Contribution of Atlantic and Pacific Multidecadal Variability to Twentieth-Century Temperature Changes

MARTIN B. STOLPE, ISELIN MEDHAUG, AND RETO KNUTTI

Institute for Atmospheric and Climate Science, ETH Zürich, Zurich, Switzerland

(Manuscript received 10 November 2016, in final form 10 May 2017)

ABSTRACT

Recent studies have suggested that significant parts of the observed warming in the early and the late twentieth century were caused by multidecadal internal variability centered in the Atlantic and Pacific Oceans. Here, a novel approach is used that searches for segments of unforced preindustrial control simulations from global climate models that best match the observed Atlantic and Pacific multidecadal variability (AMV and PMV, respectively). In this way, estimates of the influence of AMV and PMV on global temperature that are consistent both spatially and across variables are made. Combined Atlantic and Pacific internal variability impacts the global surface temperatures by up to 0.15°C from peak-to-peak on multidecadal time scales. Internal variability contributed to the warming between the 1920s and 1940s, the subsequent cooling period, and the warming since then. However, variations in the rate of warming still remain after removing the influence of internal variability associated with AMV and PMV on the global temperatures. During most of the twentieth century, AMV dominates over PMV for the multidecadal internal variability imprint on global and Northern Hemisphere temperatures. Less than 10% of the observed global warming during the second half of the twentieth century is caused by internal variability in these two ocean basins, reinforcing the attribution of most of the observed warming to anthropogenic forcings.

1. Introduction

The observed increase in global surface temperatures does not occur uniformly over time but exhibits multidecadal variability. The decadal rate of global temperature increase has changed significantly three times over the observational period (Cahill et al. 2015). A period of warming in the first half of the twentieth century was succeeded by a slight cooling. Since 1970, the global mean temperature increases again. Also, the spatial pattern of temperature trends varied throughout the twentieth century (Knutson et al. 2016). Although its statistical significance remains controversial (Cahill et al. 2015; Karl et al. 2015; Fyfe et al. 2016), an apparent slowdown in temperature increase between about 1998 and 2012—sometimes referred to as global warming hiatus or pause—received a lot of attention in both the media and scientific literature (e.g., Kosaka and Xie 2013; England et al. 2014; Medhaug et al. 2017).

While most of the global warming that occurred in the second half of the twentieth century is attributed to increases in atmospheric greenhouse gas (GHG) concentration (Bindoff et al. 2013), the role of unforced internal variability in the acceleration or slowdown of the observed warming is heavily debated. The mean of climate model simulations from phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) only partly resembles the observed spatial and temporal inhomogeneities in the rate of warming. This suggests that internal variability plays a role in the temperature record (Knutson et al. 2016). Studies trying to quantify and attribute the variations in the rate of warming to internal variability identified two important modes of internal multidecadal variability in the sea surface temperatures (SSTs; Parker et al. 2007; Deser et al. 2010): the Atlantic multidecadal variability (AMV) and the Pacific multidecadal variability (PMV); sometimes also referred to as the Pacific decadal oscillation (PDO) for the North Pacific variability or as the interdecadal Pacific oscillation (IPO) for the decadal-scale variability of the whole Pacific. On a decadal time scale...
scale, PDO and IPO are highly correlated (Henley et al. 2015). Here, we use PMV to describe a Pacific-wide manifestation of multidecadal variability. Possibly, there is also considerable multidecadal variability in the Southern Ocean (e.g., Le Bars et al. 2016), but because of a lack of historical observation (Deser et al. 2010), we do not consider Southern Ocean variability in this study.

The AMV describes large-scale SST variations in the North Atlantic Ocean with a period of roughly 60–80 years (Schlesinger and Ramankutty 1994; Drinkwater et al. 2014). Both land-based (e.g., Gray et al. 2004) and marine-based (e.g., Svendsen et al. 2014) reconstructions suggest that AMV also existed prior to the instrumental record several centuries back in time. It is unclear whether AMV is solely a mode of internal variability (e.g., Knight et al. 2005), is due to a combination of anthropogenic forcing and internal variability (e.g., Mann and Emanuel 2006), or is driven by anthropogenic forcings alone (e.g., Booth et al. 2012). Climate model simulations suggest that AMV is linked to variations in the Atlantic meridional overturning circulation (AMOC; e.g., Delworth et al. 1993; Delworth and Mann 2000; Buckley and Marshall 2016). The AMOC transports large amounts of heat northward, and variations in AMOC strength thereby have a profound impact on Earth’s climate (Buckley and Marshall 2016). AMV influences global temperatures by redistributing heat into the high northern latitudes (where local feedbacks amplify the warming) and by redistributing heat between atmosphere and ocean (Gregory 2000; Knutti and Stocker 2000; Raper et al. 2002).

The PMV describes a Pacific-wide multidecadal El Niño–Southern Oscillation (ENSO)-like climate variation in the Pacific SSTs (Mantua et al. 1997; Zhang et al. 1997; Power et al. 1999). Reconstructions indicate that multidecadal variability occurred in the Pacific Ocean during the last several hundred years (MacDonald and Case 2005; Shen et al. 2006; Felis et al. 2010), but the various reconstructions differ strongly (Newman et al. 2016). PMV during the twentieth century might also have been modulated by anthropogenic forcings, where aerosol emissions from North America and China influence the strength of the Aleutian low and thereby winds, which manifests itself as changes in Pacific SSTs (Yeh et al. 2013; Boo et al. 2015; Smith et al. 2016).

