AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2015WR017644

Key Points:
- Precipitation predictor variables are evaluated conditioned on the "state" of ENSO
- Marked difference in SST fields emerges within these states
- Forecasting results improve significantly over evaluating SST fields in only one state

Correspondence to:
B. G. Zimmerman,
bgzimmerman@wisc.edu

Citation:
Zimmerman, B. G., D. J. Vimont, and P. J. Block (2016), Utilizing the state of ENSO as a means for season-ahead predictor selection, *Water Resour. Res.*, *52*, 3761–3774, doi:10.1002/2015WR017644.

Received 1 JUN 2015
Accepted 21 APR 2016
Accepted article online 25 APR 2016
Published online 18 MAY 2016

Utilizing the state of ENSO as a means for season-ahead predictor selection

Brian G. Zimmerman[1,2,3], Daniel J. Vimont[2,3], and Paul J. Block[1]

[1]Department of Civil and Environmental Engineering, University of Wisconsin - Madison, Madison, Wisconsin, USA, [2]Department of Atmospheric and Oceanic Science, University of Wisconsin - Madison, Madison, Wisconsin, USA, [3]Now at Center for Climatic Research, University of Wisconsin - Madison, Madison, Wisconsin, USA

Abstract This paper introduces the Nino Index Phase Analysis (NIPA) framework for forecasting hydroclimatic variables on a seasonal time scale. Antecedent Sea Surface Temperatures (SSTs) are commonly used in statistical predictive frameworks for seasonal forecasting, however, the typical approach of evaluating all the years on record in one bin ("phase") does not often provide the level of skill required by decision makers. For many locations around the world, the most influential climate signal on the seasonal time scale is the El Nino Southern Oscillation (ENSO), and there are various indices used to capture the state of ENSO and provide this information. NIPA utilizes the state of ENSO to classify the years of record into four phases, operating under the hypothesis that ENSO itself is affecting the "mean state" of the atmospheric-oceanic system, and relevant teleconnections depend on and must be selected within these mean states. A case study focused on spring precipitation over the Lower Colorado River Basin (LCRB) in Texas is chosen to illustrate NIPA's potential. Results show that correlations between wintertime SST fields and spring precipitation in the LCRB improve from 0.21 to 0.47 for the typical "one phase" and the NIPA "four-phase" approach, respectively. Even greater improvements are seen across tercile-based skill scores such as the Heidke Hit Skill Score and Ranked Probability Skill Score; skill is particularly strong for years exhibiting extreme wet or dry conditions. It also outperforms the North American Multi-Model Ensemble predictions across the LCRB for the selected seasons. This is encouraging as improved predictability through NIPA may translate to better decision-making for water managers.

1. Introduction to Season-Ahead Precipitation Prediction

Water is the primary medium through which climate has an impact on people, ecosystems, and economies [*Sadoff and Muller*, 2009]. Recently, persistent droughts across parts of the west and southwest U.S. have presented considerable challenges to decision makers involved in regional water planning and management. Such challenges include implementing reductions to interruptible contract holders (agriculture and industry) and imposing reduction measures in municipalities while simultaneously maintaining fiscal solvency and managing public relations. Reservoir management is rarely simple—many reservoirs in the western U.S. serve multiple, often competing, purposes, such as drought mitigation and flood prevention, which, respectively, prioritize high and low reservoir levels.

According to projections by state-of-the-art climate models, it is likely that anthropogenic climate change will spur an increase in extreme climate events around the world [*Allan and Soden*, 2008], changing the landscape of both water supply and demand [*Jiménez Cisneros et al.*, 2014]. For water managers to properly oversee an increasingly scarce resource with large economic, environmental, and livelihood implications, it is imperative that they have sufficient flexibility to adapt, requiring access to the advanced resources and tools necessary to effectively react to current conditions and simultaneously plan for both near and long-term predicted conditions.

Thus, prospects for predicting precipitation and other hydroclimatic variables on time scales of months to decades is of interest to explore means of buffering impacts induced by climate variability. Both dynamical and statistical predictive frameworks are commonly applied, each having their own distinctive advantages and disadvantages with neither proving superior for all seasons or locations [*Block et al.*, 2009].

Regional to local-scale precipitation predictions by atmospheric-oceanic general circulation models AOGCMs are typically highly parameterized [*Neelin et al.*, 2010]. Even when an optimal combination of

multiple models, such as the North American Multi-Model Ensemble (NMME) [*Kirtman et al*., 2013] is utilized, skill for precipitation over the U.S. is generally modest [*Infanti and Kirtman*, 2014; *Mo and Lettenmaier*, 2014; *Becker et al*., 2014].

Thus, statistical prediction models (SPMs) that forgo parameterization of complex physical processes in favor of investigating lagged relationships between regional seasonal precipitation amounts and antecedent atmospheric/oceanic (A/O) conditions warrant attention as a complementary approach [*Abbot and Marohasy*, 2012]. SPMs capitalize on teleconnections (climate anomalies related to each other at large spatial and temporal scales) present in the A/O system. They typically utilize anomalies of Sea Surface Temperatures (SSTs), though they can also incorporate a wide variety of other A/O variables (or indices that represent the time evolution of a particular A/O variable, or combination of variables) as predictors to directly estimate hydroclimatic variables of interest. As such, SPMs offer an appealing complement to dynamical models, particularly for locations where dynamical models are currently underperforming. While a complete physical understanding of the climate system through statistical models may not be possible or intended [*Anderson et al*., 1999], investigating hydroclimatic relationships through SPMs can assist in developing a better mechanistic understanding of the atmosphere by illuminating regional hydroclimatic teleconnections to global A/O variables, as will be demonstrated here.

For season-ahead SPMs, SST anomalies are well documented as the best performing large-scale precipitation predictor field [*Barnston*, 1994; *Markowski and North*, 2003; *Quan et al*., 2006], based on the mechanistic premise that the lower atmosphere is forced by large-scale anomalous surface processes [*Frankignoul*, 1985; *Branković et al*., 1994; *Lloyd-Hughes and Saunders*, 2002; *Mo and Lettenmaier*, 2014]. Indeed, the most prominent teleconnection pattern shown to affect precipitation patterns in both North America and around the world is the El Nino Southern Oscillation (ENSO) [*Ropelewski and Halpert*, 1986, 1987; *Quan et al*., 2006], with the Nino 3.4 index commonly used in seasonal prediction [*Abbot and Marohasy*, 2014; *Block and Rajagopalan*, 2007; *Hartmann et al*., 2008].

