

# Geophysical Research Letters®

## RESEARCH LETTER

10.1029/2022GL099572

### Key Points:

- Predictability of monthly global rainfall using hidden states of Sea Surface Temperatures in the tropical Pacific Ocean is demonstrated
- Highly significant skill for many global regions and months is identified by sampling of means and deviations given hidden state
- A sequential application of the method leads to much higher cross-validated correlations than those based on regression on an El Niño Southern Oscillation index

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

M. Zhang,  
mz2689@columbia.edu

### Citation:

Zhang, M., Rojo-Hernández, J. D., Yan, L., Mesa, Ó. J., & Lall, U. (2022). Hidden tropical Pacific sea surface temperature states reveal global predictability for monthly precipitation for sub-season to annual scales. *Geophysical Research Letters*, 49, e2022GL099572. <https://doi.org/10.1029/2022GL099572>

Received 17 MAY 2022

Accepted 29 SEP 2022

### Author Contributions:

**Conceptualization:** Upmanu Lall

**Data curation:** Julián David Rojo-Hernández

**Formal analysis:** Mengjie Zhang

**Investigation:** Julián David Rojo-Hernández

**Methodology:** Mengjie Zhang, Óscar José Mesa, Upmanu Lall

**Resources:** Julián David Rojo-Hernández,

Óscar José Mesa, Upmanu Lall

**Supervision:** Upmanu Lall

**Validation:** Lei Yan

**Visualization:** Lei Yan

**Writing – original draft:** Mengjie

Zhang, Upmanu Lall

**Writing – review & editing:** Julián

David Rojo-Hernández, Lei Yan, Óscar José Mesa, Upmanu Lall

© 2022. American Geophysical Union.  
All Rights Reserved.

## Hidden Tropical Pacific Sea Surface Temperature States Reveal Global Predictability for Monthly Precipitation for Sub-Season to Annual Scales

Mengjie Zhang<sup>1,2</sup> , Julián David Rojo-Hernández<sup>3</sup>, Lei Yan<sup>1,2</sup>, Óscar José Mesa<sup>3</sup> , and Upmanu Lall<sup>1,2</sup>

<sup>1</sup>Columbia Water Center, Earth Institute, Columbia University, New York, NY, USA, <sup>2</sup>Department of Earth and Environmental Engineering, Columbia University, New York, NY, USA, <sup>3</sup>Departamento de Geociencias y Medio Ambiente, Universidad Nacional de Colombia, Medellín, Colombia

**Abstract** We present the first global precipitation predictability estimates corresponding to the recently discovered flavors of El Niño Southern Oscillation (ENSO) that are encoded in the hidden states of Tropical Pacific sea surface temperatures identified using a non-homogeneous hidden Markov model. For each calendar month and for each hidden state, we assess future precipitation predictability through the conditional standardized anomaly of the average and the standard deviation of monthly precipitation, at 1, 3, 6, 9, and 12 months lead times. We find statistically significant potential predictive skill for key regions for each hidden state and calendar month, even for 12-months in the future. We apply the algorithm sequentially over the period of record to identify regions that can be consistently predicted for different lead times and calendar months. The cross-validated correlation skill is demonstrably superior to that of regression with an ENSO index used in the same way.

**Plain Language Summary** El Niño Southern Oscillation (ENSO) information is routinely used by hydrologists, agriculturalists and others as a useful prognosticator of upcoming precipitation. Since the characterization of ENSO has recently been broadened to consider different “flavors,” the NINO3.4 correlations used by many people may not be the most informative. Our machine learning model identified five hidden states of the tropical Pacific sea surface temperatures that correspond to the different ENSO flavors. We explore how future monthly precipitation may be predictable knowing which hidden state is currently identified and show that this leads to a higher predictability for up to almost a year in the future for many more calendar months in more regions of the world than NINO3.4 correlations.

## 1. Introduction

Sub-seasonal to annual precipitation forecasts are of strong interest for hydrological and agricultural planning and outcomes. The El Niño Southern Oscillation (ENSO) is the strongest interannual signal in the climate system. It develops around the tropical Pacific Ocean and is the most important mode of global climate variability (McPhaden et al., 2006; Nigam & Sengupta, 2021). There is an extensive literature on the teleconnections of ENSO to regional precipitation including future predictability (Arcodia et al., 2020; Chapman et al., 2021; Dai & Wigley, 2000; Emerton et al., 2017; Henderson et al., 2020; Ropelewski & Halpert, 1987; Sun et al., 2015; Tseng et al., 2021; Van Oldenborgh & Burgers, 2005; Vicente-Serrano et al., 2011; Yan et al., 2021; Yang et al., 2021; Zhang et al., 2016).

Spatial patterns of sea surface temperature (SST) are key to each ENSO event's evolution and impacts. The NINO1.2 and NINO3.4 regions have been used most often for correlative teleconnection analyses, and also for benchmarking the performance of physics-based models in reproducing ENSO dynamics and the associated teleconnections (Lenssen et al., 2020; Nigam & Sengupta, 2021). Recent ENSO events have exhibited strong anomalies in SST regions in the Central Tropical Pacific that are different from the traditional ones. Consequently, the teleconnections have also had a different flavor. Many studies have proposed using the principal components or contrasting different ENSO regions, instead of the well-known NINO3.4 index. In particular, the El Niño Modoki index (Ashok et al., 2007) uses empirical orthogonal functions of SST in the Pacific Ocean. Tang et al. (2018) summarized the current status of ENSO prediction models. They reviewed the progress in ENSO prediction/predictability and the improvement in the theoretical study of the intrinsic predictability limit.

