https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

The CANARI HadGEM3 Large Ensemble: Design, and evaluation of historical experiments

Reinhard K. H. Schiemann[1], Grenville Lister[1], Rosalyn Hatcher[1], Dan Hodson[1], Bryan Lawrence[1], Len Shaffrey[1,2], Jeff Cole[1], Andrea Dittus[1], Jenny Mecking[2], Jon Robson[1], Simon Wilson[1], Yevgeny Aksenov[2], Adam T. Blaker[2], Ben Harvey[1], Oscar Martínez-Alvarado[1], Annette Osprey[1], Stephanie Rynders[2], Sharar Ahmadi[1], Jake Aylmer[3], Laura Baker[1], David Case[1], Emanuele Silvio Gentile[1], Steve George[1], Kevin Hodges[1], Eliza Karlowska[1], Charlotte Lang[1], Hua Lu[4], Niamh O'Callaghan[3], Weronika Osmolska[5], Scott Osprey[6], Tony Phillips[4], David Schröder[3], Robin Smith[1], Andrew G. Turner[1,3], Steve Woolnough[1], Martin Andrews[7], Andrew Coward[2], Steven Hardiman[7], James Harle[2], John W. Rostron[7], David M. H. Sexton[7], Bablu Sinha[2], Jonny Williams[1,8], and Chris Wilson[9]

[1]National Centre for Atmospheric Science, University of Reading, Reading, United Kingdom
[2]National Oceanography Centre, Southampton, United Kingdom
[3]Department of Meteorology, University of Reading, Reading, United Kingdom
[4]British Antarctic Survey, Cambridge, United Kingdom
[5]National Centre for Atmospheric Science, University of Leeds, Leeds, United Kingdom
[6]National Centre for Atmospheric Science, University of Oxford, Oxford, United Kingdom
[7]Met Office Hadley Centre, Exeter, United Kingdom
[8]Earth Sciences New Zealand, National Institute of Water and Atmospheric Research, Wellington, New Zealand
[9]National Oceanography Centre, Liverpool, United Kingdom

Correspondence: Reinhard Schiemann (r.k.schiemann@reading.ac.uk)

Abstract. This paper introduces the CANARI HadGEM3 Large Ensemble (CANARI LE), a 40-member Single-Model Initial-condition Large Ensemble (SMILE) designed to advance the study of climate variability and change. It is based on HadGEM3-GC3.1, the physical climate model used by the UK Met Office and partners for submissions to CMIP6 (Climate Model Intercomparison Project Phase 6). The CANARI LE simulations span 1950–2100 and are driven by CMIP6 historical and SSP3-7.0 forcings. Compared to existing Large Ensembles, the CANARI LE offers several distinctive features: (i) its resolution (1/4° in the ocean, 60 km in the midlatitude atmosphere), (ii) sub-daily output at the surface and on atmospheric pressure levels, and (iii) the provision of boundary conditions for regional modelling. In this study, the design of the CANARI LE is documented and key aspects of historical model performance are evaluated, while applications of the SSP3-7.0 simulations are left for forthcoming studies. Model evaluations are carried out for global mean surface temperature, North Atlantic jets and storm tracks, Northern Hemisphere stratospheric circulation, Atlantic meridional overturning circulation and subpolar gyre, Arctic sea ice, and El Niño-Southern Oscillation teleconnections. In addition, an example is provided deriving analogues of the North Atlantic 2023 summer marine heatwave showing that the CANARI LE captures unprecedented events of this magnitude and spatial scale. The CANARI LE is a community resource that lends itself for the study of physical coupled climate processes

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

and weather systems in a changing climate. The ability to drive regional models will potentially allow for risk assessments of unprecedented detail in climate impacts and adaptation studies.

1 Introduction

Internal variability is a major source of uncertainty in climate projections and is irreducible (Hawkins and Sutton, 2009; Knutti and Sedláček, 2013). These facts have triggered the development of Single-Model Initial-Condition Large Ensembles (SMILEs), hereafter Large Ensembles (LEs), of climate model simulations, which allow quantification and exploration of internal variability (Wood et al., 2021), separation of variability from forced climate change (Lehner et al., 2020), and better sampling of extremes (Fischer et al., 2021; Deser et al., 2024). The first LEs were ensembles of all-forcings experiments generated from small perturbations of the initial atmospheric state of coupled climate models (Kay et al., 2015). Over recent years, LEs have become state-of-the-art in climate modelling, and several modelling centres have produced LEs, diversifying and advancing them in the process. These advancements include so-called macro initialisation to better sample internal climate variability on different timescales across different model components (Rodgers et al., 2021), LEs for more climate scenarios, LEs for single-forcing and all-but-one forcing experiments in which single forcings are excluded (Maher et al., 2019; Smith et al., 2022; Simpson et al., 2023), and LEs using increasingly complex Earth system models to study a wider range of processes and impacts (Kirchmeier-Young et al., 2017; Maher et al., 2019; Rodgers et al., 2021).

Here, we present the CANARI HadGEM3 Large Ensemble (CANARI LE). This LE has been produced under the umbrella of the CANARI (Climate Change in the Arctic-North Atlantic Region and Impacts on the UK) research programme. The aim of CANARI is to advance understanding of the impacts on the UK arising from climate variability and change in the Arctic-North Atlantic region, with a focus on extreme weather and the potential for rapid, disruptive change. Although the CANARI LE is a general resource with global applicability, the design choices made in producing the CANARI LE reflect the aim of the CANARI research programme: Its comparatively high resolution, both in the atmosphere and ocean, as well as sub-daily output at the surface and on several atmospheric levels allows for a better representation and more detailed study of weather systems such as blocking and cyclonic storms (see Sect. 3.2 for an evaluation of midlatitude storms tracked in the CANARI LE), and Atmospheric Rivers (McGetrick et al., 2026). Furthermore, we have stored boundary conditions to enable regional atmospheric modelling driven by the CANARI LE, which will allow, in future work, to simulate the most extreme events seen in the CANARI LE, such as windstorms and heavy precipitation, at convective-scale resolution. Another key motivation for storing these boundary conditions is to use them for forced-ocean or forced-ocean-sea ice simulations.

In this paper, we document the design of the CANARI LE and focus on key aspects of historical ensemble performance, while the future simulations will mainly be assessed elsewhere. The paper is structured as follows: Section 2 details the climate model, experiments performed, and available output. In Section 3 the ensemble performance is evaluated in terms of global mean surface temperature evolution, aspects of tropospheric and stratospheric circulation, North Atlantic ocean state and circulation, Arctic sea ice, and example simulated El Niño-Southern Oscillation (ENSO) teleconnections. An application of the CANARI LE investigating analogues of observed extreme North Atlantic sea surface temperatures is presented in Section 4.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

2 Large Ensemble description

2.1 The HadGEM3-GC3.1 climate model

The model used is HadGEM3-GC3.1 (Hadley Centre Global Environmental Model 3-Global Coupled Model 3.1; Williams et al., 2018), the components of which are fully documented in Walters et al. (2019) for the atmosphere and land surface models (GA7.1 and GL7.0), in Storkey et al. (2018) for the ocean model (GO6.0) and in Ridley et al. (2018) for the sea ice model (GSI8.1). HadGEM3-GC3.1 is the physical component of UKESM1 (UK Earth System Model version 1; Sellar et al., 2019), and HadGEM3-GC3.1 and UKESM1 are the two models that have been used for the UK submissions to the Climate Model Intercomparison Project Phase 6 (CMIP6; Kuhlbrodt et al., 2018; Roberts et al., 2019).

2.2 Large Ensemble design and experiments

The CANARI LE is designed to meet several competing demands on high-performance computing and data storage. High model resolution, both in the atmosphere and ocean, benefits the representation of midlatitude circulation and weather systems (Zappa et al., 2013; Schiemann et al., 2020; Athanasiadis et al., 2022), and potentially of teleconnections (Sexton et al., 2025), but limits the number of simulated years and ensemble members that can be afforded. We choose the N216 resolution in the atmosphere (about 60 km grid spacing at midlatitudes) and 1/4° resolution in the ocean (HadGEM3-GC3.1 at these resolutions is denoted as HadGEM3-GC3.1-MM). While CMIP6 historical/ScenarioMIP experiments are normally run for 1850–2100, the CANARI LE spans 1950–2100. The ensemble size is 40; the CANARI LE thus comprises 6000 simulated years.

The CANARI LE members were generated using a hybrid initialisation method that combines so-called macro and micro initialisation, following some existing Large Ensembles (Rodgers et al., 2021). The aim of macro initialisation is to generate members that differ in the initial state of all component models, notably in the phase of decadal modes of ocean variability. To achieve this, eight HadGEM3-GC3.1-MM CMIP6 historical experiments were initialised at different times of a pre-industrial (1850 forcings) control simulation and run for 100 years to 1 January 1950, the start date of the CANARI LE (four of these eight experiments were run specifically to generate initial conditions for the CANARI LE, another four had already been run as part of CMIP6). Micro initialisation refers to a small perturbation of the initial atmospheric state. For the CANARI LE, five such micro perturbations were applied to each of the eight macro initial conditions by setting a different random seed in the stochastic physics scheme of the atmospheric model, thus obtaining $8 \times 5 = 40$ initial conditions for the CANARI LE. The model experiments used for generating the eight macro initial conditions, the time they branched off the piControl, and the corresponding CANARI LE members are listed in Table 1.

For each of the 40 members, a CMIP6 historical experiment (Eyring et al., 2016) was run from 1950 to the end of 2014, and then continued as a CMIP6 SSP3-7.0 experiment to the end of 2099. Within CMIP6 scenario pathways, SSP3-7.0 is a high forcing scenario under a regional rivalry storyline with comparatively high levels of carbon dioxide, near-term climate forcers (aerosols, methane, tropospheric ozone precursors) and deforestation (O'Neill et al., 2016). The use of high forcing scenarios, and of models with high effective climate sensitivities, is debated (Hausfather and Peters, 2020; Hausfather et al.,

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

Table 1. HadGEM3-GC3.1 CMIP6 historical simulations used to generate eight macro initial conditions at 1 January 1950. Columns show the suite identifier of the model experiment, the nominal branch off time of the experiment from the 1850 piControl, and the CANARI LE members the macro initial conditions were used for. The suite of the piControl run is u-aq281.

Model suite	Branch time in piControl	CANARI LE members
u-as042	1850-01-01	1–5
u-as043	1885-01-01	6–10
u-as044	1930-01-01	11–15
u-as045	1960-01-01	16–20
u-cw263	2115-01-01	21–25
u-cw264	2165-01-01	26–30
u-cv247	2015-01-01	31–35
u-cv281	2065-01-01	36–40

2022; Swaminathan et al., 2024). SSP3-7.0 experiments in the CANARI LE will yield large signals for studying climate-change processes and impacts but avoid the highest and arguably unlikely SSP5-8.5 pathway.

2.3 Output and user access

The output available for each member is similar to that expected from a model experiment made available through CMIP6. Several atmospheric variables (zonal, meridional and vertical wind, specific humidity, temperature, geopotential height) have been saved every 6 hours at nine pressure levels to facilitate the study of weather systems such as cyclone tracking. At the surface, a range of variables including surface and air temperature, precipitation, 10 m and 100 m wind components, specific humidity, and different surface fluxes are available at sub-daily frequency to allow for detailed characterisation of physical impacts and to drive impact models such as in hydrological and energy meteorology applications. Data access is currently described at https://ncas-cms.github.io/canari.

CANARI LE ensemble diagnostic and prognostic variables have been written in CF-compliant NetCDF, with additional metadata which conforms approximately to the CMOR standard used for CMIP6 (Taylor et al., 2022). These data include boundary conditions which conform to the COordinated Regional Downscaling EXperiment (CORDEX) requirements. Monthly atmospheric and yearly ocean restart files have been saved for those whose regional or ocean/sea ice-forced models might want to take advantage of these data. The data are organised into three datasets: the priority variables (440 TB), which are kept online, the additional variables, and the regional model-enabling variables. The latter two datasets are currently available on tape, with extracts migrated to disk as necessary. The entire dataset comprises 6 PB.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

3 Key aspects of ensemble performance

3.1 Global mean surface temperature

Global temperature is a key variable describing climate variability and change and thus essential in model evaluation. Our evaluation follows Suarez-Gutierrez et al. (2021): Global mean surface temperature is defined as as the weighted average of near-surface air temperature over land and sea surface temperature over the ocean. We use the annual-mean global-mean surface temperature anomaly with respect to the 1961–1990 mean in the HadCRUT5 observational product (Morice et al., 2021), and calculate the same temperature anomaly in the CANARI LE by re-gridding model data to the HadCRUT5 grid and taking averages only over grid boxes with data in HadCRUT5 for the year considered. (For simulated years later than the HadCRUT5 data period (1850–2023), we evaluate at all grid boxes where HadCRUT5 has data in 2023.) This approach allows comparison of results from the CANARI LE with those from some existing Large Ensembles (Suarez-Gutierrez et al., 2021). The approach therefore differs from that used in Andrews et al. (2020) for evaluating HadGEM3-GC3.1 CMIP6 historical experiments in that the reference period for anomalies is different (1880-1919 in Andrews et al., 2020) and in that we do not remove model drift.

The global mean surface temperature evolutions in HadCRUT5 and in the CANARI LE are shown in Fig. 1a. The level of agreement between CANARI LE and observations in the historical period is comparable to that seen in Large Ensembles generated with other climate models (Suarez-Gutierrez et al., 2021). The CANARI LE temperature evolution shows faster cooling than HadCRUT5 during 1950–1975 and faster warming than HadCRUT5 between 1975 and 2014. The faster cooling may be due to too strong aerosol forcing in HadGEM3-GC3.1 (Andrews et al., 2020). HadGEM3-GC3.1 is one of several CMIP6 models with a high climate sensitivity and the faster than observed warming during 1975–2014 has been attributed to several forcings and feedbacks (Andrews et al., 2019). Our CANARI LE results robustly show that these differences cannot be explained by internal variability only, i.e. these trends in any of the ensemble members are greater then the ensemble spread in the trends. Thus, these discrepancies are at least in part due to a combination of errors in forcing, model error, and observational error (see Simpson et al., 2025, for a detailed discussion on confronting simulated and observed trends).

