



Seamless seasonal to multi-annual predictions of temperature and Standardized Precipitation Index by constraining transient climate model simulations

Juan C. Acosta Navarro¹, Alvisé Aranyossy², Paolo De Luca², Markus G. Donat^{2,3},
Arthur Hraest Essenfelder¹, Rashed Mahmood⁴, Andrea Toreti¹, and Danila Volpi¹

¹European Commission, Joint Research Centre, Ispra, Italy

²Barcelona Supercomputing Center (BSC), Barcelona, Spain

³Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

⁴National Center for Climate Research (NCKF), Danish Meteorological Institute, Copenhagen, Denmark

Correspondence: Juan C. Acosta Navarro (juan.acosta-navarro@ec.europa.eu)

Received: 24 January 2025 – Discussion started: 10 February 2025

Revised: 23 May 2025 – Accepted: 5 August 2025 – Published: 15 October 2025

Abstract. Seamless climate predictions integrate forecasts across various timescales to provide actionable information in sectors such as agriculture, energy, and public health. While significant progress has been made, there is still a gap in the continuous provision of operational forecasts, particularly from seasonal to multi-annual timescales. We demonstrate that filling this gap is possible using an established climate model analog method to constrain variability in CMIP6 climate simulations. The analog method yields predictive skill for surface air temperature forecasts across timescales, ranging from seasons to several years. On average, the analog-based surface air temperature predictions provide added value over the unconstrained CMIP6 ensemble, especially on seasonal to annual timescales. Similar to operational climate prediction systems, Standardized Precipitation Index forecasts are less skillful than surface air temperature forecasts but still better than the CMIP6 unconstrained simulations. The analog-based seamless prediction system shows very similar patterns of skill compared to state-of-the-art initialized climate prediction systems and has competitive skill on annual and biennial forecast ranges. While the current prediction systems provide only 1–2 initializations per year, the analog-based system can easily provide predictions with monthly initializations, delivering seamless climate information throughout the year currently not available from traditional seasonal or decadal prediction systems. Furthermore, due to analog-based predictions being computationally inexpensive, we argue that these methods are a valuable and viable complement to existing operational prediction systems.

1 Introduction

Seamless climate prediction aims at integrating and synthesizing climate forecasts over a range of forecast times, from sub-seasonal to multi-decadal timescales (Kirtman et al., 2013; Merryfield et al., 2020; Meehl et al., 2021). It is rooted in the concept that the internal climate variability is not confined to any specific timescale but instead spans days to several decades (Schindler et al., 2015; Zhang et al., 2020). Seamless climate prediction can support various practical applications, ranging from managing agriculture or water re-

sources to better preparing for climate-related disasters and better meeting energy demands (Buontempo et al., 2018; Bett et al., 2022; Sánchez-García et al., 2022).

On subseasonal to seasonal timescales (i.e. from a few weeks to a few months), seamless climate prediction aims to inform about variability associated with phenomena such as the Madden Julian Oscillation (Kim et al., 2019a) or sudden stratospheric warming events (Sigmond et al., 2013). Climate and weather variability in these timescales can affect sectors such as agriculture, energy production, and public health (Thomson et al., 2006; Klemm and McPherson, 2017; Kim et

al., 2019b; Lledó et al., 2019; Ceglar and Toreti, 2021). Seasonal to multi-annual climate predictions (i.e. from a season to a few years) provide information to better anticipate climate variations that are externally forced or occur due to natural variability within the climate system and which include for example the El Niño–Southern Oscillation (ENSO, Lopez and Kirtman, 2014), the Indian Ocean Dipole (Shinoda and Han, 2005), or the Arctic Oscillation (Riddle et al., 2013) and Antarctic Oscillation (Seviour et al., 2014), being important in various sectors including agriculture, water resource management, energy, public health, and disaster risk management (Caron et al., 2015; Solaraju-Murali et al., 2021; Dunstone et al., 2022). Decadal and multi-decadal predictions provide information on longer climate trends and variability like the Atlantic Multidecadal Variability (Mann et al., 2014) or Pacific Decadal Oscillation (Liu and Di Lorenzo, 2018) which are essential for long-term planning in infrastructure, resource management, and climate change adaptation (Solaraju-Murali et al., 2022; Dunstone et al., 2022).

Operational seasonal and decadal climate predictions are produced by integrating forward in time an ensemble of several parallel climate model simulations forced by a likely external forcing scenario and initialized from a climate state that is representative of the observed climate (Meehl et al., 2021). The ensemble of model simulations is meant to constitute a pool of equally probable realizations of future climate. After initialization, the models are often subject to shocks followed by a drift away from the observed climate typically towards its own attractor. This can result in a reduction of forecast skill (Bilbao et al., 2021). Model simulations used to deliver climate predictions are computationally expensive, thus being produced only by a limited number of institutions around the world. Analog-based approaches are alternatives that exploit the freely available large ensembles of non-initialized climate model simulations such as the ones from the Coupled Model Intercomparison Project Phase 6 (CMIP6; Eyring et al., 2016) to produce computationally cheaper climate predictions. These approaches work by scanning for analogs of the observed climate in a large model catalog, typically selecting a subset of them in order to better constrain the variability of those simulations and provide predictability beyond the one determined by the externally forced signal alone (see methods). An incomplete representation of the climate state at initialization is likely the major disadvantage of the analog-based predictions because of the finite available states present in the multi-model catalog. These states may be less representative or “further away” from the observed target state than the initial states in an initialized climate prediction system. Despite these potential disadvantages from a lack of a more sophisticated initialization, the simulations used in the analog-based predictions are not impacted by initialization shocks, and their direct use is computationally cheap. Analog-based prediction methods have been successfully applied on seasonal/annual scales by Ding et al. (2018, 2019) to predict climate in the

tropics (e.g. multi-year ENSO forecasts) and by Mahmood et al. (2021, 2022), De Luca et al. (2023), and Donat et al. (2024) on decadal to multi-decadal timescales to give an outlook beyond the available operational decadal predictions.

Despite recent progress in existing prediction systems (Kirtman et al., 2013; Kushnir et al., 2019; Merryfield et al., 2020; Meehl et al., 2021), only operational decadal predictions provide information across these different timescales, yet this information is only available typically at the beginning of each year. Operational seasonal forecasts do provide information about once a month, but the forecast horizon is typically limited to 6 months. For example, a user interested in obtaining climate information for different timescales (e.g. seasonal to multi-annual) would currently have to combine the information from different prediction systems for the different timescales which are often inconsistent in their setup, the model used, and the predictions they provide. In this study we show that this key climate information gap on the seasonal to multi-annual timescales can be filled by exploiting the model analog method to constrain existing non-initialized CMIP6 simulations.

2 Methods

We build from the hypothesis that finding the climatic states (analogs) in simulations from a large multi-model ensemble that are closest to an observed target state can provide valuable information on the future evolution of the climate system (Mahmood et al., 2022). The CMIP6 ensemble is currently the largest available pool of simulations from multiple state-of-the-art climate models. In this study we use data from 149 climate simulations from 19 climate models covering the period 1960–2030 forced by historical emissions before 2015 and the SSP2-4.5 scenario emissions afterwards (Table S1 in the Supplement). The total number of model simulation years is 10 579. The period 1960–2030 was chosen to include both a climatically representative period of the hindcast and an extension to the future to allow for the occurrence of unprecedented climatic states in a real-time forecasting context.

More specifically, we scan across time and ensemble members for the conditions that better resemble the observed sea surface temperature (SST) anomaly pattern over oceans at a given time as a means to align the natural climate variability around the climatological state of the model to the observed one, which conceptually corresponds to the initialization of climate predictions. To do this, we first estimate the area-averaged, area-weighted (w) mean absolute error (MAE) of monthly SSTs for each member of each model (l in Eq. 1) and across years (k in Eq. 1) with respect to the observational reference (O in Eq. 1) at the desired target month of “initialization” (m in Eq. 1). For example, to produce the surface air temperature (TAS) or Standardized Precipitation Index (SPI) forecasts of June–August 2024 with 1-

month lead time, the observed SST anomalies of April 2024 (m) are compared with all the April SST anomalies between 1960 and 2030 (k) across members in the multi-model ensemble and ranked according to their respective MAE, for each member separately. The modeled SSTs from all months of April that have the greatest similarity with the target month of April 2024 (i.e. smallest MAE) are then selected (analogs of April 2024), and the forecast is constructed by taking the average conditions of the June–August following the selected April analogs. The selection is always done with 1-month lead time (unless otherwise noted) to provide information well ahead of the targeted forecast period. We found that the analogs generated with an SST pattern comparison for the whole planet are broadly superior to the analog using a reduced Indo-Pacific region as in Ding et al. (2018) (Tables S2–S3). Additional sensitivity tests also reveal that the optimal length (m and k) of SST pattern comparison is 1 month, independent of the different forecast ranges considered, in particular for seasonal to inter-annual predictions (Tables S2–S7). Please note that for longer (e.g. multi-annual to multi-decadal) forecast times, analogs based on longer-term SST averages were determined to give highest skill (e.g. Mahmood et al., 2022; Donat et al., 2024). The timescales of the analogs represent processes relevant for the predictions. While for seasonal to inter-annual predictions SST variations at higher frequency (e.g. ENSO) are most relevant, for longer prediction horizons (also reaching beyond the ENSO predictability barrier), other lower-frequency variations (e.g. Atlantic multidecadal variability) are more relevant. The number of analogs for each TAS (SPI) prediction is defined by the top (top 5) analog(s) in each one of the model simulations which cover the period 1960–2030. Note that for long predictions, the period of analog selection is slightly reduced at both extremes (e.g. 48-month predictions are based on analogs centered between 1963 and 2027). The number of selected analogs (i.e. 1 or 5 per member) and the number of models and members used have been determined by performing sensitivity tests (Tables S8–S13). More specifically, four methods were tested:

- *Method 1.* All available members from the six models that provide more than 10 members each (Table S1), 122 members in total.
- *Method 2.* Only 10 members from the same six models that provide more than 10 members each, 60 members in total.
- *Method 3.* Ten members from each model. For models that provide less than 10 members, the members are used more than once to complete a set of 10 for each model, 190 members in total – essentially increasing the relative weight of the analogs of models with fewer than 10 members.
- *Method 4.* All available members from the 19 models, 149 members in total.

