Improving seasonal drought predictions by conditioning on ENSO states

Patrick Pieper[1], André Düsterhus[2], and Johanna Baehr[3]

[1]University Hamburg
[2]ICARUS
[3]Universität Hamburg, Center for Earth System Research and Sustainability

November 23, 2022

Abstract

Significant hindcast skill for the 3-month standardized precipitation index (SPI$_{3M}$) has been so far limited to one lead month. To increase that lead time, we propose to exploit well-known El Ni\~no-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 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. Predictions conditioned on autumn (ASO)-ENSO states demonstrate significant and reliable winter (DJF) drought hindcast skill up to lead month 4 in equatorial South- and southern North America. The area of reliable drought hindcast skill is further enlarged when the respective region's dry ENSO phase is already present in the antecedent summer (JJA-ENSO-state-conditioned). In contrast to previous studies, our evaluation separates predictions and observations. Thereby, ENSO-state conditioning demonstrates genuine hindcast skill up to lead month 4.

Improving seasonal drought predictions 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, Universität Hamburg, Hamburg, Germany

[2]ICARUS, Department of Geography, Maynooth University, Maynooth, Ireland

Key Points:

- We assess genuine hindcast skill by deriving the predicted drought index entirely from hindcasts
- We improve drought predictions by utilizing expertise on ENSO–precipitation teleconnections
- ENSO-state conditioning increases lead time of significant drought hindcast skill from 1 to 4 months

Corresponding author: Patrick Pieper, patrick.pieper@uni-hamburg.de

Abstract

Significant hindcast skill for the 3-month standardized precipitation index (SPI_{3M}) has been so far limited to one lead month. To increase that lead time, we propose to 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 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. Predictions conditioned on autumn (ASO)-ENSO states demonstrate significant and reliable winter (DJF) drought hindcast skill up to lead month 4 in equatorial South- and southern North America. The area of reliable drought hindcast skill is further enlarged when the respective region's dry ENSO phase is already present in the antecedent summer (JJA-ENSO-state-conditioned). In contrast to previous studies, our evaluation separates predictions and observations. Thereby, ENSO-state conditioning demonstrates genuine hindcast skill up to lead month 4.

Plain language summary

The time horizon of skillful seasonal drought predictions was in previous studies limited to 1 month. In this study, we increase that horizon to up to 4 months by exploiting a well-known and thoroughly investigated dependence of regional precipitation on sea-surface temperature anomalies in the equatorial Pacific Ocean. Yet, seasonal drought predictions still insufficiently capitalize on this expertise. Retrospective forecasts exhibit a better ability to predict winter droughts for a longer time horizon when these sea-surface temperature anomalies are sufficiently large. The magnitude of these anomalies is observable at the start of the prediction in November and does not change fundamentally during the prediction time. Thus, the uncertainty associated with our prediction decreases when the magnitude of those observed anomalies surpasses a certain threshold, which generates a predictable precipitation signal over the target regions. Furthermore, previous studies usually combine simulated with observed precipitation to derive the predicted drought index. This facilitates the identification of skill in the prediction. Such an approach blurs the proportion of the predictive skill that is based on the prediction. In contrast to this practice, we strictly separate observations from simulations and, thereby, demonstrate the genuine skill of our prediction in parts of the Americas.

1 Introduction

Reliable seasonal drought predictions can alleviate the harm caused by droughts through timely and accurate warnings, resulting in increased preparedness. However, the time horizon of reliable drought predictions is currently strictly confined to one lead month (Mo & Lyon, 2015; Ma et al., 2015; Yuan & Wood, 2013; Quan et al., 2012; Yoon et al., 2012). Here, we analyze the potential to increase this time horizon by evaluating our predictions for times and regions known to be influenced by El Niño-Southern Oscillation (ENSO) teleconnections. Previous studies have shown that SST anomalies in the equatorial Pacific lead the response of winter precipitation anomalies on the American continent by roughly 4 to 6 months (Redmond & Koch, 1991; Harshburger et al., 2002). Despite this expertise on lagged ENSO–precipitation teleconnections, current evaluations of dynamical seasonal drought predictions still insufficiently utilize this window of opportunity. Exploiting this, the present study generates significant and reliable drought hindcast skill up to lead month 4.