The ranked histogram in Fig. 1b provides a summary evaluation of the global mean temperature evolution in the historical simulation period (1950–2014). The histogram counts the ranks of the HadCRUT5 values in the CANARI LE, e.g., there are four years for which HadCRUT5 is smaller than any of the CANARI LE members (rank 0) and there are three years for which HadCRUT5 is greater than all CANARI LE members (rank 40). If the observations were statistically indistinguishable from the CANARI LE members, the histogram of the ranks of the observations in the CANARI LE would be expected to be flat, i.e. all ranks would occur with equal probability. This is approximately the case, as can been seen by the solid line which shows a running mean of the ranks and lies broadly within the sampling uncertainty bounds indicated by the dashed lines. There is, however, a tendency of a U-shaped ranked histogram, i.e. the HadCRUT5 observations take low and high ranks more often than would be expected by chance. This could be due to an overestimation of interannual variability of global mean surface temperature in the CANARI LE and due to the differences from HadCRUT5 in the simulated trends discussed above. Inspection of the time series (Fig. 1a) suggests that variability is reasonably well represented, so that the trend differences discussed in the previous paragraph are likely the main reason for the U-shape.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

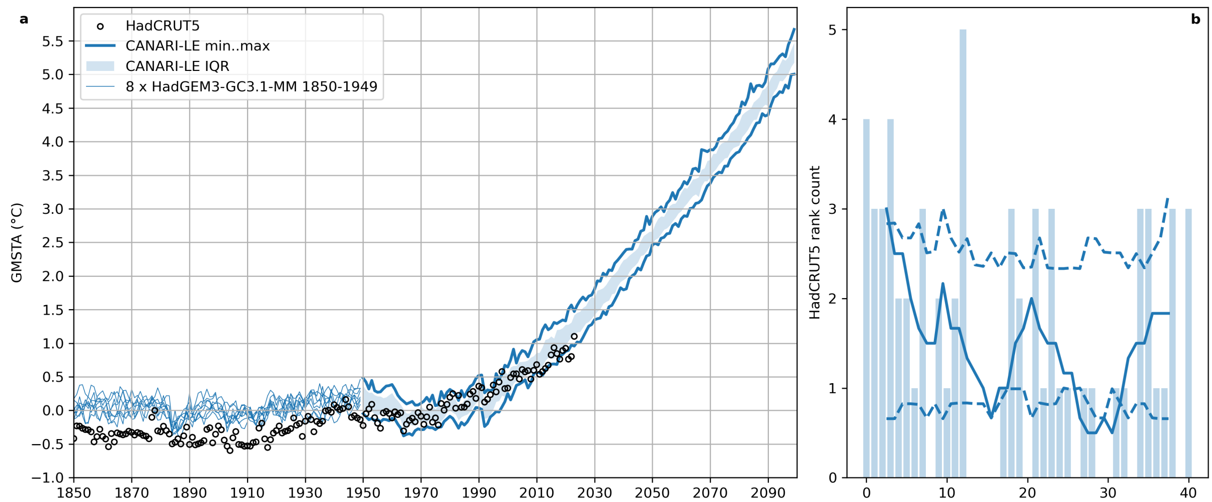

Figure 1. (a) Global-mean annual-mean surface temperature anomaly from the 1961–1990 mean in HadCRUT5 observations (black circles, 1850-2023) and in the CANARI Large Ensemble. Bold blue solid lines show the ensemble maximum and minimum, and blue shading the interquartile range, of the CANARI Large Ensemble during 1950-2099. Solid lines in 1850-1949 are for the 8 model runs used to generate macro initial conditions (Sect. 2.2). (b) Histogram counting the rank of HadCRUT5 in the CANARI Large Ensemble during 1950-2014 (bars, see text for details), the running mean of these counts over 6 adjacent bars (solid line), and a 90% resampling confidence interval (dashed lines) obtained from a perfect model approach, evaluated as the 5%–95% range of the ranks obtained when each ensemble member is compared against the rest of the ensemble, following Suarez-Gutierrez et al. (2021).

3.2 North Atlantic tropospheric jet and storm track

The North Atlantic jet plays a dominant role in the weather and climate of north-west Europe and eastern North America, and serves as a key dynamical link between the Arctic and mid-latitudes, connecting short-term variability with long-term trends. The seasonal mean lower-tropospheric winds, often termed the eddy-driven jet, are highly correlated with seasonal climate conditions and variability on shorter timescales and are intimately related to the occurrence of individual weather systems such as blocking and extra-tropical cyclones (ETCs). Here we examine the representation of the North Atlantic eddy-driven jet and ETCs in the CANARI LE, focusing on their biases, variability and trends relative to the ERA5 reanalysis (Hersbach et al., 2020).

The winter mean (DJF) lower-tropospheric winds in the CANARI LE have a spatial structure closely resembling ERA5, including a good representation of the strong SW-NE tilt characteristic of the North Atlantic jet. The ensemble-mean winds are slightly weaker than ERA5 (Fig. 2a, indicated by the 8 m s^{-1} contours) but individual ensemble members do encompass the ERA5 value (grey contours). As noted by previous authors, ERA5 exhibits a notable strengthening of wintertime North Atlantic winds during 1950-2014 (Fig. 2c; Blackport and Fyfe, 2022). This trend is not present in the CANARI LE ensemble mean which instead has a robust intensification of winds on its southeastern flank (Fig. 2b).

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

A key question is whether any individual members exhibit a trend as large as observed, and to address this we employ a simple wintertime jet speed index defined as the maximum value of the 0–60° W zonal mean DJF mean U850 profile for each year. This index exhibits a marked positive trend in ERA5 (Fig. 2d) which is not mirrored by either the LE ensemble mean (Fig. 2d) nor any of the LE members (Fig. 2e, blue). This behaviour is consistent with the CMIP6 historical simulations, none of which exhibit a trend as large as observed (Blackport and Fyfe, 2022). We note that there are large multi-decadal variations present in ERA5 to the extent that the trend over the shorter 1979-2014 period is near zero and well inside the LE spread (Fig. 2e, green). The realism and potential drivers of the observed variations over the 1950-1979 period remain under investigation, and examining analogues of any similar periods in the LE simulations will be the focus of future work.

Finally, we examine the inter-annual variability of the North Atlantic jet in the LE. Given the mismatch in multi-decadal variability already discussed, we apply the ranked histogram methodology from Section 3.1 to the detrended jet speed indices. According to this method, we find the inter-annual variability in the LE jet speed index is largely in line with ERA5 (Fig. 2f).

To examine ETCs, we employ the TRACK algorithm (Hodges, 1994, 1995) using the same methodology as Hoskins and Hodges (2002) and Priestley et al. (2023), with 850 hPa relative vorticity filtered to T42 spectral resolution as the tracking variable. We focus on the mean intensity of wintertime North Atlantic ETCs, and on trends over the satellite era (1979-2014). The mean intensity at each grid point is computed for each DJF season as the average intensity of all tracks passing within 5 degrees.

The mean intensity in the LE closely resembles ERA5 (Fig. 3a). There is a weak negative bias across the sub-polar gyre region, consistent with other CMIP models which on average underestimate ETC intensity (Zappa et al., 2013). There is notable similarity in pattern between the 1979-2014 trends in the LE ensemble mean (Fig. 3b) and ERA5 (Fig. 3c) with a weakening of mean intensity over the far north-east Atlantic in the Greenland-Iceland-Norwegian (GIN) seas region, and an intensification of mean intensity over the core of the storm track.

To examine the weakening trends in the GIN seas, we compute timeseries of area mean intensity over the region (Fig. 3d). Whilst the trend in the ensemble mean is negative, it remains weak in the sense that several members exhibit trends in the opposite direction (Fig. 3e). However, the ERA5 trend is of the same sign as the ensemble mean and is well within the spread. A potential explanation for the trend is that the reduction in sea ice in the region towards the end of the period (Section 3.5) decreases local baroclinicity, thereby weakening the intensity of cyclones.

3.3 Stratospheric circulation

The winter stratospheric polar vortex (SPV) is an important source of atmospheric variability (Baldwin and Dunkerton, 2001). Its weakening has been linked to Arctic sea-ice loss and can contribute to increased occurrences of extreme cold events in the mid-latitudes (Lee et al., 2019; Zhang et al., 2020). It is thus crucial to assess how well the CANARI LE reproduces the observed strength and variability of the SPV.

During the northern winter (DJF), the climatological zonal mean zonal winds of the CANARI LE show minimal spread across the 40 ensemble members in both the stratosphere and troposphere (Fig. 4). The LE also reproduces the broad latitude-height structure of the SPV and the easterlies near the equator. Nonetheless, the CANARI LE exhibits several systematic biases:

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

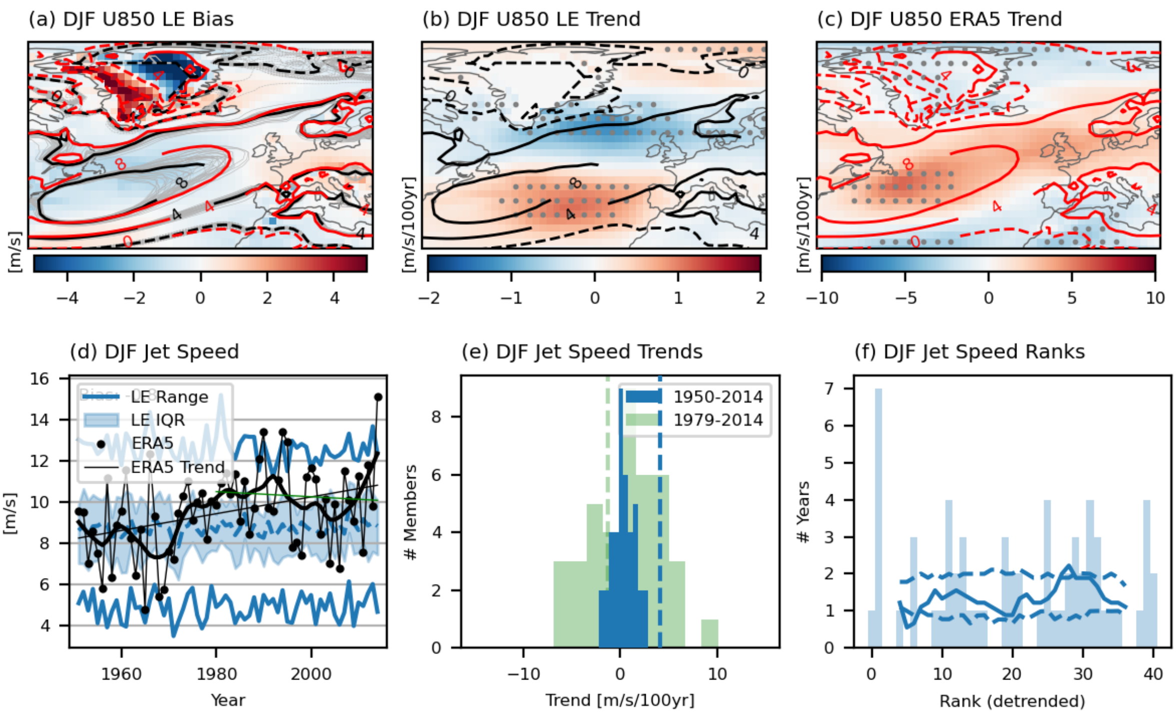

Figure 2. Tropospheric winds: bias, trends and variability. (a): DJF mean zonal wind at 850 hPa (U850) in ERA5 (red), the LE ensemble mean (black) and the LE members (grey), with shading showing the CANARI LE ensemble-mean bias. (b): Ensemble mean U850 trend in the LE over 1950-2014. (c): U850 trend in ERA5 over 1950-2014 (note wider colourbar than b). (d): Time series of 0–60° W jet speed in ERA5 (black) and the LE (blue; with median, interquartile range and full range indicated), the 11-year lowess-smoothed ERA5 timeseries (thick black; Cleveland, 1981) and the 1950-2014 and 1979-2014 linear trends in ERA5 (black, green). (e): Histogram of jet speed linear trends over 1950-2014 (blue) and 1979-2014 (green) and the corresponding ERA5 trends (dashed). (f): Histogram counting the rank of ERA5 in the CANARI Large Ensemble for the detrended jet speed timeseries during 1950-2014 (bars) with lines as in Fig. 1b.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

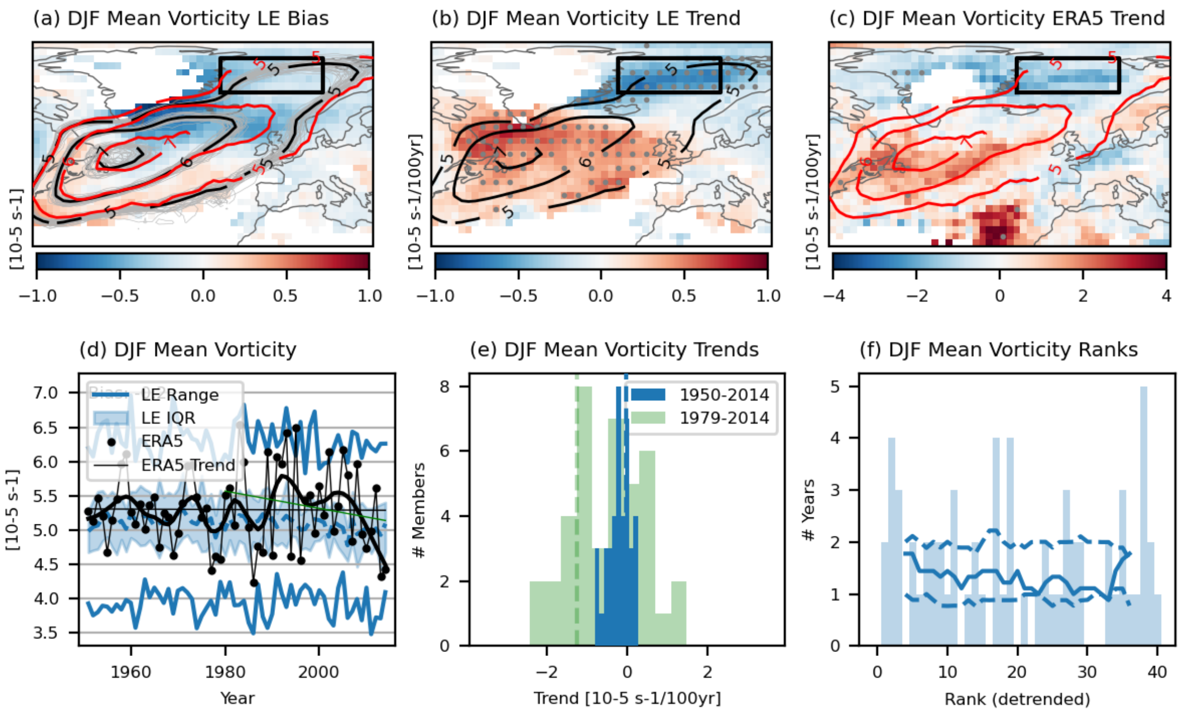

Figure 3. As Figure 2 but for the mean intensity of ETC tracks, as measured by the 850 hPa relative vorticity filtered to T42 spectral resolution. The index used in panels (d) to (f) is the area average of the mean intensity field over the box indicated in the maps.

in the lower stratosphere, the mean zonal-mean zonal wind $\overline{U}$ at 40–60° N, 70 hPa is approximately 3 m s^{-1} stronger than in ERA5 whereas in the upper stratosphere $\overline{U}$ at 30–50° N, 2 hPa is about 7 m s^{-1} weaker than observed. The simulated upper-stratospheric SPV is also 2–5° narrower than in ERA5. The CANARI-LE-simulated zero-wind line in the middle to upper stratosphere (20 to 1 hPa) differs significantly from ERA5. These upper-level biases are likely due to deficiencies in gravity wave parametrization in HadGEM3-GC3.1, as well as misrepresenting the semi-annual oscillation (Gray et al., 2022), variations in solar irradiance (Gray et al., 2010), and/or upper-stratospheric flow regimes (Lu et al., 2021). In contrast, the subtropical westerly jet near the tropopause is too strong, leading to an unrealistically tight dynamical coupling between the SPV and the subtropical jet.