Using the best overall method from the sensitivity tests (method 4), the selected analogs then constitute the forecasts and can be interpreted as ensemble members. Additionally, the trend of the ensemble-mean TAS analog-based predictions is adjusted by first removing the signal explained by external forcing as in Smith et al. (2019) and then adding the externally forced trend (i.e. the CMIP6 ensemble mean) to those residuals. This is necessary because the analogs can be selected from any year in the period 1960–2030 and do not necessarily have the right forcing state. The trend adjustment ensures that potential offsets related to selecting analogs from other forcing states are corrected to represent the forcing of the year(s) of the predictions. For the SPI predictions the trend adjustment is not needed, as they are optimal without it (not shown).

Mathematically, the criterion to rank and determine the analogs is

$$\text{MAE}_{l,k} = \frac{\sum_i \sum_j^{w_j} |T_{i,j,l,k} - O_{i,j,m}|}{N}, \quad (1)$$

where the indices i , j , k , and l run across longitudes, latitudes, months, and models/members, respectively. T stands for the model values, and N stands for the total number of ocean grid points.

We apply the methodology described above to generate retrospective predictions of SPI and TAS of 3 months, 1 year, 2 years, and 4 years and evaluate their predictive skill in the period 1962–2018, except for the 3-month predictions which are only evaluated during 1982–2018, defined by the availability of the benchmark SEAS51 predictions (see next paragraph). We compute the SPI (McKee et al., 1993) for 3-, 12-, 24-, and 48-month accumulations using the R package SPEI (Beguiria and Vicente-Serrano, 2023). Following Smith et al. (2019), the non-forced analog-based predictions and observations are by definition the residuals that contain only the variability that is not explained by the CMIP6 ensemble mean. Hence, the non-forced skill throughout the study can be interpreted as the residual skill explained after removing the externally forced signal. The skill metrics used in this study are the anomaly correlation coefficient (ACC) and the mean absolute error skill score (MAESS):

$$\text{ACC} = \frac{\sum_i^n (F_i - \bar{F})(O_i - \bar{O})}{\sqrt{\sum_i^n (F_i - \bar{F})^2} \sqrt{\sum_i^n (O_i - \bar{O})^2}} \quad (2)$$

$$\text{MAESS} = 1 - \frac{\sum_i^n |(F_i - \bar{F}) - (O_i - \bar{O})|}{\sum_i^n |(R_i - \bar{R}) - (O_i - \bar{O})|}. \quad (3)$$

In Eqs. (2) and (3) F stands for forecasted values, and O stands for observed values, and in Eq. (3) the reference R is a trivial climatological forecast based on observations, the multi-model uninitialized CMIP6 ensemble mean, or the ensemble mean forecast from an operational prediction system. The letters with the bars denote climatological values. The index i in Eqs. (2) and (3) runs across time.

The list of CMIP6 models used in this study is available in Table S1. All model and observational data have been bilinearly interpolated to a common grid of $5^\circ \times 5^\circ$ for TAS and about $2.8^\circ \times 2.8^\circ$ for both SSTs (i.e. analog search) and SPI. The observational reference datasets for SST analog search is ERSSTv5 (Huang et al., 2017), while the observational references for prediction evaluation of TAS and SPI are based on the monthly averages of Berkeley Earth Surface Temperatures (BEST, Cowtan) and the Global Precipitation Climatology Center (GPCC, Becker et al., 2013), respectively. In addition to the CMIP6 simulations, we used data from two operational climate prediction systems as a benchmark for the comparisons. For the 3-month predictions, we used the European Centre for Medium-Range Weather Forecasting SEAS51 (Johnson et al., 2019) and for the 12-, 24-, and 48-month predictions, we used the initialized climate model EC-Earth3 (Bilbao et al., 2021). Note that both dynamical prediction systems are limited to 25 members, whereas the analog-based predictions are based on the 149 members from the non-initialized CMIP6 ensemble. A key strength of the analog-based method is its ability to leverage a large-sized ensemble at minimal computational cost as opposed to the significant cost it requires to generate such large ensembles with initialized prediction systems. However, we acknowledge that a fraction of the skill of the analog-based predictions stems from exploiting large ensembles, and reducing the ensemble size to match the size of the dynamical prediction systems reduces the skill. This is demonstrated in Fig. S1, which shows that the skill of the analog-based predictions clearly increases with ensemble size, regardless of variable or forecast range.

3 Results

3.1 Seasonal predictions (3 months)

Figure 1a–f illustrate the spatial distribution of skill for boreal winter (December–February) TAS predictions initialized on 1 November, as assessed by the ACC and the MAESS. The analog method shows positive statistically significant correlation in the tropics and subtropics, most of the ocean, and the Arctic (Fig. 1a). A large fraction of skill in these areas can be attributed to the alignment of internal variability in the predictions and observations as revealed by the residual correlation after removing the externally forced signal from CMIP6 (Fig. 1b). In general, the analog-based predictions offer added value over a trivial climatological forecast (Fig. 1d) and over the uninitialized CMIP6 ensemble (Fig. 1e), especially over tropical regions. Figure 1c displays the correlation between the observations and SEAS51, an operational seasonal forecast system, while Fig. 1f displays the added value of the analog-based predictions over the SEAS51 ones according to the MAESS. Generally, SEAS51 has higher skill than the analog-based predictions, especially marked in ocean regions like the North Atlantic. However, the overall

spatial patterns are very similar between the analog-based and SEAS51 predictions, which gives some confidence that both exploit similar sources of predictability.

Although not as widespread as December–February predictions, boreal summer (June–August) TAS predictions initialized on 1 May also display high skill in the tropics. Additionally, skill is also high in many subtropical and mid-latitude regions (Fig. 2a, d). Generally, the skill in Northern Hemisphere land regions is higher in boreal summer than in winter. Specifically, the Middle East, Europe, and large parts of East Asia show high skill in terms of both ACC and MAESS, although this skill stems primarily from the response to external forcing and not from the analog initialization. Hence, the added value of analog-based predictions over the uninitialized CMIP6 ensemble is mostly limited to tropical and subtropical regions according to the residual correlation (Fig. 2b) but limited to Central America, Southeast Asia, and tropical oceans according to MAESS (Fig. 2e). Again there is a very large similarity between the spatial patterns of skill in the analog and SEAS51 predictions (Fig. 2a, c, respectively). SEAS51 shows larger correlations with observations in general, but similar to boreal winter, the disadvantage of the analog-based predictions over land areas seems limited and mostly not statistically significant with respect to SEAS51 when evaluated with the MAESS (Fig. 2f).

Figure 3 shows the correlation coefficients of the SPI3 analog seasonal forecasts and observations during the boreal winter (Fig. 3a–c) and summer (Fig. 3d–f), respectively. The analog-based predictions exhibit skill in Australasia, southern Africa, and the tropical Americas. In line with what is observed in forecasts from dynamical forecast systems, the analog technique yields predictions for SPI3 that are notably less skillful than those for TAS. Nonetheless, the spatial patterns of regions with skill in the analog and SEAS51 are very similar (Fig. 3a, c, d, f). The residual correlation of precipitation forecasts with the analog method during the boreal winter and summer is shown in Fig. 3b and e, respectively. The similarity of these maps to the full-skill maps (i.e. Fig. 3a, d) suggests that the analog-based predictions' accuracy is predominantly due to the alignment of natural climate variability in both the models and observations, with one notable exception: the Sahel region in boreal summer (Fig. 3d, e), in which the skill seems to result from the external forcing (Ndiaye et al., 2022). It is important to note that the MAESS of both boreal summer and winter analog-based predictions of SPI3 when using SEAS51 as a reference shows generally non-statistically significant differences, similarly to land TAS (not shown).

The SPI3 analog-based predictions show skill comparable to SEAS51 predictions between boreal fall and spring in terms of land area with positive and statistically significant residual correlation, while the analog-based predictions of 3-month TAS are generally less skillful than SEAS51 throughout the year (Fig. 4). Skill over land peaks around boreal summer/fall and fall/winter for TAS and SPI3, respectively,

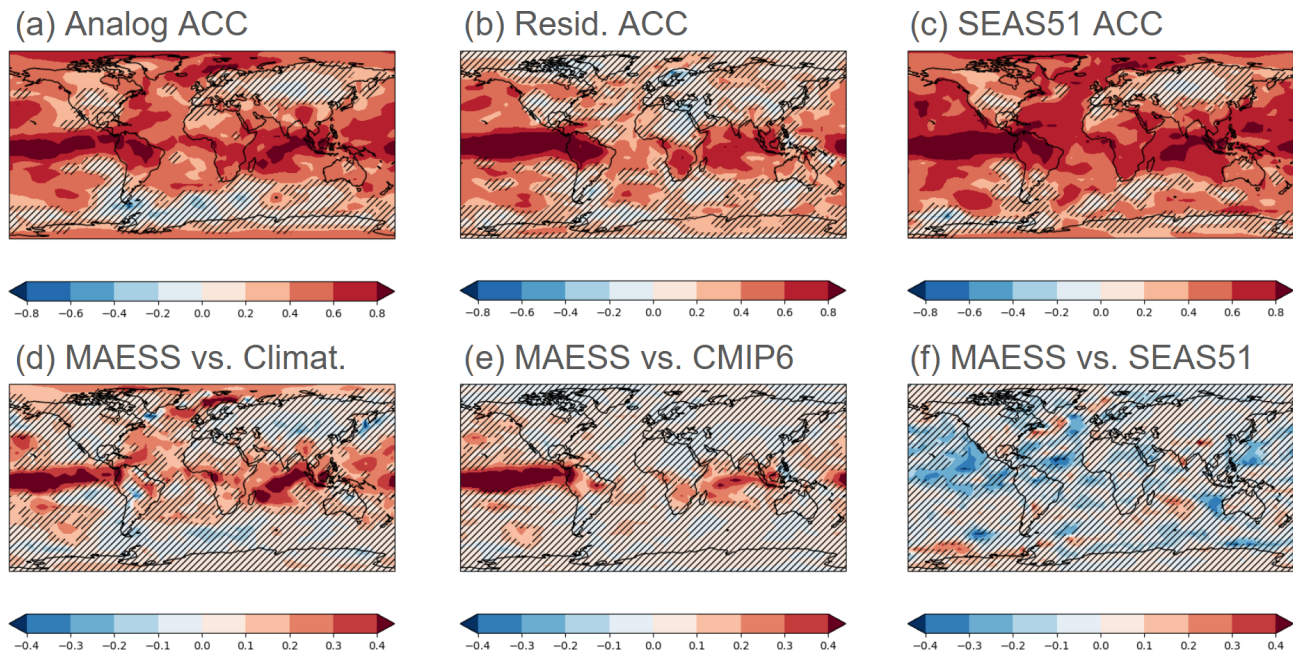


Figure 1. (a) ACC between December–February analog-based ensemble mean predictions and observations of TAS. (b) Residual ACC between December–February analog-based ensemble mean predictions and observations of TAS. (c) ACC between ECMWF–SEAS51 December–February ensemble mean predictions and observations of TAS. (d) MAESS of December–February analog-based ensemble mean predictions of TAS. The reference (R) is a climatological forecast derived from observations. (e) MAESS of December–February analog-based ensemble mean predictions of TAS. The reference (R) is the ensemble mean of CMIP6 historical simulations. (f) MAESS of December–February analog-based ensemble mean predictions of TAS. The reference (R) is the ECMWF–SEAS51 December–February ensemble mean predictions of TAS. The evaluation period is 1982–2018. The predictions are initialized each November in both analog-based predictions and SEAS51. The hatched regions in all figures indicate statistically non-significant values ($p \geq 0.1$) using a two-sided t test.