Several recent studies tried to quantify the contribution of AMV and PMV to the observed warming during the twentieth century. On a global scale, multidecadal internal variability was found to amplify or suppress global temperatures by up to $\pm 0.08 \, K \, \text{decade}^{-1}$ for 30-year-long trends (DelSole et al. 2011). Estimates of the role of internal variability in modulating observed global twentieth-century air temperatures differ substantially. On the high end, it has been estimated that a large part of the warming over the last 50 years can be attributed to AMV (Tung and Zhou 2013; Zhou and Tung 2013) and that the internal-variability-adjusted warming rate is steady (Tung and Zhou 2013; Zhou and Tung 2013) or monotonically increasing (Swanson et al. 2009) during the past century. Further, over the last three decades one-third of the warming was argued to be caused by a positive phase of the AMV (Wu et al. 2011; Chylek et al. 2014a, 2016). In contrast to these studies, Imbers et al. (2013) find only a minor AMV contribution to global warming since 1951. Dai et al. (2015), Kosaka and Xie (2016), and Meehl et al. (2016) attribute parts of the previously mentioned earlier decades of halted and accelerated warming to variations in PMV, describing the Pacific as a pacemaker for the global warming rate (Kosaka and Xie 2016). Others, however, argue for the Atlantic as the key pacemaker (Barcikowska et al. 2017). Part of the recent apparent slowdown in temperature increase has also been ascribed to a negative phase of PMV and an associated increase in heat uptake by the Pacific Ocean (Kosaka and Xie 2013; Trenberth and Fasullo 2013; England et al. 2014; Huber and Knutti 2014; McGregor et al. 2014; Meehl et al. 2014; Steinman et al. 2015).

Some studies (Tung and Zhou 2013; Zhou and Tung 2013; Chylek et al. 2014a,b) rely on regression methods that may inflate the warming attributed to AMV due to strong covariation between two predictors, such as between aerosol forcing and AMV (e.g., Booth et al. 2012). Further, these methods usually do not impose physical constraints; that is, as long as the temporal evolution of AMV or PMV is similar to the observed global temperature increase, these modes of internal variability “explain”—by construction—a substantial part of the observed warming irrespective of whether there is a physical link. Besides regression approaches, pacemaker experiments have been carried out to understand and quantify the role of internal multidecadal variability (e.g., Zhang et al. 2007; Kosaka and Xie 2013). For these experiments, observed SSTs (or other variables such as surface winds or heat fluxes) are prescribed in the regions of interest in a climate model. Such simulations resolve the imprint of internal variability on the climate system spatially and across variables and help to understand transbasin teleconnections, but they do not conserve energy and are computationally expensive.

Climate models are able to simulate both AMV and PMV patterns (Farneti 2017), but their simulated variability (to the degree it is unforced) is generally not in phase with the observed multidecadal variability. In this study, we therefore search a large suite of climate model...
simulations for periods where the observed and modeled internal variability match closest. These so-called variability analog quantify the contribution of AMV and PMV to global and regional climate change. As the pacemaker approach, this method resolves the effects of AMV and PMV on the climate spatially. At the same time, it is internally consistent across climate variables, we do not violate energy conservation, the computational cost is small as we use existing simulations, and we can explore uncertainties by utilizing the full set of state-of-the-art climate models.

2. Data and methods
a. Observations

The observed area-averaged SSTs in the Atlantic and Pacific Oceans consist of both a forced signal and internal variability (e.g., Ting et al. 2014). Various studies have developed methods to separate internal variability from the forced response in the two basins. The underlying assumption of all of these attribution-type studies is that internal variability and the forced response are linearly additive, and optimal fingerprint attribution tests this by requiring that the residual unexplained by a sum of forced responses is consistent with internal unforced variability (Bindoff et al. 2013). In the following, we summarize the indices and display them in Fig. 1.

For the North Atlantic (defined here as 0°–60°N, 5°–75°W), we use indices developed by Frankcombe et al. (2015), who advance the approach of Mann et al. (2014) and Steinman et al. (2015). The first approach of Frankcombe et al. (2015) makes a best estimate of the imprint of the forced SST component from the historical CMIP5 model simulations (containing all forcings) over the basin-averaged North Atlantic [Eq. (1)]. To account for potential biases between the “true” and simulated forced signal in the North Atlantic, they rescale the simulated response via linear regression against observed mean SSTs and subtract the scaled response from the observed North Atlantic SSTs (Frankcombe et al. 2015). We refer to this method as the Frankcombe single-factor scaling:

$$\text{AMV}(t)_{\text{Frankcombe,1}} = \text{NASST}(t) - (\alpha + \beta \text{MMM}(t)), \tag{1}$$

where NASST is the observed North Atlantic mean SST, MMM is the CMIP5 multimodel mean, $t$ is the time, and $\alpha$ and $\beta$ are constants determined by the regression. We refer to $\beta$ as the scaling factor with which the MMM is scaled to NASST. The difference between NASST and the rescaled MMM is the estimated single-factor scaling AMV.

The rescaling distinguishes this method from an earlier approach that directly subtracts the multimodel mean from the observed North Atlantic SST (Knight 2009). Ideally, the residual then represents the averaged North Atlantic SST evolution solely due to internal variability.

As an alternative method, Frankcombe et al. (2015) use two CMIP5 multimodel means, the greenhouse-gas-only and the natural-forcing-only simulations, and apply the scaling separately to the two simulations through multilinear regression (this results in two $\beta$, and therefore we refer to this as the Frankcombe two-factor scaling). These scaling methods have different strengths. The two-factor scaling accounts for the possibility that models have different sensitivities to greenhouse gas and natural forcings, while the single-factor scaling includes anthropogenic aerosol and ozone forcing, which may be important for North Atlantic SSTs (Booth et al. 2012). In a three-factor scaling approach, Frankcombe et al. (2015) also include the residual that remains after accounting for greenhouse gas and natural forcings, but note that this method may be
Table 1. Definitions of different AMV and PMV indices used in this study. The area considered in the North Atlantic is the same in all indices (0°–60°N, 5°–75°W).