Machine learning methods have been applied for the simulation and forecast of ENSO. Lima et al. (2009) demonstrated ENSO predictability beyond 1 year using Maximum Variance Unfolding with SST and thermocline data. Ham et al. (2019) provided skillful ENSO forecasts for lead times up to 18 months, relying on a convolutional neural network with SST and heat content as predictors. Rojo-Hernández et al. (2020) employed a non-homogeneous hidden Markov model (NHMM) to simulate and predict the spatiotemporal evolution of monthly tropical Pacific SST. Five SST hidden states, whose spatial patterns are similar to the so-called ENSO flavors, were identified using NHMM. This classification of the SST dynamics allows an exploration of the spatial patterns of SST and other meteorological variables associated with each hidden state, and their temporal and Markovian evolution. The SST hidden states can then be employed as concurrent or leading predictors for other variables (e.g., for regional precipitation and temperature), in addition to insights as to the ENSO flavors and their prediction. In this paper, we explore the diagnostic connection of the hidden states with the space and time expression of climate variables, globally or regionally, for the concurrent month or for a future month.

The potential predictability of gridded monthly precipitation globally using the SST hidden state, at lead times of 1–12 months starting in each calendar month, is of primary interest. We also explore the corresponding atmospheric pressure fields in the same way to assess whether there is physical consistency between the indicated precipitation and atmospheric circulation. The data and methods used are outlined in the next section, followed by an overview of the results at specific lead times. Finally, a discussion of the implications and future directions concludes the paper.

## 2. Data and Methods

### 2.1. Data

The global region discarding the polar regions (180°W–180°E, 65°S–75°N) is selected as the study area. The monthly precipitation and 700 hpa geopotential height fields are obtained for 1856–2015 from the NOAA-CIRES-DOE Twentieth Century Reanalysis (V3) data set ([https://psl.noaa.gov/data/gridded/data.20thC\\_ReanV3.html](https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html)). The standardized anomalies of precipitation at time  $t$ , denoted by  $pa_t$ , were calculated for each grid cell as

$$pa_t = \frac{p_t - \bar{p}_{i(t)}}{sp_{i(t)}} \quad (1)$$

where  $p_t$  is the observed monthly precipitation at time  $t$ ,  $\bar{p}_{i(t)}$ , and  $sp_{i(t)}$  are the mean and standard deviation of monthly precipitation for the calendar month  $i(t)$  associated with month  $t$ .

The SST anomalies used are the monthly Kaplan Extended SST V2 (Kaplan et al., 1997) from 1856 to 2015 covering the region 15°S–15°N and 150°E–80°W. The NION3.4 index was downloaded from NOAA ([https://psl.noaa.gov/gcos\\_wgsp/Timeseries/Nino34/](https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino34/)).

### 2.2. Monthly Precipitation Predictability Given the Hidden SST States Information

Through the analysis using the NHMM (see Holsclaw et al., 2017; Rojo-Hernandez et al., 2020, and Supporting Information S1 for details of the construction of NHMM and parameter estimation), five hidden states (Figure S1 in Supporting Information S1) are identified. Hidden states one–five are similar to the classical La Nina pattern, mild La Nina pattern, neutral pattern, Modoki ENSO pattern and classical El Niño pattern, respectively. The spatial pattern of the global SST for each Pacific SST hidden state for April is shown in Figure S2 of Supporting Information S1 as an example.

Our first test is to see whether knowledge of the Pacific SST hidden state in a given calendar month is informative about the probability distribution of precipitation  $m$  months into the future. We take the following approach. For each calendar month  $i(t)$ , we identify the associated hidden state  $k$ , and then compute the mean and standard deviation of the precipitation  $m$  months forward corresponding to that hidden state. Thus, for each calendar month, the predicted teleconnection to the hidden state is defined through the mean and standard deviation of the target variable of interest for the future month. As an example, January 2016 corresponds to hidden state five. In this

case, our “forecasts” for February/July 2016 would be the mean and standard deviation of February/July precipitation for all years in which January was classified as hidden state five. Symbolically, we have the forecast as:

$$pf_{t+m} = \{ \bar{p}_{i(t+m),k(t)}, sp_{i(t+m),k(t)} \} \quad (2)$$

where  $pf_{t+m}$  is the “forecast” for month  $t + m$  given that the current month is  $t$ ,  $\bar{p}_{i(t+m),k(t)}$ , and  $sp_{i(t+m),k(t)}$  are the mean and standard deviation of precipitation for month  $t + m$  given that the current calendar month is  $i(t)$ , and the current SST hidden state is  $k(t)$ . We consider  $m$  varying from 0 (i.e., concurrent) to 12 in the results presented here. This approach is applied to all the grid cells for the re-analysis given the information on the SST hidden state for each month  $t$ . We also performed a similar computation for geopotential height and temperature to check for the physical consistency of the predicted response for precipitation given the SST hidden state.

### 2.3. Significance Testing Using a Bootstrap

Once we have calculated the “forecasts”  $pf_{t+m}$  for each calendar month  $i(t)$ , each hidden state  $k(t)$  and lead time  $m$  (0, 1, 3, 6, 9, 12), we check whether these forecasted means and standard deviations are statistically different from what we expected if we knew nothing about the hidden states. Between 1856 and 2015, we have 160 observations for each calendar month. We bootstrap a sample with 160 observations with replacement for each calendar month and randomly assign bootstrapped observations into five sub-samples belonging to states one to five with sample sizes corresponding to those for the hidden state Viterbi sequence. For each sub-sample, we calculate the mean and standard deviation of the precipitation for the future month in the same way as for the original data. To get the approximate distribution of the mean and standard deviations, we generate 1,000 bootstrap random samples and get 1,000 means and standard deviations for each hidden state and each calendar month. Note that each such sample is of the same size as the number of samples for a hidden state for that calendar month, but it is randomly drawn from any year for that calendar month, without knowledge of the hidden state. This assures that the random, climatological sample from which the mean and standard deviation of precipitation are computed have the same sample size and hence the appropriate uncertainty distribution as the sample for a specific hidden state and calendar month. Next, we compare the forecasted mean and standard deviation with the bootstrapped 1,000 means and standard deviations and calculate the percentile of the forecasted mean and standard deviation relative to the corresponding random distribution of the bootstrapped 1,000 means and standard deviations.