The state of the large-scale climate system is also influenced by ENSO interactions. *Enfield et al*. [2001] illustrates how interactions between ENSO and the Atlantic Multidecadal Oscillation (AMO) can affect river flows in the U.S., while *Shabbar and Skinner* [2004] demonstrate how the interaction of ENSO and the Pacific Decadal Oscillation (PDO) plays a role in North American drought. *Gershunov* [1998] shows that year-to-year ENSO effects can be modulated by the low-frequency PDO signal, likely due to the similar spatial patterns [*Wang et al*., 2008]. *Wang et al*. [2010] confirms a cold Pacific-warm Atlantic pattern that can produce significant drought conditions across the U.S. using AOGCM simulations, supporting the data-driven findings of *McCabe et al*. [2004] that align increased drought frequency in the southwest U.S. with PDO negative phase and AMO positive phase conditions. Given the nonindependence of these large-scale phenomena, there are certainly idiosyncrasies in the manner in which they interact—perhaps not surprisingly, ENSO itself exhibits marked variations in its amplitude, temporal evolution, and spatial expression from event to event.

Interest in the diversity of ENSO events began in earnest with *Larkin and Harrison* [2005] and *Ashok et al*. [2007], into, respectively, the effect of different types of ENSO events on seasonal weather anomalies in the U.S. and on the 2004 "El Nino Modoki" and its associated hydroclimatic anomalies. Since then, according to *Capotondi et al*. [2014], "Significant research has been conducted to identify, describe, and understand these El Nino types, spurring debates on whether there are indeed two distinct modes of variability [as suggested by *Ashok et al*. [2007]], or whether ENSO can be more aptly described as a diverse continuum [as suggested by *Takahashi et al*. [2011]]." The current body of knowledge concerning these various "flavors" of ENSO is thoroughly reviewed by *Capotondi et al*. [2014], who, in conclusion, pose a set of four questions to guide continued research into the understanding of ENSO diversity and the resulting tropical and extratropical teleconnections. This paper provides insight into the second of these questions—What are the sources and limits of predictability both tropical and extratropical, associated with the differences in ENSO events?

The skill of SPMs varies from season to season and location to location. They have been found to underperform compared to multimodel ensembles [*Block and Goddard*, 2012], but have also shown success when tailored to specific regions [*Grantz et al*., 2005; *Regonda et al*., 2006], and, like dynamical models, generally perform better for temperature than precipitation [*Barnston*, 1994]. In this study, a novel technique not yet applied in a predictive manner is developed. The methodology is flexible enough to potentially provide seasonal predictions for any hydroclimatic variable for a multitude of locations; however, it is demonstrated

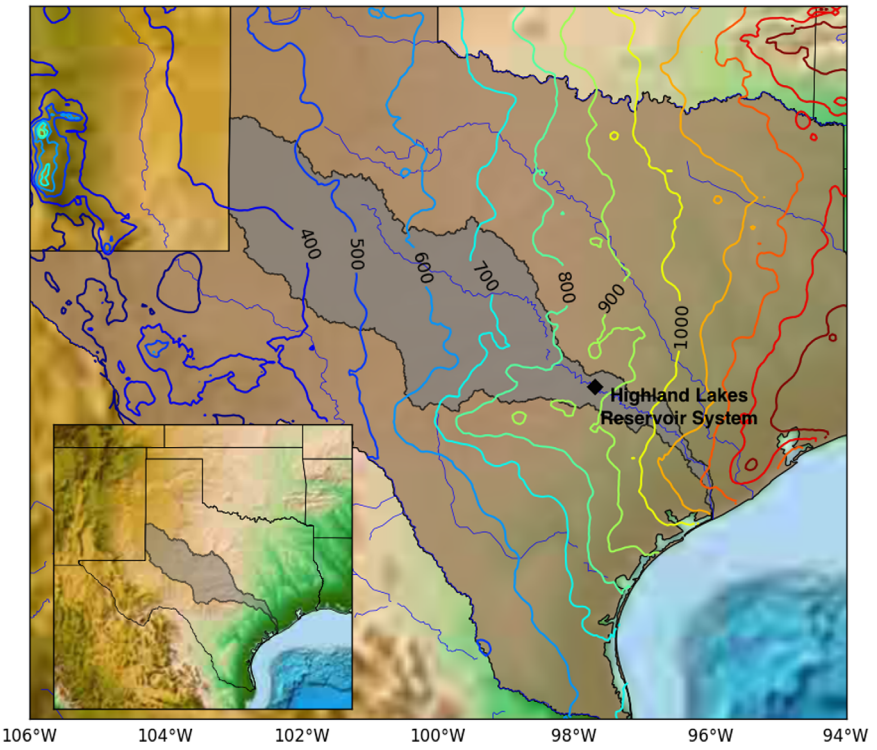

Figure 1. Texas and the LCRB, annual total precipitation (mm) contours.

here solely for precipitation in one Texas river basin, and any statements regarding potential applicability of the framework are made with this caveat in mind. In this methodology, the state of ENSO is utilized not as a predictor, but as a physical influence on the "mean state" of the A/O system to aid in uncovering otherwise overlooked hydroclimatic signals that may be informative in a statistical prediction framework. Historical years are binned into four phases based on the state (or strength) of ENSO as measured by the Multivariate ENSO Index (MEI) [*Wolter and Timlin*, 2011], and phase-specific antecedent SST fields are identified and employed as predictors in a principal component regression model. Geographical regions of winter SST anomalies that contribute predictive information toward spring precipitation will be shown to depend greatly on the state of ENSO, warranting a modeling framework that captures such intraphase variability.

2. Lower Colorado River Basin Case Study

The proposed methods utilize the state of ENSO to aid in precipitation predictor selection, and were initially developed to assist spring streamflow forecasts in the Lower Colorado River Basin (LCRB) under a NOAA grant and partnership with the Lower Colorado River Authority (LCRA). As such, motivations and the ensuing discussion will focus on the LCRB as a case study. Future work will focus on application of the methodology to a variety of basins and seasons, however, the main focus for this paper will be on introducing and validating the methodology.

The LCRB (Figure 1) supplies the Highland Lakes Reservoir system located just north of Austin, Texas, and is managed by the LCRA. Their mandate is to provide flood control, manage hydropower generation, maintain public recreational areas, and environmental flows, all while serving a wide range of customers including municipalities, power plants, and farmers. Two of the six lakes serve primarily as water supply reservoirs, with a combined storage capacity in excess of however, until recently, storage levels have hovered closer to one third of capacity given the persistent drought in Texas since 2008.

