Coupled climate response to Atlantic Multidecadal Variability in a multi-model multi-resolution ensemble

Daniel L. R. Hodson[1] · Pierre-Antoine Bretonnière[2] · Christophe Cassou[9] · Paolo Davini[4] · Nicholas P. Klingaman[1] · Katja Lohmann[5] · Jorge Lopez-Parages[8] · Marta Martín-Rey[3] · Marie-Pierre Moine[9] · Paul-Arthur Monerie[1] · Dian A. Putrasahan[5] · Christopher D. Roberts[6] · Jon Robson[1] · Yohan Ruprich-Robert[2] · Emilia Sanchez-Gomez[10] · Jon Seddon[7] · Retish Senan[6]

Abstract

North Atlantic sea surface temperatures (SSTs) underwent pronounced multidecadal variability during the twentieth and early twenty-first century. We examine the impacts of this Atlantic Multidecadal Variability (AMV), also referred to as the Atlantic Multidecadal Oscillation (AMO), on climate in an ensemble of five coupled climate models at both low and high spatial resolution. We use a SST nudging scheme specified by the Coupled Model Intercomparision Project's Decadal Climate Prediction Project Component C (CMIP6 DCPP-C) to impose a persistent positive/negative phase of the AMV in the North Atlantic in coupled model simulations; SSTs are free to evolve outside this region. The large-scale seasonal mean response to the positive AMV involves widespread warming over Eurasia and the Americas, with a pattern of cooling over the Pacific Ocean similar to the Pacific Decadal Oscillation (PDO), together with a northward displacement of the inter-tropical convergence zone (ITCZ). The accompanying changes in global atmospheric circulation lead to widespread changes in precipitation. We use Analysis of Variance (ANOVA) to demonstrate that this large-scale climate response is accompanied by significant differences between models in how they respond to the common AMV forcing, particularly in the tropics. These differences may arise from variations in North Atlantic air-sea heat fluxes between models despite a common North Atlantic SST forcing pattern. We cannot detect a widespread effect of increased model horizontal resolution in this climate response, with the exception of the ITCZ, which shifts further northwards in the positive phase of the AMV in the higher resolution configurations.

Keywords AMV · Atlantic multidecadal variability · AMO · Atlantic multidecadal oscillation · High resolution · Decadal variability

✉ Daniel L. R. Hodson
d.l.r.hodson@reading.ac.uk

[1] Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK

[2] Barcelona Supercomputing Center, Carrer de Jordi Girona, 29-31, 08034 Barcelona, Spain

[3] Departamento de Física de la Tierra y Astrofísica, Facultad de Ciencias Físicas, Universidad Complutense de Madrid, 28040 Madrid, Spain

[4] Consiglio Nazionale delle RicercheInstituto di Scienze dell'Atmosfera e del Clima (CNR-ISAC), 10133 Turin, Italy

[5] Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany

[6] ECMWF, Reading, UK

[7] Met Office, FitzRoy Road, Exeter, EX1 3PB Devon, UK

[8] CECI, Université de Toulouse, CERFACS/CNRS, Toulouse, France

[9] CERFACS Toulouse, Toulouse, France

[10] CERFACS/CECI, Toulouse, France

1 Introduction

Over the twentieth century, Sea Surface Temperatures (SSTs) in the North Atlantic underwent periods of warming and cooling. These SST variations are now commonly referred to as the Atlantic Multidecadal Variability (AMV). Earlier studies use the term Atlantic Multidecadal Oscillation (AMO: Kerr 2000), but doubts about the oscillatory nature, given the short observational record led to the revised, more general name.

Observational studies suggest that the AMV has multiple impacts on climate (Gastineau and Frankignoul 2014; O'Reilly et al. 2017; Peings and Magnusdottir 2014; Nigam et al. 2018), including the following: changes in rainfall on both sides of Atlantic (Folland et al. 1986; Hoerling et al. 2006; Zhang and Delworth 2006; Uvo et al. 1998; Folland et al. 2001; Zhou and Lau 2001; Knight et al. 2006), warming and drying over North America (Sutton and Hodson 2005, 2007; Hodson et al. 2009), changes in European summer climate (Sutton and Dong 2012), hurricane formation (Shapiro and Goldenberg 1998; Enfield et al. 2001; Goldenberg et al. 2001) and Atlantic winter atmospheric blocking (Kwon et al. 2020). Modelling studies also suggest the AMV drives changes outside the Atlantic region; the Asian and Indian monsoon (Zhang and Delworth 2006, 2005), Siberian rainfall (Sun et al. 2015), Antarctic sea ice (Li et al. 2014) and modulation of the ENSO (Dong and Sutton 2007). A number of these links between the AMV and impacts have also been seen in historical and pre-industrial coupled climate model integrations (Ting et al. 2011; Lyu and Yu 2017).

The ultimate driver of the AMV remains a matter of debate. Patterns similar to the observed AMV (Fig. 1) are associated with multidecadal variations in the Atlantic Meridional Overturning Circulation (AMOC) in coupled climate models (Delworth et al. 1993; Vellinga and Wood 2002; Zhang and Delworth 2005; Dong and Sutton 2007; Knight et al. 2006; Hodson and Sutton 2012; Ruprich-Robert and Cassou 2015) suggesting that historical variations in the AMOC may have played a role in the observed AMV. However, the 20th century also saw considerable variation in natural and anthropogenic forcings (Hodson et al. 2014); anthropogenic aerosols rose and then declined over the region, volcanic activity reduced mid-century (Sato et al. 1993) whilst total solar irradiance (TSI) underwent considerable variability in the solar cycle around an upward trend (Lean 2018). The historical variations of anthropogenic aerosols in coupled climate models modulate Atlantic SST and reproduce the observed phasing and magnitude of the AMV (Booth et al. 2012; Watanabe and Tatebe 2019; Undorf et al. 2018). Others studies suggest that volcanic stratospheric aerosols played a key role (Otterå et al. 2010;
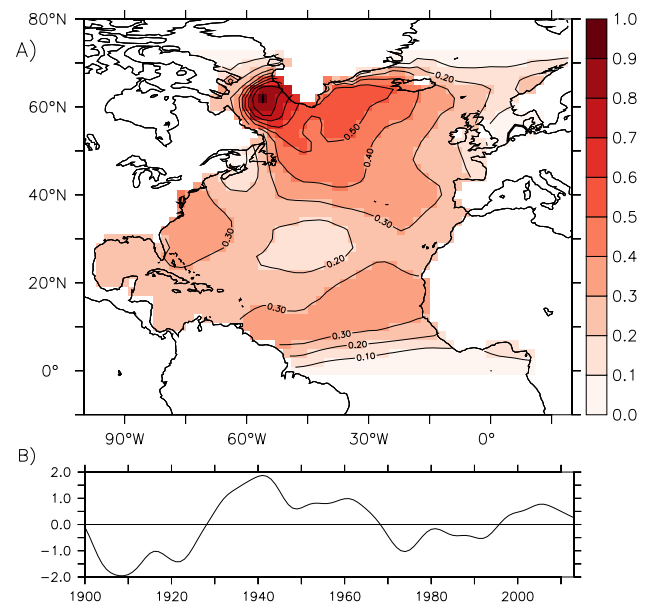

Fig. 1 A) $2 * AMV^+$ anomaly pattern used in model experiments. The AMV pattern is derived from observed Sea Surface Temperatures (SSTs), as described by the CMIP6 DCPP-C protocol (Boer et al. 2016) https://www.wcrp-climate.org/wgsip/documents/Tech-Note-1.pdf. Here we double this pattern—see Sect. 2.3 for more details. Units: K (per standard deviation). Contours are 0.2K. B) Associated AMV timeseries (low-passed and standardized)

Birkel et al. 2018), possibly by exciting multidecadal coupled modes of variability in the subpolar gyre and Labrador Sea (Swingedouw et al. 2015). The possibility also remains that the AMV may arise simply as a consequence of the ocean mixed layer integrating high-frequency atmospheric noise (Clement et al. 2015).

Whatever the ultimate origin of the AMV, the question remains: what is the climate response to the AMV? Previous studies have demonstrated significant global climate impacts using atmosphere-only models forced by AMV patterns (Sutton and Hodson 2005, 2007; Hodson et al. 2009; Davini et al. 2015; Vigaud et al. 2018; Omrani et al. 2014, 2016; Peings and Magnusdottir 2014; Mohino et al. 2011; Elsbury et al. 2019). However, such experiments do not capture the response of the ocean to this AMV forcing, and any consequent feedback on global climate. Early AMV studies using coupled atmosphere-ocean models (e.g. Zhang and Delworth 2006; Dong et al. 2006) suggested that such coupled feedbacks might be significant. Recent studies have begun to examine these feedbacks in more detail by imposing AMV forcing patterns in coupled atmosphere-ocean climate models (Kucharski et al. 2016, 2016b; Ruprich-Robert et al. 2017, 2018; Levine -et al. 2018; Qasmi et al. 2017).

These coupled model studies find widespread global impacts beyond those found in the atmosphere-only studies and hence warrant further examination. A coordinated

coupled AMV experiment was therefore proposed to assess the robustness of these responses across models and was subsequently included in the Decadal Climate Prediction Project MIP (DCPP-C Boer et al. 2016), as part of CMIP6. However, one key question that arises, is how dependent are these responses on model spatial resolution? The EU Horizon 2020 PRIMAVERA project (https://www.primavera-h2020.eu/) aimed to assess the impact of increased model resolution on a wide range of climate processes. As part of PRIMAVERA we examined the global climate impact of the AMV in five coupled climate-models in both low and high spatial resolution configurations. This allows us to assess how the modelled climate impact of the AMV is dependent on model design and model horizontal resolution.

In this paper we focus on these three key questions:

- What is the global coupled climate response to the AMV across these five models?
- How consistent is this response across the models?
- Does increasing model resolution alter this response?

This paper is set out as follows. Section 2 outlines the climate models, data and methodology used in the AMV experiments. Section 3 outlines the ANOVA technique (we provide a more detailed explanation in Appendix A). Section 4 presents the results of the analysis of these experiments. We then discuss these results in Sect. 5 and present our conclusions in Sect. 6. Additional analysis supporting this analysis can be found in the Supplementary Information.

