https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

The Influence of Atlantic Multidecadal Variability on European Summer Climate: Competing Mechanisms and Implications for Prediction

Ned C. Williams[1,2], Wolfgang A. Müller[1], and Joaquim G. Pinto[3]

[1]Max Planck Institute for Meteorology, Hamburg, Germany
[2]Met Office Hadley Centre, Exeter, UK
[3]Institute of Meteorology and Climate Research - Troposphere Research (IMKTRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Correspondence: Ned C. Williams (ned.williams@metoffice.gov.uk)

Abstract. Skilful predictions of European summer climate are increasingly relevant due to an increasing probability of temperature extremes, but prediction skill beyond the forced trend has so far proven limited. Atlantic Multidecadal Variability (AMV), characterised at the surface by North Atlantic sea surface temperatures (SSTs), is both active and predictable during boreal summer, and previous studies have linked it to surface impacts in Europe. Current understanding largely relies on the relatively short observational record of decadal variability and the predictability of impacts and associated mechanisms are poorly studied. In this study, single model large ensemble historical and decadal hindcast simulations using the MPI-ESM-LR model are used to understand the role that AMV plays for North Atlantic-Europe sector climate prediction. It is found that strong AMV-associated SST anomalies in the subpolar gyre region are better represented in the initialised hindcasts than in the uninitialised historical ensemble, and they are highly predictable at lead years 1–7. The observed cyclonic response to positive AMV in the extratropical North Atlantic is not present in historical simulations, but it is found to be predictable in decadal hindcasts, although with underestimated amplitude. The hindcast pressure anomaly nonetheless skilfully predicts observations and highlights a potential role for AMV in the yet-unsolved 'signal-to-noise paradox'. The upper tropospheric (200 hPa) geopotential height response to AMV is analysed and it is found to differ in reanalyses and models. Further investigation reveals a high frequency component relating to tropical SST anomalies and resembling a Rossby wave train emanating from the Caribbean, and a low frequency component relating to the surface level response, with an imbalance between the two mechanisms in models due to the weak surface response.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

1 Introduction

Decadal climate predictions have gained importance in recent years for their potential added value for decision making in economy and society (e.g., Meehl et al., 2009, 2014). There is an increasing body of literature focusing on skilful decadal predictions for Europe, not only regarding basic variables (Smith et al., 2020) but also for extremes and user-oriented variables (Moemken et al., 2021). Predicting decadal variability of European summer climate is increasingly vital given the increased risk of extreme heat due to anthroprogenic climate change (e.g. Rousi et al., 2023). Atlantic Multidecadal Variability (AMV)—associated with low frequency variability of North Atlantic sea surface temperatures—has been found to be highly predictable on decadal timescales and is known to play a role in European summer climate (Knight et al., 2006; Sutton and Hodson, 2005; Zhang et al., 2019). In this study, large ensemble historical and decadal hindcast simulations are used to disentangle two distinct atmospheric responses to AMV: one associated with diabatic heating due to subpolar SST anomalies (Ghosh et al., 2017), and another associated with the upper tropospheric response to tropical SST anomalies, resulting in a source of Rossby waves (Sutton and Hodson, 2005, 2007; Terray and Cassou, 2002). The capability of decadal hindcasts to skilfully and reliably predict the North Atlantic-Europe (NAE) summertime atmospheric response to AMV is assessed, taking into account the respective influence of tropical and extratropical SSTs.

Whilst initialised climate prediction for the wintertime North Atlantic-Europe (NAE) sector has proven skilful on a range of timescales (e.g. Athanasiadis et al., 2017; Scaife et al., 2014; Smith et al., 2020), summer prediction skill remains more elusive, but with episodic 'windows of opportunity' for skilful predictions under certain conditions (e.g. Dunstone et al., 2023; Wallberg et al., 2024). The increasing risk of extreme heat in Europe during summer due to anthropogenic influence means that predictions of summer climate are particularly relevant (Seneviratne et al., 2021; Suarez-Gutierrez et al., 2020). As a result, it is important to understand the impacts of relevant predictable drivers such as AMV, as well as the present capability of models in simulating them.

Positive AMV phase sea-surface temperature (SST) signals include a warmer North Atlantic subpolar gyre (SPG), and—lagging behind the SPG response—a band of warm SSTs in the tropical North Atlantic (and similar but opposite during the negative phase; Zhang et al., 2019). The atmospheric response to diabatic heating in the extratropics, such as the heating associated with the SPG component of AMV, is a shallow cyclonic circulation one quarter-wavelength downstream of the heating, which acts to cool the heating region (Hoskins and Karoly, 1981). This diabatic heating response to AMV has been found to be robust and leads to increased surface temperatures in central Europe by advecting warm air polewards (Ghosh et al., 2017). AMV-associated tropical SST anomalies lead to upper level divergence anomalies and hence a source of Rossby waves which can propagate to the extratropical NAE sector. This Rossby wave response has been associated with increased temperatures in Europe (Sutton and Hodson, 2005).

The tropical and extratropical components of AMV are strongly coupled on decadal timescales and therefore so are their impacts (Zhang et al., 2019). The time period for which high quality observational data, particularly for non-surface fields, exists is short relative to the timescale of AMV and therefore it is can be difficult to separate the tropical and extratropical contributions to AMV impacts using observations or reanalyses alone. Large ensembles of model simulations (e.g. Maher et al.,

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

2019) allow for a larger sample of AMV variability and resulting European climate impacts. Furthermore, skilful predictions of NAE sector atmospheric variability require large ensembles, to not only separate the predictable 'signal' and unpredictable 'noise' variability which exist in reality, but also as prediction systems tend to underestimate the signal-to-noise ratio in the NAE region (Scaife and Smith, 2018; Smith et al., 2020). This issue is known as the 'signal-to-noise paradox' due to the counterintuitive consequence that the correlation skill between the ensemble mean and observations is higher than the average correlation between the ensemble mean and individual members (Scaife and Smith, 2018), and there is evidence of its presence during boreal summer (Dunstone et al., 2023; Ossó et al., 2020). Using a single model large ensemble prediction system ensures that differences between members are result of chaotic 'noise' variability rather than model differences, avoiding the risk of a reduced signal-to-noise ratio arising from inter-model spread (Weigel et al., 2008).

The aim of this study is to understand the capability of a large single model ensemble decadal hindcast in simulating the summertime atmospheric circulation response to AMV. The data and methodology used in this study are presented in Section 2. AMV SSTs and the associated atmospheric response on 7 year timescales are analysed and compared in observations/reanalyses and model experiments in Section 3. Section 4 examines the timescale dependence of the upper tropospheric response. Section 5 summarises and contextualises the results of this study.

2 Data and Methods

This study uses an 80 member decadal hindcast ensemble using the low resolution version of the Max-Planck-Institute for Meteorology (MPI) Earth System Model (MPI-ESM-LR; Mauritsen et al., 2019). The ensemble consists of an original 16 member ensemble (Brune and Baehr, 2020; Hövel et al., 2022) and a more recent addition of 64 further members (Krieger et al., 2022). Members are initialised every year in November from 1960 to 2019 inclusive and are run for a minimum of 10 full calendar years beyond the initialisation date. In addition, the original 16 members have been extended to 20 years after each initialisation date (Düsterhus and Brune, 2024). Lead year 1–7 means are used except where otherwise stated. 50 historical runs from the MPI Grand Ensemble (Maher et al., 2019; Olonscheck et al., 2023), also using MPI-ESM-LR and running from 1850 to 2014, are used to provide independent long-term realisations of internal decadal variability.

MPI-ESM-LR simulations are compared to the ERA5 reanalysis (Hersbach et al., 2020) for all fields, the HadISST1 (HadISST; Rayner et al., 2003) dataset for SSTs, the HadSLP2 (HadSLP; Allan and Ansell, 2006) dataset for mean sea level pressure (MSLP), and the NOAA 20th Century Reanalysis version 3 (NOAA 20CR; Slivinski et al., 2019) for geopotential height. NOAA 20CR is forced by HadISST data and assimilates surface pressure observations. For ERA5, only data from the hindcast period is used in order to aid comparison with the hindcasts, and because this still covers most of the ERA5 period (1961–2023 instead of 1940–2023). HadISST data extends from 1870 to 2023, and both the full period and hindcast period are analysed. HadSLP data extends from 1850 to 2012, and the periods common to HadSLP and HadISST (i.e. 1870–2012) and common to HadSLP and the hindcasts (1961–2012) are analysed. NOAA 20CR data extends from 1806 to 2015, and the periods common to NOAA 20CR and HadISST (1870–2015) and common to NOAA 20CR and the hindcasts (1961–2015) are analysed.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

Throughout this study, all fields are taken as June–August (JJA) means. An AMV index is defined as the area-weighted average SST anomaly between 0–60°N and 280–360°E. Positive and negative phases of the AMV are defined as when the AMV index is greater or less than 0, respectively; results are insensitive to the choice of threshold and we found no evidence of significant asymmetry between phases in models or observations (not shown). Tropical AMV and Extratropical AMV indices are defined using the part of the AMV region south and north of 30°N respectively. Throughout this study, impacts of AMV are studied through composites of summers with a negative AMV index subtracted from those with a positive AMV index. Therefore, the results shown reflect how positive AMV summers differ from negative AMV summers, and for simplicity the text of this paper will generally refer to the conditions during positive AMV events. However, no robust asymmetries between the phases were found (not shown) and so it is appropriate to reverse the results when considering impacts of the negative phase.