Sudden stratospheric warmings (SSWs) can exert substantial impact on surface weather and climate (Bett et al., 2023). Their frequency and downward influence are key diagnostics of stratosphere-troposphere coupling. SSWs can induce widespread cooling over Eurasia and frequent cold events in North America by influencing the North Atlantic Oscillation (NAO) or the Northern Annular Mode (NAM) (Kretschmer et al., 2018; Lee et al., 2019; Zhang et al., 2020). The seasonal distribution of SSW occurrences observed in ERA5 lies within two standard deviation range of the CANARI LE during December to April

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

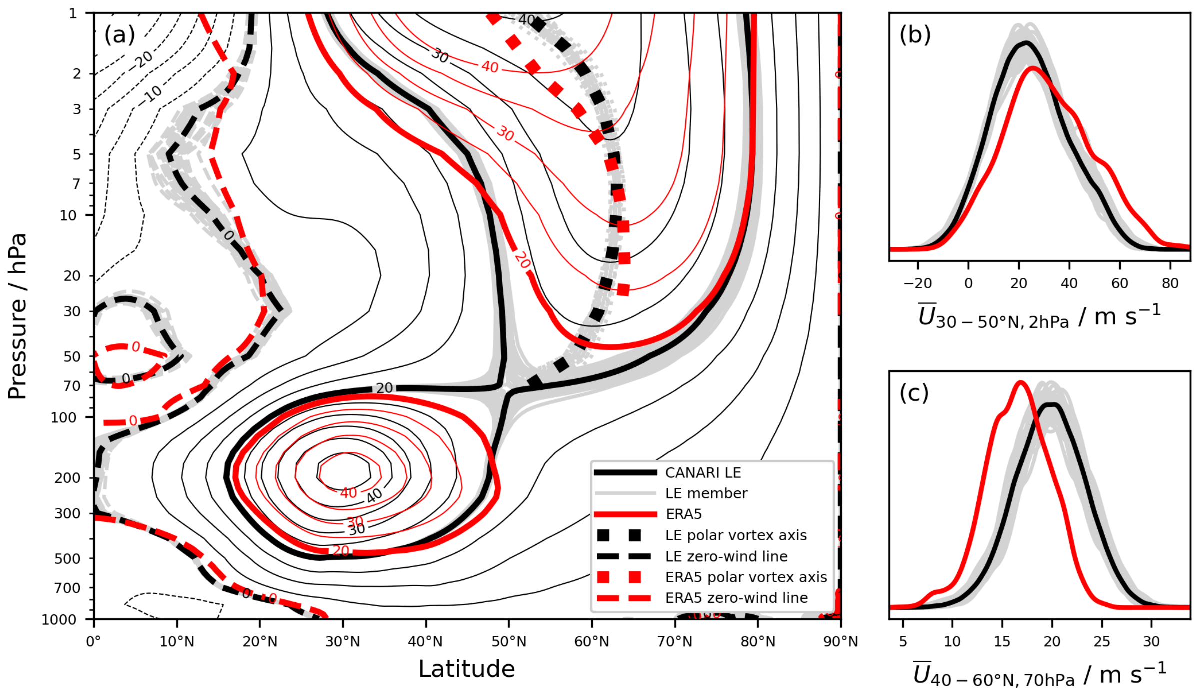

Figure 4. (a) DJF zonal-mean zonal wind $\overline{U}$ (m s^{-1}) from ERA5 (red), the CANARI LE ensemble mean (black), and individual ensemble members (grey). The dotted lines indicate the axis of the SPV (i.e., the latitude of maximum $\overline{U}$), while the dashed lines show the zero-wind line. The 20 m s^{-1} contours are thickened to highlight the SPV and the tropospheric subtropical jet. (b) Probability distribution of $\overline{U}_{30-50°\mathrm{N,2hPa}}$ estimated from ERA5 (red) and the CANARI LE (black and gray). (c) As in (b) but for $\overline{U}_{40-60°\mathrm{N,70hPa}}$.

(Fig. 5a). This suggests that the differences in mid- to late-winter SSW frequency between the CANARI LE and ERA5 can be plausibly attributed to internal variability. However, the CANARI LE significantly overestimates SSWs during October-November. Such early-winter SSW frequency biases are closely related to the seasonal development of the SPV (not shown).

SSWs are associated with negative NAM, characterized by positive height anomalies over the polar cap. The observed SSW signal in the NAM descends from the upper stratosphere to the lower stratosphere and then to the surface over a period of 3 to 25 days (Fig. 5). The CANARI LE reproduces the downward propagation of the SSW signal, although the tropospheric signal is noticeably weaker. It is important to note that Fig. 5c is based on 2329 SSWs from the CANARI LE, while Fig. 5b is based on only 35 observed events that occurred during 1979-2023, the period for which reliable upper stratospheric data are available. This difference in sample size likely accounts for the discrepancy between the LE-modeled and observed descent of the SSW signal, as well as the smoother evolution of the NAM in the CANARI LE. The weaker downward influence of SSWs in the CANARI LE may also contribute to signal-to-noise issues in NAO prediction on seasonal and decadal timescales (Scaife and Smith, 2018), as discussed by Bett et al. (2023) in the context of probabilistic analysis based on LEs.

Fig. 5a is broadly comparable to Fig. 9 in Menary et al. (2018), who analysed SSW distributions using the same model but at two different resolutions under a pre-industrial control configuration. Their high-resolution run matched the resolution

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

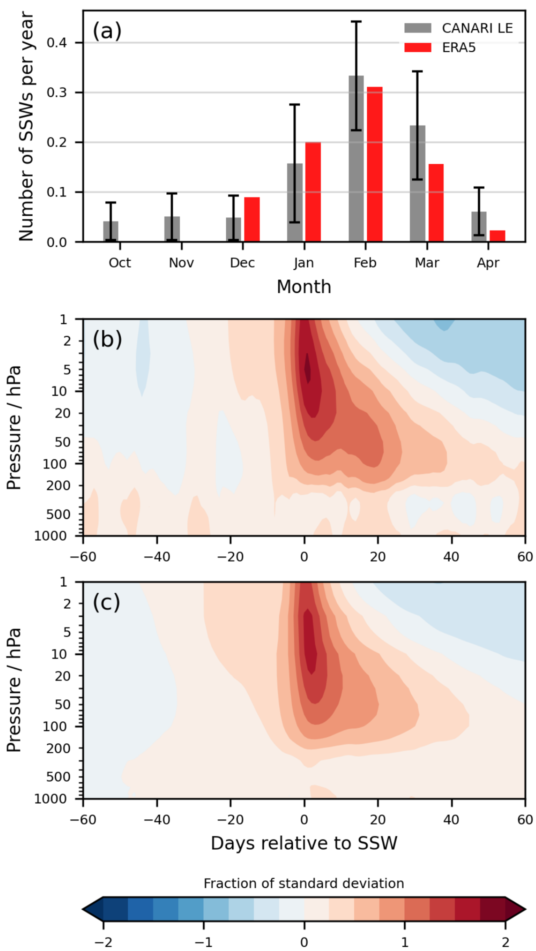

Figure 5. (a) Number of SSWs per year in ERA5 (red) and the CANARI LE (gray), with whiskers indicating the ±2-standard deviation range of the ensemble. SSWs are identified following Charlton and Polvani (2007). In total, 35 SSWs occur in ERA5 (1979-2023) and 2329 SSWs in the CANARI LE. (b–c) Composite evolution of SSW signal in the NAM for 60 days before and after the SSW onset date (lag = 0), over 1–1000 hPa, in (b) ERA5 and (c) the CANARI LE. The NAM is estimated based on the standardised zonal-mean geopotential height anomalies over the polar cap (60–90° N).

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

of the CANARI LE but consisted of a single 500-year realization. The main difference between our Fig. 5 and their Fig. 9 lies in the observational dataset: Menary et al. (2018) used ERA-Interim for the period of 1979-2016, whereas here we use ERA5 for the 1979-2023 period. Because SSWs are extreme events, changes in sample size in both observations/reanalysis and model simulations can significantly affect their seasonal distributions. The CANARI LE provides a closer comparison with observations/reanalyses in terms of SSWs due to its comparable climate and substantially large ensemble (2600 years versus 500 years in Menary et al., 2018). Menary et al. (2018) found that the simulated mid-winter SSWs tend to occur later than observed, with the delay being more pronounced in the high-resolution run. Nevertheless, higher model resolution reduces the model bias, improving both the frequency of early-winter SSWs and the strength of mid-winter SPV. These improvements are likely associated with a better representation of tropospheric processes that generate planetary-scale Rossby waves and/or a more realistic depiction of the stratospheric waveguide.

The unrealistic early concurrence of SSWs in the CANARI LE may be linked to an overly strong SPV in the lower stratosphere (Fig. 4), which generates excessive vertical wind shear near the tropopause at 40–60° N. This enhanced waveguide facilitates increased propagation of planetary-scale Rossby waves into the stratosphere during early winter, resulting in unrealistically more SSWs in October-November. The SPV then requires a few weeks to recover, which may explain the relatively lower number of SSWs in December in the CANARI LE than observed.

3.4 The North Atlantic state and circulation

The Atlantic meridional overturning circulation (AMOC) and the subpolar gyre (SPG) are the leading features of North Atlantic ocean circulation. Both transport heat and salt into and around the North Atlantic and play a significant role in maintaining and modulating the North Atlantic and European climate (e.g., Jackson et al., 2015; Moreno-Chamarro et al., 2017). We examine these circulations and transports in the CANARI LE, in particular by highlighting the members with the greatest and least post-1980 AMOC decline. We also compare with observations where available.

The ensemble mean AMOC at 26° N increases to a peak around 1980, before declining (Fig. 6a). This is consistent with the response to anthropogenic forcing, particularly the role of aerosol forcing seen in previous work (e.g., Robson et al., 2022). The 1980-2014 decline in the AMOC ranges between -0.5 to -1.6 Sv/decade (1 Sv = $1\times10^{6}\ m^{3}\ s^{-1}$) across the ensemble members; highlighting the large role of internal variability in shaping decadal-to-multidecadal change in this model. After 1980, the declining AMOC is broadly consistent with the available observations - except for the earlier part of the RAPID record. Over the common period (2004-2014) the ensemble mean AMOC is 15.5 Sv, ≈1.7 Sv weaker than RAPID (Moat, 2024) and less variable (annual mean standard deviation averaged over all members: 0.9 Sv (range: 0.88:1.01), RAPID: 1.4 Sv). The AMOC streamfunction peaks at ≈20 Sv between 50–55° N (when calculated in density space, Figure 6b), with a maximum ensemble spread in the lower limb, consistent with previous studies (e.g., Hirschi et al., 2013, 2020). The northward heat transport at 26° N evolves similarly to the AMOC (Fig. 6c), but the 2005-2014 mean is lower (1.0 PW) than the RAPID estimates (1.2 PW, Johns et al., 2023). The freshwater transport at 30° S also shows this peaked response, but the ensemble mean peak occurs later (≈2000) than the AMOC (Fig. 6d). The ensemble-mean northward transport at 30° S is also greater than all the available

