

# Water Resources Research®



## RESEARCH ARTICLE

10.1029/2021WR031302

# Hotspots of Predictability: Identifying Regions of High Precipitation Predictability at Seasonal Timescales From Limited Time Series Observations

Antonios Mamalakis<sup>1,2</sup> , Amir AghaKouchak<sup>1,3</sup> , James T. Randerson<sup>1,3</sup> , and Efi Foufoula-Georgiou<sup>1,3</sup> 

<sup>1</sup>Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA, <sup>2</sup>Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA, <sup>3</sup>Department of Earth System Science, University of California, Irvine, CA, USA

### Key Points:

- We present a rigorous probabilistic framework and novel statistical metrics to assess the predictability of seasonal precipitation totals
- Hotspots of winter precipitation predictability occur in the Southwest and Southeast of contiguous United States (CONUS)
- Sea Surface Temperature (SST) information may help with predicting dry or wet conditions in some areas, while normal conditions are difficult to predict

### Correspondence to:

A. Mamalakis and E. Foufoula-Georgiou, [amamalak@rams.colostate.edu](mailto:amamalak@rams.colostate.edu); [efi@uci.edu](mailto:efi@uci.edu)

### Citation:

Mamalakis, A., AghaKouchak, A., Randerson, J. T., & Foufoula-Georgiou, E. (2022). Hotspots of predictability: Identifying regions of high precipitation predictability at seasonal timescales from limited time series observations. *Water Resources Research*, 58, e2021WR031302. <https://doi.org/10.1029/2021WR031302>

Received 26 SEP 2021

Accepted 9 MAY 2022

**Abstract** Precipitation prediction at seasonal timescales is important for planning and management of water resources as well as preparedness for hazards such as floods, droughts and wildfires. Quantifying predictability is quite challenging as a consequence of a large number of potential drivers, varying antecedent conditions, and small sample size of high-quality observations available at seasonal timescales, that in turn, increases prediction uncertainty and the risk of model overfitting. Here, we introduce a generalized probabilistic framework to account for these issues and assess predictability under uncertainty. We focus on prediction of winter (Nov–Mar) precipitation across the contiguous United States, using sea surface temperature-derived indices (averaged in Aug–Oct) as predictors. In our analysis we identify “predictability hotspots,” which we define as regions where precipitation is inherently more predictable. Our framework estimates the entire predictive distribution of precipitation using copulas and quantifies prediction uncertainties, while employing principal component analysis for dimensionality reduction and a cross validation technique to avoid overfitting. We also evaluate how predictability changes across different quantiles of the precipitation distribution (dry, normal, wet amounts) using a multi-category  $3 \times 3$  contingency table. Our results indicate that well-defined predictability hotspots occur in the Southwest and Southeast. Moreover, extreme dry and wet conditions are shown to be relatively more predictable compared to normal conditions. Our study may help with water resources management in several subregions of the United States and can be used to assess the fidelity of earth system models in successfully representing teleconnections and predictability.

## 1. Introduction

It is well established that precipitation predictability on seasonal timescales draws upon information about ocean dynamics, which is considered as the principal forcing variable of the atmospheric circulation that ultimately drives regional hydroclimate (Charney & Shukla, 1981; DelSole & Banerjee, 2017; Goddard et al., 2001; Hao et al., 2018; Khan et al., 2017; Palmer & Anderson, 1994; Sankarasubramanian et al., 2009; Zimmerman et al., 2016). Sea surface temperatures (SSTs), which provide a partial constraint on ocean circulation, are typically used for prediction either with deterministic models (i.e., climate model simulations, also broadly known as dynamical models) or statistical models that aim to exploit physically and historically established teleconnections of precipitation with large-scale modes of climate variability (e.g., El Niño–Southern Oscillation, ENSO; see Dai, 2013; Enfield et al., 2001; Hao et al., 2018; Lindsey, 2016; McCabe & Dettinger, 1999; McCabe et al., 2004; Newman et al., 2016; Redmond & Koch, 1991; Schonher & Nicholson, 1989). However, both approaches face important limitations, and reliable prediction of seasonal precipitation remains a challenge (National Academies of Sciences, 2016; Wang et al., 2009). Limits to predictive skill of dynamical models arise from epistemic uncertainties (like incomplete knowledge of the underlying processes, uncertain initial conditions, uncertainties in model physics and limits on computational resources that place constraints on the grid resolution used in operational systems; Becker et al., 2014; Chang et al., 2000; Mamalakis et al., 2017), as well as aleatoric uncertainties due to the chaotic nature of the system. Similarly, empirical statistical models exhibit limited predictive skill, arising primarily from the complex nature of the relationship between large scale modes and regional hydroclimate, and the limited sample size that is available to build robust predictive models.

In the United States, expectations have been set in the recent report by the National Academies of Sciences, Engineering and Medicine (National Academies of Sciences, 2016), that seasonal predictions would, in 10 years,

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

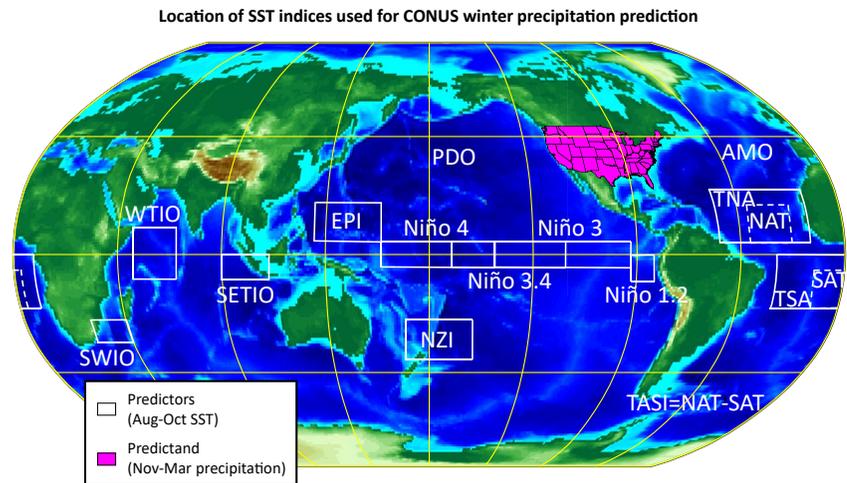
be used like weather forecasts are today. Under this premise and given the still challenging nature of the problem, it is important to recognize that in the Earth system there are periods and regions of inherently higher and lower climate predictability. For example, the so-called “forecasts of opportunity” (Mariotti et al., 2020; Mayer & Barnes, 2021) correspond to time windows of enhanced predictability for the climate system, which arise from specific climate modes like ENSO being active. When one of these climate modes is active, it can lead to persistent and consistent anomalous responses in the atmospheric circulation (e.g., quasi-stationary Rossby waves) that drive regional hydroclimate, thus, enhancing climate predictability. As recommended by National Academies of Sciences, (2016), broader acknowledgment and communication that the predictability of the climate system varies temporally and spatially would more adequately address user community expectations for skillful prediction and actionable information (Mariotti et al., 2020).

Here, we are interested in the spatial variation of predictability and recognize that there exist regions with inherently higher levels of predictability, that is, “hotspots of predictability” that are often surrounded by regions where predictive skill is considerably lower. Physically, hotspots of precipitation predictability are regions where the forcing signal of large-scale climate modes (like ENSO) to precipitation is robust and is not overwhelmed by chaotic internal variability. In these regions the forcing signal from climate indices has a high signal-to-noise ratio in explaining past precipitation observations (Wang et al., 2017). Quantitatively, we herein define hotspots of predictability as regions where the prediction skill derived from large-scale climate indices is statistically distinguishable from a climatology-based prediction (at a  $p < 0.05$  level), when evaluated considering prediction uncertainties and using an out-of-sample test data set. In contrast, surrounding regions where climate indices do not yield statistically significant predictions for out-of-sample test data are defined as low (or zero) predictability regions.

Under this setting, we revisit the problem of seasonal precipitation prediction across the contiguous United States (CONUS) and aim to gain insight into mechanisms structuring hotspots of precipitation predictability. Specifically, we focus on the winter months (Nov–Mar) – a time of the year with elevated fire risk for southeastern states (Li et al., 2020; Nowell et al., 2018) and when most precipitation occurs in western states – and attempt to answer the following questions: (a) In which CONUS regions is winter precipitation inherently more predictable? (b) How is predictability distributed across different classes of precipitation totals (dry, normal, wet)? (c) Are dry and wet extremes more or less predictable than normal conditions? (d) What are the sources of predictability, that is, the large-scale climate modes that contribute to predictability? Although prediction of precipitation across different CONUS regions has been explored in the past (Gibson et al., 2021; Liu et al., 2018; Madadgar et al., 2016; Mamalakis et al., 2018; McCabe & Dettinger, 1999; Pan et al., 2019; Schonher & Nicholson, 1989; Stevens et al., 2021; Zarekarizi et al., 2018; Zhang et al., 2018; Zimmerman et al., 2016), fewer studies have approached the problem focusing on the entire CONUS (Devineni & Sankarasubramanian, 2010a, 2010b; Gershunov & Cayan, 2003; Schubert et al., 2016), and to the best of our knowledge, our study is the first to explore hotspots of predictability and to systematically address the four questions described above. Results from our analysis may improve water resources management at a federal level, by integrating the knowledge of how predictability is distributed spatially across the US for both dry and wet precipitation extremes.

In our approach, we assess seasonal predictability by recognizing that one of the key problems in this setting is the limited sample size of high-quality time series observations. The small sample size cautions against overinterpreting predictability when prediction uncertainty has not been quantitatively assessed and calls for also addressing the high risk of overfitting (DelSole & Banerjee, 2017; Ham et al., 2019; Stevens et al., 2021). So far, no generalized framework of how to assess predictability under this setting exists in the literature. Indeed, to avoid overfitting, previous studies often have used simplified models that rely on a small number of predictors and thus may not leverage the full structure of global SST information (L’Heureux et al., 2015; Liu & Negrón Juárez, 2001; Lloyd-Hughes & Saunders, 2002; Manatsa et al., 2017; Wu et al., 2009). Also, most of these studies have not systematically evaluated prediction uncertainty, limiting assessment of predictability from a signal versus noise perspective and the use of the findings in an operational setting.

The probabilistic prediction framework we present integrates several complementary elements. First, we use copula modelling to enable prediction across the full distribution of precipitation, including dry and wet tails, and to provide quantitative estimates of uncertainty. Second, we use principal component analysis as a dimensionality reduction technique to provide a compressed representation of the climate index predictors across multiple ocean basins and avoid overfitting. Third, we use cross validation to avoid overestimating predictive skill and ensure



**Figure 1.** The 14 Sea Surface Temperature (SST) predictors (white boxes), averaged in Aug–Oct, used in this study to predict contiguous United States (CONUS) precipitation in the following Nov–Mar (magenta region) from 1967–1968 to 2016–2017. More information about the SST indices is provided at NOAA’s State of the Ocean website (<http://stateoftheocean.osmc.noaa.gov/>) and in Chen et al. (2016).

a rigorous uncertainty assessment. In our analysis, we predict the entire precipitation distribution function for each year in the last five decades over CONUS at a  $1^\circ \times 1^\circ$  spatial resolution. To answer questions (b) and (c) above, we make use of a multi-category  $3 \times 3$  contingency table (an extension of the typical  $2 \times 2$  table) that is not commonly used in seasonal prediction settings, but (as we show) can offer insights about predictability across different precipitation classes. For the expanded contingency table, we also derive a modified version of the critical success index (CSI; Schaefer, 1990) that accounts for the uncertainty in the prediction, and we assess predictability for three separate classes of precipitation totals (dry, normal, wet), which are defined based on a range of probability levels. The asymptotic, theoretical values of the modified CSI statistic under the null hypothesis of no predictive skill are presented, and their finite-sample distributions are obtained using Monte Carlo simulation. Finally, to gain physical insight on the sources of predictability, we investigate which climate indices are the most important predictors of precipitation in predictability hotspots.

The structure of the paper is the following. In Section 2, we describe the spatial domain and period of our analysis, the datasets we used to assess predictability, the copula-based predictive model, the multi-category  $3 \times 3$  contingency table, and the modified CSI. In Section 3, we present the results of our analysis across CONUS and identify hotspots of winter precipitation predictability. In Section 4, we provide context for our main findings and identify directions for future research.

