ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Improving seasonal predictions of meteorological drought by conditioning on ENSO states

To cite this article: Patrick Pieper *et al* 2021 *Environ. Res. Lett.* **16** 094027

IOP Publishing Environ. Res. Lett. 16 (2021) 094027 https://doi.org/10.1088/1748-9326/ac1cbb

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
10 February 2021

REVISED
9 August 2021

ACCEPTED FOR PUBLICATION
11 August 2021

PUBLISHED
27 August 2021

LETTER

Improving seasonal predictions of meteorological drought by conditioning on ENSO states

Patrick Pieper[1,*], **André Düsterhus**[2] and **Johanna Baehr**[1]

[1] Institute for Oceanography, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Hamburg, Germany
[2] Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography, Maynooth University, Maynooth, Ireland
* Author to whom any correspondence should be addressed.

E-mail: patrick.pieper@uni-hamburg.de

Keywords: seasonal prediction, meteorological drought, El Niño-Southern Oscillation (ENSO) conditioning, standardized precipitation index (SPI)

Abstract

Useful hindcast skill of meteorological drought, assessed with the 3-month standardized precipitation index (SPI_{3M}), has been so far limited to one lead month (time horizon of the prediction). Here, we quadruple that lead time by demonstrating useful skill up to lead month 4. To obtain useful hindcast skill of meteorological drought at these long lead times, we exploit well-known El Niño-Southern Oscillation (ENSO)–precipitation teleconnections through ENSO-state conditioning. We condition initialized seasonal SPI_{3M} hindcasts, derived from the Max-Planck-Institute Earth System Model (MPI-ESM) over the period 1982–2013, on ENSO states by exploring significant agreements between two complementary analyses: hindcast skill ENSO–composites, and observed ENSO–precipitation correlations. Such conditioned hindcast skill of meteorological drought is in MPI-ESM significant and reliable for lead months 2 to 4 in equatorial South America and southern North America during these regions' dry ENSO phases. When a region's dry ENSO phase is present at the initialization in autumn (ASO), predictions of meteorological drought show useful hindcast skill for the upcoming winter (DJF) in the respective region. The area of this useful hindcast skill is further enlarged in both regions when the respective region's dry ENSO phase is already present in the antecedent summer (conditioning on ENSO states in JJA). Active ENSO events constitute windows of opportunity for drought predictions that are insufficiently covered by typical predictability analyses. For these windows, we demonstrate predictive skill at unprecedented lead times with a single model whose output is not bias corrected. This contribution exemplifies the value of ENSO-state conditioning in identifying these windows of opportunity for regions that are arguably most affected by ENSO–precipitation teleconnections. During these regions' dry ENSO phases, reliable predictive skill of meteorological drought is at long lead times particularly valuable and moves the frontier of meteorological drought predictions.

1. Introduction

Reliable seasonal meteorological drought predictions can alleviate the harm caused by droughts through timely and accurate warnings, resulting in increased preparedness. However, the time horizon of reliable meteorological drought predictions is currently confined to one lead month (Yuan and Wood 2013). Here, we analyze the potential to increase this time horizon by evaluating our predictions at times and in regions known to be influenced by El Niño-Southern Oscillation (ENSO) teleconnections. While the imprint of ENSO on regional precipitation is well-known, current evaluations of dynamical seasonal predictions of meteorological drought still insufficiently utilize the window of opportunity that arises from this statistical insight. Exploiting this window, the present study scrutinizes the idea that dynamical

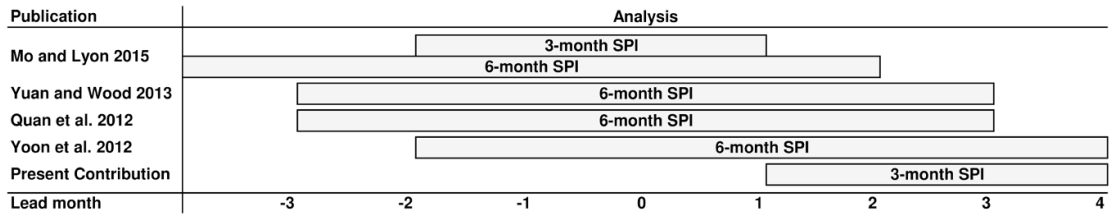

Figure 1. Useful hindcast skill identified in previous studies. Initialization is performed at lead month 0. Negative (positive) lead months employ observed (predicted) precipitation to compute the predicted SPI.

seasonal predictions of meteorological drought are during active ENSO states more skillful at larger lead times than expected.

The predictive skill of precipitation is usually unreliable over land on seasonal timescales (Kim *et al* 2012). Nevertheless, ENSO teleconnections affect regional precipitation and are known to generate seasonal prediction skill (Kumar *et al* 2013). Nowadays, ENSO–precipitation teleconnections are recognized as dominant forcing of regional precipitation over many areas in observations (Ropelewski and Halpert 1986, 1987, Dai and Wigley 2000, Seager *et al* 2005) and simulations (Schubert *et al* 2008, 2016). Unsurprisingly, ENSO-conditioned seasonal predictions of hydrological drought demonstrate skill at comparably long lead times (Wood and Lettenmaier 2006). Surprisingly, in contrast to hydrological drought, the predictive skill of ENSO-conditioned seasonal predictions of meteorological drought has to authors' best knowledge not yet been investigated at long lead times; although teleconnections between ENSO and meteorological drought indices, such as the standardized precipitation index (SPI) (McKee *et al* 1993), are nowadays well established for observations (Hallack-Alegria *et al* 2012, Manatsa *et al* 2017) and simulations (Mo *et al* 2009, Ma *et al* 2015) over many regions.

SPI is recommended by the WMO (Hayes *et al* 2011) and widely used for meteorological drought predictions (e.g. Yoon *et al* 2012, Ma *et al* 2015, Mo and Lyon 2015). The index quantifies the standardized deficit (or surplus) of precipitation during a predefined accumulation period. Here, we analyze SPI with an accumulation period of 3 months to investigate the hindcast skill of meteorological drought.

Most of the previous studies that investigated SPI hindcast skill evaluated the hindcast skill of overall SPI variability; rather than the intuitively more useful hindcast skill of meteorological drought (figure 1). However, attaining useful hindcast skill of meteorological drought, extreme SPI values, is more challenging (and arguably more relevant) than attaining useful hindcast skill of overall SPI variability (Ma *et al* 2015). The present contribution tackles this challenge of predicting the occurrence of meteorological drought.

