https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development Discussions
Open Access

The CMCC Decadal Prediction System

Dario Nicolì[1], Alessio Bellucci[1,a], Paolo Ruggieri[1,b], Panos Athanasiadis[1], Stefano Materia[1], Daniele Peano[1], Giusy Fedele[1] and Silvio Gualdi[1]

[1]Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici, Bologna, 40127, Italy
[a]now at: Consiglio Nazionale delle Ricerche, Istituto di Scienze dell'Atmosfera e del Clima (CNR-ISAC), Bologna, 40129, Italy
[b]now at: Department of Physics and Astronomy, University of Bologna, Bologna, 40126, Italy.

Correspondence to: Dario Nicolì (dario.nicoli@cmcc.it)

Abstract. Decadal climate predictions, obtained by constraining the initial condition of a dynamical model through a truthful estimate of the observed climate state, provide an accurate assessment of climate change in the near-term range and a useful tool to inform decision-makers on future climate-related risks.

Here we present results from the CMIP6 DCPP-A decadal hindcasts produced with the operational CMCC decadal prediction system (CMCC DPS), based on the fully-coupled CMCC-CM2-SR5 dynamical model. A 15-member suite of 10-year retrospective forecasts, initialized every year from 1960 to 2020, is performed using a full-field initialization strategy.

The predictive skill for key variables is assessed and compared with the skill of an ensemble of non-initialized historical simulations so as to quantify the added value of initialization. In particular, the CMCC DPS is able to skilfully reproduce past-climate surface temperature fluctuations over large parts of the globe. The North Atlantic Ocean is the region that benefits the most from initialization, with the largest skill enhancement occurring over the subpolar region compared to historical simulations. On the other hand, the predictive skill over the Pacific Ocean rapidly decays with forecast time, especially over the North Pacific. In terms of precipitation, the CMCC DPS skill is significantly higher than that of the historical simulations over a few specific regions, including Sahel, Northern Eurasia and over the Western and Central Europe.

The Atlantic Multidecadal Variability is also skilfully predicted, and this likely contributes to the skill found over remote areas through downstream influence, circulation changes and teleconnections. Considering the relatively small ensemble size, a remarkable prediction skill is also found for the North Atlantic Oscillation, with maximum correlations obtained in the 1–9 lead-year range.

Model systematic errors also affect the forecast quality of the CMCC DPS, featuring a prominent cold bias over the Northern Hemisphere, which is not found in the historical runs. This lack of agreement suggests that in some areas the adopted full-value initialization strategy likely perturbs the equilibrium state of the model climate quite significantly.

1 Introduction

Climate fluctuations are the end result of a number of processes, acting on a multitude of timescales. For a long time, century-scale climate change projections, initialized with using a physical state of the climate system obtained from a long simulation of the pre-industrial period and subject to prescribed anthropogenic and natural forcings, have been the only available product to inform decision makers on future climate-related risks. A major limitation of non-initialized climate projections is their lack of information about the ongoing natural variability that may affect climate changes in the near future, which is, at least in part, linked to the current state of the Earth's climate system. Decadal predictions, obtained by constraining the initial condition of a dynamical model (Coupled Global Circulation Model / Earth System Model) through a realistic estimate of the observed climate state, provide a more accurate assessment of climate change in the near-term (decadal) range, where both external and internal drivers contribute to the climate evolution (Smith et al., 2007; Kushnir et al., 2019).

Starting from the 2000s, initialized decadal predictions have been assessed in multiple projects, from the first pioneering efforts up to the 5th Coupled Model Intercomparison Project (CMIP5) (Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009; Meehl et al., 2009; Doblas-Reyes et al., 2011), in which coordinated experiments allowed multi-system comparison to reduce single-model uncertainties (Taylor et al., 2012; Bellucci et al., 2015), contributing to the Intergovernmental Panel on Climate Change fifth assessment report (AR5, Chapter 11) (Kirtman et al., 2013).

Years of coordinated research and development led to an established experiment protocol that has overcome some of the limitations (e.g. limited ensemble size, initialization every 5 years) of the decadal prediction simulations produced in the CMIP5 framework. This protocol is extensively described in the CMIP6 Decadal Climate Prediction Project (DCPP), a coordinated multi-model effort within the World Climate Research Programme (WCRP) which aims to investigate climate predictions, predictability and variability from annual to decadal timescales (Boer et al., 2016). The DCPP is intended to make skilful forecasts and predictions on these timescales using state-of-the-art climate models and statistical approaches. The core of the DCPP is its "component A" that includes a set of retrospective forecasts (hindcasts). This framework has laid the groundwork for a number of single-model (Bethke et al., 2021; Bilbao et al., 2021; Kataoka et al., 2020; Robson et al., 2018; Sospedra-Alfonso et al., 2021; Xin et al., 2019; Yang et al., 2021; Yeager et al., 2018) and multi-model studies (Borchert et al., 2021a, 2021b; Delgado-Torres et al., 2022).

Climate anomalies on annual–to–multi-decadal timescales are determined from both the internal and the externally forced variability (Boer et al., 2016). External contributions derive from solar irradiance variations, volcanic aerosols and anthropogenic activities, including land use, aerosols and greenhouse gas emissions, accounting for the global warming trend. On the other hand, the oceans and their interaction with the atmosphere are the primary source of internal variability within the climate system on decadal time scales. Low-frequency fluctuations of the North Atlantic sea surface temperature (SST), known as the Atlantic Multidecadal Variability (AMV), affect the global climate through local impacts and remote teleconnections (e.g. Sutton and Hodson, 2005; Zhang and Delworth, 2005; Knight et al., 2006; Sun et al., 2015; Nicolì et al., 2020; Ehsan et al., 2020; Ruprich-Roberts et al., 2021). The long-term AMV evolution is well captured in most state-of-the-art forecast systems and represents one of the primary sources of predictability and, possibly, of skill at decadal timescale, generally attributed to the initialization of the Atlantic Meridional Overturning Circulation (AMOC) (Zhang et al., 2019). Other predictability sources arise from the Pacific Ocean. Interestingly, decadal El-Niño Southern Oscillation (ENSO) impacts may be modulated by the Interdecadal Pacific Oscillation (IPO), which features both the ENSO SST region and extends to

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development Discussions

other extratropical areas (Henley et al., 2015). In addition, the initial state of some land-surface characteristics, stratosphere, snow cover and sea-ice may also impact the predictability of the climate system (e.g. Bellucci et al., 2015; Meehl, 2021). The aforementioned initialized components provide additional predictability for the atmospheric circulation and, in particular, for the North Atlantic Oscillation (NAO) affecting boreal winter climate over Europe (Smith et al., 2019; Athanasiadis et al.,2020). In this paper we present the Decadal Prediction System (DPS) developed at the Euro-Mediterranean Center on Climate Change (CMCC) using the CMCC-CM2-SR5 state-of-the-art climate model (Cherchi et al., 2019) and contributing to the CMIP6 DCPP project. In particular, the study aims to assess the skill in predicting the observed anomalies in key meteorological variables, testing the ability of the DPS in simulating main climate variations from annual to the decadal timescale.

The article is structured as follows: Section 2 provides details on the model configuration, the experimental protocol, the evaluation metrics and the data used to check the benefits of the initialization. Section 3 presents results on the predictions' skill for key quantities and their evolution in time using deterministic and probabilistic approaches and assesses the evolution of some relevant model biases. Section 4 focuses on the skill for selected regional climate variability indices. Finally, Section 5 summarizes and discusses the main findings of the study, also drawing some conclusions.

2 Data and Methodology

2.1 Description of the CMCC DPS model

The CMCC decadal prediction system is based on the CMCC-CM2-SR5 coupled model, shortly described below (see Cherchi et al., 2019 for additional details). The atmospheric component is the Community Atmosphere Model version 5 (CAM5) with a regular grid of 0.9°–1.25° and 30 hybrid levels including 17 levels below 200 hPa and extending up to 2 hPa. The finite-volume configuration has been chosen for the dynamical core. The ocean model is the Nucleus for European Modelling of the Ocean version 3.6 (NEMOv3.6), using a tripolar ORCA grid with a horizontal resolution of about 1° (with a varying latitudinal resolution ranging from 1/3° near the Equator up to 1° at high latitudes) and 50 levels in the vertical. The sea-ice component is the Community Ice CodE in its version 4 (CICE4). The DPS configuration of CICE4 uses a single category to characterize the sea-ice thickness, for consistency with the respective reanalysis used for the initialization. Finally, the Community Land Model version 4.5 (CLM4.5) is used for the simulation of the land surface at the same horizontal grid used by the atmospheric component. Finally, the River Transport Model (RTM, Branstetter 2001) routes liquid and ice runoff from the land surface model to the active ocean to simulate a closed hydrological cycle.

A suite of retrospective forecasts (hindcasts) consisting of 15-member ensembles of 10-year long hindcasts, initialized every year from 1960 to 2020 has been completed, following the CMIP6 DCPP-A protocol (Boer et al., 2016). All members are initialized on November 1st, starting from full-field estimates of the observed state of the ocean, sea-ice, land surface and atmosphere. For each start date, two initial conditions for the atmosphere are obtained from the ERA-40 (1960–1978, Uppala et al., 2005) and ERA-Interim (1979–2020, Berrisford et al., 2019) reanalyses, taking the atmospheric states of November 1st and 2nd. The ocean, sea-ice and land surface states are initialized with an ensemble of global data assimilation products (ocean

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

and sea-ice) and analyses constrained with observed fluxes (land surface). Specifically, land surface is initialized using two different analyses obtained from a land-only configuration of the CLM4.5 land model integrated offline with two different atmospheric forcing datasets: CRUNCEP version 7 (Viovy, 2016) and GSWP3 (Kim, 2017). These datasets provide four instantaneous 2-meter air temperature and humidity, 10-meter winds and surface pressure every six hours, and 3-hourly-accumulated radiation and precipitation. Ocean initial states derive from CHOR (for the period 1960–2010, Yang et al. 2016) and CGLORSv7 reanalysis (for the period 2011–2020, Storto and Masina, 2016). An ensemble of 5 ocean initial states is used to initialize the ocean and sea-ice components: 3 initial estimates originate from global ocean reanalysis characterized by different assimilation strategies of SST and in-situ profiles of temperature and salinity in a 0.5° configuration of the NEMO ocean model, while the remaining 2 initial states are derived through linear combinations of the former 3 initial states.
The time-evolving radiative forcings (including solar radiation, greenhouse gasses concentrations, anthropogenic and volcanic aerosols) are prescribed during the historical period (1960–2014) and follow the ssp2-4.5 scenario (O'Neil et al., 2016) from 2015 onwards, in agreement with the CMIP6 DCPP protocol.