The SST variability is dominated by the anthropogenic forced trend, which cannot be expected to lead to the same atmospheric response as regionally constrained North Atlantic SST variability. Therefore, it is desirable to remove the globally coherent forced trend. The use of a single model large ensemble of historical simulations allows for the externally forced trend to be removed, by regressing out co-variability with the ensemble mean Global Mean SST (GMSST; Deser and Phillips, 2023). This method is used here, with the ensemble mean GMSST timeseries extended beyond the end of the historical period using future simulations from MPI-ESM-LR under the SSP245 scenario where necessary. GMSST has been calculated as the area-weighted average SST between 60°S and 60°N. Although this method was developed in Deser and Phillips (2023) for the specific purpose of AMV, it is equally applicable to other variables and so it is used consistently in this study.

In the study of decadal variability and when taking rolling means, it is clear that consecutive summers are not independent and that their interdependence cannot be neglected when estimating statistical significance. To calculate statistical significance of skill maps, an existing method to calculate 'effective sample size' is used in order to estimate degrees of freedom (Guemas et al., 2014). Autocorrelations with successive lags are used in this method and several others used to calculate effective sample size (e.g. Trenberth, 1984), but with increased lag time, autocorrelation values are increasingly uncertain. This method reduces the series of autocorrelation values used to a single parameter α, where the timeseries is assumed to be a sample of an AR(1) process defined by α, with α computed by minimising the mean square error between theoretical AR(1) and sample autocorrelations, where the errors are inversely weighted by the autocorrelation lag. Details of the implementation in this study can be found in the Supplementary Information.

The most prominent surface feature in the summer atmospheric response to AMV found in Ghosh et al. (2017) is a negative pressure anomaly over the North Atlantic centred west of Great Britain and Ireland, which is consistent with simple linear theory for the atmospheric response to extratropical diabatic heating (Hoskins and Karoly, 1981). This is also the main feature of the 'East Atlantic Pattern' mode of variability in the North Atlantic, including during summer (Barnston and Livezey, 1987; Cassou et al., 2005; Wulff et al., 2017). For this reason, we define East Atlantic (EA) MSLP as the area-weighted mean MSLP anomaly between 45–60°N and 330–350°E. This region is shown in Figure 2 (c). The same region is also used to define East Atlantic 200 hPa geopotential height anomalies.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

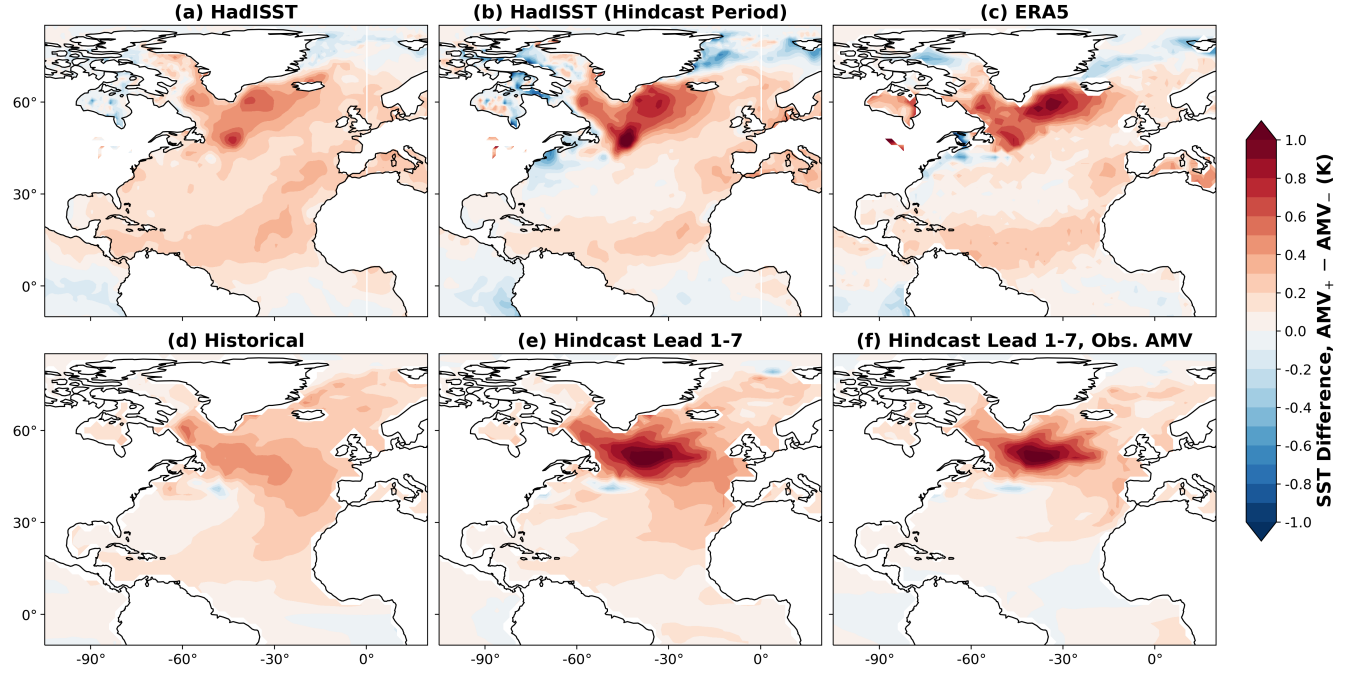

Figure 1. Patterns of AMV SSTs in observations, reanalysis and model experiments. Composite difference between SSTs during positive and negative AMV using SSTs and AMV index from (a) HadISST; (b) HadISST during the period of the decadal hindcasts only; (c) ERA5 during the hindcast period; (d) historical runs using MPI-ESM-LR; and (e) MPI-ESM-LR decadal hindcasts. (f) as with (e), but using the ERA5 AMV index rather than that of the hindcasts.

Signal-to-noise errors associated with the signal-to-noise paradox are typically diagnosed using the 'Ratio of Predictable Components' (RPC; Eade et al., 2014) metric, defined in Scaife and Smith (2018) as:

$$\mathrm{RPC} = \frac{r_{mo}}{r_{mm}} \tag{1}$$

where r_{mo} is the correlation between the ensemble mean and observations (i.e. the correlation skill), and r_{mm} is the mean correlation between the ensemble mean and individual ensemble members. This quantity should ideally be one, but underestimated predictable signals or overestimated noise variability will lead to values greater than one. In this study, the quantity r_{mi} is defined as the correlation between the ensemble mean and an individual ensemble member i, so that r_{mm} is the mean of r_{mi} for all members.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

3 Summer Atmospheric Response to AMV

Composites differences in positive and negative AMV phase North Atlantic SSTs from different data sources are shown in Figure 1. The 'horseshoe' pattern of SSTs associated with AMV is clearly present in all observational/reanalysis datasets (Figure 1 (a–c)); there is a strong positive signal in the SPG region south of Greenland and a band of positive SST anomalies in the tropical North Atlantic, along with warm anomalies at mid-latitudes in the eastern North Atlantic whilst in the west at mid-latitudes, anomalies are relatively small. The strongest anomalies are in the SPG region and this is particularly true during the hindcast period.

The historical large ensemble shows a similar pattern in composites defined using model AMV (Figure 1 (d)), although the SPG response is considerably weaker and with the strongest anomalies centred further south and east compared to observations/reanalysis. The large ensemble hindcast (Figure 1 (e)) shows a considerably stronger SPG response which is comparable to hindcast period observations/reanalysis, although the tropical signal is weak. This implies that initialisation has constrained the SPG signature of AMV towards that of observations. By using ERA5 to define AMV phases for the hindcast (Figure 1 (f)), the predictability of AMV SSTs can be assessed. The strong SPG signal remains, suggesting that it is highly predictable (consistent with previous studies, e.g. Borchert et al. (2021)). However, the tropical signal vanishes, suggesting that AMV-driven tropical SST anomalies are poorly predicted at lead 1–7.

Figure 2 shows composites of MSLP for different AMV phases using different data sources. In HadSLP during both the full period and period common to the hindcasts (Figure 2 (a) and (b) respectively), the strongest response is a negative anomaly west of Great Britain and Ireland, as previously found in Ghosh et al. (2017). This feature is also present in ERA5 (Figure 2 (c)), but a comparable positive anomaly exists to its north, resembling the summertime North Atlantic Oscillation (SNAO; Folland et al., 2009). In ERA5 and hindcast period HadSLP, there is a distinct band of negative MSLP anomalies in the tropical North Atlantic, which is a typical low level response to positive SST anomalies in the tropics; this response is weaker in the full HadSLP period.

In the historical MPI-ESM-LR large ensemble (Figure 2 (d)), the MSLP response to (model) AMV is very different to that of reanalysis/observations. The composite response is negligible in the extratropics where the observed response is strongest. The tropical response is comparable to observations, although it extends further northwards. In the large ensemble hindcast (Figure 2 (e)), the tropical response is weaker than in the uninitialised simulations, but a negative pressure anomaly comparable to that in Ghosh et al. (2017) and Figure 2 (a–c) is present, although notably weaker than observed. The weaker tropical response but improved extratropical response compared to the historical ensemble is consistent with the difference in SSTs found in Figure 1. When defining based on ERA5 AMV, the hindcast MSLP composite is consistent with the corresponding SST composite: the tropical signal vanishes but the extratropical signal remains: a cyclonic anomaly with a position consistent with observations, indicating the potential for predictability albeit with a model signal which is weaker than observed.