<table>
<thead>
<tr>
<th>AMV index</th>
<th>Methodology</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frankcombe single-factor scaling</td>
<td>Subtract scaled multimodel mean (historical simulations with all forcings) from the North Atlantic mean SSTs.</td>
<td>Frankcombe et al. (2015)</td>
</tr>
<tr>
<td>Frankcombe two-factor scaling</td>
<td>Subtract scaled multimodel means (of the GHG-only and natural simulations) from the North Atlantic mean SSTs.</td>
<td>Frankcombe et al. (2015)</td>
</tr>
<tr>
<td>Frankcombe three-factor scaling</td>
<td>Subtract scaled multimodel means (of the GHG-only and natural simulations and of the remaining unexplained variability) from the North Atlantic mean SSTs.</td>
<td>Frankcombe et al. (2015)</td>
</tr>
<tr>
<td>Enfield and Cid-Serrano</td>
<td>Subtract a quadratic fit from the North Atlantic mean SSTs.</td>
<td>Enfield and Cid-Serrano (2010)</td>
</tr>
<tr>
<td>Trenberth and Shea</td>
<td>Subtract global mean SSTs (60°S–60°N) from the North Atlantic mean SSTs.</td>
<td>Trenberth and Shea (2006)</td>
</tr>
<tr>
<td>Enfield</td>
<td>Subtract a linear trend from the North Atlantic mean SSTs.</td>
<td>Enfield et al. (2001)</td>
</tr>
<tr>
<td>PMV index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Henley</td>
<td>Difference in mean SST between three regions in the Pacific (between central equatorial, northwestern, and southwestern Pacific).</td>
<td>Henley et al. (2015)</td>
</tr>
<tr>
<td>Mantua</td>
<td>Subtract global mean SSTs (60°S–60°N) from the North Pacific (≥20°N) and calculate the leading EOF of North Pacific SSTs.</td>
<td>Mantua et al. (1997)</td>
</tr>
</tbody>
</table>

more prone to misattributing internal variability to the forced signal than their other methods.

We compare the three Frankcombe factor scaling AMV estimates with the approaches of Enfield et al. (2001; in the following, Enfield), Enfield and Cid-Serrano (2010; in the following, Enfield and Cid-Serrano), and Trenberth and Shea (2006; in the following, Trenberth and Shea). We summarize all AMV definitions used in Table 1 and show the time series in Fig. 1a.

All AMV indices agree that the Atlantic was in a cold state in the early twentieth century and in the 1970–80s and in a warm phase in the mid- and late-twentieth century (Fig. 1a). Differences between AMV indices are largest in the late nineteenth/early twentieth century and during recent decades (note that this also depends on the chosen reference period). The three Frankcombe scaling approaches slightly shift the phase of the AMV compared to the other indices. This can be seen in the 1960s and 1970s.

During the late nineteenth/early twentieth century, the Trenberth and Shea AMV shows the weakest signal. SST observations during this period are sparse, with the North Atlantic relatively well sampled compared to other ocean basins (Deser et al. 2010). Global mean SST biased toward the North Atlantic might be a reason for the small signal in the Trenberth and Shea AMV during these years.

Recent decades experienced rapid global warming and therefore the method to isolate the internal variability component of North Atlantic SSTs makes a large difference. The Enfield AMV shows the largest amplitude, followed by Enfield and Cid-Serrano and Trenberth and Shea. Of the three model-based indices, the two-factor scaling AMV increases the most, reaching a similar anomaly as the “purely” observational SSTs, followed by the single-factor scaling AMV. The three-factor scaling AMV shows only a minor increase in unforced North Atlantic mean SSTs since the 1980s.

Frankcombe et al. (2015) compare different approaches to separate internal variability from the forced signal in the North Atlantic and found the scaling approaches to work better in both restoring the amplitude and phase of the unforced variability compared to methods that use a linear trend as an estimate of the forced signal (i.e., Enfield) or subtract the climate model mean without scaling. Considerable uncertainties in isolating the internal variability, however, remain, which is reflected in the large spread between the three different scaling AMVs during recent decades although they follow similar approaches. In the following, we will use the Frankcombe two-factor scaling as our primary index but compare it with the other AMV indices.

To capture the Pacific-wide climate variability, we use the Henley et al. (2015; in the following, Henley) tripole index. The Henley index is defined as the difference between the area-averaged SST anomaly in the central equatorial Pacific (10°S–10°N, 170°E–90°W) and the mean over the northwestern (25°–45°N, 140°E–145°W) and southwestern (50°–15°S, 150°E–160°W) Pacific. Under the assumption that the forced signal is of similar strength in the three regions, it cancels and the index represents the pure signal of internal variability. By
prescribing the geographic location of the centers of action, selecting the variability analogs based on the Henley index has the advantage of both matching the PMV time series of the variability analogs with observations and also ensuring the correct location of the centers of action. We compare the results with the Mantua et al. (1997; in the following, Mantua) index of North Pacific variability (see Table 1 for details and Fig. 1b). Since it additionally takes into account information in the tropical and southern Pacific, we use Henley as our primary PMV index.

Mantua and Henley do not describe exactly the same mode of Pacific variability. Mantua depicts the North Pacific variability while Henley attempts to represent the Pacific-wide variability. Indices of North Pacific and Pacific-wide variability are, however, strongly correlated (Henley et al. 2015; Fig. 1b). Henley shows more high-frequency variability than Mantua, probably because of the imprint of the ENSO region that is included in Henley. In general, between around 1920 and 1940 and from the 1980s to the end of the 1990s both PMV indices were in a warm state, while from the 1940s to 1970s and again during the first decade of the twenty-first century they were in a cold state.