2.2 Verification Data

Uninitialized historical simulations covering the 1850–2014 period are used to assess the added value of realistic model initialization in decadal predictions. We use a 10-member ensemble of historical simulations initialized with different states of a multi-century pre-industrial climate simulation. Each member of the historical ensemble is extended until 2030 under the ssp2-4.5 scenario, thus allowing a fair comparison with the decadal forecast ensemble initialized in year 2020.
The predictive skill for both initialized reforecasts and uninitialized projections is assessed against observational products. To verify the skill for SST, we rely on the Met Office Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) dataset version 1.1 (Rayner et al., 2003 while for 2-meter air temperature (T2m) the CRU TS v4.05 dataset (Harris et al., 2021) is used. Precipitation is assessed by means of the GPCC Full Data Monthly Product Version 2020 (Schneider et al., 2020), while for mean sea level pressure the HadSLP2 dataset (Allan and Ansell, 2006) is used.

2.2 Verification metrics

Initializing decadal predictions from estimates of the observed states of the Earth system may generate spurious responses, since the climate model used to produce the simulations, after initialization, tends to drift towards its own "attractor" (mean climate), deviating from the observed climatology, a consequence of the model's systematic error (bias). This issue is particularly pronounced in the prediction systems adopting a full-field initialization strategy, as in the present case. The spurious drift can be removed a posteriori by subtracting a lead-time dependent climatology at each grid point, assuming a constant drift throughout the time record (Goddard et al., 2013; Boer et al., 2016).
To evaluate the skill of the prediction system, both deterministic and probabilistic metrics are used. The anomaly correlation coefficient (ACC) and the mean square skill score (MSSS) are deterministic metrics, measuring the accuracy of the ensemble mean prediction in reproducing the observed variability over the historical 1960–2014 period targeted by the decadal

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

reforecasts. More specifically, the ACC is a dimensionless measure evaluating the phase agreement between predicted and observed anomalies, ranging from -1 to 1 (Wilks, 2011). The MSSS, instead, quantifies the magnitudes between the predicted and observed anomalies (Goddard et al., 2013). This metric evaluates the skill of the ensemble mean prediction with respect to a reference prediction.

Specifically, the MSSS is defined as:

$$\mathrm{MSSS_{HPO}} = 1 - \frac{\mathrm{MSE_{HO}}}{\mathrm{MSE_{PO}}}, \tag{1}$$

where, $\mathrm{MSE_{HO}}$ ($\mathrm{MSE_{PO}}$) is the mean square error evaluated for the initialized (uninitialized) ensemble mean against observations. The MSSS takes a maximum value of one (1.0), while it does not have a lower limit. Positive MSSS values mean more accurate predictions in the initialized runs and one may speculate that the opposite is also true. However, since the MSSS is not symmetric around zero, the same absolute values of positive and negative MSSS do not have the same meaning in terms of variance.

Probabilistic skill scores provide a useful complement to deterministic metrics in assessing the quality of a prediction system. In this study, Relative Operating Characteristic (hereafter ROC) score maps have been assessed for the hindcasts (Kharin and Zwiers 2003, Wilks 2011). Each grid-point in these maps shows the area under the ROC curve, equal to the probability of a certain anomaly to exceed a specific threshold. When the ROC score approaches the perfect forecast (i.e. equal to one), the DPS is able to discriminate the occurrence of predetermined events. On the other hand, no skill emerges when the score is close to 0.5. Here, we have considered three equiprobable categories: upper tercile, lower tercile and between lower and upper terciles (neutral). Note that the ROC score outcome is not dependent on forecast biases (i.e. calibration) (Kharin and Zwiers, 2003).

The DPS ability to reproduce the dominant climate variability patterns is also tested, focussing in particular on the North Atlantic and North Pacific sectors. Decadal variability in the Atlantic region is well described by the Atlantic Multidecadal Variability (AMV), estimated as the detrended anomalies of SSTs area-weighted over the North Atlantic basin, following the definition adopted in Trenberth and Shea (2006). The skill in predicting the NAO index is also tested using the definition in Li and Wang, 2003 (Fig. S1).

We characterize the low-frequency variability in the Pacific basin through the IPO, which is in turn expressed in terms of the IPO tripolar index (TPI), accounting for the difference between the averaged SST anomalies over the equatorial zone and over the extratropical lobes of the IPO (Henley et al., 2015). At shorter timescales, the ENSO prediction is evaluated through the NINO 3.4 index, representing the spatially average SST anomaly over the respective region of the equatorial Pacific.

The statistical significance is assessed with a one-tailed Student's T test (Wilks, 2011), accounting for auto-correlation in the time series (Bretherton et al., 1999). The anomalies of the observations and the historical simulations are computed with respect to their climatologies (reference period 1981–2010, in agreement with WMO recommendation). In the initialized runs, the climatology for each forecast year is computed considering the highest possible number of initialization years. This approach allows to maximize statistical robustness, even if the skill may depend on the targeted verification years.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development Discussions

3 Skill evaluation

3.1 Near-surface air temperature

Skill in predicting the global mean surface temperature (GMST; based on 2-meter air temperature over land and SST over the ocean) is assessed against observed anomalies combining CRU TS4.05 (Harris et al., 2020) over land and HadISST 1.1 for SSTs (Rayner et al., 2003). Figure 1 shows GMST for initialized hindcasts (in red; "Init", hereafter), non-initialized historical simulations (in blue; "NoInit", hereafter) and observations (in black).

At lead-year 1, the initialized ensemble reproduces quite closely the observed GMST anomalies, showing higher correlation (ACC=0.95) compared to NoInit (ACC=0.89), mainly explained by the strong impact of the imposed initial state at the beginning of the forecasts. Looking at the lead-year range 1–5, the Init still resembles the observed variability (ACC=0.96) within a range of 0.05 °C. Even if the NoInit displays relatively high correlation (ACC=0.94), its skill is substantially due to the global warming trend driven by external forcings and, in this case, the anomaly time series does not reproduce the observed interannual variability. The time evolution of the near-surface temperature over the 6–10 lead-year range exhibits comparable correlations for both initialized and historical ensembles (ACC=0.96), indicating that the radiative forcing has a dominant role at longer lead times.

The ensemble spread envelope of predicted GMST (shown in orange) encompasses the observations, especially at lead-year 1 and 1–5, successfully capturing the multi-year variability, including the cooling effect of major volcanic eruptions, such as El Chichón and Pinatubo occurred in 1982 and 1991, respectively. The initialization contributes to the reduction of the Init ensemble spread, which is about half the envelope of the NoInit for lead-year 1. This is not so unexpected since, when a simulation is initialized the observed internal variability is imposed thus reducing the uncertainty related to systematic errors (Doblas-Reyes et al., 2013).

3.2 Deterministic metrics

Predictive skill at the regional scale is assessed through ACC maps. Figure 2 shows ACC for annual surface temperatures evaluated at different lead-year intervals, as well as the corresponding differences with respect to NoInit. The MSSS maps provide further details on the skill improvement determined by initialization (Fig. 4) assessing the consistency between the magnitudes of the predicted and the observed anomalies (see section 2.3 Verification metrics).

At lead-year 1, significant predictive skill is found over most of the globe, reaching the highest values (ACC=0.80) over tropical Indian Ocean, northern and equatorial Africa, north-eastern part of South America, subpolar North Atlantic and western tropical Pacific. Lack of skill, instead, characterizes western subtropical North Atlantic, Eastern Europe, central part of South America and part of the western North Pacific and Southern Ocean. The added value of initialization (Fig. 2b) is particularly prominent over the tropical and the eastern subpolar North Atlantic, as well as over the tropical and the extratropical North Pacific. In addition, Init exhibits higher skills (up to 0.5) over the American continent, central Africa and the Indian subcontinent. The corresponding MSSS pattern (Fig. 4a) clearly indicates that the Init outperforms NoInit in

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development Discussions
Open Access EGU

reproducing the magnitude and the sign of the observed anomalies over approximately the same areas, showing improved ACC in respect to NoInit (Fig. 2b).

In the 1–5 lead-year range, skill is generally higher than for lead-year 1, likely due to the effect of averaging over a longer interval (5 years) and to the emerging warming trend. In contrast, the skill undergoes a clear deterioration over the tropical and northern part of the Pacific Ocean (Fig. 2c). Significant skill is found over the continental areas of North America, Eurasia, Africa and over the Maritime continent. A large fraction of the skill seems to derive from the warming trend that increases predictability, at this lead-year range, over land and over the Indian Ocean (Van Oldenborg et al., 2012). Over the North Atlantic Ocean the emerging AMV footprint is recognizable with high predictive skill associated with the typical horse-shoe pattern emerging from the Init vs NoInit comparison (Fig. 2d). This pattern is also noticeable in the relative MSSS map (Fig. 4c) suggesting improved predictability for the AMV tropical lobe whilst the extratropical lobe may be affected by strong biases as it is characterized by high ACC values and neutral MSSS. Near-term prediction skill is improved especially over the eastern Mediterranean and the Arabian Peninsula (ACC=0.3), reaching high correlation values (ACC=0.90 in Figure 2c) also reflected in the MSSS (Fig. 4c).

Lead-year range 6–10 exhibits a pattern very similar to the one of lead-year 1–5 with some regional changes such as Eastern Europe and Siberian region. Areas with non-statistically significant skill cover part of the eastern Pacific Ocean (Fig. 2e). The ACC decreases over the Siberian region while increases over Eastern Europe by 0.2. The generally higher skill attributable to initialization (Fig. 2f) is substantially consistent with the pattern obtained for the lead-year range 1–5, even if it is not reproduced in the MSSS analysis (Fig. 4e), suggesting that surface temperature variations are not well captured.

Compared to surface temperatures, skill in precipitation is generally lower and less spatially coherent (Collins, 2002; Doblas-Reyes et al., 2013). At lead-year 1 significant skill is found only in limited areas, including the North-Eastern Brazil, South-Western U.S., Southern Africa, eastern Australia, Turkey and the Balkan Peninsula, as reflected also by the MSSS values (Fig. 4b). For the lead-year ranges 1–5 (Fig. 3c) and 6–10 (Fig. 3e), significant ACC values can be attributed to the Northern part of the Eurasian continent, the Sahel and Europe, including the Iberian Peninsula, the British Isles and central Europe. However, comparing Init with NoInit reveals that the skill is largely due to trends in the radiative forcing, with slight improvements associated with initialization (Gaetani and Mohino, 2013; Bellucci et al., 2015).

3.3 Mean bias assessment

The full-value strategy is used to initialize the forecasts, providing best estimates of the observed state to each model component. It does have an important drawback: it generates spurious, transient signals determined by the model's tendency to drift towards its own climatological mean state after being initialized from a realistic state around the observed climatology. Following the recommendation of the International CLIVAR Project Office (ICPO 2011), the mean bias is defined as the lead-time dependent ensemble mean deviation from the observed mean state defined throughout the whole time record (1960–2020). Assessing the mean bias is an important part of evaluating decadal predictions. The time-dependent SST bias for decadal hindcasts (lead-years 1, 2–5 and 10) and the bias in the historical simulations are shown in Figure 5.