The LCRA's customers hold either firm or interruptible water contracts. On 1 March of each year, prior to the first wet season (March–July, MAMJ) in the basin, the LCRA issues commitments to their interruptible

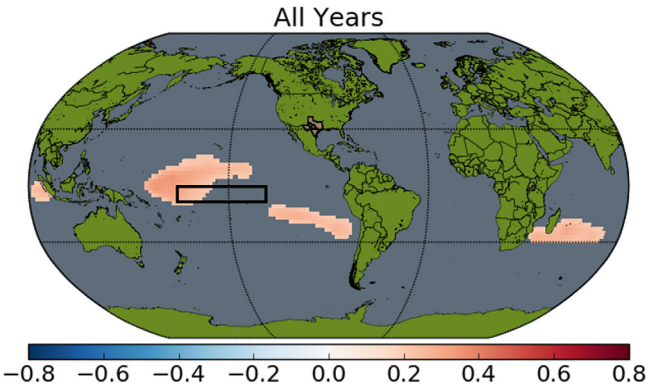

Figure 2. Correlation map of MAMJ precipitation with NDJF SSTs for 1921–2010, Nino 3.4 region bounded in black.

contract holders. The current decision support system is rigid; if the reservoir is below a certain level on the decision date, specific curtailments will occur. In three recent consecutive years (2012–2014), downstream farmers holding interruptible contracts were cut off, causing significant financial strains and distress among stakeholders. The LCRA has shown interest in utilizing seasonal forecasts to enhance their decision support system, and are considering formally including the use of seasonal forecasts in their next Water Management Plan. Skillful forecasts could enable the LCRA to optimize their management decisions in both the short and long term. This could be especially useful as water demand increases and supply in the LCRB, known to be under the influence of ENSO [*Gershunov*, 1998; *Rajagopalan et al*., 2000], varies.

SSTs within the Nino 3.4 region (120°W–165°W, 5°S–5°N) are among the statistically significant regions of wintertime (NDFJ) SST anomalies to correlate with springtime precipitation in the LCRB (Figure 2), however, the correlation is modest (0.28–0.32). Regions of statistically significant correlated SSTs change dramatically when ENSO phase information is taken into account; the data and methodologies used to incorporate this type of information into a novel forecasting framework and create skillful predictions are presented in the following sections. Figure 3 provides further justification that the mean precipitation (through composite analysis) during El Nino years is statistically significantly higher (by the student's t test, $p = 0.05$) than the mean precipitation during La Nina years. However, the within-phase variance is large for both phases; 7 of 14 El Nino years fall below the all-year mean, while 5 of 17 La Nina years fall above the all-year mean. It is these types of within-phase variations that the proposed methodology aims to predict.

3. Hydroclimatic Variables and Sources

Hydroclimate observations and model inputs utilized in this study span 1921–2010. Each year is split into two seasons to match LCRB climatology—winter (NDJF, November and December start in 1920) and spring (MAMJ).

The PRISM precipitation data set [*Daly et al*., 2008], gridded at 2.5 arc min ($\frac{1}{8}$ km) with monthly resolution, is adopted here to represent observations. The precipitation data are averaged across all grids within the basin and aggregated to seasonal totals. A gridded data set was selected in lieu of station data for two reasons: first, seasonal forecasts will eventually be used to force a distributed (gridded) hydrologic model to produce inflows into the Highland Lakes reservoirs, thus aligning well spatially. Second, a large east-west gradient in total annual precipitation exists across the basin (Figure 1) ranging from 850 mm/yr near the Highland Lakes Reservoir system to 350 mm/yr on the western edge of the watershed. The gridded data, at a higher spatial resolution, better captures this pattern.

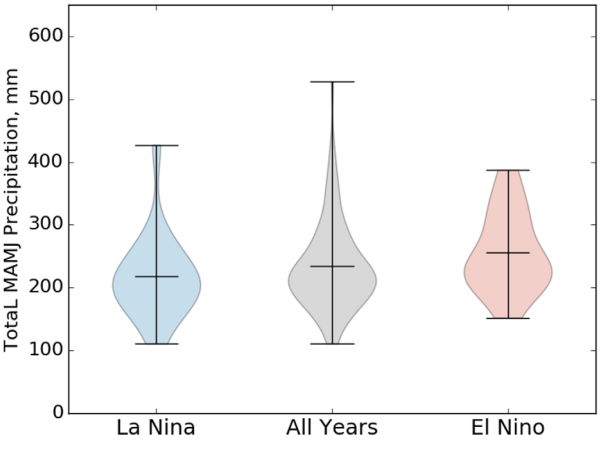

Figure 3. Violin plots of MAMJ precipitation by phase.

To ensure that the seasonal precipitation signal is spatially homogeneous across the

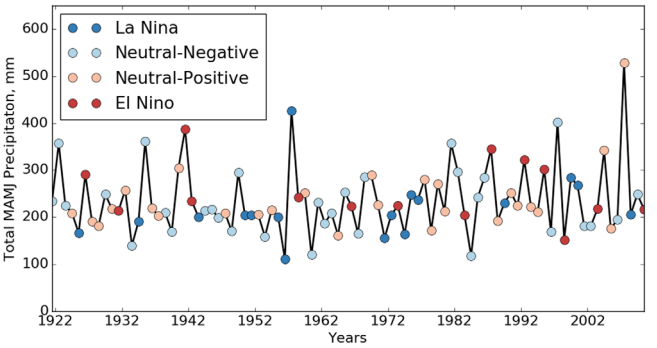

Figure 4. Historical precipitation binned by phase.

basin on the seasonal scale, a principal component analysis (PCA) was performed on the gridded data. PCA is a statistical procedure commonly used to extract dominant modes of variability from large, multidimensional data sets [*Wilks*, 1995]. The leading mode explains 69% of the variance in the data, and the leading principal component correlates almost perfectly ($r = 0.99$) with the areal average of precipitation, justifying the use of spatially averaged basin data in building a forecast model.

NOAA's Extended Reconstructed SST Version 3b [*Smith et al.*, 2008] is used for global monthly SST anomalies, gridded at 2.5°.