While the predictive skill of precipitation is usually unreliable over land (Kim et al., 2012), ENSO teleconnections affect regional precipitation and are known to generate seasonal prediction skill (Kumar et al., 2013). Several studies established ENSO teleconnections as a dominant forcing for observed precipitation over many regions (Seager et al., 2005; Dai & Wigley, 2000; Ropelewski & Halpert, 1987, 1986). Additionally, the same patterns of teleconnections were identified with similar strength in simulations (Schubert et al., 2016, 2008). The insights about ENSO–precipitation teleconnections were also successfully transferred to teleconnections between ENSO and specific drought indices.

One such drought index is the Standardized Precipitation Index (SPI) (McKee et al., 1993), which we use in this study. SPI is recommended by the WMO (Hayes et al., 2011) and widely in use (e.g., Mo & Lyon, 2015; Ma et al., 2015; Yoon et al., 2012). 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 predictability of meteorological droughts. Analog to ENSO-precipitation teleconnections, ENSO–SPI teleconnections are nowadays equally well established for observations (Manatsa et al., 2017; Hallack-Alegria et al., 2012) and simulations (Ma et al., 2015; Mo et al., 2009) over many regions. In summary, models usually capture ENSO–precipitation and ENSO–SPI teleconnections properly.

Ma et al. (2015) evaluated the seasonal forecast skill of SPI in ENSO composites. However, they focused on the relationship between seasonal drought predictability and forecast skill. Among several sensitivities, they also illustrate this relationship through ENSO composites. Their results indicate promising impacts of an active ENSO state on the forecast skill of general SPI variability. Yet, their results suggest that the impacts of active ENSO states on forecast skill of extremes, such as droughts, are less robust. However, Ma et al. (2015) investigated forecast skill over southern China. With this contribution, we want to investigate drought hindcast skill in northern South America and southern North America. Both regions display more pronounced ENSO–precipitation teleconnections than China (Dai & Wigley, 2000). We attempt to expand this expertise by investigating the predictive potential of ENSO–SPI teleconnections during active ENSO states. Our investigation focuses on opportunities to increase the lead time of reliable drought hindcast skill.

A remaining key challenge for seasonal predictions of meteorological droughts is to increase the lead time of skillful seasonal precipitation and drought index predictions (Wood et al., 2015). Several studies (e.g., Mo & Lyon, 2015; Yuan & Wood, 2013; Quan et al., 2012; Yoon et al., 2012) have demonstrated significant SPI hindcast skill up to lead month 1 with an accumulation period of 3 months and/or up to lead month 3 with an accumulation period of 6 months. In these studies, hindcast skill usually drops below the significance threshold when the lead time exceeds half of SPI's accumulation period. This implies that significant prediction skill has been achieved only when the precipitation output of the model accounts for not more than half of the data of the predicted SPI, while the other half stems from observations. The predicted SPI with an accumulation period of 3 (6) months employs observed precipitation in 2 (3) months. On one hand, this is a valid approach to exploit the memory of the drought index introduced by its accumulation period. On the other hand, 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, our predicted drought index is solely forecast based and does not use observations.

Consequently, we analyze drought hindcast skill using SPI with an accumulation period of 3 months (SPI_{3M}), which comprises lead months 2 to 4. Instead of relying on a blend of observations and simulations in the predicted drought index, we attempt to extend predictive skill through ENSO teleconnections. We investigate the lagged impacts of an active ENSO state on winter (DJF) drought hindcast skill 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 ENSO-state-conditioned drought hindcast skill in our analysis. To showcase the potential of ENSO-state conditioning, we investigate the lead time 2-4 months using SPI with an accumulation period of 3 months. With this investigation, we attempt to quadruple the time horizon of skillful drought predictions.

2 Data and methods

2.1 Data

Our seasonal prediction system (Baehr et al., 2015; Bunzel et al., 2018; Pieper et al., 2020a) 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 & 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°x1.875°) with 47 vertical layers in the atmosphere between the surface and 0.01 hPa, and GR15 (maximum 1.5°x1.5°) 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 to 4).

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°x2.5°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

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

While analyzing these timeseries, we differentiate between two target regions that display strong ENSO–precipitation teleconnections: the southern USA and northern Mexico (henceforth simply referred to as North America), and northern South America (henceforth simply referred to as South America).

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 the Brier-Score of a random prediction that uses theoretical climatological occurrence probabilities to predict the likelihood of drought and non-drought conditions (against the null-hypotheses that the correlation is zero). We use well-known theoretical occurrence probabilities of the standard normal distribution for this random prediction since Pieper et al. (2020b) 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 SPI to extremes in the signal. By exploring grid-cell-wise significant congruences of both analyses, we establish the robustness of our investigation. Henceforth, 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.