The SST bias in the Init is very rapidly established during year 1, followed by a slower adjustment occurring in the following years, since the Init curves of zonal mean bias for lead-years 1, 2–5 and 10 remain relatively close to each other (Fig. 5e). Bias patterns featured by Init and NoInit substantially differ over the Northern Hemisphere, with the former presenting a prominent cold bias, which is not found at all longitudes in NoInit. In the Southern Hemisphere, Init and NoInit are much more similar. This lack of agreement between Init and NoInit suggests that initialization likely perturbs the equilibrium state of the model climate quite significantly. Interestingly, the same kind of departures from the observed state have been found also in several other decadal prediction systems, including some contributing to the CMIP5 decadal prediction ensemble that adopted a full-value initialization strategy (Bellucci et al, 2015). This fact indicates that the time adjustment period following the initialization shock typically exceeds 10 years over the subpolar North Atlantic.

The Init bias for precipitation is comparable to the NoInit time-mean bias (Figure 6). Major departures occur in the tropical Pacific (20ºS–20ºN). Here, precipitation is overly strong in both Init and NoInit, especially south of the Equator, where the spurious occurrence of a southern ITCZ is a common bias in coupled models (Tian and Dong, 2020; Bellucci et al., 2010). The double ITCZ is enhanced by initialization, with a rainfall overestimation at lead-year 1 (Fig. 6e). This may lead to artificial increase of precipitation over the South-Western U.S. and drier conditions over the Mediterranean Sea (Dong et al., 2021), as seems to be the case in both the Init and NoInit. Finally, the Init bias for precipitation shows that no strong drift occurs outside the tropical zone as the bias remains rather stable in lead time.

3.4 ROC score

The Relative Operating Characteristic (ROC) Score is used to assess the probabilistic properties of the ensemble Init (see Sect. 2.3 Verification metrics). In this paper, the ROC score analysis focuses on near-surface temperature, because of its high predictability, and considers the occurrences of tercile categories.

The Init well reproduces below-tercile and above-tercile anomalies (Fig. 7), featuring patterns similar to the ACC ones throughout the lead times (Fig. 2), which might represent an upper boundary of the ROC score (Wilks, 2011). Specifically, ROC scores are close to one over land, Western Pacific and North Atlantic for multi-year predictions (lead years 1–5 and 6–10), confirming anomaly direction within the ensemble spread. Predictions of anomalies falling in the middle tercile category exhibit less skill compared to the lower and upper tercile cases. Nevertheless some skill emerges over Africa, northern and eastern part of South America, North Atlantic and Indian Ocean.

4 Assessing the prediction skill for the main climate indices

Predictability of selected regional climate indices is investigated in this section, since they influence climate variability on global and regional scale through the action of teleconnections. The Atlantic Multidecadal Variability (AMV) represents the dominant climate variability pattern of the multi-decadal SST fluctuations in the North Atlantic basin (Knight et al., 2005; Smith et al., 2012; O'Reilly et al., 2019).

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development Discussions

Figure 8a shows the ACC for the AMV index for different lead-time ranges, thus helping to identify the lead-year range that exhibits the maximum skill (see Sect. 2.3 Verification metrics). The largest ACC values are found for lead-year ranges longer than 4 years (ACC>0.80), in agreement with previous works (Van Oldenborgh et al., 2012; Garcia-Serrano et al., 2012), reaching a peak for the lead-year range 4–10 (ACC=0.91). At this lead-year range, the AMV index in Init reproduces well the observed low-frequency variability of the North Atlantic SST, including the '80s negative phase, the sharp increase during the '90s, the peak occurring in the 2000s and the subsequent decline (Fig. 8b). The AMV spatial pattern is also well captured by the DPS, as depicted in ACC of linearly detrended near-surface temperature (Fig. S2). ACC patterns in the North Atlantic reveal the AMV footprint, with correlations ranging from 0.50 (in the subtropics) to 0.91 (at high latitudes). It is worth noting that large ACC values are also found over regions linked to the AMV through remote teleconnections: the Eastern Mediterranean region (Mariotti and Dell'Aquila, 2012; Bellucci et al., 2017), Arabian Peninsula (Van Oldenborgh et al., 2012; Ehsan et al., 2020), Southern Eurasia (Li et al., 2021) and the Western Tropical Pacific (Kucharski et al., 2016; Sun et al., 2017), suggesting that the skillful AMV prediction has also non-local impacts in regions affected by the AMV teleconnection pattern.

The predictive skill for the NAO is also analyzed focussing on the boreal winter season (see Sect. 2.3 Verification metrics). Significant skill is primarily found for lead-year ranges, longer than 5 years (Fig. 8c), although lead-year 1 (coinciding with the first winter season after initialization, essentially a seasonal forecast) also features significant skill (ACC=0.42). The NAO predictive skill is maximum for the lead-year range 1–9 (ACC=0.58). At this lead-year range, the observed NAO phases are well reproduced in Init (Fig. 8d) and in particular, starting from the satellite era, decadal hindcasts realistically simulate the growing trend of the '80s and the following decline that occurred after 1995. The Init well captures the NAO variability despite the limited ensemble size (ACC=0.80 applying an 8-year running mean to the model index).

In the Pacific sector, the TPI is considered a proxy of decadal variability (Power et al., 1999; Henley et al., 2015) accounting for both the equatorial and extratropical SST fluctuations (see Sect. 2.3 Verification metrics). Figure 9a displays the Init ACC for TPI with significant skill peaking at lead-years 4–10 (ACC=0.56). As one may expect, Init is not able to reproduce most of the variance in the Pacific Ocean due to the lack of skill over that domain, also inferable by the ACC and MSSS maps (Fig. 2 and Fig. 4, respectively). In fact, the TPI evolution for lead years 4–10 shows that the respective predicted and observed anomalies are broadly consistent until the late '80s but significantly diverge afterwards (Fig. 9b).

Decadal variability in the equatorial Pacific still remains widely unpredictable (Doblas-Reyes et al., 2013) with very limited predictability beyond lead-year 2. We quantify the DJF Nino 3.4 index prediction skill in Figure 9c,d. As mentioned, the respective maximum occurs during the first winter season after initialization (ACC=0.95, Fig. 9d), although there is some predictability up to lead-year 3 (Fig. 9c).

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development
Discussions
Open Access

5 Summary and conclusions

In this study we analyzed the predictive capabilities of the CMCC DPS, using a set of 15-member hindcasts, covering the period 1960–2020, performed with the CMCC-CM2-SR5 coupled model and a full-value initialization strategy. These hindcasts contributed to the CMIP6 DCPP-A multi-model ensemble. As the results have shown, the DPS skilfully reproduces the observed variability of surface temperature, with a large fraction of the total skill being explained by trends in the boundary forcing (Van Oldenborg et al., 2012). In particular, significant skill is found for different variables over large parts of the globe (ACC up to 0.80, Fig. 2). The North Atlantic Ocean is the region that benefits the most from initialization, with the largest skill enhancement (compared to NoInit) over the subpolar region. The ROC score index in Init highlights the DPS ability to correctly discriminate the occurrences of below-tercile and above-tercile temperature anomalies throughout different lead-year intervals. Lack of skills covers the whole Pacific Ocean, over which significant ACC values are bound to lead year 1.

The MSSS analysis reveals that beyond the first lead-year, initialization improves the predictions of surface temperature variations in a few areas, including the Eastern Mediterranean, the Arabian Peninsula and the subtropical Atlantic, while significant deterioration (compared to NoInit) is found over the Southern Ocean and some continental areas. The poor response in MSSS is also confirmed by a cold bias covering the Northern hemisphere, with maximum departures from the observed state over the North Pacific, the subtropical Eastern Atlantic and the subpolar gyre in the North Atlantic. The latter, already documented in several decadal CMIP5 systems (e.g. Bellucci et al., 2015) and CMIP6 (e.g. Bilbao et al., 2021), seems to be caused by the slow ocean response to the prescribed initial conditions which alter advective dynamics.

Regarding precipitation, the CMCC DPS shows limited skill, with statistically significant correlations only over specific areas, a feature shared with other state-of-the-art decadal prediction systems. Indeed, significant skill is only found over Sahelian Africa, Northern Eurasia and over the Western and Central Europe. This spatially confined skill may derive from the AMV, which is known to be a source of predictability for these regions, influencing rainfall variability on annual-to-decadal timescales (Doblas-Reyes et al., 2013; Ehsan et al., 2020; Ruggieri et al., 2020). On the other hand, no significant skill for precipitation is found over the rest of the globe, probably also due to the relatively small size of our ensemble (Yeager et al., 2018), to the very high spatial variability (Goddard et al., 2013) and absence of a strong trend in precipitation (Gaetani and Mohino 2013; Bellucci et al., 2015). Improvements compared to the NoInit are bound to some regional features, suggesting no substantial benefits from initialization in terms of precipitation skill.

Certain climate variability patterns in the North Atlantic sector feature significant predictability. In particular, the observed AMV signal is skillfully predicted, and this likely contributes to obtaining significant skills also in remote areas through downstream influence, circulation changes and teleconnections. A remarkable prediction skill is also found for the NAO index, with maximum correlations obtained in the 1–9 lead-year range (ACC=0.58). This promising result, which has been demonstrated also for other decadal prediction systems (Athanasiadis et al., 2020; Smith et al. 2020) means that the decadal NAO index can be used as a statistical predictor to improve forecasts of poorly simulated variables influenced by the observed NAO (Dunstone et al., 2022). Over the tropical Pacific Ocean, ENSO variability exhibits some predictability up to year 3, with highest values of ACC bound to the first winter after initialization (ACC=0.95). Moreover, the TPI shows higher predictive skill on longer timescales despite the scarce skill that the DPS exhibits over most of the Pacific Ocean.

Model systematic errors also affect the forecast quality. The strong cold bias occurring over the Northern Hemisphere could be related to the initialization shock and the subsequent drift, typical of the full-value initialization (He et al., 2017) approach used in the CMCC DPS. Admittedly, AMOC is particularly affected by the initialization strategy, with the full-value approach inducing a long-term adjustment due to the bias in the representation of the large-scale ocean circulation (Polkova et al., 2014). In this context, another sensitive region is the Equatorial Pacific in which full-value initialization seems to give the highest benefit in skill compared to anomalous initialization (Bellucci et al., 2015). From another perspective, model errors may be mitigated by enhancing spatial resolution (both horizontal and vertical) in the oceanic and atmospheric model components, since course resolution limits a realistic representation of key physical processes (e.g. realistic SST front in the Gulf-Stream region), impacting the atmospheric circulation downstream (Athanasiadis et al., 2022, Paolini et al., 2022). For instance, an eddy-permitting ocean model (i.e. 0.25° horizontal resolution) in a fully-coupled system led to improved decadal predictions over the whole equatorial zone (Shaffrey et al., 2017). Moreover, model uncertainty may be also reduced by increasing the ensemble size, allowing the predictable signal to emerge from the noisy background (Athanasiadis et al., 2020).