To assess predictability, correlation skill for hindcast MSLP relative to ERA5 (Figure 3 (a)) and HadSLP (Figure 3 (b)) was computed. Note that as HadSLP does not extend as far as ERA5 and the hindcasts, the years used in Figure 3 (a) and (b) are not identical, although most of the differences are present when using only the HadSLP/hindcast common period (not shown).

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

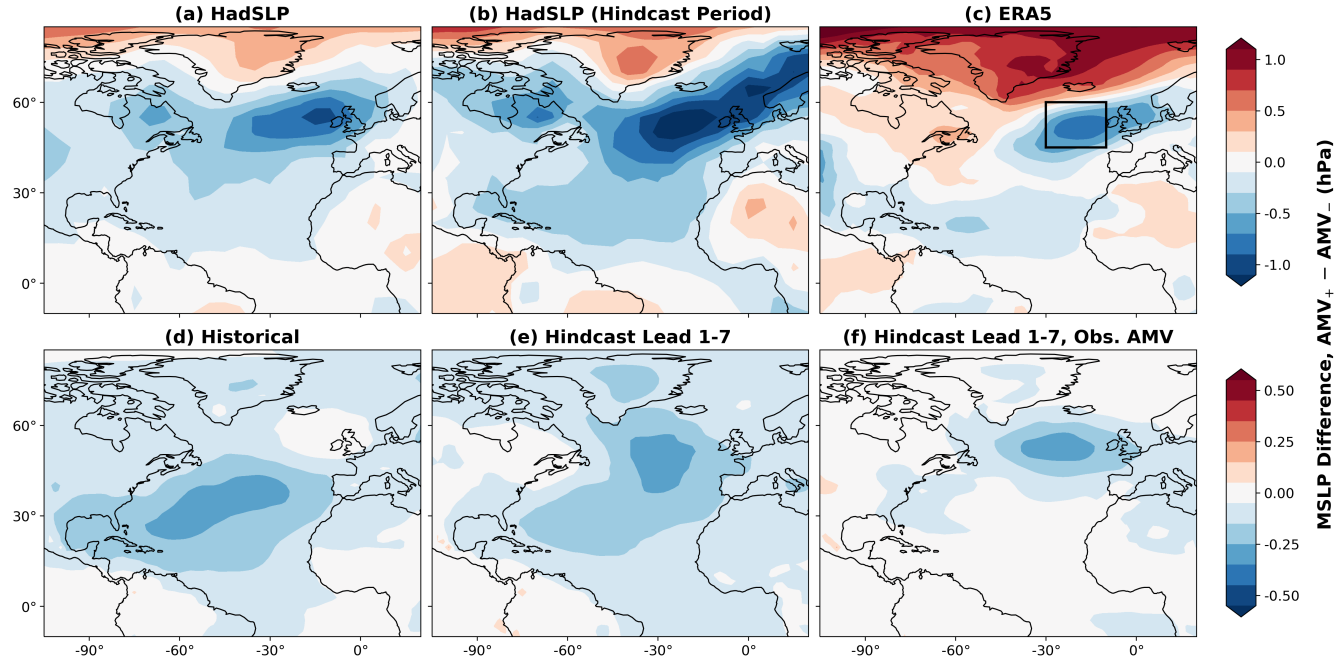

Figure 2. MSLP response to AMV in observations, reanalysis and model experiments. Composite difference in MSLP during positive and negative AMV using MSLP and AMV index from (a) HadSLP/HadISST; (b) HadSLP/HadISST during the common period of the decadal hindcasts and HadSLP only; (c) ERA5 during the hindcast period; (d) historical runs using MPI-ESM-LR and e) MPI-ESM-LR decadal hindcasts. (f) as with (e), but using the ERA5 AMV index rather than that of the hindcasts. The black box in (c) shows the region used to define East Atlantic MSLP. Note that (a–c) and (d–f) use different colour scales.

The highest skill levels are to be found off the west coast of Europe, where the negative MSLP response to AMV is found. This skill relative to ERA5 is highly significant in some regions. In the western North Atlantic, there are regions of negative skill, which has particularly high magnitude and is significant in ERA5. It is not clear if (and if so, how) this relates to AMV, but the high significance implies *potential* for predictability, that is nonetheless not realised by the MPI-ESM-LR hindcasts due to model error. In the tropics, MSLP skill is generally lower and differs depending on the observational reference used.

Although the hindcasts are capable of predicting EA MSLP, the response to AMV is nonetheless weak. Figure 4 (a) shows the histogram of regression slopes for sampled single member timeseries of EA MSLP with the ERA5 AMV index. Despite considerable spread, the regression slope for ERA5 EA MSLP is more negative/stronger than 99% of sampled hindcast slopes. An underestimated response to a predictable driver can contribute to the signal-to-noise paradox, which is examined in Figure 4 (b). The correlation between ensemble mean EA MSLP and that of sampled individual members r_{mi} is calculated and plotted as a histogram; the ensemble mean correlation skill r_{mo} computed relative to ERA5 is greater than 99% of all values of r_{mi}, and the ratio of predictable components (RPC) is above 2, suggesting that the reliability issues which define the signal-to-noise paradox are present with high significance. Finally, 4 (c) shows the joint distribution (as a 2-dimensional histogram) of the

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

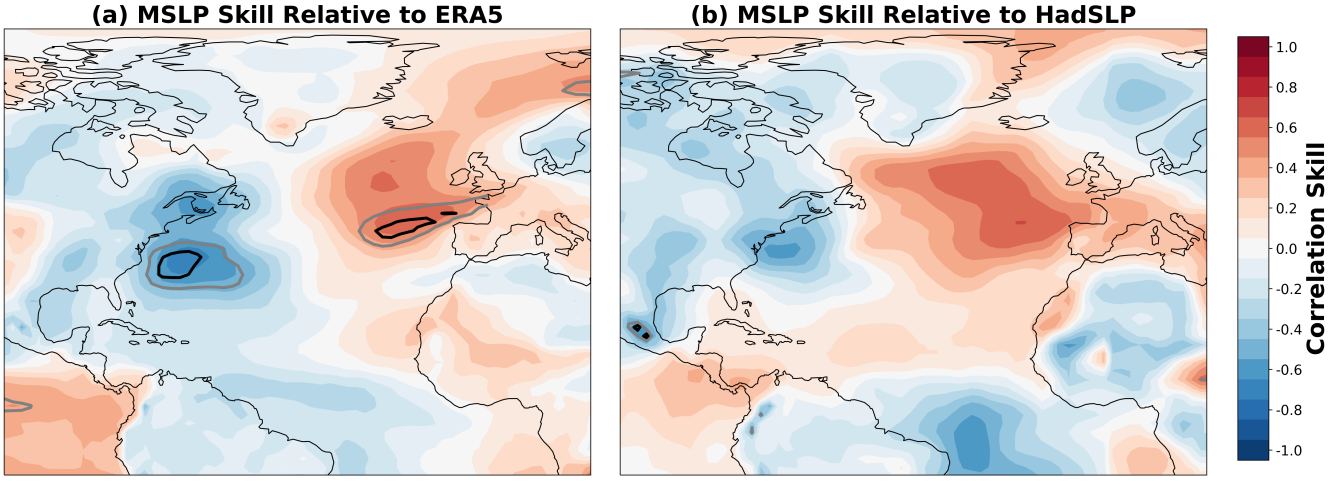

Figure 3. Decadal hindcast skill for MSLP. Temporal correlation between the decadal hindcast ensemble mean MSLP for lead years 1–7 and 7 year rolling mean MSLP from (a) ERA5 and (b) HadSLP. Grey and black contours enclose regions where zero correlation is outside of the 10–90 and 5–95 confidence interval for the correlation skill, respectively.

EA MSLP/ERA5 AMV regression slopes and r_{mi} values. The sloped pattern demonstrates that the two quantities are related, with a correlation of -0.50, demonstrating that the weak atmospheric response to AMV relates to the signal-to-noise paradox for summer decadal prediction. Overall, Figure 4 demonstrates that the EA MSLP response to AMV is underestimated in MPI-ESM-LR decadal hindcasts and as a consequence, the ensemble signal-to-noise ratio is significantly lower than that of observations, affecting the reliability of the hindcasts.

The upper tropospheric response to AMV is important for downstream impacts. In Figure 5, composites differences of zonally anomalous 200 hPa geopotential height for positive and negative AMV phases are shown for different data sources. The zonal mean is removed as there are large discrepancies between NOAA 20CR and ERA5 in the zonal mean component, particularly in the tropics/subtropics. The cause of this discrepancy is unclear and it was found to not be an artifact of the detrending method used. The atmospheric response to extratropical heating is often shallow (Hoskins and Karoly, 1981; Gastineau and Frankignoul, 2015; Ghosh et al., 2017). Despite this, the upper tropospheric response in the extratropical eastern Atlantic in NOAA 20CR and ERA5 reanalyses (Figure 5 (a–c)) appears to be an extension of the negative response at the surface and considerable negative anomalies exist at levels between the surface and 200 hPa in this region (not shown). In this particular region the reanalysis results were found to be the same when the zonal mean is not removed (not shown). In the tropical/subtropical North Atlantic, the upper level response is positive and hence baroclinic through the troposphere, which is typical in response to tropical SST anomalies.