4. Nino Index Phase Analysis

The Nino Index Phase Analysis (NIPA) statistical framework is proposed to predict total LCRB MAMJ precipitation conditioned on prior season A/O variables. This novel framework utilizes information from the mean state of the A/O system in the tropical Pacific in the months prior to the season of interest to divide the historical record (90 years for LCRB) into four distinct "mean states," or phases. Though ENSO is the dominant signal used in precipitation prediction for many global locations, North American seasonal predictability has been previously associated with different phases of ENSO [*Kumar and Hoerling*, 2011], and there is much evidence that localized (in close proximity to the basin of interest) SST anomalies, outside of the tropical Pacific, can provide skill in seasonal forecasting [*Rajagopalan et al.*, 2000; *Camberlin et al.*, 2001; *Markowski and North*, 2003; *Phillips and Thorpe*, 2006; *Block and Rajagopalan*, 2007; *White et al.*, 2008; *Bazo et al.*, 2013; *Chen and Georgakakos*, 2014]. It will be shown for the LCRB that, conditioned on the phase of ENSO, NIPA excels at illuminating these types of teleconnections.

4.1. Multivariate ENSO Index

Changes in the "mean state" of the A/O system caused by tropical hydroclimatic anomalies associated with ENSO are of interest, and so the Multivariate ENSO Index (MEI) is selected to bin the years [*Wolter and Timlin*, 1993]. The MEI is constructed as the first principle component of six variables associated with ENSO-SST, SLP, surface air temperature, zonal wind, meridional wind, and cloudiness fraction. A spectra of the MEI illustrate power in the frequency expected for ENSO, 4–7 years, indicating that the typical ENSO signal is captured. Figure 4 shows the annual precipitation record binned into the four phases utilized for this study: La Nina, Neutral-Negative, Neutral-Positive, and El Nino. Thresholds between bins are selected based on *Wolter and Timlin* [2011], using percentile ranks to categorize El Nino or La Nina events into strong (tenth percentile), moderate (twentieth percentile), and weak (thirtieth percentile), with the middle 40% considered neutral. This study utilizes four phases; as such MEI = 0 divides the positive from negative phases, and the thirtieth percentile within positive/negative phases is used to delineate El Nino (La Nina) years from neutral-positive (negative) years. A summary can be seen in Table 1. These divisions produce unique NDJF SST patterns for each phase that are significantly correlated with MAMJ precipitation in the same phase.

Table 1. MEI Ranges for Phase Bins

Phase	MEI Range	Number of Years
La Nina	−1.80 to −0.84	17
Neutral Negative	−0.79 to 0.00	32
Neutral Positive	0.00–1.06	27
El Nino	1.07–2.12	14

4.2. Steps for Modeling Each Nino Index Phase

Each phase is evaluated individually, resulting in four unique predictive models. At the end of February of each year, the appropriate phase may be observed based on the MEI, and the associated MEI phase model is selected to predict the following spring months.

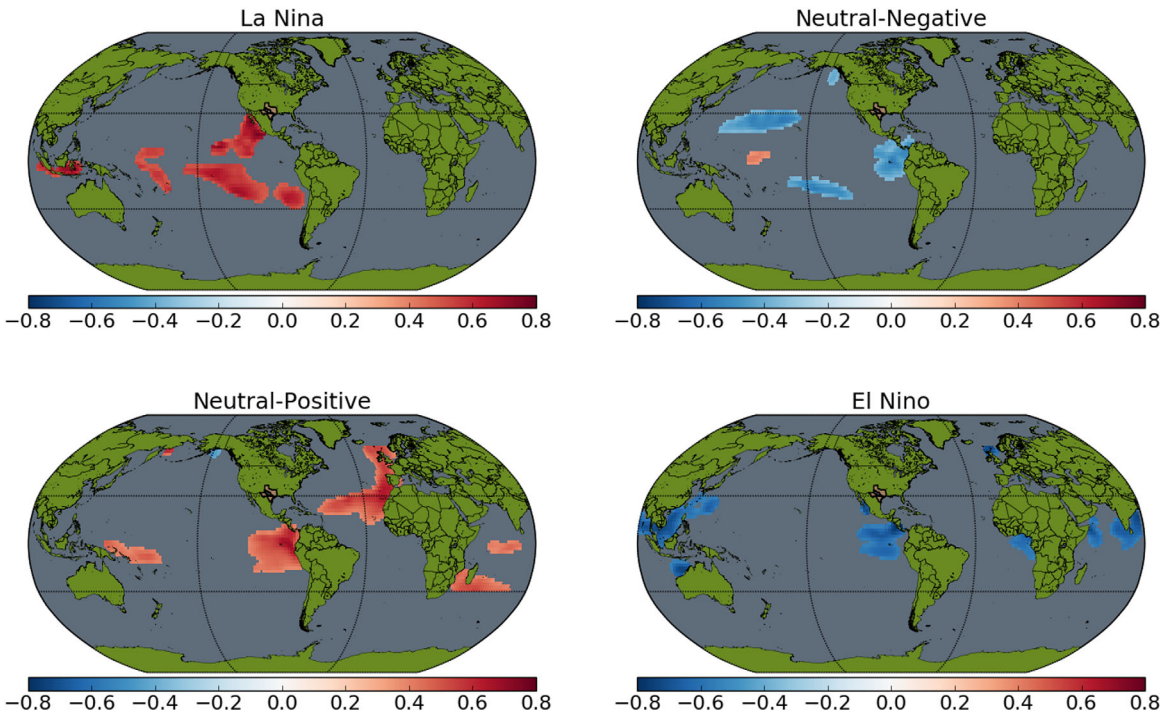

Figure 5. Correlation maps of MAMJ precipitation with NDJF SSTs for individual phases, clockwise from top left: La Nina, Neutral Negative, Neutral Positive, and El Nino.

The steps to create models for each phase are outlined below. A separate phase is also evaluated as a baseline comparison—the "all-years" phase (Figure 2), which includes all the years without binning as is more common when constructing these types of statistical models.

4.2.1. Identify SST Predictor Regions

Using correlation maps between total seasonal precipitation and preseason gridded SSTs, regions correlated at the 95% significance level for each phase are identified (Figure 5). The authors acknowledge that this methodology is likely to produce some spurious correlations; even though the SST grid cells are certainly not entirely independent realizations, it would not be surprising to see "significant" grid points emerge when correlating a time series with a huge matrix of variables (16,020 grid cells). To address this issue and increase confidence that the correlations shown are indicative of an underlying physical connection, Monte Carlo (MC) techniques are employed (detailed in section 4.3).

4.2.2. Principal Component Regression

Using the suite of predictors identified in section 4.2.1, a principal component regression (PCR) prediction modeling approach is adopted. PCA is conducted on the entire predictor field (all SST grid points that are not masked), and a subset (n) of the resulting PCs are retained as predictors in a linear regression model (equation (1)). North's rule of thumb [*North et al*., 1982] is utilized to determine the number of PCs to retain; however, in each phase, only the first PC is selected, explaining between 65% and 78% of the variance in selected SST grid points.