3 ENSO-state-conditioned drought hindcast skill

In agreement with prior studies (Mo & Lyon, 2015; Wood et al., 2015; Yoon et al., 2012), BSS-assessed drought hindcast skill is poor for lead months 2 to 4 in climate models such as MPI-ESM-LR almost everywhere around the globe (Fig 1a). Still, the best drought hindcast skill emerges in North and South America (black boxes in Fig 1a). In particular, those parts of North and South America, where observed precipitation is strongly coupled to variations of the ENSO-index (Fig 1b). Grid cells that demonstrate comparable high hindcast skill concurrently show large correlation values between the ENSO-index and precipitation (compare Fig 1c with 1d). The more skillful the model's prediction of droughts, 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.

Confining our hindcast skill analysis to start years that exhibit La Niña (Fig 1e) or El Niño (Fig 1f) conditions in ASO (at the initialization at the start of November) substantially improves drought hindcast skill. However, some grid cells (e.g. in western South America, and East North Central USA) show significant BSS hindcast skill 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 over-confidence. To condition our drought hindcast skill on ENSO states, we highlight grid cells (Fig 1g and 1h) exhibiting both: significant correlations between ENSO-index with precipitation (Fig 1d) and significant drought hindcast skill in the respective ENSO composite analysis (Fig 1e and 1f). Thereby, we achieve reliable (significant in both analyses) ENSO-state-conditioned drought hindcast skill (Fig 1g and 1h).

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 SPI_{3M} hindcast skill during a region's dry ENSO phase (indicated by brown grid cells in Fig 1g and 1h). Second, we obtain reliable SPI_{3M} hindcast skill during a region's wet ENSO phase (indicated by green grid cells in Fig 1g and 1h). Since we investigate drought hindcast skill, we focus on the dry ENSO phase for the remainder of this study.

Next, we maximize the area of reliable drought hindcast skill during the dry ENSO phase of our target regions. We maximize that area by examining its sensitivity to the prescribed lag of the ENSO signal in our analysis. Instead of selecting composites based on (and correlating DJF precipitation with) the ENSO signal in ASO, this sensitivity analysis investigates the ENSO signal in an earlier season than ASO. In this process, we identify that conditioning our drought hindcast skill on JJA-ENSO states maximizes the area of each region's reliable drought hindcast skill (the count of brown grid cells in Fig 1g and 1h).

In North America (Fig 2a - c) and South America (Fig 2d - f), ENSO-index variability imprints similar during JJA as during ASO on observed DJF precipitation (compare Fig 2a and 2d against Fig 1d). This result agrees well with the lag identified by other studies (Redmond & 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 drought hindcast skill increases in Fig 2 relative to Fig 1 by 60% (42%) in North (South) America. Consequently, also the count of grid cells in which we achieve reliable drought hindcast skill through ENSO-state conditioning increases in Fig 2 relative to Fig 1 by 44% and 46% in North- and South America, respectively. Consequently, ENSO-state conditioning leads to reliable drought hindcast skill for lead months 2 to 4 in large parts of our target regions during their respective dry ENSO phases.

Illustrating why MPI-ESM-LR represents ENSO–precipitation teleconnections better, when they are present in JJA than those present in ASO, finalizes our results. Time-series demonstrate that active ENSO events in JJA develop a stronger ENSO signal than active ENSO events in ASO. This stronger ENSO signal leads, via stronger ENSO–precipitation teleconnections, to a more pronounced precipitation signal in observations. MPI-ESM-LR captures this stronger signal easier than weaker signals, stemming from active ENSO events in ASO. Consequently, MPI-ESM-LR represents ENSO-precipitation teleconnections better when they are present in JJA than those only present in ASO.

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 Fig 3a against 3d). These comparable weak events, that developed in between JJA and ASO, often coincided with ordinary drought-prone conditions (SPI values close to -1 in Fig 3b and 3c). The classification of these ordinary drought-prone conditions as drought or non-drought sensitively depends on SPI's threshold used by BSS. 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 drought hindcast skill as seen before. As a result of omitting these weak events, SPI's DJF ensemble mean prediction demonstrates a better agreement with observations during the remaining stronger events (compare highlighted years in Fig 3b and 3c against 3e and 3f). 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 droughts occurred in both regions, coinciding with particularly strong La Niña or El Niño events. The model seems to skillfully capture distinct teleconnections during these strong events. Yet, these distinct teleconnections may still vary temporally and do not necessarily cause droughts (see also Patricola et al., 2020). These 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.