For all AMV and PMV definitions except for Mantua we use the Hadley Centre Ice and Sea Surface Temperature dataset, version 1.1 (HadISST1.1; Rayner et al. 2003). (We obtain the Mantua index directly from http://research.jisao.washington.edu/pdo/PDO.latest.) We include data until year 2015 except for the Frankcombe scaling approaches that end in 2005 (limited by the availability of historical climate model experiments).

b. Climate models

We use the preindustrial control (piControl) simulations of 35 general circulation models (GCMs; see Table SM1 in the supplemental material) that participated in CMIP5. Using a bilinear weighting, we interpolate all GCM results onto a common regular 2.5° × 2.5° grid. To eliminate spurious model drift, we follow the recommendation by Sen Gupta et al. (2013) and remove a linear trend fitted to the whole control simulation from all variables except for the subsurface seawater temperatures. As the deep ocean temperatures need very long to equilibrate, the ocean is prone to strong drift. Hence, we find that a linear trend is not sufficient (cf. Fig. SM1) and subtract a quadratic polynomial fit instead (similar results are obtained with a cubic polynomial drift estimate).

c. Methods

The models provide in total more than 19 000 years of unforced internal variability. We search for variability analogs that most closely resemble the observed realizations of internal variability in the Atlantic and Pacific Oceans. Huber and Knutti (2014) introduced this method first to estimate the ENSO imprint on global temperatures, and Risbey et al. (2014) used a related approach by picking only models in the same ENSO phase as observed to reconcile simulated and observed temperatures during the global warming hiatus.

We select the simulated variability analogs with the smallest root-mean-square error (RMSE) compared to the observed annual mean Atlantic and Pacific internal variability time series (as defined in Table 1) for 6-year-long periods. We shift this time window and repeat the search for the best variability analogs until we have sampled the whole twentieth century. We keep the 40 best matching analogs in each time window and do not use a specific maximum RMSE threshold. In addition to the RMSE also other metrics can be used to select analogs, such as the Pearson correlation coefficient in Huber and Knutti (2014). A disadvantage of the correlation coefficient is that it does not factor in an offset in the mean anomaly between the analogs and the observed index (e.g., a short analog from a cold AMV state could show a high correlation with the observed AMV in the warm state).

To suppress high-frequency variability, we smooth our results by applying local regressions using locally weighted scatterplot smoothing (LOESS) with a second-order polynomial (Cleveland 1979). Each regression is calculated using 23 years of data.

In principle, one can select analogs in two different ways: (i) select variability analogs from the piControl simulations that agree with both the observed Atlantic and Pacific variability simultaneously or (ii) select them separately for the two ocean basins and then add the two patterns of variability. The agreement between the simulated analogs and observed variability is better for (ii) than for (i). This is because (ii) only has one constraint of temporal coherence in the Atlantic or Pacific. However, the validity of selecting analogs separately depends on whether the climate imprint of the two modes of variability is additive. This question is addressed in the evaluation section of this paper.

Two major issues arise by using the analog method. First, no index of internal variability is perfect in a sense that it represents the “true” unforced variability. There is both a risk of underestimating internal variability (by overestimating the forced signal) or of overestimating internal variability (by underestimating the forced signal). We try to mitigate this by comparing different indices. Second, the magnitude (and possibly even the sign) of the imprint of Atlantic and Pacific modes of variability onto other climate variables varies across models and depends on the number and length of
We perform sensitivity experiments using different numbers of variability analogs for each period (from 5 up to 200 analogs) and different lengths of the analogs (from 5 to 80 years). The method is not very sensitive to these settings but behaves as expected: With increasing length and number of variability analogs, it gets harder to find matching analogs, which results in a growing underestimation of the observed amplitude in AMV and PMV (Fig. 2). This in turn reduces the amplitude of the estimated AMV and PMV imprint on observed global air temperature changes. These results are shown in Figs. SM2–SM4. Laepple and Huybers (2014) and Cheung et al. (2017) found better agreement between observed and modeled internal variability at higher frequencies, while interdecadal to multidecadal variability tends to be underestimated in models. This underestimation of multidecadal variability in longer variability analogs could be either because the 19,000 years of control simulations are still not sufficient, because there is a remaining forced signal in the isolated observed multidecadal variability, or because climate models indeed simulate too-weak multidecadal variability.

The sensitivity of global mean air temperatures to internal variability might depend on the time scale of the variability, with possibly different temperature responses to higher-frequency than to lower-frequency variability since physical mechanisms causing the variability are also time-scale dependent. To therefore test if our results are crucially dependent on the length of analogs, we calculate the sensitivity of global air temperatures to a unit change in AMV and PMV using different lengths of variability analogs. The sensitivities of global mean temperatures to AMV tend to increase by around 20% with increasing length of analogs when we compare 6-year-long analogs with 80-year-long analogs (for the 80-year-long analogs we smooth the AMV time series before we pick analogs to suppress high-frequency variability; additional information in Fig. SM5). For the response of global temperatures to PMV, we observe a decrease of about 20% in the sensitivity (Fig. SM5). This indicates that models that do well on a longer time scale simulate a larger global temperature response to SST anomalies in the Atlantic but a smaller temperature response to variations in the Pacific compared to models.

Fig. 2. Observed and simulated natural multidecadal variability in the (a),(c) North Atlantic and (b),(d) Pacific Oceans from 1880 to 2005. The observed variability is shown in blue (Frankcombe two-factor scaling for the Atlantic and Henley for the Pacific), the 6-year-long variability analogs are shown in red, and the 20-year-long analogs are shown in green. For (a) and (b) the analogs were selected separately for the Atlantic and Pacific Oceans, while they were selected simultaneously for (c) and (d). The time series have been filtered smoothed using LOESS.
that do well on a shorter time scale (and are then preferentially included into the 6-year-long analogs). Varying the number of analogs does not have a strong influence once more than around 20 analogs are selected (Fig. SM6).