Interestingly, the results obtained from the CMCC DPS are broadly consistent with similar assessments from other CMIP6 decadal prediction systems (Bethke et al., 2021; Bilbao et al., 2021; Kataoka et al., 2020; Robson et al., 2018; Sospedra-Alfonso et al., 2021; Xin et al., 2019; Yang et al., 2021; Yeager et al., 2018) and multi-model studies (Borchert et al., 2021a, 2021b; Delgado-Torres et al., 2022). In particular, all the DPSs feature high predictive skills over the Atlantic Ocean, the Indian Ocean and continental areas, where a large fraction of predictability and of the skill are provided by external forcings. The added value of initialization is most noted over the subpolar gyre and the subtropical Atlantic (in most of the DPSs), confirming the North Atlantic as the region which benefits more from decadal predictions. In fact, skillful prediction of long-term SST fluctuations over the subpolar gyre is substantially driven by the AMOC initialization and the associated ocean heat transport (Yeager et al., 2018). The non-significant skill found over the Southern Ocean could be, at least in part, due to the lack of oceanic observations covering that region, limiting our capability to create accurate oceanic initial conditions, and increasing the uncertainty related to the skill assessment over this area. The predictive skill of the Pacific Ocean rapidly decays with forecast lead-time, especially over the North Pacific, because of the inability of the DPS to reproduce the observed variability.

Finally, we remind that the CMCC DPS operatively provides decadal forecasts from 2020 and contributes to the annual release of the WMO (World Meteorological Organization) Global Annual-to-Decadal Climate Update, a multi-model assessment of next five-year climate for societal applications (Hermanson et al., 2022). This coordinated effort represents an important step forward to issue reliable predictions and stresses the need to further improve future decadal systems for better understanding near-term climate.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Code and Data availability

The code relative to the CMCC-CM2-SR5 climate model used for DPS is available on the Zenodo repository (Cherchi et al. 2019, https://doi.org/10.5281/zenodo.6810749). The Init and NoInit data are available on ESGF data portal (https://esgf-node.llnl.gov/projects/cmip6/). HadISST was downloaded from https://www.metoffice.gov.uk/hadobs/hadisst/ (last access: 30 May 2022, Rayner et al., 2003). CRU TS4.05 was downloaded from https://catalogue.ceda.ac.uk/uuid/c26a65020a5e4b80b20018f148556681 (last access: 30 May 2022, Harris et al. 2021). GPCC Full Data Monthly Product version 2020 at 1° for horizontal resolution was downloaded from https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-monthly_v2020_doi_download.html (last access: 30 May 2022, Schneider et al., 2020). HadSLP2 was downloaded from https://www.metoffice.gov.uk/hadobs/hadslp2/ (last access: 30 May 2022, Allan and Ansell, 2006). ERA40 was downloaded from https://apps.ecmwf.int/datasets/data/era40-moda/levtype=sfc/ (last access: 30 May 2022, Uppala et al., 2005). CRUNCEP version 7 was downloaded from https://rda.ucar.edu/datasets/ds314.3/#!description (last access: 30 May 2022, Viovy 2016). GSWP3 was downloaded from https://svn-ccsm-inputdata.cgd.ucar.edu/trunk/inputdata/atm/datm7/atm_forcing.datm7.GSWP3.0.5d.v1.c170516/ (last access: 30 May 2022, Kim 2017). Oceanic and sea-ice initial conditions from NEMO CHOR (Yang et al. 2016) and CGLORSv7 (Storto and Masina, 2016) analysis are available on Zenodo repository (doi: https://doi.org/10.5281/zenodo.6866295).

Author contributions

AB, PR and DN conceived the study, setup the model and led the analysis. PR prepared the initial conditions. DN performed the decadal predictions. GF contributes to produce the decadal predictions. DP and SM performed the land analysis for the land initial conditions. All the authors contributed to the discussion and interpretation of the results. DN prepared the manuscript with contributions from all co-authors.

References

Allan, R., & Ansell, T. (2006). A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. Journal of Climate, 19(22), 5816-5842.

Athanasiadis, P. J., Yeager, S., Kwon, Y. O., Bellucci, A., Smith, D. W., & Tibaldi, S. (2020). Decadal predictability of North Atlantic blocking and the NAO. NPJ Climate and Atmospheric Science, 3(1), 1-10.

Bellucci, A., Gualdi, S., & Navarra, A. J. J. O. C. (2010). The double-ITCZ syndrome in coupled general circulation models: The role of large-scale vertical circulation regimes. Journal of Climate, 23(5), 1127-1145.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Bellucci, A., Haarsma, R., Gualdi, S., Athanasiadis, P. J., Caian, M., Cassou, C., ... & Yang, S. (2015). An assessment of a multi-model ensemble of decadal climate predictions. Climate Dynamics, 44(9), 2787-2806

Bellucci, A., Mariotti, A., & Gualdi, S. (2017). The role of forcings in the twentieth-century North Atlantic multidecadal variability: The 1940–75 North Atlantic cooling case study. Journal of Climate, 30(18), 7317-7337.

Berrisford, P., Kallberg, P., Kobayashi, S., Dee, D., Uppala, S., Simmons, A. J., ... & Sato, H. (2019). Coauthors, 2011: The ERA-interim archive version 2.0. ERA Rep. Series, 1, 23.

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., ... & Eldevik, T. (2021). NorCPM1 and its contribution to CMIP6 DCPP. Geoscientific Model Development, 14(11), 7073-7116

Bilbao, R., Wild, S., Ortega, P., Acosta-Navarro, J., Arsouze, T., Bretonnière, P. A., ... & Vegas-Regidor, J. (2021). Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth. Earth System Dynamics, 12(1), 173-196.

Borchert, L. F., Menary, M. B., Swingedouw, D., Sgubin, G., Hermanson, L., & Mignot, J. (2021a). Improved decadal predictions of North Atlantic subpolar gyre SST in CMIP6. Geophysical Research Letters, 48(3), e2020GL091307.

Borchert, L. F., Koul, V., Menary, M. B., Befort, D. J., Swingedouw, D., Sgubin, G., & Mignot, J. (2021b). Skillful decadal prediction of unforced southern European summer temperature variations. Environmental Research Letters, 16(10), 104017

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., ... & Eade, R. (2016). The decadal climate prediction project (DCPP) contribution to CMIP6. Geoscientific Model Development, 9(10), 3751-3777.

Branstetter, M. L. (2001). Development of a parallel river transport algorithm and applications to climate studies. The University of Texas at Austin.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., & Bladé, I. (1999). The effective number of spatial degrees of freedom of a time-varying field. Journal of climate, 12(7), 1990-2009.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., et al.,(2019). Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. Journal of Advances in Modeling Earth Systems, 11, 185–209. https://doi.org/10.1029/2018MS001369

Collins, M. (2002). Climate predictability on interannual to decadal time scales: The initial value problem. Climate dynamics, 19(8), 671-692.

Delgado-Torres, C., Donat, M. G., Gonzalez-Reviriego, N., Caron, L. P., Athanasiadis, P. J., Bretonnière, P. A., ... & Doblas-Reyes, F. J. (2022). Multi-model forecast quality assessment of CMIP6 decadal predictions. Journal of Climate, 1-46.

Doblas-Reyes, F. J., Balmaseda, M. A., Weisheimer, A., & Palmer, T. N. (2011). Decadal climate prediction with the European Centre for Medium-Range Weather Forecasts coupled forecast system: Impact of ocean observations. Journal of Geophysical Research: Atmospheres, 116(D19).

Doblas-Reyes, F. J., Andreu-Burillo, I., Chikamoto, Y., García-Serrano, J., Guemas, V., Kimoto, M., ... & Van Oldenborgh, G. J. (2013). Initialized near-term regional climate change prediction. Nature communications, 4(1), 1-9.

Dong, L., Leung, L. R., Lu, J., & Song, F. (2021). Double-ITCZ as an emergent constraint for future precipitation over Mediterranean climate regions in the North Hemisphere. Geophysical Research Letters, 48(3), e2020GL091569.

Dunstone, N., Lockwood, J., Solaraju-Murali, B., Reinhardt, K., Tsartsali, E. E., Athanasiadis, P. J., ... & Thornton, H. E. (2022). Towards useful decadal climate services. Bulletin of the American Meteorological Society.

Ehsan, M. A., Nicoli, D., Kucharski, F., Almazroui, M., Tippett, M. K., Bellucci, A., ... & Kang, I. S. (2020). Atlantic Ocean influence on Middle East summer surface air temperature. NPJ Climate and Atmospheric Science, 3(1), 1-8.

Gaetani, M., & Mohino, E. (2013). Decadal prediction of the Sahelian precipitation in CMIP5 simulations. Journal of Climate, 26(19), 7708-7719.

García-Serrano, J., & Doblas-Reyes, F. J. (2012). On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. Climate dynamics, 39(7), 2025-2040.

Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., ... & Delworth, T. (2013). A verification framework for interannual-to-decadal predictions experiments. Climate Dynamics, 40(1-2), 245-272.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Harris, I.C.; Jones, P.D.; Osborn, T. (2021): CRU TS4.05: Climatic Research Unit (CRU) Time-Series (TS) version 4.05 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2020). NERC EDS Centre for Environmental Data Analysis. University of East Anglia Climatic Research Unit;

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scientific data, 7(1), 1-18.

He, Y., Wang, B., Liu, M., Liu, L., Yu, Y., Liu, J., ... & Li, F. (2017). Reduction of initial shock in decadal predictions using a new initialization strategy. Geophysical Research Letters, 44(16), 8538-8547.

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., & Folland, C. K. (2015). A tripole index for the interdecadal Pacific oscillation. Climate Dynamics, 45(11), 3077-3090.

Hermanson, L., Bilbao, R., Dunstone, N., Ménégoz, M., Ortega, P., Pohlmann, H., ... & Danabasoglu, G. (2020). Robust multiyear climate impacts of volcanic eruptions in decadal prediction systems. Journal of Geophysical Research: Atmospheres, 125(9), e2019JD031739.

Hermanson, L., Smith, D., Seabrook, M., Bilbao, R., Doblas-Reyes, F., Tourigny, E., ... & Kumar, A. (2022). WMO global annual to decadal climate update: a prediction for 2021–25. Bulletin of the American Meteorological Society, 103(4), E1117-E1129.

ICPO (International CLIVAR Project Office) (2011) Data and bias correction for decadal climate predictions. International CLIVAR Project Office, CLIVAR Publication Series No. 150, 6 pp.

Kataoka, T., Tatebe, H., Koyama, H., Mochizuki, T., Ogochi, K., Naoe, H., ... & Watanabe, M. (2020). Seasonal to decadal predictions with MIROC6: Description and basic evaluation. Journal of Advances in Modeling Earth Systems, 12(12), e2019MS002035.

Keenlyside, N. S., Latif, M., Jungclaus, J., Kornblueh, L., & Roeckner, E. (2008). Advancing decadal-scale climate prediction in the North Atlantic sector. Nature, 453(7191), 84-88.

Kharin, V. V., & Zwiers, F. W. (2003). On the ROC score of probability forecasts. Journal of Climate, 16(24), 4145-4150.

Geoscientific Model Development Discussions Open Access EGU

Kim, H. J. Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1). Data Integr. Anal. Syst. (2017).

Kirtman, B., Power, S. B., Adedoyin, A. J., Boer, G. J., Bojariu, R., Camilloni, I., ... & Wang, H. J. (2013). Near-term climate change: projections and predictability. Climate Change 2013: The Physical Science Basis. IPCC Working Group I Contribution to AR5., 953-1028.

Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., & Mann, M. E. (2005). A signature of persistent natural thermohaline circulation cycles in observed climate. Geophysical Research Letters, 32(20).

Knight, J. R., Folland, C. K., & Scaife, A. A. (2006). Climate impacts of the Atlantic multidecadal oscillation. Geophysical Research Letters, 33(17).

Kucharski, F., Parvin, A., Rodriguez-Fonseca, B., Farneti, R., Martin-Rey, M., Polo, I., ... & Mechoso, C. R. (2016). The teleconnection of the tropical Atlantic to Indo-Pacific sea surface temperatures on inter-annual to centennial time scales: a review of recent findings. Atmosphere, 7(2), 29.

Kushnir, Y., Scaife, A. A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., ... & Wu, B. (2019). Towards operational predictions of the near-term climate. Nature Climate Change, 9(2), 94-101.

Li, G., & Xie, S. P. (2014). Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems. Journal of Climate, 27(4), 1765-1780.

Li J., & Wang, J. X. (2003). A new North Atlantic Oscillation index and its variability. Advances in Atmospheric Sciences, 20(5), 661-676.

Li, J., Li, F., He, S., Wang, H., & Orsolini, Y. J. (2021). The Atlantic multidecadal variability phase dependence of teleconnection between the North Atlantic Oscillation in February and the Tibetan Plateau in March. Journal of Climate, 34(11), 4227-4242.

Mariotti, A., & Dell'Aquila, A. (2012). Decadal climate variability in the Mediterranean region: roles of large-scale forcings and regional processes. Climate Dynamics, 38(5), 1129-1145.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., ... & Stockdale, T. (2009). Decadal prediction: can it be skillful?. Bulletin of the American Meteorological Society, 90(10), 1467-1486.

Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblas-Reyes, F., ... & Xie, S. P. (2021). Initialized Earth System prediction from subseasonal to decadal timescales. Nature Reviews Earth & Environment, 2(5), 340-357.

Nicolì, D., Bellucci, A., Iovino, D., Ruggieri, P., & Gualdi, S. (2020). The impact of the AMV on Eurasian summer hydrological cycle. Scientific reports, 10(1), 1-11.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... & Sanderson, B. M. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. Geoscientific Model Development, 9(9), 3461-3482.

O'Reilly, C. H., Zanna, L., & Woollings, T. (2019). Assessing external and internal sources of Atlantic multidecadal variability using models, proxy data, and early instrumental indices. Journal of Climate, 32(22), 7727-7745.

Paolini, L. F., Athanasiadis, P. J., Ruggieri, P., & Bellucci, A. (2022). The atmospheric response to meridional shifts of the Gulf Stream SST front and its dependence on model resolution. Journal of Climate, 1-57.

Pohlmann, H., Jungclaus, J. H., Köhl, A., Stammer, D., & Marotzke, J. (2009). Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic. Journal of Climate, 22(14), 3926-3938.

Polkova, I., Köhl, A., & Stammer, D. (2014). Impact of initialization procedures on the predictive skill of a coupled ocean–atmosphere model. Climate dynamics, 42(11), 3151-3169.

Power, S., Casey, T., Folland, C., Colman, A., & Mehta, V. (1999). Inter-decadal modulation of the impact of ENSO on Australia. Climate dynamics, 15(5), 319-324.

Rayner, N., Parker, D. E., Folland, C. K., Horton, E. B., Alexander, L. V., & Rowell, D. P. (2003). The global sea-ice and sea surface temperature (HadISST) data sets. J. Geophys. Res, 108, 2-22.

Robson, J. I., Sutton, R. T., & Smith, D. M. (2012). Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s. Geophysical Research Letters, 39(19).

Robson, J., Polo, I., Hodson, D. L., Stevens, D. P., & Shaffrey, L. C. (2018). Decadal prediction of the North Atlantic subpolar gyre in the HiGEM high-resolution climate model. Climate dynamics, 50(3), 921-937.

Ruprich-Robert, Y., Moreno-Chamarro, E., Levine, X., Bellucci, A., Cassou, C., Castruccio, F., ... & Tourigny, E. (2021). Impacts of Atlantic multidecadal variability on the tropical Pacific: a multi-model study. npj climate and atmospheric science, 4(1), 1-11.

Ruggieri, P., Bellucci, A., Nicolí, D., Athanasiadis, P. J., Gualdi, S., Cassou, C., ... & Zampieri, M. (2021). Atlantic multidecadal variability and North Atlantic jet: a multimodel view from the decadal climate prediction project. Journal of Climate, 34(1), 347-360.

Schneider, U., Becker, A., Finger, P., Rustemeier, E., & Ziese, M. (2020). GPCC Full Data Monthly Product Version 2020 at 1.0°: Monthly land-surface precipitation from rain-gauges built on GTS-based and historical data. Deutscher Wetterdienst.

Shaffrey, L. C., Hodson, D., Robson, J., Stevens, D. P., Hawkins, E., Polo, I., ... & Stephens, A. (2017). Decadal predictions with the HiGEM high resolution global coupled climate model: description and basic evaluation. Climate Dynamics, 48(1), 297-311.

Simpson, I. R., Yeager, S. G., McKinnon, K. A., & Deser, C. (2019). Decadal predictability of late winter precipitation in western Europe through an ocean–jet stream connection. Nature Geoscience, 12(8), 613-619.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., & Murphy, J. M. (2007). Improved surface temperature prediction for the coming decade from a global climate model. science, 317(5839), 796-799.

Smith, D. M., Scaife, A. A., & Kirtman, B. P. (2012). What is the current state of scientific knowledge with regard to seasonal and decadal forecasting?. Environmental Research Letters, 7(1), 015602.

Smith, D. M., Eade, R., Scaife, A. A., Caron, L. P., Danabasoglu, G., DelSole, T. M., ... & Yang, X. (2019). Robust skill of decadal climate predictions. Npj Climate and Atmospheric Science, 2(1), 1-10.

Smith, D. M., Scaife, A. A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I., ... & Zhang, L. (2020). North Atlantic climate far more predictable than models imply. Nature, 583(7818), 796-800.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Sospedra-Alfonso, R., Merryfield, W. J., Boer, G. J., Kharin, V. V., Lee, W. S., Seiler, C., & Christian, J. R. (2021). Decadal climate predictions with the Canadian Earth System Model version 5 (CanESM5). Geoscientific Model Development, 14(11), 6863-6891.

Storto, A., & Masina, S. (2016). C-GLORSv5: an improved multipurpose global ocean eddy-permitting physical reanalysis, Earth Syst. Sci. Data, 8, 679–696.

Sun, C., Li, J., & Zhao, S. (2015). Remote influence of Atlantic multidecadal variability on Siberian warm season precipitation. Scientific Reports, 5(1), 1-9.

Sun, C., Kucharski, F., Li, J., Jin, F. F., Kang, I. S., & Ding, R. (2017). Western tropical Pacific multidecadal variability forced by the Atlantic multidecadal oscillation. Nature Communications, 8(1), 1-10.

Sutton, R. T., & Hodson, D. L. (2005). Atlantic Ocean forcing of North American and European summer climate. science, 309(5731), 115-118.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. Bulletin of the American meteorological Society, 93(4), 485-498.

Tian, B., & Dong, X. (2020). The double-ITCZ bias in CMIP3, CMIP5, and CMIP6 models based on annual mean precipitation. Geophysical Research Letters, 47(8), e2020GL087232.

Trenberth, K. E., & Shea, D. J. (2006). Atlantic hurricanes and natural variability in 2005. Geophysical research letters, 33(12).

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., ... & Woollen, J. (2005). The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 131(612), 2961-3012.

Van Oldenborgh, G. J., Doblas-Reyes, F. J., Wouters, B., & Hazeleger, W. (2012). Decadal prediction skill in a multi-model ensemble. Climate dynamics, 38(7), 1263-1280.

Viovy, N. (2018), CRUNCEP Version 7 - Atmospheric Forcing Data for the Community Land Model, http://rda.ucar.edu/datasets/ds314.3/, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, Colo., Accessed, 15 July 2016.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Wang, C., Deser, C., Yu, J. Y., DiNezio, P., & Clement, A. (2017). El Niño and southern oscillation (ENSO): a review. Coral reefs of the eastern tropical Pacific, 85-106.

Wilks, D. S. (2011). Statistical methods in the atmospheric sciences (Vol. 100). Academic press.

Yang, C., S. Masina, and A. Storto (2016), Historical ocean reanalyses (1900–2010) using different data assimilation strategies, Q. J. of Royal Met. Soc., https://doi.org/10.1002/qj.2936.

Xin, X., Wei, M., Li, Q., Zhou, W., Luo, Y., & Zhao, Z. (2019). Decadal prediction skill of BCC-CSM1. 1 with different initialization strategies. Journal of the Meteorological Society of Japan. Ser. II.

Yang, C., Masina, S., Bellucci, A., & Storto, A. (2016). The rapid warming of the North Atlantic Ocean in the Mid-1990s in an eddy-permitting ocean reanalysis (1982–2013). Journal of Climate, 29(15), 5417-5430.

Yang, X., Delworth, T. L., Zeng, F., Zhang, L., Cooke, W. F., Harrison, M. J., ... & McColl, C. (2021). On the Development of GFDL's Decadal Prediction System: Initialization Approaches and Retrospective Forecast Assessment. Journal of Advances in Modeling Earth Systems, 13(11), e2021MS002529.