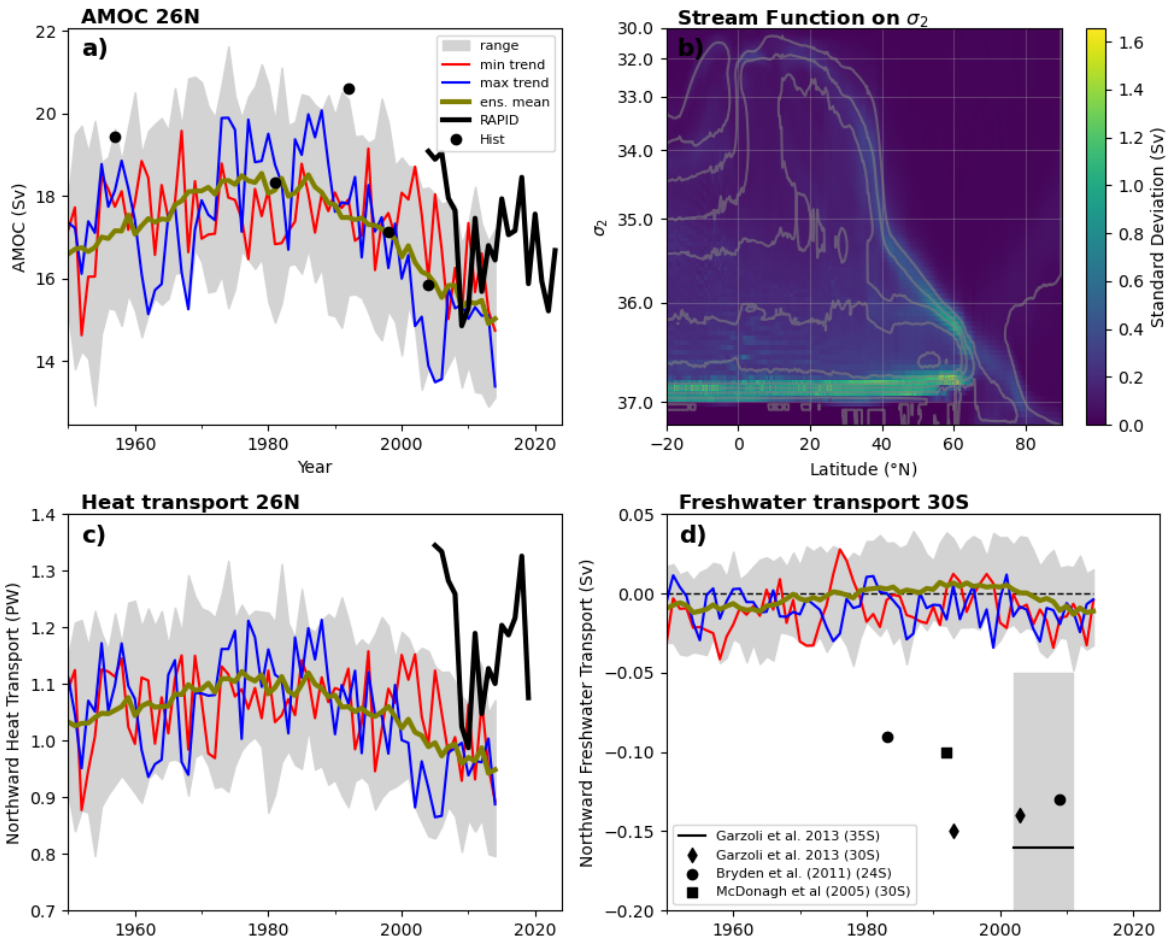

Figure 6. a) AMOC at 26° N (Sv). Olive green: ensemble mean. Shading is the ensemble range. Red (blue) member with least (greatest) AMOC decline over 1980-2014 :-0.5 (-1.6) Sv/decade. Black line: Observation-based estimate of AMOC strength (RAPID). Black dots: Historical section-based estimates of AMOC from Bryden et al. (2005), seasonally adjusted following the Kanzow et al. (2010) method using the April 2004 - February 2023 RAPID data. b) AMOC stream function on density levels (σ_2) (Sv) in grey contours, ensemble standard deviation of time mean (1950-2014) in colours c) Atlantic Ocean Heat transport at 26° N (PW). Red (blue) member with weakest (strongest) AMOC decline as in a) also shading as a). Black: observed heat transport from RAPID (Moat, 2024). d) As c, but for Northward Freshwater Transport (F_{ov}) at 30° S (colours and shading as c). The uncertainty range for the Garzoli et al. (2013) estimate extends below the y-axis minimum. Black symbols: F_{ov} estimates from observational studies (Garzoli et al., 2013; McDonagh and King, 2005; Bryden et al., 2011). 35° S estimate shows the 2002-2011 time mean and range.

observation-based estimates (by ≈0.1 Sv), becoming slightly positive around the peak. However, this bias is not unusual for coupled climate models, due to common salinity biases (e.g., Mecking et al., 2017).

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

Salinity in the SPG (defined here as the region 50–65° N, 0–60° W) peaks around 1980 before declining (Fig. 7b). The peak in temperature is much less pronounced (Fig. 7a). The member with the strongest post-1980 AMOC decline (blue) becomes warmer and more saline up until ≈2005, before rapidly cooling and freshening. The member with the weakest post-1980s AMOC decline warms without a strong trend in salinity. The model is colder (-0.1° C) and more saline (0.2 PSU) than observations (black 1960-2014), although this is not an unusual bias for climate models (Menary et al., 2015). Neither ensemble mean temperature or salinity are similar to the observed variability, indeed there is a high anti-correlation between the two. Horizontal ocean circulation can be characterized by the barotropic streamfunction (Fig. 7c), which displays significant variability of the western boundary of the SPG. The ensemble mean SPG spins up to a peak around the 1970s (Fig 7d) (preceding the AMOC peak - Fig. 6a) followed by a decline, somewhat consistent with the post-1995 decline seen in recent observed estimates of the SPG circulation (Häkkinen and Rhines, 2004; Berx and Payne, 2017), although there remains some uncertainty about how to best characterize SPG circulation (Koul et al., 2020). The decline in gyre circulation is larger in the member with the weakest AMOC decline (red) than the member with the strongest AMOC decline, suggesting that there is an anticorrelation between gyre and overturning circulations. Both these members also display large multidecadal signals that are not obvious in the ensemble mean.

3.5 Arctic sea ice

Time series of the Arctic sea ice extent are shown in Fig. 8 for September (Fig. 8a, annual minimum) and March (Fig. 8b, annual maximum) from 1950 to 2014. The CANARI LE is characterized by the shaded interquartile range and the minimum and maximum values for each year. The sea ice extent slightly increases from 1950 to the 1970s, with a maximum extent of around 8 million km^2 in September and 16 million km^2 in March. Afterwards, the ensemble mean decreases to 5-6 million km^2 and 14.5-15 million km^2 in the 2010s, respectively. The sea ice extent in the CANARI LE is consistent with satellite observations from 1979 to 2014 (Cavalieri et al., 1996; Comiso, 1999) with respect to mean values, seasonal cycle and trend. The annual-mean sea ice volume ((Fig. 8c) shows a similar temporal evolution with an ensemble mean maximum of 32 Mill km^3 around 1980 and a minimum of 17 million km^3 in 2014. The decrease of ice volume of 15 million km^3 is larger than the ensemble spread of 10 million km^3. Data from the Pan-Arctic Ice-Ocean Modelling and Assimilation System (PIOMAS; Zhang and Rothrock, 2003; Schweiger et al., 2011) reanalysis indicate that the volume is rather high in the CANARI LE, particularly in the 1980s.

The trend in Arctic sea ice extent is correlated with the trend in Arctic near-surface air temperature. This has been shown for CMIP6 models by Aylmer et al. (2024) and can be seen in the CANARI LE (Fig. 8d). All 40 members show a decrease in sea ice extent varying between 0.2 and 0.8 million km^2 per decade and an increase in air temperature of 0.5 to 1.1 K per decade. The spread is remarkable for a 35-year period and reveals that a verification of simulated trends for individual model simulations is problematic due to internal variability even on this time scale. While observations are within the range, the correlation between sea ice and temperature trend demonstrates that members with the most accurate sea ice trend overestimate the temperature increase. It seems plausible that the overestimation of ice thickness requires more atmospheric heating to reduce the observed sea ice extent.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

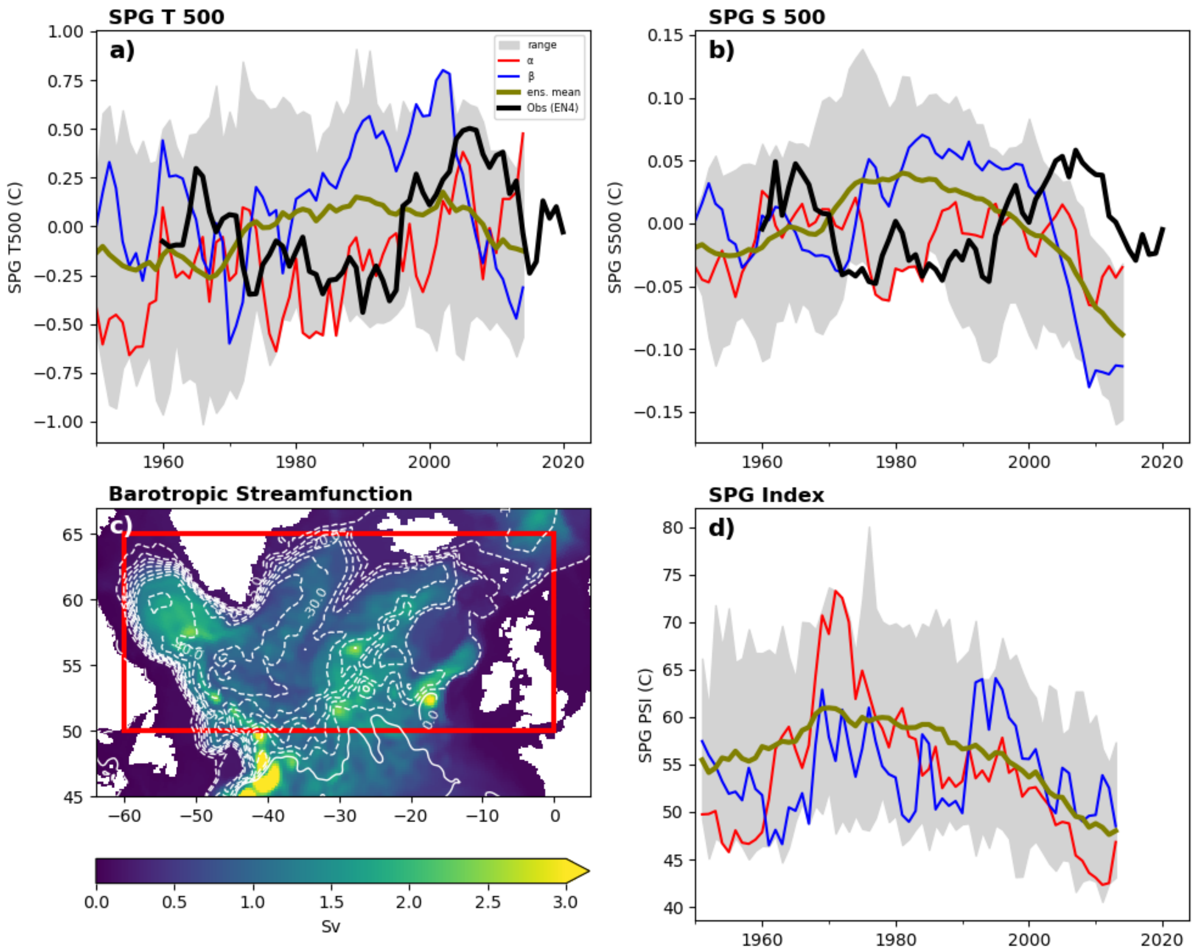

Figure 7. a) SPG (50–65° N, 0–60° W - red box in panel c) ocean temperature (mean over 0–500 m). Olive green: ensemble mean. Shading is the ensemble range. Red, α (blue, β) member with weakest (strongest) AMOC decline over 1980-2014. Black: observations (EN4) b) as a) but for salinity. c) Time-mean ensemble-mean barotropic stream function across the SPG (contours, interval 5 Sv). Shading: ensemble standard deviation of the time mean (2000-2014) barotropic stream function across the SPG. d) SPG circulation index (-1* the minimum of the barotropic stream function across the SPG at 60° N)

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

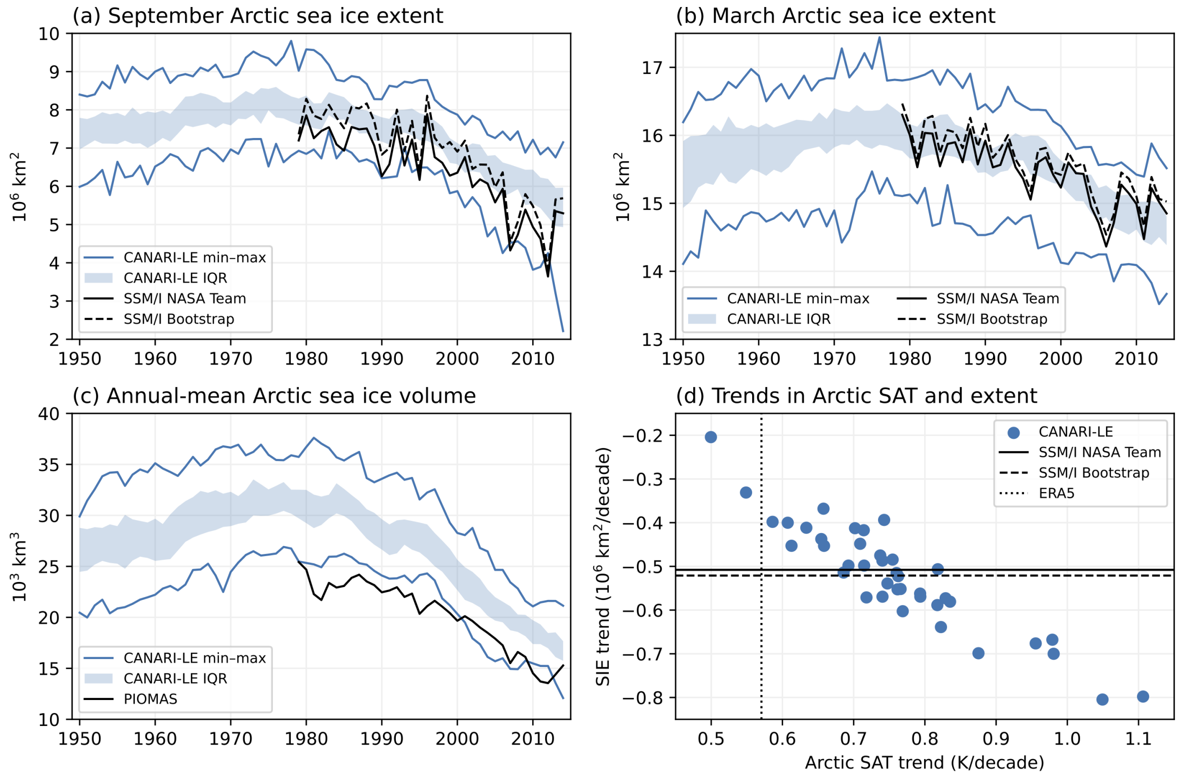

Figure 8. Arctic sea ice extent in (a) September and (b) March in the CANARI LE (blue) and as derived from passive microwave observations (black, solid and dashed). (c) Annual-mean Arctic sea ice volume in the CANARI LE (blue) and from PIOMAS (black; Schweiger et al. (2011)). In (a–c), blue shading indicates the ensemble interquartile range (IQR), and blue lines indicate the ensemble minimum and maximum with time. (d) Trend in annual-mean Arctic sea ice extent over 1979–2014 plotted against that of the Arctic surface air temperature (SAT), averaged northwards of 60° N. Each point is one ensemble member. Black lines indicate the corresponding sea ice trends from passive microwave observations (solid, dashed) and from ERA5 (dotted).

We conduct further regional analysis to compare the observed extreme retreat of sea ice in the Arctic Siberian Shelf Seas with the CANARI LE. Recession can potentially increase shelf ocean mixing, coastal erosion, and fluxes of nutrients and contaminants from land to ocean.