4 Discussion

ENSO-state conditioning reliably improves drought hindcast skill in MPI-ESM-LR over North and South America during their respective dry ENSO phases. For ENSO-state conditioning to improve drought hindcast skill, strong, large-scale ENSO–precipitation teleconnections need to be present. We confirm their existence and relevance 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 drought hindcast skill in the composite analysis. ENSO-state conditioning classifies this drought hindcast skill as reliable only in those grid cells that concurrently also display significant correlations.

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 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 1983-2013, respectively) than in JJA (5 and 6 times, respectively), MPI-ESM-LR demonstrates reliable drought predictions more often when they are ENSO-state-conditioned on ASO-ENSO events. Yet, when active ENSO events persist in JJA, they usually cause more distinct teleconnections that cover 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 & 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 4 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, usually demonstrate an exceedance of the threshold in at least

6 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 before the time of our prediction in DJF. 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 & Koch, 1991; Harshburger et al., 2002).

We also checked for ENSO-state-conditioned drought hindcast skill outside of our target regions. Elsewhere in the world, ENSO-state conditioning only leads in single, scattered grid cells to reliable drought hindcast skill during ENSO's dry phase (not shown). In MPI-ESM-LR, ENSO-state conditioning improves drought hindcast skill only in the investigated target regions. This indicates a plausible reason for our drought hindcast skill to improve stronger for longer lead times than Ma et al. (2015) were able to identify over south China during an active ENSO.

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. Aside, there could be scope to employ ENSO-state conditioning in a similar manner, as demonstrated here, to improve the hindcast skill of surplus precipitation extremes (by suitably adapting the BSS threshold).

Our seasonal hindcasts start – as usually with the satellite era – in 1982 spanning 31 years. The composite analysis, which considers only years exhibiting a certain ENSO state, further reduces our dataset to 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 completely independent of each other. Thus, the problem of a scarce database would be further exacerbated if we had e.g. analyzed different ENSO flavors. Different ENSO flavors are certainly promising to capture variations in ENSO–precipitation teleconnections. However, such an analysis is 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 drought conditions. The threshold we use here is disputed within the literature. Svoboda et al. (2002) proposed to identify drought conditions in the *US Drought Monitor* by an SPI threshold of -0.8 – rather than -1, as used in this study. On one hand, a lower absolute value of this threshold would increase the number of (modeled and observed) 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 droughts. Disentangling these two competing effects is difficult, and has to the authors' best knowledge not been investigated up to now.

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 droughts (Mo & Lyon, 2015). Despite these uncertainties, monthly precipitation analyses remain one of our most powerful tools for the task at hand.

5 Conclusions

This study investigates drought hindcast skill of DJF SPI_{3M}, which comprises lead months 2 to 4, in an initialized MPI-ESM seasonal hindcast ensemble. The evaluation process of SPI hindcasts usually combines predicted and observed precipitation. Such a combination artificially generates predictive skill. In contrast, our evaluation strictly separates simulations and observations and, thereby, quantifies genuine hindcast skill of the forecast system. To demonstrate reliable drought hindcast skill despite this more challenging evaluation process, we exploit well-known ENSO–precipitation teleconnections. During ENSO's dry phase – when skillful drought predictions are particularly valuable –, we achieve reliable drought hindcast skill up to 4 lead months ahead with SPI_{3M} in DJF. When the dry ENSO phase is already present in the preceding JJA, the area of reliable drought hindcast skill covers large parts of northern South America and southern North America. Ultimately, this study reveals the potential of ENSO-state conditioning in uncovering the predictive potential of dynamical models by exploiting ENSO–

precipitation teleconnections. That revelation might excite further progress towards reliable and timely drought warnings.