Continuing with the example, consider that the hidden state for January 2016 is 5. We are interested in the precipitation for March 2016. Then the mean and standard deviation of precipitation for all March such that the January hidden state is five are computed. The percentiles associated with these values relative to bootstrapped March mean and standard deviations of precipitation for the sample size for each random draw equal to the frequency of hidden state five occurring in January are assessed. The percentile tells us how unusual the calculated forecasted mean and standard deviation is compared with what one could get by chance. For example, suppose the forecasted statistic has a percentile of 0.01 or 0.99. In that case, one could say that at the 1% significance level, the forecasted statistic is different from what may be expected by chance.

This process is applied to each statistic for each grid cell, and the potential predictability of monthly precipitation can be assessed. Of interest are situations where the mean of the forecasted precipitation for a given hidden state may be unusually high or unusually low for a given calendar month and forecast horizon. If at the same time, the corresponding standard deviation is low, high predictability of the future teleconnection is indicated. A situation where the standard deviation is very high may also be of interest since it indicates a situation when there is high uncertainty/low predictability as to the outcome, even if on average an extreme response is indicated.

### 2.4. Performance of a Continuously Applied $m$ Month Ahead Monthly Precipitation Forecast Using the Pacific SST Hidden State Information

The question explored in this section is whether the conditional “forecast” of global precipitation using the hidden SST states for a given lead time and calendar month, provides more information than using the usual NINO3.4 - precipitation correlations that are commonly used. If the answer is yes, then an additional diagnostic for the teleconnection is provided as a direct comparison to a commonly used current method. To address this question, we consider a continuous application of the precipitation conditional mean as a “forecast” based on the SST hidden states applied sequentially for each month over the historical record for each grid. To ensure that the predictions

are independent of the sample used to define the conditional mean precipitation for a given month, lead time and hidden state, leave-one-out cross-validation procedures were carried out for both forecasts using hidden states and NINO3.4. Once this is done, we can compute the correlation between the “forecasted out of sample” and observed precipitation for that lead time for each grid, and this can then be compared with the correlation of the NINO3.4 for the current month with the  $m$  month ahead precipitation at the same location.

#### 2.4.1. Continuous $m$ Month Ahead Forecast Using Hidden States Information

We applied the cross-validated forecast procedure described in Section 2.2 sequentially for each month and each grid cell. A correlation map of the forecast against observations was then generated for each lead time across all months of record. Note that since no predictive model was calibrated using the precipitation data, this is an out of sample skill assessment.

As an example, consider that the hidden state of April 1950 is state two. To forecast the May 1950 precipitation, the mean values were calculated based on the May precipitation for all years whose April hidden state is state two, excluding data from May 1950. This algorithm is applied sequentially for each May from May 1870 to May 2015 (146 samples), based on the identified state for April. Then we calculate the Pearson correlation between the forecast precipitation anomaly time series and the observed precipitation anomaly series for each location and each calendar month. For example, suppose we would like to assess the 1-month ahead forecast skill for May. We compute the Pearson correlation between forecasted precipitation in May using the hidden state of April and the observed precipitation in May for the whole period.

#### 2.4.2. Comparison With Conditioning on the NINO3.4 Index

As a simple test to see if we learn anything more than would be indicated by using a cross-validated regression with the NINO3.4 index, we compare our results with the correlation of the observed precipitation with the cross-validated regression forecast of precipitation for each grid box from a  $m$  month lagged NINO3.4 index.

### 3. Results

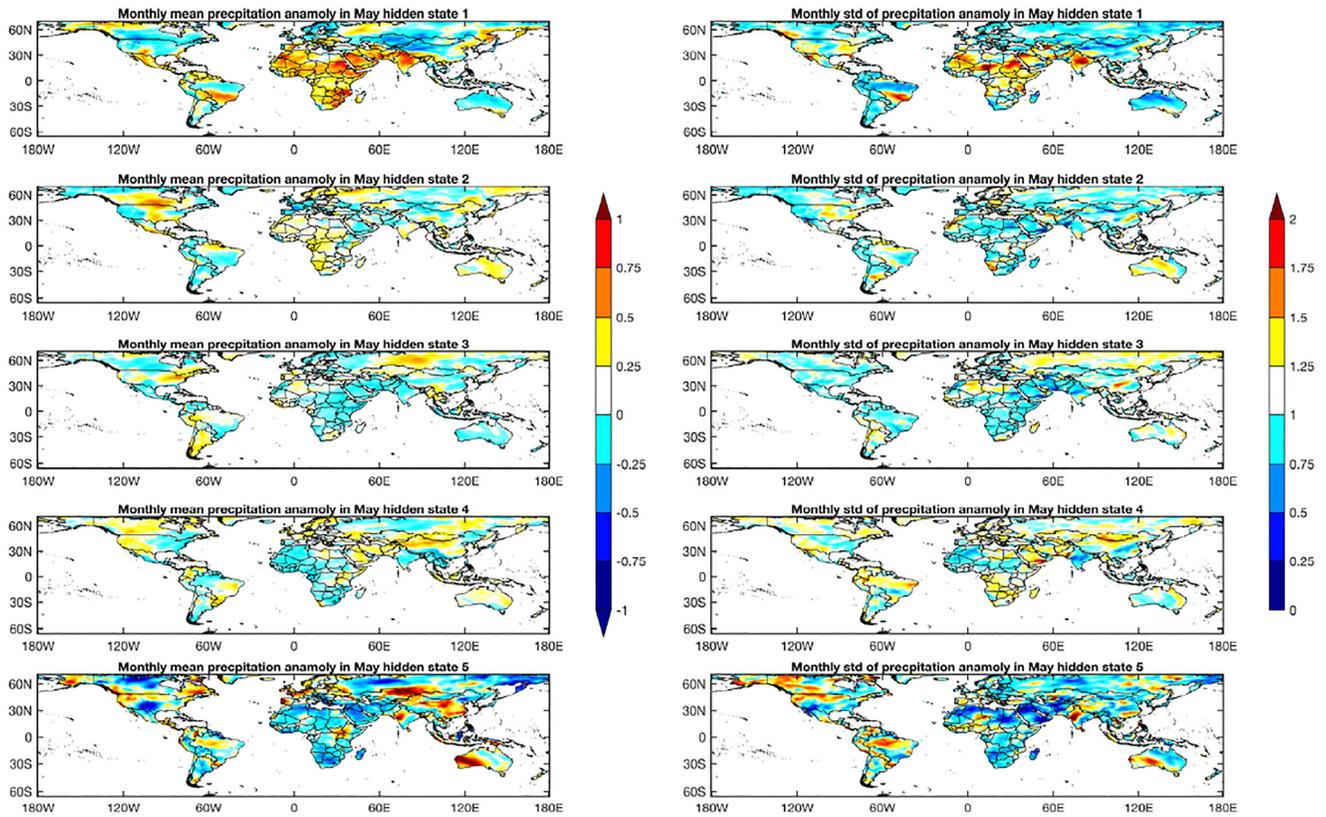
#### 3.1. Monthly Precipitation Predictability Using the Hidden ENSO States