To conduct this analysis, we have assessed LE simulation quality by comparing sea ice extent in these regions calculated from the CANARI LE historical simulations 1950-2014 with data from digitised navigational sea ice charts (Mahoney et al., 2008) and from satellites (Cavalieri et al., 1996). The regions for comparison are the Kara, Laptev, East Siberian Seas and the western part of the Chukchi Sea – the areas where the accelerated sea ice retreat has been most pronounced in the last decades. The monthly average sea ice extent time series for October of individual ensemble members along with the ensemble mean have been chosen to represent sea ice conditions before the beginning of the freeze-up season.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

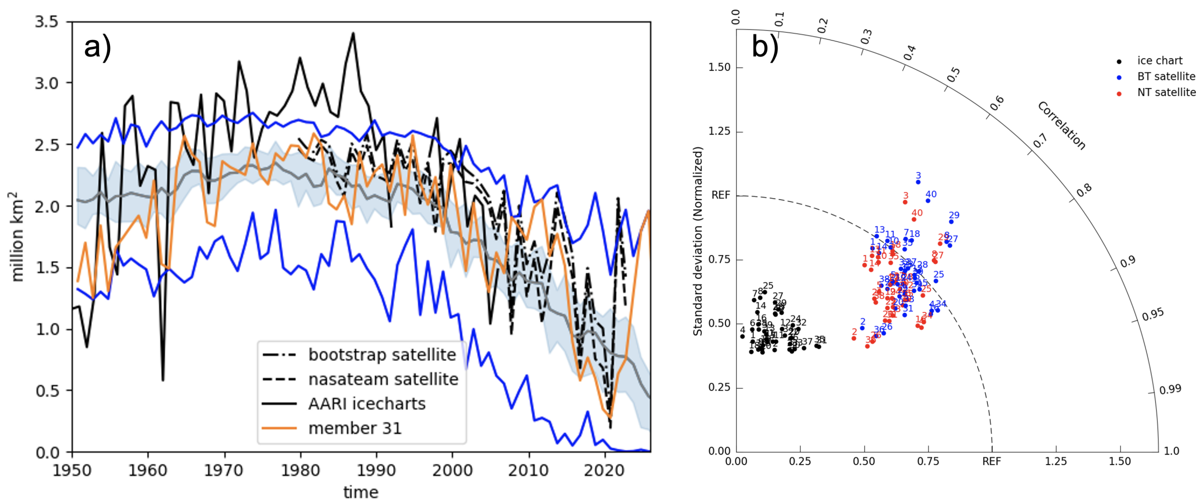

Figure 9. (a) Sea ice extent monthly-mean time series in October in the Siberian Seas (Kara, Laptev, East-Siberian and West Chukchi Seas) in CANARI LE simulations (1950–2025) and from the gridded AARI sea ice charts 1933-2006 and satellite products 1978-2022 (Mahoney et al., 2008). The grey line shows the ensemble mean; the dark blue lines and the light blue shading are the minimum, maximum and the range of one standard deviation for the ensemble members; the orange line is the ensemble member 31 with the best fit to the satellite data; the black lines depict observations. b) Taylor Diagram to compare model simulations and observations for the normalised standard deviation and correlation. Black symbols with ensemble member number compare the model with ice charts; blue and read symbols compare model with satellite data from the Bootstrap and NASA Team products.

We have checked model biases in mean, standard deviation and trends and occurrence of extreme low ice conditions. We have found that the simulations have lower standard deviation and higher mean for the pre-2000s period and less steep trend after that (Fig. 9).

3.6 Tropical El Niño–Southern Oscillation teleconnections

ENSO represents the most prominent mode of interannual climate variability with profound impacts across the tropics and some extending into the extra tropics. We test this here using linear regression and explore the ensemble spread. Fig. 10 shows the regression pattern of boreal summer precipitation onto the summer Niño-3.4 index as a measure of ENSO strength, meaned across all ensemble members. Since the focus here is on internal climate variability, the data is first detrended. The figure shows the expected response to ENSO in the Maritime Continent and equatorial Pacific, with convection shifting to the central Pacific, as well as the familiar pattern of weakened Indian monsoon rainfall associated with El Niño events (*e.g.* Shukla and Paolino, 1983; Webster and Yang, 1992).

The figure also shows the cross-ensemble variations in the strength of the regression pattern, as represented by the standard deviation of the regression coefficients at each location. Notable signals are found over the eastern Indian Ocean in the upwelling region off the coast of Sumatra, the Maritime Continent and West Pacific warm pool region, as well as over India.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

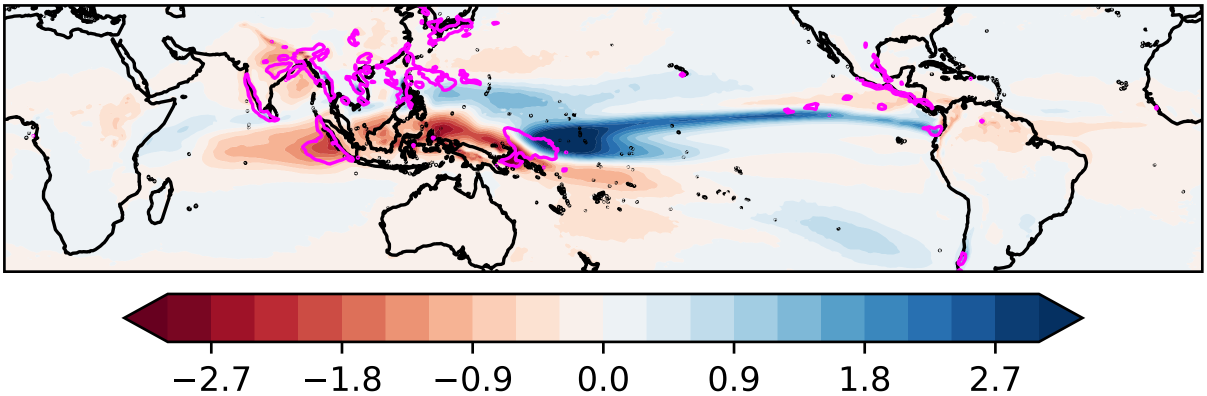

Figure 10. Linear regression of JJAS precipitation onto the JJAS Niño-3.4 index over time from 1950 to 2014 (units are $\mathrm{mm\ day^{-1}\ °C^{-1}}$). The ensemble mean across 40 members was removed from both the precipitation field and Niño-3.4 index prior to performing the regression, as a way to de-trend the data and to partly counteract the impact of external forcing. The regression was performed separately for each ensemble member before the mean was taken. The pink line corresponds to the $0.5\mathrm{mm\ day^{-1}\ °C^{-1}}$ contour of the standard deviation across the 40 regression patterns.

This suggests that there is considerable variation in the strength or representation of tropical ENSO teleconnection patterns dependent on the ensemble member chosen, that is, internal variability plays an important role.

To demonstrate this at the regional level, we use the Indian monsoon as an example, and show in Fig. 11 the correlation between summer (JJAS) all-India rainfall and the Niño-3.4 index. The histogram shows a range of negative correlation values ($i.e.$, ENSO warming associated with reduced Indian rainfall). An observation-based estimate is also shown, which lies within the ensemble. The use of a large ensemble demonstrates that an assessment of model skill of ENSO teleconnections should not be performed based on a signal ensemble member (a single individual member may appear to demonstrate a poor model, or we may have undue confidence in a model that performs well against observations). We acknowledge there is also observational uncertainty (not tested here) that covers a broad spread across estimates of monsoon-ENSO teleconnection strength (Collins et al., 2013).

Fig. 11 also highlights uncertainties associated with estimating and removing the forced response in observations. Methods such as removing a linear trend from the observations are frequently used, but historical external forcing is non-linear. By definition, the multi-member mean of a large ensemble represents the true forced response in the model (with some error), provided the ensemble is large enough. However, this method is not applicable to observations since we only have one realisation of the real world. The impact of choosing a linear de-trending method, compared to the more accurate estimate of removing the ensemble mean, is shown by comparing the two histograms in Fig. 11. The magenta histogram shows a distribution that is more positively skewed than the histogram in which the data were linearly de-trended, highlighting that de-trending choices have an impact on estimates of the ENSO-monsoon teleconnection.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

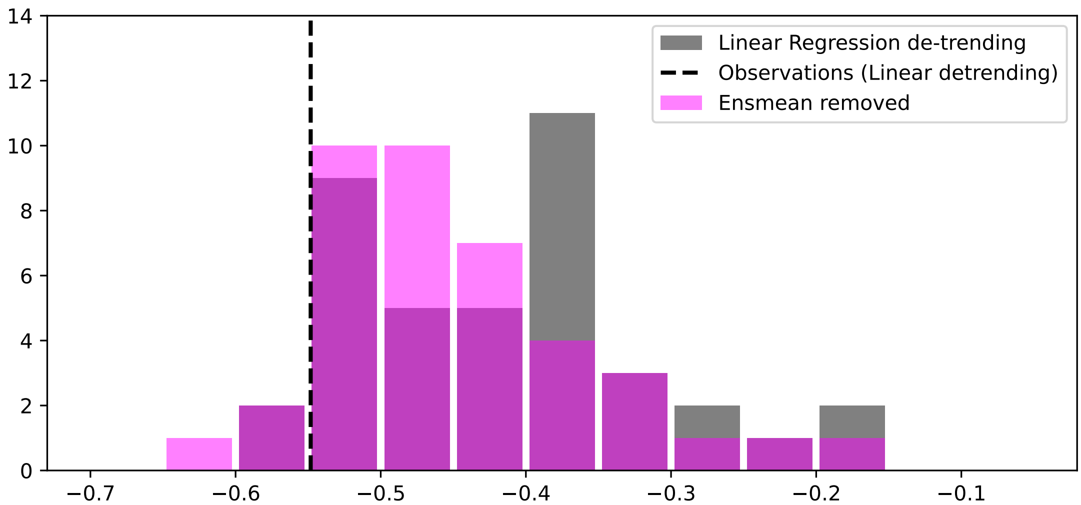

Figure 11. Correlation coefficients between the JJAS Niño-3.4 index and JJAS all-India rainfall for 1950–2014. The histograms show the range of the correlation coefficients obtained across the 40 members of the CANARI LE. Two methods of de-trending were applied to the CANARI LE: removing the multi-member ensemble mean (magenta) and removing a linear trend from each ensemble member (black). The latter serves as the reference for the comparison with the observed correlation coefficient, which was calculated on linearly de-trended data (vertical dashed line). The observed Niño-3.4 index was obtained from NOAA Physical Sciences Laboratory, itself calculated using HadISST1.1 data (Rayner et al., 2003). All-India rainfall in the model is calculated by applying a mask of Indian gridpoints; the observed All-India Rainfall is a rain-gauge product following the method of Parthasarathy et al. (1994).

Finally, the use of a large ensemble also serves as a reminder that observations are but one realisation of internal variability, and may not necessarily lie at the centre of the distribution arising from sampling multiple realisations of internal variability.

4 Application: analogues of North Atlantic 2023 extreme temperature

The summer of 2023 saw extremely high sea surface temperatures (SSTs) over the North Atlantic (Kuhlbrodt et al., 2024). The 2023 June–September (JJAS) SST anomaly from the 1961–1990 mean attained nearly 1.4° C over a large North Atlantic domain (0–80° W, 0–60° N; Fig. 12a) and exceeded that of the second and third warmest summers, 2010 and 2022, by about 0.6° C. We use this event as an example to test if the CANARI LE contains similar unprecedented marine heatwaves.

To this end, we find analogues of the 2023 event by considering both the similarity of the domain-mean SST anomaly, and the similarity of the SST anomaly pattern within the domain. This is done by ranking the domain-mean SST anomaly difference from 2023, then ranking the spatial correlation with the observed 2023 spatial anomaly pattern, and finally averaging the mean-based rank and the correlation-based rank to obtain a combined measure to decide on how ‘close’ a given observed or simulated SST anomaly field is to 2023.

The closest analogues we find in the 74-year instrumental record (1950–2023) are shown in Fig. 12b,d,f,h. We can see that our method works as expected: all four of the closest analogues correspond to anomalously warm summers across the domain,

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

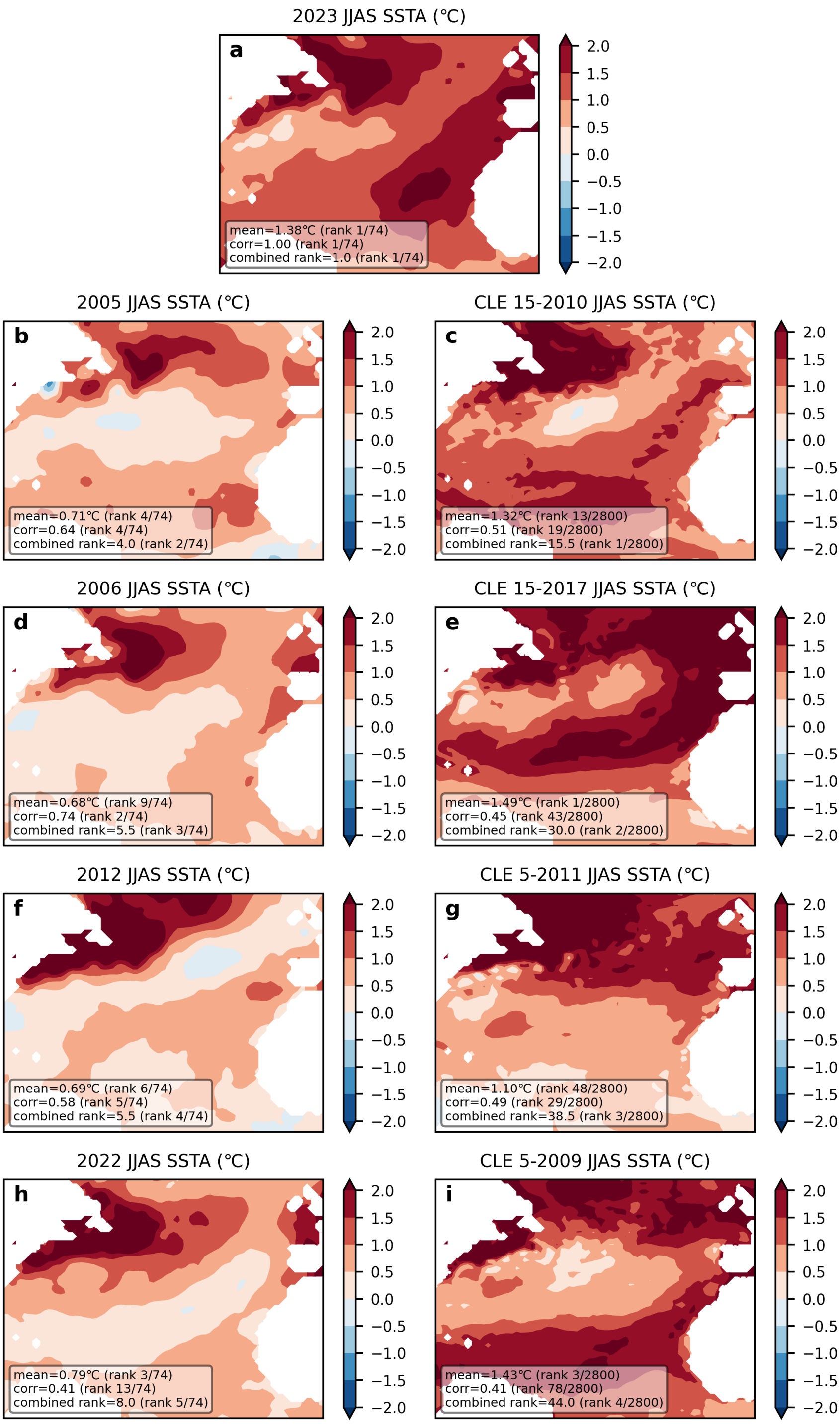

Figure 12. (a) JJAS 2023 SST anomaly in the North Atlantic (HadISST1.1; Rayner et al., 2003). (b,d,f,h) Closest observed analogues of the 2023 event in panel (a), (c,e,g,i) closest simulated analogues in the CANARI LE. See text for details.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

and the analogues also display a broadly similar pattern to the 2023 event with one warming centre to the east of Labrador and a second warm region along the coasts of Europe and Africa and across the south of the domain. As discussed, the main difference between 2023 and the closest observed analogues is the magnitude of the warming. Finding analogues of extreme events in the instrumental record is strongly limited by sampling variability, especially in a non-stationary climate.