4.2.3. Cross-Validated Model Hindcasts

To evaluate model performance, it is common to conduct a hindcast (prediction of historical years for which observations are available for comparison) and apply performance metrics. To eliminate artificial skill and potential persistence, a cross-validation procedure is utilized [*Barnston*, 1994]. Thus, for one phase at a time, the raw data for the year being hindcast are dropped and new PCs and EOFs are generated using the remaining data from other years in that particular phase. A regression coefficient (β) is obtained via equation (1) for the PC (predictor) retained as prescribed in section 4.2.2, along with an intercept (α) and associated error(ε).

$$Observation = \beta \times PC_1 + \alpha + \varepsilon \tag{1}$$

It should be noted that region selection is *not* cross validated—that is, all the years in a phase are used to generate the correlation maps in section 4.2.1, which is the selection mechanism for grid points that enter

Table 2. Regression Diagnostics for La Nina Phase

	Max	Min	Median
β	17.0	9.8	14.6
p value	0.0001	0.001	0.0003

the PCR. The cross-validation procedure generates a new regression slope coefficient (β) every round; diagnostics for the La Nina phase are presented in Table 2, indicating that β is consistently significant. No phase exhibits a median p value greater than 0.007 for their respective slope or intercept coefficient, indicating significant regression parameters for each phase. Intercept coefficients (α) were also significant by the same criteria.

The raw data initially dropped are projected onto the appropriate EOF pattern to produce a predictor value (X), which is utilized by the regression model to produce a hindcast for the given year via equation (2). This is repeated for all years in each phase.

$$Hindcast = \beta * X + \alpha \tag{2}$$

4.2.4. Create Hindcast Ensembles

To characterize model error, a kernel density estimator is used to create a probability density function (PDF) based on model residuals (which are slightly non-Gaussian due to the leave-one-out cross-validation procedure). For each year, a hindcast ensemble is generated by randomly sampling 1000 times from the model residual PDF and adding the samples to the deterministic hindcast value generated by equation (2) on a phase-by-phase basis. The result is a PDF for each hindcast, or an "ensemble prediction" [*Block and Rajagopalan*, 2007]. Distributions for each yearly hindcast may then be compared with the climatological (historical average) distribution to judge the model skill.

4.2.5. Create Tercile Hindcast Predictions

Tercile predictions, in which three intervals are delineated (below normal, normal, and above normal) based on the climatological precipitation distribution, are common in seasonal climate forecasts. Frequently, equally sized categories are selected, as is the case in this study, however, unequal category size is also permissible. Figure 6a shows a PDF generated by a kernel density estimator based on the 90 years of LCRB data. Analytical integration techniques are then used to determine the precipitation values that create equally sized categories.

Without any forecast information, the probability that any of the three terciles would be observed in a given year is considered to be about one third. However, the distributions created in section 4.2.4 can be compared to the climatological PDF, and shifts in the probabilities of being in one of the terciles can be computed (Figure 6b). These categorical forecasts are commonly given in terms of [B N A], [6 38 57], is

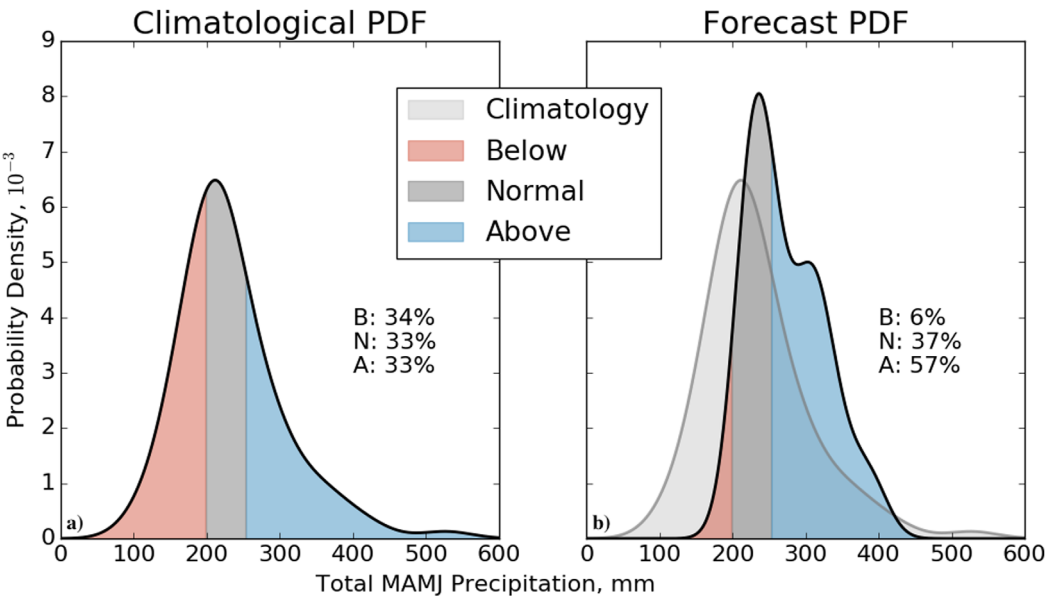

Figure 6. (a) Climatological PDF, and (b) an example Forecast PDF with the climatological underlaid in light gray.