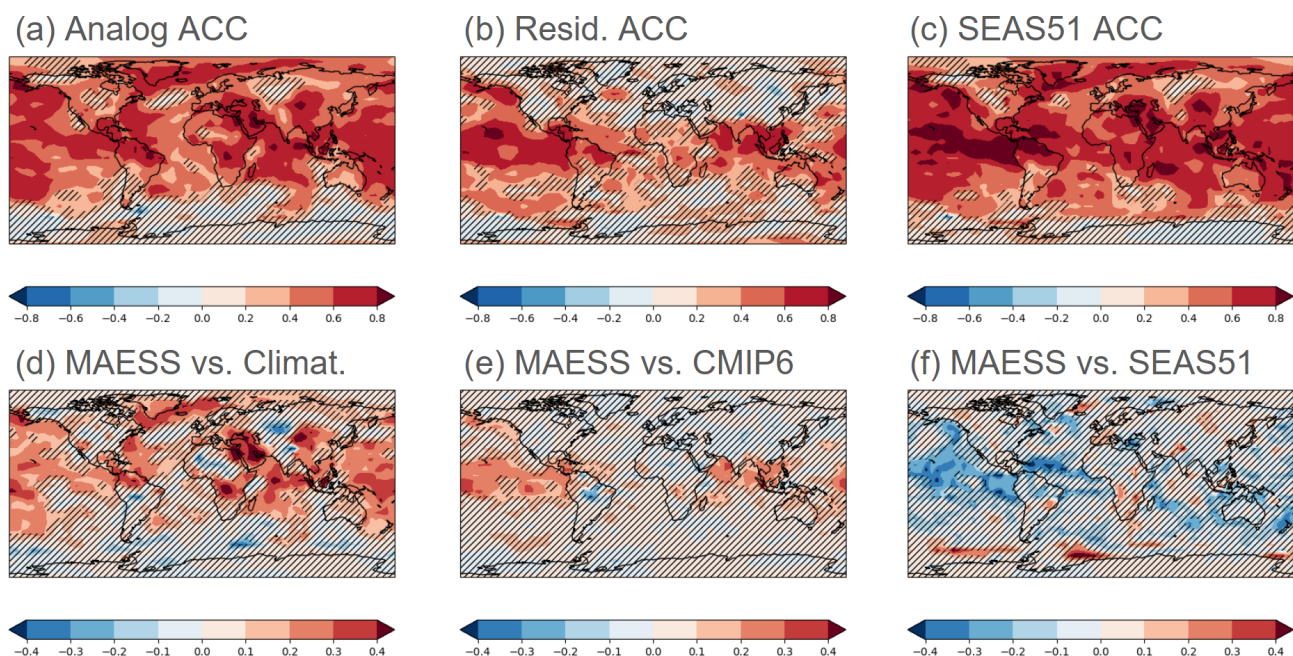


Figure 2. The same as Fig. 1 but for June–August TAS predictions. The predictions are initialized each May.

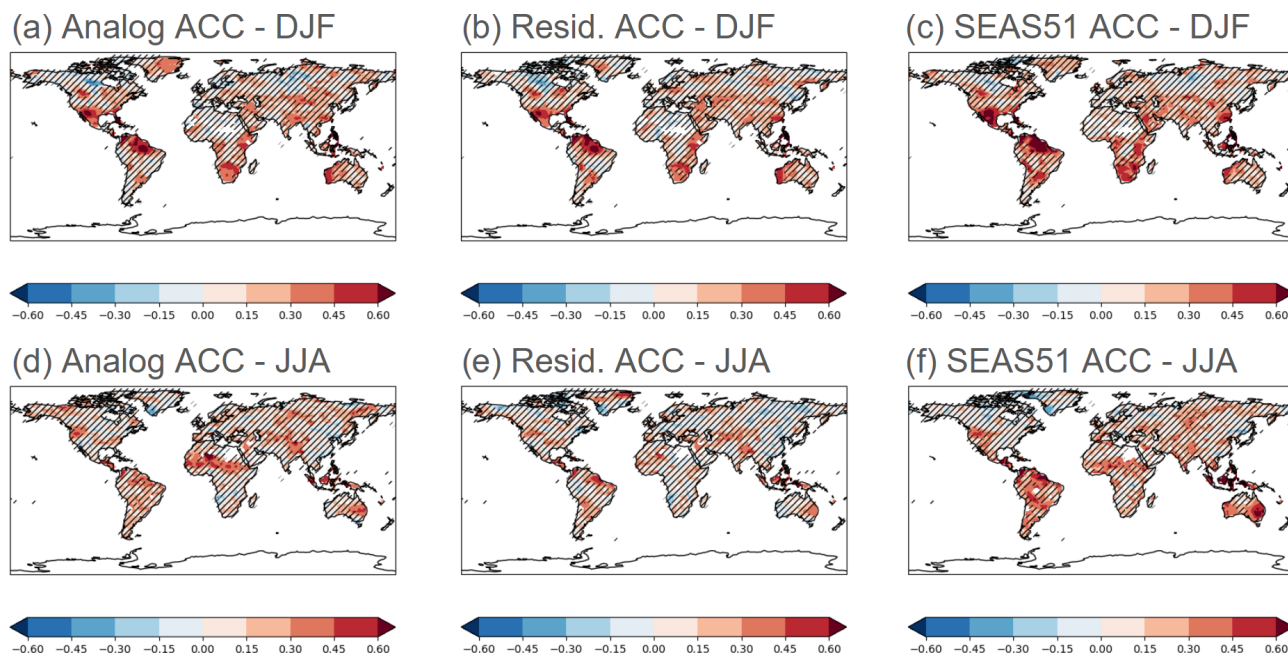


Figure 3. (a) ACC between December–February analog-based ensemble mean predictions and observations of SPI3. (b) Residual ACC between December–February analog-based ensemble mean predictions and observations of SPI3. (c) ACC between ECMWF-SEAS51 December–February ensemble mean predictions and observations of SPI3. (d) ACC between June–August analog-based ensemble mean predictions and observations of SPI3. (e) Residual ACC between June–August analog-based ensemble mean predictions and observations of SPI3. (f) ACC between ECMWF-SEAS51 June–August ensemble mean predictions and observations of SPI3. The evaluation period is 1982–2018. The predictions are initialized each November (a–c) and each May (d–f). The hatched regions in all figures indicate statistically non-significant values ($p \geq 0.1$) using a two-sided t test.

in both analog and SEAS51 predictions. This difference between the two variables can most likely be attributed to a more dominant influence of external forcing on TAS predictability, while for SPI3 the primary driver is natural variability.

Figure 5 shows the temporal evolution of four key seasonal climate indices of the tropical oceans: the December–February NINO34 (Fang and Xie, 2020) in the tropical Pacific, the June–August tropical Atlantic index (ATL3, Zebiak, 1993), and the March–May Western (WIO) and September–November Eastern (EIO, Saji et al., 1999) tropical Indian Ocean indices. All forecasts are initialized 1 month before the target season. The indices in the respective seasons are important because they measure oceanic variability that induces remote impacts on hydroclimatic conditions over land. They are a small selection to highlight and confront the analog and the SEAS51 predictions in these particular areas. Although not as skillful as SEAS51, the analog-based predictions of NINO34 and the WIO show high skill and closely follow the observed year-to-year variability and trend. The ATL3 and EIO follow the observed trend but largely underestimate the magnitude of year-to-year variability as opposed to SEAS51. This underestimation of the local variability may be a result of the sampling and averaging of hundreds of analogs, with some of them being less representative of the

observed conditions in the particular area. Despite this, at the global scale, the larger the ensemble of analogs the higher the skill (Fig. S1). As commonly done for other types of climate predictions, recalibrating the ensemble could make the analog-based forecasts more valuable.