The 9-month ahead “forecast” of precipitation anomaly in May is shown in Figure 1. Conditioning on hidden state five, it is seen that for East Canada, East Europe, Central China, the Korean peninsula etc., the standard deviation (0.5–1) is low, while the mean is high for specific hidden states, which means we are very confident that the May precipitation may be very high in these regions. For the Southwest U.S. and Middle East, both the mean and standard deviation (0–0.75) are low, so we are confident that the precipitation is expected to be low in these regions. On the contrary, both the mean and standard deviation are high for Northwest India (larger than 1.75), Southeast China, Western Europe etc., which means the variation across years in these regions is high, even conditional on the SST hidden state. The results for other lead times can be found in the supplement ([https://github.com/Ivyzhangmj/Mengjie\\_Global-Predictability-for-Monthly-Precipitation-Revealed-by-SST\\_Supporting-Information.git](https://github.com/Ivyzhangmj/Mengjie_Global-Predictability-for-Monthly-Precipitation-Revealed-by-SST_Supporting-Information.git)). As may be expected, the strength of the teleconnection decays by lead time, and varies by calendar month. We use 9-month ahead teleconnections for May for our examples in the main body of the paper to highlight a lead time that is not expected to have skill and also crosses the spring barrier of ENSO predictability.

Given the hidden state and a calendar month, we can look at the “forecast” of any other climate variable (e.g., temperature or atmospheric circulation variables) in the same way. The atmospheric circulation is of particular interest to help establish the physical pathways of causality associated with the precipitation forecast. To demonstrate this, we present the spatial maps of predicted geopotential height of May using April (1 month) Pacific Hidden states in the supplement, and note that the directional changes are consistent with the directional changes one would expect for precipitation in most cases, reinforcing the physical interpretability of the results.

#### 3.2. Significance Test

To see if the anomalies in the conditional means and standard deviations are statistically different from what one would get from a random distribution without the hidden state's information, we constructed a bootstrap test. Figure 2 displays the significance test results of monthly mean and standard deviation of precipitation anomaly in May for 9-month ahead “forecast.” In Figure 2, 0.1/0.9 means the “forecasted” statistic is below/above the



**Figure 1.** Monthly mean (left panel) and standard deviation (std) (right panel) of precipitation anomaly forecast in May using the Pacific hidden states in last August for 9-month ahead teleconnection.

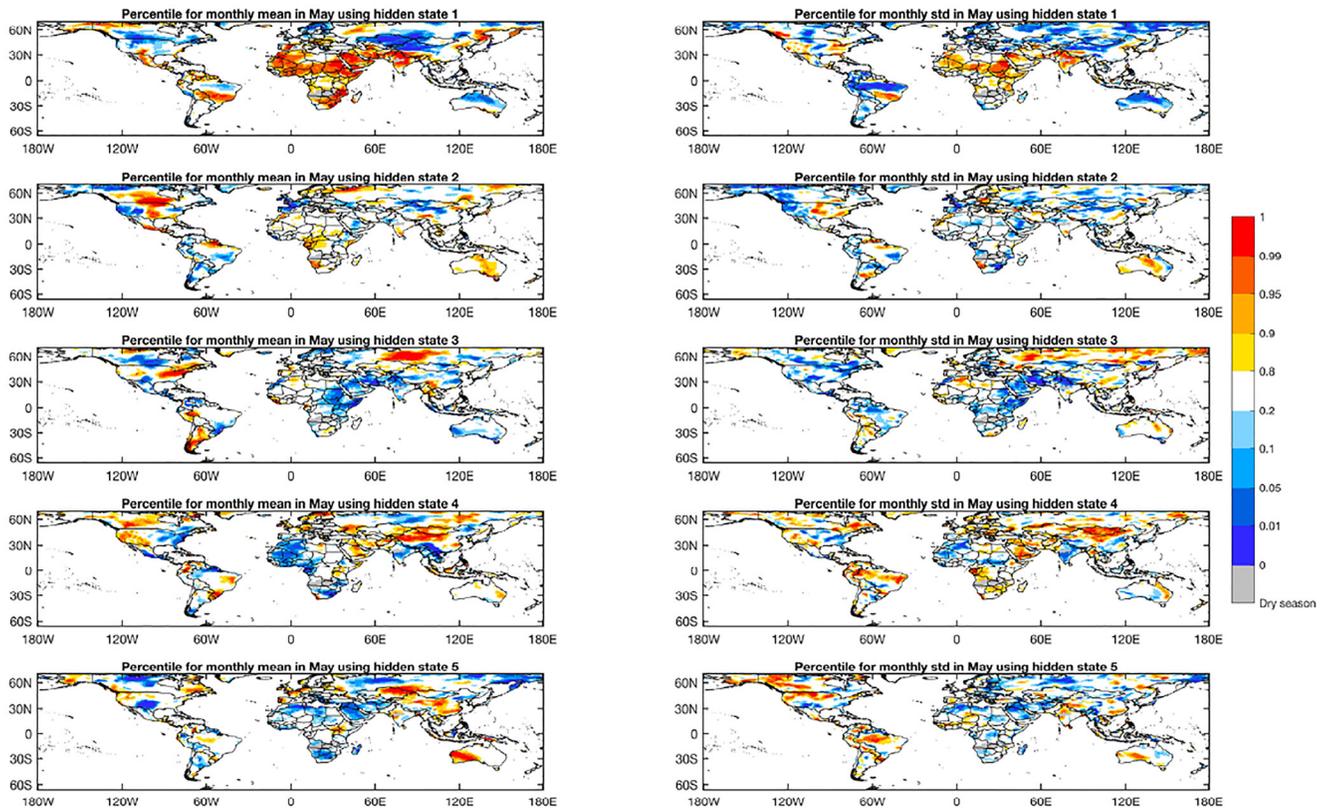
10th/90th percentile of the random distribution of 1,000 bootstrapped means or standard deviations. For the presentation, the percentiles are classified into three categories, which are beyond 90th or 10th percentile (category 1, C1), 20th/80th–10th/90th percentile (category 2, C2), and 20th–80th percentile (category 3, C3).