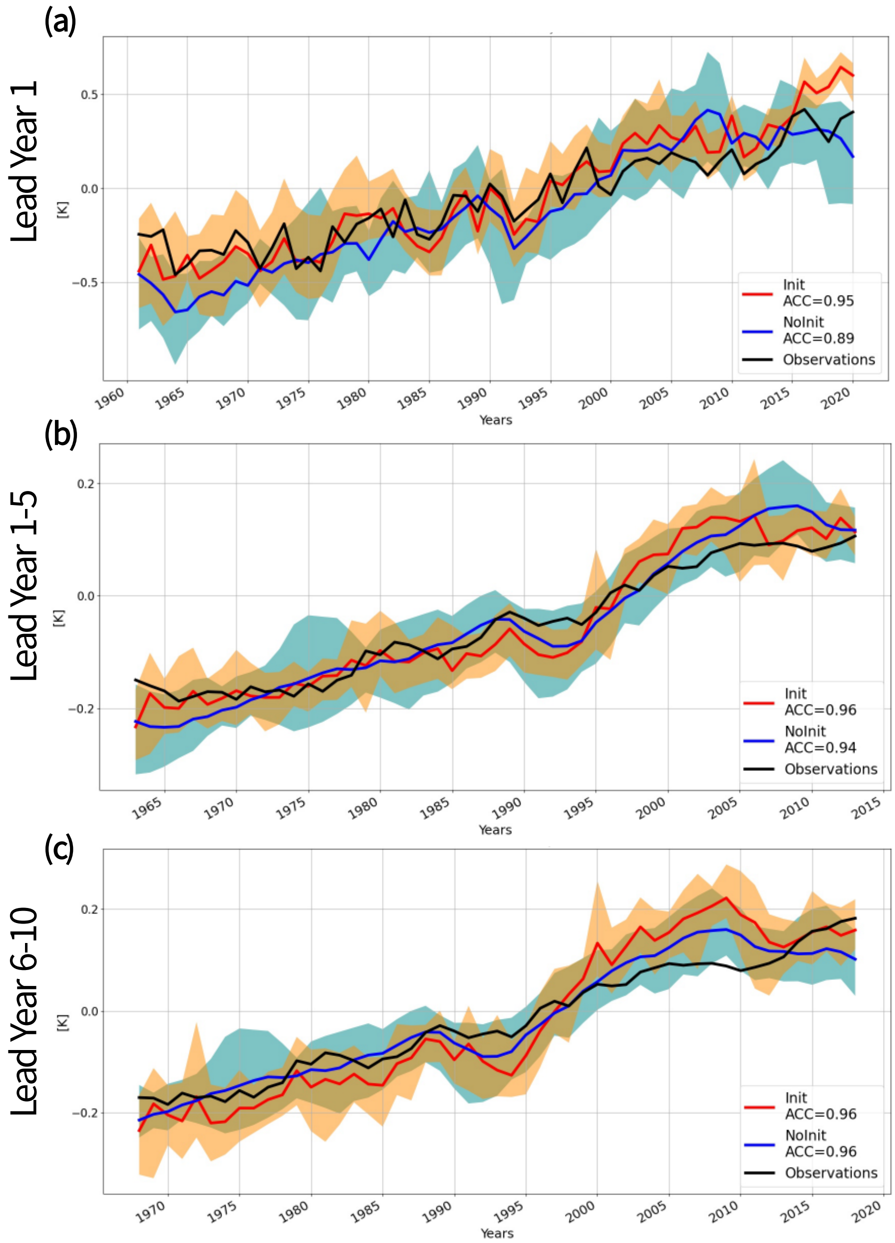
Yeager, S. G., Danabasoglu, G., Rosenbloom, N. A., Strand, W., Bates, S. C., Meehl, G. A., ... & Lovenduski, N. S. (2018). Predicting near-term changes in the Earth System: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. Bulletin of the American Meteorological Society, 99(9), 1867-1886.

Zhang, R., & Delworth, T. L. (2005). Simulated tropical response to a substantial weakening of the Atlantic thermohaline circulation. Journal of climate, 18(12), 1853-1860.

Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y. O., Marsh, R., Yeager, S. G., ... & Little, C. M. (2019). A review of the role of the Atlantic meridional overturning circulation in Atlantic multidecadal variability and associated climate impacts. Reviews of Geophysics, 57(2), 316-375.

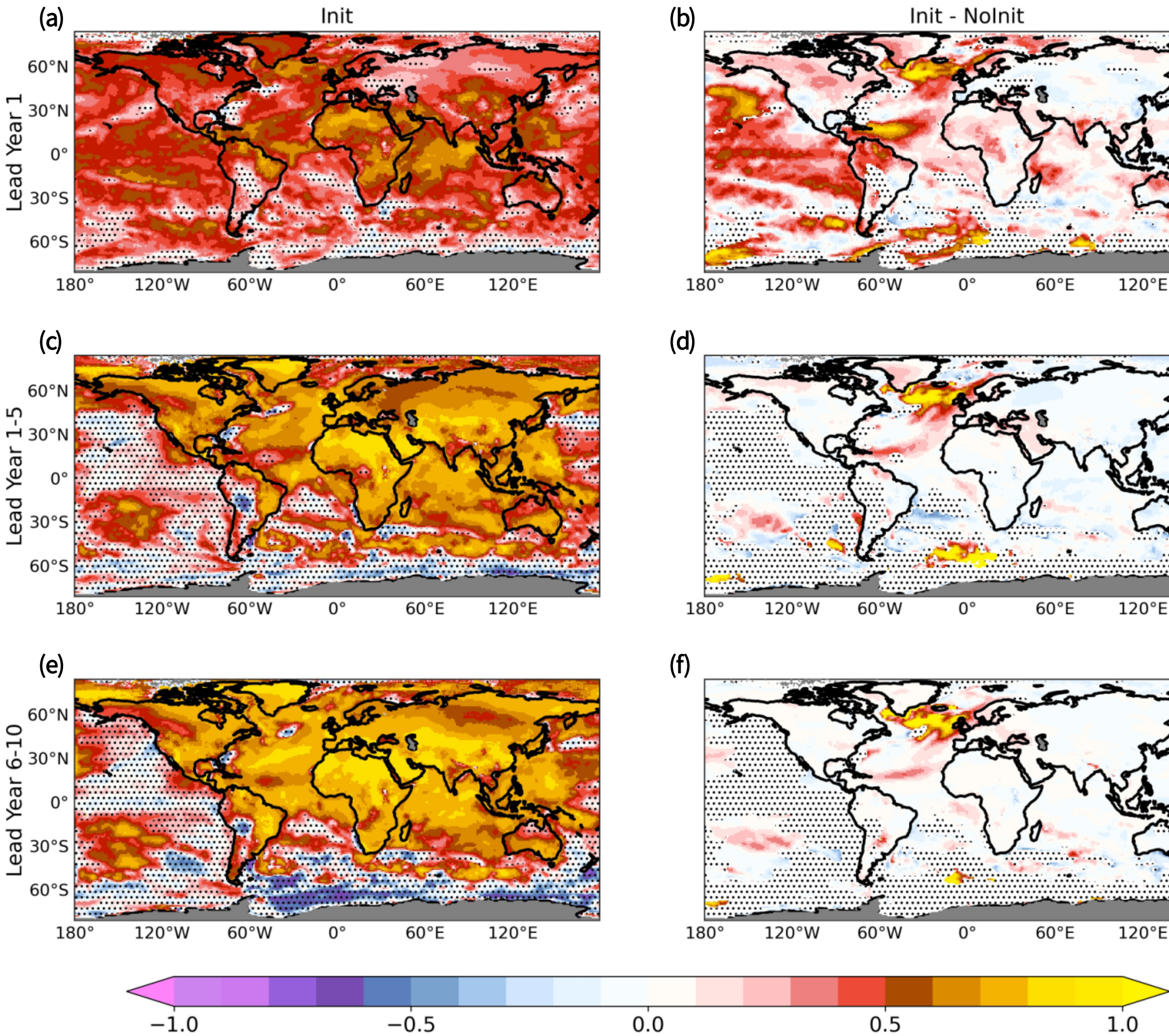
https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development
Discussions



655 **Figure 1: Global mean near-surface temperature (T2m+SST) annual average anomaly time series [K] for the hindcast (Init, in red), historical+ssp2 scenario (NoInit, in blue) and CRU ts4.05 and HadISST1.1 (in black) for (a) forecast years 1, (b) 1–5 and (c) 6–10. Orange (cyan) colored envelope denotes intra-ensemble spread for Init (NoInit). The time series are centered at the lead-year interval (e.g.: 1963 corresponds to 1961–1965 mean in panel (b).**

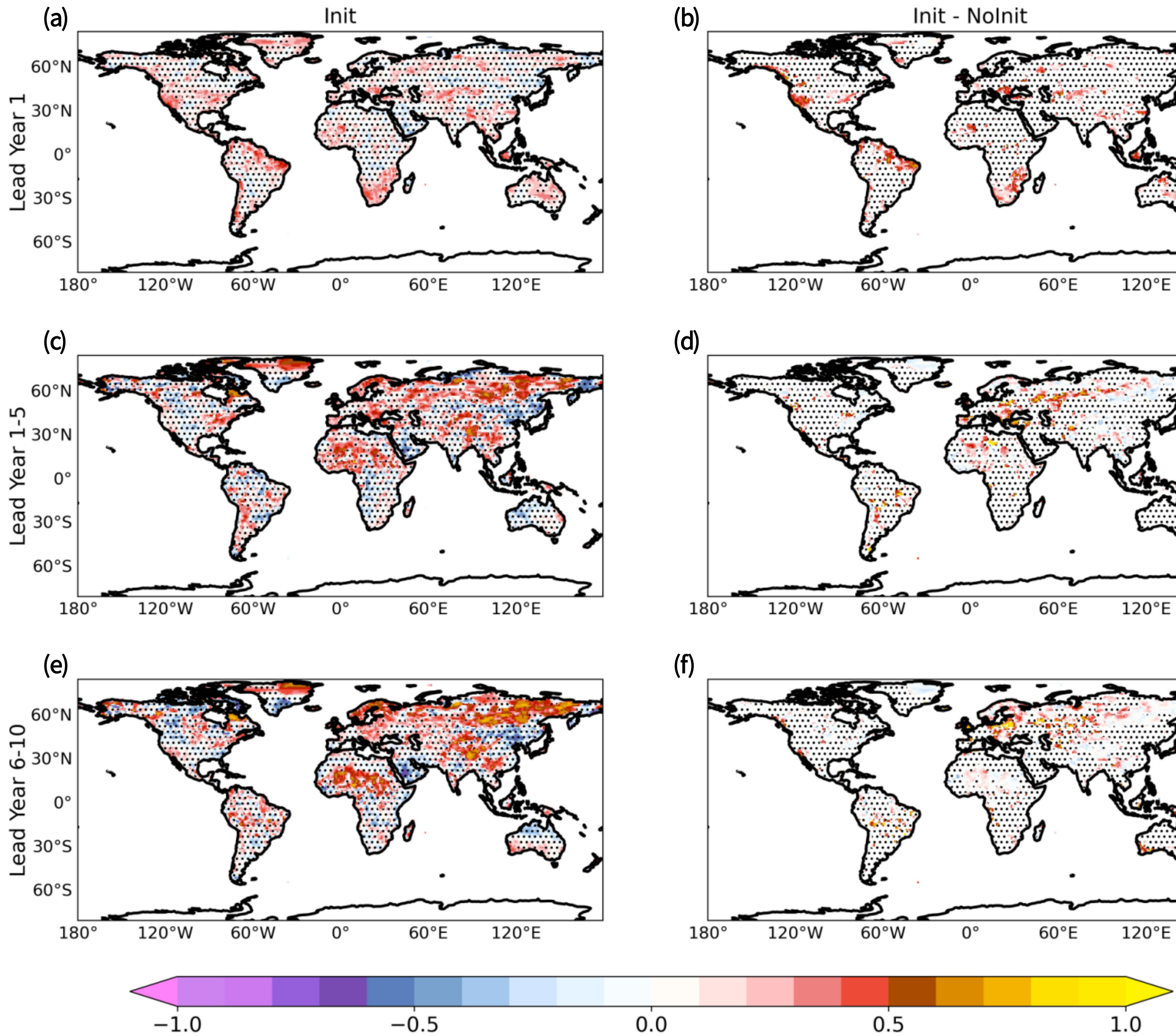
https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022