When using model AMV, model composites (Figure 5 (d–e)) are similar to each other and markedly different from reanalysis. The response resembles a Rossby wave train emanating from the Caribbean, leading to a negative response east/south east of Newfoundland and a positive response in approximately the position of the strong negative response in reanalyses. The

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

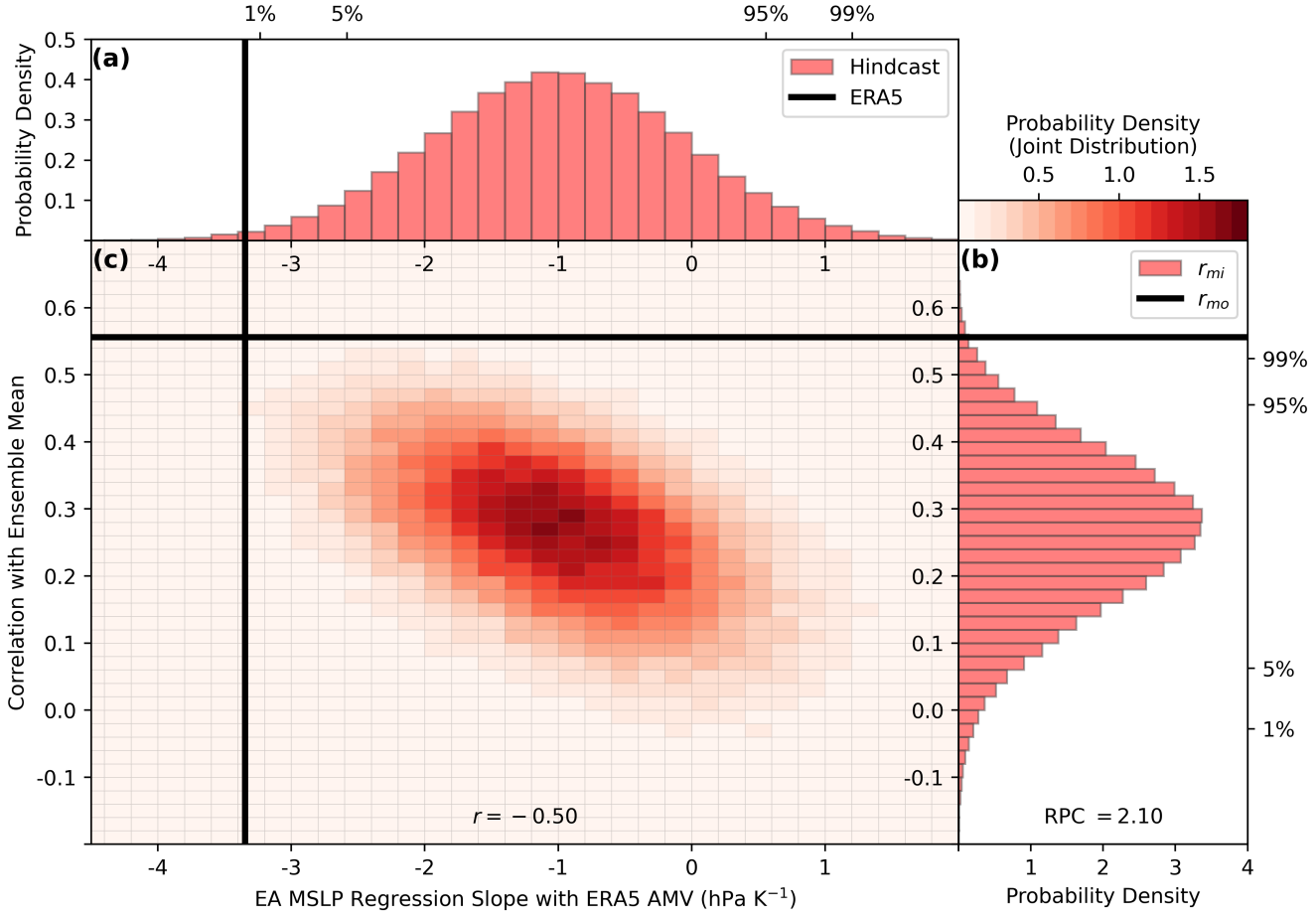

Figure 4. Teleconnection strength deficiencies lead to signal-to-noise errors. (a) (top) histogram of regression slopes for 100,000 samples of single member East Atlantic (EA) MSLP timeseries with the ERA5 AMV Index, along with the same slope but for ERA5 EA MSLP (black line). (b) (right) histogram of correlations r_{mi} of sampled hindcast member EA MSLP with ensemble mean EA MSLP, along with the correlation between ERA5 and the ensemble mean (black line). (c) (main) joint histogram for the sampled hindcast member statistics in (a) and (b). Black lines in (a) and (b) extend in to (c). In (a) and (b), cumulative 1, 5, 95 and 99% levels for the histograms are shown.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

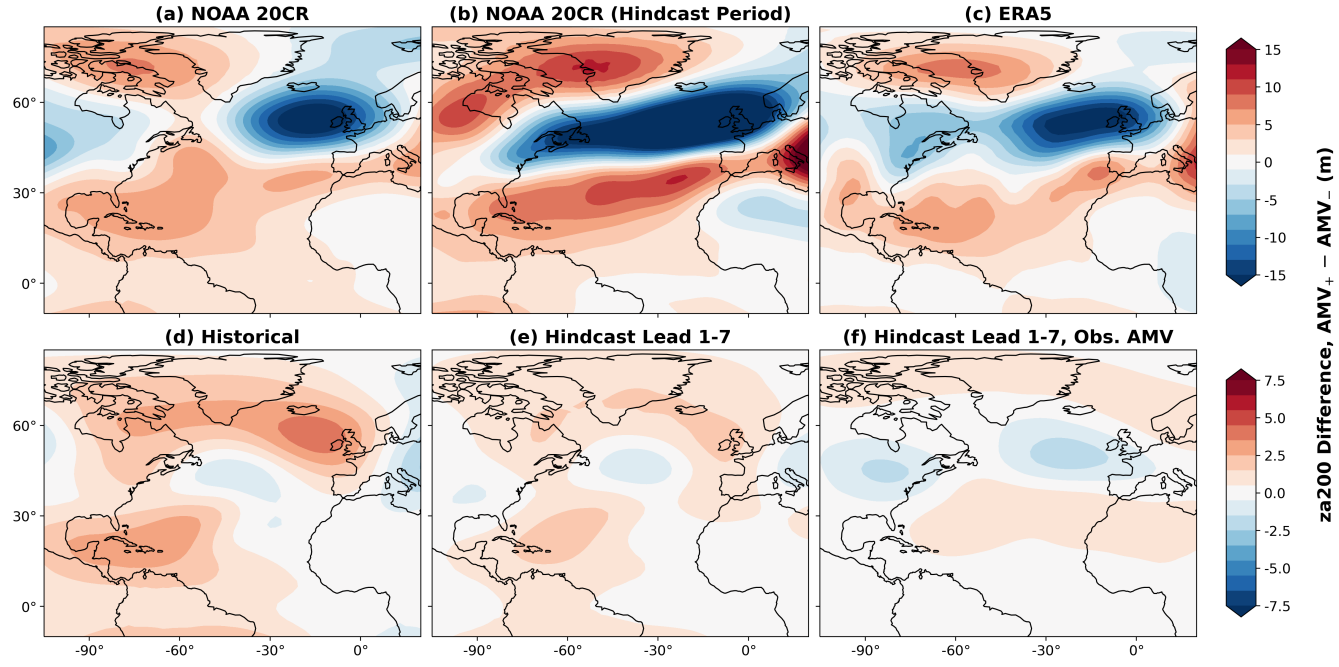

Figure 5. 200 hPa geopotential height response to AMV in reanalyses and model experiments. Composite difference in MSLP during positive and negative AMV using MSLP and AMV index from (a) NOAA 20CR/HadISST; (b) NOAA 20CR/HadISST during the common period of the decadal hindcasts and NOAA 20CR only; (c) ERA5 during the hindcast period; (d) historical runs using MPI-ESM-LR and (e) MPI-ESM-LR decadal hindcasts. (f) as with (e), but using the ERA5 AMV index rather than that of the hindcasts. Note that the zonal mean is removed in each case (see text) and (a–c) and (d–f) use different colour scales.

hindcast response is weaker than in the historical simulations, which is consistent with the weaker tropical SST signal. The hindcast composite using ERA5 AMV phases (Figure 5 (f)) shows a weak signal which resembles the reanalysis response in the extratropics. Figure 1 (f) demonstrated that tropical SSTs are poorly predicted by the hindcasts, which likely explains why the tropical-origin Rossby wave response is not present in Figure 5 (f).