A remaining key challenge for seasonal predictions of meteorological drought is to increase the lead time of skillful seasonal precipitation and drought index predictions (Wood *et al* 2015). At currently skillful lead times, initial conditions massively contribute to the evaluated hindcast skill (figure 1). Several studies (Quan *et al* 2012, Yoon *et al* 2012, Yuan and Wood 2013, Mo and Lyon 2015) have demonstrated significant SPI hindcast skill up to lead month 1 with an accumulation period of 3 months and up to lead month 3 with an accumulation period of 6 months. In these studies, hindcast skill is only significant if the lead time is about half as long as SPI's accumulation period. Consequently, only one half of the predicted SPI stems from the precipitation output of the model, while observations account for the other half. Demonstrating useful, significant hindcast skill of SPI that is derived only from the predicted precipitation output of the model constitutes the next frontier of meteorological drought predictions (Wood *et al* 2015). This study tackles that frontier.

Meteorological drought predictions need to merge several sources of information to be skillful (Wood *et al* 2015). Merging the dynamical prediction with observed precipitation is a valid, typical approach to exploit the memory of the drought index introduced by its accumulation period. However, this approach introduces two major drawbacks. First, the chosen accumulation period of the drought index prescribes scrutable lead times of the prediction. That confines the lead time at which predictions can demonstrate skill. Second, using observations in the calculation of the predicted drought index obscures the quantification of the model's predictive skill. That may lead to over-confidence in the performance of the model because the actual skill might originate from observations. Depending on the prediction time, these observations may impact the predicted drought index stronger than predicted precipitation. To avoid such obscurities and overconfidence, our predicted drought index is solely forecast-based and does not use observations. That also facilitates the investigation of ambitious lead times. Thus, the present contribution investigates dormant opportunities to reliably predict meteorological drought at large lead times through exploring the predictive potential of dynamical seasonal forecast systems during active ENSO years.

Instead of relying on a blend of observations and simulations in the predicted drought index, we attempt to extend predictive skill through ENSO teleconnections. Thus, we merge the dynamical prediction with ENSO as a second source of information. We investigate the lagged impacts of an active ENSO state on meteorological drought hindcast skill during winter (DJF) for the period 1982–2013 in seasonal hindcasts of the Max-Planck-Institute Earth System Model (MPI-ESM), which were initialized each start of November. The analysis conditions our prediction on active ENSO states by exploring significant agreements between two complementary analyses: hindcast skill composites of ENSO states, and ENSO–precipitation correlations. In this process, we investigate the sensitivity of our ENSO-state-conditioned prediction by considering different lead times of the ENSO signal and determine which of those lead times maximizes the area of reliable, ENSO-state-conditioned hindcast skill of meteorological drought in our analysis. To showcase the potential of ENSO-state conditioning, we use SPI with an accumulation period of 3 months to investigate the prediction's lead time of 2–4 months. With this investigation, we attempt to quadruple the time horizon of skillful SPI_{3M} predictions.

2. Data and methods

2.1. Data

Our seasonal prediction system (Baehr *et al* 2015, Bunzel *et al* 2018) is based on MPI-ESM, which is also used in the Coupled Model Intercomparison Project 5 (CMIP5). MPI-ESM couples general circulation components for the ocean (Jungclaus *et al* 2013) and the atmosphere (Stevens *et al* 2013). Moreover, MPI-ESM additionally contains subsystem components for terrestrial processes (Hagemann and Stacke 2015) and the marine bio-geochemistry (Ilyina *et al* 2013). For this study the model runs with 10 ensemble members in the same resolution as in CMIP5—MPI-ESM-LR (low-resolution): T63 (approx. $1.875° \times 1.875°$) with 47 vertical layers in the atmosphere between the surface and 0.01 hPa, and GR15 (maximum $1.5° \times 1.5°$) with 40 vertical layers in the ocean. Except for an extension of the simulation to cover the period 1982–2013, the analyzed simulations are identical to the ensemble investigated by Bunzel *et al* (2018). In hindcasts, initialized each start of November, we evaluate the precipitation output from December till February (lead months 2–4).

The German Weather Service used the seasonal prediction system employed in this study to issue operational forecasts until recently; when a successor version with a higher resolution was implemented (Fröhlich *et al* 2021) that is based on MPI-ESM1.2 (Mauritsen *et al* 2019). This increased resolution requires new parameterization schemes to estimate large-scale and convective precipitation amounts. Since these new schemes still need refinements, we test the conservative model version in this study.

Observed monthly precipitation is obtained from the global precipitation climatology project (GPCP). GPCP's dataset combines observations and satellite precipitation data into a $2.5° \times 2.5°$ global grid spanning 1979 to present (Adler *et al* 2003). To evaluate our hindcasts against these observations, the precipitation output of the model is interpolated to GPCP's grid.

2.2. Methods

ENSO conditioning explores significant agreements between two complementary analyses. First, obtaining significant Brier-Skill-Scores (BSS) hindcast skill in an ENSO composite analysis ensures the quality of the model's prediction. Attaining also significant observed correlations in an ENSO–precipitation correlation analysis safeguards the afore ascertained quality of the model. Correlation and composite analyses are both linked to a sound, well-understood physical mechanism and, thus, complement each other in our study. Moreover, while the correlation analysis quantifies precipitation variations relative to fluctuations in the signal, the composite analysis investigates the response of hindcast skill of dry extremes to extremes in the signal. By exploring grid-cell-wise significant congruences of both analyses, we establish robustness for our investigation.

We calculate the 3-month SPI during DJF (SPI_{DJF}) (McKee *et al* 1993) for observations and simulations to evaluate modeled against observed SPI_{DJF} timeseries. SPI timeseries ought to be normally distributed and it is important to note that non-normally distributed SPI_{DJF} timeseries would impair our evaluation process; same as differences between observations' and simulations' goodness of fit in SPI's calculation algorithm. Consequently, SPI's calculation algorithm ought to establish comparability between observed and modeled SPI_{DJF} timeseries by maximizing their normality both individually as well as concurrently. To ensure such comparability, we employ the methodology proposed by Pieper *et al* (2020), which uses the exponentiated Weibull distribution, to compute SPI_{3M} timeseries.

While evaluating hindcast skill of meteorological drought, we differentiate between two target regions that display strong ENSO–precipitation teleconnections: the southern USA and northern Mexico (henceforth referred to as North America), and northern South America (henceforth referred to as South America). It is noteworthy that all global data sets of observed precipitation data carry considerable uncertainties over South America (Mo and Lyon 2015). We address these uncertainties in the discussion of the results. Technical details how we condition our hindcast skill on active ENSO states can be found in appendix A.