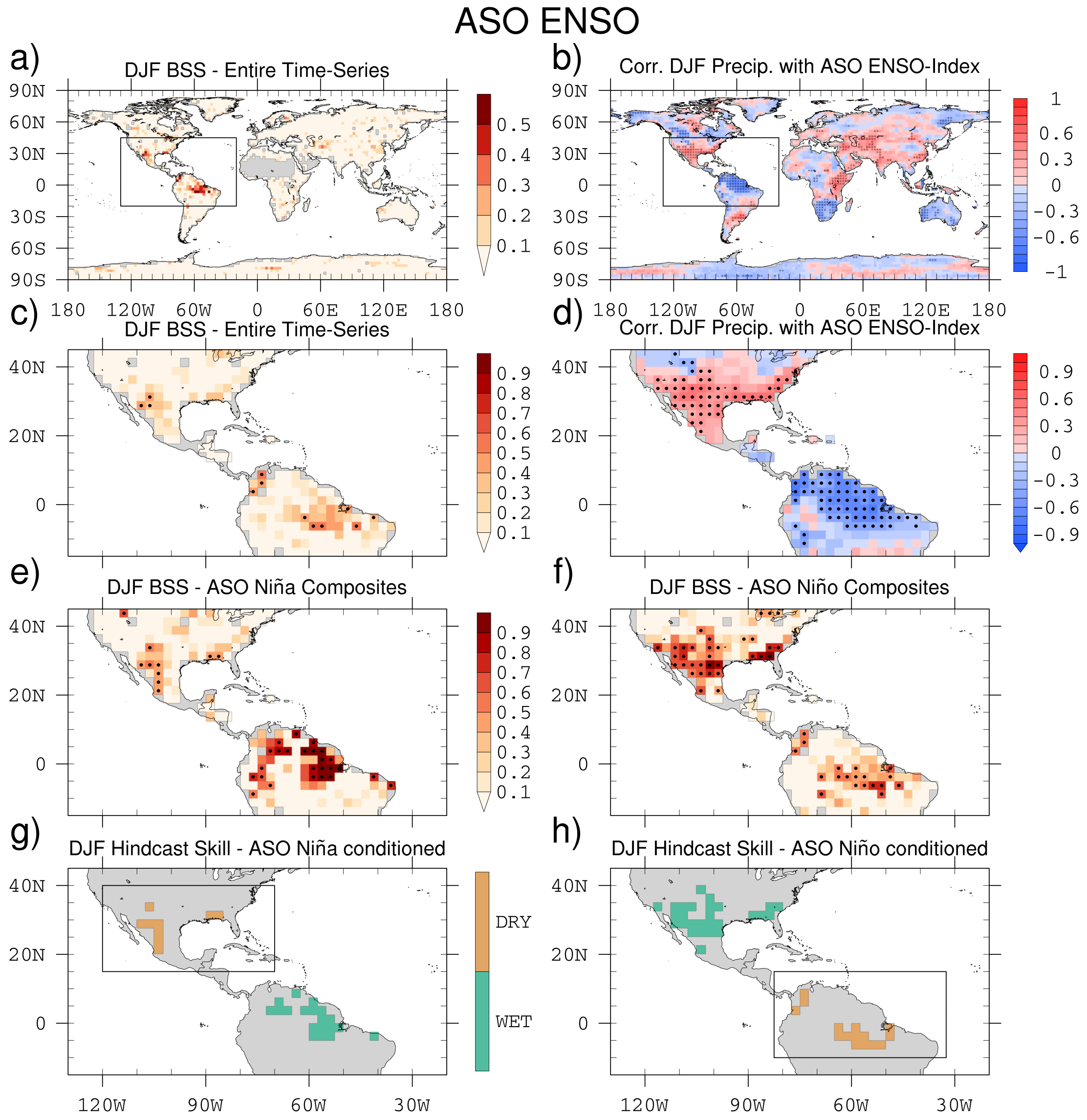

Figure 1. The BSS-assessed skill of the model in predicting droughts at lead-months 2 to 4 and Pearson correlations between DJF precipitation and ASO ENSO-index on a global map (**a** and **b**, respectively) and in our target regions (**c** and **d**, respectively). BSS for a composite analysis which 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 during DJF achieved through conditioning the prediction on La Niña (**g**) or El Niño (**h**) states in ASO (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.