With knowledge of the 200 hPa geopotential height response to AMV found in model experiments, it is useful to re-consider whether the reanalysis responses may also relate to a Rossby wave train from the Caribbean. The response over the Caribbean is similar in reanalyses and model experiments, whilst the extratropical response in reanalyses is likely to at least in-part relate to the surface level response to extratropical diabatic heating. As the hindcast extratropical MSLP response is much weaker than that in observations/reanalyses, it is not surprising that it is less likely to extend to the upper troposphere; furthermore in the case of Figure 1 (f) where the confounding effect of tropical SST anomalies is reduced, the pattern is remarkably similar, albeit much weaker, to that in reanalyses. It is therefore hypothesised that the upper tropospheric response in reanalyses is some combination of the Rossby wave train response to tropical SSTs which is evident in the model experiments, and the surface response to diabatic heating in the SPG region extending upwards. This hypothesis is tested in the next section.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

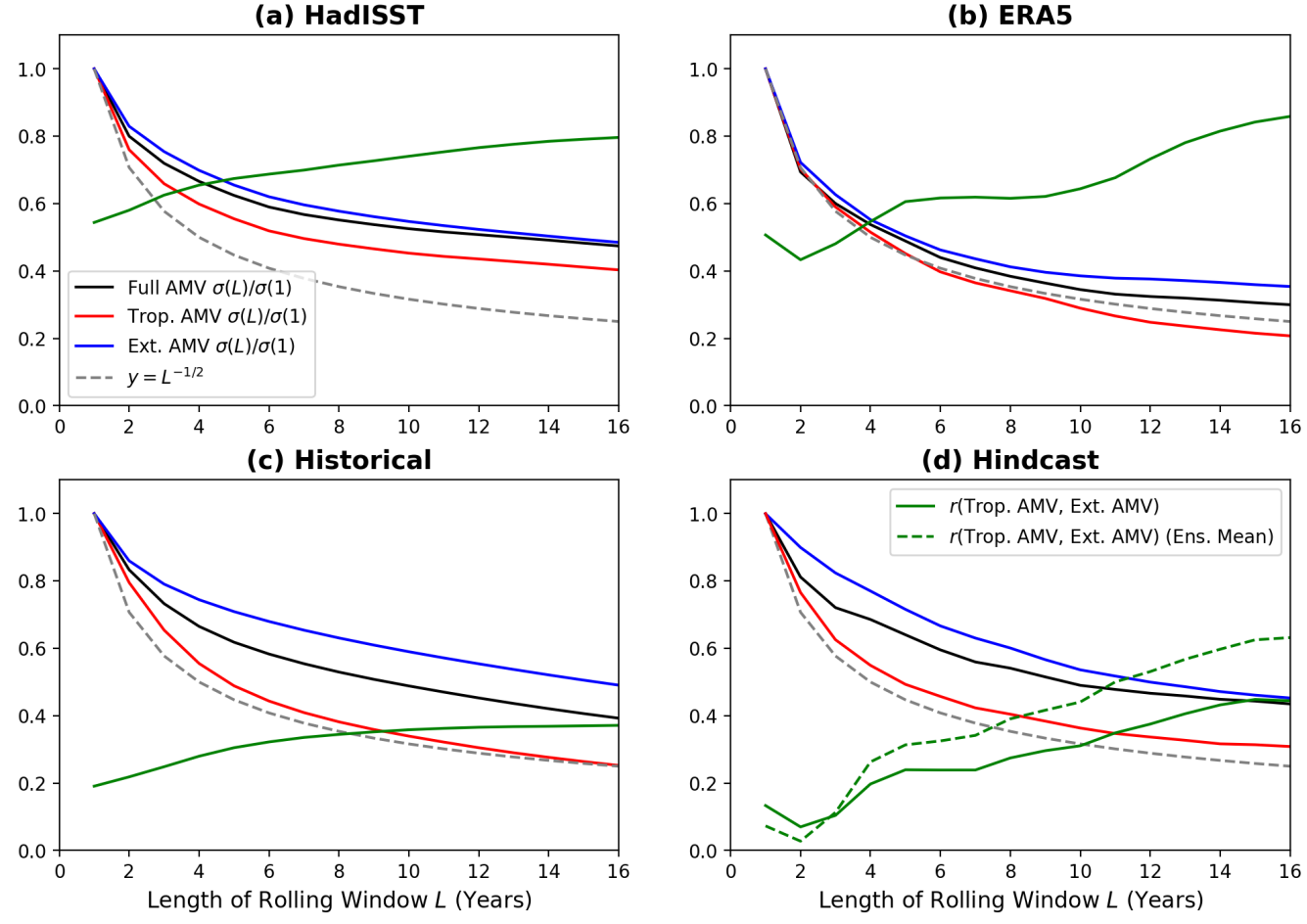

Figure 6. Tropical and extratropical North Atlantic SST variability for different rolling windows. Ratio of rolling mean to interannual standard deviation for the full AMV region and its tropical (south of 30°N) and extratropical (north of 30°N) components in (a) HadISST, (b) ERA5, (c) Historical MPI-ESM-LR simulations and (d) MPI-ESM-LR hindcasts, where each rolling window L is the mean of leads 1 to L. Grey dashed lines show the theoretical relationship for white noise. Solid green lines show correlations between tropical and extratropical AMV against rolling window length, while the dashed green line in (d) shows the same for the hindcast ensemble mean.

4 Timescale Dependence of Tropical and Extratropical Pathways

Subpolar gyre SST variability is known to have a large multidecadal component associated with AMV, whilst considerable interannual variability exists in tropical Atlantic SSTs. Therefore the relative contributions of the tropics and extratropics to North Atlantic SST variability are likely to depend on the timescale of interest. In Figure 6, the ratio of standard deviations of rolling-averaged SSTs to that of interannual SSTs is shown for different rolling window lengths, using the full, tropical and extratropical AMV indices as defined in Section 2. Low-frequency variability beyond that of white noise exists for the

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

full AMV region in all datasets. In all cases, extratropical North Atlantic SSTs have a greater ratio than the full region, and in turn tropical SSTs have a lower ratio than the full region. It should therefore be expected that in both model experiments and reanalyses, the longer the rolling window used, the larger the relative contribution of extratropical SSTs to AMV region variability. As a result, it is expected that the diabatic heating response to extratropical SSTs should play an increasingly important role in the upper level atmospheric response with increasing rolling window length.

Figure 6 also shows correlations between tropical and extratropical AMV for different rolling window lengths. In all cases, the correlation tends to increase with increasing rolling window length. The HadISST and ERA5 correlations are large, demonstrating that the extratropical and tropical SSTs are strongly coupled; this, alongside the relatively short length of data availability, means that it is difficult to separate the effects of tropical and extratropical SSTs by defining composites based on them rather than the full AMV index. In both the historical and hindcast simulations, the tropical-extratropical SST correlations are much lower than in observations, although in both cases the correlation increases with rolling window length as in observations. The correlation curve for the hindcasts appears 'noisy' despite the large ensemble size; this is explained by a significant contribution from the shared ensemble mean component.

Figure 7 shows the composite difference response of zonally anomalous geopotential height to year-on-year AMV region SST variability. Year-on-year variability is computed as half of the difference between successive summers; this is a simple high-pass filter which damps the low frequency contribution to interannual variability (Stephenson et al., 2000). Remarkably, a Rossby wave train response emanating from the Caribbean like that found using 7 year rolling means in model experiments is found in all reanalyses and model experiments. This implies that this Rossby wave response is a component of the overall atmospheric response to AMV in reality as well as in model experiments, but it is only the dominant component on the 7 year timescale for the model case. The strength of associated geopotential height anomalies are comparable in all data sources. The zonal mean component was found to be similar amongst reanalyses (not shown), but zonal anomalies are used for consistency with Figure 5. For the hindcast, results for both lead 1 only and all individual leads up to lead 15 are shown (Figure 7 (e) and (f) respectively). There are no major differences, demonstrating that this feature of the hindcast atmospheric response to North Atlantic SSTs is not dependent on lead time.

Figure 8 shows the regression slope for EA zonally anomalous 200 hPa geopotential height against the AMV index for different rolling window lengths and using different data sources. Only the 16 extended length ensemble members are considered for the hindcasts, and up to lead year 15 is used; results up to lead year 10 are robust when using the full ensemble (not shown). As demonstrated in Figures 5 and 7, for single years the slope is positive in ERA5 and NOAA 20CR reanalyses, but becomes increasingly negative with increased rolling window length L. The mean regression slope in historical and hindcast simulations is positive for all rolling window lengths tested (up to 15 years), but in both cases, there is a robust decrease in the regression slope with increased rolling window length, and for all multiyear rolling window lengths, the hindcast slope is significantly lower than that of the historical simulations. This likely relates to the improved response to extratropical diabatic heating in the hindcasts, due to improvements in SPG SST simulation. By including all leads—for rolling window length $L = 1$, leads 1 to 15 are used, for $L = 3$, leads 1–3, 2–4 up to 13–15 are used, and so on—it is shown that the dependence on rolling window length in the hindcasts is not caused by the inclusion of longer leads with longer rolling windows. Leads 1–L (i.e. the first