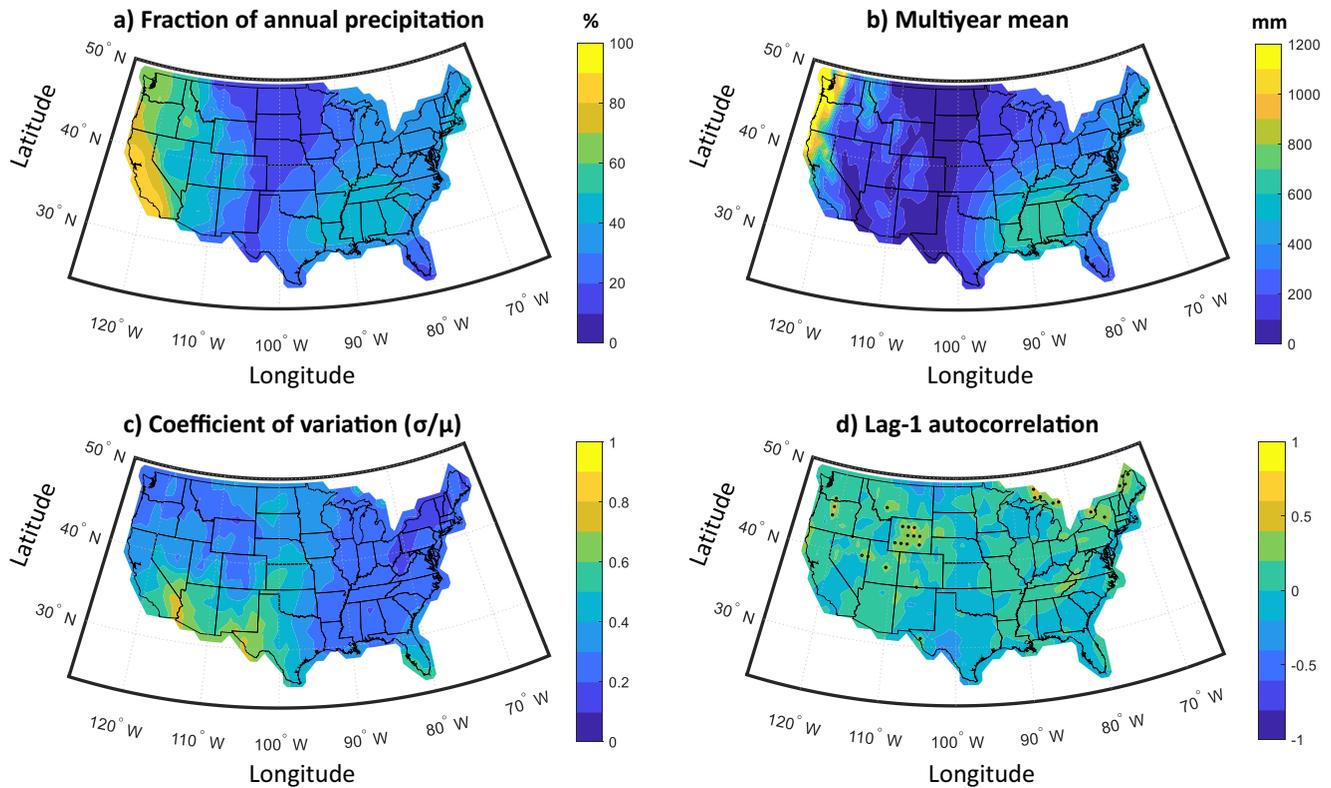
## 2. Prediction Setting and Methodology

### 2.1. Overview and Data Sets

The predictability of winter (Nov–Mar) precipitation totals across the contiguous US (CONUS) is of high practical and scientific interest. Nov–Mar is when most precipitation occurs for many states in the West, while it coincides with the fire season (or its onset) for many states in the Southeast (Li et al., 2020; Nowell et al., 2018). Thus, early and accurate prediction of precipitation is directly related to the ability of practitioners to efficiently manage water resources and mitigate fire risk, with important consequences for agricultural planning, food security and air quality. Indeed, limited prediction skill has led to considerable economic damage in the past (on the order of billions of US dollars), with the recent multi-year California drought in 2012–2016 serving as a prominent example (AghaKouchak et al., 2015; Howitt et al., 2014, 2015; Medellín-Azuara et al., 2016).

To assess predictive skill of the CONUS precipitation, we used 14 SST indices (see white boxes in Figure 1) averaged over Aug–Oct to predict CONUS precipitation totals in the following Nov–Mar (i.e., there was no temporal overlap between the predictors and the predictands). We focused on the 50-year period from 1967–1968 to 2016–2017 to build and test our predictive models. We avoided considering earlier years in the record because

Basic Statistics of Nov-Mar total precipitation from 1967-68 to 2016-17 (50 seasons)



**Figure 2.** Summary statistics of Nov–Mar precipitation over contiguous United States (CONUS), during the years from 1967–1968 to 2016–2017 (50 seasons). (a) Fraction of annual precipitation that occurs during the Nov–Mar period. (b) Multiyear average of Nov–Mar precipitation. (c) Coefficient of variation of Nov–Mar precipitation (defined as the standard deviation divided by the multiyear average). (d) Temporal autocorrelation of the precipitation series for a 1 year lag. Stippling indicates where there is statistically significant autocorrelation ( $p < 0.05$ ).

of the low spatial and temporal resolution of the early SST observations (see Figure 3 in Deser et al., 2010), and because the 1970s is roughly the time that many studies have pinpointed as the start of major changes in climate teleconnections (Johnson et al., 2019; Mamalakis et al., 2018, 2019; Swain et al., 2016; Wang et al., 2014). Both the precipitation and SST time series were linearly detrended before they were used in the analysis, so that long term trends do not impact our assessment.

For our analysis, we used gridded observations at a  $1^\circ \times 1^\circ$  resolution of Nov–Mar precipitation amount (upscaled using spatial averaging from the original  $0.25^\circ \times 0.25^\circ$  resolution), which are freely available from the NOAA Climate Prediction Center (CPC) at <https://psl.noaa.gov/data/gridded/data.unified.daily.conus.html> (Chen et al., 2008). In Figure 2, we show several basic statistics of precipitation across CONUS, during the period of the analysis. For western states, Nov–Mar is when the most precipitation occurs (see Figure 2a), while for many eastern and southeastern states, this season accounts for less than 50% of the total annual precipitation and coincides with the regional fire season (or its onset) (Li et al., 2020; Nowell et al., 2018). This means that for most of the CONUS, predicting Nov–Mar precipitation may aid with effective water planning and/or mitigation of fire risk. On average, the highest precipitation during Nov–Mar occurs over the western states, often exceeding 1,000 mm per season, while the highest year-to-year variability is shown over the southwestern US (Figures 2b and 2c; also see Dettinger et al., 2011; Dettinger & Cayan, 2014). Our results also show that there is no statistically significant temporal autocorrelation in the precipitation series for the majority of CONUS grid points (see Figure 2d). This implies that there is no significant year-to-year dependence, and practically, there is no predictive potential in using the past values of precipitation to predict in the next year.

As predictor variables, we examined various SST indices in three ocean basins (Figure 1). Specifically, we examined Niño 4, Niño 3.4, Niño 3, and Niño 1.2 indices from the Pacific Ocean, TNA, TSA, AMO, and TASI indices

from the Atlantic Ocean, and WTIO, SETIO, and SWIO indices from Indian Ocean. More information about these indices is provided at NOAA's State of the Ocean website (<http://stateoftheocean.osmc.noaa.gov/>) and in Chen et al. (2016). We also considered two recently introduced Pacific indices, that is, the New Zealand Index (NZI; 40°S–25°S and 170°E–200°E) and the East of the Philippines Index (EPI; 5°N–20°N and 130°E–160°E), which have been recently shown to exhibit strong statistical relationship with precipitation over the southwestern US (Mamalakis et al., 2018, 2019). The series of all above indices were constructed using monthly SSTs on a 1° × 1° grid that were obtained from version 2 of the Centennial In Situ Observation-Based Estimates (COBE SST2; <https://www.esrl.noaa.gov/psd/data/gridded/data.cobe2.html>; Hirahara et al., 2014). Additionally, a monthly time series of PDO was directly obtained from the NOAA website <https://www.esrl.noaa.gov/psd/data/climateindices/list>. For our analysis, we used the series of all 14 SST indices (temporally averaged during Aug–Oct), considering them both individually and combined, while we also built predictive models using their principal components (PCs). This allowed us to retain most of the SST variability (e.g., the first and second PCs capture 38.5% and 19.7% of the variance of the 14 indices, respectively, i.e., yielding a total of nearly 60%), without considerably increasing the complexity of the predictive model, and thus, avoid overfitting (i.e., when thinking in terms of a linear regression model, using the first two PCs required the estimation of only two regression coefficients, rather than the much larger number of parameters required to access the same level of variance from the set of uncompressed ocean indices). In Figure 3, we present the spatial patterns of the first five PCs. This analysis revealed that each PC originates primarily from a different ocean basin, which helped us to draw physical insight about the different sources of predictability. The first PC characterizes SST variability in the Pacific (and resembles an ENSO pattern), the second and third PCs characterize variability in the Atlantic (resembling the Atlantic meridional and equatorial modes, respectively), while the fourth and fifth PCs capture considerable SST variability in the Indian Ocean (see an Indian Ocean dipole pattern in PC4). Higher order PCs correspond to only 10% of the total SST variability cumulatively, and so they were not further analyzed in our study.

## 2.2. Predictive Model to Assess Predictability

For each year  $t = 1967 - 1968, 1968 - 1969, \dots, 2016 - 2017$ , we used SST predictors to estimate the entire predictive Probability Density Function (PDF) of precipitation; see schematic in Figure 4a. Specifically, if  $f_{Y|X}^t$  is the conditional predictive PDF of the random variable  $Y$  (here, precipitation in Nov–Mar) in year  $t$ , given the random vector  $\mathbf{X}$  ( $M$  predictors in Aug–Oct), then:

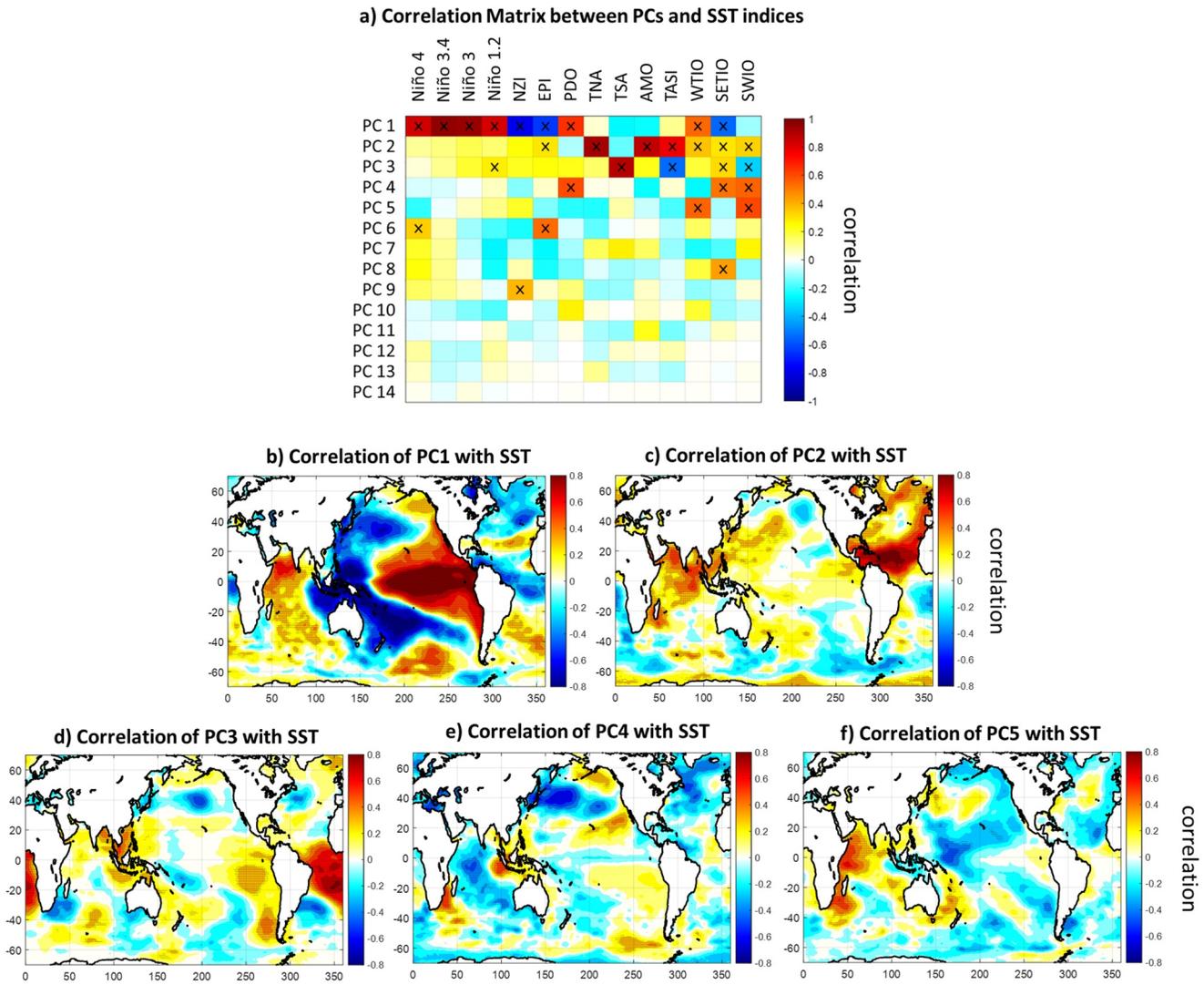
$$f_{Y|X}^t(y) = \frac{f_{Y,X}(y, \mathbf{x}^t)}{f_X(\mathbf{x}^t)} \quad (1)$$