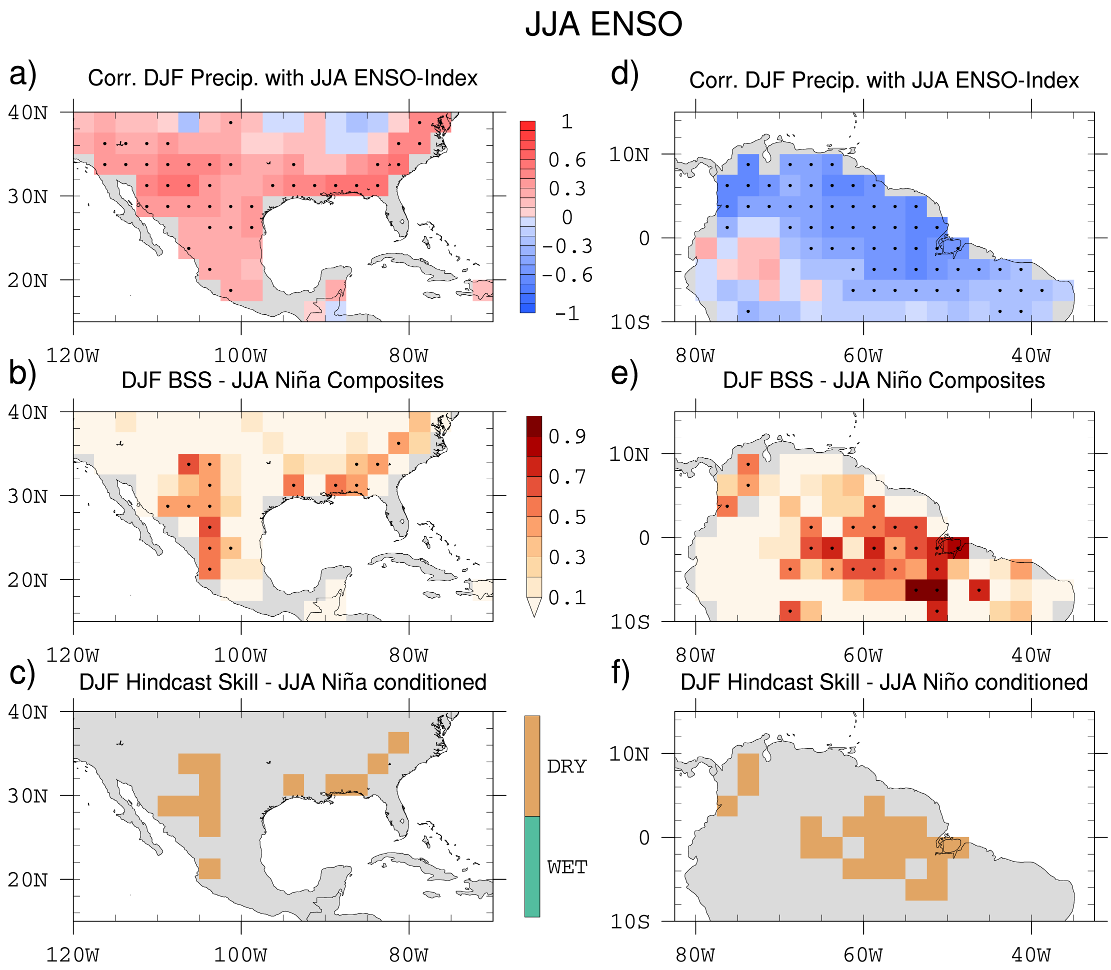

Figure 2. Correlations between DJF precipitation and JJA ENSO-index over North America (**a**) and South America (**d**). BSS for a composite analysis that only considers years exhibiting La Niña (**b**) or El Niño (**e**) states present in JJA. Dots indicate again BSS (Pearson correlations) significantly greater than (different from) 0. Reliable hindcast skill during DJF achieved through conditioning the prediction on La Niña (**c**) and El Niño (**f**) states present in JJA.