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

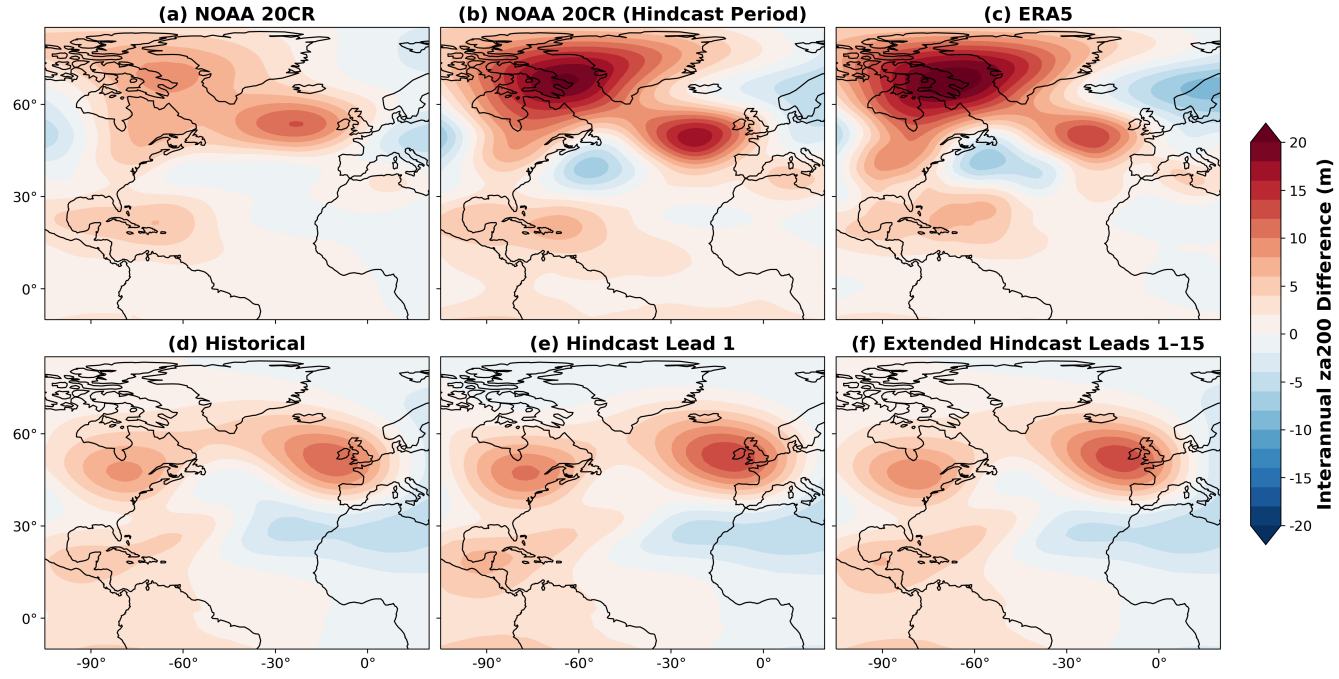

Figure 7. 200 hPa geopotential height response to year-on-year AMV index variability in reanalysis and MPI-ESM-LR model experiments, where year-to-year variability is computed as the difference between consecutive years. Composite difference in MSLP during positive and negative interannual AMV index years using MSLP and SSTs from (a) NOAA 20CR/HadISST; (b) NOAA 20CR/HadISST during the common period of the decadal hindcasts and NOAA 20CR only; (c) ERA5 during the hindcast period; (d) historical runs using MPI-ESM-LR, (e) MPI-ESM-LR decadal hindcasts for lead year 1 and (f) single years for leads 1 to 15 in the 16 extended length members of the MPI-ESM-LR decadal hindcasts. Note that the zonal mean is removed in each case, as with the 7 year rolling mean results shown in Figure 5.

L summers) are also shown; for $L = 3$ and $L = 7$, the slope is significantly more negative than when considering all leads, demonstrating that the improvements are greater at short leads; note that as L tends towards 15, the red squares (lead 1–L) and red circles (all leads) necessarily converge.

Figure 8 also shows the hindcast regression slopes using ERA5 AMV. This approximates the component of the response to AMV which is predictable. When considering all leads, the slopes are significantly less than zero for rolling window lengths of 3, 5 and 7 years, whilst the lead 1–L slope is significantly more negative than for all leads for L from 1 to 9. This is consistent with the skilfully predicted SPG SSTs and poorly predicted tropical SSTs shown in Figure 1.

Overall, the observed EA 200 hPa geopotential height response to North Atlantic SST anomalies is positive on interannual timescales, but becomes negative for multiyear windows due the increasing importance of extratropical SSTs relative to tropical SSTs. This behaviour is present in model experiments, but due to the weaker/shallower response to extratropical diabatic heating, the overall response remains positive for all rolling window lengths considered. However, the hindcast response to the

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

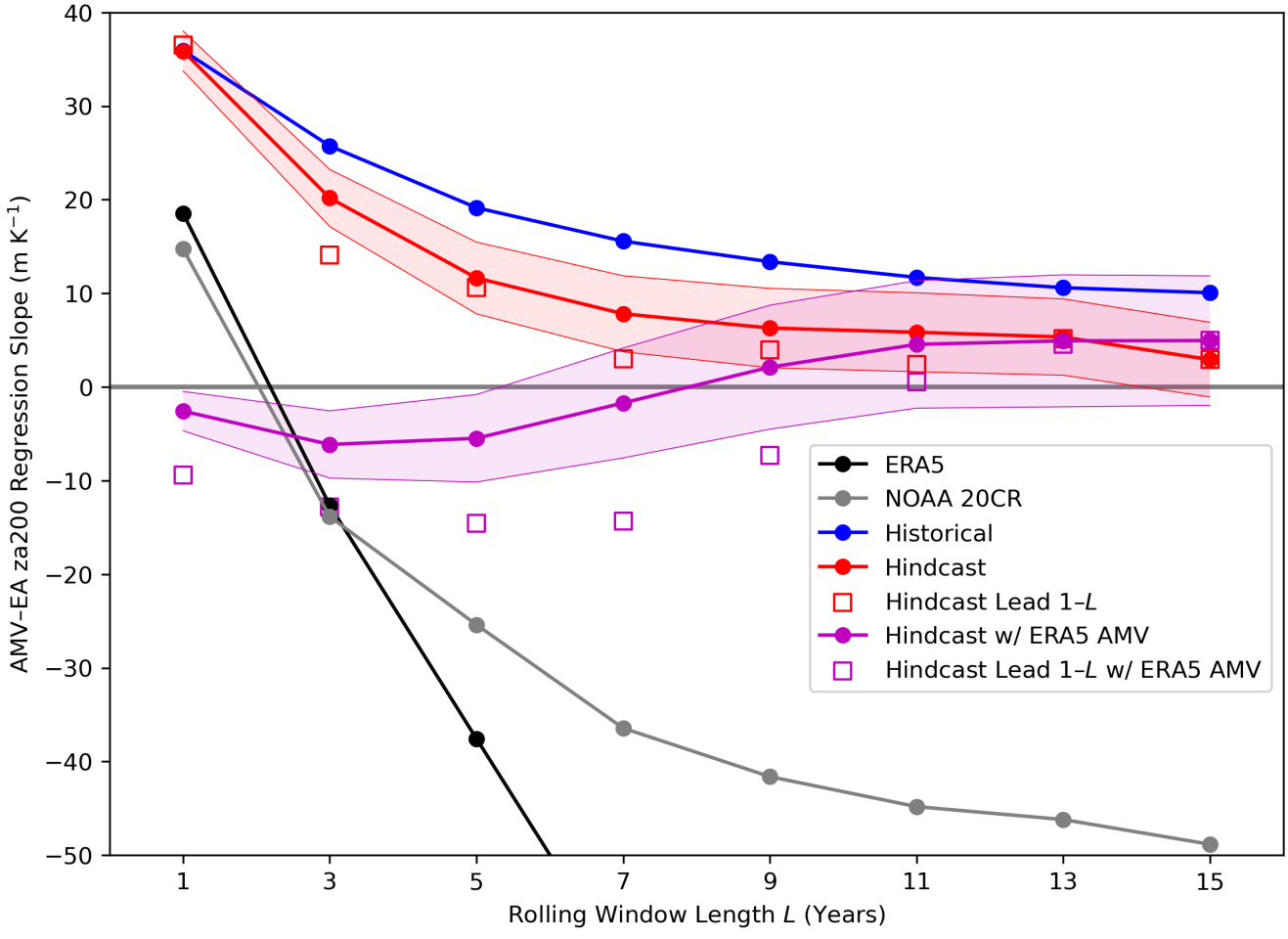

Figure 8. Regression slope between AMV and EA zonally anomalous 200 hPa geopotential height for different rolling window lengths L. Black: ERA5; Grey: NOAA 20CR with HadISST AMV; Blue: MPI-ESM-LR historical simulations; Red: MPI-ESM-LR hindcasts, for all leads (line and filled circles) and lead 1–L (squares); Purple: Hindcasts with ERA5 used for AMV, for all leads (line and filled circles) and lead 1–L (squares). Shaded red and purple regions show the 5–95 confidence interval for the hindcasts at all leads with hindcast and ERA5 AMV respectively, computed as the mean value ± the standard error of slopes from individual member timeseries scaled by 1.644.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

ERA5 AMV index is negative for rolling window lengths up to 9 years, due to skilful prediction of extratropical SSTs at these leads.

5 Discussion and Conclusions

The capability of MPI-ESM-LR historical and decadal hindcast simulations at capturing spatial patterns of SSTs associated with Atlantic Multidecadal Variability (AMV) was tested. We find that both can capture the typical 'horseshoe' pattern of SSTs, but only the hindcasts accurately simulate the strength of subpolar gyre (SPG) SST anomalies. However, tropical SST anomalies associated with AMV are weaker and hence less accurate in the hindcasts compared to the historical simulations, and whilst the SPG response in hindcasts is present when defining composites using ERA5 AMV phases (indicating predictability), the tropical SST response largely vanishes. The previously reported negative EA MSLP response to AMV during boreal summer was found in observations and reanalyses, which is associated with diabatic heating due to subpolar gyre SST anomalies. This response is not present in the historical simulations, but it is present—although with underestimated amplitude—in the hindcasts, in agreement with the differences in SST anomalies. This response is also found in the hindcasts when using ERA5 AMV phases, suggesting that this response is predictable; the correlation skill between the hindcast ensemble mean and observational references was computed which further demonstrated predictability in the extratropical North Atlantic.