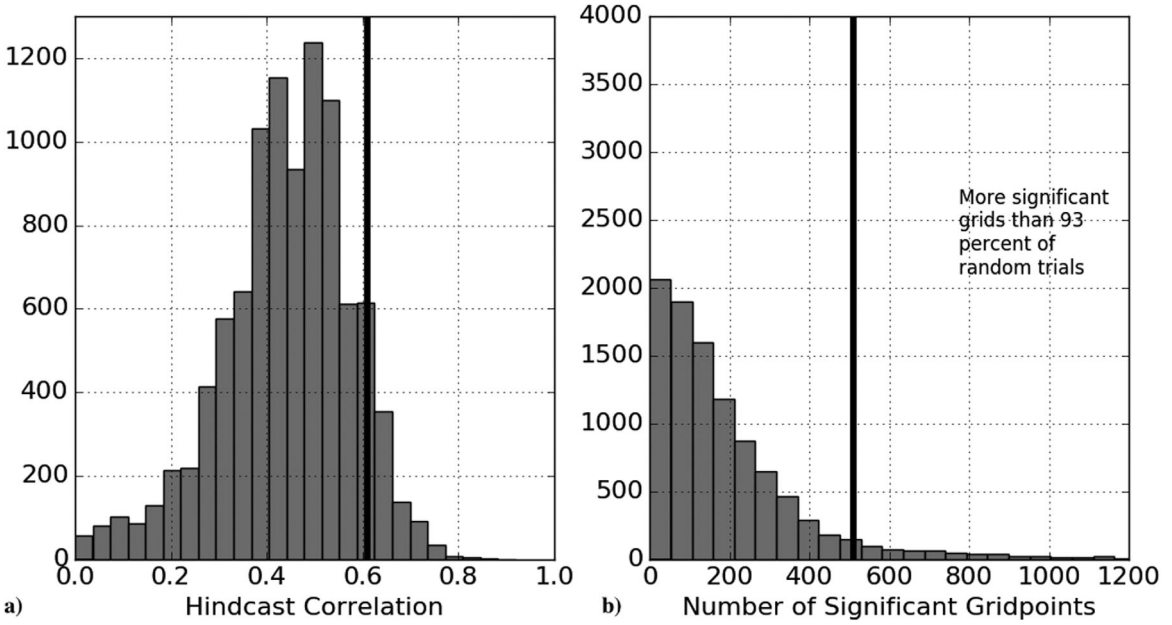

Figure 7. Monte Carlo results for the La Nina phase.

interpreted as 6% chance of below normal, 38% chance of near normal, and 57% chance of above normal precipitation for the upcoming season.

4.2.6. Evaluate Model Performance

Three metrics are evaluated to gauge the probabilistic model performance as compared to climatology: the standard Pearson correlation coefficient, the Heidke Skill Score (HSS) [*Heidke*, 1926], and Ranked Probability Skill Score (RPSS) [*Epstein*, 1969]. The skill scores are performance measures interpreted as a percentage improvement over some "reference" forecast. A standard reference forecast—the climatological forecast—is adopted here, with equal chances of B, N, or A, [34 33 33]. Values range from $-\infty$ to 1, with positive values indicating a model forecast that is more skillful than the reference forecast. A score of 0 indicates that the model shows the same skill as the reference forecast (climatology); a score of 1 indicates a perfect model forecast. For the unfamiliar reader, *Wilks* [1995] offers a detailed yet approachable explanation of the mechanics of HSS and RPSS. For this application, it is sufficient for the reader to understand that the HSS is based on model "hits," and that a "near miss" (i.e., A forecast, N observed) is penalized the same as a "far miss" (A forecast, B observed); RPSS penalizes forecasts increasingly as more probability is assigned to event categories farther from the actual outcome.

4.3. Monte Carlo Test

As briefly mentioned in section 4.2.1, caution must be exercised when correlating a single time series (LCRB precipitation) with a massive matrix of variables (180 $\times$ 89 SST grid cells). The MC method presented here demonstrates that the regions selected are indeed significant at a high level of confidence. The routine is as follows:

1. Randomly shuffle **only** the precipitation data (MEI and SST unchanged).
2. Bin the shuffled precipitation data using the unshuffled MEI (section 4.1).
3. Generate correlation maps and extract significant grid points (SGs) (section 4.2.1).
4. Record number of SGs per phase, record total for all four phases.
5. Run principal component regression (section 4.2.3).
6. Record the hindcast correlation.
7. Repeat 10,000 times.

The results from the MC are most easily interpreted by examining the histograms in Figure 7 (results from only one phase are displayed, the other phases exhibit similar characteristics). The histograms illustrate the distribution of hindcast correlations and number of SGs for the randomized data, and the vertical black line

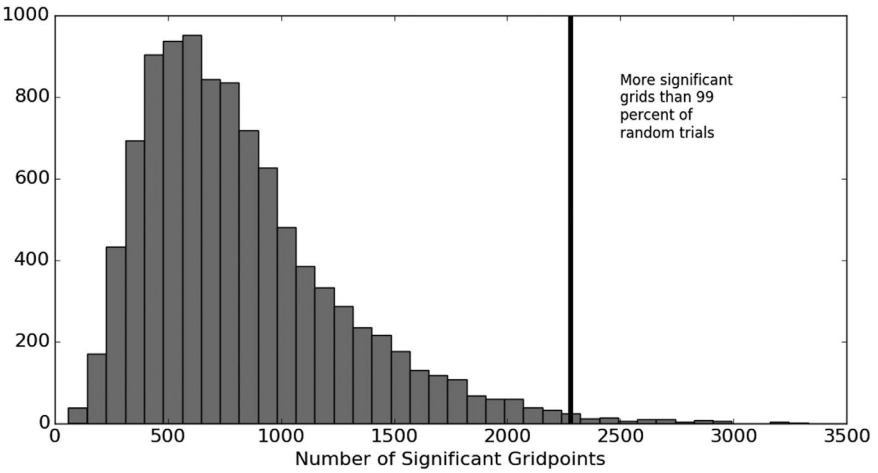

Figure 8. Monte Carlo results for total significant grid points among the four phases for each trial.

indicates the actual correlation and number of grid points for the La Nina phase. Looking at only the correlation, it may be difficult to justify that the actual results are significant; however, the number of SGs clearly indicates that large regions of SST anomalies are statistically significant. Spuriously correlated grid cells may produce similar hindcast correlations in the PCR scheme—this is the susceptibility of the method—however, the likelihood of large regions correlating by chance per phase is low. Depending on the phase, the SGs produced by the LCRB data are more numerous than those produced in 88–97% of the random trials. Additionally, Figure 8 indicates that on a per trial basis, the total number of SGs summed across all four phases is greater than the total SGs for the actual data only 97 out of 10,000 times (99% significant), indicating that if one phase in a random trial shows a large number of SGs, the likelihood that all the other phases also show a large number of SGs is low. However, the unshuffled data from each of the four phases show large regions of SGs. These results provide a high degree of confidence that the relationships outlined here are indeed significant and potentially indicative of an underlying physical process (large regions of anomalous SSTs) driving the correlation, however, the exact physical processes responsible are not explored here.

5. NIPA Framework Performance

First, NIPA performance will be examined through the per-phase correlation coefficient and a comparison to the "one-phase" or "all-years" model, where the data are not split into phases but all other NIPA steps are followed. Next, the skill scores (comparison to "climatological" forecast) will be presented. Last, NIPA performance will be compared to that of the North American Multi-Model Ensemble via the RPSS skill score.

5.1. Skill Scores

For each phase, the first PC was consistently the only PC selected as a predictor using North's rule of thumb, indicating that in each phase it represents the dominant mode of variability among the predictor field. For example, in the La Nina phase, PC-1 explains 65% of the variability, and correlates better with precipitation ($r = 0.72$) than any individual SST grid point; PC-2 explains 9% of the variability and does not significantly correlate ($r = 0.21$) with observed precipitation. Correlations between cross-validated model hindcasts and observations vary by phase (Figure 9).