3. ENSO-state-conditioned hindcast skill of meteorological drought

In agreement with prior studies (Yoon *et al* 2012, Mo and Lyon 2015, Wood *et al* 2015), hindcast skill of meteorological drought, assessed with BSS, is poor for lead months 2 to 4 in climate models such as MPI-ESM-LR almost everywhere around the globe (figure 2(a)). Still, the best hindcast skill of meteorological drought emerges in North and South America (black boxes in figure 2(a)). In particular, those parts of North and South America, where observed precipitation is strongly coupled to variations of the ENSO-index (figure 2(b)). Grid cells that demonstrate comparable high hindcast skill concurrently show large correlation values between the ENSO-index and precipitation (compare figures 2(c) with (d)). The more skillful the model's prediction of meteorological drought, the higher is the correlation value between observed precipitation and ENSO-index. This co-occurrence affirms our presumption that MPI-ESM-LR captures strong ENSO–precipitation teleconnections in our target regions. Yet, neutral ENSO states might conceal significant skill during active ENSO states.

Confining our hindcast skill analysis to start years that exhibit La Niña (figure 2(e)) or El Niño (figure 2(f)) conditions in ASO (the latest information available at the initialization at the start of November) substantially improves hindcast skill of meteorological drought. However, some grid cells (e.g. in western South America, and East North Central USA) show significant BSS hindcast skill (tested at the 5% confidence level; see appendix A for more information) in this composite analysis but weak ENSO–precipitation correlations. In those grid cells, we cannot maintain the claim that ENSO–precipitation teleconnections depict the physical basis for the skill improvement. Therefore, ENSO-state conditioning safeguards our analysis against overconfidence. To condition our hindcast skill of meteorological drought on ENSO states, we highlight grid cells (figures 2(g) and 1(h)) exhibiting both: significant correlations (also tested at the 5% confidence level; see appendix A for more information) between ENSO-index with precipitation (figure 2(d)) and significant hindcast skill of meteorological drought in the respective ENSO composite analysis (figures 2(e) and (f)). Thereby, we achieve reliable (significant in both analyses) ENSO-state-conditioned hindcast skill of meteorological drought (figures 2(g) and (h)).

Because a specific ENSO state contributes to either drying or wettening of our target regions, we separate our results into two cases. First, we obtain reliable hindcast skill of meteorological drought during a region's dry ENSO phase (indicated by brownish colored grid cells in figures 2(g) and (h)). Second, we obtain reliable hindcast skill of meteorological drought during a region's wet ENSO phase (indicated by greenish colored grid cells in figures 2(g) and (h)). We focus on the dry ENSO phase for the remainder of this study because skillful meteorological drought predictions are particularly important during this phase (Wilhite 1992, Wood *et al* 2015, Crimmins and McClaran 2016, Madadgar *et al* 2016, Baek *et al* 2019).

Next, we maximize the area of reliable hindcast skill of meteorological drought during the dry ENSO phase of our target regions. We maximize that area by examining its sensitivity to the prescribed lead time of the ENSO signal in our analysis. Instead of selecting composites based on (and correlating DJF precipitation with) the ASO ENSO signal, this sensitivity analysis investigates the ENSO signal in an earlier season than ASO. In this process, we identify that conditioning our hindcast skill of meteorological drought on JJA-ENSO states maximizes the area of each region's reliable hindcast skill of meteorological drought (the count of brown grid cells in figure 2(g) and (h)).

In North America (figures 3(a)–(c)) and South America (figures 3(d)–(f)), ENSO-index variability imprints similar during JJA as during ASO on observed DJF precipitation (compare figures 3(a) and (d) against figure 2(d)). This result agrees well with the lag identified by other studies (Redmond and Koch 1991, Harshburger *et al* 2002). Yet, when an ENSO event is present in the preceding boreal summer (JJA), MPI-ESM-LR captures ENSO–precipitation teleconnections better (see next paragraph). As a result of exploiting this lagged relationship, the count of grid cells showing significant BSS-assessed hindcast skill of meteorological drought increases in figure 3 relative to figure 2 by 60% (42%) in North (South) America. Consequently, also the count of grid cells in which ENSO-state conditioning ascertains reliable hindcast skill of meteorological drought during ENSO's dry phase (brownish colored grid cells) increases in figure 3 relative to figure 2 by 44% and 46% in North- and South America, respectively. Consequently, ENSO-state conditioning leads to reliable hindcast skill of meteorological drought at lead months 2 to 4 in large parts of our target regions during their respective dry ENSO phases.

But why does MPI-ESM-LR represent ENSO–precipitation teleconnections better when they are already present in JJA relative to when they present in ASO? Active JJA ENSO events typically develop into a stronger ENSO signal by the end of the year than active ASO ENSO events. Thus, they lead to more pronounced precipitation signals that affect larger parts of our target regions. Since MPI-ESM-LR generally under-represents spatial precipitation variability (not shown), the model benefits from these more pronounced signal after JJA ENSO events that uniformly affects a larger area than the signal of ENSO events that developed in between JJA and ASO.