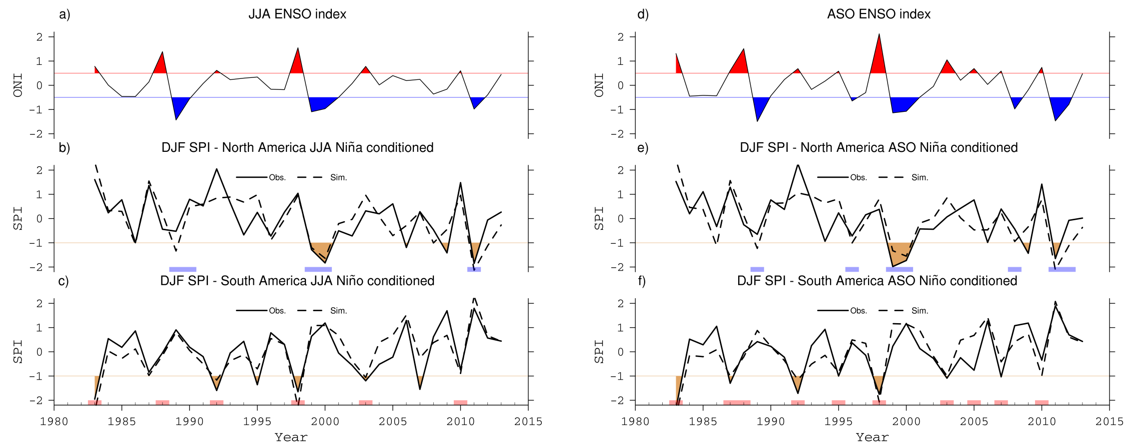

Figure 3. ENSO-index during JJA (**a**) and ASO (**d**). DJF SPI averaged and standardized over the brownish colored grid points in Fig 2c (**b**), 2f (**c**), 1g (**e**), and 1h (**f**). Observations are depicted by solid lines, while the ensemble mean is indicated by dashed lines. In JJA, the Pearson correlation between ENSO-index and observations (simulations) amounts to -0.67 (-0.7) in South and 0.56 (0.7) in North America, while the correlation between the ensemble mean and observations is 0.86 and 0.79 in South and North America, respectively. In ASO, the correlation between ENSO-index and observations (simulations) 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.

Acknowledgments

This work was funded by the BMBF-funded joint research projects RACE – Regional Atlantic Circulation and Global Change and RACE – Synthesis. P.P. is supported by the Stiftung der deutschen Wirtschaft (SDW, German Economy Foundation). A.D. and J.B. are supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC 2037 "Climate, Climatic Change, and Society"–Project: 390683824, contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg. A.D. is also supported by A4 (Aigéin, Aeráid, agus athrú Atlantaigh), funded by the Marine Institute and the European Regional Development fund (grant: PBA/CC/18/01). The authors declare that they have no conflict of interest. The model simulations were performed at the Deutsche Klimarechenzentrum (DKRZ, German Climate Computing Centre) and are available at the World Data Center for Climate (WDCC): `http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=DKRZ_LTA_1075_ds00001` maintained by DKRZ. The authors thank David Nielsen for helpful discussions and comments on this manuscript.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., … others (2003). The version-2 global precipitation climatology project (gpcp) monthly precipitation analysis (1979–present). *Journal of hydrometeorology*, *4*(6), 1147–1167.

Baehr, J., Fröhlich, K., Botzet, M., Domeisen, D. I., Kornblueh, L., Notz, D., … Mueller, W. A. (2015). The prediction of surface temperature in the new seasonal prediction system based on the mpi-esm coupled climate model. *Climate Dynamics*, *44*(9-10), 2723–2735.

Bunzel, F., Müller, W. A., Dobrynin, M., Fröhlich, K., Hagemann, S., Pohlmann, H., … Baehr, J. (2018). Improved seasonal prediction of european summer temperatures with new five-layer soil-hydrology scheme. *Geophysical Research Letters*, *45*(1), 346–353.

Climate Prediction Center. (2015). *Oceanic Niño Index (ONI): Cold & warm episodes by season.* Retrieved from `https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php`

Dai, A., & Wigley, T. (2000). Global patterns of enso-induced precipitation. *Geophysical Research Letters*, *27*(9), 1283–1286.

Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., … others (2011). The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553–597.

Hagemann, S., & Stacke, T. (2015). Impact of the soil hydrology scheme on simulated soil moisture memory. *Climate Dynamics*, *44*(7-8), 1731–1750.

Hallack-Alegria, M., Ramirez-Hernandez, J., & Watkins Jr, D. (2012). Enso-conditioned rainfall drought frequency analysis in northwest baja california, mexico. *International Journal of Climatology*, *32*(6), 831–842.

Harshburger, B., Ye, H., & Dzialoski, J. (2002). Observational evidence of the influence of pacific ssts on winter precipitation and spring stream discharge in idaho. *Journal of Hydrology*, *264*(1-4), 157–169.

Hayes, M., Svoboda, M., Wall, N., & Widhalm, M. (2011). The lincoln declaration on drought indices: universal meteorological drought index recommended. *Bulletin of the American Meteorological Society*, *92*(4), 485–488.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., & 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. *Journal of Advances in Modeling Earth Systems*, *5*(2), 287–315.

Jungclaus, J., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., ... 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. *Journal of Advances in Modeling Earth Systems*, *5*(2), 422–446.

Kim, H.-M., Webster, P. J., & Curry, J. A. (2012). Seasonal prediction skill of ecmwf system 4 and ncep cfsv2 retrospective forecast for the northern hemisphere winter. *Climate Dynamics*, *39*(12), 2957–2973.

Kumar, A., Chen, M., & Wang, W. (2013). Understanding prediction skill of seasonal mean precipitation over the tropics. *Journal of Climate*, *26*(15), 5674–5681.