2 Models, regridding, experimental design, and data

2.1 Models

We employed five coupled climate models in this study (CNRM-CM6-1, EC-Earth, ECMWF-IFS, MetUM-GOML2 and MPI-ESM1.2). The AMV experiments were performed at two groups of horizontal resolutions for each atmosphere model: low resolution (LR: 250–100 km) and high resolution (HR 100–50 km) for each model (summarized in Table 1) . Brief descriptions of the five climate models and their formulations are given below.

2.1.1 CNRM-CM6-1

CNRM-CM6-1 is coupled climate model consisting of the ARPEGE-Climat (Déqué et al. 1994) atmospheric model coupled to the NEMO v3.6 ocean model (Madec et al. 2017) via the OASIS3-MCT coupler (Craig et al. 2017). The model also includes a land surface scheme (ISBA—Noilhan and Planton 1989), the GELATO v6 (Salas Mélia 2002) sea ice

Table 2 AMV experiments—number of ensemble members for the 2AMV+ (2AMV−) experiment for each model and resolution. For each model, each realization was integrated for 10 years. We consider each year of a given realization to be independent (see SI section 9), hence combine realizations*years to give the total available number of ensemble members

Model	Low resolution	High resolution
CNRM-CM6-1	250	100
EC-Earth	250	170
ECMWF-IFS	300	150
MetUM-GOML2	150	150
MPIESM1.2	100	100

model, the SURFEX (Masson et al. 2013) externalized surface interface model, and the CTRIP (Decharme et al. 2019) river routing scheme. For full details see Voldoire et al. (2019). CNRM-CM6-1-LR uses a spectral model atmosphere with 91 vertical levels and a horizontal truncation of T127, resulting in a resolution at the equator of about 1.4°. The ocean has 75 vertical levels and a horizontal resolution of about 1°, reducing to 1/3° in the tropics. CNRM-CM6-1-HR uses a spectral model with 91 vertical levels and a horizontal truncation of T359, resulting in a resolution at the equator of about 0.5°. The ocean has 75 vertical levels and a horizontal resolution of about 0.25°.

2.1.2 EC-Earth

EC-Earth3P and EC-Earth3P-HR (Haarsma et al. 2020) are coupled climate models consisting of an atmospheric component based on the IFS (cycle 36r4 of the Integrated Forecast System (IFS) atmosphere-land-wave model of the European Centre for Medium Range Weather Forecasts (ECMWF)) coupled to NEMO (v3.6 Madec et al. 2017). The H-TESSEL model is used for the land surface (Balsamo et al. 2009) and is an integral part of IFS: for more details see Hazeleger and Bintanja (2012). The atmosphere and ocean/sea ice parts are coupled through the OASIS (Ocean, Atmosphere, Sea Ice, Soil) coupler (Valcke 2013). The ice model, embedded in NEMO, is the Louvain la Neuve sea ice model version 3 (LIM3, Vancoppenolle et al. 2012), which is a dynamic-thermodynamic sea ice model with 5 thickness categories. EC-Earth3P uses a reduced Gaussian-grid with 91 vertical levels and a T255 horizontal truncation/N128 grid resolution (∼100 km) for the IFS atmosphere. The NEMO ocean has 75 vertical levels and a horizontal resolution of about 1°, reducing to 1/3° in the tropics. The higher resolution EC-Earth3P-HR uses a T511 horizontal truncation/N256 grid resolution (∼50 km) IFS atmosphere, together with a 0.25° NEMO ocean and 75 vertical levels.

2.1.3 ECMWF-IFS

ECMWF-IFS is a global Earth system model that includes dynamic representations of the atmosphere, sea ice, ocean, land surface, and ocean waves. A detailed description of the ECMWF-IFS-HR and ECMWF-IFS-LR configurations used in this study, including scientific assessment of the coupled model performance, is provided in Roberts et al. (2018). ECMWF-IFS is based on the IFS atmosphere-land-wave model (cycle 43r1) coupled to NEMO (v3.4 Madec et al. 2017) and the Louvain-la-Neuve Sea Ice Model (LIM2; Bouillon et al. 2009; Fichefet and Maqueda 1997). ECMWF-IFS-LR uses a reduced Gaussian octahedral grid (Tco199 ~100 km) in the atmosphere and NEMO ORCA1 grid (~100 km) for ocean-sea ice. ECMWF-IFS-HR uses (Tco399~50 km) in the atmosphere and NEMO ORCA025 grid (~25 km) for ocean-sea ice. One of the significant differences between these configurations is the use of the Gent and Mcwilliams (1990) parameterization for the effect of mesoscale eddies with the ORCA1 grid, which is disabled when using the ORCA025 grid. Both ocean configurations use the same vertical discretization, which consists of 75 z-levels and partial cells at the ocean floor.

2.1.4 MetUM-GOML2

MetUM-GOML2 is an ocean mixed-layer coupled configuration of the Met Office Unified Model (MetUM-GOML2; Hirons et al. 2015); combining the atmosphere component from HadGEM3 (GA6.0; Walters et al. 2017) coupled to a Multi-Column K Profile (MC-KPP) mixed layer Ocean model (Hirons et al. 2015) via the Ocean Atmosphere Sea Ice Soil (OASIS) coupler (Valcke 2013)). For full details of MetUM-GOML2 see Hirons et al. (2015). The atmosphere and ocean have a horizontal resolution of either $1.25 \times 1.87°$ (~250 km, N96—LR) or $0.833 \times 0.55°$ (~100 km, N216—HR). The Atmosphere has 85 vertical levels whilst the ocean mixed-layer component extends to 1 km depth with 100 vertical levels. Sea ice fraction is prescribed from 1976–2005 mean climatology, as is Sea Surface Temperature in regions that are not ice-free all year. Although there is vertical ocean mixing, there is no horizontal advection or mixing in the model; these terms are replaced by seasonally-varying 3d temperature and salinity flux corrections, diagnosed from seasonal climatologies. Consequently, MetUM-GOML2 has small sea surface temperature biases and small model drifts (Hirons et al. 2015). In this paper, a 1976–2005 mean ocean temperature and salinity reference climatology is used, derived from the Met Office ocean analysis (Smith and Murphy 2007). Anthropogenic greenhouse gases concentration, aerosol emissions, volcanic activity are imposed and kept constant to their mean value of the period 1976–2005.

2.1.5 MPI-ESM-1-2

MPI-ESM (version 1.2.01), consisting of the atmosphere component ECHAM6.3 (Stevens et al. 2013) including the land-surface scheme JSBACH, the combined ocean and sea ice component MPIOM1.6.3 (Jungclaus et al. 2013) including the ocean biogeochemical component HAMOCC. Ocean and atmosphere are coupled through the OASIS3 coupler (Valcke 2013) with a coupling frequency of one hour. The atmosphere component applies a spectral grid at truncation T127 (about 1°, LR) or T255 (about 0.5°, HR) and 95 hybrid levels. The ocean component applies a tripolar grid (two northern poles) with a nominal resolution of 0.4° and 40 unevenly spaced z-levels. The first 20 levels are distributed over the upper 700 m of the water column. A partial grid cell formulation is used to better represent the bottom topography.

2.2 Regridding

Each model and resolution (Table 1) uses a different horizontal model grid. In order to compare a model variable between models and resolutions on a grid-point basis, the variables are regridded to a common grid using bilinear interpolation. For all variables we regrid to a $1° \times 1°$ longitude latitude grid. This will potentially result in the loss of some small-scale differences between high and low resolutions, but our primary focus in this study is the impact of increased resolution on the large-scale climate response to the AMV, rather than such smaller scale impacts.

2.3 Experimental design

The goal of this study is to assess the global impact of the AMV using coupled climate models. Previous atmosphere-only studies (e.g. Sutton and Hodson 2005; Hodson et al. 2009; Davini et al. 2015; Vigaud et al. 2018) used AMV SST anomalies (derived from observations) to drive atmosphere models. In a coupled ocean-atmosphere model, a balance must be preserved between maintaining the AMV SST pattern (on a background SST climatology), and permitting the ocean to respond to anomalous fluxes from the atmosphere. To achieve this, we followed a modified form of the experimental design proposed for the CMIP6 DCPP-C AMV experiments. Full details of this design are given in Boer et al. (2016). We briefly outline this experimental design and the necessary modifications that were required for the experiments presented here.

The AMV forcing pattern used in these experiments is shown in Fig. 1. This pattern was generated from observed SSTs (ERSST4 Huang et al. 2015) after removing an estimate of the global forced trend (from CMIP5, by regression). (Full details of this method is given in https://www.

wcrp-climate.org/wgsip/documents/Tech-Note-1.pdf). This AMV pattern is similar to those diagnosed in other studies; the horseshoe structure with higher values in the subpolar gyre (SPG) (e.g. Sutton and Hodson 2005; Deser et al. 2010; Buckley and Marshall 2016). Figure 1 does display a more prominent small maximum in the Labrador Sea than these other studies, this may arise due to the different methods used to remove the global mean SST trend.

For each model experiment, SSTs within the North Atlantic (using a predefined mask) are nudged towards the target SST field constructed from an SST climatology plus or minus the 2*AMV spatial pattern in Fig. 1: $SST_{target} = Climatology \pm 2 * AMV$. *Climatology* is a model climatology derived from a control run (or observations, in the case of MetUM-GOML2—see section 2.1.4). SST nudging is achieved through an additional surface heat flux term *hfcorr*, defined as:

$$hfcorr = -40(SST_{model} - SST_{target}) \qquad (1)$$

The coefficient of -40 $(W/m^2/K)$ was chosen based on a range of sensitivity studies (for details see http://www.wcrp-climate.org/wgsip/documents/Tech-Note-2.pdf).

Each model is then integrated multiple times, each starting from atmosphere and ocean initial conditions taken from different points in a control integration. Each realization is integrated for a maximum of 10 years, to prevent *hfcorr* from generating significant anomalies in the unconstrained subsurface ocean, which could significantly impact on deep ocean circulations and hence overwhelm the AMV forced signal. The resulting difference between the $2AMV^+$ and $2AMV^-$ experiments can then be used to assess the impact of AMV forcing on global climate. We assume that each year in a realization is statistically independent, hence we consider each year to be an *ensemble member*. (We examine this assumption further in SI section 9.)