The forecast results are expected to be helpful based on the following combination of the mean and standard deviation categories:

- Mean > 90th percentile (C1), with standard deviation < 10th percentile (C1).
- Mean > 90th percentile (C1), with standard 10th < standard deviation < 20th percentile (C2).
- 80th < mean < 90th percentile (C2), with 10th < standard deviation < 20th percentile (C2).
- Mean > 90th percentile (C1), with 80th < standard deviation < 90th percentile (C2).
- Mean < 10th percentile (C1), with standard deviation < 10th percentile (C1).

Taking the 9-month ahead forecast using hidden state five as an example (Figure 2), we note that for some areas in East Europe, Central China, and Indonesia, both the mean and standard deviation meet the criteria for C1. This strongly supports that hidden state five is associated with significantly higher monthly precipitation for these regions at 9-months lead time. Similarly, we are more confident that the conditional precipitation is low in the Central U.S. On the contrary, the variation across years is high in North Germany, Spain, Southeast China etc., where the mean meets C1 criteria but the standard deviation falls into the C2 category. It should also be noted that the results for East Australia and the Korean peninsula are not as confident either since the percentile of mean is in C1 while the percentile of the standard deviation is in C3.

Figure 3 displays the regions with high confidence high or low precipitation forecasts using the five hidden states for 9-month lead time. Note that “high confidence” is defined as forecasts with a low standard deviation (C1 or C2) and a forecast mean that is greater than the 80th/90th percentile or less than the 10th/20th percentile, and we are confident that the precipitation is high or low in that area. The results for other lead times can be found in the supplement. We noticed that the strong ENSO teleconnections indicated in our monthly analysis often vary



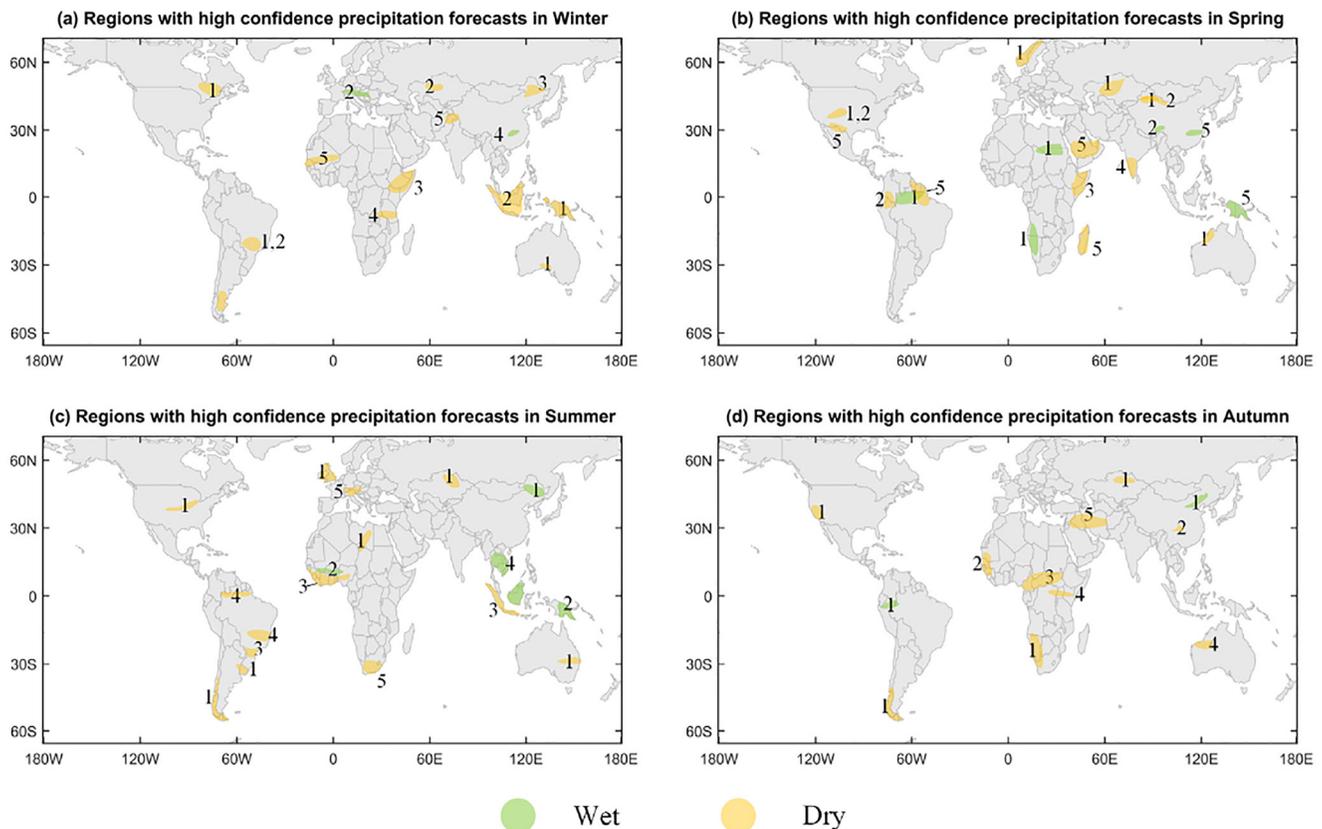
**Figure 2.** Percentiles relative to bootstrapped distributions for monthly mean (left panel) and standard deviation (right panel) of precipitation anomaly forecast in May using the Pacific hidden states in the previous August for 9-month ahead teleconnection.