With 40 analogs each having a length of 6 years, we avoid that analogs from one particular model influence the results too much and at the same time ensure that the amplitude of the observed AMV and PMV is well reproduced. There is a trade-off between tracking the short-term internal variability in the indices well by using short analogs and favoring models that simulate multidecadal variability comparatively well (by using long analogs). To account for the large spread in simulated global air temperature response to a unit AMV or PMV change across the CMIP5 models (Fig. SM7), we subsample the CMIP5 models. We repeat the search for variability analogs 60 times by randomly picking subsets of 10 CMIP5 models each from the 35 models. We use the 90% range (from 5th to 95th percentile) across these 60 realizations as a measure of structural model uncertainty with the limitation that these samples are not fully independent from each other. We add the uncertainty range to the best estimate using the full CMIP5 ensemble. When studying for example lead–lag relationships, longer analogs than used here might be necessary, but for investigating the direct effect of the analogs on the temperatures, 6 years is sufficient.

With these criteria, we achieve a good agreement between the observed time series and the variability analogs both for the Atlantic (Fig. 2a) and the Pacific (Fig. 2b). When selected simultaneously for both basins, the analogs underestimate the amplitude of the observed variability (Figs. 2c,d) somewhat. In general, the 6-year-long analogs track the observed variability better than the 20-year-long analogs, which in particular underestimate the PMV amplitude. In Fig. SM8, we show how many variability analogs each CMIP5 model contributes.

3. Evaluation

To address the risk that the chosen variability analogs do not represent a physically consistent depiction of the Atlantic and Pacific variability, we evaluate the intervariable relationships in the selected segments for consistency with the known underlying mechanisms for PMV and AMV. Since not all variables are available for all CMIP5 climate models, we use reduced model ensembles (see Table SM1 for the included models) for the evaluation.

a. Air temperature patterns

To verify whether the local temperature trends from the variability analogs associated with a change in AMV and PMV correspond well with observations, we look at the surface air temperature (SAT) trend maps for four periods of large changes in AMV and PMV, respectively (Fig. 3). For the Atlantic we choose the periods from 1921 to 1940 (AMV increases) and from 1955 to 1974 (AMV decreases) and for the Pacific the periods from 1931 to 1950 (PMV decreases) and from 1971 to 1990 (PMV increases). For this, we use 20-year-long variability analogs and smooth the AMV and PMV time series prior to selecting the analogs since we need continuous analogs over the periods examined.

For the period from 1921 to 1940, we find considerable (i.e., at least 75% of the selected variability analogs...
agree on the sign of the trend) warming that is centered in the North Atlantic (Fig. 3c) and extends over the Arctic. While the Northern Hemisphere is warming, the Southern Hemisphere is less affected by AMV, where the surface air temperature trends are either close to zero or negative (i.e., a dipole structure with no warming or slight cooling in the South Atlantic and the strong warming in the North Atlantic). The CMIP5 models, however, do not fully capture the tropical arm of the North Atlantic variability that is seen in observations (Yuan et al. 2016). An almost opposite temperature trend pattern can be seen from 1955 to 1974 (Fig. 3d), when the North Atlantic cooled. The cooling is strongest along Greenland, possibly amplified by growing sea ice in this region.

For the PMV, the three previously described centers of action in the Pacific are clearly visible during both periods of Pacific cooling and warming (Figs. 3a,b). Note, however, while the Henley index was constructed as such, we do not prescribe the SST variations in each region but only the difference between the three regions. We confirm that the observed and simulated mean SST time series are similar in each region. The variability is closest matched in the equatorial Pacific but somewhat underestimated in the North and South Pacific (Fig. SM9). The analogs also capture the temperature trends in the Pacific sector of the Southern Ocean—where we do not prescribe SSTs—associated with ENSO (e.g., Folland and Salinger 1995). Biases in the simulated Pacific variability, however, remain. The climate models tend to overestimate ENSO SST variability in the western equatorial Pacific compared to observations (Bellenger et al. 2014). PMV also affects temperatures over northern North America and most of the variability analogs indicate an imprint in a band around the tropics.

b. Atlantic Ocean

Variability in North Atlantic SSTs is believed to be related to the strength of the AMOC, which controls the amount of heat transported northward (Delworth et al. 1993). Hence, we examine the behavior of the AMOC in the selected variability analogs. We define the AMOC as the maximum volume transport streamfunction in the North Atlantic at 30°N and below 500-m depth. We find a close temporal agreement between the AMOC and AMV time series (Fig. 4a), where periods of positive SST anomalies in the North Atlantic are associated with increased AMOC strength (Tandon and Kushner 2015) confirming that the method of variability analogs allows us to study relations across different variables. The peak-to-peak amplitude of the AMOC variations is around 1.5 Sverdrups (Sv; 1 Sv = 10^6 m^3 s^-1). Our results show a strengthened AMOC during the last decades, in agreement with Tett et al. (2014). Further, we find a weak AMOC in the 1970s and 1980s, consistent with Zhang (2007). Observational constraints on the AMOC are, however, weak and show conflicting results (Buckley and Marshall 2016).

An AMOC weakening leads to increased ocean heat content (Gregory 2000; Knutti and Stocker 2000; Raper et al. 2002) due to enhanced stability in the high-latitude ocean. The increased stability reduces heat loss from the high-latitude ocean (i.e., the ocean retains more heat; Raper et al. 2002). To verify whether this is captured by the variability analogs, we integrate the potential water temperature of different ocean basins (cf. Fig. SM10 for a map of the boundaries of the ocean basins) and different depths and multiply with constant water density and specific heat capacity of seawater [constants taken from Hobbs et al. (2016)]. For the upper 300 m of the Atlantic, ocean heat content varies in phase with North Atlantic SSTs (i.e., the AMV), consistent with increased northward heat transport. At greater depths, the ocean heat storage is, however, reversed, and an increased AMOC strength is associated with less ocean heat storage (Fig. 4b).

c. Pacific Ocean

To evaluate the Pacific variability, we study the relationship between the PMV and the Aleutian low. The multidecadal variability in the strength of the Aleutian low during the twentieth century covaries with the North Pacific manifestation of the PMV (Deser et al. 2004; Newman et al. 2016). Using the variability analogs, we define the strength of the Aleutian low as the anomalous and standardized sea level pressure in the boreal winter in the North Pacific (Deser et al. 2004). In agreement with observations, we find a deepened Aleutian low (i.e., lower sea level pressure) during a positive PMV period and the opposite during a negative PMV period (Fig. SM11).