We now turn to the CANARI LE and repeat the identification of analogues of the observed 2023 event. We sample the period 1950-2019 and can thus choose the closest analogues from $70 \times 40 = 2800$ simulated summers. While the exact choice of the data period is subjective, 2019 is a reasonable endpoint since the CANARI LE global warming level in 2019 is similar to the observed global warming level in 2023 (Fig. 1). The four closest CANARI LE analogues are shown in Fig. 12c,e,g,i. In contrast to the observed analogues, we find CANARI LE analogues with a similar or greater mean SST anomaly than in the observed event. The greater sample size allows analogues to be found that are much closer to the event, although the correspondence is not perfect: The simulated analogues have an SST pattern that is similar to the observed event but differs especially in the second warm region, which tends to be displaced or partially absent in the simulated analogues. The closest simulated analogues attain pattern correlations of about 0.5, which is smaller than the highest correlations in the observation-based analogues. These limitations are likely due to systematic SST biases in the HadGEM3-GC3.1 model. (These biases are shown in Williams et al., 2018).

We proceed by examining the evolution of the domain-mean SST anomaly and the pattern correlation prior to the 2023 event and its closest simulated analogues. Fig. 13a shows how the SST anomaly evolves prior to the 2023 event and its analogues, respectively. The observed SST anomaly (black line) rises by a large amount over the historical record in 2023. This behaviour is also seen in all but one of the seven closest CANARI LE analogues. The exception is the analogue found in member 5, year 2011, for which the SST anomalies are very high for about a decade prior to the event (orange line). The evolution of the pattern correlation is shown in Fig. 13b. In both the observations and the CANARI LE members, these correlations oscillate around zero and only rise to values greater than the historical envelope in the year of the event itself. There is, perhaps, an indication of slightly greater multi-decadal variability in these correlations in the observations than in the CANARI LE; further investigation of this is beyond the scope of this study. In summary, the North Atlantic 2023 marine heatwave is a so-called record-shattering event (Fischer et al., 2021) at the scales of an entire season and ocean basin. The occurrence of this type of event is captured by the CANARI LE.

5 Summary and conclusions

Over recent years, the generation of climate model large ensembles has facilitated a deeper exploration of internal variability in the climate system, allowing a cleaner separation of variability from forced climate change and a better sampling of extremes. The 40-member CANARI HadGEM3 Large Ensemble (CANARI LE), described here, is a new community addition to the growing set of Single-Model Initial-Condition Large Ensembles (SMILEs), key features being its comparatively high resolution, both in the atmosphere and ocean, and its provision of sub-daily output and of CORDEX-like boundary conditions allowing for detailed process-based studies of high-impact weather systems such as mid-latitude storms and blocks. Key as-

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

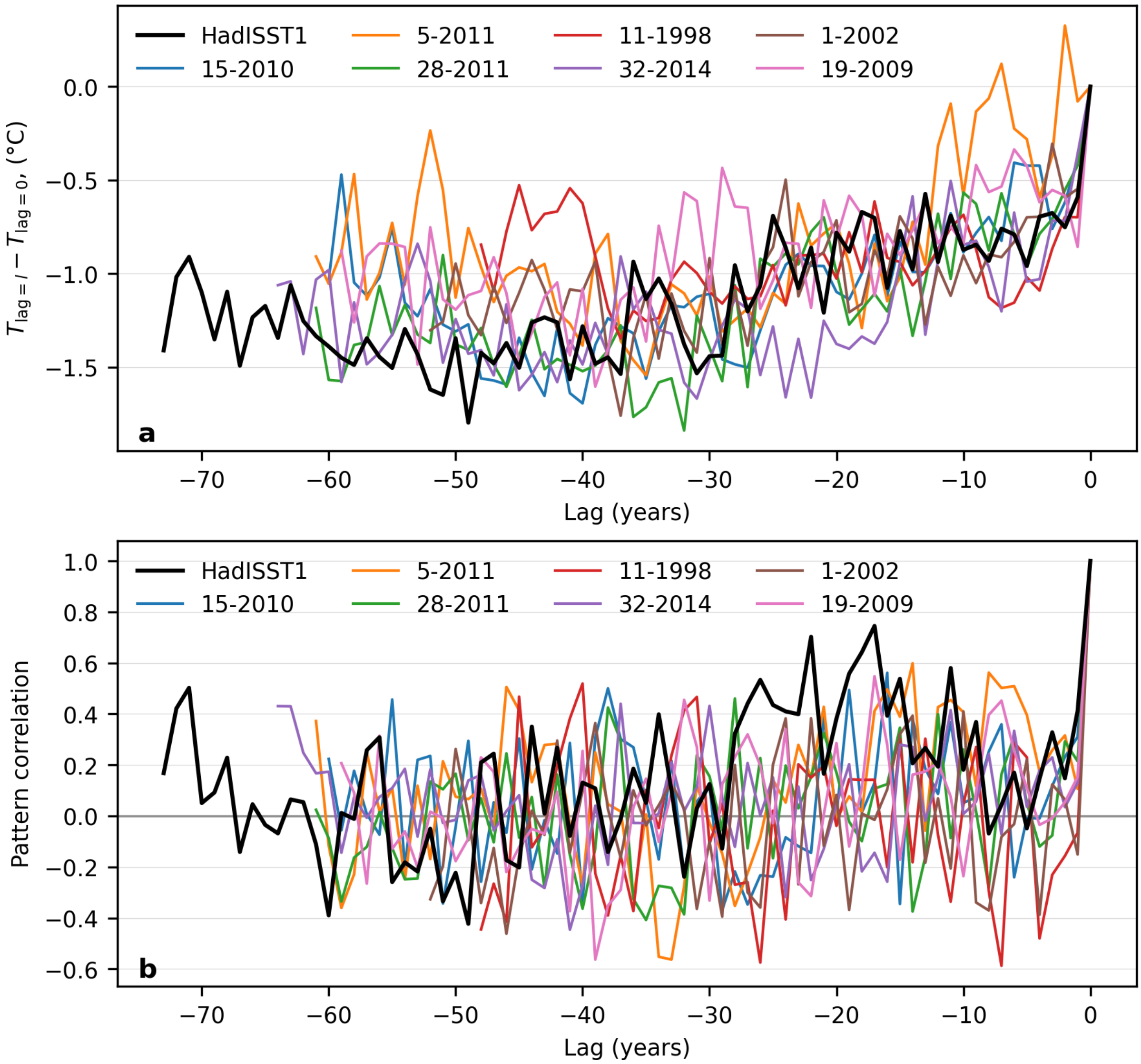

Figure 13. (a) Evolution of domain-mean JJAS SST prior to the 2023 marine heatwave event in the instrumental record (black) and in the closest simulated analogues (colours). (b) Evolution of the pattern correlation with the 2023 event and its analogues, respectively. The seven closest analogues are included, with the additional constraint that they are not from the same CANARI LE ensemble member. The legend identifies the ensemble member and year of these analogues.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

pects of ensemble performance have been presented here, focusing on a comparison of the historical period against reanalyses and observations.

The assessment of mean climate features shows that the CANARI LE reproduces many characteristics of the observed climate system while also exhibiting known issues of the underlying HadGEM3-GC3.1 model (see main text for quantitative results):

– In terms of the global mean surface temperature evolution, the level of agreement between the CANARI LE and an observational estimate is within the range of agreement seen for other LEs. We confirm a faster cooling than observed during 1950–1975, followed by a faster-than-observed warming.

– The representation of the mean North Atlantic eddy driven jet is consistent with the ERA5 reanalysis, both in terms of the strength and tilt of the jet. Likewise, tracking extratropical cyclones reveals good agreement of the simulated and observed North Atlantic storm track location and intensity, albeit somewhat too weak storms in the subpolar-gyre region, as seen in other models.

– The CANARI LE broadly reproduces the zonal-mean stratospheric circulation, capturing the stratospheric polar vortex and the tropical easterlies. Biases include too strong midlatitude westerlies in the upper troposphere/lower stratosphere so that the stratospheric polar vortex and subtropical tropospheric jet appear too tightly dynamically coupled.

– Biases in the Atlantic meridional overturning circulation (AMOC) and subpolar gyre in the CANARI LE are not unusual for climate models: the AMOC is slightly weaker than available observational estimates, and the simulated subpolar gyre is a little too cold and saline.

– The sea-ice extent in the CANARI LE is consistent with observations with respect to mean values, seasonal cycle, and trend. The sea-ice volume is rather high, with the smallest member of the LE close to the observational estimate.

Going further, several assessments in this study highlight the value of the CANARI LE for evaluating trends, variability, and extremes, as well as for interpreting model-observation differences in the presence of internal variability:

– Model-observation discrepancies in the global mean surface temperature evolution cannot be explained by internal variability only. Interannual variability in this variable is well represented.

– Confirming previous results for other models, we show that there is a positive trend in the North Atlantic winter eddy driven jet speed in 1950–2014 in the ERA5 reanalysis that is not seen in any CANARI LE member. We also show that this result is very sensitive to the period chosen to evaluate the trend: the CANARI LE and ERA5 are fully consistent when trends are re-evaluated in the satellite era (1979–2014), pointing to a significant role of multi-decadal variability. Interannual variability in the jet speed is well represented.

– There is a notable similarity in satellite-era trends in extratropical cyclone intensity in the North Atlantic and Arctic between the CANARI LE and ERA5: both show a strengthening over the jet core, and a weakening over the Greenland-Iceland-Norwegian (GIN) seas. While it is possible that the GIN seas weakening is physically-linked to sea-ice retreat in

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

this area, some CANARI LE members show a strengthening so that internal variability could also explain the observed trend.

- The large sample afforded by the CANARI LE enables the robust evaluation of the annual cycle of simulated sudden stratospheric warmings (SSWs). The number of SSWs is consistent with ERA5 from December to April but there are too many simulated SSWs in October and November, which is dynamically consistent with the stratospheric polar vortex mean bias. The CANARI LE captures the downward propagation of the SSW signal to the troposphere. The propagation is weaker than in ERA5, yet this discrepancy could at least in part be due to the small sample of historical SSWs in ERA5.

- The CANARI LE-mean, i.e. forced, evolution of the AMOC shows a peak around 1980, which has been shown in previous work to be consistent with anthropogenic forcing. The CANARI LE also highlights a large range in post-1980 AMOC decline. Contrasting the members with the weakest and strongest AMOC decline shows that the co-evolution of the North Atlantic subpolar gyre and the AMOC is complex, specifically the subpolar gyre declines more quickly in the ensemble member with the weakest AMOC decline than in the member with the strongest AMOC decline.

- Satellite-era sea ice decline also exhibits a remarkable spread across the CANARI LE and is positively correlated with Arctic warming. Members that reproduce the observed decline warm more than the real world, which is consistent with the positive sea-ice thickness bias. The extreme recent decline in the Arctic Siberian Shelf is captured qualitatively by the CANARI LE, yet the magnitude of both the decline and of interannual variability is underestimated.

- The CANARI LE captures the known interannual ENSO teleconnection with tropical rainfall. In addition, the variability in the teleconnection pattern can be examined across the LE, identifying regions where this variability is particularly large such as parts of the Maritime Continent and coastal regions of South and East Asia. The observed correlation between ENSO and Indian Summer Monsoon rainfall is consistent with, but at the negative end of, correlations in the CANARI LE.

- Finally, we have demonstrated the application of the CANARI LE for studying unprecedented heat extremes, using the example of the North Atlantic 2023 marine heatwave. Using an analogue method, we show that the CANARI LE contains events that are similar in scale and magnitude to the observed event; whereas there are no such similar analogues in the observational record. The CANARI LE also includes several so-called record-shattering events that exceed all previous hot anomalies by far.

In summary, the CANARI LE provides a comparatively high-resolution, process-rich, and internally consistent framework for quantifying internal variability, isolating forced climate responses, evaluating model performance, and assessing climate risks. The dataset is intended as a long-term community asset, and future work will exploit the SSP3-7.0 simulations to investigate emerging climate-change signals such changes in mid-latitude storminess, Arctic–mid-latitude linkages, and the drivers of high-impact extremes in a warming world.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

Code and data availability. Analysis code has been archived on Zenodo at https://doi.org/10.5281/zenodo.19824683 (Schiemann et al., 2026a). The priority variables (440 TB) from the CANARI LE are available in the Centre for Environmental Data Analysis (CEDA) archive at https://dx.doi.org/10.5285/4b62fafc68ae4c3784fb2819f652c537 (Schiemann et al., 2026b). All results presented in this paper utilise the priority variables. Additional output is currently stored on CANARI team tape, and can be accessed on request to the author team provided disk space can be made available. These data will be migrated to CEDA archive near-line access in due course.

Author contributions. Developing the CANARI LE was made possible by the collaborative efforts of a team of scientists whose expertise spans physical climate science, computational science, and data science. The individuals listed below have contributed to the project over periods ranging from days to years; the extent of these contributions is broadly reflected in the author order. We use the Creditor Role Taxonomy (CRediT, https://credit.niso.org) to characterise the nature of these contributions. Within each contribution category below, authors are listed alphabetically.

Conceptualization: B. Lawrence, J. Mecking, S. Osprey, J. Robson, R. Schiemann, L. Shaffrey.

Data Curation: R. Hatcher, B. Lawrence, G. Lister, R. Schiemann.