where  $f_{Y,X}$  is the joint PDF of  $Y$  and  $\mathbf{X}$ , and  $f_X$  is the joint PDF of  $\mathbf{X}$ . Since the marginal PDFs of the predictor and the predictand variables are in principle different (e.g., SST vs. precipitation), simple theoretical multivariate distributions (like the multivariate Normal or Gamma distributions) cannot be used to model the joint distributions of  $Y$  and  $\mathbf{X}$ . To address this issue and to keep our approach simple, we used copula functions to represent the above joint distributions (see also Madadgar & Moradkhani, 2013, 2014; Madadgar et al., 2016). Equation 1 becomes (Nelsen, 1999):

$$f_{Y|X}^t(y) = \frac{c_{Y,X}(v, \mathbf{u}^t)}{c_X(\mathbf{u}^t)} f_Y(y) \quad (2)$$

where  $f_Y$  is the marginal PDF of  $Y$ ,  $c_{Y,X}$  and  $c_X$  are the PDFs of the copulas,  $v = \Pr[Y \leq y]$  is the marginal cumulative distribution function (CDF) of  $Y$  evaluated at  $y$ , and  $\mathbf{u}^t = [u_1^t, u_2^t, \dots, u_i^t, \dots, u_M^t]^T$ , where  $u_i^t = \Pr[X_i \leq x_i^t]$  is the marginal CDF of the  $i$ th predictor  $X_i$  evaluated at its value  $x_i^t$ . As previously mentioned, to fit this model we first linearly detrended all series and then fit parametric distributions to the historical precipitation totals and the indices series, to model the marginal distributions. The fitted distributions were Gamma and Gaussian, respectively, both statistically significant at  $\alpha = 0.05$ , based on the Kolmogorov-Smirnov test. To model the joint distributions, we chose to use Gaussian copulas in Equation 2, because the Gaussian copula is among the simplest models to preserve high dimensional dependencies (Joe, 1997).

To ensure that our assessment of the predictability was rigorous, we adopted a 5-fold cross validation approach. Under this setting, the 50-year period was separated into five, non-overlapping and continuous 10-year sets (i.e., the first set included the years [1967 – 1968, 1968 – 1969, ..., 1976 – 1977], the second set included the years



**Figure 3.** Spatial patterns of the principal components (PCs) of the sea surface temperature (SST) indices. (a) Correlation matrix of the PCs with the SST indices. (b)-(f) Correlation maps of the first five PCs with SST. SST series in each grid point corresponds to the average SST in the Aug-Oct season. All series are detrended before the analysis. Symbols “x” in panel (a) and stippling in other panels indicate statistically significant correlation ( $p < 0.05$ ).

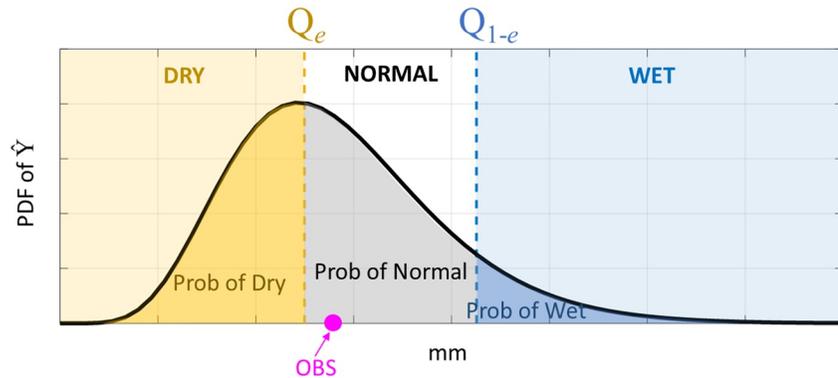
[1977 – 1978, 1978 – 1979, ..., 1986 – 1987], etc.). For each year  $t$  in the first 10-year set, an out-of-sample prediction was made by fitting the predictive model (i.e., fitting the marginal and copula distributions) to the data from the remaining 40 years. This was repeated for the four other 10-year sets, until a single out-of-sample prediction was made for all 50 years. Our performance metrics (described below) were applied to the out-of-sample predictions. Note that the latter is a much stricter approach relative to the leave-one-out cross validation, which can lead to unrealistically high predictive performance (usually from overfitting), especially as the complexity of the model increases (see discussion in Section 3 and DelSole & Banerjee, 2017).

### 2.3. Statistical Metrics of Prediction Skill

In this section we describe the metrics we used to assess the predictability of precipitation across the CONUS. Predictability was assessed in each of the grid points separately, using both continuous and categorical metrics. For the continuous case, we used the coefficient of determination:

$$R^2 = 1 - \frac{\sum_t \left\{ y'_{obs} - \int_0^{+\infty} y f'_{Y|X}(y) dy \right\}^2}{\sum_t \left\{ y'_{obs} - \bar{y}_{obs} \right\}^2} \quad (3)$$

a) Schematic representation of precipitation prediction each year



b) Contingency Table

		OBSERVATION Y		
		$Y \leq Q_e$	$Q_e < Y \leq Q_{1-e}$	$Y > Q_{1-e}$
PREDICTION $\hat{Y}$	$\hat{Y} \leq Q_e$	DRY HIT	FALSE DRY ALARM	WET MISS
	$Q_e < \hat{Y} \leq Q_{1-e}$	DRY MISS	NORMAL HIT	WET MISS
	$\hat{Y} > Q_{1-e}$	FALSE WET ALARM		WET HIT

**Figure 4.** (a) Schematic of the predictive probability density function (PDF) of precipitation for each year. The prediction is made using Equation 2. (b) A  $3 \times 3$  contingency table including all possible outcomes when predicting precipitation.  $Q_e$  denotes the  $e$ th quantile of the historical precipitation distribution.

where  $\overline{y_{obs}}$  is the average observed precipitation in the corresponding grid point, and the term  $\int_0^{+\infty} y f'_{Y|X}(y) dy$  is equivalent to the least-squares prediction and represents the expected value of the conditional predictive distribution  $f'_{Y|X}$  that comes from Equation 2. In linear settings as in our study and when predicting within sample,  $R^2$  is equal to the square of the Pearson correlation between the observations and the predictions and is thus a non-negative number. However, when predicting out-of-sample (e.g., in a cross validation setting), overfitting the data may lead to very poor predictions (e.g., worse than a climatology-based prediction), in which case  $R^2$  can be negative (see Section 3).

For the categorical case, over each grid point and for a probability of nonexceedance  $e$ , we defined historical “dry conditions” as the Nov-Mar seasons for which the precipitation total did not exceed the  $e$ -quantile  $Q_e$  of the historical probability distribution. Similarly, historical “wet conditions” correspond to Nov-Mar seasons when the precipitation total was above the historical  $Q_{1-e}$ . Note that as described below in our analysis, we assessed the prediction skill for several different levels of extremes by allowing  $e$  to vary.

Based on these definitions, we created a table with all possible outcomes with regard to correctly or incorrectly predicting dry, normal, or wet conditions (see Figure 4b). The latter is known as “the contingency table” (Wilks, 2006), and has been widely used in many hydrologic and weather forecasting applications, to assess the

skill of correctly predicting/monitoring events in different spatiotemporal scales (Behrangi et al., 2011; Haile et al., 2012; Hao et al., 2013). A typical contingency table consists of four possible outcomes (hit, miss, false alarm, and correct rejection), which originate from predicting two possible categorical states: The occurrence or lack of occurrence of an event. In our case, we had three states (classes), that is, dry, normal, or wet totals of seasonal precipitation, thus yielding a multi-category  $3 \times 3$  contingency table. Multi-category contingency tables have been discussed in the past (Brooks & Doswell, 1996; Iglesias et al., 2014; Murphy & Winkler, 1987; Zhou et al., 2011), and our study highlights their utility for identifying asymmetries in predictability of different classes (see Section 3).

Many different metrics to assess predictive skill can be defined, even when considering only two classes (Agha-Kouchak & Mehran, 2013; Wilks, 2006), however, the probabilities of a hit (known as probability of detection; POD) and of a false alarm (known as the false alarm ratio; FAR) are most commonly used in the literature (Anagnostou et al., 2010; Gourley et al., 2012; Hao et al., 2013). Here we used the critical success index (CSI), which combines the latter two metrics, as follows (Schaefer, 1990):

$$\text{CSI} = \frac{1}{\frac{1}{\text{POD}} + \frac{1}{1 - \text{FAR}} - 1} = \frac{1}{\frac{1}{\text{POD}} + \frac{1}{\text{SR}} - 1} \quad (4)$$

where  $\text{SR} = 1 - \text{FAR}$  is the success ratio. The CSI is unitless and ranges from 0 (corresponds to no prediction skill) to 1 (corresponds to perfect prediction skill). The POD and the SR of dry/normal/wet conditions or in general of a year when precipitation belongs to a set of outcomes  $A$  are, respectively:

$$\begin{aligned} \text{POD} &= \Pr[\hat{Y} \in A | Y \in A] \\ \text{SR} &= \Pr[Y \in A | \hat{Y} \in A] = \frac{\text{POD} * \Pr[Y \in A]}{\Pr[\hat{Y} \in A]} \end{aligned} \quad (5)$$

where  $Y$  is the observed precipitation value,  $\hat{Y}$  is the prediction, and the last equation arises from Bayes' theorem. In other words, in both metrics one considers the number of correct predictions or hits of an event (e.g., dry conditions), and divides it by the number of total observed events in POD and by the number of total predicted events in SR (Schaefer, 1990).

When incorporating the uncertainty of the prediction (as in the present study), the POD and SR cannot be trivially calculated as described above, but rather need to be based on the integration of the probability distribution between thresholds that define different precipitation classes (e.g., the quantiles  $Q_e$  and  $Q_{1-e}$ ). Specifically, by inserting Equation 5 into 4, we derived the modified dry, wet and normal CSI metrics to assess the predictability of dry, wet and normal conditions, respectively. We examined different values of  $e$  (i.e., 0.25, 0.333), and for each case, we calculated the values of the metrics in the period from 1967–1968 to 2016–17.

$$\text{Dry CSI} = \left[ \frac{1}{\Pr[\hat{Y} \leq Q_e | Y \leq Q_e]} + \frac{1}{\frac{\Pr[\hat{Y} \leq Q_e | Y \leq Q_e] \Pr[Y \leq Q_e]}{\Pr[\hat{Y} \leq Q_e]}} - 1 \right]^{-1} \quad (6)$$

The term  $\Pr[\hat{Y} \leq Q_e | Y \leq Q_e]$  corresponds to the probability of the predictive model detecting an actual drought (i.e., dry POD), and it can be easily calculated as the integral  $\int_0^{Q_e} f_{Y|\hat{Y}}^{\text{Dry}}(y) dy$ , where  $f_{Y|\hat{Y}}^{\text{Dry}}$  is the composite PDF of all predictive PDFs during the actual dry Nov-Mar seasons within the period from 1967–1968 to 2016–2017. The term  $\Pr[\hat{Y} \leq Q_e]$  represents how often the predictive model predicts dry conditions. Lastly, note that the term  $\Pr[Y \leq Q_e]$  is by definition equal to  $e$ .