The hindcast precipitation from all phases, once reassembled (Figure 10a, $r = 0.47$), illustrates a stronger correlation with observed MAMJ seasonal precipitation than a typical one-phase model (Figure 10b, $r = 0.21$), and better captures observed inter-annual variability (Figure 11). As is common with linear models, hindcast values have a smaller standard deviation than the original time series, generally underestimating extreme years. The categorical skill scores can provide useful information despite this reduced variance.

The PDFs generated from the hindcast ensembles for each year (section 4.2.4) can be compared with the climatological distribution, and probabilities of drier or wetter conditions can be computed (Figure 6).

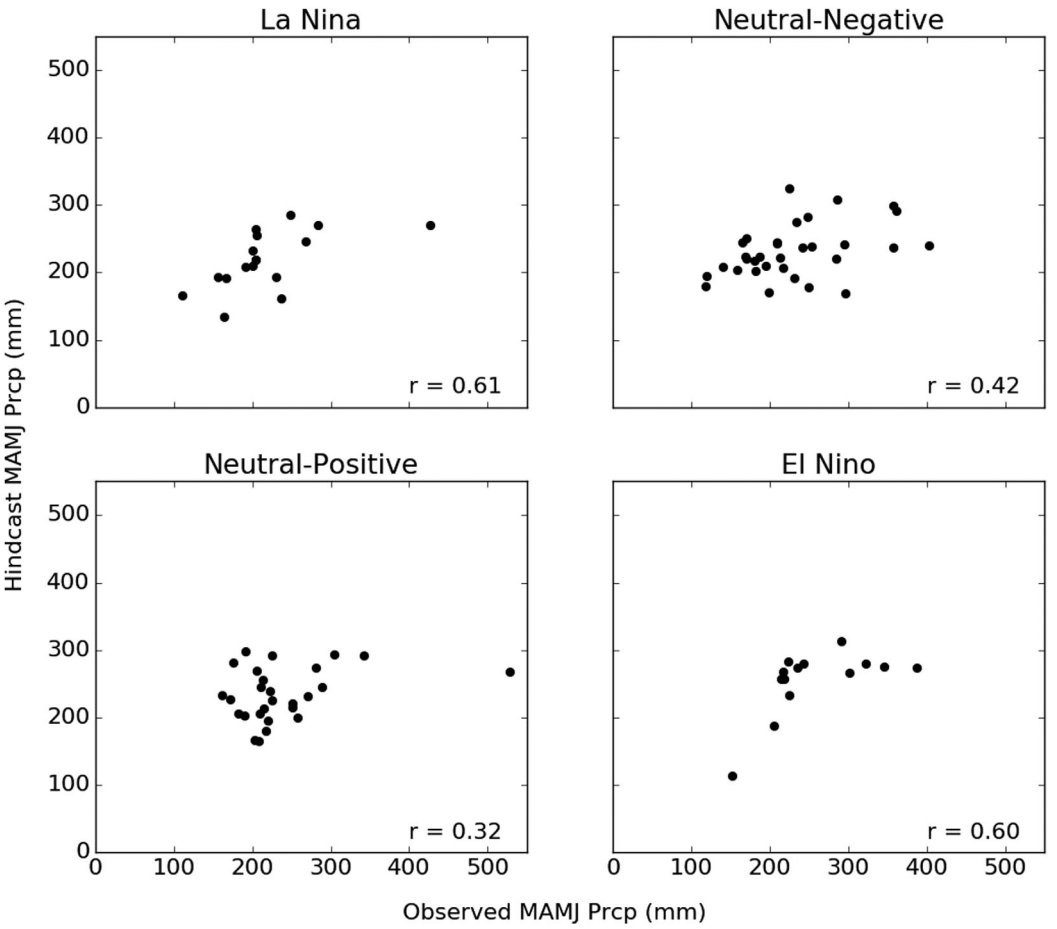

Figure 9. Scatterplots of model hindcasts and observations for individual phases, clockwise from top left: La Nina, Neutral Negative, Neutral Positive, and El Nino.

Though the deterministic hindcast may underestimate or overestimate precipitation, ensembles, representing model uncertainty, can indicate the relative expected shift in precipitation as compared with climatology.

For the 90 reassembled years, skill scores illustrate that utilizing NIPA shows mild improvement over simply using the reference climatological forecasts (median RPSS = 0.12, HSS = 0.28.) However, of primary concern

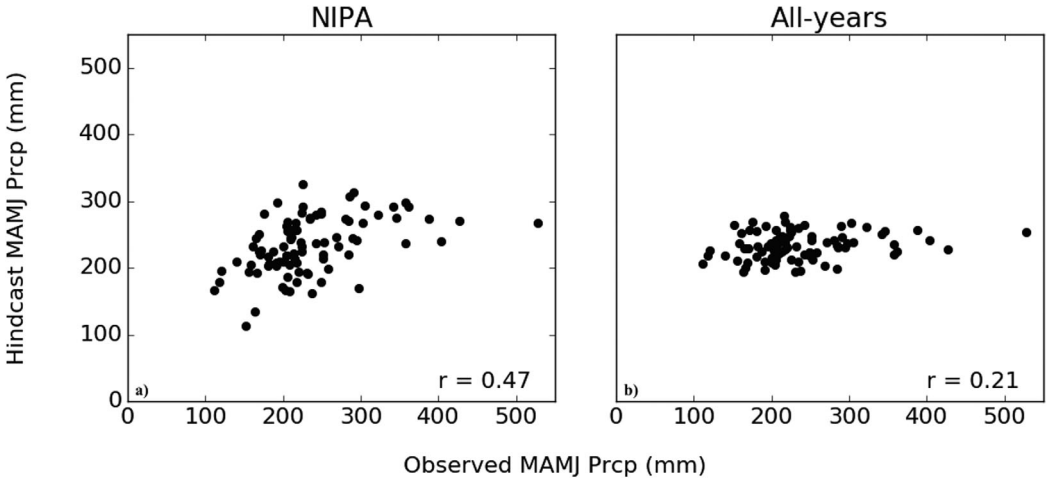

Figure 10. Scatterplots of (a) composite hindcasts from NIPA and (b) hindcasts from a traditional "one-phase" or "all-years" model.