One of the goals of this study (and a focus of the PRIMAVERA project), is to assess the impact of increased model resolution. Integrating a climate model at a higher resolution requires extra computational resources. In order to meet this goal within the available computing resources, we used a modified form of the DCPP-C AMV experimental design by doubling the originally specified AMV forcing pattern (hence 2AMV). This enhanced forcing reduces the number of realizations required to produce a detectable response. For consistency, we use this enhanced forcing (2AMV) for both LR and HR experiments. The DCPP-C protocol defined the use of a pre-industrial control climatology as the background climatology. However, none of the models had (or were likely to have) high resolution pre-industrial control integrations performed, as required by the DCPP-C protocol. Instead, each model (at both resolutions) used a 1950s constant-forcing control that had already been produced

as part of the PRIMAVERA project. One model (MetUM-GOML2), used a later climatology (1976–2014), derived from observations, due to the nature of the model design (see Sect. 2.1.4).

2.4 Data availability

All the model data used in this study, together with the code used to analyse and plot the results are available to download via https://doi.org/10.5281/zenodo.5884227. The full experimental dataset is available at https://prima-dm1.jasmin.ac.uk/ and at the ESGF https://esgf-index1.ceda.ac.uk/search/esgf-ceda/ (search for *primwp5* or *dcpp*—for MPI-ESM). Please contact the authors for further help accessing these datasets.

2.4.1 Observations

For observational comparisons we use the HadCRUT4 (https://www.metoffice.gov.uk/hadobs/hadcrut4/ Morice et al. 2012) surface air temperature (tas) dataset and the HadSLP2 (https://www.metoffice.gov.uk/hadobs/hadslp2/ Allan and Ansell 2006) mean sea level pressure (psl) dataset.

3 ANOVA

In the following section we use Analysis of Variance (ANOVA) to examine the global climate response to the AMV and how this response varies across the models and between resolutions. ANOVA is a common statistical technique that simultaneously examines the influence of a set of predictors or factors on a dependent variable. It is closely related to multiple linear regression, but can also be used with categorical factors, such as choice of climate model. It is also a generalization of the two-sample *t* test (see the appendix (A)) and is better at detecting significant impacts from multiple factors than the application of multiple *t* tests, as it simultaneously accounts for all sources of variance.

ANOVA has widespread use in many fields, but has only had limited application to climate science to date (e.g. Hodson and Sutton 2008; Yip et al. 2011; Christensen and Kjellström 2020). We provide a brief outline of the basis for ANOVA and its relation to *t*-tests in the Appendix (A). For a more detailed explanation and the application of ANOVA to climate models see Storch and Zwiers (1999), Zwiers (1996), Hodson and Sutton (2008), Yip et al. (2011) or Wilks (2019).

Much like linear regression, ANOVA begins by proposing a statistical model, consisting of a predictor, or combination of predictors, to explain the variance in a given variable. For example, a model variable X_{ej} (e.g. mean sea level

pressure—where e is the experiment, and j is the ensemble member) can be represented by:

$$X_{ej} = \mu + \alpha_e + \epsilon_{ej}. \tag{2}$$

Here μ is the mean (over both e and j), and ϵ_{ej} is a residual noise term, which we assume is *independently* and *normally* distributed with a variance σ_ϵ^2 (i.e. $\epsilon_{ej} \sim N(0, \sigma_\epsilon^2)$). Hence a predictor ($\alpha_e$) can then be assessed to see if it explains a significant fraction of the variance of variable X_{ej}, with respect to the residual noise term ϵ_{ej}. The similarity with linear regression is clear (e.g. $y = c + mx + \epsilon$).

In this study our dataset X_{emrj}, consists of two experiments (e: 2AMV^+,2AMV^-), five models (m), two resolutions (r: LR, HR), and multiple ensemble members (j). We can propose the following statistical model for this dataset combining these predictors and their interactions:

$$X_{emrj} = \mu + \alpha_e + \beta_m + \gamma_r + A_{em} + G_{er} + Z_{mr} + W_{emr} + \epsilon_{emrj} \tag{3}$$

Here α_e represents the variation in X_{emrj} due to the experiments (e: 2AMV^+, 2AMV^-) averaged across all models (m), resolutions (r) and ensemble members (j). A_{em} is an *interaction* term, and represents the variation in X_{emrj} due to both the experiments (e) and models (m). In this way we can examine the impact of the AMV (α_e), and also both how this impact varies across the models (m: A_{em}), and changes with model resolution (r: G_{er}). The remaining terms describe other aspects of the dataset, for example, β_m describes the spread of climatologies between models (m) (see the appendix (A)). We do not consider these further terms in the remainder of this study, but focus on the impact of the experiment (α_e), the impact of the choice of model on the experiment (A_{em}), and the impact of the choice of resolution on the experiment (G_{er}).

4 Results

We now present the results of the AMV experiments. Rather than presenting the results for each model at each resolution, we use the above ANOVA analysis to examine the climate response to the AMV for each season and the influence of resolution and model on this response. We begin by examining the multi-model, multi-resolution mean responses (α_e) for each season. We then examine where the models significantly differ in their AMV response (A_{em}). Finally we examine whether the AMV response changes with increased resolution (G_{er}). To ensure the analysis is not biased towards any single model, we use the same number of ensemble members for each experiment, model and resolution in the analysis—100 (10 realizations × 10 years) members for each season (90 for DJF, as most models start on the 1st January,

only 9 consecutive winters are available), e.g. a total of 1000 (DJF: 900) across models, resolution and members for each experiment. When a model has more that 100 (DJF: 90) ensemble members (see Table 2), we randomly subsample 100 (DJF: 90) from the full ensemble. The analysis results are mostly the same for different random subsamples (we examine this further in the Supplementary Information: SI Sect. 1—Figures S1–S9). There is some sensitivity to subsampling when examining the impact of resolution, which we discuss below.

We also attempt to assess the *field significance* of these results (Storch and Zwiers 1999; Livezey and Chen 1983) by counting all grid points where a result is *significant* ($p < 0.05$) and dividing this by the total number of gridpoints. This ratio is shown in each panel of the subsequent spatial field figures. For data drawn from the same random distribution we would expect this fraction to be 0.05 for a gridpoint threshold of $p < 0.05$. In the absence of covariances between neighbouring gridpoints, we could therefore classify any field where this fraction is > 0.05 to be *field significant*, hence reject the null hypothesis of it occurring purely by chance. In practice, there will be such covariances, and this will lead to an elevated threshold for field significance (i.e. greater than 0.05). There is no obvious objective method of assessing the resulting reduction of the degrees of freedom in this set of experiments, therefore we use this threshold as a guide; if a result falls below the field significance threshold, we can reject it, if it far exceeds the threshold, we can accept it. We should be more cautious, however if a result only just exceeds the threshold.

4.1 Multi-model multi-resolution mean response

We first examine the seasonal mean climate response to the AMV for three key climate variables, surface air temperature (tas), precipitation (pr) and mean sea-level pressure (psl). For each variable, we average over models, resolutions and ensemble members and compute the seasonal mean difference between the two experiments ($2 * AMV^+ - 2 * AMV^-$). These differences between the experiments considered are significant where α_e is significant in the ANOVA ($p < 0.05$, Eq. (3)). The following results are robust to subsampling (see SI section 1.1—Figures S1-S3).

4.1.1 Surface air temperature

Figure 2 shows the seasonal mean surface air temperature response to the AMV (α_e, $2 * AMV^+ - 2 * AMV^-$). The imposed time-invariant AMV SST pattern (Fig. 1a) is generally well maintained in the North Atlantic throughout the year (Fig. 2), although there is some seasonal variation. (A very similar pattern is seen in the model SSTs—SI section 2—Figure S10). The northern

sub-polar maximum is weakest in spring (MAM), perhaps due to deeper mixed layer depths that occur during late winter and early spring (Montégut et al. 2004); the larger heat capacity of the deeper mixed ocean column leading to a longer adjustment timescale for a given nudging heat flux (Eq. 1).

The AMV drives a downstream warming of surface air temperatures to the east of the Atlantic basin; much of Eurasia and northern Africa is warmer throughout the year. Over Europe, this warming is largely confined to the western and southern edges. Central Europe shows a much weaker warming, particularly in summer (JJA, Fig. 2c). This weaker warming is not what would be expected from a simple advection by the mean flow of the warm North Atlantic anomalies (e.g. $\bar{V}T'$), and may arise due to colder polar air being advected into the region by the atmospheric circulation response to the AMV (e.g. $V'\bar{T}$ (Fig. 3c)). The balance between such dynamic ($V'\bar{T}$) and thermodynamic ($\bar{V}T'$) contributions to the climate response to the AMV have been examined in both observations (O'Reilly et al. 2017) and

model studies (Qasmi et al. 2017) using analog reconstructions to separate the two processes. O'Reilly et al. (2017) demonstrates that the circulation response to the AMV acts to *reinforce* the warming of summer temperatures in Europe in observations. This difference may arise due to multiple confounders present in the observations with respect to the AMV forcing experiments, or climate model deficiencies, but highlights the potential regional sensitivities of the response to the detailed pattern of the circulation response.