We similarly defined:

$$\text{Wet CSI} = \left[ \frac{1}{\Pr[\hat{Y} > Q_{1-e} | Y > Q_{1-e}]} + \frac{1}{\frac{\Pr[\hat{Y} > Q_{1-e} | Y > Q_{1-e}] \Pr[Y > Q_{1-e}]}{\Pr[\hat{Y} > Q_{1-e}]}} - 1 \right]^{-1} \quad (7)$$

$$\text{Normal CSI} = \left[ \frac{1}{\Pr[Q_e < \hat{Y} \leq Q_{1-e} | Q_e < Y \leq Q_{1-e}]} + \frac{1}{\frac{\Pr[Q_e < \hat{Y} \leq Q_{1-e} | Q_e < Y \leq Q_{1-e}] \Pr[Q_e < Y \leq Q_{1-e}]}{\Pr[Q_e < \hat{Y} \leq Q_{1-e}]} - 1 \right]^{-1} \quad (8)$$

Apart from the CSI for each of the different classes of precipitation (dry, normal, wet), we also introduced a summarizing statistical metric which considers the skill with regard to all classes in combination. This statistic represents, in logarithmic scale, the probability assigned by the considered model to the correct (observed) precipitation class per year, that is, the model-based likelihood of the observed class per year. The categorical log-likelihood of the data is defined as:

$$\text{Log}L = \frac{1}{N} \text{Log} \left[ \prod_t \left( I_D^t \int_0^{Q_e} f_{Y|X}^t(y) dy + I_N^t \int_{Q_e}^{Q_{1-e}} f_{Y|X}^t(y) dy + I_W^t \int_{Q_{1-e}}^{+\infty} f_{Y|X}^t(y) dy \right) \right] \quad (9)$$

where  $N$  is the number of years (in our case  $N = 50$ ). The three integrals in Equation 9 represent the probability of dry, normal, and wet conditions in year  $t$ , as estimated from the model given  $\mathbf{x}^t$ . For any  $t$ , the coefficients (indicator functions) of the integrals indicate whether the year was actually dry, normal, or wet:

$$I_D^t = \begin{cases} 1, & 0 \leq y_{obs}^t \leq Q_e \\ 0, & \text{otherwise} \end{cases}, \quad I_N^t = \begin{cases} 1, & Q_e < y_{obs}^t \leq Q_{1-e} \\ 0, & \text{otherwise} \end{cases}, \quad I_W^t = \begin{cases} 1, & y_{obs}^t > Q_{1-e} \\ 0, & \text{otherwise} \end{cases}$$

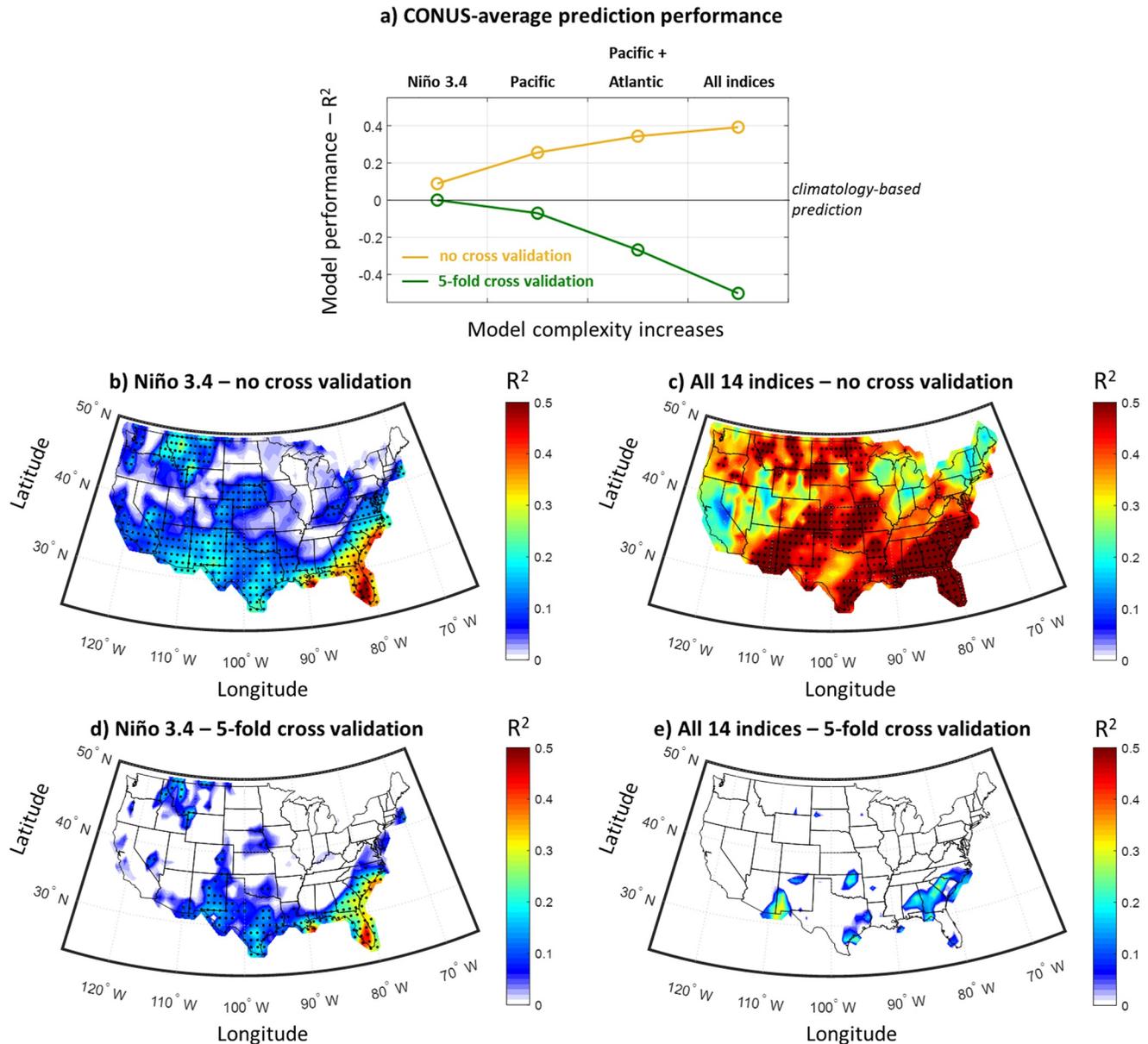
To be able to assess statistical significance when using the above metrics, we defined the null hypothesis,  $H_0$ : “The predictand  $Y$  is independent of the predictors  $X$ , and thus,  $X$  are non-informative for prediction”.

Under  $H_0$ , the copula ratio in Equation 2 is equal to unity and the predictive distribution of precipitation for any  $t$  and for any grid point is equal to the historical marginal PDF of precipitation,  $f_{Y|X}^t(y) = f_Y(y)$ ; that is, predictions under  $H_0$  are based on the “baseline climatology” of  $Y$ . As such, the asymptotic ( $N \rightarrow \infty$ ) values of the considered statistical metrics under  $H_0$ , can be easily derived:

$$\begin{aligned} R_{H_0}^2 &= 0 \\ \text{Dry CSI}_{H_0} &= \frac{e}{2-e} \\ \text{Wet CSI}_{H_0} &= \frac{e}{2-e} \\ \text{Normal CSI}_{H_0} &= \frac{1-2e}{1+2e} \end{aligned}$$

$$\text{Log}L_{H_0} = 2e \text{Log}[e] + (1-2e) \text{Log}[1-2e]$$

However, because of the limited sample size,  $N = 50$ , the asymptotic values cannot be used to assess significance, that is, to reject the  $H_0$ . In other words, for a predictive model with no true predictive skill, the value of a metric may randomly result in a value that is higher than its asymptotic value. Thus, we used the  $F$  statistic for the  $R^2$  and Monte Carlo simulations to obtain the critical values of the rest of the statistics, for a significance level of  $\alpha = 0.05$ . Particularly, we simulated 5,000 pairs of independent  $Y$  and  $X$  series (of sample size  $N = 50$ ), and determined the empirical 95%-quantile for each of the above categorical statistics and for different definitions of the extremes ( $e$  values), which we then used to assess significance.



**Figure 5.** Performance of seasonal prediction models for contiguous United States (CONUS) precipitation using the 14 SST indices. (a) Coefficient of determination ( $R^2$ ) of the prediction mean for increasingly complex predictive models, with and without applying a 5-fold cross validation technique. The CONUS-average  $R^2$  is presented. (b) Coefficient of determination of the prediction mean across CONUS when using the Niño 3.4 index, and without applying a 5-fold cross validation technique. Stippling indicates statistical significance ( $p < 0.05$ ) as assessed using the  $F$  statistic. (c) Same as in (b), but when using all 14 SST indices. (d and e) Same as in (b and c), but when applying a 5-fold cross validation technique. Results in this figure clearly show the overestimation of the predictive skill when not applying a cross validation technique, and the drastic effect of overfitting when the model complexity increases.

### 3. Results

We start by assessing the predictive skill when using the 14 SST indices to predict precipitation across the CONUS and under two settings: When no cross validation is performed and when a 5-fold cross validation is performed (Figure 5). For our assessment, we used the coefficient of determination ( $R^2$ ) between the mean of the prediction and the observations. Our results show that when no cross validation is performed, adding SST indices in the predictive model and increasing model complexity leads to a much better predictive skill (Figures 5a–5c). In fact, when using all 14 indices, precipitation predictability seems very high over the entire CONUS, with a CONUS-average  $R^2$  on the order of 0.4 (i.e., the predictive model explains 40% of precipitation variability, on average). However, when applying a 5-fold cross validation technique that resembles an out-of-sample prediction

setting, predictive performance is much lower, highlighting the effect of overfitting (Figures 5a, 5d and 5e). Even when using a single SST index for prediction (e.g., the Niño 3.4 index), the cross validation technique leads to lower estimate of predictability (compare panels d and b), which highlights the importance of applying cross validation techniques to avoid overestimating predictability (see also DelSole & Banerjee, 2017). Thus, for the remainder of the study we only present results where the 5-fold cross validation has been implemented. In the cross validation case and when only using the Niño 3.4 index, one can see some regions mainly in the Southwest and Southeast where there is significant predictive skill, specifically reaching almost  $R^2 = 0.5$  over Florida (Figure 5d). When adding all Pacific indices, the predictive skill slightly decreases (Figure 5a), and continues to decrease when all 14 indices are used for prediction (Figure 5e), illustrating the increasing effect of overfitting that occurs when more predictors are added into the model.