Pacific variability also influences the ocean heat storage. A negative PMV is associated with reduced ocean heat storage in the upper few hundred meters of the Pacific Ocean and increased heat storage beneath due to an anomalous strengthening of the subtropical Pacific cells during such periods (Meehl et al. 2011; Nieves et al. 2015). In agreement, the selected variability analogs show reduced ocean heat content in the upper 150 m and increased ocean heat content in the 150–300-m layer during a negative PMV phase and the opposite during a positive PMV. Below 300 m, the ocean heat content variability is small in the Pacific (Fig. 4c).

d. Combined Atlantic and Pacific variability

Among others, McGregor et al. (2014), Kucharski et al. (2015), Zanchettin et al. (2016), and Barcikowska et al. (2017) found that SST variability in the Atlantic...
might create a transbasin sea level pressure gradient to the Pacific and accordingly modify the Walker circulation. A change in the Walker circulation in turn influences SSTs in the equatorial Pacific. Such a physical connection would invalidate method (ii) and we would overestimate the imprint of multidecadal internal variability on global climate trends. To test this, we select variability analogs for the Atlantic and examine how the Pacific responds, and vice versa.

We do not see an imprint of the Atlantic onto the Pacific or of the Pacific onto the Atlantic (Figs. SM12 and SM13). This may be related to a too-weak or missing tropical arm of the North Atlantic variability in CMIP5 models (Yuan et al. 2016). This tropical branch appears to be crucial for the Atlantic–Pacific connection (Chikamoto et al. 2016), and this might explain why such a connection is seen in assimilation experiments where the Atlantic temperature field is explicitly prescribed. Further, the time lag with which the Atlantic forces the Pacific variability might differ between climate models and thereby suppress the multimodel signal. Alternatively, the connection between the two basins may indeed be weak, and apparent relationships could be partly a result of the particular short time period that we happen to observe.

To further examine if AMV and PMV are linearly additive, we compare the 6-year-long analogs for which we sample the Atlantic and Pacific separately with short analogs (3–6 years; everything else as described in the methods section) for which AMV and PMV are matched simultaneously. By searching for short analogs, we achieve good agreement between the analogs and the observed variability also for the approach of simultaneous selection. The combined AMV and PMV imprint on global air temperatures is similar for methods (i) and (ii), suggesting that we can search for analogs for the two basins separately and then add the two patterns linearly (see Fig. SM14). Therefore, we will select analogs separately for the Atlantic and Pacific in the following.

4. Results and discussion

To quantify AMV and PMV influence on observed global temperature increase, we examine the changes in global
near-surface air temperatures of the selected variability analogs. As we select the analogs separately for the Atlantic and the Pacific, we are able to study both the imprint of AMV and PMV on global temperatures independently.

First, we examine whether it matters which AMV and PMV indices are used. To quantify the difference, we use each index and search for analogs to then compare the imprints on global surface temperatures (Fig. 5a for AMV and Fig. 5b for PMV).

The AMV indices agree on the general progression of unforced Atlantic SSTs (Fig. 1a), and accordingly, the estimated imprint on global mean air temperatures has a similar evolution over time (cf. Fig. 5a). During the early twentieth century when the Trenberth and Shea AMV shows the smallest signal (Fig. 1a), this also converts to the weakest imprint on global mean temperatures. The AMV indices again differ significantly in the period after the 1980s, and the question arises if there has been an AMV contribution to the observed global warming.

The Enfield definition shows the strongest AMV increase and thereby also the largest global temperature increase from the 1980s to the 2000s attributable to AMV: the linear trend between 1981 and 2005 is $0.08^\circ C$ decade$^{-1}$, which corresponds to roughly 38% of the observed global warming over the same period (cf. Table 2). We take Hadley Centre/Climatic Research Unit, version 4.5 (HadCRUT4.5), as the observational dataset (Morice et al. 2012). These numbers decrease to 0.07$^\circ C$ decade$^{-1}$ and 33%, respectively, when we use Enfield and Cid-Serrano instead. It is further reduced to 0.06$^\circ C$ decade$^{-1}$ and 30%, respectively, when we use Trenberth and Shea’s method. Applying the Frankcombe scaling approaches, the contribution is down to 0.03$^\circ$ and 0.05$^\circ$C decade$^{-1}$ (i.e., 15% for single-factor scaling and 25% two-factor scaling, respectively). With the three-factor scaling it is even less than 0.01$^\circ$C decade$^{-1}$ (4%). This spread of a factor of 10 solely due to different AMV definitions highlights the relevance of the index used (Mann et al. 2014; van der Werf and Dolman 2014).

Previous studies that found a large AMV contribution to recent warming often rely on the Enfield AMV (Tung and Zhou 2013; Zhou and Tung 2013; Chylek et al. 2014a). Since a linear or quadratic fit is not adequate to fully remove the forced signal from the Atlantic SSTs, it is likely that the Enfield and the Enfield and Cid-Serrano AMVs overestimate the AMV increase in recent decades and thereby also overestimate the contribution of AMV to observed warming (Mann et al. 2014; Frankcombe et al. 2015; Steinman et al. 2015; Tandon and Kushner 2015). The Trenberth and Shea AMV, for which global mean SSTs are subtracted from the average North Atlantic SST, relies on the assumption that there is no strong influence of AMV on global temperatures. However, our results suggest that this assumption is not fulfilled (cf. Fig. 5a). Therefore, the Trenberth and Shea AMV might not represent the true unforced signal either.