3.2 Annual and multi-annual predictions (1–4 years)

The skill of annual TAS analog-based predictions is very high ($\text{ACC} > 0.8$, $\text{MAESS} > 0.3$) across most tropical areas and the North Atlantic, as well as being high, positive, and statistically significant over land regions outside of the tropics, as shown in Fig. 6a, d. However, skill in areas like central Asia, central North America, and southern South America exhibits mostly low to moderate skill ($\text{ACC} > 0.2$) but still surpasses that of a climatological forecast ($\text{MAESS} > 0$). The results indicate a distinct improvement of the annual analog forecasts when compared to the CMIP6 ensemble across the Pacific region, the tropical Atlantic and Indian oceans, Australasia and East Asia, and large parts of the Americas and Africa. However, based on MAESS, these improvements are limited mostly to the Caribbean, southern Africa, and the Maritime Continent (Fig. 6e). This discrepancy between ACC and MAESS is likely the result of analog-based predictions being capable of estimating the variability

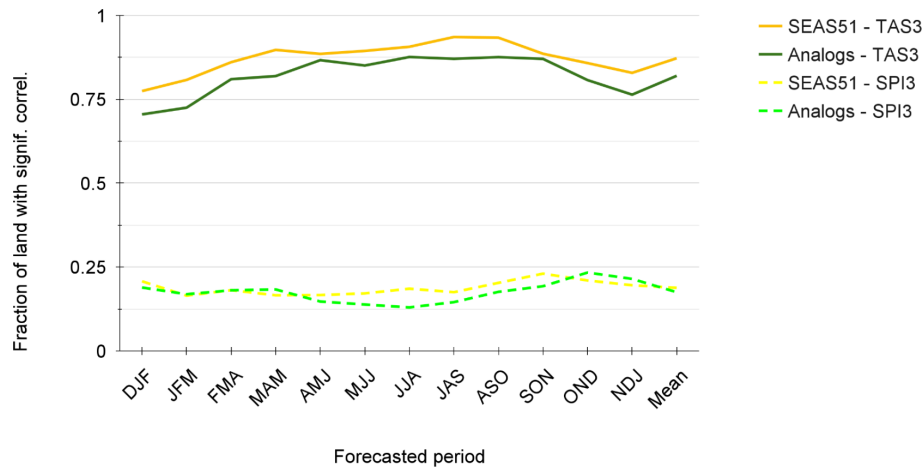


Figure 4. Fraction of land area with statistically significant positive correlation ($p < 0.1$) between the 3-month TAS (solid lines) and SPI3 (dashed lines) from the analog, the SEAS51 predictions, and the respective observations. Statistical significance is assessed using a two-tailed t test. The evaluation period is 1982–2018.

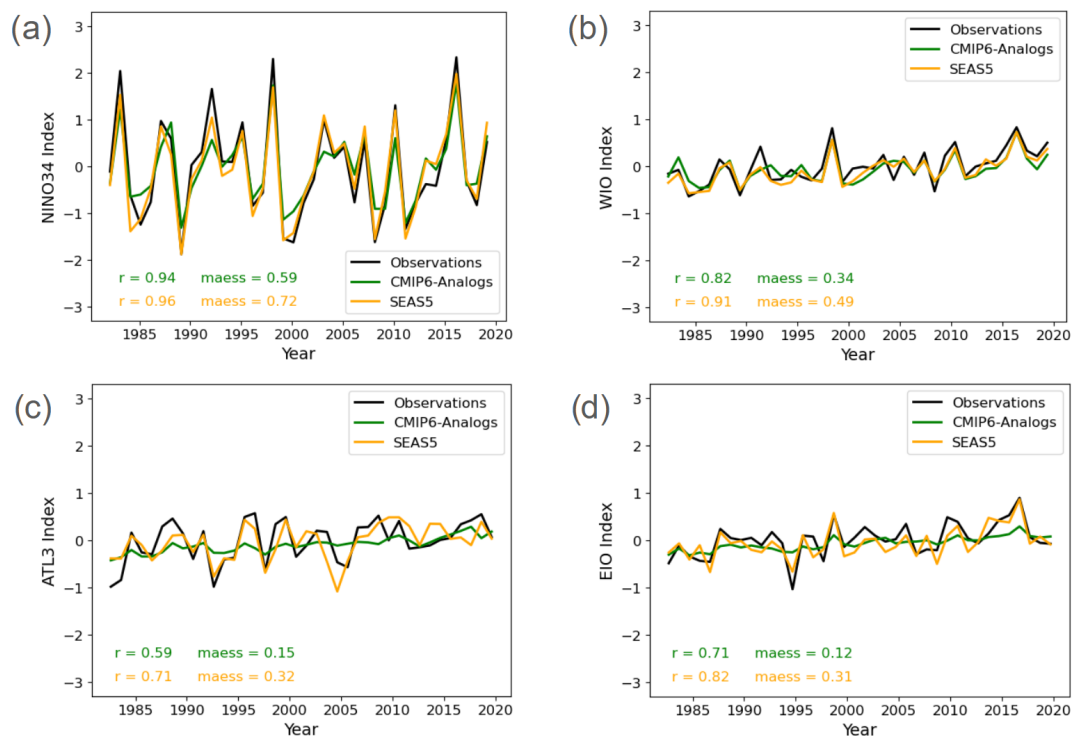


Figure 5. Observed and predicted evolution of the indices estimated as the area-averaged TAS time series of the (a) December–February NINO34 (170–120° W, 5° S–5° N; Fang and Xie, 2020) in the tropical Pacific Ocean, (b) the March–May WIO (50–70° E, 10° S–10° N; Saji et al., 1999) in the tropical western Indian Ocean, (c) the June–August ATL3 (20° W–0° E, 3° S–3° N; Zebiak, 1993) in the tropical Atlantic Ocean, and the September–November EIO (90–110° E, 10° S–0° N; Saji et al., 1999) in the tropical eastern Indian Ocean.

around the forced signal (positive ACC), but due to biases in the analog-based predictions, MAESS may be affected to the point of making it zero or negative. There is also a broad similarity in the spatial distribution of skill (correlation) of the analog-based and the operational decadal predictions from EC-Earth3 (Fig. 6a vs. 6c). The analog-based pre-

dictions slightly outperform EC-Earth3 over land according to MAESS, as seen in southern Africa or Australia (Fig. 6f)

When extending the forecast to a 2-year period, the analog-based predictions of TAS continue to show high skill across most land areas. The skill in the extratropical regions such as the Mediterranean or East Asia is comparable with

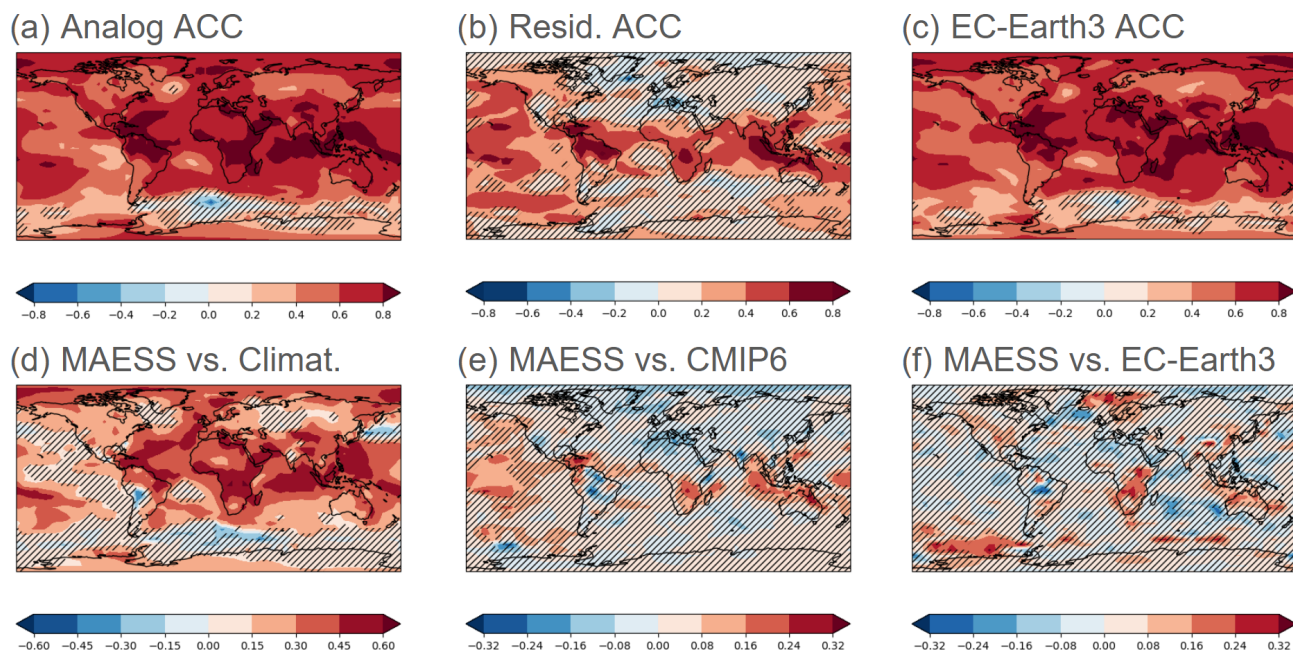


Figure 6. The same as Fig. 1 but for annual TAS predictions (January–December). The evaluation period is 1962–2018. The predictions are initialized each November in both analog-based predictions and EC-Earth3. The dynamic forecast system evaluated in (c) and (f) is EC-Earth3.

the skill in the tropical zones, as shown by Fig. 7a and d. The residual skill after removing the forced signal is slightly less pronounced for forecasts spanning 2 years than for 1-year forecasts, as shown in Figs. 7b and 6b, respectively. The benefit from initialization (residual correlation) of the analog-based predictions can still be observed in several areas like tropical South America, South Asia, Australia, and sub-Saharan Africa. Most of the Pacific and the Indian oceans also show benefits from initialization in the analog-based predictions. Subtropical regions tend to show reduced skill in the TAS analog-based predictions according to the MAESS (Fig. 7e), with the Mediterranean region under-performing CMIP6 but still generally outperforming EC-Earth3 predictions in northern South America, sub-Saharan Africa, and Australia (Fig. 7f). As for annual predictions, possible biases present in the analog-based biennial predictions are likely behind the little to no advantage of the analog-based predictions over CMIP6 ensemble based on MAESS (Fig. 7e), despite some clear advantages measured by the residual correlation (Fig. 7b), which only estimates the variability around the forced signal.

The results for quadrennial predictions of TAS show higher overall ACC and MAESS than the biennial or annual predictions (Fig. 8a, c). However, residual correlation is generally smaller in the quadrennial predictions (e.g. the tropical Atlantic and western Africa no longer show added skill). This implies that despite higher overall ACC, the benefit from initialization is smaller for quadrennial predictions than for biennial and annual. Furthermore, the analog pre-

dictions clearly underperform CMIP6 when measured by MAESS (Fig. 8e) as opposed to the overall clear advantage measured by residual ACC (Fig. 8b). This is especially clear in the Northern Hemisphere subtropics and again most likely the result of biases. The added value of the analog prediction over EC-Earth3 is however still visible in many regions, with the analog-based predictions under-performing EC-Earth3 predictions only in the North Atlantic, northern Africa, and southwestern Asia but outperforming it in northern South America, Sub-Saharan Africa, South and South-east Asia, Australia, and parts of Europe (Fig. 8f).