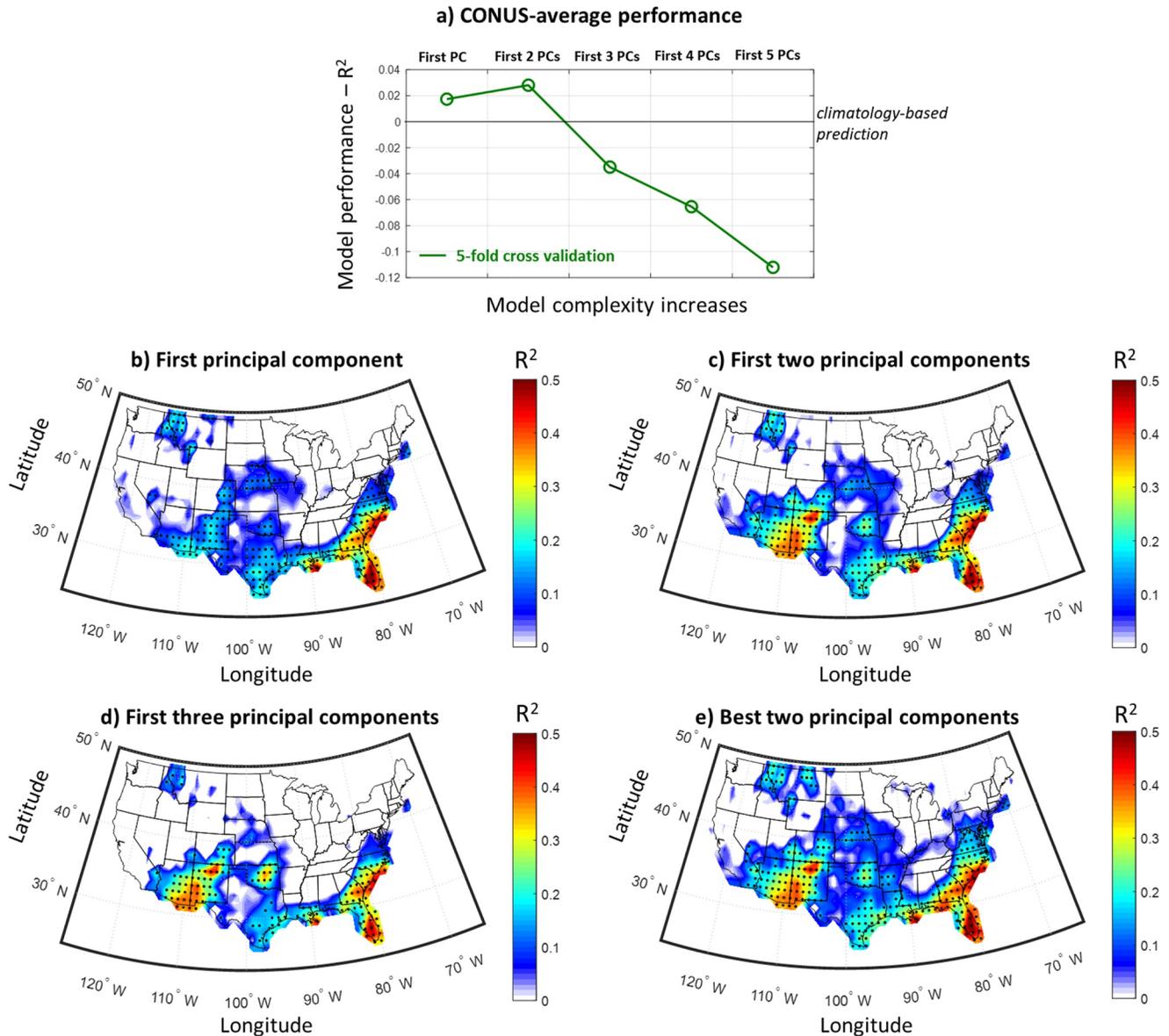
To limit the effect of overfitting, we subsequently examined the use of PCs of the SST indices as predictors of precipitation. Since the PCs are compressed representations of the information contained in the SST indices, their use allowed us to retain most of the SST information, without considerably increasing the model complexity. Our results show that the predictive skill slightly increases when using the first two PCs instead of only the first PC (see Figure 6 panel a, and compare panels b and c) and it decreases when using the first three PCs (panel d). When we added more components as predictors, the predictive skill continued to decrease (panel a). This result implies that for our study design examining seasonal precipitation predictability using a sample size of 50 years, models with more than two predictors appear prone to overfitting. To investigate the upper limit of prediction skill, we considered all possible combinations of predictive models using two out of the first five PCs (i.e.,  $5 \times 4/2 = 10$  combinations). In Figure 6e, we present the  $R^2$  of the prediction for the best performing combination, for each grid point separately. This figure shows that predictive skill increases slightly in some regions compared to panel (c), including western Texas, Arkansas and Montana.

Independent of the predictive model, Figure 6 shows that there are several coherent regions where precipitation is most predictable. Hotspots with relatively high levels of predictability are (a) Arizona and New Mexico in the Southwest with  $R^2$  ranging from 0.2 to 0.4 and (b) Florida, Georgia, South Carolina, North Carolina and Virginia in the Southeast with  $R^2$  ranging from 0.2 to 0.5. Hotspots with moderate levels of predictability include (a) Texas, Louisiana, and Oklahoma across the South with  $R^2$  mostly on the order of 0.2, and (b) northern Idaho and western Montana in the Northern Rockies with  $R^2$  on the order of 0.2. With our approach, we were unable to identify high levels of predictability for most of the Northeast and Midwest and for coastal states in the West.

To better understand the predictability of extremes, we then examined the skill in predicting the occurrence of dry, normal and wet conditions, using the models that combine the best two PCs in each grid point (i.e., the suite of predictive models shown in Figure 6e). We present the dry and wet CSI for different values of  $e$  (0.25 or 0.333, that is, dry and wet conditions correspond to the lowest and highest 1/4 or 1/3 of the historical distribution) across the CONUS in Figure 7. Our results indicate that over the hotspots of predictability identified above, both dry and wet conditions are significantly predictable, independent of the value of  $e$ . Particularly, in the Southwest and Southeast, the prediction exhibits a CSI on the order of 30%–35%, for  $e = 0.333$  (assessed as statistically significant;  $p < 0.05$ ). Over the Northern Rockies region, the CSI is on the order of 30%. Moreover, considering the full spatial pattern of predictability shown in Figure 7, results indicate that there are some regions where the predictability exhibits an asymmetry, that is, wet conditions are more predictable than dry or vice versa. For example, wet conditions seem to be more predictable relative to dry conditions over the Southeast (most notably in Florida) and in Nebraska, while dry conditions are relatively more predictable over western Texas, Oklahoma, and Idaho.

Analysis of the predictability of normal conditions (Figure 8) shows that normal conditions are much less predictable than dry/wet conditions, and that the prediction skill is for most regions statistically indistinguishable from the skill that corresponds to a climatology-based prediction (see results considering either value for  $e$  in Figures 8a and 8c). Intuitively, this result of virtually no predictability for normal conditions is consistent with the concept of forecasts of opportunity, which asserts that there are time windows of little or no predictability, when large-scale climate oscillations (i.e., potential drivers of regional hydroclimate) are inactive. On the contrary, when the climate oscillations become active (e.g., with ENSO moving into either a strong El Niño or La Niña phase), atmospheric forcing and teleconnections may be stronger, providing a more robust signal to mid-latitude precipitation that is less likely to be masked by internal noise, which, in turn, may enhance predictability on seasonal timescales for precipitation extremes (see also Mayer & Barnes, 2021). The maps of the categorical log-likelihood in Figures 8b and 8d clearly highlight the hotspots of predictability that were identified earlier. On

Prediction of winter precipitation totals over CONUS

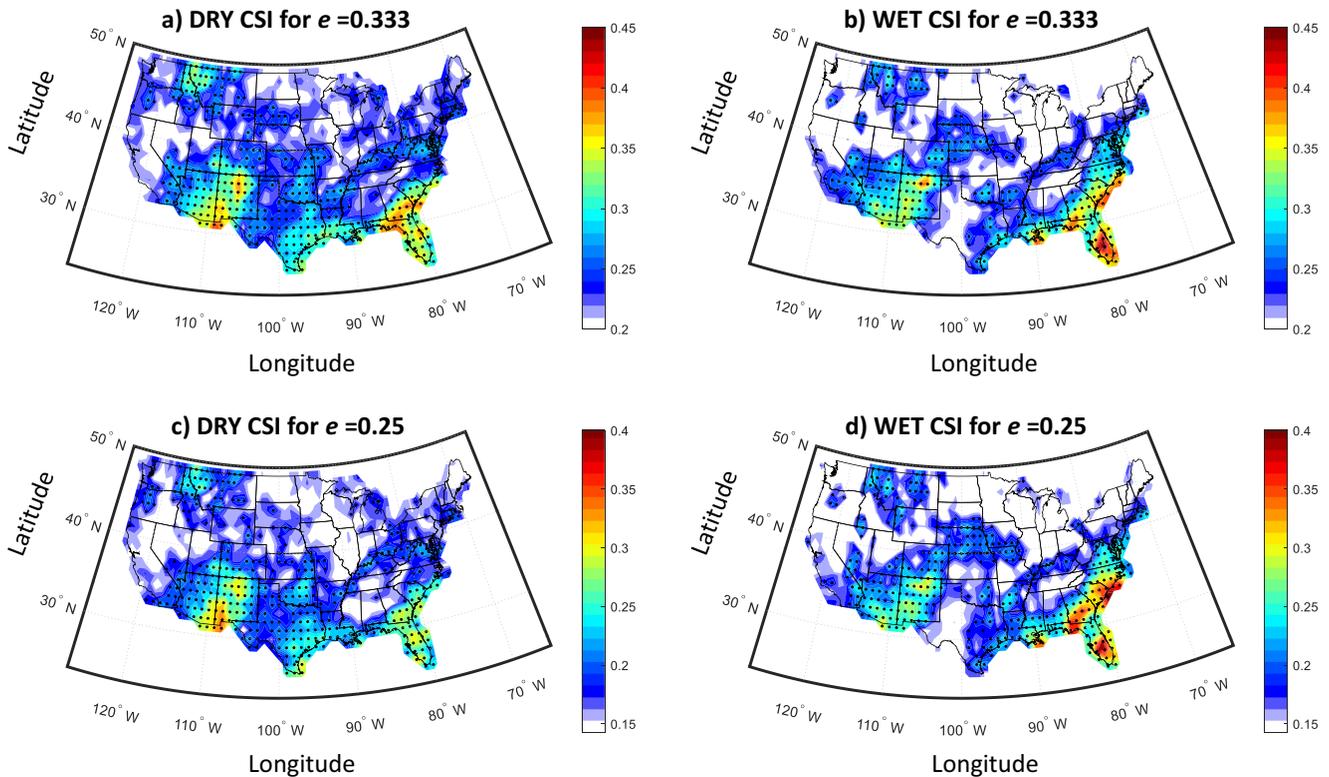


**Figure 6.** Performance of seasonal prediction models using the PCs of the climate indices. (a) Coefficient of determination ( $R^2$ ) of the prediction mean for increasingly complex predictive models, using the indices' principal components (PCs) and applying a 5-fold cross validation technique. The CONUS-average  $R^2$  is presented. (b)  $R^2$  of the prediction mean across contiguous United States (CONUS) when using the first PC as a predictor. Stippling indicates statistical significance ( $p < 0.05$ ) as assessed using the  $F$  statistic. (c) Same as in (b), but when using the first two PCs as predictors. (d) Same as in (b), but when using the first three PCs as predictors. (e) Same as in (b), but when using all possible combinations of two predictors selected from the first five PCs and presenting the performance of the suite of models that lead to the best prediction. Results in this figure highlight predictability hotspots over the Southwest (Arizona, New Mexico), the Southeast (Florida, Georgia, South and North Carolina, Virginia), the South (u) and the northern Rockies (Idaho, Montana).

average, results show that predictive models assign a probability of 0.35–0.4 to the correct class of precipitation over these regions. Note that despite being statistically significant, the latter predictive skill is limited, thus, new methods for improving prediction are warranted (see the discussion in the concluding section).

In a final step, we explored the sources of predictability. Specifically, for each grid point, we created a bivariate linear regression model using the two predictors identified in Figure 6e. We present the PCs that received the highest (lowest) regression coefficient in terms of absolute values in Figure 9a (Figure 9b). Our results show that on average PC1 and PC2 are contributing the most to predictability. This means that precipitation variability

Prediction of dry and wet winters over CONUS



**Figure 7.** Predictability of dry and wet winters using the suite of predictive models in Figure 6e (a) Dry Critical Success Index (CSI) across contiguous United States (CONUS) for  $e = 0.333$  (the dry and wet conditions correspond to the lowest and highest 1/3 of the distribution). (b) Same as in (a), but predicting wet winters. (c and d) Same as in (a and b), but for  $e = 0.25$  (the dry and wet conditions correspond to the lowest and highest 1/4 of the distribution). In all panels, stippling indicates statistical significance ( $p < 0.05$ ).

