

380 resolution ocean components to investigate, for example, the origin of these so-called hiatuses (e.g. Maher et al. (2014)). In light of our findings, estimates of the frequency and magnitude of excursions from the forced trend may be systematically low biased in low-resolution models, as internal multidecadal variability is underestimated. The increased internal variability also implies that the attribution of forced signals becomes more difficult and the issue of the origin of the recent warming trend slowdown may therefore never be satisfactorily resolved (Hedemann et al., 2017). Finally, our finding of stronger than

385 previously modelled multidecadal OHC variability underlines the necessity of continued observations of the ocean heat content
with the ARGO program.

Code and data availability. The analysis scripts are available at `zenodo link to final code repository`, while the model output is stored at SURFsara and available upon request to the corresponding author.

Author contributions. AJ, AvdH, and HD conceived the presented ideas in this study. AJ performed the analysis and wrote the manuscript.
390 AvdH and HD contributed to writing the paper.

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

Acknowledgements. This work was carried out under the program of the Netherlands Earth System Science Centre (NESSC), financially supported by the Ministry of Education, Culture and Science (OCW) (Grantnr. 024.002.001). The computations were performed on the Cartesius high performance computer at SURFsara in Amsterdam. Use of the Cartesius computing facilities was sponsored by the Netherlands
395 Science Foundation (NWO) under the project 15502. We thank Michael Kliphuis (IMAU) for carrying out the computations.

References

Berloff, P., McC. Hogg, A., and Dewar, W.: The turbulent oscillator: A mechanism of low-frequency variability of the wind-driven gyres, Journal of Physical Oceanography, 37, 2363–2386, https://doi.org/10.1175/JPO3118.1, 2007a.

Berloff, P. S., Dewar, W., Kravtsov, S., and McWilliams, J. C.: Ocean Eddy Dynamics in a Coupled Ocean–Atmosphere Model*, Journal of
400   Physical Oceanography, 37, 1103–1121, https://doi.org/10.1175/JPO3041.1, 2007b.

Biastoch, A., Lutjeharms, J. R. E., Böning, C. W., and Scheinert, M.: Mesoscale perturbations control inter-ocean exchange south of Africa, Geophysical Research Letters, 35, L20 602, https://doi.org/10.1029/2008GL035132, 2008.

Brown, P. T., Li, W., and Xie, S.-P.: Regions of significant influence on unforced global mean surface air temperature variability in climate models, Journal of Geophysical Research: Atmospheres, 120, 480–494, https://doi.org/10.1002/2014JD022576, 2015.

405   Buckley, M. W. and Marshall, J.: Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review, Reviews of Geophysics, 54, 5–63, https://doi.org/10.1002/2015RG000493, 2016.

Cheung, A. H., Mann, M. E., Steinman, B. A., Frankcombe, L. M., England, M. H., and Miller, S. K.: Comparison of Low-Frequency Internal Climate Variability in CMIP5 Models and Observations, Journal of Climate, 30, 4763–4776, https://doi.org/10.1175/JCLI-D-16-0712.1, 2017.

410   Chylek, P., Folland, C. K., Dijkstra, H. A., Lesins, G., and Dubey, M. K.: Ice-core data evidence for a prominent near 20 year time-scale of the Atlantic Multidecadal Oscillation, Geophysical Research Letters, 38, n/a–n/a, https://doi.org/10.1029/2011GL047501, 2011.

Couvelard, X., Dumas, F., Garnier, V., Ponte, A. L., Talandier, C., and Treguier, A. M.: Mixed layer formation and restratification in presence of mesoscale and submesoscale turbulence, Ocean Modelling, 96, 243–253, https://doi.org/10.1016/j.ocemod.2015.10.004, 2015.

Delworth, T., Manabe, S., and Stouffer, R. J.: Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model,
415   https://doi.org/10.1175/1520-0442(1993)006<1993:IVOTTC>2.0.CO;2, 1993.

Delworth, T. L. and Mann, M. E.: Observed and simulated multidecadal variability in the Northern Hemisphere, Climate Dynamics, 16, 661–676, https://doi.org/10.1007/s003820000075, 2000.