spatially by calendar month. This is to be expected given the seasonal variations in the Intertropical Convergence Zone and the jet stream dynamics that influence monsoonal dynamics as well as tropical moisture exports to the mid-latitudes and the position and strength of frontal and cyclonic systems. Thus, contrary to the typical practice of seasonal aggregation with moving season boundaries, we believe that the monthly analysis is more informative of the underlying dynamics of the system, and therefore better informs the sub-season to season trajectory that is of interest.

### 3.3. Continuous Monthly Resolution Precipitation Forecast for the Full Record

There is a vast literature on the correlation between seasonal precipitation and climate indices. However, very few focus directly on monthly precipitation. In this section, we report the correlation of 1-, 3-, 6-, 9-, and 12-month ahead continuous forecasts of monthly mean precipitation. To remind the reader, a 9-month ahead continuous “forecast” considers the hidden SST state assigned to month  $t$ , and assigns the conditional mean of precipitation for month  $t + 9$  given that hidden state in month  $t$ . The same process is then repeated for each month ( $t + 9$ ) of 1870–2015, and the correlation between the cross-validated predicted mean and the observed value is then computed using all forecasts of that lead time for each grid cell.

Figure 4 presents the 9-month ahead continuous forecast correlations for May over the period of 1870–2015. The correlations between observed and cross-validated forecasted precipitation in May for some places in Northwest China, West, and North Europe, Southern Canada, Southern Africa, Southern USA, West Australia etc., are above 0.3 (Figure 4a). To account for spatial autocorrelation among grids, Walker's test and false discovery rate (Wilks, 2006, 2016) were used to examine the field significance of correlations, and our results are field significant using both methods at the 0.05 global significance level (Figures 4a–4c). Figure 4c presents the difference between the correlation of observed and cross-validated forecast precipitation using hidden states and the correlation of observed and cross-validated forecast precipitation using NINO3.4. It is found that the forecast correlation using hidden states is significantly better than that using NINO3.4 for about 70% of land area. The



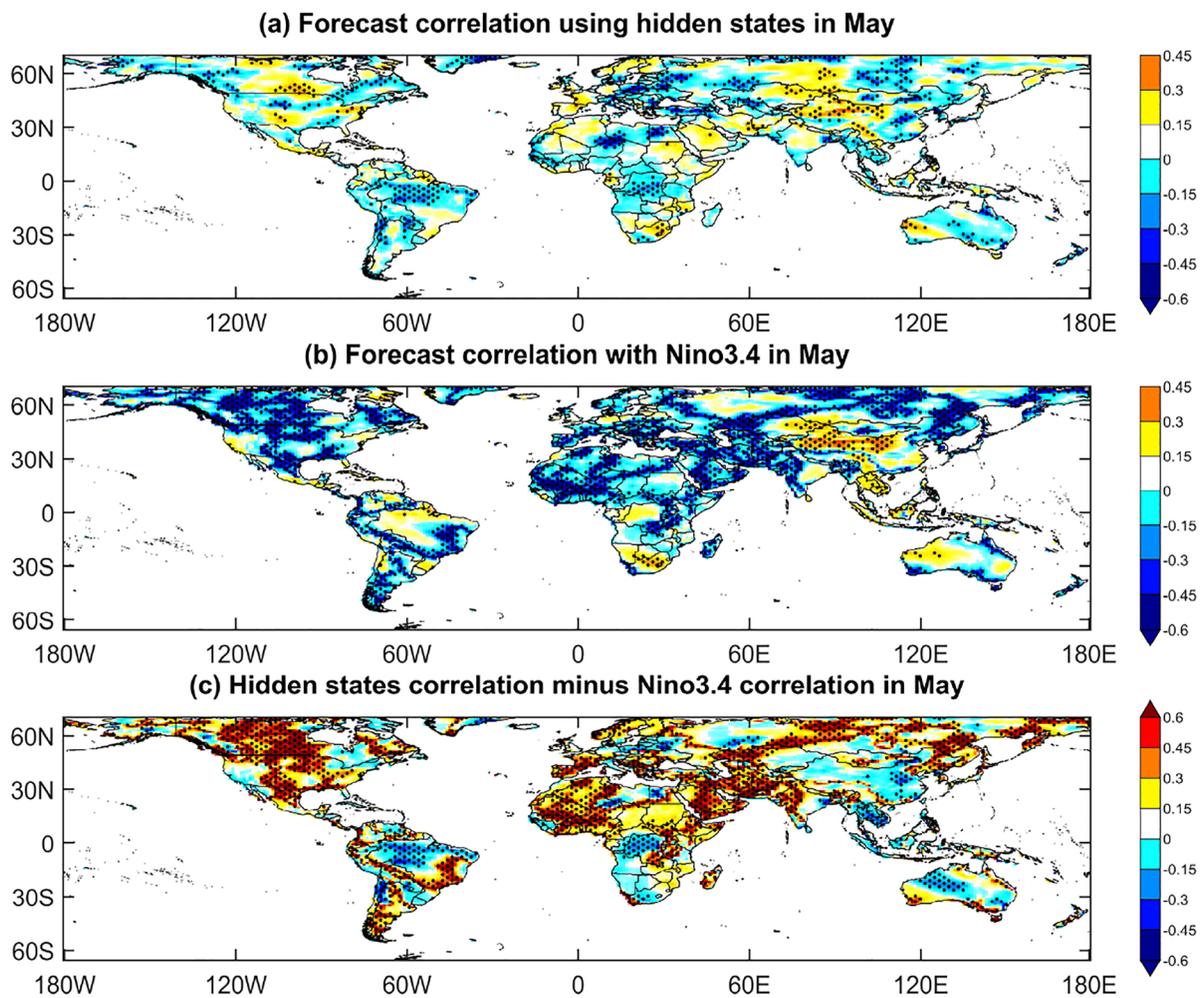
**Figure 3.** Regions with high confidence high or low precipitation forecasts for 9-month lead time. The numbers 1–5 indicate which hidden states (1–5) contributes to the strength of the information.