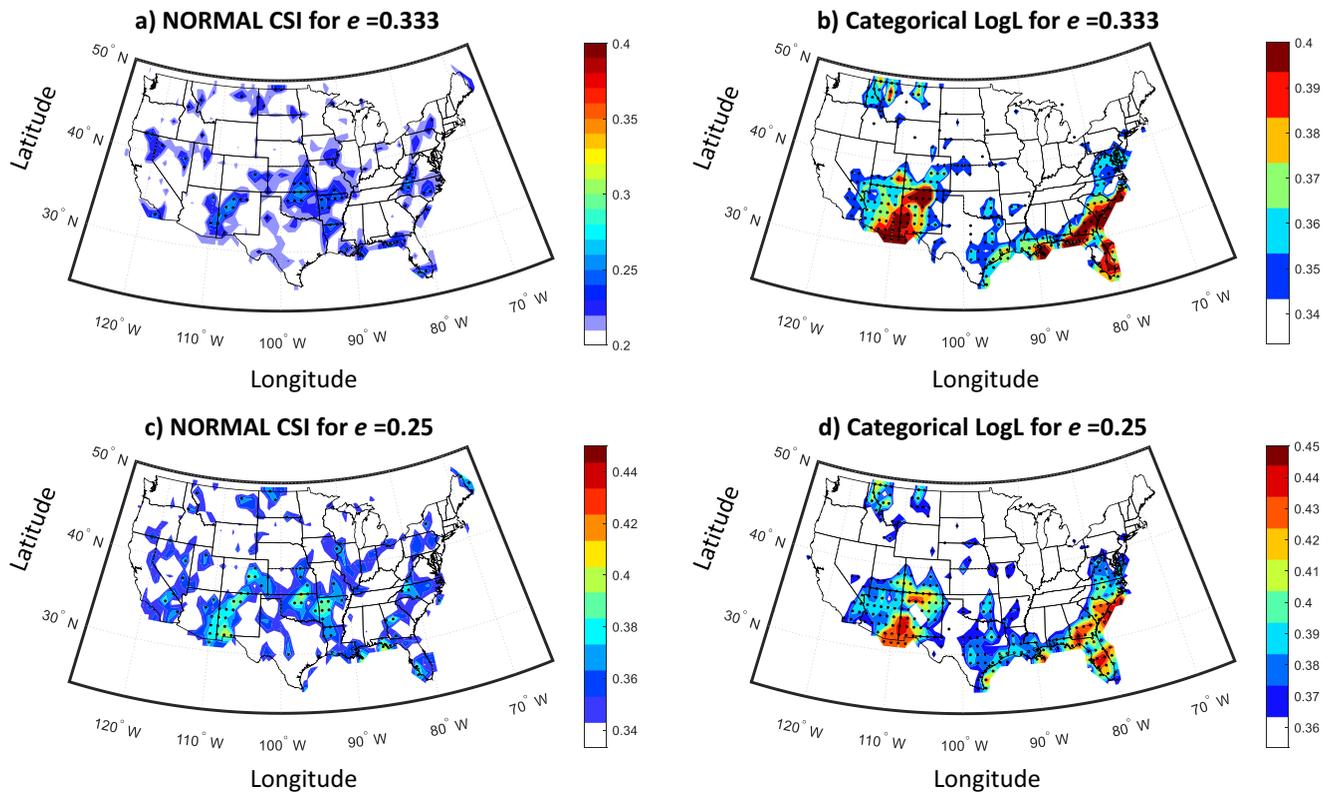
across the CONUS is linked with SST over the Pacific basin (PC1) and over the North Atlantic basin (PC2), which is consistent with past literature, for example, see Dai (2013); Enfield et al. (2001); McCabe et al. (2004); Newman et al. (2016). Specifically, over all predictability hotspots that were identified herein, Pacific Ocean SST is the most important predictor of precipitation. Over the Southwest, the Atlantic Ocean SST is also an important contributor to precipitation predictability; we find that the regression coefficients of PC1 and PC2 are almost of the same order. Over the Northern Rockies and the Southeast, the second most important contributor is PC4 (Figure 9b), although the regression coefficient with this PC is relatively small (not shown) in both regions. This means that over these regions, the Pacific SST is the main source of predictability.

#### 4. Discussion and Concluding Remarks

In this study, we present a rigorous probabilistic framework for assessing precipitation (or other climate variables) predictability under uncertainty. Our specific aim has been to gain insight about hotspots of predictability, which are defined as regions where precipitation is inherently more predictable than in surrounding regions. Our framework uses copula modeling to estimate the entire predictive distribution and quantify uncertainty, and principal component analysis and a cross validation technique to reduce dimensionality and avoid the impact of overfitting. In our analysis, we also evaluate the predictability of dry, normal and wet conditions using a multi-category  $3 \times 3$  contingency table, a modified critical success index, and a categorical loglikelihood evaluation metric.

We applied our framework focusing on the problem of seasonal prediction of precipitation over CONUS. We used SST information that is averaged during late boreal summer and early fall (Aug–Oct) to assess predictability of CONUS precipitation during the following Nov–Mar – a time of the year with elevated fire risk for southeastern states and when most precipitation occurs over the western states. Our results indicate that Nov–Mar precipitation

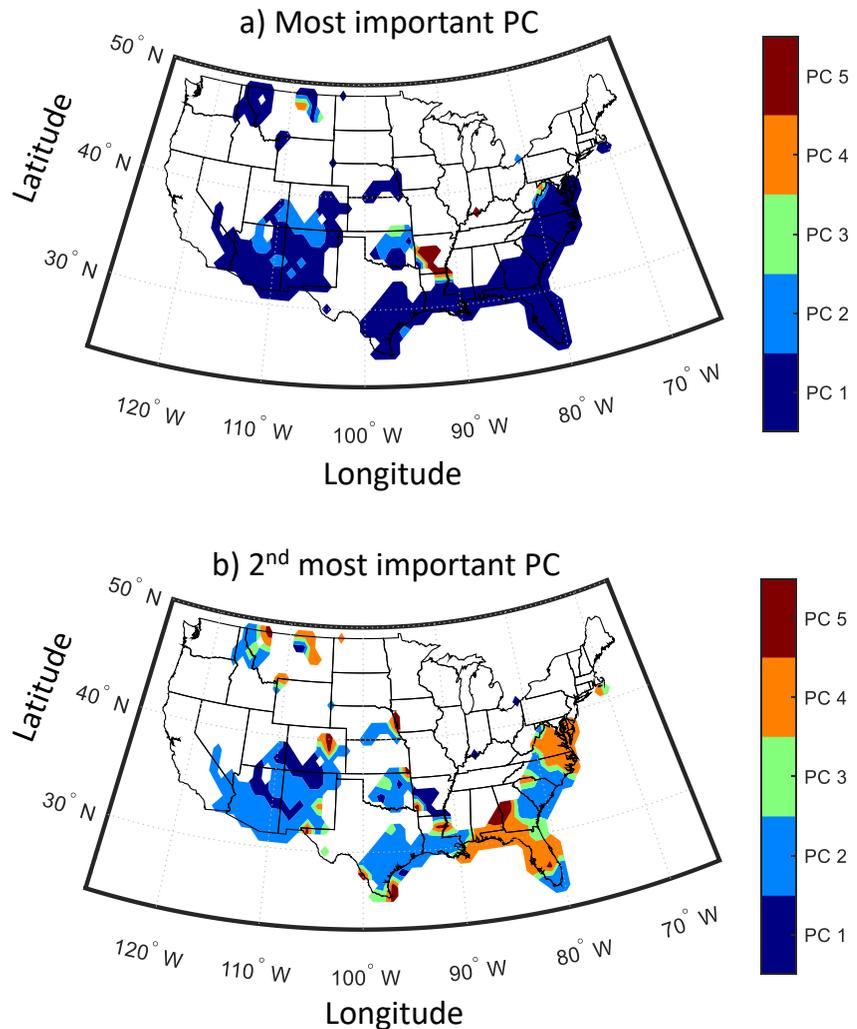
Prediction of normal winters and categorical log-likelihood over CONUS



**Figure 8.** Predictability of normal winters and categorical log-likelihood using the suite of predictive models in Figure 6e (a) Normal Critical Success Index (CSI) across contiguous United States (CONUS) for  $e = 0.333$  (the dry and wet conditions correspond to the lowest and highest 1/3 of the distribution). (b) Same as in (a), but the categorical log-likelihood across CONUS is presented. (c and d) Same as in (a and b), but for  $e = 0.25$  (the dry and wet conditions correspond to the lowest and highest 1/4 of the distribution). In all panels, stippling indicates statistical significance ( $p < 0.05$ ).

totals are most predictable for several states in the Southwest, including Arizona and New Mexico, and across several states in the Southeast, including Florida, Georgia, South Carolina, North Carolina, and Virginia. Additional hotspots with moderate levels of predictability occur in the Northern Rockies and in southern states, including Texas, Oklahoma, and Louisiana. We find that SST information can be used to predict the occurrence of dry and wet conditions, while normal conditions are more difficult to predict. This result illustrated the advantage of using multi-category contingency tables (not usually employed in the seasonal prediction literature) for identifying asymmetries in predictability of different states. We also establish that the biggest source of precipitation predictability during the last 50 years has been the SST over the Pacific and north Atlantic Oceans. The identified hotspots as well as sources of predictability align reasonably well with previous reports. Specifically, many past studies have identified statistically significant skill in prediction of precipitation totals for the Southwest (Gibson et al., 2021; Liu et al., 2018; Madadgar et al., 2016; Mamalakis et al., 2018; McCabe & Dettinger, 1999; Pan et al., 2019; Schonher & Nicholson, 1989; Stevens et al., 2021; Zhang et al., 2018 among many others), for Gulf Coast and Southeast regions (Becker et al., 2014; Devineni & Sankarasubramanian, 2010a, 2010b; Kirtman et al., 2014), and for precipitation extremes in the Northwest (Gershunov & Cayan, 2003; Zarekarizi et al., 2018). Moreover, previous studies have highlighted the Pacific and North Atlantic Ocean basins as the main sources of precipitation predictability over North America (see e.g., Dai, 2013; Enfield et al., 2001; McCabe et al., 2004; Newman et al., 2016).

Our results defining the spatial pattern of precipitation predictability may help improve water resources management in the US (as well as other regions; Delorit & Block, 2019; Giuliani et al., 2020), by informing the design of operational systems. Moreover, the probabilistic framework introduced herein can be applied not only to historical data (retrospectively), but also to climate model outputs in order to explore possible future changes in the strength of climate teleconnections and predictability, which are important for assessment of climate change impacts on regional hydroclimate.



**Figure 9.** Investigation of sources of predictability. (a) Map of the most important PC (corresponding to the highest regression coefficient) used for prediction in Figure 6e (b) Same as in (a), but for the second most important principle component (PC). In both panels, we present results only for points where the value of  $R^2$  in Figure 6e is statistically significant ( $p < 0.05$ ).

Future extensions of our framework may explore using different marginal distributional models or different copula models (other than the Gaussian) that exhibit tail dependence. It is possible that previously reported asymmetries in teleconnections to the US hydroclimate (Feng et al., 2017; Zhang et al., 2014) are more efficiently modeled with non-Gaussian copulas. Second, although in this study we aimed to limit the effect of non-stationarity by detrending and focusing on the most recent 50 years, non-stationary dependencies and their effect on statistically based predictive skill cannot be totally ruled out. Thus, future work should focus on developing ways to incorporate changes in for example, ENSO teleconnections in a warming climate (Beverley et al., 2021; Fasullo et al., 2018; Perry et al., 2017; Sun et al., 2020; Yan et al., 2020). Last, it may be relatively straightforward to extend our approach to assess predictability of other aspects of the precipitation distribution, even beyond the ones considered here, for example, by defining extremes based on specified magnitudes (instead of quantiles), and/or by using additional statistical metrics, important for risk quantification and water management decisions.

We conclude by highlighting that our results show that predictability is low over many areas on the CONUS west coast (i.e., Washington, Oregon and northern California), for which Nov-Mar is the season when the majority of annual precipitation occurs (see Figure 2a). This low predictability is a consequence of the weak relationship between the predefined SST indices and western US precipitation, and has also been confirmed in work

using dynamical models (see Becker et al., 2014; Kirtman et al., 2014). This low predictability places important constraints on decision making, mitigation of hazards and water resources management. Toward improving predictive skill, future work could use data-driven (going beyond predefined indices) machine learning methods that are designed to account for high dimensionality and spatiotemporal dependencies of predictor variables. Such methods have shown considerable potential (e.g., see data-driven and machine/deep learning methods in DelSole & Banerjee, 2017; Liu et al., 2018; Giuliani et al., 2019; Ham et al., 2019; Stevens et al., 2021; Gibson et al., 2021; Peng et al., 2021 among others), and are suited for investigating seasonal precipitation predictability in nonlinear settings. Further improvements may be possible by combining statistical and dynamical model predictions in a hybrid, post processing setting (Hao et al., 2018; Khajehi et al., 2018; Khajehi & Moradkhani, 2017; Madadgar et al., 2016).

### Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

### Data Availability Statement

The data used in our analysis are all freely available online. Precipitation data over CONUS can be found at <https://psl.noaa.gov/data/gridded/data.unified.daily.conus.html>. SST data are downloaded from <https://www.esrl.noaa.gov/psd/data/gridded/data.cobe2.html>, while PDO series can be found at <https://www.esrl.noaa.gov/psd/data/climateindices/list>. The code that was used for this work can be found at: <https://doi.org/10.5281/zenodo.6316579>.