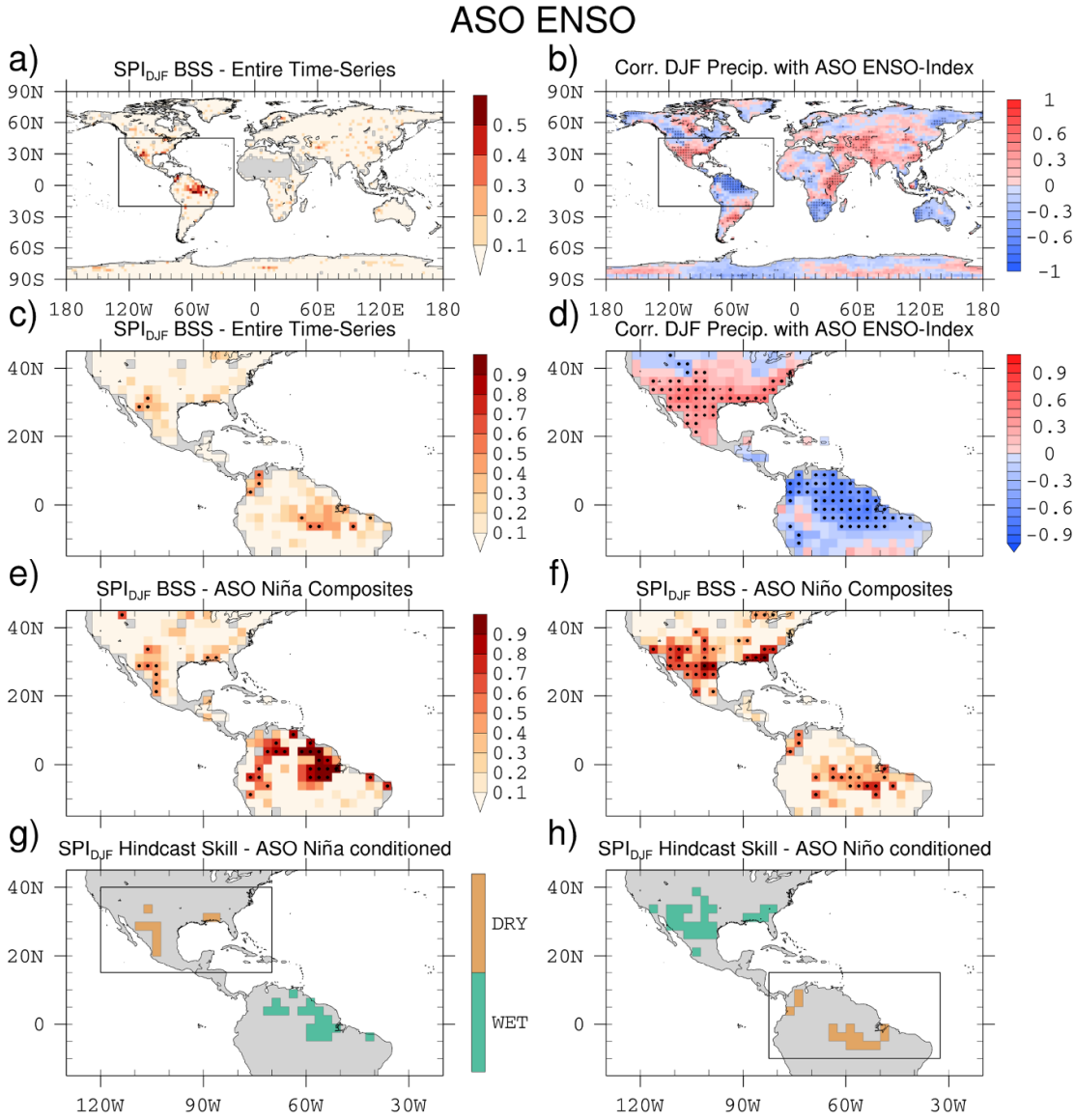

Figure 2. The BSS-assessed skill of the model in predicting meteorological drought at lead-months 2 to 4 and Pearson correlations between DJF precipitation and ASO ENSO-index on a global map ((a) and (b), respectively) and in our target regions ((c) and (d), respectively). BSS-assessed skill of predicting dry SPI_{DJF} extremes for a composite analysis which only considers years exhibiting La Niña (e) or El Niño (f) states present in ASO. Dots indicate BSS values significantly greater than 0 (which translates to Brier-Scores significantly greater than the ones of the random reference prediction) and Pearson correlations that significantly differ from 0. Reliable hindcast skill of dry SPI_{DJF} extremes achieved through conditioning the prediction on La Niña (g) or El Niño (h) states in ASO (i.e. significant correlations (d) that spatially coincide with significant BSS (e)/(f)). Colors indicate whether reliable hindcast skill is obtained during the region's wet (greenish) or dry (brownish) ENSO phase.

Between 1983–2013, La Niña and El Niño events observable in JJA became the strongest events in ASO. In contrast, comparable weak ASO events developed later than JJA (compare figure 4(a) against 4(d)). These comparable weak events, that developed in between JJA and ASO, often coincided with ordinary drought-prone conditions (SPI values close to −1 in figures 4(b) and (c)). The classification of these ordinary drought-prone conditions as drought or non-drought sensitively depends on SPI's threshold used in the BSS calculation. Such threshold sensitivity is highly unfavorable for any model tasked with the demonstration of BSS-assessed predictive skill. Consequently, omitting these comparably weak events from our analysis maximizes the area of reliable hindcast skill of meteorological drought as seen before. As a result of omitting these weak events, the ensemble mean prediction of SPI_{DJF} demonstrates a better agreement with observations during the remaining stronger events (compare highlighted years in figures 4(b) and (c) against 4(e) and (f)). This improved agreement during strong events is apparent e.g. in North America during the years 1999, 2000, and 2011 and in South America during the years 1983, 1992, 1998. During these years also the most intense meteorological droughts occurred in both regions, coinciding with particularly strong La Niña or El Niño events. The model seems to skillfully capture distinct