North and South America show large regions of warming throughout the year—peaking in summer (JJA) in the United States (Fig. 2c), consistent with previous studies (e.g. Sutton and Hodson 2005; Ruprich-Robert et al. 2017). Over North America, the summer warming may arise as a response to the increased descent (and hence reduced cloud cover and increased insolation) in the atmosphere associated with the Gill-type response in mean sea level pressure (psl—Fig. 3—see Sect. 4.1.3), as described by Ruprich-Robert et al. (2018). In South America, the warming peaks in spring (SON, Fig. 2d). Warming of northeast Asia peaks in autumn
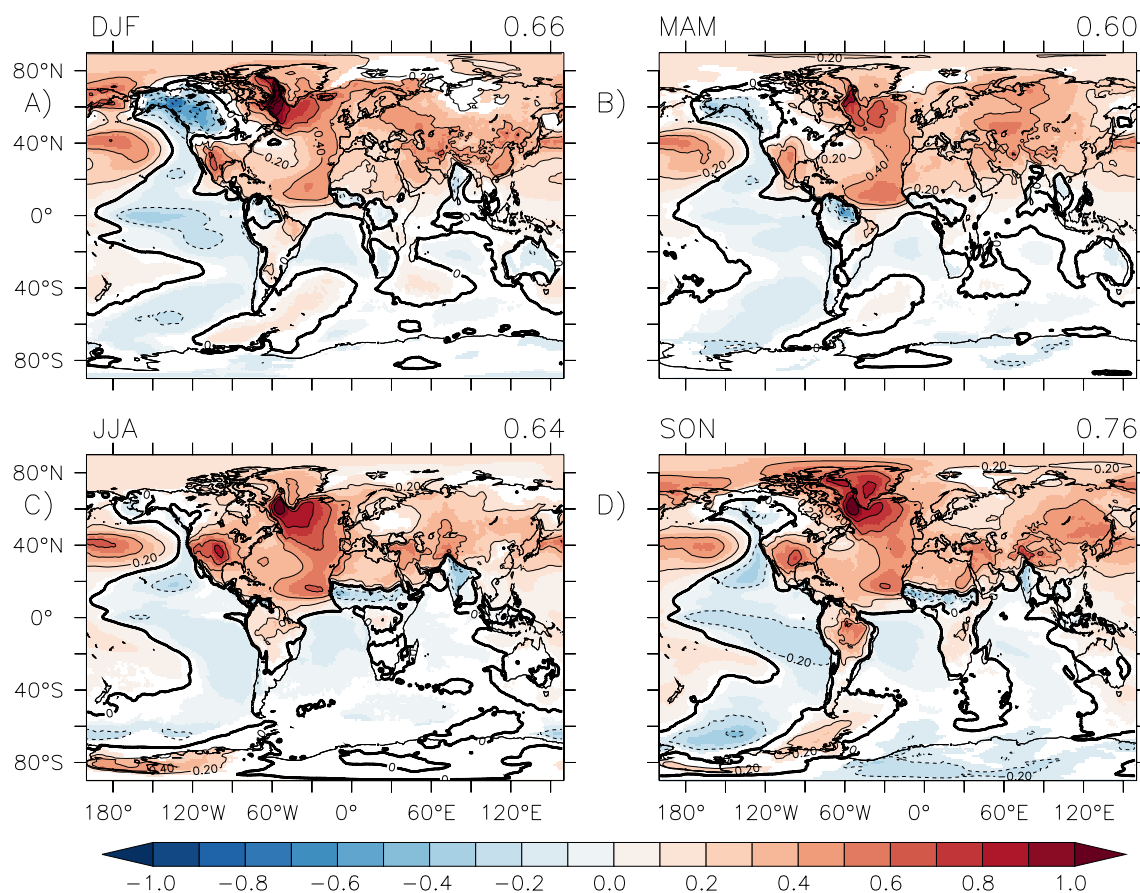

Fig. 2 Seasonal Surface Air Temperature (tas) mean (averaged over all models, resolutions and ensemble members) AMV response (2AMV$^+$-2AMV$^-$). A) Winter (Dec–Jan)). B) Spring (Mar–May). C) Summer (Jun–Aug). D) Autumn (Sep–Nov). Regions where the difference is significant (see Sect. 3, Eq. 3: α_e, $p < 0.05$) are shaded. Units: K. Contours are 0.2K. Top right of each panel: fraction of the total number of gridpoints that are significant ($p < 0.05$)
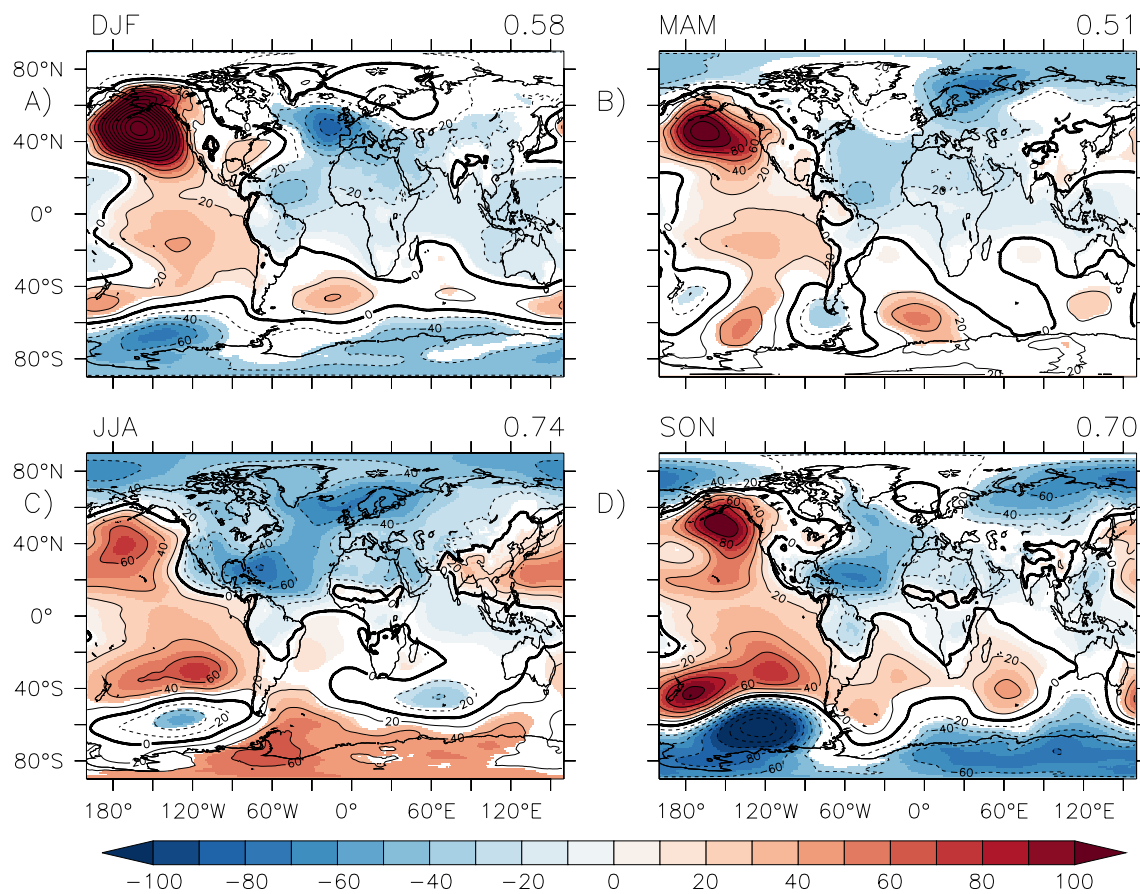
Fig. 3 As Fig. 2, but for mean sea level pressure (psl). Units: Pa. Contours show the mean sea level pressure difference for all regions—contour interval: 20 Pa

(SON) and may be driven by the changes in atmospheric circulation as described in Monerie et al. (2020).

Whilst most impacts over land are warmings, the AMV drives a cooling over some land regions: northern South America, western Canada and Alaska during winter and spring (Fig. 2 a, b) and India, northern sub-Saharan Africa during summer and autumn (Fig. 2 c, d). These coolings may be due to reduced net surface shortwave radiation from increased cloud cover associated with the enhanced precipitation (northern South America, India, northern sub-Saharan Africa—see Fig. 4 and Sects. 4.1.2 and 5) or strong changes in circulation (western Canada and Alaska—see Fig. 3 and Sect. 4.1.3).

The AMV also drives significant surface air temperature changes over the oceans beyond the North Atlantic (the same pattern is seen in SSTs (tos) SI section 2—Figure S10). The Aleutian low region in the north Pacific is warm throughout the year and the eastern and central Pacific experiences a widespread cooling, peaking in winter (DJF, Fig. 2a) and autumn (SON, Fig. 2d). This pattern of warm and cool anomalies is close to the Pacific Decadal Oscillation (PDO: Zhang and Delworth 2015) and is a common Pacific response in Atlantic-forced coupled models (Ruprich-Robert

et al. 2017; McGregor et al. 2014; Kucharski et al. 2011; Li et al. 2016); we discuss this further in Sect. 4.3.

4.1.2 Precipitation

The global precipitation response (Fig. 4) is dominated by dipolar anomalies within the tropics, which imply a northward shift in the Inter Tropical Convergence Zone (ITCZ). In the North Atlantic, the ITCZ undergoes a northward shift in all seasons. This is likely due to the enhanced heating within the northern hemisphere modifying the northward global energy transports, a modified Hadley circulation, and a consequent northward displacement of the ITCZ (see Kang et al. 2008).

This ITCZ displacement response is accompanied by an enhanced African monsoon in summer (JJA, Fig. 4c), increased precipitation in northern South America in winter (DJF, Fig. 4a) and spring (MAM, Fig. 4b), followed by reduced precipitation throughout South America during the remainder of the year.
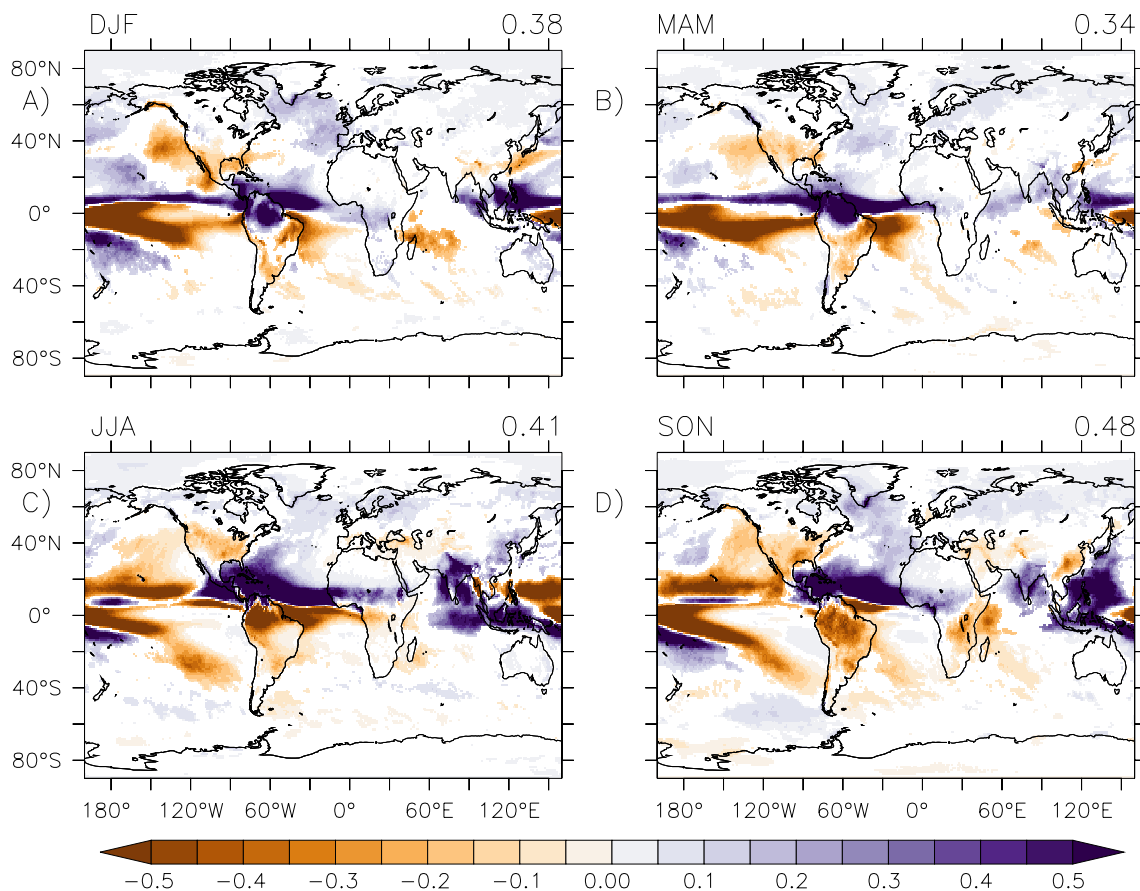
Fig. 4 As Fig. 2, but for precipitation (pr). Units: mm/day