The correlation maps of SPI12 (Fig. 9a, c), SPI24 (Fig. 9d, f), and SPI48 (Fig. 9g, i) predictions using the analog method and EC-Earth3 are again spatially similar. There is an increase in correlation in Northern Hemisphere high latitudes with longer precipitation accumulations (SPI48) but comparable skill elsewhere in SPI12, 24, and 48, except for a few regions such as southern Africa or western North America which exhibit lower skill at longer accumulations. An important fraction of the skill for SPI24 and especially SPI12 predictions stems from the synchronization of unforced variability in the models and observations, similar to the seasonal predictions. This can be implied by the broad similarities between Fig. 9a and d and b and e. Contrastingly, for SPI48 predictions, the forced signal largely dominates over the non-forced one. MAESS maps of SPI12, 24, and 48 reveal mostly non-significant values when comparing analog-based predictions with both CMIP6 and EC-Earth3 (not shown).

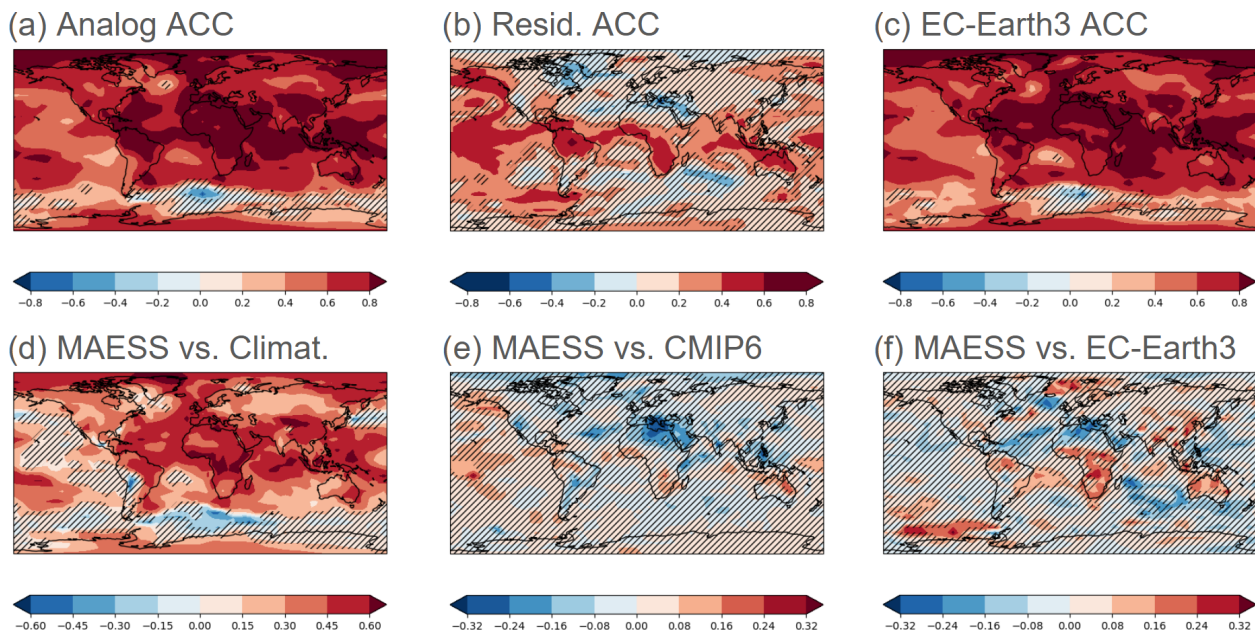


Figure 7. The same as Fig. 1 but for biennial TAS predictions (January–December + 1 year). The evaluation period is 1962–2018. The predictions are initialized each November in both analog-based predictions and EC-Earth3. The dynamic forecast system evaluated in (c) and (f) is EC-Earth3.

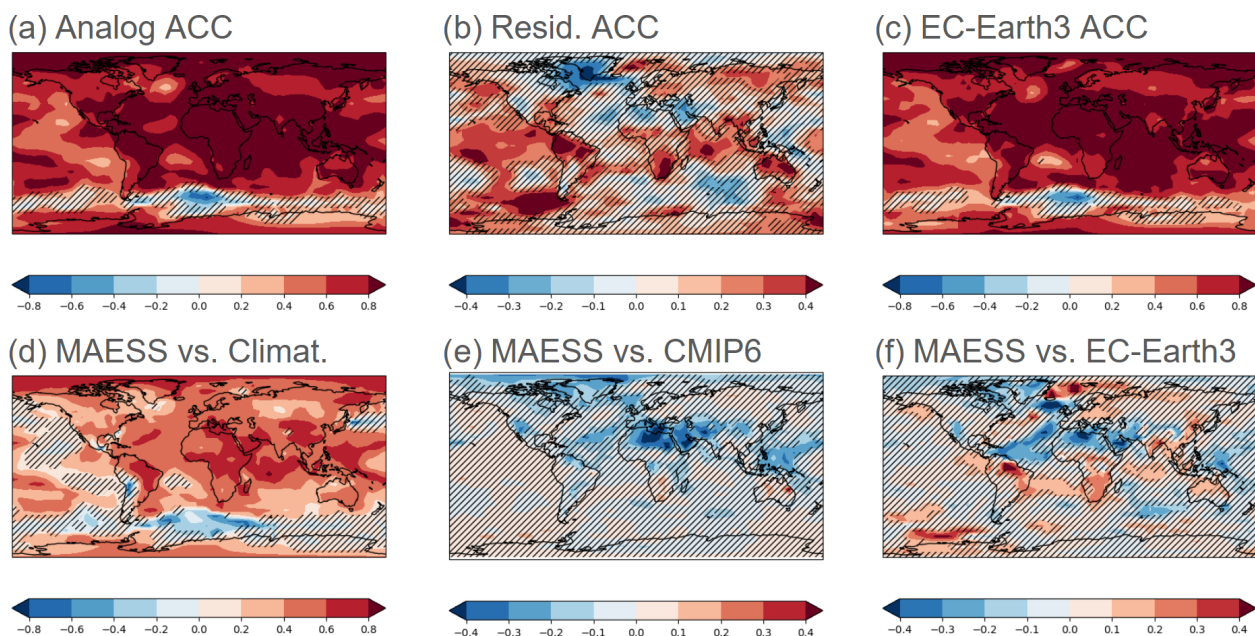


Figure 8. The same as Fig. 1 but for quadrennial TAS predictions (January–December + 3 years). The evaluation period is 1962–2018. The predictions are initialized each November in both analog-based predictions and EC-Earth3. The dynamic forecast system evaluated in (c) and (f) is EC-Earth3.

Figure 10 presents the time series for annual (a, b), biennial (c), and quadrennial (d) averages of global TAS in the analog-based predictions, the EC-Earth3 predictions, and the observations. Annual predictions initialized each November for the following January–December (Fig. 10a) show very

similar results, with comparable performance metrics for the analog and EC-Earth3 predictions. When considering the 12-month period from July–June (Fig. 10b), the analog-based predictions initialized in June (1-month lag) are superior to the EC-Earth initialized the previous November (8-month

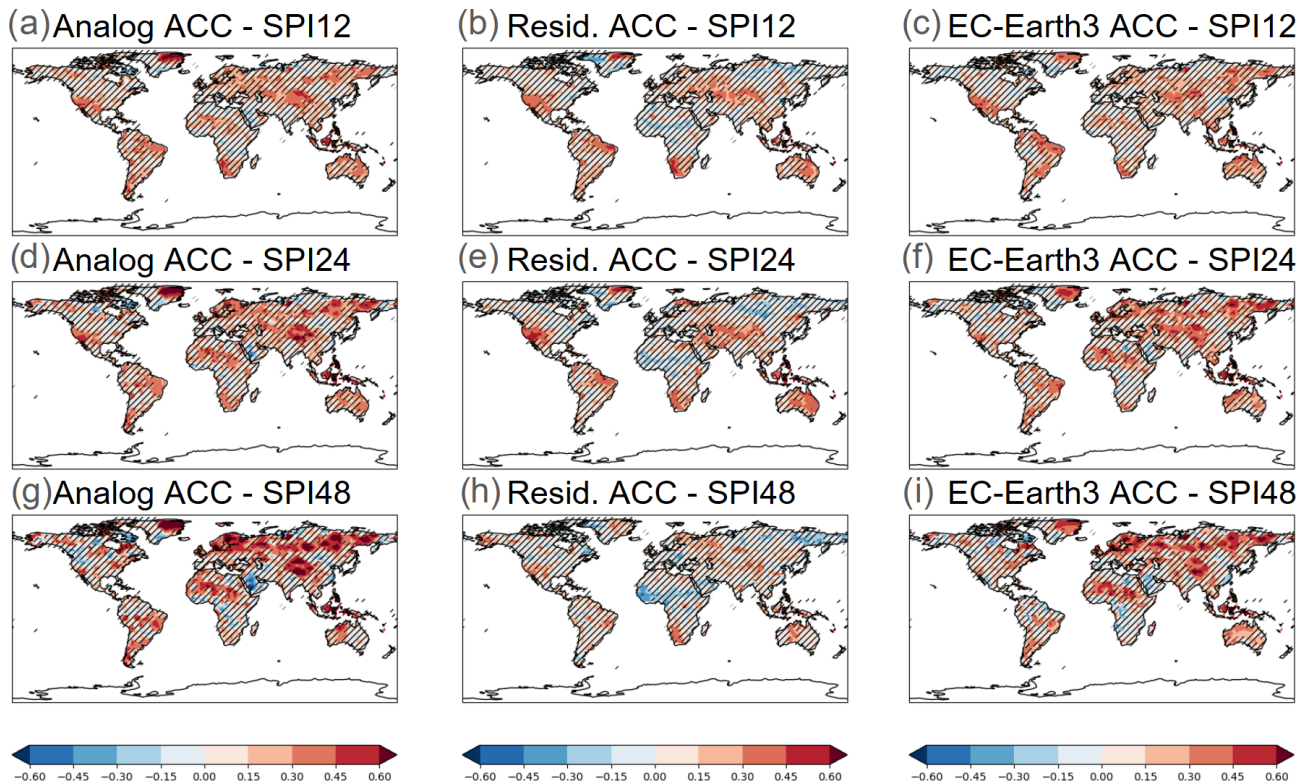


Figure 9. The same as Fig. 3a–c but for (a–c) SPI12, (d–f) SPI24, and (g–i) SPI48 predictions initialized every November. The reference period is 1962–2018. The dynamic forecast system evaluated in (c), (f), and (i) is EC-Earth3.

lag), showing residual correlations of 0.74 and 0.26 for the analog-based and EC-Earth3 predictions, respectively. This example highlights a key advantage of the analog-based predictions over the dynamical prediction systems, as the analog ones can be produced every month throughout the year without large computational cost as opposed to the dynamical ones. Biennial and quadrennial predictions (Fig. 10c, d), initialized each November (1-month lead), show very similar values of correlation for both analog and EC-Earth3 predictions, with residual correlations higher in analog-based predictions and MAESS slightly higher in EC-Earth3 predictions.