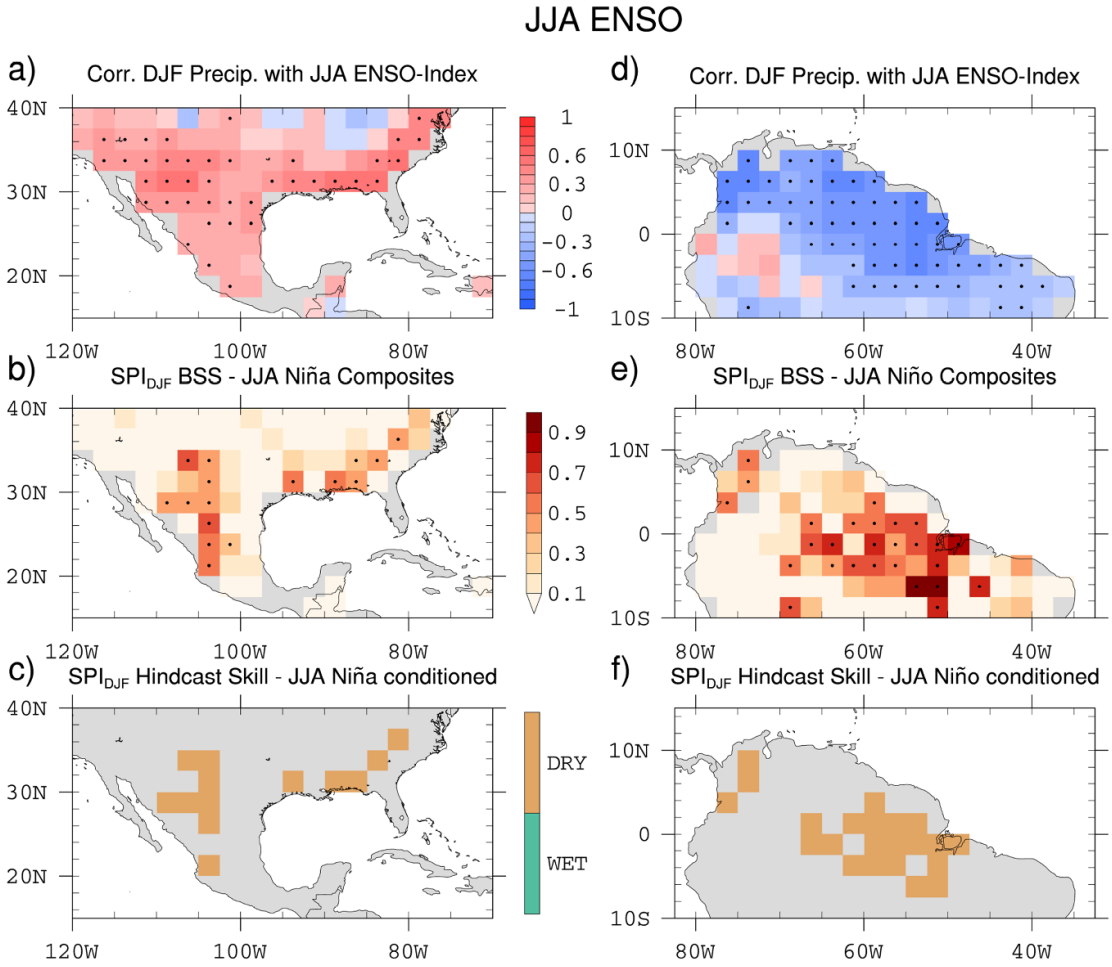

Figure 3. Correlations between DJF precipitation and JJA ENSO-index over North America (a) and South America (d). BSS-assessed skill of predicting dry SPI_{DJF} extremes for a composite analysis that only considers years exhibiting La Niña (b) or El Niño (e) states present in JJA. Dots indicate again BSS (Pearson correlations) significantly greater than (different from) 0. Reliable hindcast skill of SPI_{DJF} achieved through conditioning the prediction on La Niña (c) and El Niño (f) states present in JJA.

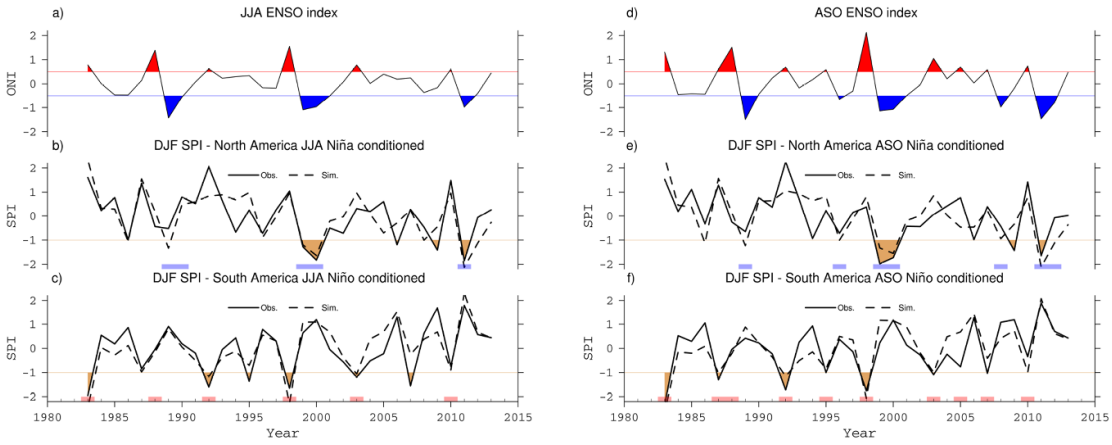

Figure 4. ENSO-index during JJA (**a**) and ASO (**d**). SPI_{DJF} averaged and standardized over the brownish colored grid points in figure 3(c) (b), 3(f) (c), 2(g) (e), and 2(h) (f). Observed averaged SPI_{DJF} is depicted by solid lines, while the ensemble mean is indicated by dashed lines. In JJA, the Pearson correlation between ENSO-index and observed (predicted) averaged SPI_{DJF} amounts to −0.67 (−0.7) in South and 0.56 (0.7) in North America, while the correlation between the ensemble mean and observed average SPI_{DJF} is 0.86 and 0.79 in South and North America, respectively. In ASO, the correlation between ENSO-index and observed (predicted) averaged SPI_{DJF} amounts to −0.75 (−0.77) in South and 0.57 (0.73) in North America, while the correlation between the ensemble mean and observations is 0.83 and 0.77 in South and North America, respectively.

teleconnections during these strong events. Yet, these distinct teleconnections may still vary inter-annually and do not necessarily cause meteorological droughts (see also Patricola *et al* 2020). These inter-annual variations are also captured by the model. The model correctly predicts normal conditions e.g. in South America during the strong El Niño event of 1988 or in North America during the phase-out of a strong La Niña event in 1990.

Despite capturing these inter-annual variations, the results beg the question whether the dynamical prediction provides any additional value beyond the statistical ENSO–SPI relationship. To answer this question, we tested the predictive skill of our ENSO-state-conditioned prediction against the prediction of a statistical model for all highlighted grid cells of figures 3(c) and (f). This statistical model linearly regresses the JJA ENSO index onto precipitation in each grid cell separately (not shown; see appendix B for more information). It is noteworthy that this statistical model does not separate between training and test period and derives individual regression coefficients for each grid cell. Thus, the results from the statistical model likely over-estimate the predictive skill of the statistical relationship due to over-fitting.

Averaged over all highlighted grid cells of figures 3(c) and (f), the dynamical model conclusively out-performs the statistical model despite the explained over-estimation. Over the entire time series, the statistical model predicts in North (South) America overall SPI_{DJF} variability with a correlation of 0.56 (0.67). In contrast, the ensemble mean of the dynamical SPI_{DJF} prediction exhibits over the entire time-series correlations of 0.79 and 0.86 in North and South America, respectively. The dynamical model also out-performs the statistical model in the prediction of meteorological droughts. Averaging their predictions over North (South) America, the statistical model predicts meteorological drought during the dry phase of ENSO with a BSS of 0.54 (0.31), while the dynamical, ENSO-state-conditioned prediction demonstrates BSS of 0.64 (0.68). In North (South) America, correlations (and BSS) of the dynamical model are significantly (tested at the 1% significance (99% confidence) level; see appendix B for more information) higher than those of the statistical prediction.

4. Discussion

ENSO-state conditioning reliably improves hindcast skill of meteorological drought in MPI-ESM-LR over North and South America during their respective dry ENSO phases. For ENSO-state conditioning to improve hindcast skill of meteorological drought, strong, large-scale ENSO–precipitation teleconnections need to be present. We confirm their existence through significant correlations between local precipitation and a lagged ENSO-index. Moreover, the forecast system needs to capture these ENSO–precipitation teleconnections. We ascertain this ability through significant hindcast skill of meteorological drought in an ENSO composite analysis. ENSO-state conditioning classifies hindcast skill of meteorological drought in those grid cells as reliable that concurrently pass significance tests of both analyses.

We condition our prediction on the state of ENSO in two different seasons (ASO and JJA). Depending on the season, on which we condition, the meteorological drought prediction of MPI-ESM-LR exhibits different strengths. Since La Niña and El Niño events generally occur more often in ASO (7 and 10 times in between 1983 and 2013, respectively) than in JJA (5 and 6 times, respectively), MPI-ESM-LR demonstrates reliable meteorological drought predictions more often when the prediction is conditioned on ASO-ENSO states. Yet, when active ENSO events are already present in JJA, they typically develop into stronger events by December that usually cause more distinct teleconnections covering a larger area. Therefore, MPI-ESM-LR captures the teleconnections of these stronger events (which are detectable in JJA) in more grid cells than the teleconnections of the weaker events (which are only detectable in ASO).