Precipitation is reduced over North America in all seasons, with summer (JJA, Fig. 4c) and autumn (SON, Fig. 4d) showing the most widespread reductions.

Such reductions may be driven by enhanced subsidence over the region associated with the circulation response over North America (Fig. 3) as noted above, and described by Ruprich-Robert et al. (2018). There are small increases in precipitation over Europe (Fig. 4), mainly over southern and western Europe, but peaking over northern Europe and Scandinavia in summer (JJA). These increases may be the result of the circulation response driving a northward displacement of the Atlantic jet (Figs. 3 and SI section 3—Figure S12).

The AMV drives an increased Indian monsoon (JJA, SON; Fig. 4c, d), possibly driven by the increased land-sea temperature contrast between the Indian Ocean and the Tibetan Plateau (Zhang et al. 2004) (Fig. 2), together with an increased rainfall over eastern China in summer (JJA, Fig. 4c) followed by a reduction in autumn (SON, Fig. 4 d). The impacts over eastern China are consistent with the circulation response (JJA, Fig. 3c) modifying seasonal rainfall patterns and are consistent with an earlier studies using a single model (Monerie et al. 2019, 2020).

There are widespread precipitation anomalies across the Pacific Ocean. These are predominantly characterized as a northwards shift of the ITCZ and a southward shift of the South Pacific convergence zone (SPCZ). The reduced precipitation over the South and tropical Pacific is likely related to the cooler surface temperatures over the region (Fig. 2)—most likely this is a response to increased subsidence over the Pacific (Monerie et al. 2020). We discuss the link between this displacement of the ITCZ and the Pacific cooling further in section 4.3. These widespread monsoonal changes in rainfall are examined further in detail using a precipitation budget analysis (for a single model in the ensemble) by Monerie et al. (2019).

4.1.3 Mean sea level pressure

Figure 3 shows the widespread global impact of the AMV on global circulation. The large scale response is a low pressure anomaly over much of the AMV forcing region in the North Atlantic and high pressure anomalies over large parts
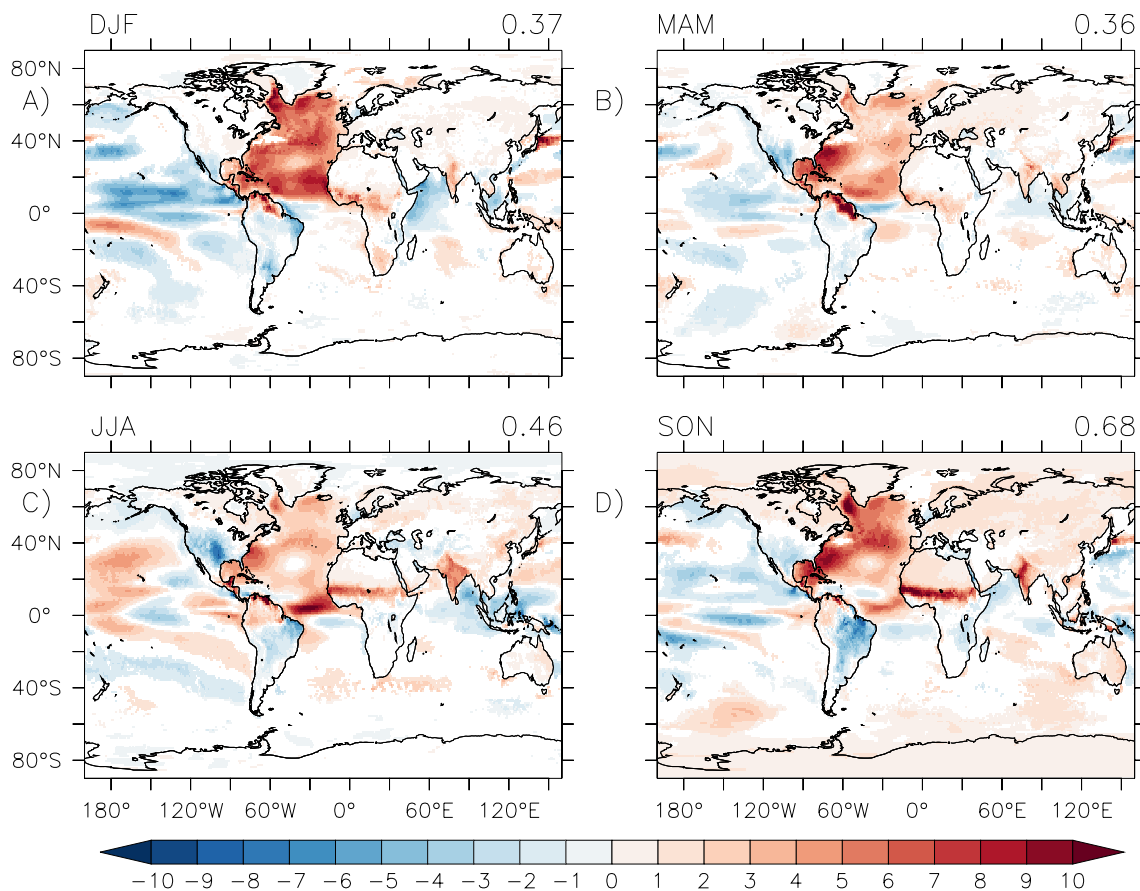
Fig. 5 As Fig. 2, but for surface upward latent heat flux. Units: W/m^2 (Note: values over land have *not* been multiplied by 10—unlike Fig. 6)

of the Pacific ocean; the most notable feature of which is the large high pressure anomaly in the Aleutian low region of the north Pacific. This feature persists throughout the year, peaking in winter (DJF, Fig. 3a) with a minimum in summer (JJA, Fig. 3c). Similar, but weaker, pressure anomalies over the north Pacific have been seen in previous atmosphere-only studies (e.g. Hodson et al. 2009; Sutton and Hodson 2007), suggesting the ocean may play a role in intensifying this response. Ruprich-Robert et al. (2017) and Lyu et al. (2017) demonstrate that that ocean atmosphere coupled feedbacks over the Tropical Pacific do enhance this AMV Pacific response

There are widespread circulation impacts over land. The circulation changes over Europe arising from the low pressure response (Fig. 3) as discussed above, may partly explain the weak response in surface air temperature in the region during summer and autumn (Fig. 2c, d); the increased northerly flow bringing colder polar air in the region, partly counteracting any local warming. The large Atlantic low pressure anomaly at 40°N during winter (DJF, Fig. 3a) leads to an equator-wards displacement of adjustment of the winter jet (SI section 2—Figure S12). Ruggieri et al. (2021) see a similar displacement of the jet in AMV-forced

experiments; although they find the extratropical Atlantic circulation response is a weak North Atlantic Oscillation (NAO-) pattern, whereas Fig. 3a projects more onto the East Atlantic Ridge Pattern (Cassou et al. 2004).