Delworth, T. L., Rosati, A., Anderson, W., Adcroft, A. J., Balaji, V., Benson, R., Dixon, K., Griffies, S. M., Lee, H.-C., Pacanowski, R. C., Vecchi, G. A., Wittenberg, A. T., Zeng, F., and Zhang, R.: Simulated Climate and Climate Change in the GFDL CM2.5 High-Resolution
420   Coupled Climate Model, Journal of Climate, 25, 2755–2781, https://doi.org/10.1175/JCLI-D-11-00316.1, 2012.

Deser, C., Alexander, M. a., Xie, S.-P., and Phillips, A. S.: Sea surface temperature variability: patterns and mechanisms., vol. 2, https://doi.org/10.1146/annurev-marine-120408-151453, 2010.

Dijkstra, H. A.: Nonlinear Climate Dynamics, Cambridge University Press, New York, USA, 2013.

Dijkstra, H. A.: A Normal Mode Perspective of Intrinsic Ocean-Climate Variability, Annual Review of Fluid Mechanics, 48, 341–363,
425   https://doi.org/10.1146/annurev-fluid-122414-034506, 2016.

Dufour, C. O., Morrison, A. K., Griffies, S. M., Frenger, I., Zanowski, H., and Winton, M.: Preconditioning of the Weddell Sea polynya by the ocean mesoscale and dense water overflows, Journal of Climate, 30, 7719–7737, https://doi.org/10.1175/JCLI-D-16-0586.1, 2017.

Durbin, J. and Koopman, S. J.: Time series analysis by state space methods, Oxford University Press, 2012.

Eden, C. and Jung, T.: North Atlantic Interdecadal Variability: Oceanic Response to the North Atlantic Oscillation (1865–1997), Journal of
430   Climate, 14, 676–691, https://doi.org/10.1175/1520-0442(2001)014<0676:NAIVOR>2.0.CO;2, 2001.

England, M. H., McGregor, S., Spence, P., Meehl, G. a., Timmermann, A., Cai, W., Gupta, A. S., McPhaden, M. J., Purich, A., and Santoso, A.: Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus, Nature Climate Change, 4, 222–227, https://doi.org/10.1038/nclimate2106, 2014.

Frajka-Williams, E., Beaulieu, C., and Duchez, A.: Emerging negative Atlantic Multidecadal Oscillation index in spite of warm subtropics,
435   Scientific Reports, 7, 1–8, https://doi.org/10.1038/s41598-017-11046-x, 2017.

Frankcombe, L. M., Dijkstra, H. A., and von der Heydt, A.: Noise-Induced Multidecadal Variability in the North Atlantic: Excitation of Normal Modes, Journal of Physical Oceanography, 39, 220–233, https://doi.org/10.1175/2008JPO3951.1, 2009.

Frankcombe, L. M., England, M. H., Mann, M. E., and Steinman, B. A.: Separating internal variability from the externally forced climate response, Journal of Climate, 28, 8184–8202, https://doi.org/10.1175/JCLI-D-15-0069.1, 2015.

440 Frankcombe, L. M., England, M. H., Kajtar, J. B., Mann, M. E., and Steinman, B. A.: On the Choice of Ensemble Mean for Estimating the Forced Signal in the Presence of Internal Variability, Journal of Climate, 31, 5681–5693, https://doi.org/10.1175/JCLI-D-17-0662.1, 2018.

Gebbie, G. and Huybers, P.: The Little Ice Age and 20th-century deep Pacific cooling, Science, 363, 70–74, https://doi.org/10.1126/science.aar8413, 2019.

445 Gent, P. R. and McWilliams, J. C.: Isopycnal Mixing in Ocean Circulation Models, J. Phys. Oceanogr., 20, 150–155, https://doi.org/10.1175/1520-0485, 1990.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F., and Yiou, P.: Advanced spectral methods for climatic time series, Reviews of Geophysics, 40, https://doi.org/10.1029/2000RG000092, 2002.

Gray, S. T., Graumlich, L. J., Betancourt, J. L., and Pederson, G. T.: A tree-ring based reconstruction of the Atlantic Multidecadal Oscillation
450   since 1567 A.D., Geophysical Research Letters, 31, n/a–n/a, https://doi.org/10.1029/2004GL019932, 2004.

Griffies, S. M., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., Dunne, J. P., Goddard, P., Morrison, A. K., Rosati, A., Wittenberg, A. T., Yin, J., and Zhang, R.: Impacts on Ocean Heat from Transient Mesoscale Eddies in a Hierarchy of Climate Models, Journal of Climate, 28, 952–977, https://doi.org/10.1175/JCLI-D-14-00353.1, 2015.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Catherine, A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., Von
455   Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J. J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and Von Storch, J. S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geoscientific Model Development, 9, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016, 2016.