This explanation agrees with previous studies (Redmond and Koch 1991, Harshburger *et al* 2002) and with *NOAA Climate Prediction Center's* definition of an ENSO event: 5 consecutive overlapping seasons of ± 0.5 °C in the 3-month running mean Niño3.4-index (ONI) (Climate Prediction Center 2015). Active ENSO events detected at initialization in ASO may demonstrate an exceedance of this threshold only in four consecutive overlapping seasons by our prediction time in DJF. Since ENSO events generally peak around December, events present in JJA usually strengthen over the following months. Those events, present in JJA, universally demonstrate an exceedance of the threshold in at least six consecutive overlapping seasons by DJF, our prediction time. In the time-period analyzed here, we identify a single exception to this pattern in 1990. In 1990, one La Niña event was still present in JJA, while a neutral ENSO state emerged by ASO later that year. Still, this La Niña event persisted for more than 5 consecutive overlapping seasons beforehand. According to previous studies, the imprint of this La Niña event on precipitation over the American continent should be notable during our prediction time in DJF (Redmond and Koch 1991, Harshburger *et al* 2002).

Unsurprisingly, we identify strong ENSO–SPI_{DJF} teleconnections over the target regions investigated here. Yet, despite probable over-fitting, a statistical linear regression model predicts SPI_{DJF} distinctly worse than the dynamical model. For the entire time series, the statistical model predicts overall SPI_{DJF} variability significantly worse in both target regions.

In South America during the dry ENSO phase, the dynamical model demonstrates significantly better skill in predicting meteorological drought than the statistical model. While the dynamical model also better predicts meteorological drought than the statistical model South America, the improvement does not pass significance tests. These insights, that the ENSO-state-conditioned prediction conclusively out-performs the statistical prediction, additionally accentuate the potential value of ENSO-state conditioning.

It seems noteworthy that we do not perform any bias correction to the precipitation output of the model. The aim of this study is to demonstrate the skill of the ENSO-state conditioning. As a consequence of not performing any bias correction, the identified useful hindcast skill can be fully attributed to ENSO-state conditioning.

Our seasonal hindcasts span 31 years. The composite analysis, which considers only years exhibiting a certain ENSO state, further reduces our dataset to a minimum of 5 to 6 independent years, which arguably constitutes a scarce database. This issue is partially mitigated by the fact that BSS evaluates the entire probabilistic ensemble space of the prediction. Since our ensemble space is spanned by 10 different ensemble members, we rely on at least 50 to 60 events for our BSS-evaluation. Yet, an increasing ensemble size cannot arbitrarily compensate for a limited temporal length of dynamical seasonal hindcasts, because different ensemble members are not independent of each other. Thus, the problem of a scarce database would have been further exacerbated if we had conditioned our analysis on different ENSO flavors or on several climate oscillations. Different ENSO flavors and additional climate oscillations are certainly promising to capture a variety of precipitation teleconnections. However, such conditioning approaches are not feasible with current dynamical seasonal hindcasts initialized with satellite observations.

One way to alleviate the issue of statistical reliability is to decrease the SPI threshold that BSS uses to classify meteorological drought conditions. The threshold we use here is disputed within the literature. Svoboda *et al* (2002) proposed to identify meteorological drought conditions in the *US Drought Monitor* by an SPI threshold of −0.8—instead of −1, as used in this study. On one hand, a lower absolute value of this threshold would increase the number of (modeled and observed) meteorological droughts and would thereby increase statistical reliability. On the other hand, a lower absolute value of that threshold would result in a reduced extremity of the analyzed meteorological droughts. Disentangling these two competing effects has to the authors' best knowledge not been investigated up to now and is beyond the scope of this study.

While GPCP's precipitation data set is generally reliable, estimating South American precipitation is principally delicate. Observational datasets are notably sparse in South America. Consequently, uncertainties might be too large to reliably classify meteorological droughts (Mo and Lyon 2015). Despite these uncertainties, monthly precipitation analyses remain one of our most powerful tools for the task at hand.

This contribution attempts to highlight the potential and prove the concept of ENSO-state conditioning. During our analysis, we also checked for reliable ENSO-state-conditioned hindcast skill of meteorological drought outside of our target regions. Elsewhere in the world, ENSO-state conditioning only leads in single, scattered grid cells to reliable hindcast skill of meteorological drought during ENSO's dry phase (not shown). Thus, there appears to be little scope to extend ENSO-state conditioning to other regions that are characterized by strong ENSO–precipitation teleconnections with MPI-ESM-LR. MPI-ESM-LR seems to insufficiently capture these teleconnections elsewhere. Still, multi-model ensembles might compensate for emergent deficiencies through averaging which would lead to a better representation of teleconnections in these ensembles than in MPI-ESM-LR. Therefore, ENSO-state conditioning may improve meteorological drought predictions of multi-model ensembles also in other regions than those scrutinized in this investigation. Additionally, our analysis that uses ENSO-state conditioning to identify hotspots of meteorological drought predictability could be extended to soil-moisture drought (also known as agricultural drought). The Standardized Precipitation Evapotranspiration Index (SPEI; Vicente-Serrano *et al* 2010) measures soil-moisture drought but is calculated similarly as SPI and, thus, their substitution would require few methodological adjustments. Yet, an analysis of both indices can illuminate further windows of opportunity for predicting the propagation from meteorological to soil-moisture drought.

5. Conclusions

This study investigates hindcast skill of meteorological drought during DJF with 3-month SPI_{DJF}, which comprises lead months 2 to 4 of an initialized MPI-ESM seasonal hindcast ensemble. In previous studies, the predicted drought index usually merges predicted and observed precipitation. This approach artificially generates predictive skill. Additionally, this approach also links scrutable lead times to the chosen accumulation period of SPI. In contrast, our evaluation strictly separates simulations and observations and, thereby, quantifies genuine hindcast skill of the forecast system. To demonstrate reliable hindcast skill of meteorological drought despite this more challenging evaluation process, we exploit well-known ENSO–precipitation teleconnections. During ENSO's dry phase—when skillful meteorological drought predictions are particularly valuable—we achieve reliable

hindcast skill of meteorological drought up to four lead months ahead with SPI_{DJF}. Disentangling the accumulation period of SPI from the lead time of the prediction enables us to quadruple the lead time of reliable hindcast skill of dry SPI_{3M} extremes, meteorological droughts. At this unprecedented lead time for skillful meteorological drought predictions, the area of reliable hindcast skill of meteorological drought is further extended to cover large parts of northern South America and southern North America when the dry ENSO phase is already present in the preceding JJA. Thereby, this study reveals the potential of ENSO-state conditioning in uncovering the predictive potential of dynamical models by exploiting ENSO–precipitation teleconnections. During active ENSO states, dynamical seasonal meteorological drought predictions are more skillful at larger lead times than widely expected from typical predictability analyses. Exploiting this window of opportunity, we quadruple the lead time of skillful seasonal drought predictions with a single model whose output is not bias corrected. That revelation might encourage other analyses into windows of opportunities for meteorological drought predictions and excite further progress towards reliable and timely drought warnings.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=DKRZ_LTA_1075_ds00001.