### Acknowledgments

The authors acknowledge the helpful reviews and constructive comments provided by the Editor and three anonymous reviewers that greatly improved the manuscript. The work of A. Mamalakis was supported by the National Science Foundation under Grant OAC-1934668. E. Foufoula-Georgiou acknowledges support from the National Science Foundation (TRIPODS + grant DMS-1839336 and EAGER grant ECCS-1839441) and by NASA's Global Precipitation Measurement (GPM) Program (Grant 80NSSC19K0684). J.T. Randerson acknowledges support from DOE's Office of Science Regional and Global Modeling Analysis program to the RUBISCO SFA, NASA's IDS, CMS, and MAP programs, and a UCNL Lab Fees grant. Foufoula-Georgiou and Randerson also acknowledge support from California's Strategic Growth Council to the Center for Ecosystem Climate Solutions.

### References

- AghaKouchak, A., Feldman, D., Hoerling, M., Huxman, T., & Lund, J. (2015). Recognize anthropogenic drought. *Nature*, 524(7566), 409–411. <https://doi.org/10.1038/524409a>
- AghaKouchak, A., & Mehran, A. (2013). Extended contingency table: Performance metrics for satellite observations and climate model simulations. *Water Resources Research*, 49(10), 7144–7149. <https://doi.org/10.1002/wrcr.20498>
- Anagnostou, E. N., Maggioni, V., Nikolopoulos, E. I., Meskele, T., Hossain, F., & Papadopoulos, A. (2010). Benchmarking high-resolution global satellite rainfall products to radar and rain-gauge rainfall esti-mates. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4), 1667–1683. <https://doi.org/10.1109/tgrs.2009.2034736>
- Becker, E., van den Dool, H., & Zhang, Q. (2014). Predictability and forecast skill in NMME. *Journal of Climate*, 27(15), 5891–5906. <https://doi.org/10.1175/jcli-d-13-00597.1>
- Behrangi, A., Khakbaz, B., Jaw, T., AghaKouchak, A., Hsu, K., & Sorooshian, S. (2011). Hydrologic evaluation of satellite precipitation products at basin scale. *Journal of Hydrology*, 397(3–4), 225–237. <https://doi.org/10.1016/j.jhydrol.2010.11.043>
- Beverley, J. D., Collins, M., Lambert, F. H., & Chadwick, R. (2021). Future changes to El Niño teleconnections over the north Pacific and North America. *Journal of Climate*, 34(15), 6191–6205. <https://doi.org/10.1175/jcli-d-20-0877.1>
- Brooks, H. E., & Doswell, C. A. (1996). A comparison of measures-oriented and distributions-oriented approaches to forecast verification. *Weather and Forecasting*, 11(3), 288–303. [https://doi.org/10.1175/1520-0434\(1996\)011<0288:acomoa>2.0.co;2](https://doi.org/10.1175/1520-0434(1996)011<0288:acomoa>2.0.co;2)
- Chang, Y., Schubert, S. D., & Suarez, M. J. (2000). Boreal winter predictions with the GEOS-2 GCM: The role of boundary forcing and initial conditions. *Quarterly Journal of the Royal Meteorological Society*, 126(567), 2293–2321. <https://doi.org/10.1256/smsqj.56714>
- Charney, J. G., & Shukla, J. (1981). Predictability of monsoons. *Cambridge Univ. Press*, 99–110.
- Chen, M., Shi, W., Xie, P., Silva, V. B. S., Koucky, V. E., Wayne, H. R., & Janowiak, J. E. (2008). Assessing objective techniques for gauge-based analyses of global daily precipitation. *Journal of Geophysical Research*, 113(D4), D04110. <https://doi.org/10.1029/2007JD009132>
- Chen, Y., Morton, D. C., Andela, N., Giglio, L., & Randerson, J. T. (2016). How much global burned area can be forecast on seasonal time scales using sea surface temperature? *Environmental Research Letters*, 11(4), 045001. <https://doi.org/10.1088/1748-9326/11/4/045001>
- Dai, A. (2013). The influence of the inter-decadal Pacific oscillation on US precipitation during 1923-2010. *Climate Dynamics*, 41(3–4), 6–3646. <https://doi.org/10.1007/s00382-012-1446-5>
- Delorit, J. D., & Block, P. J. (2019). Using seasonal forecasts to inform water market-scale option contracts. *Journal of Water Resources Planning and Management*, 145(5), 04019018. [https://doi.org/10.1061/\(asce\)wr.1943-5452.0001068](https://doi.org/10.1061/(asce)wr.1943-5452.0001068)
- DelSole, T., & Banerjee, A. (2017). Statistical seasonal prediction based on regularized regression. *Journal of Climate*, 30, 1345–1361. <https://doi.org/10.1175/JCLI-D-16-0249.1>
- Deser, C., Alexander, M. A., Xie, S. P., & Phillips, A. S. (2010). Sea surface temperature variability: Patterns and mechanisms. *Annual Review of Marine Science*, 115143. <https://doi.org/10.1146/annurev-marine-120408-151453>
- Dettinger, M. D., & Cayan, D. (2014). Drought and the California Delta – a matter of extremes. *San Francisco Estuary and Watershed Science*, 12(2). <https://doi.org/10.15447/sfews.2014v12iss2art4>
- Dettinger, M. D., Ralph, F. M., Das, T., Neiman, P. J., & Cayan, D. (2011). Atmospheric rivers, floods, and the water resources of California. *Water*, 3(2), 455–478. <https://doi.org/10.3390/w3020445>
- Devineni, N., & Sankarasubramanian, A. (2010a). Improved categorical winter precipitation forecasts through multimodel combinations of coupled GCMs. *Geophysical Research Letters*, 37(24), L24704. <https://doi.org/10.1029/2010GL044989>
- Devineni, N., & Sankarasubramanian, A. (2010b). Improving the prediction of winter precipitation and temperature over the continental United States: Role of ENSO state in developing multimodel combinations. *Monthly Weather Review*, 138(6), 2447–2468. <https://doi.org/10.1175/2009MWR3112.1>

- Enfield, D. B., Mestas-Nuñez, A. M., & Trimble, P. J. (2001). The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophysical Research Letters*, 28(10), 2077–2080. <https://doi.org/10.1029/2000gl012745>
- Fasullo, J. T., Otto-Bliesner, B. L., & Stevenson, S. (2018). ENSO's changing influence on temperature, precipitation, and wildfire in a warming climate. *Geophysical Research Letters*, 45(17), 9216–9225. <https://doi.org/10.1029/2018gl079022>
- Feng, J., Chen, W., & Li, Y. (2017). Asymmetry of the winter extratropical teleconnections in the Northern Hemisphere associated with two types of ENSO. *Climate Dynamics*, 48(7–8), 2135–2151. <https://doi.org/10.1007/s00382-016-3196-2>
- Gershunov, A., & Cayan, D. R. (2003). Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. *Journal of Climate*, 16, 2752–2765. [https://doi.org/10.1175/1520-0442\(2003\)016<2752:hdpfot>2.0.co;2](https://doi.org/10.1175/1520-0442(2003)016<2752:hdpfot>2.0.co;2)
- Gibson, P. B., Chapman, W. E., Altinok, A., Delle Monache, L., DeFlorio, M. J., & Waliser, D. E. (2021). Training machine learning models on climate model outputs yields skillful interpretable seasonal precipitation forecasts. *Communications Earth and Environment*, 2(159). <https://doi.org/10.1038/s43247-021-00225-4>
- Giuliani, M., Crochemore, L., Pechlivanidis, I., & Castelletti, A. (2020). From skill to value: Isolating the influence of end user behavior on seasonal forecast assessment. *Hydrology and Earth System Sciences*, 24(12), 5891–5902. <https://doi.org/10.5194/hess-24-5891-2020>
- Giuliani, M., Zaniolo, M., Castelletti, A., Davoli, G., & Block, P. (2019). Detecting the state of the climate system via artificial intelligence to improve seasonal forecasts and inform reservoir operations. *Water Resources Research*, 55(11), 9133–9147. <https://doi.org/10.1029/2019wr025035>
- Goddard, L., Mason, S. J., Zebiak, S. E., Ropelewski, C. F., Basher, R., & Cane, M. A. (2001). Current approaches to seasonal to interannual climate predictions. *International Journal of Climatology*, 21(9), 1111–1152. <https://doi.org/10.1002/joc.636>
- Gourley, J. J., Erlingis, J. M., Hong, Y., & Wells, E. B. (2012). Evaluation of tools used for monitoring and forecasting flash floods in the United States. *Weather and Forecasting*, 27(1), 158–173. <https://doi.org/10.1175/waf-d-10-05043.1>
- Haile, A. T., Habib, E., & Rientjes, T. (2012). Evaluation of the climate prediction center (CPC) morphing technique (CMORPH) rainfall product on hourly time scales over the source of the Blue Nile river. *Hydrological Processes*, 27(12), 1829–1839. <https://doi.org/10.1002/hyp.9330>
- Ham, Y. G., Kim, J. H., & Luo, J. J. (2019). Deep learning for multi-year ENSO forecasts. *Nature*, 573(7775), 568–572. <https://doi.org/10.1038/s41586-019-1559-7>
- Hao, Z., AghaKouchak, A., & Phillips, T. J. (2013). Changes in concurrent monthly precipitation and temperature extremes. *Environmental Research Letters*, 8(3), 034014. <https://doi.org/10.1088/1748-9326/8/3/034014>
- Hao, Z., Singh, V. P., & Xia, Y. (2018). Seasonal drought prediction: Advances, challenges, and future prospects. *Reviews of Geophysics*, 56(1), 108–141. <https://doi.org/10.1002/2016rg000549>
- Hirahara, S., Ishii, M., & Fukuda, Y. (2014). Centennial-scale sea surface temperature analysis and its uncertainty. *Journal of Climate*, 27(1), 57–75. <https://doi.org/10.1175/jcli-d-12-00837.1>
- Howitt, R. E., MacEwan, D., Medellín-Azuara, J., Lund, J. R., & Sumner, D. A. (2015). *Economic analysis of the 2015 drought for California agriculture* (p. 16). Center for Watershed Sciences, University of California – Davis.
- Howitt, R. E., Medellín-Azuara, J., MacEwan, D., Lund, J. R., & Sumner, D. A. (2014). *Economic analysis of the 2014 drought for California agriculture*. University of California.
- Iglesias, I., Lorenzo, M. N., & Taboada, J. J. (2014). Seasonal predictability of the east atlantic pattern from sea surface temperatures. *PLoS One*, 9(1), e86439. <https://doi.org/10.1371/journal.pone.0086439>
- Joe, H. (1997). *Multivariate models and dependence concepts*. Chapman and Hall.
- Johnson, N. C., L'Heureux, M. L., Chang, C.-H., & Hu, Z.-Z. (2019). On the delayed coupling between ocean and atmosphere in recent weak El Niño episodes. *Geophysical Research Letters*, 46(20), 11416–11425. <https://doi.org/10.1029/2019GL084021>
- Khajehei, S., Ahmadalipour, A., & Moradkhani, H. (2018). An effective post-processing of the North American multi-model ensemble (NMME) precipitation forecasts over the continental US. *Climate Dynamics*, 51(1–2), 457–472. <https://doi.org/10.1007/s00382-017-3934-0>
- Khajehei, S., & Moradkhani, H. (2017). Towards an improved ensemble precipitation forecast: A probabilistic post-processing approach. *Journal of Hydrology (Amsterdam)*, 546, 476–489. <https://doi.org/10.1016/j.jhydrol.2017.01.026>
- Khan, M. Z. K., Sharma, A., & Mehrotra, R. (2017). Global seasonal precipitation forecasts using improved sea surface temperature predictions. *Journal of Geophysical Research: Atmospheres*, 122(9), 4773–4785. <https://doi.org/10.1002/2016JD025953>
- Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., et al. (2014). The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *BAMS*, 95(4), 585–601. <https://doi.org/10.1175/bams-d-12-00050.1>
- L'Heureux, M. L., Tippett, M. K., & Barnston, A. G. (2015). Characterizing ENSO coupled variability and its impact on north American seasonal precipitation and temperature. *Journal of Climate*, 28(10), 4231–4245. <https://doi.org/10.1175/jcli-d-14-00508.1>
- Li, F., Zhang, X., Kondragunta, S., Schmidt, C. C., & Holmes, C. D. (2020). A preliminary evaluation of GOES-16 active fire product using Landsat-8 and VIIRS active fire data, and ground-based prescribed fire records. *Remote Sensing of Environment*, 237, 111600. <https://doi.org/10.1016/j.rse.2019.111600>
- Lindsey, R. (2016). *How El Niño and La Niña affect the winter jet stream and U.S. climate*. Climate.gov. Retrieved from <https://www.climate.gov/news-features/featured-images/how-el-ni-c3-b1o-and-la-ni-c3-b1a-affect-winter-jet-stream-and-us-climate>
- Liu, T., Schmitt, R. W., & Li, L. (2018). Global search for autumn-lead sea surface salinity predictors of winter precipitation in southwestern United States. *Geophysical Research Letters*, 45(16), 8445–8454. <https://doi.org/10.1029/2018GL079293>
- Liu, W. T., & Negrón Juárez, R. I. (2001). ENSO drought onset prediction in northeast Brazil using NDVI. *International Journal of Remote Sensing*, 22(17), 3483–3501. <https://doi.org/10.1080/014311600100006430>
- Lloyd-Hughes, B., & Saunders, M. A. (2002). Seasonal prediction of European spring precipitation from El Niño–southern oscillation and local sea-surface temperatures. *International Journal of Climatology*, 22, 1–14. <https://doi.org/10.1002/joc.723>
- Madadgar, S., AghaKouchak, A., Shukla, S., Wood, A. W., Cheng, L., Hsu, K.-L., & Svoboda, M. (2016). A hybrid statistical-dynamical framework for meteorological drought prediction: Application to southwestern United States. *Water Resources Research*, 52(7), 5095–5110. <https://doi.org/10.1002/2015wr018547>
- Madadgar, S., & Moradkhani, H. (2013). A Bayesian framework for probabilistic seasonal drought forecasting. *Journal of Hydrometeorology*, 14(6), 1685–1705. <https://doi.org/10.1175/jhm-d-13-010.1>
- Madadgar, S., & Moradkhani, H. (2014). Spatio-temporal drought forecasting within Bayesian networks. *Journal of Hydrology*, 512, 134–146. <https://doi.org/10.1016/j.jhydrol.2014.02.039>
- Mamalakis, A., Langousis, A., Deidda, R., & Marrocu, M. (2017). A parametric approach for simultaneous bias correction and high-resolution downscaling of climate model rainfall. *Water Resources Research*, 53(3), 2149–2170. <https://doi.org/10.1002/2016wr019578>
- Mamalakis, A., Yu, J.-Y., Randerson, J. T., AghaKouchak, A., & Foufoula-Georgiou, E. (2018). A new interhemispheric teleconnection increases predictability of winter precipitation in southwestern US. *Nature Communications*, 9(1), 2332. <https://doi.org/10.1038/s41467-018-04722-7>