forecast correlations for other months and lead times are provided in supplement, and we summarize the key features for each lead time and each season in Table S1 of Supporting Information S1. It should be mentioned that our monthly forecast correlations are higher compared with many studies typically analyzing the teleconnection between ENSO and regional seasonal aggregated precipitation (Lee, 2015; Pandey et al., 2020; Yoon & Leung, 2015), even for a 12-month lead time.

#### 4. Summary and Discussion

Since we used precipitation re-analysis and retrospective SST data, what we presented is really an exercise in examining the potential predictability of the re-analysis data given the SST characterization and not a forward forecasting exercise. In this context, we find that potentially there is substantial predictability of monthly precipitation at lead times up to 1 year in many global regions conditional on a classification of the tropical Pacific SST. The predictability is highly state dependent, varies by the calendar month used to choose the conditioning state, and as expected, it decreases with increasing lead time. The predictability is asymmetric between the La Niña and the El Niño states, suggesting that linear statistical models typically used with ENSO indices are not likely to be as effective as conditioning on the states identified by the NHMM. The approach taken here was to compute the first two moments of the conditional distribution of precipitation at each global location, given each hidden state and calendar month. Regions with promising predictability were identified based on the relative anomalies of the mean and standard deviation of precipitation given a hidden state. A more detailed conditional model for precipitation that also considers spatial correlation could be developed targeting these regions individually or collectively.

The findings suggest that an approach similar to the one in Rojo-Hernández, (2018) where the Pacific SST is modeled by a NHMM and the resulting monthly SST hidden states are used as predictors of the hidden states of a NHMM for daily precipitation in Colombia may be a promising building block for global precipitation



**Figure 4.** Correlation of cross-validated continuous 9-month ahead forecast for May (from May 1870 to May 2015). (a) Correlation between observed and “forecast” monthly mean precipitation for May using previous August hidden states; (b) correlation between observed and “forecast” precipitation for May using previous August NINO3.4; (c) difference in correlations between using hidden state and NINO3.4. The significant correlations at the 5% level are highlighted using the stippling.

from monthly to annual time scales. We have considered a coupled, multilevel NHMM where regional or global daily precipitation could be modeled conditionally on global SST, with a simultaneous, Bayesian selection of model parameters and uncertainty characterization. However, the associated computational requirements for model fitting and testing at the global scale proved to be formidable. Consequently, we restricted ourselves to the experiment reported here which is computationally inexpensive, as a proof of concept of the idea that the SST hidden states from a NHMM or similar model that classifies SST dynamics, could be a useful building block for a predictive model.

Our future work will focus on addressing the computational issues with the multi-level global machine learning model for SST and precipitation, as well as explore some of the direct modeling of the conditional precipitation and temperature fields given the SST hidden states. We intend to explore emerging deep learning methods in addition to the NHMM for the purpose.

### Data Availability Statement

NOAA-CIRES-DOE Twentieth Century Reanalysis (V3) data set used in this study can be accessed from [https://psl.noaa.gov/data/gridded/data.20thC\\_ReanV3.html](https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html); The SST anomaly data extracted from the monthly Kaplan Extended SST V2 data set is from [https://www.psl.noaa.gov/data/gridded/data.kaplan\\_sst.html](https://www.psl.noaa.gov/data/gridded/data.kaplan_sst.html); The NINO3.4 index data is available at [https://psl.noaa.gov/gcos\\_wgsp/Timeseries/Nino34/](https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino34/).

**Acknowledgments**

This paper is financially supported by National Science Foundation Convergence Accelerator Track D America's Water Risk: Water System Data Pooling for Climate Vulnerability Assessment and Warning System Project Number 2040613.

**References**

Arcodia, M. C., Kirtman, B. P., & Siqueira, L. S. (2020). How MJO teleconnections and ENSO interference impacts US precipitation. *Journal of Climate*, 33(11), 4621–4640. <https://doi.org/10.1175/JCLI-D-19-0448.1>

Ashok, K., Behera, S. K., Rao, S. A., Weng, H., & Yamagata, T. (2007). El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research*, 112, C11007. <https://doi.org/10.1029/2006jc003798>

Chapman, W. E., Subramanian, A. C., Xie, S. P., Sierks, M. D., Ralph, F. M., & Kamae, Y. (2021). Monthly modulations of ENSO teleconnections: Implications for potential predictability in North America. *Journal of Climate*, 34(14), 5899–5921. <https://doi.org/10.1175/JCLI-D-20-0391.1>

Dai, A., & Wigley, T. M. L. (2000). Global patterns of ENSO-induced precipitation. *Geophysical Research Letters*, 27(9), 1283–1286. <https://doi.org/10.1029/1999GL011140>

Emerton, R., Cloke, H. L., Stephens, E. M., Zsoter, E., Woolnough, S. J., & Pappenberger, F. (2017). Complex picture for likelihood of ENSO-driven flood hazard. *Nature Communications*, 8(1), 1–9. <https://doi.org/10.1038/ncomms14796>

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>

Henderson, S. A., Vimont, D. J., & Newman, M. (2020). The critical role of non-normality in partitioning tropical and extratropical contributions to PNA growth. *Journal of Climate*, 33(14), 6273–6295. <https://doi.org/10.1175/JCLI-D-19-0555.1>