Acknowledgments

The authors declare that they have no conflict of interest. P P and J B are funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2037 'CLICCS—Climate, Climatic Change, and Society'—Project Number: 390683824, contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg. P P was also funded by the BMBF-funded joint research projects RACE – Regional Atlantic Circulation and Global Change and RACE – Synthesis. A D is supported by A4 (Aigéin, Aeráid, agus athrú Atlantaigh), funded by the Marine Institute (Grant No. PBA/CC/18/01). The model simulations were performed at the Deutsche Klimarechenzentrum (DKRZ, German Climate Computing Centre). The authors thank David Nielsen for helpful discussions and comments on this manuscript.

Appendix A. ENSO-state conditioning

To quantify the strength of the ENSO signal, we calculate an ENSO-index by averaging SST anomalies, from the ERA-Interim reanalysis (Dee *et al* 2011), in the Niño3.4 region (5° S–5° N, 120° W–170° W). El Niño and La Niña events, used in the process of conditioning our prediction on active ENSO states, are identified analog to *NOAA Climate Prediction Center*, based on a threshold of ±0.5 °C in the 3-month running mean Niño3.4-index (ONI) (Climate Prediction Center 2015).

We condition our prediction on active ENSO states by exploring significant agreements between hindcast skill composites of active ENSO states and ENSO–precipitation correlations. In this process, we calculate Brier-Skill-Scores (BSS) (Murphy 1973) and Pearson correlations. BSS needs to distinguish between a drought and a non-drought event to quantify the hindcast skill. For this differentiation a threshold is set in accordance with WMO's *SPI User Guide* (Svoboda *et al* 2012) to an SPI value of −1. Significances of BSS (Pearson correlations) are computed with a one- (two-)sided 500-sample bootstrap which is evaluated at the 5% significance level against a null-hypothesis. The null-hypothesis for Pearson correlations assumes that the correlation is zero. The null-model for the Brier-Score is a random prediction that uses theoretical climatological occurrence probabilities to predict the likelihood of drought and non-drought conditions. We use well-known theoretical occurrence probabilities of the standard normal distribution for this random prediction since Pieper *et al* (2020) demonstrated the normality of the here employed calculation algorithm of SPI_{3M}.

Obtaining significant BSS hindcast skill in an ENSO composite analysis ensures the quality of the model's prediction. Attaining also significant observed correlations in an ENSO–precipitation correlation analysis safeguards the afore ascertained quality of the model. Correlation and composite analyses are both linked to a sound, well-understood physical mechanism and, thus, complement each other in our study. Moreover, while the correlation analysis quantifies precipitation variations relative to fluctuations in the signal, the composite analysis investigates the response of hindcast skill of dry SPI_{DJF} to extremes in the signal. By exploring grid-cell-wise significant congruences of both analyses, we establish the robustness of our investigation. We refer to this procedure as conditioning our hindcast skill on ENSO states. Since the hindcasts are initialized at the start of November, we consequently use the ENSO information available by November to condition our hindcast skill.

Appendix B. Evaluating the dynamical against a statistical prediction

We construct a linear regression model. Using this model we compute a statistical SPI prediction by linearly regressing the ENSO index time series onto the time series of observed local SPI. We perform

this regression for each brownish colored grid-cell in figures 3(c) and (f). Without any differentiation between training and test period, we use the entire time series to find the optimal regression coefficients for each grid cell.

To obtain single time series for all North (South) American grid cells, we average the statistical and the dynamical predictions over all brownish colored grid cells in figures 3(c) and (f). After the averaging process, we standardize the resulting time series (division by the resulting standard deviation). To evaluate these standardized time-series of our predictions against observations, we also average and standardize the observed SPI (derived from GPCP's precipitation analysis).

We evaluate the statistical and the dynamical prediction against observations by computing two different skill metrics. First, Pearson correlations over the entire time series indicate the skill of each model to predict overall SPI variability. Second, a BSS composites analysis evaluates the skill of each model to predict dry SPI extremes, meteorological drought, during the respective region's dry ENSO phase.

To evaluate whether the identified skill difference between the dynamical and the statistical prediction is significant, we compute one-sided 500-sample bootstraps. We evaluate these bootstraps at the 1% confidence level against the null-hypotheses that both predictions are identical. We test the significance in both directions. First, by bootstrapping the dynamical prediction, we evaluate whether the dynamical prediction is significantly better than the statistical prediction. Second, by bootstrapping the statistical prediction, we evaluate whether the statistical prediction is significantly worse than the dynamical prediction. Significance tests of both directions deliver the same results.

ORCID iDs

Patrick Pieper https://orcid.org/0000-0002-5555-4700
André Düsterhus https://orcid.org/0000-0003-2192-175X
Johanna Baehr https://orcid.org/0000-0003-4696-8941