Over North America the circulation response peaks in summer (JJA, Fig. 3c). This response, a low pressure anomaly extending over much of the southern US, with a maximum over the Gulf of Mexico, is characteristic of the Gill off-equatorial pressure response to off-equatorial heating (Gill 1980); where a latent heat anomaly (due to increased precipitation) north of the equator drives a stationary Rossby Wave, a surface low pressure anomaly, northwest of the heating source. Such a summertime circulation response has been seen in previous atmosphere-only studies (Sutton and Hodson 2005; Hodson et al. 2009). Whilst, the intensification of the western Pacific subtropical high (Zhou et al. 2009) near 20°N in the Northwest Pacific during summer (JJA, Fig. 3c) is likely to drive significant changes in East Asian climate (Monerie et al. 2020).
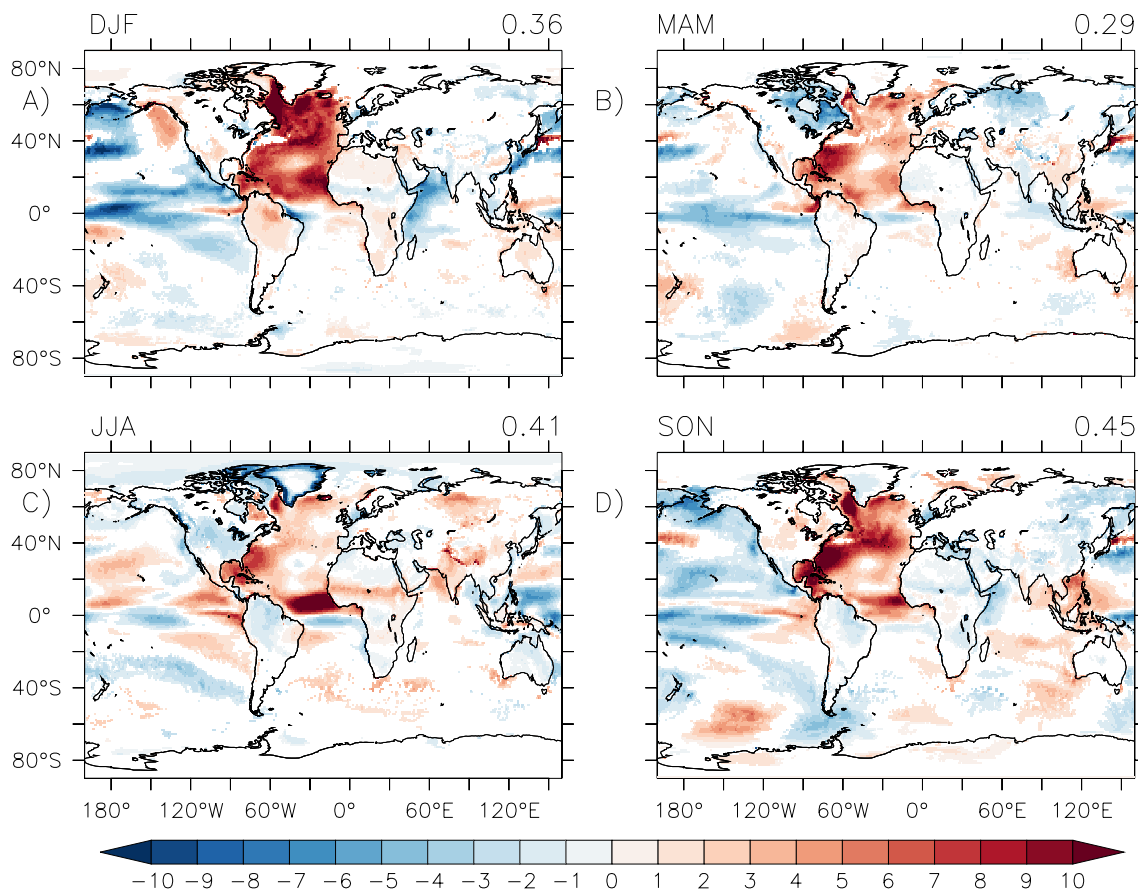
Fig. 6 As Fig. 2, but for net upward heat flux from the surface. Values over *land* have been **multiplied by 10** to aid comparison. Units: W/m^2

4.2 Surface heat fluxes

The heat flux nudging used to maintain the $2AMV^{+/-}$ forcing pattern (Sect. 2.3) adds heat to the top ocean model layer at each timestep. Some of this additional heat will be mixed down into the ocean layers below, but the majority will be released into the atmosphere above the Atlantic Ocean. Across the models, the AMV drives a strong surface upward latent heat flux from the Atlantic (Fig. 5), with a strong seasonal cycle and a maximum in the winter months (DJF). The net surface upward heat fluxes from all fluxes (Fig. 6—note, land values have been multiplied by 10), which also peaks in the winter months, are generally greater in magnitude than the surface latent heat fluxes, although the latent heat fluxes makes the largest contribution in most regions (SI section 4—Figure S13).

The amplitude of net upward heat fluxes over land are smaller than over the ocean, but generally consistent with the pattern of surface air temperature changes (Fig. 2); suggesting that changes in surface heat fluxes drive the surface air temperature response. However, in western Eurasia, Alaska,

and eastern Canada (Fig. 2a, b), where the surface flux contribution is weak, horizontal heat flux convergences within the atmosphere must play a greater role.

Changing surface shortwave fluxes drive surface air temperature responses in some regions. The summer and autumn cooling over sub-Saharan Africa and the Indian subcontinent, together with the northern South American cooling during winter and spring (Fig. 2) are all associated with a reduction in downwelling surface shortwave (Fig. 7). Similarly, the year-round warming response of North America and the spring (SON) warming over South America (Fig. 2) are both accompanied by increased surface shortwave. Both these responses are co-located with corresponding responses in precipitation (Fig. 4) and hence the positive (negative), surface air temperature anomalies are likely driven by reduced (increased) precipitation and cloud cover, leading to an increase (decrease) in downwelling shortwave, together with reduced (increased) upward surface latent heat fluxes due to reduced (increased) soil moisture which follow the circulation response to the AMV (see Ruprich-Robert et al. 2018).
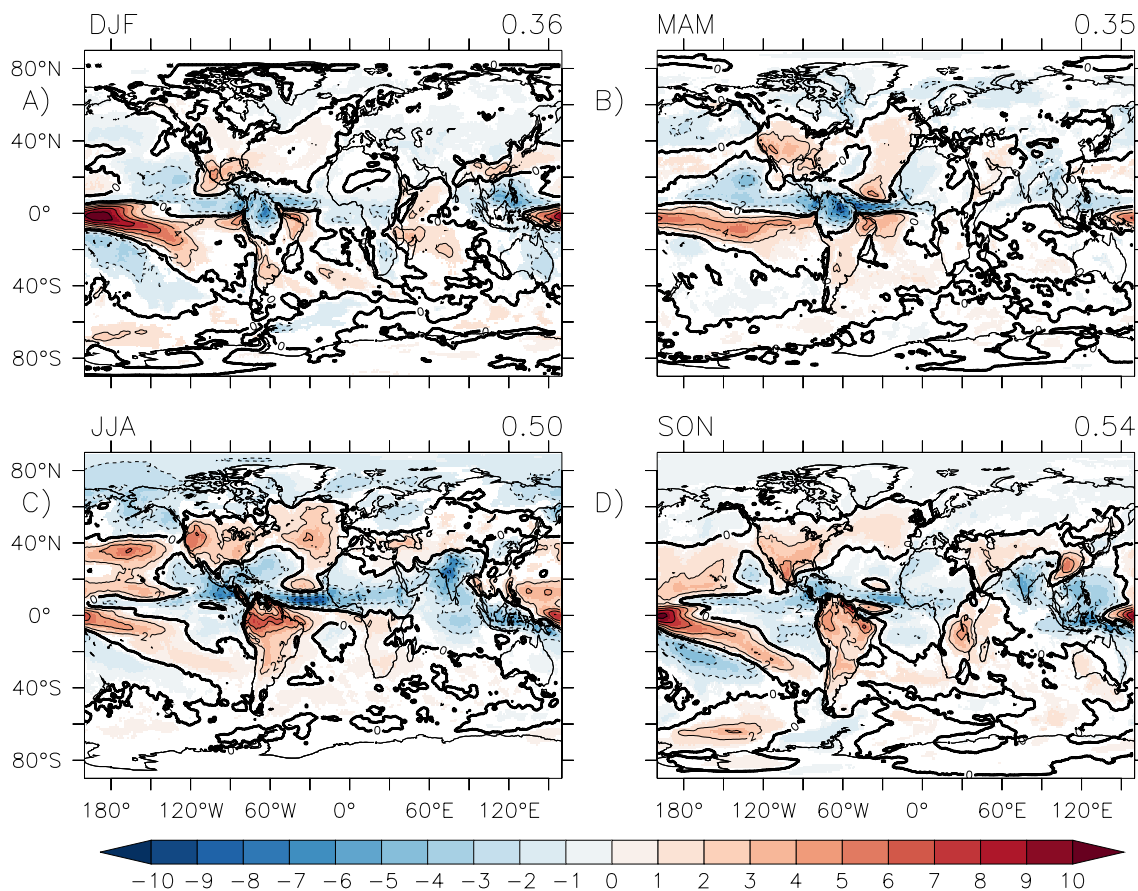
Fig. 7 As Fig. 2, but for downwelling surface shortwave (rsds). Units: W/m^2. Contours: $2\ W/m^2$

4.3 Pacific ocean response

Over much of the Pacific Ocean (especially the tropical region), there is a net downward flux into the Pacific Ocean (Fig. 6), demonstrating that the heat flux forcing from the Atlantic is largely being absorbed by the Pacific Ocean (stronger in winter (DJF), weaker in summer (JJA)), with the western Indian Ocean also contributing in winter (DJF). Comparing the surface air temperature (Fig. 2—or SST SI section 2—Figure S10) and net heat flux in the region, it is clear that the eastern and central Pacific cooling is not a direct response to net surface heat fluxes—which are generally acting to *warm* the ocean surface. The cooling must partly arise from the ocean response—for example, increased upwelling of colder subsurface waters, driven by changes in the surface winds (via Ekman pumping).

This widespread east Pacific cooling (Fig. 2) is a robust feature of many coupled model studies examining the impact of a warmer Atlantic (McGregor et al. 2018; Kucharski et al. 2011; Li et al. 2016; Ruprich-Robert et al. 2021; Meehl et al. 2021). Overall, the Pacific response to the AMV can be understood as follows. Warmer North Atlantic SSTs (Fig. 2) increase the latitudinal SST gradient, driving the

ITCZ further north (Fig. 4). This displacement results in anomalies in convection and latent heating release aloft over the Tropical Atlantic. Such anomalies will drive changes in the Tropical Walker Circulation (Rodríguez-Fonseca et al. 2009; Kucharski et al. 2011), leading to enhanced descent, increased surface pressure (Fig. 3) and surface easterly wind anomalies over the eastern Pacific Ocean (Li et al. 2016; Meehl et al. 2021). These surface easterly wind anomalies drive surface cooling and increased Ekman upwelling of colder subsurface waters, leading to cooler SSTs (England et al. 2014; Li et al. 2016; Ruprich-Robert et al. 2021; Meehl et al. 2021). The resulting east-west temperature gradient may then be intensified further via the Bjerknes feedback (Bjerknes 1969; Wang 2018). Such a widespread Pacific cooling could be viewed as a global negative feedback to the imposed SST forcing; attempting to return the climate system to its mean state.