- Mamalakis, A., Yu, J.-Y., Randerson, J. T., AghaKouchak, A., & Foufoula-Georgiou, E. (2019). Reply: A critical examination of a newly proposed interhemispheric teleconnection to southwestern US winter precipitation. *Nature Communications*, *10*(1), 2918. <https://doi.org/10.1038/s41467-019-10531-3>
- Manatsa, D., Mushore, T., & Lenouo, A. (2017). Improved predictability of droughts over southern Africa using the standardized precipitation index and ENSO. *Theoretical and Applied Climatology*, *127*(1–2), 259–274. <https://doi.org/10.1007/s00704-015-1632-6>
- Mariotti, A., Baggett, C., Barnes, E. A., Becker, E., Butler, A., Collins, D. C., et al. (2020). Windows of Opportunity for Skillful Forecasts Subseasonal and Seasonal and Beyond. *Bulletin of the American Meteorological Society*, *101*(5), E608–E625.
- Mayer, K. J., & Barnes, E. A. (2021). Subseasonal forecasts of opportunity identified by an explainable neural network. *Geophysical Research Letters*, *48*(10), e2020GL092092. <https://doi.org/10.1029/2020gl092092>
- McCabe, G. J., & Dettinger, M. D. (1999). Decadal variations in the strength of ENSO teleconnections with precipitation in the Western United States. *International Journal of Climatology*, *19*(13), 1399–1410. [https://doi.org/10.1002/\(sici\)1097-0088\(19991115\)19:13<1399::aid-joc457>3.0.co;2-a](https://doi.org/10.1002/(sici)1097-0088(19991115)19:13<1399::aid-joc457>3.0.co;2-a)
- McCabe, G. J., Palecki, M. A., & Betancourt, J. L. (2004). Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(12), 4136–4141. <https://doi.org/10.1073/pnas.0306738101>
- Medellín-Azuara, J., MacEwan, D., Howitt, R. E., Sumner, D. A., & Lund, J. R. (2016). *A report for the California department of food and agriculture, Center for watershed Sciences*. University of California – Davis.
- Murphy, A. H., & Winkler, R. L. (1987). A general framework for forecast verification. *Monthly Weather Review*, *115*(7), 1330–1338. [https://doi.org/10.1175/1520-0493\(1987\)115<1330:agffv>2.0.co;2](https://doi.org/10.1175/1520-0493(1987)115<1330:agffv>2.0.co;2)
- National Academies of Sciences. (2016). *Engineering, and medicine, NASEM*. The National Academies Press. <https://doi.org/10.17226/21873>
- Nelsen, R. B. (1999). *An introduction to copulas*. Springer.
- Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., et al. (2016). The Pacific decadal oscillation, revisited. *Journal of Climate*, *29*(12), 4399–4427. <https://doi.org/10.1175/jcli-d-15-0508.1>
- Nowell, H. K., Holmes, C. D., Robertson, K., Teske, C., & Hiers, J. K. (2018). A new picture of fire extent, variability, and drought interaction in prescribed fire landscapes: Insights from Florida government records. *Geophysical Research Letters*, *45*(15), 7874–7884. <https://doi.org/10.1029/2018gl078679>
- Palmer, T., & Anderson, D. (1994). The prospects for seasonal forecasting—a review paper. *Quarterly Journal of the Royal Meteorological Society*, *120*(518), 755–793. <https://doi.org/10.1256/smsqj.51801>
- Pan, B., Hsu, K., AghaKouchak, A., Sorooshian, S., & Higgins, W. (2019). Precipitation prediction skill for the west coast United States: From short to extended range. *Journal of Climate*, *32*(1), 161–182. <https://doi.org/10.1175/JCLI-D-18-0355.1>
- Peng, X., Li, T., & Albertson, J. D. (2021). Investigating predictability of the TRHR seasonal precipitation at long lead times using a generalized regression model with regularization. *Frontiers of Earth Science*, *9*, 724599. <https://doi.org/10.3389/feart.2021.724599>
- Perry, S. J., McGregor, S., Gupta, A. S., & England, M. H. (2017). Future changes to El Niño-southern oscillation temperature and precipitation teleconnections. *Geophysical Research Letters*, *44*(20), 10608–10616. <https://doi.org/10.1002/2017gl074509>
- 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. <https://doi.org/10.1029/91wr00690>
- Sankarasubramanian, A., Lall, U., Souza Filho, F., & Sharma, A. (2009). Improved water allocation utilizing probabilistic climate forecasts: Short-term water contracts in a risk management framework. *Water Resources Research*, *45*(11), W11409. <https://doi.org/10.1029/2009WR007821>
- Schaefer, J. T. (1990). The Critical Success Index as an indicator of warning skill. *Weather and Forecasting*, *5*(4), 570–575. [https://doi.org/10.1175/1520-0434\(1990\)005<0570:tcsiaa>2.0.co;2](https://doi.org/10.1175/1520-0434(1990)005<0570:tcsiaa>2.0.co;2)
- Schonher, T., & Nicholson, S. E. (1989). The relationship between California rainfall and ENSO events. *Journal of Climate*, *2*(11), 1258–1269. [https://doi.org/10.1175/1520-0442\(1989\)002<1258:trbcra>2.0.co;2](https://doi.org/10.1175/1520-0442(1989)002<1258:trbcra>2.0.co;2)
- Schubert, S., Chang, Y., Wang, H., Koster, R., & Suarez, M. (2016). A modeling study of the causes and predictability of the spring 2011 extreme US weather activity. *Journal of Climate*, *29*(21), 7869–7887. <https://doi.org/10.1175/jcli-d-15-0673.1>
- Stevens, A., Willett, R., Mamalakis, A., Foufoula-Georgiou, E., Tejedor, A., Randerson, J. T., et al. (2021). Graph-guided regularized regression of Pacific Ocean climate variables to increase predictive skill of southwestern US winter precipitation. *Journal of Climate*, *34*(2), 737–754. <https://doi.org/10.1175/jcli-d-20-0079.1>
- Sun, Q., Miao, C., Aghakouchak, A., Mallakpour, I., Ji, D., & Duan, Q. (2020). Possible increased frequency of ENSO-related dry and wet conditions over some major watersheds in a warming climate. *Bulletin of the American Meteorological Society*, (pp. E409–E426). <https://doi.org/10.1175/BAMS-D-18-0258.1>
- Swain, D. L., Horton, D. E., Singh, D., & Diffenbaugh, N. S. (2016). Trends in atmospheric patterns conducive to seasonal precipitation and temperature extremes in California. *Science Advances*, *2*(4). <https://doi.org/10.1126/sciadv.1501344>
- Wang, B., Lee, J. Y., Kang, I. S., Shukla, J., Park, C. K., Kumar, A., et al. (2009). Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004). *Climate Dynamics*, *33*(1), 93–117. <https://doi.org/10.1007/s00382-008-0460-0>
- Wang, S., Anichowski, A., Tippett, M. K., & Sobel, A. H. (2017). Seasonal noise versus subseasonal signal: Forecasts of California precipitation during the unusual winters of 2015–2016 and 2016–2017. *Geophysical Research Letters*, *44*(18), 9513–9520. <https://doi.org/10.1002/2017gl075052>
- Wang, S. Y., Hipps, L., Gillies, R. R., & Yoon, J. H. (2014). Probable causes of the abnormal ridge accompanying the 2013–2014 California drought: ENSO precursor and anthropogenic footprint. *Geophysical Research Letters*, *41*(9), 3220–3226. <https://doi.org/10.1002/2014gl059748>
- Wilks, D. (2006). *Statistical methods in the atmospheric Sciences* (2nd ed.). Academic.
- Wu, Z., Wang, B., Li, J., & Jin, F.-F. (2009). An empirical seasonal prediction model of the east Asian summer monsoon using ENSO and NAO. *Journal of Geophysical Research*, *114*(D18), D18120. <https://doi.org/10.1029/2009JD011733>
- Yan, Z., Wu, B., Li, T., Collins, M., Clark, R., Zhou, T., et al. (2020). Eastward shift and extension of ENSO-induced tropical precipitation anomalies under global warming. *Science Advances*, *6*(2), eaax4177. <https://doi.org/10.1126/sciadv.aax4177>
- Zarekarizi, M., Rana, A., & Moradkhani, H. (2018). Precipitation extremes and their relation to climatic indices in the Pacific Northwest USA. *Climate Dynamics*, *50*, 4519–4537. <https://doi.org/10.1007/s00382-017-3888-2>
- Zhang, T., Hoerling, M. P., Wolter, K., Eischeid, J., Cheng, L., Hoell, A., et al. (2018). Predictability and prediction of southern California rains during strong El Niño events: A focus on the failed 2016 winter rains. *Journal of Climate*, *31*(2), 555–574. <https://doi.org/10.1175/JCLI-D-17-0396.1>

- Zhang, T., Perlwitz, J., & Hoerling, M. P. (2014). What is responsible for the strong observed asymmetry in teleconnections between El Niño and La Niña? *Geophysical Research Letters*, *41*(3), 1019–1025. <https://doi.org/10.1002/2013gl058964>
- Zhou, H., Tang, G., Li, N., Wang, F., Wang, Y., & Jian, D. (2011). Evaluation of precipitation forecasts from NOAA global forecast system in hydropower operation. *Journal of Hydroinformatics*. 1318195. <https://doi.org/10.2166/hydro.2010.005>
- Zimmerman, B., Vimont, D., & Block, P. (2016). Utilizing the state of El Niño/southern oscillation as a means for season-ahead precipitation predictor selection. *Water Resources Research*, *52*(5), 3761–3774. <https://doi.org/10.1002/2015WR017644>