The estimate of the PMV contribution to global air temperature variability also depends on the chosen index (Fig. 5b), but overall Mantua and Henley lead to similar imprints onto global mean temperatures that scale directly with the PMV indices themselves: a warming imprint in the early twentieth century, cooling in the midcentury, and then warming again until the 1990s. The Henley index produces a global temperature imprint that is slightly larger than that for the Mantua PMV definition. The PMV imprint on temperatures reached a local minimum in the first decade of the twenty-first century, contributing to the warming hiatus, in
agreement with earlier findings (Kosaka and Xie 2013; Steinman et al. 2015). Since then, the PMV contribution to global temperatures is increasing rapidly, fueling the record temperatures in 2014 and 2015 (Thoma et al. 2015). PMV offsets part of the AMV-induced warming from 1981 to 2005 as its warming contribution to global temperatures decreased over this period. This is the case for both PMV indices (Table 2): 0.01°C and 0.02°C decade⁻¹ (4–11%) for Mantua and Henley, respectively.

In Figs. 6a and 6b we compare the global and Northern Hemisphere temperature effects of AMV and PMV and show their combined temperature signature. Here, we use the Henley and Frankcombe two-factor scaling as the best estimates, but we construct uncertainty intervals that contain both the climate model uncertainty (from subsampling the CMIP5 archive) as well as the uncertainty from isolating the internal variability (therefore we treat all indices and their combinations as equally likely). Consequently, the uncertainty bands are wider than those shown in Figs. 5a and 5b, which only include the structural model uncertainty. Globally and for the Northern Hemisphere, the peak-to-peak amplitude of their combined effect is approximately 0.15° and 0.23°C, respectively, when the peak-to-peak amplitude is measured as the difference between the 5th and 95th percentiles of the smoothed time series. The AMV contributes up to 0.12° and 0.21°C for the global and Northern Hemisphere temperatures, respectively, while the PMV contributes up to 0.09°C on both spatial scales. On multidecadal time scales, we find that AMV is the main contributor to the multidecadal internal variability in Northern Hemisphere temperatures during most of the twentieth century. This is also true in global air temperatures, though less clear since the Henley PMV imprint on temperatures is similar on both the global and Northern Hemispheric scale, while the AMV imprint is larger for the Northern Hemisphere than globally. The overall shape of the combined AMV and PMV effect on multidecadal air temperatures is therefore determined by AMV during

![Fig. 6. (a) The imprint of the Henley PMV (blue), the Frankcombe two-factor scaling AMV (orange), and their combined effect (gray) on global SATs. (b) As in (a), but for Northern Hemisphere mean SATs. The shading indicates the 90% range related to both structural model uncertainty and uncertainty related to isolating the unforced variability in the Atlantic and Pacific Oceans (as represented by the various indices of AMV and PMV). The Henley PMV and Frankcombe two-factor scaling AMV are therefore not in the center of the uncertainty ranges. The time series have been smoothed using LOESS.](attachment:image)
most of the twentieth century and PMV modulates this imprint.

To quantify the multidecadal imprint of AMV and PMV on the measured warming, we use an observational twentieth-century land–ocean temperature record (HadCRUT4.5) and subtract the estimated effect of internal variability from global (Fig. 7a) and Northern Hemisphere (Fig. 7b) temperatures. Again, we use the combined temperature imprint of the Henley PMV and the Frankcombe two-factor scaling AMV as the best estimate but construct an uncertainty interval that combines the uncertainty from structural climate model differences and from the isolated internal variability. We assume linear additivity of the forced and unforced climate variations. This results in temperature time series (HadCRUT4.5 adjusted) where multidecadal variability is (partly) removed and therefore better represents the forced signal. It has to be noted that the HadCRUT4.5 record is not spatially complete and blends SSTs over the oceans with near-surface air temperatures over the continents. The SSTs tend to warm slightly slower than the air temperatures above (Cowtan et al. 2015; Richardson et al. 2016). For the variability analogs, we do not account for these two effects, and we might therefore slightly overestimate the amplitude of AMV and PMV within the observational record. Some multidecadal variability remains in the adjusted observations, possibly because the method does not remove all variability (e.g., of the Southern or Indian Oceans), because of uncertainties in the interpolated SST dataset we use to construct the AMV and PMV indices from, and because part of the variability is forced (e.g., by aerosols, solar variability, and volcanic eruptions).

![Fig. 7. HadCRUT4.5 (a) global and (b) Northern Hemisphere mean SAT anomalies during the twentieth century (yellow). The combined effect of AMV and PMV is shown in gray and the mean temperatures adjusted for internal variability are shown in orange (‘HadCRUT4.5 adjusted’). The linearly detrended HadCRUT4.5 temperature change for (c) global and (d) Northern Hemisphere SATs (yellow) along with the estimated internal variability signal. The shading indicates the 90% range related to both structural model uncertainty and uncertainty related to isolating the unforced variability in the Atlantic and Pacific Oceans. All time series are smoothed as described in the methods section, and are given relative to the reference period of 1900–2000.]
The overall warming from 1900 to present day is not strongly influenced by multidecadal internal variability centered in the Atlantic and Pacific Oceans (cf. Figs. 7a,b). We calculate the global warming signal and the AMV and PMV contributions as the difference between the period from 1995 to 2005, with 1890–1910 as the reference period (for the Mantua PMV we use 1900–10). Less than 10% of the measured 0.79°C global warming between these two periods can be attributed to AMV and PMV (range across all possible AMV and PMV indices combinations is from −2% to 9%). For the Northern Hemisphere, we find a wider range from −1% to 14% contribution to the 0.85°C warming. For the second half of the twentieth century (1940–60 reference period), we find the maximum AMV and PMV contribution to be ±10% (from −10% to 9% for global temperatures and from −16% to 10% for Northern Hemisphere temperatures).