660 **Figure 2: Near surface temperature (T2m+SST) Anomaly Correlation Coefficient (ACC) of the hindcast ensemble ("Init", left column) and its difference with the NoInit ensemble ("Init - NoInit", right column) for lead years 1 (top panels), 1–5 (middle panels) and 6–10 (bottom panels). Stippling denotes points where 95% statistical significance is not reached, according to a one-tailed t test. The actual number of degrees of freedom has been computed following Bretherton et al., 1999.**

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development
Discussions



665 **Figure 3: Same as Figure 2, but for precipitation field.**

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

Geoscientific Model Development Discussions

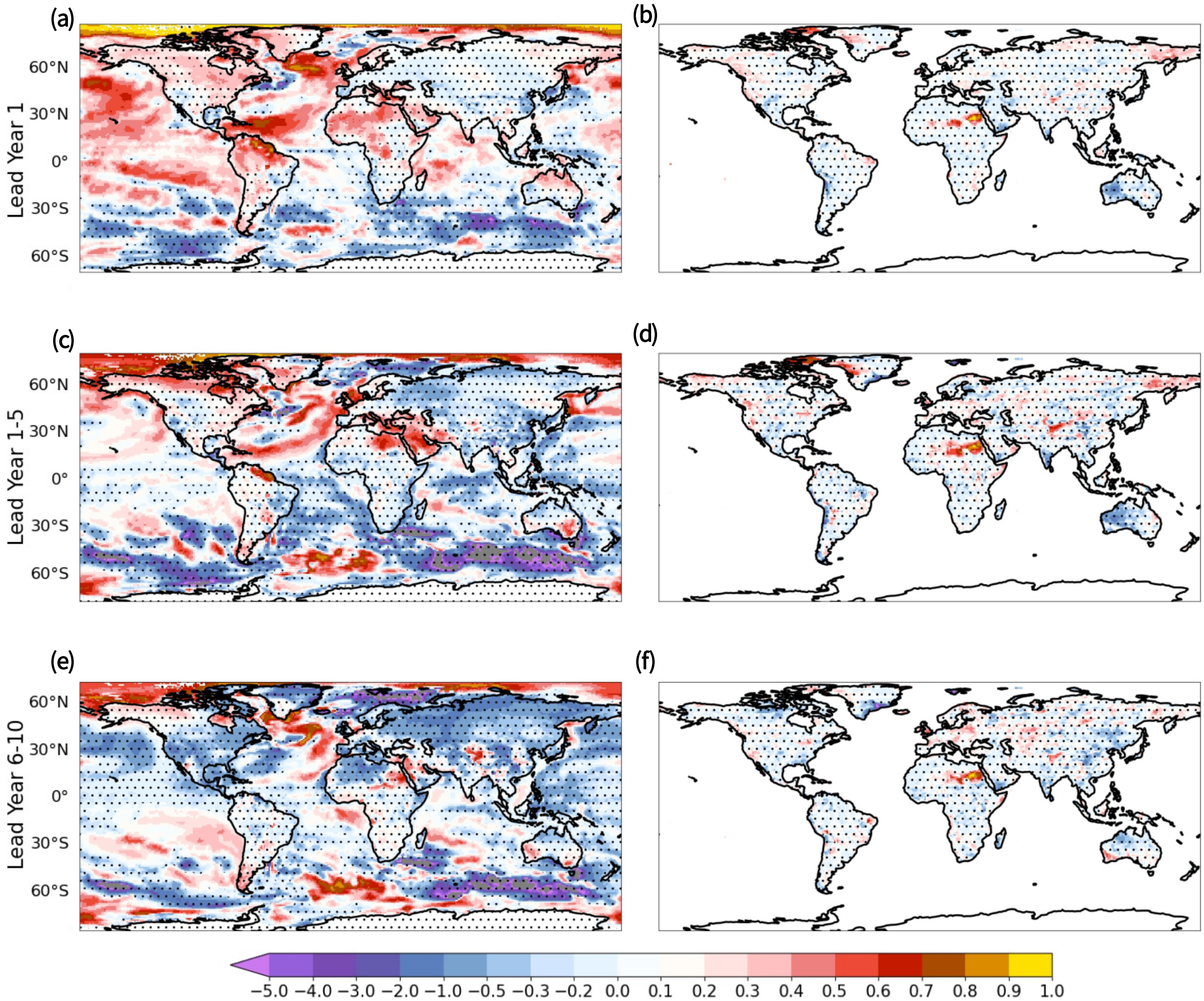


Figure 4: Near-surface temperature (T2m+SST, left column) and precipitation (right column) mean squared skill score (MSSS) of the hindcasts using NoInit runs as the reference forecast to beat. Note that the colorbar is not symmetric around zero. Stippling is used to indicate points where 95% statistical significance is not reached, according to a one–tailed t test. The actual number of degrees of freedom has been computed following Bretherton et al., 1999.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

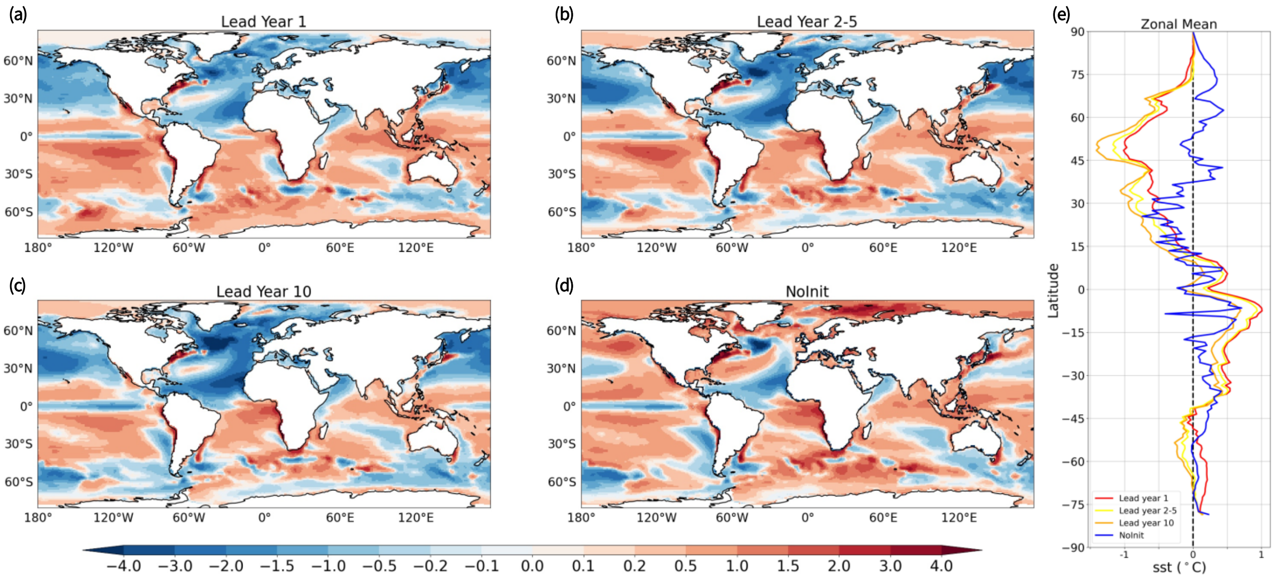


Figure 5: Mean SST Bias for (a) year 1, (b) year 2–5, (c) year 10, (d) NoInit runs and (e) their zonal mean (scaled with the cosine of the latitude) with respect to the 1960–2020 period from HadISST1.1 dataset (Rayner etal., 2003). Unit is °C.

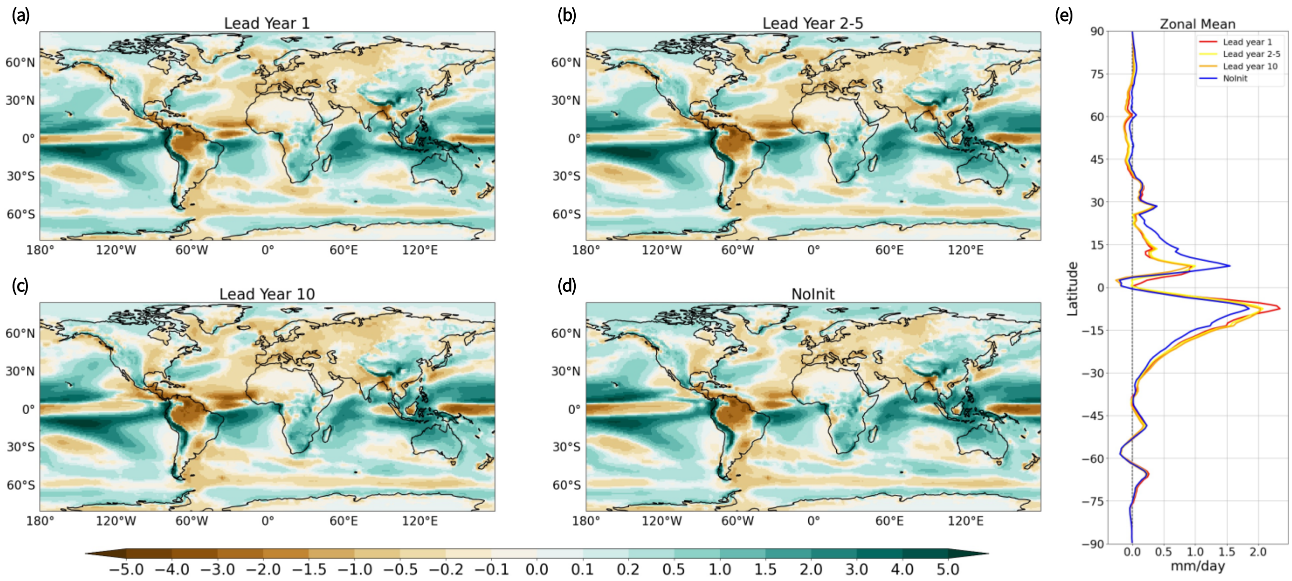


Figure 6: Mean Precipitation Bias for (a) year 1, (b) year 2–5, (c) year 10, (d) NoInit runs and (e) their zonal mean (scaled with the cosine of the latitude) with respect to the 1960–2020 period from GPCC dataset [REF]. Unit is mm/d.