Figure 11 summarizes the results of the analog and EC-Earth3 predictions in terms of total land area fraction with significant positive correlation for annual, biennial, and quadrennial predictions of TAS and SPI. Due to a saturation of correlation with observations, the residual correlation after removing CMIP6 signal is used for TAS, while the anomaly correlation of the actual time series is used for SPI. The analog-based and EC-Earth3 predictions initialized every November have a comparable skill of predicting 12-month mean TAS and SPI (Fig. 11a, d) that decreases with increasing lead time from 2 months up to 13 months (dark-green and yellow lines). However, when using 12-month analog-based TAS and SPI predictions initialized every month always with 1-month lead time, the skill is broadly superior

to the benchmark (light-green line vs. yellow lines), and the skill increases when the forecasts are initialized after the spring ENSO predictability barrier. For 24-month forecasts, the November-initialized analog-based predictions are also slightly superior to EC-Earth for TAS (Fig. 11b, e), and the analog-based predictions initialized every month are consistently superior to the November initialized ones, similar to the 12-month forecasts. Analog-based predictions of 48-month TAS and SPI initialized every November are less skillful than the EC-Earth3 counterpart, while the analog-based predictions initialized every month have a similar skill closer to EC-Earth3 on average but exhibit more variability throughout the year depending on the month of initialization and variable predicted (Fig. 11c, f).

4 Summary and conclusions

The analog-based predictions provide skillful forecasts on seasonal to multi-annual timescales and show in general similar spatial patterns of skill to initialized numerical predictions. Furthermore, the analog-based predictions are competitive with existing annual and multi-annual predictions from initialized numerical predictions. On seasonal timescales the analog-based predictions demonstrate high skill for boreal winter and summer TAS forecasts with 1-month lead time, particularly in the tropics, North Atlantic, and most of the

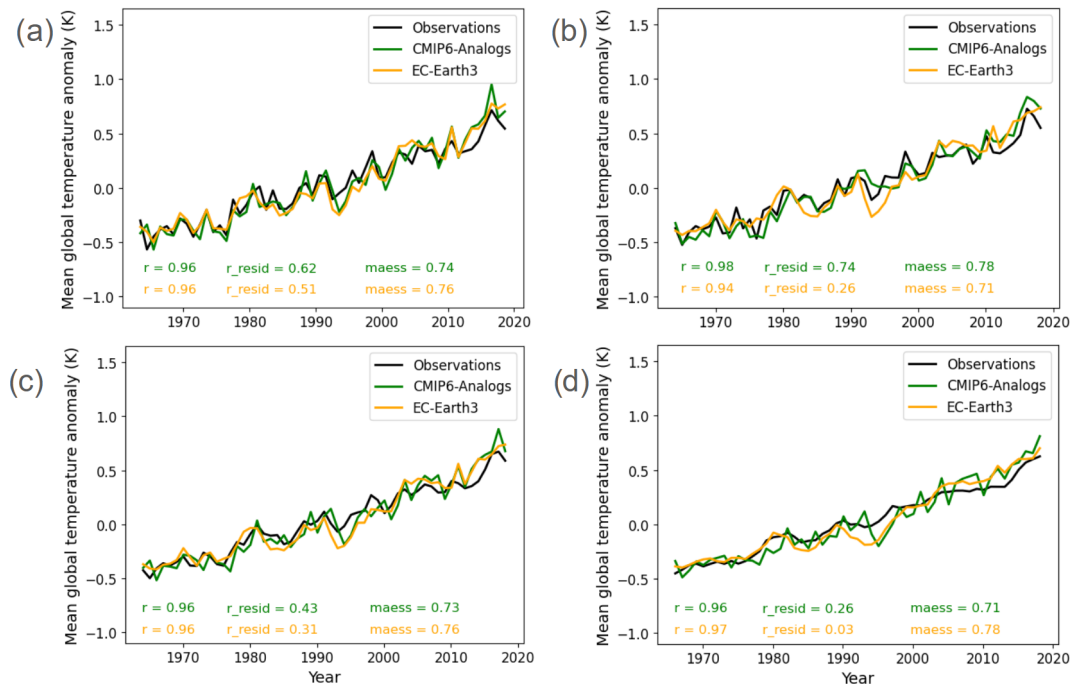


Figure 10. Observed and predicted evolution of the global area-averaged TAS time series of (a) annual January–December predictions initialized in November for both the analog-based method and EC-Earth3 (2-month lead), (b) the annual July–June (+1 year) predictions initialized in November (8-month lead) for EC-Earth3 and in June (1-month lead) for the analog-based method, (c) the biannual January–December (+1 year) predictions initialized in November for both the analog-based method and EC-Earth3 (2-month lead), and (d) the quadrennial January–December (+3 years) predictions initialized in November for both the analog-based method and EC-Earth3 (2-month lead).

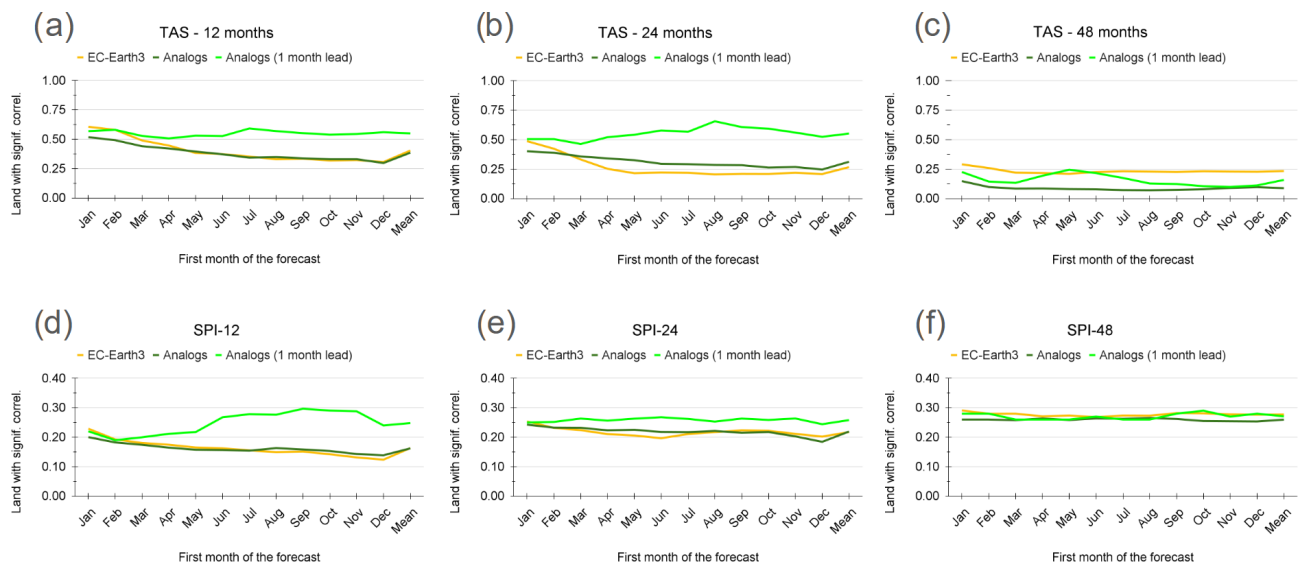


Figure 11. Fraction of global land area with statistically significant ($p < 0.1$) positive residual correlation between TAS predictions and observations for (a) 12-month, (b) 24-month, and (c) 48-month forecasts. Panels (d), (e), and (f) are the same as (a), (b), and (c), respectively, but for the statistically significant ($p < 0.1$) positive correlation between SPI predictions and observations using a two-sided t test. The dark-green and yellow lines in all panels show the skill of the analog-based and EC-Earth3 predictions, respectively, initialized every November as the lead time increases from 2 to 13 months. The light-green line shows the skill of the analog-based predictions, initialized always with a 1-month lead time. The x axis always shows the first month of the forecasted period evaluated. For example, in panels (b) and (e) the values of August indicate the skill for predictions between August in the first forecast year and July 2 years later.

Arctic, offering substantial added value over the CMIP6 ensemble (e.g. residual skill). As for boreal summer, the skill extends into subtropical and mid-latitude regions, with the Northern Hemisphere land showing greater skill in summer than winter. The improvement over the non-initialized CMIP6 ensemble is less pronounced in the tropical Pacific during summer, likely due to the peak activity of the El Niño Southern Oscillation (ENSO) in winter. Like climate predictions from dynamical forecasting systems, analog SPI3 forecasts are generally less skillful than 3-month TAS predictions but still show higher skill than the non-initialized CMIP6 ensemble and skill peaks around boreal winter. We show that skill in SPI3 predictions primarily stems from internal climate variability alignment, while for TAS predictions, external forcing also plays an important role. Seasonal TAS and SPI3 predictions for all initializations throughout the year display clear added value over the non-initialized CMIP6 ensemble but are generally less skillful than operational predictions from SEAS51 (Johnson et al., 2019). Furthermore, the spatial patterns of skill are very similar between the analog-based predictions and the state-of-the-art benchmark prediction system SEAS51, suggesting that both predictions have skill due to similar physical processes.