The upper tropospheric (200 hPa) geopotential height response to AMV in MPI-ESM-LR historical and hindcast simulations was also analysed and compared to reanalyses. The negative EA MSLP response was found to extend upwards to this level in reanalyses. Modelled responses are consistent amongst historical and hindcast simulations, and resemble a Rossby wave train emanating from the Caribbean and leading to a positive anomaly in the extratropical eastern North Atlantic. The positive anomalies in the Caribbean which appear to be the origin of the Rossby wave train exist in the reanalyses, leading to the hypothesis that the upper tropospheric reanalysis response to AMV is a combination of the extratropical surface response due to diabatic heating extending upwards, and the tropical-to-extratropical Rossby wave train evident in model simulations.

By varying the number of years used for rolling means, it becomes clear that the relative contribution to AMV SSTs from the extratropics compared to the tropics increases for longer rolling means, indicating a larger proportion of low-frequency variability in the extratropics. As the upper tropospheric response to AMV in reanalyses is hypothesised to consist of two components with one due to tropical SSTs and one due to extratropical SSTs, the 200 hPa geopotential height response is investigated for interannual North Atlantic SST variability. In this case, both the reanalysis and model responses to AMV region SSTs resemble the modelled response to 7 year rolling mean AMV SSTs, highlighting the importance of tropical SST anomalies at shorter timescales, and supporting the hypothesis that the model 7 year Rossby wave response is a component of the reanalysis response. The regression slope of the AMV index on EA 200 hPa geopotential height is calculated for different data sources and rolling window lengths, and a robust reduction in the slope (to negative levels in the case of reanalyses) with increasing rolling window length is found, with the decadal hindcasts performing better than historical simulations. Furthermore, the predictable part of the hindcast slope (calculated using ERA5 AMV) is found to have the correct sign for rolling windows lengths from 3 to 9 years.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

It is found that MPI-ESM-LR is capable (in historical and hindcast simulations) of representing the tropical Rossby wave mechanism, whilst the extratropical diabatic heating mechanism is not captured at all without initialisation, and is present and predictable but weaker than observed in initialised lead 1–7 hindcasts. In the extratropical North Atlantic, the upper tropospheric response to AMV is dominated by the response to diabatic heating at the 7 year timescale, but the underestimation of this mechanism in hindcasts means that the higher-frequency Rossby wave response still dominates, although to a decreasing extent with increasing rolling window length.

Skilful prediction but with underestimated strength of predictable signals is a common phenomenon in the North Atlantic-Europe sector known as the 'signal-to-noise paradox' (Scaife and Smith, 2018). By sampling individual member hindcasts we find that the reliability errors which define the signal-to-noise paradox are present in MPI-ESM-LR decadal predictions of JJA East Atlantic MSLP, and that these errors are closely linked to the underestimated atmosperic response to AMV.

With an 80 member hindcast, it has been possible to demonstrate robust MSLP predictability in the region where the observed response to AMV is strongest, despite a) the weak amplitude of the model response to AMV in this region and b) the limited degrees of freedom available to assess skill given the short hindcast period. Unless the causes of the signal-to-noise paradox are identified and eliminated, large ensemble forecasts which are post-processed to address signal-to-noise deficiencies are necessary in order to predict summertime decadal variability, including surface impacts. It is notable that the Rossby wave mechanism does not appear to be underestimated in MPI-ESM-LR simulations (except where tropical SST variability is underestimated in the decadal hindcasts); this may be relevant to understand the causes of the signal-to-noise paradox. However, the fact that one mechanism is underestimated whilst the other is not means that skill is reduced where they interfere in the upper troposphere, which is an issue which cannot be easily rectified by increased ensemble size and/or post-processing of forecasts.

The dominance of the response to extratropical diabatic heating in the upper troposphere in reanalyses suggests that the impact of heating associated with SPG SSTs is less shallow than previously reported in Ghosh et al. (2017), but at 7 year timescales it is weak and baroclinic in hindcasts. Ossó et al. (2020) identify a summertime atmospheric response to seasonal springtime Gulf Stream SST anomalies which is equivalent-barotropic in nature, which they attribute to reinforcement through ocean-atmosphere coupling. As with the response to diabatic heating in this study, this response was found to be present but weak in models and hence was identified as potentially relevant to the signal-to-noise paradox; furthermore, model deficiencies were associated with underestimated ocean-atmosphere coupling. Baroclinic anomalies can also lead to stronger, sustained barotropic anomalies through eddy feedback (e.g. O'Reilly and Czaja, 2015), which has also been found to be deficient in models and linked to the signal-to-noise paradox (Hardiman et al., 2022). The role of positive ocean-atmosphere and eddy feedbacks for the observed summer atmospheric response to AMV warrants further study.

The atmospheric response to subpolar gyre diabatic heating has previously been identified as leading to increased precipitation in northwestern Europe and increased temperatures in central Europe (Ghosh et al., 2017). The aim of this study has been to understand mechanisms behind the summer atmospheric response to AMV and the capability of MPI-ESM-LR model experiments to simulate and predict them; further work is necessary to understand whether models are capable of simulating their surface impacts over Europe and other land regions.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

The surface level decadal atmospheric response to AMV in the North Atlantic-Europe (NAE) sector during summertime is dominated by a low pressure anomaly to the west of Great Britain and Ireland. This response is a result of diabatic heating due to subpolar gyre SST anomalies, and it is captured by a large ensemble of MPI-ESM-LR decadal hindcasts, resulting in statistically significant MSLP skill. Despite this, the amplitude of the surface level response is strongly underestimated in MPI-ESM-LR decadal hindcasts, which is consistent with the 'signal-to-noise paradox' issue in NAE sector climate prediction. The upper level response in reanalyses can be explained by the surface level response extending upwards, whilst in hindcast and historical simulations, the dominant response is a Rossby wave train associated with tropical SST anomalies. By adjusting rolling window lengths between 1 and 15 years, it becomes clear that both the high-frequency tropical Rossby wave and low-frequency extratropical diabatic heating response exist in observations and reanalyses, but deficiencies in the strength of the surface level response in hindcasts means that its role at upper levels increases more slowly with rolling window length than in observations. Overall, this study demonstrates the existence of dynamically-driven skill for decadal prediction of European summers, which is best realised using large ensembles and may be enhanced further by future model improvements.

Data Availability

The MPI-ESM-LR large ensemble free-running model data is openly available through the Earth System Grid Federation (ESGF), accessible at https://esgf-metagrid.cloud.dkrz.de/search under the CMIP6 project and the CMIP activity ID. The first 16 members of MPI-ESM-LR decadal hindcasts up to a 10 year lead are also available through ESGF under the DCPP activity ID. These members and their extension to 20 years are openly available from DKRZ via the following link: https://www.wdc-climate.de/ui/entry?acronym=DKRZ_LTA_1075_ds00008. The remaining 64 members are openly available from DKRZ via the following link: https://www.wdc-climate.de/ui/entry?acronym=DKRZ_LTA_1075_ds00016. HadISST sea surface temperature data is openly available from https://www.metoffice.gov.uk/hadobs/hadisst/. HadSLP sea level pressure data is openly available from https://www.metoffice.gov.uk/hadobs/hadslp2/ NOAA 20th Century Analysis Version 3 data is openly available from https://www.psl.noaa.gov/data/gridded/data.20thC_ReanV3.html. ERA5 reanalysis data is openly available from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels-monthly-means (surface fields) and https://cds.climate.copernicus.eu/datasets/reanalysis-era5-pressure-levels-monthly-means (pressure level fields).

Author Contribution

WAM and JGP conceived the original study and sourced project funding. NCW performed the analysis with input from WAM. NCW prepared the manuscript with contributions from all authors.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

Financial Support

This research was supported by the BMBF funded project *Coming Decade* (Grants 01LP2327D, 01LP2327E; all authors) and the Horizon Europe funded project *ASPECT* (Grant 101081460; NCW and WAM). JGP thanks the AXA Research Fund for support.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

References

Allan, R. and Ansell, T.: A New Globally Complete Monthly Historical Gridded Mean Sea Level Pressure Dataset (HadSLP2): 1850–2004, Journal of Climate, 19, 5816–5842, 2006.

Athanasiadis, P. J., Bellucci, A., Scaife, A. A., Hermanson, L., Materia, S., Sanna, A., Borrelli, A., MacLachlan, C., and Gualdi, S.: A Multisystem View of Wintertime NAO Seasonal Predictions, Journal of Climate, 30, 1461–1475, 2017.

Barnston, A. G. and Livezey, R. E.: Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns , Monthly weather review, 115, 1083–1126, 1987.

Borchert, L. F., Menary, M. B., Swingedouw, D., Sgubin, G., Hermanson, L., and Mignot, J.: Improved decadal predictions of North Atlantic subpolar gyre SST in CMIP6, Geophysical Research Letters, 48, e2020GL091 307, 2021.

Brune, S. and Baehr, J.: Preserving the coupled atmosphere–ocean feedback in initializations of decadal climate predictions, Wiley Interdisciplinary Reviews: Climate Change, 11, e637, 2020.

Cassou, C., Terray, L., and Phillips, A. S.: Tropical Atlantic Influence on European Heat Waves, Journal of climate, 18, 2805–2811, 2005.

Deser, C. and Phillips, A. S.: Spurious Indo-Pacific Connections to Internal Atlantic Multidecadal Variability Introduced by the Global Temperature Residual Method, Geophysical Research Letters, 50, e2022GL100 574, 2023.