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

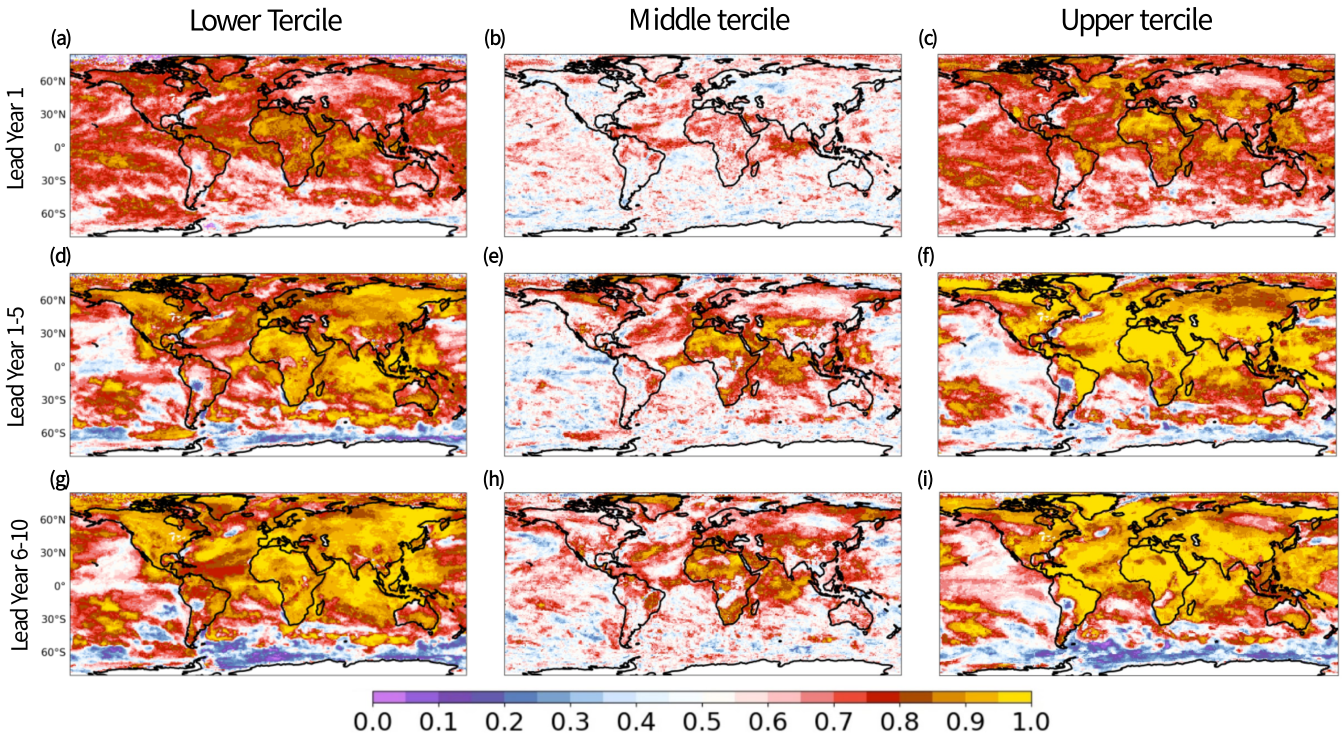


Figure 7: Relative Operating Characteristic (ROC) for near-surface temperatures (SST/TAS) for lead years 1, 1–5 and 6–10, considering three tercile categories: lower tercile (left column), middle tercile (central column) and upper tercile (right column).

https://doi.org/10.5194/gmd-2022-181
Preprint. Discussion started: 21 July 2022

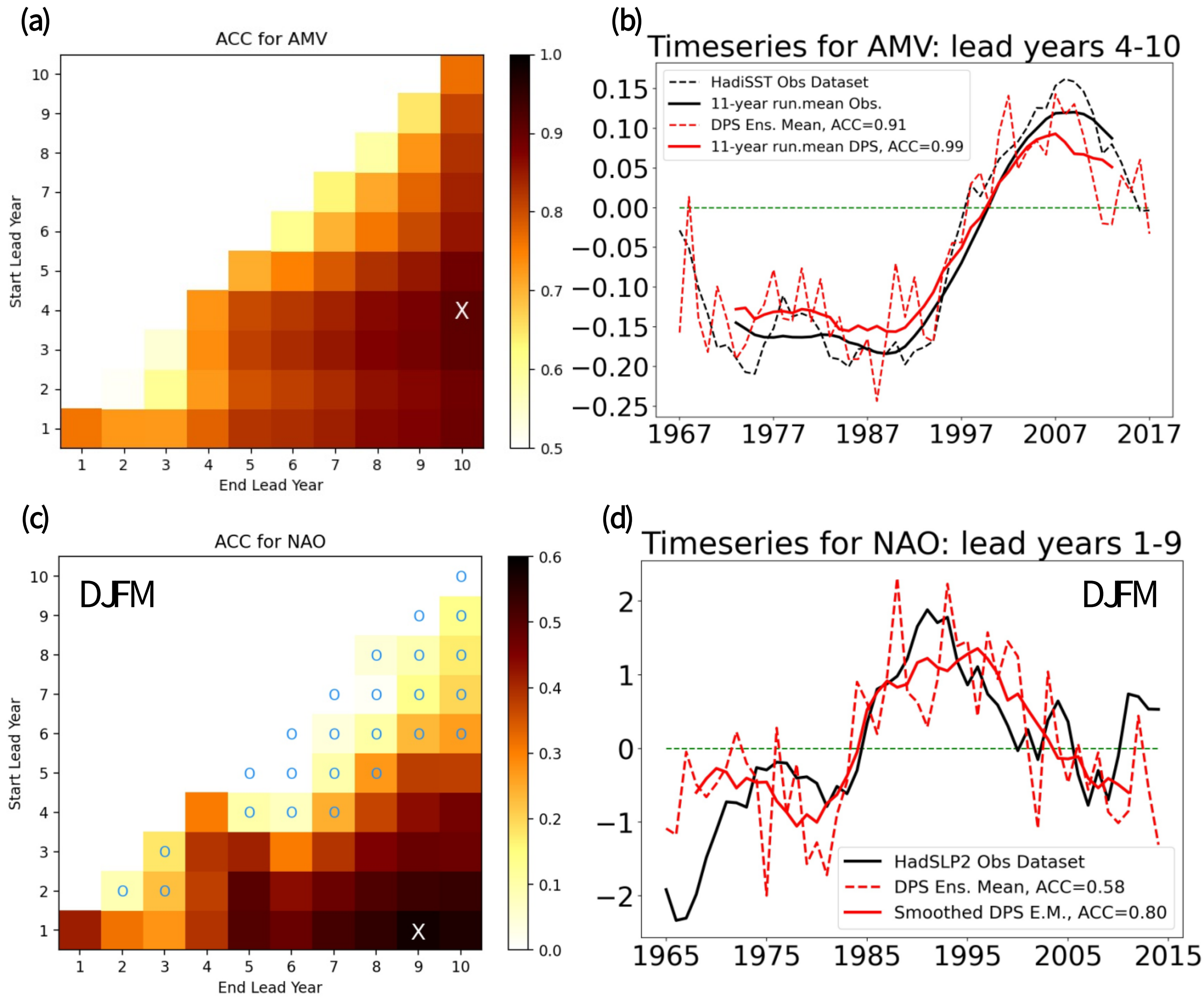
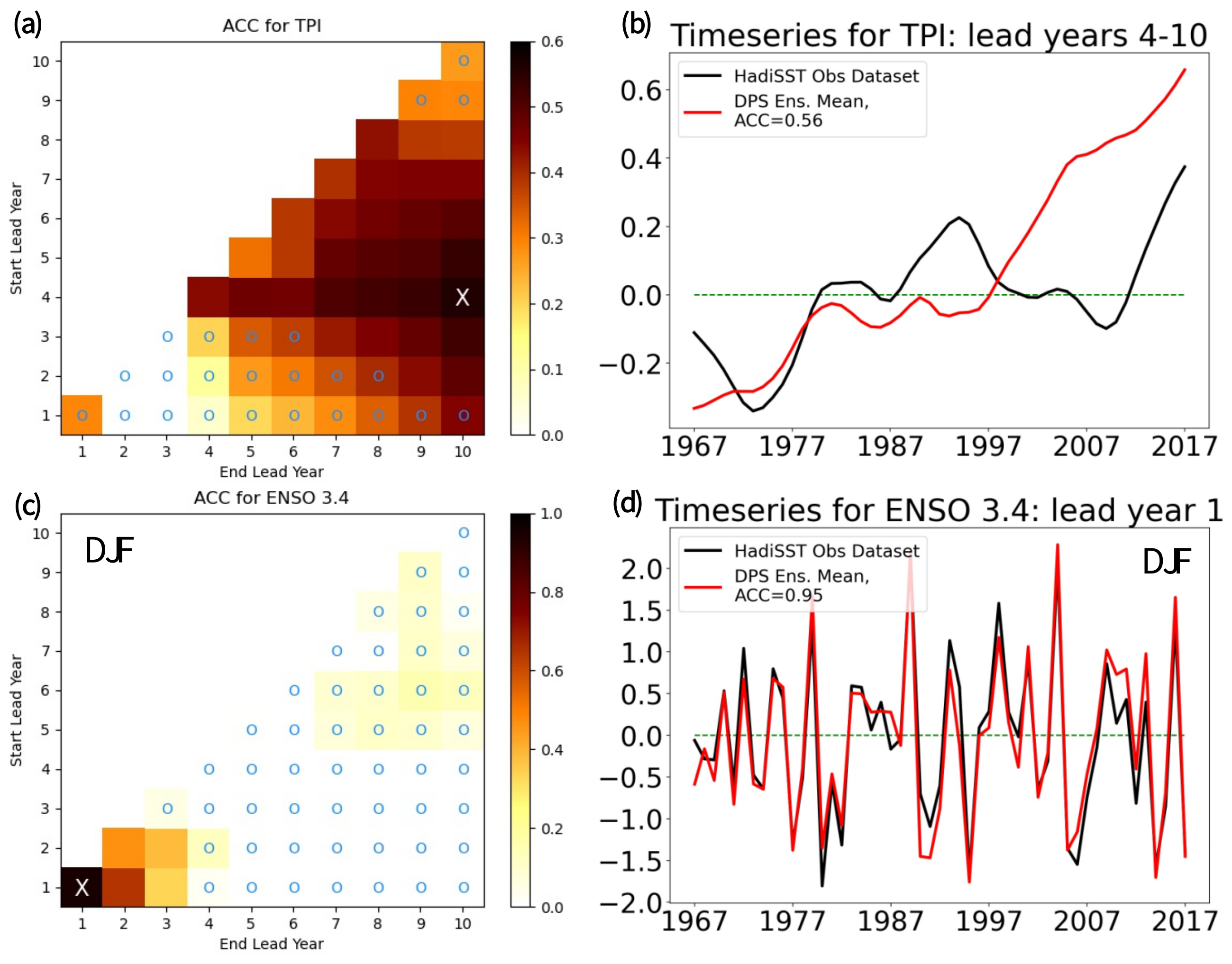


Figure 8: (a) ACC for AMV index. The colorbar ranges from 0.5 to 1. The cyan markers indicate not statistically significant correlations. White cross denotes the maximum value (ACC=0.91). (b) Observed and predicted AMV index for lead year 4–10 (in which ACC is maximum). (c) same of (a) but for DJFM NAO index. The colorbar ranges from 0 to 0.6 (maximum ACC=0.58). (d) Observed and predicted NAO index for lead year 1–9 (in which ACC is maximum).



695 **Figure 9: (a) ACC for TPI index. The colorbar ranges from 0 to 0.6. The cyan markers indicate not statistically significant correlations. White cross denotes the maximum value (ACC=0.56). (b) Observed and predicted TPI index for lead years 4–10 (in which ACC is maximum). (c) same of (a) but for DJF ENSO3.4 index. The colorbar ranges from 0 to 1 (maximum ACC=0.95). (d) Observed and predicted ENSO 3.4 index for lead year 1 (in which ACC is maximum).**