Ma, F., Yuan, X., & Ye, A. (2015). Seasonal drought predictability and forecast skill over china. *Journal of Geophysical Research: Atmospheres*, *120*(16), 8264–8275.

Manatsa, D., Mushore, T., & Lenouo, A. (2017). Improved predictability of droughts over southern africa using the standardized precipitation evapotranspiration index and ENSO. *Theoretical and applied climatology*, *127*(1-2), 259–274.

McKee, T. B., Doesken, N. J., Kleist, J., et al. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th conference on applied climatology* (Vol. 17, pp. 179–183).

Mo, K. C., & Lyon, B. (2015). Global meteorological drought prediction using the north american multi-model ensemble. *journal of Hydrometeorology*, *16*(3), 1409–1424.

Mo, K. C., Schemm, J.-K. E., & Yoo, S.-H. (2009). Influence of enso and the atlantic multidecadal oscillation on drought over the united states. *Journal of Climate*, *22*(22), 5962–5982.

Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology*, *12*(4), 595–600.

Patricola, C. M., O'Brien, J. P., Risser, M. D., Rhoades, A. M., O'Brien, T. A., Ullrich, P. A., … Collins, W. D. (2020). Maximizing enso as a source of western us hydroclimate predictability. *Climate Dynamics*, *54*(1-2), 351–372.

Pieper, P., Düsterhus, A., & Baehr, J. (2020a). *MPI-ESM-LR seasonal precipitation hindcasts.* World Data Center for Climate (WDCC) at DKRZ. Retrieved from http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=DKRZ_LTA_1075_ds00001

Pieper, P., Düsterhus, A., & Baehr, J. (2020b). Global and regional performances of spi candidate distribution functions in observations and simulations. *Hydrology and Earth System Sciences Discussions*, 1–33.

Quan, X.-W., Hoerling, M. P., Lyon, B., Kumar, A., Bell, M. A., Tippett, M. K., & Wang, H. (2012). Prospects for dynamical prediction of meteorological drought. *Journal of Applied Meteorology and Climatology*, *51*(7), 1238–1252.

Redmond, K. T., & Koch, R. W. (1991). Surface climate and streamflow variability in the western united states and their relationship to large-scale circulation indices. *Water Resources Research*, *27*(9), 2381–2399.

Ropelewski, C. F., & Halpert, M. S. (1986). North american precipitation and temperature patterns associated with the el niño/southern oscillation (enso). *Monthly Weather Review*, *114*(12), 2352–2362.

Ropelewski, C. F., & Halpert, M. S. (1987). Global and regional scale precipitation patterns associated with the el niño/southern oscillation. *Monthly weather review*, *115*(8), 1606–1626.

Schubert, S. D., Stewart, R. E., Wang, H., Barlow, M., Berbery, E. H., Cai, W., … others (2016). Global meteorological drought: a synthesis of current understanding with a focus on sst drivers of precipitation deficits. *Journal of Climate*, *29*(11), 3989–4019.

Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D., & Bacmeister, J. T. (2008). Potential predictability of long-term drought and pluvial conditions in the us great plains. *Journal of Climate*, *21*(4), 802–816.

Seager, R., Harnik, N., Robinson, W., Kushnir, Y., Ting, M., Huang, H.-P., & Velez, J. (2005). Mechanisms of enso-forcing of hemispherically symmetric precipitation variability. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *131*(608), 1501–1527.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., . . . others (2013). Atmospheric component of the mpi-m earth system model: Echam6. *Journal of Advances in Modeling Earth Systems*, *5*(2), 146–172.

Svoboda, M., Hayes, M., & Wood, D. (2012). Standardized precipitation index user guide. *World Meteorological Organization Geneva, Switzerland*.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., . . . others (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181–1190.

Wood, E. F., Schubert, S. D., Wood, A. W., Peters-Lidard, C. D., Mo, K. C., Mariotti, A., & Pulwarty, R. S. (2015). Prospects for advancing drought understanding, monitoring, and prediction. *Journal of Hydrometeorology*, *16*(4), 1636–1657.

Yoon, J.-H., Mo, K., & Wood, E. F. (2012). Dynamic-model-based seasonal prediction of meteorological drought over the contiguous united states. *Journal of Hydrometeorology*, *13*(2), 463–482.

Yuan, X., & Wood, E. F. (2013). Multimodel seasonal forecasting of global drought onset. *Geophysical Research Letters*, *40*(18), 4900–4905.