Dunstone, N., Smith, D. M., Hardiman, S. C., Hermanson, L., Ineson, S., Kay, G., Li, C., Lockwood, J. F., Scaife, A. A., Thornton, H., Ting, M., and Wang, L.: Skilful predictions of the Summer North Atlantic Oscillation, Communications Earth & Environment, 4, 409, 2023.

Düsterhus, A. and Brune, S.: The effect of initialisation on 20 year multi-decadal climate predictions, Climate Dynamics, 62, 831–840, 2024.

Eade, R., Smith, D. M., Scaife, A. A., Wallace, E., Dunstone, N. J., Hermanson, L., and Robinson, N.: Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?, Geophysical Research Letters, 41, 5620–5628, 2014.

Folland, C. K., Knight, J., Linderholm, H. W., Fereday, D., Ineson, S., and Hurrell, J. W.: The Summer North Atlantic Oscillation: Past, Present, and Future , Journal of Climate, 22, 1082–1103, 2009.

Gastineau, G. and Frankignoul, C.: Influence of the North Atlantic SST Variability on the Atmospheric Circulation during the Twentieth Century, Journal of Climate, 28, 1396–1416, 2015.

Ghosh, R., Müller, W. A., Baehr, J., and Bader, J.: Impact of observed North Atlantic multidecadal variations to European summer climate: A linear baroclinic response to surface heating, Climate Dynamics, 48, 3547–3563, 2017.

Guemas, V., Auger, L., and Doblas-Reyes, F. J.: Hypothesis Testing for Autocorrelated Short Climate Time Series, Journal of Applied Meteorology and Climatology, 53, 637–651, 2014.

Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Comer, R., Nie, Y., and Ren, H.-L.: Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions, npj Climate and Atmospheric Science, 5, 2022.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keely, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hoskins, B. J. and Karoly, D. J.: The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing, Journal of the Atmospheric Sciences, 38, 1179–1196, 1981.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

Hövel, L., Brune, S., and Baehr, J.: Decadal Prediction of Marine Heatwaves in MPI-ESM, Geophysical Research Letters, 49, e2022GL099 347, 2022.

Knight, J. R., Folland, C. K., and Scaife, A. A.: Climate impacts of the Atlantic multidecadal oscillation, Geophysical Research Letters, 33, 2006.

Krieger, D., Brune, S., Pieper, P., Weisse, R., and Baehr, J.: Skillful decadal prediction of German Bight storm activity, Natural Hazards and Earth System Sciences, 22, 3993–4009, 2022.

Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., Kröger, J., Takano, Y., Ghosh, R., Hedemann, C., 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, 2019.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E., Nam, C. C., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1. 2) and its response to increasing CO2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, 2019.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., and Stockdale, T.: Decadal Prediction: Can It Be Skillful?, Bulletin of the American Meteorological Society, 90, 1467–1486, 2009.

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Reyes, F., Hawkins, E., Karspeck, A., Kimoto, M., Kumar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M., Rosati, T., Schneider, E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G. J., Vecchi, G., and Yeager, S.: Decadal Climate Prediction: An Update from the Trenches, Bulletin of the American Meteorological Society, 95, 243–267, 2014.

Moemken, J., Feldmann, H., Pinto, J. G., Buldmann, B., Laube, N., Kadow, C., Paxian, A., Tiedje, B., Kottmeier, C., and Marotzke, J.: The regional MiKlip decadal prediction system for Europe: Hindcast skill for extremes and user-oriented variables, International Journal of Climatology, 41, E1944–E1958, 2021.

Olonscheck, D., Suarez-Gutierrez, L., Milinski, S., Beobide-Arsuaga, G., Baehr, J., Fröb, F., Ilyina, T., Kadow, C., Krieger, D., Li, H., Marotzke, J., Plésiat, É., Schupfner, M., Wachsmann, F., Wallberg, L., Wieners, K.-H., and Brune, S.: The New Max Planck Institute Grand Ensemble With CMIP6 Forcing and High-Frequency Model Output, Journal of Advances in Modeling Earth Systems, 15, e2023MS003 790, 2023.

O'Reilly, C. H. and Czaja, A.: The response of the Pacific storm track and atmospheric circulation to Kuroshio Extension variability, Quarterly Journal of the Royal Meteorological Society, 141, 52–66, 2015.

Ossó, A., Sutton, R., Shaffrey, L., and Dong, B.: Development, Amplification, and Decay of Atlantic/European Summer Weather Patterns Linked to Spring North Atlantic Sea Surface Temperatures, Journal of Climate, 33, 5939–5951, 2020.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

Rayner, N., Parker, D. E., Horton, E., Folland, C. K., Alexander, L. V., Rowell, D., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research: Atmospheres, 108, 2003.

Rousi, E., Fink, A. H., Andersen, L. S., Becker, F. N., Beobide-Arsuaga, G., Breil, M., Cozzi, G., Heinke, J., Jach, L., Niermann, D., Petrovic, D., Richling, A., Riebold, J., Steidl, S., Suarez-Gutierrez, L., Tradowsky, J. S., Coumou, D., Düsterhus, A., Ellsäßer, F., Fragkoulidis, G., Gliksman, D., Handorf, D., Haustein, K., Kornhuber, K., Kunstmann, H., Pinto, J. G., Warrach-Sagi, K., and Xoplaki, E.: The extremely hot and dry 2018 summer in central and northern Europe from a multi-faceted weather and climate perspective, Natural Hazards and Earth System Sciences, 23, 1699–1718, 2023.

Scaife, A. A. and Smith, D. M.: A signal-to-noise paradox in climate science, npj Climate and Atmospheric Science, 1, 1–8, 2018.

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R., Dunstone, N. J., Eade, R., Fereday, D., Folland, C., Gordon, M., Knight, J., Lea, D., MacLachlan, C., Maidens, A., M, M., Peterson, K., Smith, D., Vellinga, M., Wallace, E., Waters, J., and Williams, A.: Skillful long-range prediction of European and North American winters, Geophysical Research Letters, 41, 2514–2519, 2014.

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. D., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., and Zhou, B.: Weather and climate extreme events in a changing climate, Climate Change 2021: The Physical Science Basis: Working Group I contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 2021.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Jourdain, S., Kaplan, A., Kubota, H., Le Blancq, F., Lee, T.-C., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C. J., Moore, G. K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V. C., Smith, C. A., Tinz, B., Trewin, B., Valente, M. A., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P.: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system, Quarterly Journal of the Royal Meteorological Society, 145, 2876–2908, 2019.

Smith, D. M., Scaife, A. A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I., Bilbao, R., Borchert, L., Caron, L.-P., Counillon, F., Danabasoglu, G., Delworth, T., Doblas-Reyes, F., Dunstone, N., Estella-Perez, V., Flavoni, S., Hermanson, L., Keenlyside, N., Kharin, V., Kimoto, M., Merryfield, W., Mignot, J., Mochizuki, T., Modali, K., Monerie, P.-A., WA, M., Nicoli, D., Ortega, P., Pankatz, K., Pohlmann, H., Robson, J., Ruggieri, P., Sospedra-Alfonso, R., Swingedouw, D., Wang, Y., Wild, S., Yeager, S., Yang, X., and Zhang, L.: North Atlantic climate far more predictable than models imply, Nature, 583, 796–800, 2020.

Stephenson, D. B., Pavan, V., and Bojariu, R.: Is the North Atlantic Oscillation a random walk?, International Journal of Climatology: A Journal of the Royal Meteorological Society, 20, 1–18, 2000.

Suarez-Gutierrez, L., Müller, W. A., Li, C., and Marotzke, J.: Dynamical and thermodynamical drivers of variability in European summer heat extremes, Climate Dynamics, 54, 4351–4366, 2020.

Sutton, R. T. and Hodson, D. L.: Atlantic Ocean forcing of North American and European summer climate, science, 309, 115–118, 2005.

Sutton, R. T. and Hodson, D. L.: Climate Response to Basin-Scale Warming and Cooling of the North Atlantic Ocean, Journal of Climate, 20, 891–907, 2007.

Terray, L. and Cassou, C.: Tropical Atlantic Sea Surface Temperature Forcing of Quasi-Decadal Climate Variability over the North Atlantic–European region, Journal of Climate, 15, 3170–3187, 2002.

https://doi.org/10.5194/egusphere-2025-6330
Preprint. Discussion started: 10 February 2026

Trenberth, K. E.: Some Effects of Finite Sample Size and Persistence on Meteorological Statistics. Part I: Autocorrelations, Monthly Weather Review, 112, 2359–2368, 1984.

Wallberg, L., Suarez-Gutierrez, L., Matei, D., and Müller, W. A.: Extremely warm European summers preceded by sub-decadal North Atlantic ocean heat accumulation, Earth System Dynamics, 15, 1–14, 2024.

Weigel, A. P., Liniger, M., and Appenzeller, C.: Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 134, 241–260, 2008.

Wulff, C. O., Greatbatch, R. J., Domeisen, D. I., Gollan, G., and Hansen, F.: Tropical Forcing of the Summer East Atlantic Pattern, Geophysical Research Letters, 44, 11–166, 2017.

Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y.-O., Marsh, R., Yeager, S. G., Amrhein, D. E., and Little, C. M.: A Review of the Role of the Atlantic Meridional Overturning Circulation in Atlantic Multidecadal Variability and Associated Climate Impacts, Reviews of Geophysics, 57, 316–375, 2019.