References

Adler R F *et al* 2003 The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present) *J. Hydrometeorol.* **4** 1147–67
Baehr J *et al* 2015 The prediction of surface temperature in the new seasonal prediction system based on the MPI-ESM coupled climate model *Clim. Dyn.* **44** 2723–35
Baek S H, Smerdon J E, Seager R, Williams A P and Cook B I 2019 Pacific ocean forcing and atmospheric variability are the dominant causes of spatially widespread droughts in the contiguous united states *J. Geophys. Res.: Atmos.* **124** 2507–24
Bunzel F, Müller W A, Dobrynin M, Fröhlich K, Hagemann S, Pohlmann H, Stacke T and Baehr J 2018 Improved seasonal prediction of European summer temperatures with new five-layer soil-hydrology scheme *Geophys. Res. Lett.* **45** 346–53
Climate Prediction Center 2015 Oceanic Niño Index (ONI): cold & warm episodes by season
Crimmins M A and McClaran M P 2016 Where do seasonal climate predictions belong in the drought management toolbox? *Rangelands* **38** 169–76
Dai A and Wigley T 2000 Global patterns of ENSO-induced precipitation *Geophys. Res. Lett.* **27** 1283–6
Dee D P *et al* 2011 The era-interim reanalysis: configuration and performance of the data assimilation system *Q. J. R. Meteorol. Soc.* **137** 553–97
Fröhlich K *et al* 2021 The German climate forecast system: GCFS *J. Adv. Model. Earth Syst.* **13** e2020MS002101
Hagemann S and Stacke T 2015 Impact of the soil hydrology scheme on simulated soil moisture memory *Clim. Dyn.* **44** 1731–50
Hallack-Alegria M, Ramirez-Hernandez J and Watkins J, D 2012 Enso-conditioned rainfall drought frequency analysis in Northwest Baja California, Mexico *Int. J. Climatol.* **32** 831–42
Harshburger B, Ye H and Dzialoski J 2002 Observational evidence of the influence of pacific ssts on winter precipitation and spring stream discharge in idaho *J. Hydrol.* **264** 157–69
Hayes M, Svoboda M, Wall N and Widhalm M 2011 The lincoln declaration on drought indices: universal meteorological drought index recommended *Bull. Am. Meteorol. Soc.* **92** 485–8
Ilyina T, Six K D, Segschneider J, Maier-Reimer E, Li H and Núñez-Riboni I 2013 Global ocean biogeochemistry model HAMOCC: model architecture and performance as component of the MPI-earth system model in different cmip5 experimental realizations *J. Adv. Model. Earth Syst.* **5** 287–315
Jungclaus J, Fischer N, Haak H, Lohmann K, Marotzke J, Matei D, Mikolajewicz U, Notz D and Storch J 2013 Characteristics of the ocean simulations in the max Planck institute ocean model (MPIOM) the ocean component of the MPI-earth system model *J. Adv. Model. Earth Syst.* **5** 422–46
Kim H-M, Webster P J and Curry J A 2012 Seasonal prediction skill of ECMWF system 4 and NCEP cfsv2 retrospective forecast for the northern hemisphere winter *Clim. Dyn.* **39** 2957–73
Kumar A, Chen M and Wang W 2013 Understanding prediction skill of seasonal mean precipitation over the tropics *J. Clim.* **26** 5674–81
Ma F, Yuan X and Ye A 2015 Seasonal drought predictability and forecast skill over China *J. Geophys. Res.: Atmos.* **120** 8264–75
Madadgar S, aghakouchak A, Shukla S, Wood A W, Cheng L, Hsu K-L and Svoboda M 2016 A hybrid statistical-dynamical framework for meteorological drought prediction: application to the southwestern united states *Water Resour. Res.* **52** 5095–110
Manatsa D, Mushore T and Lenouo A 2017 Improved predictability of droughts over southern Africa using the standardized precipitation evapotranspiration index and ENSO *Theor. Appl. Climatol.* **127** 259–74
Mauritsen T *et al* 2019 Developments in the MPI-m earth system model version 1.2 (mpi-esm1. 2) and its response to increasing co2 *J. Adv. Model. Earth Syst.* **11** 998–1038
McKee T B, Doesken N J and Kleist J *et al* 1993 The relationship of drought frequency and duration to time scales *Proc. 8th Conf. Applied Climatology* vol 17 (Boston, MA: American Meteorological Society) pp 179–83
Mo K C and Lyon B 2015 Global meteorological drought prediction using the North American multi-model ensemble *J. Hydrometeorol.* **16** 1409–24
Mo K C, Schemm J-K E and Yoo S-H 2009 Influence of ENSO and the Atlantic multidecadal oscillation on drought over the united states *J. Clim.* **22** 5962–82

Murphy A H 1973 A new vector partition of the probability score *J. Appl. Meteorol.* **12** 595–600

Patricola C M, O'Brien J P, Risser M D, Rhoades A M, O'Brien T A, Ullrich P A, Stone D A and Collins W D 2020 Maximizing enso as a source of western us hydroclimate predictability *Clim. Dyn.* **54** 351–72

Pieper P, Düsterhus A and Baehr J 2020 A universal standardized precipitation index candidate distribution function for observations and simulations *Hydrol. Earth Syst. Sci.* **24** 4541–65

Quan X-W, Hoerling M P, Lyon B, Kumar A, Bell M A, Tippett M K and Wang H 2012 Prospects for dynamical prediction of meteorological drought *J. Appl. Meteorol. Climatol.* **51** 1238–52

Redmond K T and Koch R W 1991 Surface climate and streamflow variability in the western united states and their relationship to large-scale circulation indices *Water Resour. Res.* **27** 2381–99

Ropelewski C F and Halpert M S 1986 North american precipitation and temperature patterns associated with the El niño/Southern Oscillation (ENSO) *Mon. Weather Rev.* **114** 2352–62

Ropelewski C F and Halpert M S 1987 Global and regional scale precipitation patterns associated with the El niño/Southern Oscillation *Mon. Weather Rev.* **115** 1606–26

Schubert S D *et al* 2016 Global meteorological drought: a synthesis of current understanding with a focus on SST drivers of precipitation deficits *J. Clim.* **29** 3989–4019

Schubert S D, Suarez M J, Pegion P J, Koster R D and Bacmeister J T 2008 Potential predictability of long-term drought and pluvial conditions in the US great plains *J. Clim.* **21** 802–16

Seager R, Harnik N, Robinson W, Kushnir Y, Ting M, Huang H-P and Velez J 2005 Mechanisms of enso-forcing of hemispherically symmetric precipitation variability *Q. J. R. Meteorol. Soc.* A **131** 1501–27

Stevens B *et al* 2013 Atmospheric component of the MPI-M earth system model: ECHAM6 *J. Adv. Model. Earth Syst.* **5** 146–72

Svoboda M *et al* 2002 The drought monitor *Bull. Am. Meteorol. Soc.* **83** 1181–90

Svoboda M, Hayes M and Wood D 2012 *Standardized Precipitation Index User Guide* (Geneva: World Meteorological Organization)

Vicente-Serrano S M, Beguería S and López-Moreno J I 2010 A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index *J. Clim.* **23** 1696–1718

Wilhite D A 1992 Drought *Encyclopedia of Earth System Science* vol 2(San Diego, CA: Academic) pp 81–92

Wood A W and Lettenmaier D P 2006 A test bed for new seasonal hydrologic forecasting approaches in the western united states *Bull. Am. Meteorol. Soc.* **87** 1699–1712

Wood E F, Schubert S D, Wood A W, Peters-Lidard C D, Mo K C, Mariotti A and Pulwarty R S 2015 Prospects for advancing drought understanding, monitoring and prediction *J. Hydrometeorol.* **16** 1636–57

Yoon J-H, Mo K and Wood E F 2012 Dynamic-model-based seasonal prediction of meteorological drought over the contiguous united states *J. Hydrometeorol.* **13** 463–82

Yuan X and Wood E F 2013 Multimodel seasonal forecasting of global drought onset *Geophys. Res. Lett.* **40** 4900–5