On annual to multi-annual timescales, the annual TAS analog-based predictions are highly skillful across most tropical and many extratropical land regions. Central Asia, central North America, and southern South America show lower skill but still better than climatological forecasts. The analog-based predictions generally outperform the CMIP6 ensemble, while the added value over the CMIP6 ensemble decreases with increasing forecast range (i.e. biennial and quadrennial), indicating that external forcing drives most of the skill particularly at quadrennial timescales. Spatially, the skill of annual and biennial SPI forecasts is generally similar to that of seasonal ones, with positive statistically significant correlations in several tropical regions being a common feature. High-latitude regions in Eurasia exhibit enhanced skill, particularly for quadrennial predictions, with external forcing contributing significantly to the skill in these areas. A comparison with the operational decadal prediction system EC-Earth3 (Bilbao et al., 2021) reveals that the analog method can provide comparable annual and biennial predictions of TAS and SPI when the predictions are initialized at the same month (i.e. every November) and the lead time increases. While decadal prediction systems are typically initialized only once per year, the analog-based predictions can however be easily generated every month in an operational context, and the skill of those predictions is broadly superior to the skill of the EC-Earth3 decadal predictions initialized only once a year. The 48-month analog-based predictions of TAS and SPI are less skillful than the EC-Earth3 counterpart when initialized in November but become comparable if the analog-based predictions are produced every month. We have chosen EC-Earth3 as a representative model of the typical decadal prediction system. It is possible that other

decadal prediction systems perform better in particular regions and for particular timescales, but EC-Earth3 forecasts quality metrics reveal it to be a good representative of these systems. For a thorough evaluation of several decadal prediction systems including EC-Earth3, the reader is referred to Figs. S7–S11 in Delgado-Torres et al. (2022).

Building on the established concept of climate analogs, our research demonstrates that by sampling through time and model of a large CMIP6 multi-model ensemble based on their similarity with observed SST patterns, one can extract valuable information on the future evolution of TAS and SPI, spanning a forecast range of seasons to multiple years. In other words, the analog-based forecasts can provide seamless predictions for different forecast times, which have traditionally been addressed with specific forecast systems (seasonal or decadal). Despite some potential limitations related to the lack of a more sophisticated model initialization, these analog-based forecasts have no initialization shock nor drift and are competitive with the existing prediction systems on annual to multi-annual forecast ranges. This methodology offers a complementary source of climate information to existing seasonal and decadal climate predictions, filling an existing gap across timescales and doing so in a seamless manner. Crucially, the method is computationally inexpensive and based on a straightforward approach that facilitates the generation of seamless climate predictions reproducible at low computational cost once the multi-model ensemble of transient simulations has been produced.

Code and data availability. The CMIP6 and EC-Earth simulation data that support the findings of this study are available at <https://esgf-node.llnl.gov/search/cmip6/> (last access: 14 June 2024), SEAS51 data are available at <https://cds.climate.copernicus.eu/> (last access: 14 June 2024), BEST data are available at <https://climatedataguide.ucar.edu/climate-data/global-surface-temperatures-best-berkeley-earth-surface-temperatures> (last access: 14 June 2024) (Cowtan and National Center for Atmospheric Research Staff), GPCC data are available at <https://psl.noaa.gov/data/gridded/data.gpcc.html> (last access: 14 June 2024), and ERSSTv5 data are available at <https://www.psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html> (last access: 14 June 2024).

Supplement. The supplement related to this article is available online at <https://doi.org/10.5194/esd-16-1723-2025-supplement>.

Author contributions. JCAN, AT, and AHE conceived the idea for this study. JCAN conducted the analysis and prepared all figures. PDL, MGD, RM, DV, and AA contributed with the preparation of CMIP6 and EC-Earth3 data. All authors were actively involved in the analysis of the results and the writing process.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Financial support. This research has been supported by EU HORIZON EUROPE Climate, Energy and Mobility (grant no. 101081460).

Review statement. This paper was edited by Yun Liu and reviewed by three anonymous referees.

References

- Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, *Earth Syst. Sci. Data*, 5, 71–99, <https://doi.org/10.5194/essd-5-71-2013>, 2013.
- Beguiría S. and Vicente-Serrano, S. M.: SPEI: Calculation of the Standardized Precipitation – Evapotranspiration Index, <https://github.com/sbegueria/SPEI> (last access: June 2024), 2023.
- Bett, P. E., Thornton, H. E., Troccoli, A., De Felice, M., Suckling, E., Dubus, L., Saint-Drenan, Y.-M., and Brayshaw, D. J.: A simplified seasonal forecasting strategy, applied to wind and solar power in Europe, *Climate Services*, 27, 100318, <https://doi.org/10.1016/j.cliser.2022.100318>, 2022.
- Bilbao, R., Wild, S., Ortega, P., Acosta-Navarro, J., Arsouze, T., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., Cruz-García, R., Cvijanovic, I., Doblas-Reyes, F. J., Donat, M., Dutra, E., Echevarría, P., Ho, A.-C., Loosveldt-Tomas, S., Moreno-Chamarro, E., Pérez-Zanon, N., Ramos, A., Ruprich-Robert, Y., Sicardi, V., Tourigny, E., and Vegas-Regidor, J.: Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth, *Earth Syst. Dynam.*, 12, 173–196, <https://doi.org/10.5194/esd-12-173-2021>, 2021.
- Buontempo, C., Hanlon, H. M., Soares, M. B., Christel, I., Soubeyroux, J.-M., Viel, C., Calmanti, S., Bosi, L., Falloon, P., Palin, E. J., Vanvyve, E., Torralba, V., Gonzalez-Reviriego, N., Doblas-Reyes, F., Pope, E. C. D., Newton, P., and Liggins, F.: What have we learnt from EUPORIAS climate service prototypes?, *Climate Services*, 9, 21–32, <https://doi.org/10.1016/j.cliser.2017.06.003>, 2018.
- Caron, L., Hermanson, L., and Doblas-Reyes, F. J.: Multiannual forecasts of Atlantic U.S. tropical cyclone wind damage potential, *Geophys. Res. Lett.*, 42, 2417–2425, <https://doi.org/10.1002/2015gl063303>, 2015.
- Ceglar, A. and Toreti, A.: Seasonal climate forecast can inform the European agricultural sector well in advance of harvesting, *Npj Climate and Atmospheric Science*, 4, 42, <https://doi.org/10.1038/s41612-021-00198-3>, 2021.
- Cowan, K. and National Center for Atmospheric Research Staff (Eds.): The Climate Data Guide: Global surface temperatures: BEST: Berkeley Earth Surface Temperatures, NCAR [data set], <https://climatedataguide.ucar.edu/climate-data/global-surface-temperatures-best-berkeley-earth-surface-temperatures>, last access: 14 June 2024.
- Delgado-Torres, C., Donat, M. G., Gonzalez-Reviriego, N., Caron, L., Athanasiadis, P. J., Bretonnière, P., Dunstone, N. J., Ho, A., Nicoli, D., Pankatz, K., Paxian, A., Pérez-Zanón, N., Cabré, M. S., Solaraju-Murali, B., Soret, A., and Doblas-Reyes, F. J.: Multi-Model Forecast Quality Assessment of CMIP6 Decadal Predictions, *J. Climate*, 35, 4363–4382, <https://doi.org/10.1175/JCLI-D-21-0811.1>, 2022.
- De Luca, P., Delgado-Torres, C., Mahmood, R., Samso-Cabre, M., and Donat, M. G.: Constraining decadal variability regionally improves near-term projections of hot, cold and dry extremes, *Environ. Res. Lett.*, 18, 094054, <https://doi.org/10.1088/1748-9326/acf389>, 2023.
- Ding, H., Newman, M., Alexander, M. A., and Wittenberg, A. T.: Skillful climate forecasts of the tropical Indo-Pacific Ocean using Model-Analogs, *J. Climate*, 31, 5437–5459, <https://doi.org/10.1175/jcli-d-17-0661.1>, 2018.
- Ding, H., Newman, M., Alexander, M. A., and Wittenberg, A. T.: Diagnosing secular variations in retrospective ENSO seasonal forecast skill using CMIP5 Model-Analogs, *Geophys. Res. Lett.*, 46, 1721–1730, <https://doi.org/10.1029/2018gl080598>, 2019.
- Donat, M. G., Mahmood, R., Cos, P., Ortega, P., and Doblas-Reyes, F.: Improving the forecast quality of near-term climate projections by constraining internal variability based on decadal predictions and observations, *Environmental Research Climate*, 3, 035013, <https://doi.org/10.1088/2752-5295/ad5463>, 2024.
- Dunstone, N., Lockwood, J., Solaraju-Murali, B., Reinhardt, K., Tsartsali, E. E., Athanasiadis, P. J., Bellucci, A., Brookshaw, A., Caron, L.-P., Doblas-Reyes, F. J., Früh, B., González-Reviriego, N., Gualdi, S., Hermanson, L., Materia, S., Nicodemou, A., Nicoli, D., Pankatz, K., Paxian, A., Scaife, A., Smith, D., and Thornton, H. E.: Towards useful decadal climate services, *B. Am. Meteorol. Soc.*, 103, E1705–E1719, <https://doi.org/10.1175/bams-d-21-0190.1>, 2022.
- 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, *Geosci. Model Dev.*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Fang, X. and Xie, R.: A brief review of ENSO theories and prediction, *Science China Earth Sciences*, 63, 476–491, <https://doi.org/10.1007/s11430-019-9539-0>, 2020.
- Huang, B., Banzon, V. F., Freeman, E., Lawrimore, J., Liu, W., Peterson, T. C., Smith, T. M., Thorne, P. W., Woodruff, S. D., and Zhang, H. M.: NOAA extended reconstructed sea surface temperature (ERSST), version 5, NOAA National Centers for Environmental Information, 30, 8179–8205, <https://doi.org/10.7289/V5T72FNM> (last access: June 2024), 2017.

- Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremier, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., and Monge-Sanz, B. M.: SEAS5: the new ECMWF seasonal forecast system, *Geosci. Model Dev.*, 12, 1087–1117, <https://doi.org/10.5194/gmd-12-1087-2019>, 2019.
- Kim, H., Richter, J. H., and Martin, Z.: Insignificant QBO-MJO prediction skill relationship in the SubX and S2S subseasonal reforecasts, *J. Geophys. Res.-Atmos.*, 124, 12655–12666, <https://doi.org/10.1029/2019jd031416>, 2019a.
- Kim, Y., Ratnam, J. V., Doi, T., Morioka, Y., Behera, S., Tsuzuki, A., Minakawa, N., Sweijid, N., Kruger, P., Maharaj, R., Imai, C. C., Ng, C. F. S., Chung, Y., and Hashizume, M.: Malaria predictions based on seasonal climate forecasts in South Africa: A time series distributed lag nonlinear model, *Sci. Rep.*, 9, 17882, <https://doi.org/10.1038/s41598-019-53838-3>, 2019b.
- Kirtman, B., Anderson, D., Brunet, G., Kang, I.-S., Scaife, A. A., and Smith, D.: Prediction from Weeks to Decades, in: Springer eBooks, 205–235, https://doi.org/10.1007/978-94-007-6692-1_8, 2013.
- Klemm, T. and McPherson, R. A.: The development of seasonal climate forecasting for agricultural producers, *Agric. Forest Meteorol.*, 232, 384–399, <https://doi.org/10.1016/j.agrformet.2016.09.005>, 2017.
- Kushnir, Y., Scaife, A. A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., Hawkins, E., Kimoto, M., Kolli, R. K., Kumar, A., Matei, D., Matthes, K., Müller, W. A., O’Kane, T., Perlwitz, J., Power, S., Raphael, M., Shimp, A., Smith, D., Tuma, M., and Wu, B.: Towards operational predictions of the near-term climate, *Nat. Clim. Change*, 9, 94–101, <https://doi.org/10.1038/s41558-018-0359-7>, 2019.
- Liu, Z. and Di Lorenzo, E.: Mechanisms and Predictability of Pacific Decadal variability, *Current Climate Change Reports*, 4, 128–144, <https://doi.org/10.1007/s40641-018-0090-5>, 2018.
- Lledó, L., Torralba, V., Soret, A., Ramon, J., and Doblas-Reyes, F. J.: Seasonal forecasts of wind power generation, *Renewable Energy*, 143, 91–100, <https://doi.org/10.1016/j.renene.2019.04.135>, 2019.
- Lopez, H. and Kirtman, B. P.: WWBs, ENSO predictability, the spring barrier and extreme events, *J. Geophys. Res.-Atmos.*, 119, 10114–10138, <https://doi.org/10.1002/2014jd021908>, 2014.
- Mahmood, R., Donat, M. G., Ortega, P., Doblas-Reyes, F. J., and Ruprich-Robert, Y.: Constraining decadal variability yields skillful projections of Near-Term climate change, *Geophys. Res. Lett.*, 48, e2021GL094915, <https://doi.org/10.1029/2021gl094915>, 2021.
- Mahmood, R., Donat, M. G., Ortega, P., Doblas-Reyes, F. J., Delgado-Torres, C., Samsó, M., and Bretonnière, P.-A.: Constraining low-frequency variability in climate projections to predict climate on decadal to multi-decadal timescales – a poor man’s initialized prediction system, *Earth Syst. Dynam.*, 13, 1437–1450, <https://doi.org/10.5194/esd-13-1437-2022>, 2022.
- Mann, M. E., Steinman, B. A., and Miller, S. K.: On forced temperature changes, internal variability, and the AMO, *Geophys. Res. Lett.*, 41, 3211–3219, <https://doi.org/10.1002/2014gl059233>, 2014.
- McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration of time scales, Eighth Conference on Applied Climatology, American Meteorological Society, 17–23 January 1993, Anaheim CA, 179–186, 1993.
- Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblas-Reyes, F., Donat, M. G., England, M. H., Fyfe, J. C., Han, W., Kim, H., Kirtman, B. P., Kushnir, Y., Lovenduski, N. S., Mann, M. E., Merryfield, W. J., Nieves, V., Pegion, K., Rosenbloom, N., Sanchez, S. C., Scaife, A. A., Smith, D., Subramanian, A. C., Sun, L., Thompson, D., Ummenhofer, C. C., and Xie, S.-P.: Initialized Earth System prediction from subseasonal to decadal timescales, *Nat. Rev. Earth Environ.*, 2, 340–357, <https://doi.org/10.1038/s43017-021-00155-x>, 2021.
- Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A. S., Danabasoglu, G., Dirmeyer, P. A., Doblas-Reyes, F. J., Domeisen, D. I. V., Ferranti, L., Ilynia, T., Kumar, A., Müller, W. A., Rixen, M., Robertson, A. W., Smith, D. M., Takaya, Y., Tuma, M., Vitart, F., White, C. J., Alvarez, M. S., Ardilouze, C., Attard, H., Baggett, C., Balmaseda, M. A., Beraki, A. F., Bhattacharjee, P. S., Bilbao, R., De Andrade, F. M., DeFlorio, M. J., Díaz, L. B., Ehsan, M. A., Fragkoulidis, G., Gonzalez, A. O., Grainger, S., Green, B. W., Hell, M. C., Infanti, J. M., Isensee, K., Kataoka, T., Kirtman, B. P., Klingaman, N. P., Lee, J.-Y., Mayer, K., McKay, R., Mecking, J. V., Miller, D. E., Neddermann, N., Ng, C. H. J., Ossó, A., Pankatz, K., Peatman, S., Pegion, K., Perlwitz, J., Recalde-Coronel, G. C., Reintges, A., Renkl, C., Solaraju-Murali, B., Spring, A., Stan, C., Sun, Y. Q., Tozer, C. R., Vigaud, N., Woolnough, S., and Yeager, S.: Current and emerging developments in subseasonal to decadal prediction, *B. Am. Meteorol. Soc.*, 101, E869–E896, <https://doi.org/10.1175/bams-d-19-0037.1>, 2020.
- Ndiaye, C. D., Mohino, E., Mignot, J., and Sall, S. M.: On the Detection of Externally Forced Decadal Modulations of the Sahel Rainfall over the Whole Twentieth Century in the CMIP6 Ensemble, *J. Climate*, 35, 6939–6954, <https://doi.org/10.1175/JCLI-D-21-0585.1>, 2022.
- Riddle, E. E., Butler, A. H., Furtado, J. C., Cohen, J. L., and Kumar, A.: CFSv2 ensemble prediction of the wintertime Arctic Oscillation, *Clim. Dynam.*, 41, 1099–1116, <https://doi.org/10.1007/s00382-013-1850-5>, 2013.
- Saji, N. H., Goswami, B. N., Vinayachandran, P. N., and Yamagata, T.: A dipole mode in the tropical Indian Ocean, *Nature*, 401, 360–363, <https://doi.org/10.1038/43854>, 1999.
- Sánchez-García, E., Rodríguez-Camino, E., Bacciu, V., Chiarle, M., Costa-Saura, J., Garrido, M. N., Lledó, L., Navascués, B., Paranzunzio, R., Terzago, S., Bongiovanni, G., Mereu, V., Nigrelli, G., Santini, M., Soret, A., and Von Hardenberg, J.: Co-design of sectoral climate services based on seasonal prediction information in the Mediterranean, *Climate Services*, 28, 100337, <https://doi.org/10.1016/j.cliser.2022.100337>, 2022.
- Schindler, A., Toreti, A., Zampieri, M., Scoccimarro, E., Gualdi, S., Fukutome, S., Xoplaki, E., and Luterbacher, J.: On the Internal Variability of Simulated Daily Precipitation, *J. Climate*, 28, 3624–3630, <https://doi.org/10.1175/jcli-d-14-00745.1>, 2015.
- Seiour, W. J. M., Hardiman, S. C., Gray, L. J., Butchart, N., MacLachlan, C., and Scaife, A. A.: Skillful seasonal prediction of the southern annular mode and Antarctic ozone, *J. Climate*, 27, 7462–7474, <https://doi.org/10.1175/jcli-d-14-00264.1>, 2014.
- Shinoda, T. and Han, W.: Influence of the Indian Ocean Dipole on atmospheric subseasonal variability, *J. Climate*, 18, 3891–3909, <https://doi.org/10.1175/jcli3510.1>, 2005.

- Sigmond, M., Scinocca, J. F., Kharin, V. V., and Shepherd, T. G.: Enhanced seasonal forecast skill following stratospheric sudden warmings, *Nat. Geosci.*, 6, 98–102, <https://doi.org/10.1038/ngeo1698>, 2013.
- Smith, D. M., Eade, R., Scaife, A. A., Caron, L.-P., Danabasoglu, G., DelSole, T. M., Delworth, T., Doblas-Reyes, F. J., Dunstone, N. J., Hermanson, L., Kharin, V., Kimoto, M., Merryfield, W. J., Mochizuki, T., Müller, W. A., Pohlmann, H., Yeager, S., and Yang, X.: Robust skill of decadal climate predictions, *Npj Climate and Atmospheric Science*, 2, 13, <https://doi.org/10.1038/s41612-019-0071-y>, 2019.
- Solaraju-Murali, B., Gonzalez-Reviriego, N., Caron, L.-P., Ceglar, A., Toreti, A., Zampieri, M., Bretonnière, P.-A., Cabré, M. S., and Doblas-Reyes, F. J.: Multi-annual prediction of drought and heat stress to support decision making in the wheat sector, *Npj Climate and Atmospheric Science*, 4, 34, <https://doi.org/10.1038/s41612-021-00189-4>, 2021.
- Solaraju-Murali, B., Bojovic, D., Gonzalez-Reviriego, N., Nicodemou, A., Terrado, M., Caron, L.-P., and Doblas-Reyes, F. J.: How decadal predictions entered the climate services arena: an example from the agriculture sector, *Climate Services*, 27, 100303, <https://doi.org/10.1016/j.cliser.2022.100303>, 2022.
- Thomson, M. C., Doblas-Reyes, F. J., Mason, S. J., Hagedorn, R., Connor, S. J., Phindela, T., Morse, A. P., and Palmer, T. N.: Malaria early warnings based on seasonal climate forecasts from multi-model ensembles, *Nature*, 439, 576–579, <https://doi.org/10.1038/nature04503>, 2006.
- Zebiak, S. E.: Air–Sea interaction in the Equatorial Atlantic region, *J. Climate*, 6, 1567–1586, 1993.
- Zhang, C., Adames, Á. F., Khouider, B., Wang, B., and Yang, D.: Four Theories of the Madden-Julian Oscillation, *Rev. Geophys.*, 58, e2019RG000685, <https://doi.org/10.1029/2019rg000685>, 2020.