Formal Analysis: S. Ahmadi, Y. Aksenov, J. Aylmer, L. Baker, A. T. Blaker, D. Case, J. Cole, A. Dittus, E. S. Gentile, S. George, B. Harvey, R. Hatcher, K. Hodges, D. Hodson, E. Karlowska, C. Lang, G. Lister, H. Lu, O. Martínez-Alvarado, J. Mecking, N. O'Callaghan, W. Osmolska, A. Osprey, S. Osprey, T. Phillips, J. Robson, S. Rynders, R. Schiemann, D. Schröder, R. Smith, A. G. Turner, S. Wilson, S. Woolnough.

Funding Acquisition: J. Mecking, S. Osprey, J. Robson, R. Schiemann, L. Shaffrey.

Investigation: Y. Aksenov, J. Aylmer, A. T. Blaker, J. Cole, A. Dittus, E. S. Gentile, B. Harvey, R. Hatcher, K. Hodges, D. Hodson, G. Lister, H. Lu, O. Martínez-Alvarado, J. Mecking, N. O'Callaghan, S. Osprey, T. Phillips, J. Robson, S. Rynders, R. Schiemann, D. Schröder, A. G. Turner, S. Wilson, S. Woolnough.

Methodology: Y. Aksenov, J. Aylmer, A. T. Blaker, J. Cole, A. Dittus, E. S. Gentile, S. Hardiman, B. Harvey, R. Hatcher, K. Hodges, D. Hodson, G. Lister, H. Lu, O. Martínez-Alvarado, J. Mecking, N. O'Callaghan, S. Osprey, T. Phillips, J. Robson, J. Rostron, S. Rynders, R. Schiemann, D. Schröder, D. Sexton, A. G. Turner, J. Williams, S. Wilson, S. Woolnough.

Project Administration: B. Lawrence, G. Lister, R. Schiemann, L. Shaffrey.

Resources: M. Andrews, A. Coward, R. Hatcher, B. Lawrence, G. Lister, R. Schiemann, J. Williams.

Software: J. Cole, S. Hardiman, R. Hatcher, G. Lister, A. Osprey, S. Wilson.

Supervision: B. Lawrence, G. Lister, R. Schiemann, L. Shaffrey.

Validation: Y. Aksenov, J. Harle, R. Hatcher, D. Hodson, G. Lister, S. Rynders, R. Schiemann, B. Sinha, C. Wilson.

Visualization: Y. Aksenov, J. Aylmer, A. Dittus, E. S. Gentile, B. Harvey, K. Hodges, D. Hodson, G. Lister, H. Lu, O. Martínez-Alvarado, J. Mecking, N. O'Callaghan, S. Osprey, T. Phillips, J. Robson, S. Rynders, R. Schiemann, D. Schröder, A. G. Turner, S. Woolnough.

Writing – original draft: Y. Aksenov, J. Aylmer, A. Dittus, E. S. Gentile, B. Harvey, K. Hodges, D. Hodson, B. Lawrence, H. Lu, O. Martínez-Alvarado, J. Mecking, N. O'Callaghan, S. Osprey, T. Phillips, J. Robson, S. Rynders, R. Schiemann, D. Schröder, A. G. Turner, S. Woolnough.

Writing – review editing: A. T. Blaker, J. Harle, B. Harvey, H. Lu, O. Martínez-Alvarado, R. Schiemann, D. Schröder, D. Sexton, A. G. Turner, J. Williams, C. Wilson.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

Competing interests. All authors declare that they have no competing interests.

Acknowledgements. This work was funded by the NERC Climate change in the Arctic–North Atlantic region and impacts on the UK (CANARI) programme, grant number NE/W004984/1. A. G. Turner was funded by the National Centre for Atmospheric Science through the NERC National Capability International Programme Award (NE/X006263/1). We acknowledge the invaluable support provided by the teams at the ARCHER2 UK National Supercomputing Service (https://www.archer2.ac.uk) and at the JASMIN data analysis facility (https://jasmin.ac.uk) at the UK Centre for Environmental Data Analysis (CEDA).

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

References

Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., Burke, E., Dittus, A. J., Florek, P., Gray, L. J., Haddad, S., Hardiman, S. C., Hermanson, L., Hodson, D., Hogan, E., Jones, G. S., Knight, J. R., Kuhlbrodt, T., Misios, S., Mizielinski, M. S., Ringer, M. A., Robson, J., and Sutton, R. T.: Historical Simulations With HadGEM3-GC3.1 for CMIP6, Journal of Advances in Modeling Earth Systems, 12, 1–34, https://doi.org/10.1029/2019MS001995, 2020.

Andrews, T., Andrews, M. B., Bodas-Salcedo, A., Jones, G. S., Kuhlbrodt, T., Manners, J., Menary, M. B., Ridley, J., Ringer, M. A., Sellar, A. A., Senior, C. A., and Tang, Y.: Forcings, Feedbacks, and Climate Sensitivity in HadGEM3-GC3.1 and UKESM1, Journal of Advances in Modeling Earth Systems, 11, 4377–4394, https://doi.org/10.1029/2019MS001866, 2019.

Athanasiadis, P. J., Ogawa, F., Omrani, N.-E., Keenlyside, N., Schiemann, R., Baker, A. J., Vidale, P. L., Bellucci, A., Ruggieri, P., Haarsma, R., Roberts, M., Roberts, C., Novak, L., and Gualdi, S.: Mitigating Climate Biases in the Midlatitude North Atlantic by Increasing Model Resolution: SST Gradients and Their Relation to Blocking and the Jet, Journal of Climate, 35, 6985–7006, https://doi.org/10.1175/JCLI-D-21-0515.1, 2022.

Aylmer, J. R., Ferreira, D., and Feltham, D. L.: Impact of ocean heat transport on sea ice captured by a simple energy balance model, Communications Earth & Environment, 5, 406, https://doi.org/10.1038/s43247-024-01565-7, 2024.

Baldwin, M. P. and Dunkerton, T. J.: Stratospheric Harbingers of Anomalous Weather Regimes, Science, 294, 581–584, https://doi.org/10.1126/science.1063315, 2001.

Berx, B. and Payne, M. R.: The Sub-Polar Gyre Index – a community data set for application in fisheries and environment research, Earth System Science Data, 9, 259–266, https://doi.org/10.5194/essd-9-259-2017, publisher: Copernicus GmbH, 2017.

Bett, P. E., Scaife, A. A., Hardiman, S. C., Thornton, H. E., Shen, X., Wang, L., and Pang, B.: Using large ensembles to quantify the impact of sudden stratospheric warmings and their precursors on the North Atlantic Oscillation, Weather and Climate Dynamics, 4, 213–228, https://doi.org/10.5194/wcd-4-213-2023, 2023.

Blackport, R. and Fyfe, J. C.: Climate models fail to capture strengthening wintertime North Atlantic jet and impacts on Europe, Science Advances, 8, eabn3112, 2022.

Bryden, H., King, B., and McCarthy, G.: South Atlantic overturning circulation at 24°S, Journal of Marine Research, 69, https://elischolar.library.yale.edu/journal_of_marine_research/294, 2011.

Bryden, H. L., Longworth, H. R., and Cunningham, S. A.: Slowing of the Atlantic meridional overturning circulation at 25° N, Nature, 438, 655–657, https://doi.org/10.1038/nature04385, publisher: Nature Publishing Group, 2005.

Cavalieri, D., Parkinson, C., Gloersen, P., and Zwally, H. J.: Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I- SSMIS Passive Microwave Data, Boulder, Colorado, USA, NASA DAAC at the National Snow and Ice Data Center, (updated 2014), 1996.

Charlton, A. J. and Polvani, L. M.: A New Look at Stratospheric Sudden Warmings. Part I: Climatology and Modeling Benchmarks, Journal of Climate, 20, 449–469, https://doi.org/10.1175/JCLI3996.1, 2007.

Cleveland, W. S.: LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression, The American Statistician, 35, 54, https://doi.org/10.2307/2683591, 1981.

Collins, M., AchutaRao, K., Ashok, K., Bhandari, S., Mitra, A. K., Prakash, S., Srivastava, R., and Turner, A.: Observational challenges in evaluating climate models, Nature Climate Change, 3, 940–941, https://doi.org/10.1038/nclimate2012, 2013.

Comiso, J.: Bootstrap Sea Ice Concentrations From NIMBUS-7 SMMR and DMSP SSM/I, available at: http://nsidc.org/data/ nsidc-0079.html (last access: December 2014), Natl. Snow and Ice Data Cent., Boulder, CO, (updated 2014), 1999.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

Deser, C., Phillips, A. S., Alexander, M. A., Amaya, D. J., Capotondi, A., Jacox, M. G., and Scott, J. D.: Future Changes in the Intensity and Duration of Marine Heat and Cold Waves: Insights from Coupled Model Initial-Condition Large Ensembles, Journal of Climate, 37, 1877–1902, https://doi.org/10.1175/JCLI-D-23-0278.1, 2024.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fischer, E. M., Sippel, S., and Knutti, R.: Increasing probability of record-shattering climate extremes, Nature Climate Change, 11, 689–695, https://doi.org/10.1038/s41558-021-01092-9, 2021.

Garzoli, S. L., Baringer, M. O., Dong, S., Perez, R. C., and Yao, Q.: South Atlantic meridional fluxes, Deep Sea Research Part I: Oceanographic Research Papers, 71, 21–32, https://doi.org/10.1016/j.dsr.2012.09.003, 2013.

Gray, L. J., Beer, J., Geller, M., Haigh, J. D., Lockwood, M., Matthes, K., et al.: Solar influences on climate, Reviews of Geophysics, 48, RG4001, https://doi.org/10.1029/2009RG000282, 2010.

Gray, L. J., Lu, H., Brown, M. J., Knight, J. R., and Andrews, M. B.: Mechanisms of influence of the Semi-Annual Oscillation on stratospheric sudden warmings, Quarterly Journal of the Royal Meteorological Society, 148, 1223–1241, https://doi.org/10.1002/qj.4256, 2022.

Hausfather, Z. and Peters, G. P.: Emissions - the 'business as usual' story is misleading, Nature, 577, 618–620, 2020.

Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W., and Zelinka, M.: Climate simulations: recognize the 'hot model' problem, Nature, 605, 26–29, https://doi.org/10.1038/d41586-022-01192-2, 2022.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, Bulletin of the American Meteorological Society, 90, 1095–1107, https://doi.org/10.1175/2009BAMS2607.1, 2009.

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., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hirschi, J. J.-M., Blaker, A. T., Sinha, B., Coward, A., de Cuevas, B., Alderson, S., and Madec, G.: Chaotic variability of the meridional overturning circulation on subannual to interannual timescales, Ocean Science, 9, 805–823, https://doi.org/10.5194/os-9-805-2013, publisher: Copernicus GmbH, 2013.

Hirschi, J. J.-M., Barnier, B., Böning, C., Biastoch, A., Blaker, A. T., Coward, A., Danilov, S., Drijfhout, S., Getzlaff, K., Griffies, S. M., Hasumi, H., Hewitt, H., Iovino, D., Kawasaki, T., Kiss, A. E., Koldunov, N., Marzocchi, A., Mecking, J. V., Moat, B., Molines, J.-M., Myers, P. G., Penduff, T., Roberts, M., Treguier, A.-M., Sein, D. V., Sidorenko, D., Small, J., Spence, P., Thompson, L., Weijer, W., and Xu, X.: The Atlantic Meridional Overturning Circulation in High-Resolution Models, Journal of Geophysical Research: Oceans, 125, e2019JC015 522, https://doi.org/10.1029/2019JC015522, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019JC015522, 2020.

Hodges, K.: Feature tracking on the unit-sphere, Monthly Weather Review, 123, 3458–3465, 1995.

Hodges, K. I.: A general method for tracking analysis and its application to meteorological data, Monthly Weather Review, 122, 2573–2586, 1994.

Hoskins, B. J. and Hodges, K. I.: New perspectives on the Northern Hemisphere winter storm tracks, Journal of the Atmospheric Sciences, 59, 1041–1061, 2002.

Häkkinen, S. and Rhines, P. B.: Decline of Subpolar North Atlantic Circulation During the 1990s, Science, 304, 555–559, https://doi.org/10.1126/science.1094917, publisher: American Association for the Advancement of Science, 2004.

Jackson, L. C., Kahana, R., Graham, T., Ringer, M. A., Woollings, T., Mecking, J. V., and Wood, R. A.: Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM, Climate Dynamics, 45, 3299–3316, https://doi.org/10.1007/s00382-015-2540-2, 2015.

Johns, W. E., Elipot, S., Moat, B., Smeed, D. A., King, B., Volkov, D., and Smith, R. H.: Atlantic Meridional Overturning Circulation (AMOC) Heat Transport Time Series between April 2004 and December 2020 at 26.5°N, https://doi.org/10.17604/3nfq-va20, publisher: University of Miami Libraries, 2023.

Kanzow, T., Cunningham, S. A., Johns, W. E., Hirschi, J. J.-M., Marotzke, J., Baringer, M. O., Meinen, C. S., Chidichimo, M. P., Atkinson, C., Beal, L. M., Bryden, H. L., and Collins, J.: Seasonal Variability of the Atlantic Meridional Overturning Circulation at 26.5N, Journal of Climate, 23, 5678 – 5698, https://doi.org/10.1175/2010JCLI3389.1, 2010.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J. F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The community earth system model (CESM) large ensemble project : A community resource for studying climate change in the presence of internal climate variability, Bulletin of the American Meteorological Society, 96, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1, 2015.

Kirchmeier-Young, M. C., Zwiers, F. W., and Gillett, N. P.: Attribution of extreme events in Arctic Sea ice extent, Journal of Climate, 30, 553–571, https://doi.org/10.1175/JCLI-D-16-0412.1, 2017.

Knutti, R. and Sedláček, J.: Robustness and uncertainties in the new CMIP5 climate model projections, Nature Climate Change, 3, 369–373, https://doi.org/10.1038/nclimate1716, 2013.

Koul, V., Tesdal, J.-E., Bersch, M., Hátún, H., Brune, S., Borchert, L., Haak, H., Schrum, C., and Baehr, J.: Unraveling the choice of the north Atlantic subpolar gyre index, Scientific Reports, 10, 1005, https://doi.org/10.1038/s41598-020-57790-5, publisher: Nature Publishing Group, 2020.

Kretschmer, M., Coumou, D., Agel, L., Barlow, M., Tziperman, E., and Cohen, J.: More-persistent weak stratospheric polar vortex states linked to cold extremes, Bulletin of the American Meteorological Society, 99, 49–60, https://doi.org/10.1175/BAMS-D-16-0259.1, 2018.

Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., Hill, R., Graham, T., Ridley, J., Blaker, A., Calvert, D., Copsey, D., Ellis, R., Hewitt, H., Hyder, P., Ineson, S., Mulcahy, J., Siahaan, A., and Walton, J.: The Low-resolution version of HadGEM3 GC3.1: Development and evaluation for global climate, Journal of Advances in Modeling Earth Systems, 10, 2865–2888, https://doi.org/10.1029/2018MS001370, 2018.

Kuhlbrodt, T., Swaminathan, R., Ceppi, P., and Wilder, T.: A Glimpse into the Future: The 2023 Ocean Temperature and Sea Ice Extremes in the Context of Longer-Term Climate Change, Bulletin of the American Meteorological Society, 105, E474–E485, https://doi.org/10.1175/BAMS-D-23-0209.1, 2024.

Lee, S. H., Furtado, J. C., and Charlton-Perez, A. J.: Wintertime North American Weather Regimes and the Arctic Stratospheric Polar Vortex, Geophysical Research Letters, 46, 14 892–14 900, https://doi.org/10.1029/2019GL085592, 2019.

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., Knutti, R., and Hawkins, E.: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6, Earth System Dynamics, 11, 491–508, https://doi.org/10.5194/esd-11-491-2020, 2020.

Lu, H., Gray, L. J., Martineau, P., King, J. C., and Bracegirdle, T. J.: Regime behavior in the upper stratosphere as a precursor of stratosphere-troposphere coupling in the northern winter, Journal of Climate, 34, 1–53, https://doi.org/10.1175/JCLI-D-20-0831.1, 2021.

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, https://doi.org/10.1029/2019MS001639, 2019.

Mahoney, A. R., Barry, R. G., Smolyanitsky, V., and Fetterer, F.: Observed sea ice extent in the Russian Arctic, 1933–2006, Journal of Geophysical Research: Oceans, 113, https://doi.org/10.1029/2008JC004830, 2008.

McDonagh, E. L. and King, B. A.: Oceanic Fluxes in the South Atlantic, Journal of Physical Oceanography, 35, 109–122, https://doi.org/10.1175/JPO-2666.1, publisher: American Meteorological Society Section: Journal of Physical Oceanography, 2005.

McGetrick, S., Lu, H., Muszynski, G., Martínez-Alvarado, O., Osman, M., Mattingly, K., and Galea, D.: A Deep Learning Approach to Detecting Atmospheric Rivers in the Arctic, Atmosphere, 17, 61, https://doi.org/10.3390/atmos17010061, 2026.

Mecking, J. V., Drijfhout, S. S., Jackson, L. C., and Andrews, M. B.: The effect of model bias on Atlantic freshwater transport and implications for AMOC bi-stability, Tellus A: Dynamic Meteorology and Oceanography, 69, https://a.tellusjournals.se/articles/10.1080/16000870.2017.1299910, 2017.

Menary, M. B., Hodson, D. L. R., Robson, J. I., Sutton, R. T., Wood, R. A., and Hunt, J. A.: Exploring the impact of CMIP5 model biases on the simulation of North Atlantic decadal variability, Geophysical Research Letters, 42, 5926–5934, https://doi.org/10.1002/2015GL064360, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL064360, 2015.

Menary, M. B., Kuhlbrodt, T., Ridley, J., Andrews, M. B., Dimdore-Miles, O. B., Deshayes, J., Eade, R., Gray, L., Ineson, S., Mignot, J., Roberts, C. D., Robson, J., Wood, R. A., and Xavier, P.: Preindustrial Control Simulations With HadGEM3-GC3.1 for CMIP6, Journal of Advances in Modeling Earth Systems, 10, 3049–3075, https://doi.org/10.1029/2018MS001495, 2018.

Moat, B. I.: Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS array at 26N from 2004 to 2023 (v2023.1), https://doi.org/10.5285/223b34a3-2dc5-c945-e063-7086abc0f274, archive Location: North Atlantic Ocean,Straits of Florida,Atlantic Ocean,Northeast Atlantic Ocean (40W),Northwest Atlantic Ocean (40W) Publisher: NERC EDS British Oceanographic Data Centre NOC, 2024.

Moreno-Chamarro, E., Zanchettin, D., Lohmann, K., Luterbacher, J., and Jungclaus, J. H.: Winter amplification of the European Little Ice Age cooling by the subpolar gyre, Scientific Reports, 7, 9981, https://doi.org/10.1038/s41598-017-07969-0, publisher: Nature Publishing Group, 2017.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, Journal of Geophysical Research: Atmospheres, 126, 1–28, https://doi.org/10.1029/2019JD032361, 2021.

NOAA Physical Sciences Laboratory: Niño-3.4 SST Index from the HadISST1.1, https://psl.noaa.gov/data/timeseries/month/data/nino34.long.anom.data, accessed: 2025-10-10.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

Parthasarathy, B., Munot, A. A., and Kothawale, D. R.: All-India monthly and seasonal rainfall series: 1871–1993, Theoretical and Applied Climatology, 49, 217–224, https://doi.org/10.1007/BF00867461, 1994.

Priestley, M. D., Ackerley, D., Catto, J. L., and Hodges, K. I.: Drivers of biases in the CMIP6 extratropical storm tracks. Part I: Northern Hemisphere, Journal of Climate, 36, 1451–1467, 2023.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., 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, https://doi.org/10.1029/2002JD002670, 2003.

Ridley, J. K., Blockley, E. W., Keen, A. B., Rae, J. G. L., West, A. E., and Schroeder, D.: The sea ice model component of HadGEM3-GC3.1, Geoscientific Model Development, 11, 713–723, https://doi.org/10.5194/gmd-11-713-2018, 2018.

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., Jackson, L. C., Kuhlbrodt, T., Mathiot, P., Roberts, C. D., Schiemann, R., Seddon, J., Vannière, B., and Vidale, P. L.: Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments, Geoscientific Model Development, 12, 4999–5028, https://doi.org/10.5194/gmd-12-4999-2019, 2019.

Robson, J., Menary, M. B., Sutton, R. T., Mecking, J., Gregory, J. M., Jones, C., Sinha, B., Stevens, D. P., and Wilcox, L. J.: The Role of Anthropogenic Aerosol Forcing in the 1850–1985 Strengthening of the AMOC in CMIP6 Historical Simulations, Journal of Climate, 35, 6843–6863, https://doi.org/10.1175/JCLI-D-22-0124.1, publisher: American Meteorological Society Section: Journal of Climate, 2022.

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., Edwards, J., Kim, J.-E., Simpson, I. R., Stein, K., Stuecker, M. F., Yamaguchi, R., Bódai, T., Chung, E.-S., Huang, L., Kim, W. M., Lamarque, J.-F., Lombardozzi, D. L., Wieder, W. R., and Yeager, S. G.: Ubiquity of human-induced changes in climate variability, Earth System Dynamics, 12, 1393–1411, https://doi.org/10.5194/esd-12-1393-2021, 2021.

Scaife, A. A. and Smith, D.: A signal-to-noise paradox in climate science, npj Climate and Atmospheric Science, 1, 28, https://doi.org/10.1038/s41612-018-0038-4, 2018.

Schiemann, R., Athanasiadis, P., Barriopedro, D., Doblas-Reyes, F., Lohmann, K., Roberts, M. J., Sein, D. V., Roberts, C. D., Terray, L., and Vidale, P. L.: Northern Hemisphere blocking simulation in current climate models: evaluating progress from the Climate Model Intercomparison Project Phase 5 to 6 and sensitivity to resolution, Weather and Climate Dynamics, 1, 277–292, https://doi.org/10.5194/wcd-1-277-2020, 2020.

Schiemann, R., Aylmer, J., Harvey, B., Hodson, D. L. R., Phillips, T., Rynders, S., Dittus, A., and Turner, A. G.: The CANARI HadGEM3 Large Ensemble: Design, and evaluation of historical experiments. Analysis code and data., https://doi.org/10.5281/zenodo.19854507, 2026a.

Schiemann, R., Lister, G., Hatcher, R., Hodson, D., Lawrence, B., Shaffrey, L., Cole, J., Dittus, A., Mecking, J., Robson, J., Wilson, S., Aksenov, Y., Blaker, A., Harvey, B., Martinez-Alvarado, O., Osprey, A., Rynders, S., Ahmadi, S., Aylmer, J., Baker, L., Case, D., Gentile, E., George, S., Hodges, K., Karlowska, E., Lang, C., Lu, H., O'Callaghan, N., Osmolska, W., Osprey, S., Phillips, T., Schröder, D., Smith, R., Turner, A., Woolnough, S., Andrews, M., Coward, A., Hardiman, S., Harle, J., Rostron, J., Sexton, D., Sinha, B., Williams, J., and Wilson, C.: CANARI-LE Priority Variables from the HadGEM3.1 large ensemble of climate model simulations for the CANARI project, https://catalogue.ceda.ac.uk/uuid/4b62fafc68ae4c3784fb2819f652c537, 2026b.

Schweiger, A., Lindsay, R., Zhang, J., Steele, M., Stern, H., and Kwok, R.: Uncertainty in modeled Arctic sea ice volume, Journal of Geophysical Research: Oceans, 116, 1–21, https://doi.org/10.1029/2011JC007084, 2011.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N.,

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

Sexton, D. M. H., Yamazaki, K., Rostron, J. W., Dunstone, N. J., Fereday, D. R., Hardiman, S. C., Ineson, S., and Knight, J. R.: Effect of resolution on simulated teleconnections to winter North Atlantic circulation inferred from a causal network derived from expert elicitation, Climate Dynamics, 63, 32, https://doi.org/10.1007/s00382-024-07497-4, 2025.

Shukla, J. and Paolino, D. A.: The Southern Oscillation and Long-Range Forecasting of the Summer Monsoon Rainfall over India, Monthly Weather Review, 111, 1830–1837, https://doi.org/10.1175/1520-0493(1983)111<1830:TSOALR>2.0.CO;2, 1983.

Simpson, I. R., Rosenbloom, N., Danabasoglu, G., Deser, C., Yeager, S. G., McCluskey, C. S., Yamaguchi, R., Lamarque, J.-F., Tilmes, S., Mills, M. J., and Rodgers, K. B.: The CESM2 Single-Forcing Large Ensemble and Comparison to CESM1: Implications for Experimental Design, Journal of Climate, 36, 5687–5711, https://doi.org/10.1175/JCLI-D-22-0666.1, 2023.

Simpson, I. R., Shaw, T. A., Ceppi, P., Clement, A. C., Fischer, E., Grise, K. M., Pendergrass, A. G., Screen, J. A., Wills, R. C., Woollings, T., Blackport, R., Kang, J. M., and Po-Chedley, S.: Confronting Earth System Model trends with observations, Science Advances, 11, 1–20, https://doi.org/10.1126/sciadv.adt8035, 2025.

Smith, D. M., Gillett, N. P., Simpson, I. R., Athanasiadis, P. J., Baehr, J., Bethke, I., Bilge, T. A., Bonnet, R., Boucher, O., Findell, K. L., Gastineau, G., Gualdi, S., Hermanson, L., Leung, L. R., Mignot, J., Müller, W. A., Osprey, S., Otterå, O. H., Persad, G. G., Scaife, A. A., Schmidt, G. A., Shiogama, H., Sutton, R. T., Swingedouw, D., Yang, S., Zhou, T., and Ziehn, T.: Attribution of multi-annual to decadal changes in the climate system: The Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP), Frontiers in Climate, 4, https://doi.org/10.3389/fclim.2022.955414, 2022.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., Kuhlbrodt, T., Rae, J. G. L., and Sinha, B.: UK Global Ocean GO6 and GO7: a traceable hierarchy of model resolutions, Geoscientific Model Development, 11, 3187–3213, https://doi.org/10.5194/gmd-11-3187-2018, 2018.

Suarez-Gutierrez, L., Milinski, S., and Maher, N.: Exploiting large ensembles for a better yet simpler climate model evaluation, Climate Dynamics, 57, 2557–2580, https://doi.org/10.1007/s00382-021-05821-w, 2021.

Swaminathan, R., Schewe, J., Walton, J., Zimmermann, K., Jones, C., Betts, R. A., Burton, C., Jones, C. D., Mengel, M., Reyer, C. P. O., Turner, A. G., and Weigel, K.: Regional Impacts Poorly Constrained by Climate Sensitivity, Earth's Future, 12, https://doi.org/10.1029/2024EF004901, 2024.

Taylor, K. E., Durack, P. J., Ames, S., and Stockhouse, M.: CMIP6 Guidance for Data Users, https://pcmdi.llnl.gov/CMIP6/Guide/dataUsers.html, 2022.

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, Geoscientific Model Development, 12, 1909–1963, https://doi.org/10.5194/gmd-12-1909-2019, 2019.

Webster, P. J. and Yang, S.: Monsoon and ENSO: Selectively Interactive Systems, Quarterly Journal of the Royal Meteorological Society, 118, 877–926, https://doi.org/10.1002/qj.49711850705, 1992.

https://doi.org/10.5194/egusphere-2026-2281
Preprint. Discussion started: 7 August 2026

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H. T., Hill, R., Hyder, P., Ineson, S., Johns, T. C., Keen, A. B., Lee, R. W., Megann, A., Milton, S. F., Rae, J. G. L., Roberts, M. J., Scaife, A. A., Schiemann, R., Storkey, D., Thorpe, L., Watterson, I. G., Walters, D. N., West, A., Wood, R. A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations, Journal of Advances in Modeling Earth Systems, 10, 357–380, https://doi.org/https://doi.org/10.1002/2017MS001115, 2018.

Wood, R. R., Lehner, F., Pendergrass, A. G., and Schlunegger, S.: Changes in precipitation variability across time scales in multiple global climate model large ensembles, Environmental Research Letters, 16, 084 022, https://doi.org/10.1088/1748-9326/ac10dd, 2021.

Zappa, G., Shaffrey, L. C., and Hodges, K. I.: The Ability of CMIP5 Models to Simulate North Atlantic Extratropical Cyclones*, Journal of Climate, 26, 5379–5396, https://doi.org/10.1175/JCLI-D-12-00501.1, 2013.

Zhang, J. and Rothrock, D. A.: Modeling Global Sea Ice with a Thickness and Enthalpy Distribution Model in Generalized Curvilinear Coordinates, Monthly Weather Review, 131, 845–861, https://doi.org/10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2, 2003.

Zhang, P., Wu, Y., Chen, G., and Yu, Y.: North American cold events following sudden stratospheric warming in the presence of low Barents-Kara Sea sea ice, Environmental Research Letters, 15, 124 017, 2020.