While their contribution to the long-term warming trend is minor, both AMV and PMV contribute to periods of accelerated and reduced warming. After correcting for internal variability, the early twentieth-century warming is reduced for both global and Northern Hemisphere temperatures. After accounting for internal variability, the mid-twentieth-century observed warming is also reduced, and the slight cooling trend largely vanishes. This is more pronounced for Northern Hemisphere temperature. The onset of warming in the 1970s starts earlier after subtracting AMV and PMV indicating that internal variability masked the warming for some years. This becomes clearer when the observed temperature increase is linearly detrended (Figs. 7c,d). For Northern Hemisphere temperatures, there is a relatively good agreement between the linearly detrended temperatures and the isolated AMV and PMV internal variability signal in terms of amplitude and phase, indicating a role of internal variability in the periods of halted and accelerated warming.

We do not, however, find a linear or monotonically accelerating residual global warming during the twentieth-century global warming after removing internal variability associated with the Atlantic and Pacific Oceans as earlier studies indicated (Swanson et al. 2009; Tung and Zhou 2013; Zhou and Tung 2013), but the variability-adjusted global warming rate still changes markedly around 1940. For the Northern Hemisphere, the trend change from early to mid-twentieth-century warming is less clear after accounting for internal variability. Other factors, such as changes in the aerosol loading (Wilcox et al. 2013), contribute to this period of paused warming in the mid-twentieth century too. However, accounting for AMV and PMV reduces the potential role of aerosols and other factors in modulating twentieth-century temperature trends somewhat.

5. Conclusions

In this study, we describe, evaluate, and apply a method of using variability analogs to quantify the contribution of multidecadal internal variability in the Atlantic and Pacific Oceans to the observed twentieth-century global and local warming. The variability analogs are simulated segments from unforced control simulations where the modeled internal variability agrees closely with what is observed. Other approaches that try to link AMV and PMV with air temperatures exist, most notably, regression approaches and climate model experiments with partial data assimilation. Contrary to our approach, the regression approach is not capable of resolving physical mechanisms, while assimilation does not conserve energy, comes at high computational costs, and is only available for a few models. Our method allows us to quantify the AMV and PMV contribution to global temperature variations at very small computational cost and in a physically consistent way (to the degree that the models are capable of resolving those modes of variability).

We use relatively short 6-year-long variability analogs since the AMV and PMV amplitudes are increasingly underestimated with increasing length of the analogs. Hence, we assume that higher-frequency internal variability projects similarly onto global temperatures as multidecadal variability. When increasing the length of analogs up to 80 years, we find that the sensitivities of how strongly air temperatures respond to these two modes of variability remain relatively stable (approximately ±20%). Wang et al. (2017) found a larger time-scale dependency in the global air temperature sensitivity when regressing surface temperatures against interannual versus interdecadal tropical Pacific variability. This might be related to the choice of index used to describe the variability or the way we weight the models.

The analogs confirm the previously established link between AMV, AMOC, and ocean heat storage. During a positive AMV period, the AMOC tends to be stronger than usual and less heat is stored in the deep Atlantic. In the Pacific, the analogs affirm the previously established relationship between PMV and the magnitude of the Aleutian low and between PMV and ocean heat storage. The anomalous ocean heat content of the upper Pacific follows the PMV (increased during a positive PMV period), but beneath the ocean heat content signal is reversed. The influence of AMV and PMV on global air temperatures is approximately linearly additive.

Previous studies found conflicting results concerning the relevance of AMV to the observed late twentieth-century warming, ranging from a marginal to a substantial warming contribution over the last decades. We
show that these differences can largely be explained by the index used to describe AMV: Using the period from 1981 to 2005 as an example, the combined contribution of AMV and PMV ranges from −0.01° to 0.07°C decade⁻¹, solely due to the choice of indices. The historical evolution of Atlantic and Pacific sea surface temperatures is the sum of internal variability and a forced signal. If the forced signal is not adequately removed, then the resulting estimate of AMV/PMV contribution to observed climate change will be biased high (if not the full forced signal is removed) or biased low (if in addition to the forced signal also part of the internal variability is removed). A careful choice of index is therefore central, and further developments and refinements of these indices are necessary.

We find that internal variability contributes more to the Northern Hemisphere temperatures than when averaged over the whole globe, which is expected as both the AMV and PMV are geographically largely focused on the Northern Hemisphere. The cooling seen in observations in the mid-twentieth century is reduced after accounting for the AMV and PMV contributions, and the multidecadal variations in the rate of Northern Hemisphere warming are dampened. During most of the twentieth century the role of AMV in modulating temperatures on a multidecadal time scale is more important than the PMV, especially for the Northern Hemisphere, in agreement with Steinman et al. (2015). This holds for multidecadal variability, but for higher-frequency variability the Pacific plays a larger role (Cheung et al. 2017). Recent studies (Kosaka and Xie 2016; Meehl et al. 2016) find improved agreement between observed and modeled twentieth-century global warming rates after accounting for PMV. The effect, however, is rather weak in some cases (Meehl et al. 2016), and our results suggest that such an analysis cannot be conclusive without considering the AMV. With a contribution of less than 10%, the measured global warming during the second half of the twentieth century is not much affected by Atlantic and Pacific variability, underlining that most of this observed warming is caused by anthropogenic forcings. Internal variability not located in these two basins might also play a role in modulating global mean air temperatures, and future research needs to clarify this contribution.

Acknowledgments. We are grateful to Leela Frankcombe and Benjamin Henley for providing data. We thank the three anonymous reviewers for their helpful comments. We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table SM1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison (PCMDI) provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. CMIP5 data are provided by the PCMDI (http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html).

REFERENCES


Unauthenticated | Downloaded 09/18/23 07:02 AM UTC