Holsclaw, T., Greene, A. M., Robertson, A. W., & Smyth, P. (2017). Bayesian nonhomogeneous Markov models via Pólya-Gamma data augmentation with applications to rainfall modeling. *Annals of Applied Statistics*, 11(1), 393–426. <https://doi.org/10.1214/16-AOAS1009>

Kaplan, A., Kushnir, Y., Cane, M. A., & Blumenthal, M. B. (1997). Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *Journal of Geophysical Research*, 102(C13), 27835–27860. <https://doi.org/10.1029/97JC01734>

Lee, H. S. (2015). General rainfall patterns in Indonesia and the potential impacts of local seas on rainfall intensity. *Water*, 7(4), 1751–1768. <https://doi.org/10.3390/w7041751>

Lenssen, N. J., Goddard, L., & Mason, S. (2020). Seasonal forecast skill of ENSO teleconnection maps. *Weather and Forecasting*, 35(6), 2387–2406. <https://doi.org/10.1175/WAF-D-19-0235.1>

Lima, C. H., Lall, U., Jebara, T., & Barnston, A. G. (2009). Statistical prediction of ENSO from subsurface sea temperature using a nonlinear dimensionality reduction. *Journal of Climate*, 22(17), 4501–4519. <https://doi.org/10.1175/2009JCLI2524.1>

McPhaden, M. J., Zebiak, S. E., & Glantz, M. H. (2006). ENSO as an integrating concept in Earth science. *Science*, 314(5806), 1740–1745. <https://doi.org/10.1126/science.1132588>

Nigam, S., & Sengupta, A. (2021). The full extent of El Niño's precipitation influence on the United States and the Americas: The sub-optimality of the Niño 3.4 SST index. *Geophysical Research Letters*, 48(3), e2020GL091447. <https://doi.org/10.1029/2020gl091447>

Pandey, P., Dwivedi, S., Goswami, B. N., & Kucharski, F. (2020). A new perspective on ENSO-Indian summer monsoon rainfall relationship in a warming environment. *Climate Dynamics*, 55(11), 3307–3326. <https://doi.org/10.1007/s00382-020-05452-7>

Rojo-Hernández, J. D. (2018). *Spatial and temporal characterization of Colombia's hydroclimatology*. Escuela de Geociencias y Medio Ambiente.

Rojo-Hernández, J. D., Mesa, Ó. J., & Lall, U. (2020). ENSO dynamics, trends, and prediction using machine learning. *Weather and Forecasting*, 35(5), 2061–2081. <https://doi.org/10.1175/WAF-D-20-0031.1>

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. [https://doi.org/10.1175/1520-0493\(1987\)115<0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<0.CO;2)

Sun, X., Renard, B., Thyer, M., Westra, S., & Lang, M. (2015). A global analysis of the asymmetric effect of ENSO on extreme precipitation. *Journal of Hydrology*, 530, 51–65. <https://doi.org/10.1016/j.jhydrol.2015.09.016>

Tang, Y., Zhang, R. H., Liu, T., Duan, W., Yang, D., Zheng, F., et al. (2018). Progress in ENSO prediction and predictability study. *National Science Review*, 5(6), 826–839. <https://doi.org/10.1093/nsr/nwy105>

Tseng, K. C., Johnson, N. C., Maloney, E. D., Barnes, E. A., & Kapnick, S. B. (2021). Mapping large-scale climate variability to hydrological extremes: An application of the linear inverse model to subseasonal prediction. *Journal of Climate*, 34(11), 4207–4225. <https://doi.org/10.1175/JCLI-D-20-0502.1>

Van Oldenborgh, G. J., & Burgers, G. (2005). Searching for decadal variations in ENSO precipitation teleconnections. *Geophysical Research Letters*, 32(15), L15701. <https://doi.org/10.1029/2005GL023110>

Vicente-Serrano, S. M., López-Moreno, J. I., Gimeno, L., Nieto, R., Morán-Tejeda, E., Lorenzo-Lacruz, J., et al. (2011). A multiscale global evaluation of the impact of ENSO on droughts. *Journal of Geophysical Research*, 116, D20109. <https://doi.org/10.1029/2011JD016039>

Wilks, D. (2016). “The stippling shows statistically significant grid points”: How research results are routinely overstated and overinterpreted, and what to do about it. *Bulletin of the American Meteorological Society*, 97(12), 2263–2273. <https://doi.org/10.1175/BAMS-D-15-00267.1>

Wilks, D. S. (2006). On “field significance” and the false discovery rate. *Journal of Applied Meteorology and Climatology*, 45(9), 1181–1189. <https://doi.org/10.1175/JAM2404.1>

Yan, L., Xiong, L., Jiang, C., Zhang, M., Wang, D., & Xu, C.-Y. (2021). Updating intensity–duration–frequency curves for urban infrastructure design under a changing environment. *Wiley Interdisciplinary Reviews: Water*, 8(3), e1519. <https://doi.org/10.1002/wat2.1519>

Yang, Y. M., Park, J. H., An, S. I., Wang, B., & Luo, X. (2021). Mean sea surface temperature changes influence ENSO-related precipitation changes in the mid-latitudes. *Nature Communications*, 12(1), 1–9. <https://doi.org/10.1038/s41467-021-21787-z>

Yoon, J. H., & Leung, L. R. (2015). Assessing the relative influence of surface soil moisture and ENSO SST on precipitation predictability over the contiguous United States. *Geophysical Research Letters*, 42(12), 5005–5013. <https://doi.org/10.1002/2015GL064139>

Zhang, Q., Wang, Y., Singh, V. P., Gu, X., Kong, D., & Xiao, M. (2016). Impacts of ENSO and ENSO Modoki+ A regimes on seasonal precipitation variations and possible underlying causes in the Huai River basin, China. *Journal of Hydrology*, 533, 308–319. <https://doi.org/10.1016/j.jhydrol.2015.12.003>