The Pacific cooling intensification will be accompanied by changes in subsidence and convection, which may drive mid-latitude changes. Further examination of the upper level circulation (Fig. 8) reveals a wavetrain originating from the Tropical Pacific with a path extending into the north Pacific and over North America; reminiscent of the Pacific-North
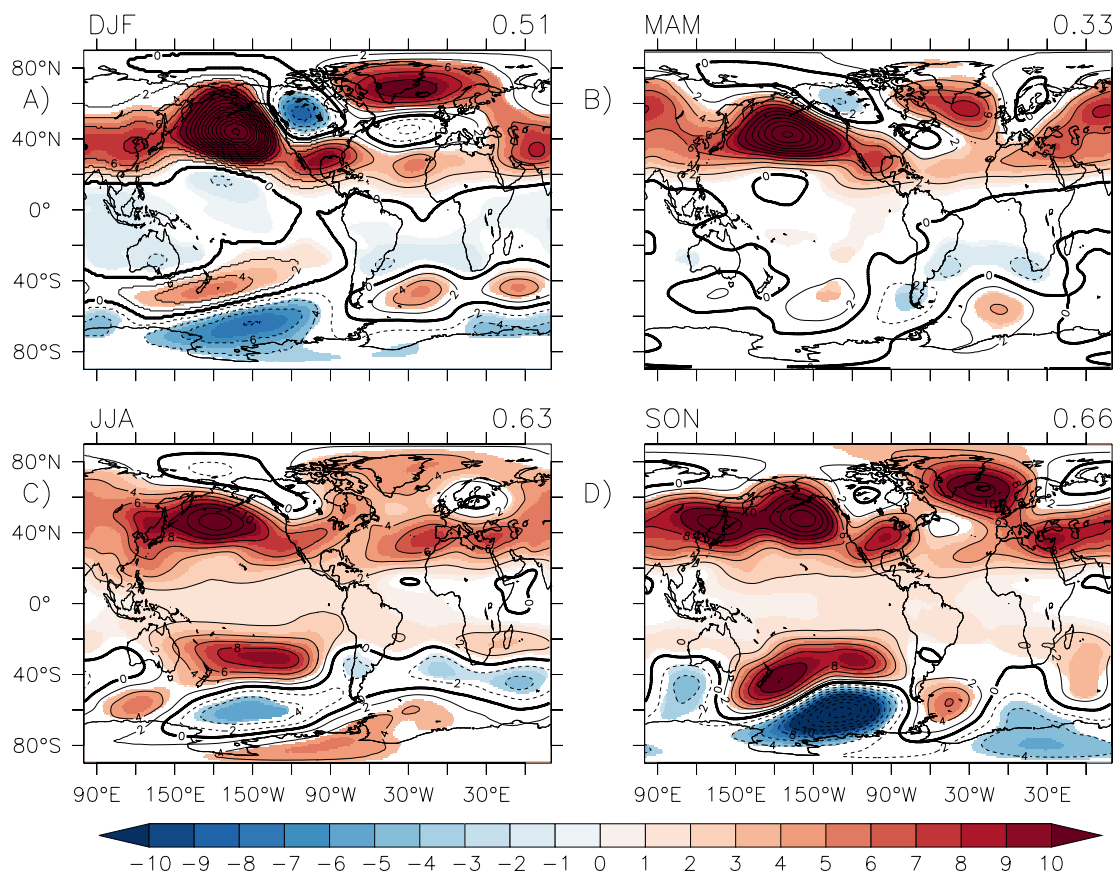
Fig. 8 As Fig. 2, but for 500 hpa geopotential height (zg). Units: *m*. Contours: 2*m*

American (PNA) pattern (Wallace and Gutzler 1981), and mirrored to a certain extent in the Southern Hemisphere. This upper level response is partly reflected in the surface pressure response in winter and spring (strongly over the north Pacific, weakly over Canada and Greenland—Fig. 3). Similar wavetrain responses were found in both CM2.1 and CESM1 by Ruprich-Robert et al. (2017), but there was a notable differences in the two models responses.

4.4 Model dependence

We now examine how the AMV response varies across the five models (averaging over resolution and ensemble members)—where do they significantly disagree in their response to the AMV? We must acknowledge that our five models are not a representative subsample of all current global climate models (e.g. Knutti et al. 2013), and so our assessment here is not a precise estimate of the model uncertainty in the AMV response, but more an assessment of which regions and aspects of the response show significant model uncertainty. We examine this dependence using the A_{em} term in the ANOVA model (Sect. 3). If $A_{em} > 0$, this implies that

the *spread* of the individual model responses around the multi-model mean response (Figs. 2, 3, 4) is greater than the internal variability (ϵ—Eq. 3). Since there are more than two models, we examine the fraction of the total variance (FVE_A—see Appendix A) in X_{emrj} that is explained by A_{em}, rather than the mean difference between $2AMV^+$ and $2AMV^-$ as we examined above. When this fraction is large compared to the fraction of the total variance explained (FVE_α) by the experiments (α_e: $2AMV^+$, $2AMV^-$) we can conclude that there is considerable model spread around the multi-model mean response. We therefore show the ratio of the two (FVE_A/FVE_α) in the following figures. We plot this ratio where A_{em} is a significant effect in the ANOVA ($p < 0.05$). Additionally, we stipple shade where A_{em} is a significant ($p < 0.05$), but α_e is not ($p \geq 0.05$). This allows us to examine where the significant responses in Figs. (2, 3, 4) are highly consistent across models, or subject to greater model spread. The following results are also robust to sub-sampling (see SI section 1.2).
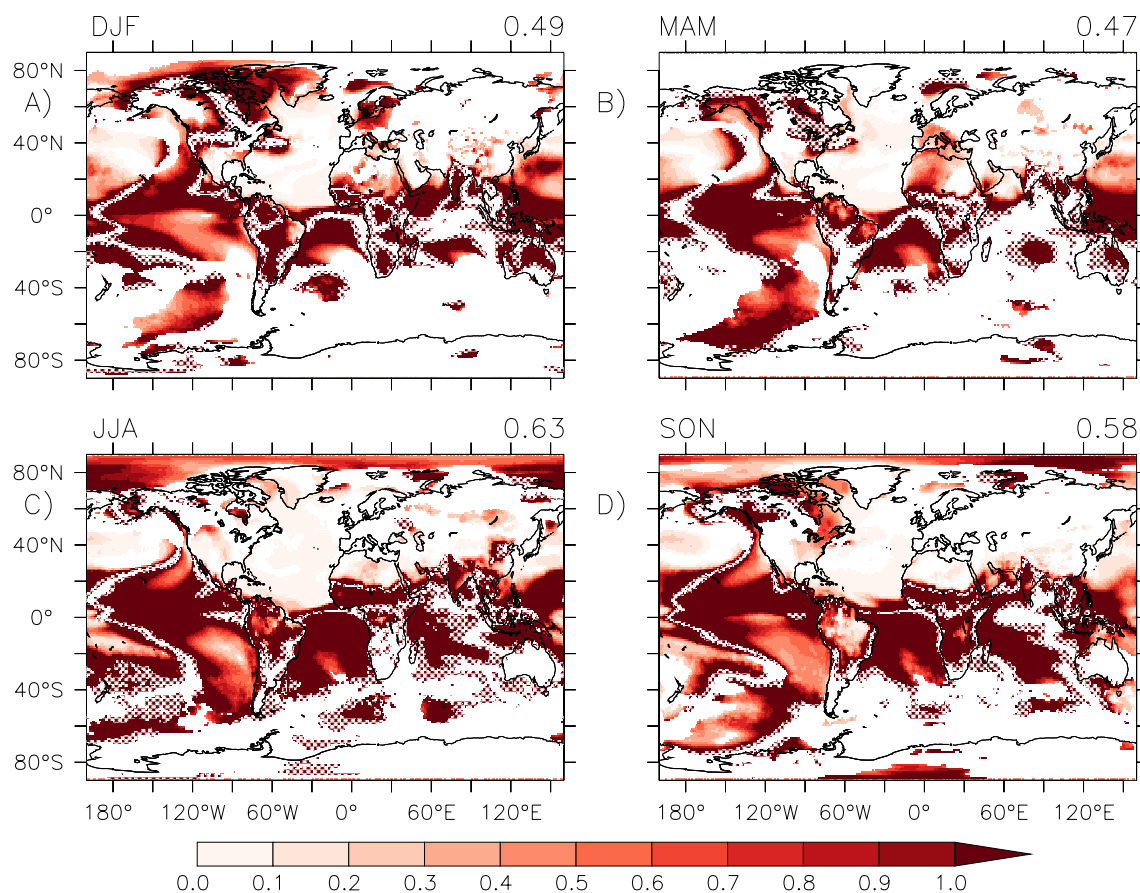
Fig. 9 Impact of model choice on the seasonal mean Surface Air Temperature (tas) AMV response. For each season we plot the fraction of the total variance in the ensemble explained by the model-experiment term (A_{em}, Sect. 3 eq. 3) divided by the fraction of the total variance explained by the experiment term (α_e). Dark red regions hence show where the models disagree on the AMV response. Shading shows where A_{em} is significant ($p < 0.05$). Stippling show where A_{em} is significant, but α_e is *not*. **A** Winter (Dec–Jan). **B** Spring (Mar–May). **C** Summer (Jun–Aug). **D** Autumn (Sep–Nov). Top right of each panel: fraction of the total number of gridpoints that are significant ($p < 0.05$)

4.4.1 Surface air temperature

Figure 9 shows variation of the surface air temperature (tas) responses to the AMV across the five models as discussed above, larger values highlight where the model spread in the AMV response is large compared to the model mean response.

Over the North Atlantic, model spread is low—implying that the SST nudging methodology (Sect. 2.3) is generally successful in applying a consistent SST anomaly across the models (assuming that SST and surface air temperature are closely coupled, which they appear to be (SI section 2—Figures S10)).

The tropics see significant model divergence in the response, particularly over the oceans. The Central Pacific and South Atlantic coolings responses (Fig. 2) show notable variation across models.

Over northern hemisphere land, models generally agree on the temperature (tas) response, with some exceptions over

North America in winter (Fig. 9a) and Europe and northern Africa in winter and spring (Fig. 9a, b).

There is larger model spread over tropical and southern Africa, and northern South America. These may be the result of variations in the extent of the northward displacement of the ITCZ across the models in response to the AMV. This northward displacement (Fig. 4) is associated with a band of cooling (Fig. 2)—a latitudinal spread in these cooling regions across the models will result in considerable spread in the model temperatures in the affected regions. The Arctic shows considerable spread in summer and autumn, suggesting a divergent response of the seasonal sea ice melting across the models.