Hallberg, R.: Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects, Ocean Modelling, 72, 92–103, https://doi.org/10.1016/j.ocemod.2013.08.007, 2013.

460 Hallberg, R. and Gnanadesikan, A.: The Role of Eddies in Determining the Structure and Response of the Wind-Driven Southern Hemisphere Overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) Project, Journal of Physical Oceanography, 36, 2232–2252, https://doi.org/10.1175/JPO2980.1, 2006.

Hasselmann, K.: Stochastic climate models Part I. Theory, Tellus, 28, 473–485, https://doi.org/10.3402/tellusa.v28i6.11316, 1976.

Hedemann, C., Mauritsen, T., Jungclaus, J., and Marotzke, J.: The subtle origins of surface-warming hiatuses, Nature Climate Change, 7,
465   336–339, https://doi.org/10.1038/nclimate3274, 2017.

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., and Folland, C. K.: A Tripole Index for the Interdecadal Pacific Oscillation, Climate Dynamics, 45, 3077–3090, https://doi.org/10.1007/s00382-015-2525-1, 2015.

Hogg, A. M. C. and Blundell, J. R.: Interdecadal variability of the Southern Ocean, Journal of Physical Oceanography, 36, 1626–1645, https://doi.org/10.1175/JPO2934.1, 2006.

470 Huck, T., Arzel, O., and Sévellec, F.: Multidecadal Variability of the Overturning Circulation in Presence of Eddy Turbulence, Journal of Physical Oceanography, 45, 157–173, https://doi.org/10.1175/jpo-d-14-0114.1, 2014.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J., and Marshall, S.: The Community Earth System Model: A Framework for Collaborative Research, Bulletin of the

475 American Meteorological Society, 94, 1339–1360, https://doi.org/10.1175/BAMS-D-12-00121.1, 2013.

Jüling, A., Viebahn, J. P., Drijfhout, S. S., and Dijkstra, H. A.: Energetics of the Southern Ocean Mode, Journal of Geophysical Research: Oceans, https://doi.org/10.1029/2018jc014191, 2018.

Kajtar, J. B., Collins, M., Frankcombe, L. M., England, M. H., Osborn, T. J., and Juniper, M.: Global mean surface temperature response to large-scale patterns of variability in observations and CMIP5, Geophysical Research Letters, https://doi.org/10.1029/2018GL081462,

480 2019.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J. F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The community earth system model (CESM) large ensemble project : A community resource for studying climate change in the presence of internal climate variability, Bulletin of the American Meteorological Society, 96, 1333–1349, https://doi.org/10.1175/BAMS-

485 D-13-00255.1, 2015.

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., Kinter, J. L., Loft, R., Rousset, C., Siqueira, L., Stan, C., Tomas, R., and Vertenstein, M.: Impact of ocean model resolution on CCSM climate simulations, Climate Dynamics, 39, 1303–1328, https://doi.org/10.1007/s00382-012-1500-3, 2012.

Knudsen, M. F., Seidenkrantz, M.-S., Jacobsen, B. H., and Kuijpers, A.: Tracking the Atlantic Multidecadal Oscillation through the last 8,000

490 years, Nature Communications, 2, 178, https://doi.org/10.1038/ncomms1186, 2011.

Kosaka, Y. and Xie, S.-P.: Recent global-warming hiatus tied to equatorial Pacific surface cooling., Nature, 501, 403–7, https://doi.org/10.1038/nature12534, 2013.

Kushnir, Y.: Interdecadal Variations in North Atlantic Sea Surface Temperature and Associated Atmospheric Conditions, Journal of Climate, 7, 141–157, https://doi.org/10.1175/1520-0442(1994)007<0141:IVINAS>2.0.CO;2, 1994.

495 Le Bars, D., Viebahn, J. P., and Dijkstra, H. A.: A Southern Ocean mode of multidecadal variability, Geophysical Research Letters, 43, 2102–2110, https://doi.org/10.1002/2016GL068177, 2016.

Maher, N., Gupta, A. S., and England, M. H.: Drivers of decadal hiatus periods in the 20th and 21st centuries, Geophysical Research Letters, 41, 5978–5986, https://doi.org/10.1002/2014GL060527, 2014.

Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., Kröger, J., Takano, Y., Ghosh, R., Hedemann, C.,

500 Li, C., Li, H., Manzini, E., Notz, D., Putrasahan, D., Boysen, L., Claussen, M., Ilyina, T., Olonscheck, D., Raddatz, T., Stevens, B., and Marotzke, J.: The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability, Journal of Advances in Modeling Earth Systems, 11, 2050–2069, https://doi.org/10.1029/2019MS001639, 2019.

Mann, M. E., Steinman, B. A., and Miller, S. K.: Absence of internal multidecadal and interdecadal oscillations in climate model simulations, Nature Communications, 11, 49, https://doi.org/10.1038/s41467-019-13823-w, 2020.

505 Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C.: A Pacific Interdecadal Climate Oscillation with
Impacts on Salmon Production, Bulletin of the American Meteorological Society, 78, 1069–1079, https://doi.org/10.1175/1520-
0477(1997)078<1069:APICOW>2.0.CO;2, 1997.

Manucharyan, G. E., Thompson, A. F., and Spall, M. A.: Eddy Memory Mode of Multidecadal Variability in Residual-Mean Ocean Circula-
tions with Application to the Beaufort Gyre, Journal of Physical Oceanography, 47, 855–866, https://doi.org/10.1175/JPO-D-16-0194.1,
510 2017.

Martin, P. E., Arbic, B. K., McC. Hogg, A., Kiss, A. E., Munroe, J. R., and Blundell, J. R.: Frequency-domain Analysis of the Energy Budget
in an Idealized, Coupled, Ocean-Atmosphere Model, Journal of Climate, pp. 707–726, https://doi.org/10.1175/jcli-d-19-0118.1, 2019.

McWilliams, J. C.: The nature and consequences of oceanic eddies, in: Eddy-Resolving Ocean Modeling, vol. 1, pp. 5–15, Cambridge
University Press, https://doi.org/10.1029/177GM03, 2008.

515 Medhaug, I., Stolpe, M. B., Fischer, E. M., and Knutti, R.: Reconciling controversies about the 'global warming hiatus', Nature, 545, 41–47,
https://doi.org/10.1038/nature22315, 2017.

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., Mantua, N. J., Miller, A. J., Minobe, S., Nakamura, H.,
Schneider, N., Vimont, D. J., Phillips, A. S., Scott, J. D., and Smith, C. A.: The Pacific decadal oscillation, revisited, Journal of Climate,
29, 4399–4427, https://doi.org/10.1175/JCLI-D-15-0508.1, 2016.

520 Penduff, T., Juza, M., Barnier, B., Zika, J., Dewar, W. K., Treguier, A.-M. M., Molines, J.-M. M., and Audiffren, N.: Sea Level Expression
of Intrinsic and Forced Ocean Variabilities at Interannual Time Scales, Journal of Climate, 24, 5652–5670, https://doi.org/10.1175/JCLI-
D-11-00077.1, 2011.

Penduff, T., Barnier, B., Terray, L., Bessières, L., Sérazin, G., S.Gregorio, Brankart, J.-M., Moine, M.-P., Molines, J.-M., and P.Brasseur:
Ensembles of eddying ocean simulations for climate, CLIVAR Exchanges No. 65, 19, 26–29, 2014.

525 Power, S., Casey, T., Folland, C., Colman, A., and Mehta, V.: Inter-decadal modulation of the impact of ENSO on Australia, Climate
Dynamics, 15, 319–324, https://doi.org/10.1007/s003820050284, 1999.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., and Rowell, D. P.: Global analyses of sea surface tem-
perature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research, 108, 4407,
https://doi.org/10.1029/2002JD002670, 2003.

530 Simpkins, G. R., Ciasto, L. M., and England, M. H.: Observed variations in multidecadal Antarctic sea ice trends during 1979-2012, Geo-
physical Research Letters, 40, 3643–3648, https://doi.org/10.1002/grl.50715, 2013.

Small, R. J., Bacmeister, J., Bailey, D., Baker, A., Bishop, S., Bryan, F., Caron, J., Dennis, J., Gent, P., Hsu, H., Jochum, M., Lawrence,
D., Muñoz, E., DiNezio, P., Scheitlin, T., Tomas, R., Tribbia, J., Tseng, Y., and Vertenstein, M.: A new synoptic scale resolving
global climate simulation using the Community Earth System Model, Journal of Advances in Modeling Earth Systems, 6, 1065–1094,
535 https://doi.org/10.1002/2014MS000363, 2014.