Part of the large spread in the Pacific and South Atlantic temperature response may be due to a larger cooling response in MPI-ESM 1.2—excluding this model from the ensemble results in a smaller model spread and reduced Pacific cooling (SI section 6—Figures S29 and S32).
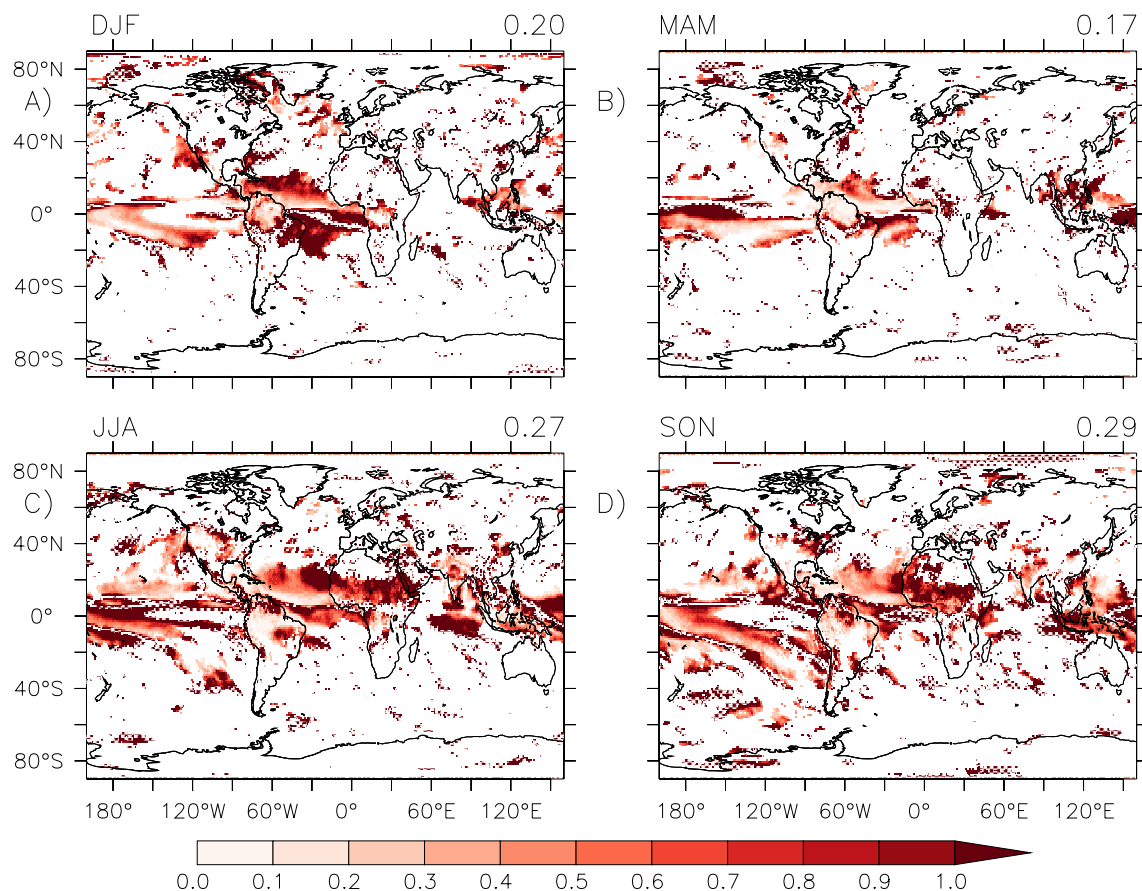
Fig. 10 As Fig. 9, but for precipitation (pr)

Overall, the spread in temperature responses suggests the models generally agree on the northern hemisphere extratropical response to the AMV (A_{em} is *not* significant—hence the spread in the responses across models is smaller than the internal variability ϵ in 3—see Appendix Eq. 10), however there is significantly less agreement over the tropical response (A_{em} *is* significant—the spread in the responses across models is larger than the internal variability ϵ.)

4.4.2 Precipitation

Outside the tropics, and particularly over land, the precipitation response to the AMV is generally consistent across the models (Fig. 10). Over the US, there is some notable model spread, particularly during summer (JJA).Within the tropics, model spread is greatest over the Tropical Atlantic throughout the year. This spread extends over Sub-Saharan northern Africa in summer (JJA) and autumn (SON) (Fig. 10c, d) and throughout the year over northern South America. This is likely due to variations in the northward displacement of the ITCZ across the models (Fig. 4), a common feature in AMV studies (e.g. Sutton and Hodson 2005; Mohino et al. 2011; Zhang and Delworth 2006). As ITCZ displacements

are driven by changes in the cross-equator SST gradient, model spread in the ITCZ may be driven by the spread in the South Atlantic SST response (Fig. 2).

The model spread over the Tropical Pacific and the Maritime continent, key regions of subsidence and ascent, suggest a spread in the models tropical Walker circulation responses to the AMV. Excluding the MPI-ESM1.2 model from the ensemble reduces this spread, but the overall pattern of spread is similar (SI section 6—Figures S31 and S34).

4.4.3 Mean sea level pressure

The local response to the AMV forcing over the North Atlantic is generally consistent between models (Fig. 11), but consistency is weaker over the Tropical Atlantic, particularly in summer and autumn (Fig. 11c & d); this may be related to the spread of the ITCZ response across the models (see above).

There is significant diversity across the models over the Maritime Continent, possibly as a result of varying ascent in the region due to a divergence in the response of the Walker Circulation to the AMV across models.
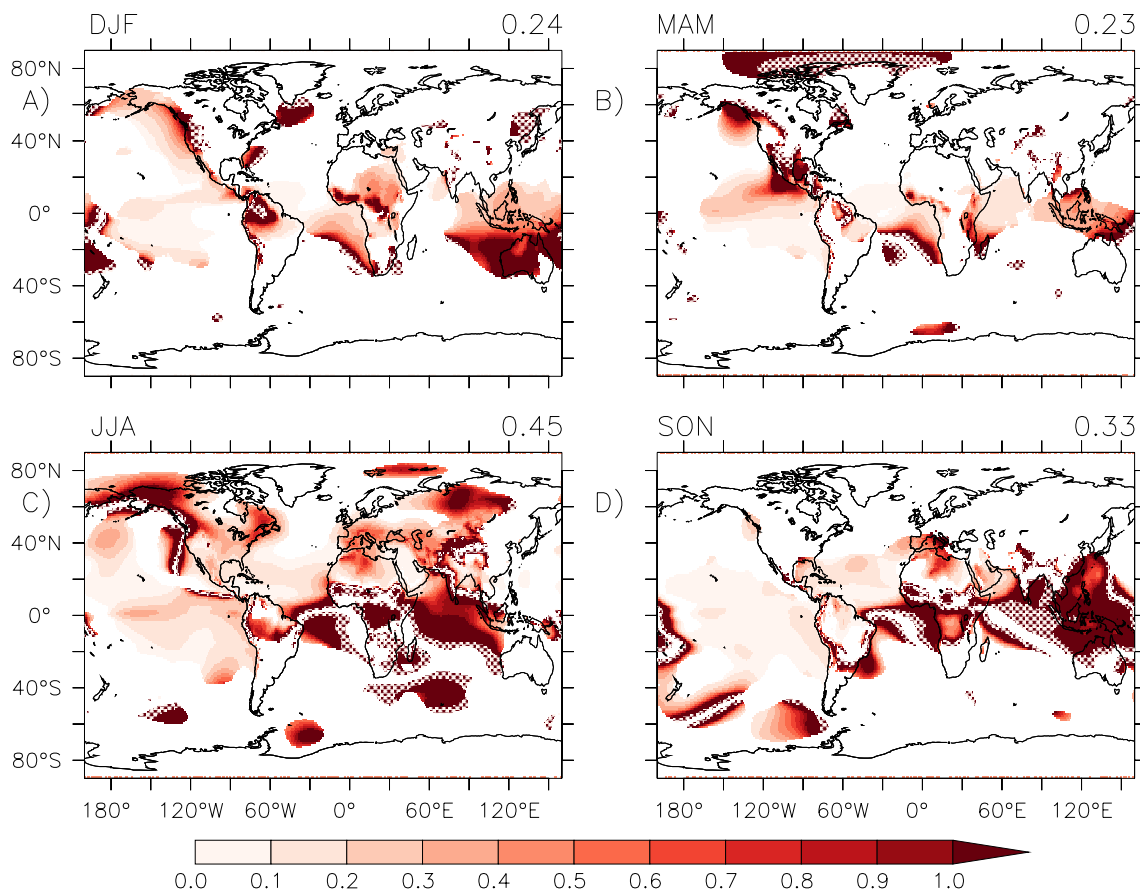
Fig. 11 As Fig. 9, but for mean sea level pressure (psl)

Summer (JJA, Fig. 11) is the season showing the greatest model diversity in the response, this season also shows the greatest model diversity in the circulation over the Indian Ocean, which may imply a diversity in the South Asian monsoon response.

4.4.4 Extratropical Wavetrain

The extratropical wavetrain response in our ensemble (Fig. 8) is generally consistent across models, as shown by the ANOVA for geopotential height in Fig. 12, but there are some inter-model variations in the strength and position of the negative Canadian and positive Icelandic 500 hpa height lobes in winter. These variations may arise from the differences in the strength of the Pacific response to the AMV across the models (Fig. 9), or also differences between models in their mean state upper level flow, which may modify the path of atmospheric Rossby wave propagation from the Tropical Pacific (Scaife et al. 2017). The impact of the AMV on the Pacific in these and other AMV experiments is discussed in depth in Ruprich-Robert et al. (2021).

4.5 Impact of resolution

Finally, we examine the impact that model resolution has on the modelled response to the AMV.

As horizontal model resolution is increased, both in the atmosphere and ocean, more small-scale processes are better resolved, sharp gradients associated with frontal systems and topography are better represented; models generally begin to better represent the observed climate. Increasing atmosphere resolution improves model climate; leading to reduced tropical biases (Jung et al. 2012), better representation of northern hemisphere blocking (Berckmans et al. 2013; Schiemann et al. 2017), leading to changes in eddy feedbacks and the representation of frontal structures (which can influence the position of the jet (Czaja et al. 2019)) and leading to enhanced moisture transport from ocean to land (Demory et al. 2014). Increasing the ocean resolution also improves multiple aspects of the ocean mean state compared to observations (Hewitt et al. 2017). The improved representation of ocean eddies and resolution of topography (Hurlburt et al. 2008) can lead to improved position and representation of the western boundary currents (Marzocchi et al. 2015) leading