Steinman, B. A., Mann, M. E., and Miller, S. K.: Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures,
Science, 347, 988–991, https://doi.org/10.1126/science.1257856, 2015.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., And, V. B., and P.M. Midgley (eds.):
IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of
540 the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,
https://doi.org/10.1017/CBO9781107415324, 2013.

Stolpe, M. B., Medhaug, I., and Knutti, R.: Contribution of Atlantic and Pacific Multidecadal Variability to Twentieth-Century Temperature Changes, Journal of Climate, 30, 6279–6295, https://doi.org/10.1175/JCLI-D-16-0803.1, 2017.

Sutton, R. T. and Hodson, D. L.: Ocean science: Atlantic Ocean forcing of North American and European summer climate, Science, 309,
545   115–118, https://doi.org/10.1126/science.1109496, 2005.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Te Raa, L. A. and Dijkstra, H. A.: Instability of the thermohaline ocean circulation on interdecadal timescales, Journal of Physical Oceanography, 32, 138–160, https://doi.org/10.1175/1520-0485(2002)032<0138:IOTTOC>2.0.CO;2, 2002.

550 Trenberth, K. E.: Some Effects of Finite Sample Size and Persistence on Meteorological Statistics. Part I: Autocorrelations, Monthly Weather Review, 112, 2359–2368, https://doi.org/10.1175/1520-0493(1984)112<2359:SEOFSS>2.0.CO;2, 1984.

Trenberth, K. E. and Shea, D. J.: Atlantic hurricanes and natural variability in 2005, Geophysical Research Letters, 33, 1–4, https://doi.org/10.1029/2006GL026894, 2006.

Tulloch, R., Marshall, J., Hill, C., and Smith, K. S.: Scales, Growth Rates, and Spectral Fluxes of Baroclinic Instability in the Ocean, Journal
555   of Physical Oceanography, 41, 1057–1076, https://doi.org/10.1175/2011JPO4404.1, 2011.

van Westen, R. M. and Dijkstra, H. A.: Southern Ocean Origin of Multidecadal Variability in the North Brazil Current, Geophysical Research Letters, 44, 10,540–10,548, https://doi.org/10.1002/2017GL074815, 2017.

Viebahn, J. P., Crommelin, D., and Dijkstra, H. A.: Towards a turbulence closure based on energy modes, Journal of Physical Oceanography, https://doi.org/10.1175/jpo-d-18-0117.1, 2018.

560 Wang, J., Yang, B., Ljungqvist, F. C., Luterbacher, J., Osborn, T. J., Briffa, K. R., and Zorita, E.: Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years, Nature Geoscience, 10, 512–517, https://doi.org/10.1038/ngeo2962, 2017.

Weijer, W. and van Sebille, E.: Impact of Agulhas leakage on the Atlantic overturning circulation in the CCSM4, Journal of Climate, 27, 101–110, https://doi.org/10.1175/JCLI-D-12-00714.1, 2014.

Wieners, C. E., Dijkstra, H. A., and de Ruijter, W. P. M.: The interaction between the Western Indian Ocean and ENSO in CESM, Climate
565   Dynamics, 52, 5153–5172, https://doi.org/10.1007/s00382-018-4438-2, 2019.

Zanna, L., Porta Mana, P., Anstey, J., David, T., and Bolton, T.: Scale-aware deterministic and stochastic parametrizations of eddy-mean flow interaction, Ocean Modelling, 111, 66–80, https://doi.org/10.1016/j.ocemod.2017.01.004, 2017.

Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., and Heimbach, P.: Global reconstruction of historical ocean heat storage and transport, Proceedings of the National Academy of Sciences, p. 201808838, https://doi.org/10.1073/pnas.1808838115, 2019.

570 Zhang, L. and Wang, C.: Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations, Journal of Geophysical Research: Oceans, 118, 5772–5791, https://doi.org/10.1002/jgrc.20390, 2013.

Zhang, L., Delworth, T. L., Cooke, W., and Yang, X.: Natural variability of Southern Ocean convection as a driver of observed climate trends, Nature Climate Change, 9, 59–65, https://doi.org/10.1038/s41558-018-0350-3, 2019.