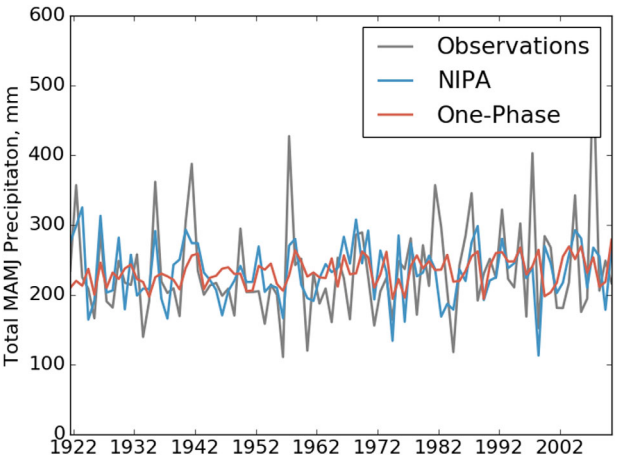

Figure 11. Time series comparing observations to "one-phase" and NIPA hindcast results.

to planners are skillful forecasts of extreme years. The median RPSS for the 10 driest (wettest) years is 0.65 (0.81), indicating very strong performance in the most extreme years (Table 3); HSS is 0.70 (0.55). Both RPSS and HPSS performance deteriorates somewhat as the "extreme" categories are enlarged, however, only to a relatively minimal degree, again indicating that the prediction model tends to shift the odds consistently well for both dry and wet observed conditions. The near-normal category does not produce comparably high skill score values, however, from a planning perspective, this is less critical than skillfully predicting extreme conditions.

When applying the traditional approach of including all years in one phase (Figure 10b), the median RPSS for the 10 driest (wettest) years drops to 0.04 (0.26); overall, RPSS falls below zero. Thus, the NIPA framework illustrates a marked improvement over climatological forecasts as well as the traditional approach of examining all years together in one "phase," especially for dry and wet years.

5.2. North American Multi-Model Ensemble Comparison

The NMME is an experimental multimodel seasonal forecasting system consisting of coupled models from U.S. modeling centers including NOAA/NCEP, NOAA/GFDL, IRI, NCAR, NASA, and Canada's CMC. The coupled models produce a total of 129 ensemble members, and the data are available from 1982 to 2010. For consistency with the NIPA approach, NMME ensembles are compiled from predictions issued at the end of February, and comprise total MAMJ precipitation over the region encompassing the LCRB. Figure 12 illustrates the comparison for 2 years—one where NIPA excels and one where the NMME excels. However, when looking at the RPSS scores for each method (Table 4), NIPA is clearly superior to NMME overall, with particularly significant performance improvements over the NMME when considering the 10 driest and 10 wettest years observed between 1982 and 2010.

6. Summary and Discussion

In this paper, the Nino Index Phase Analysis (NIPA) framework for seasonally forecasting hydroclimatic variables is introduced, and is demonstrated for spring precipitation in the LCRB. In utilizing the MEI to establish four distinct phases indicative of the "mean state" of the A/O system, correlation maps combined with principal component regression can uncover informative regions of antecedent SST anomalies that may otherwise be overlooked. The maps illustrate clear distinctions among phases in tropical Pacific and North Pacific SSTs, with several geographic regions expressing opposing signs in some cases (Figure 5). These nonuniformities and varying patterns indicate differing predictive SST signals for precipitation in the LCRB within each MEI phase.

Table 3. RPSS and HSS Values for Various Bins of Dry, Normal, and Wet Years

	Number of Driest/Wettest Years					
	10		20		30	
	RPSS	HSS	RPSS	HSS	RPSS	HSS
Dry	0.65	0.70	0.35	0.33	0.55	0.45
Wet	0.81	0.55	0.80	0.48	0.41	0.40

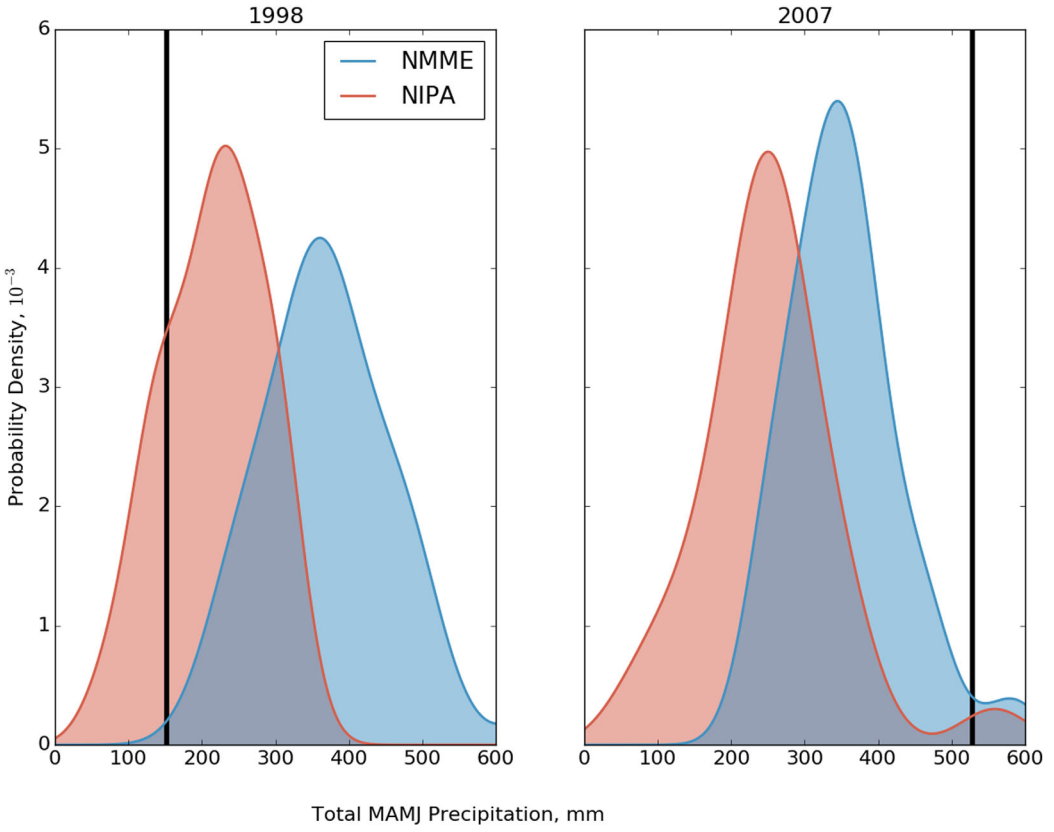

Figure 12. Comparison of NIPA to NMME ensembles.

Though linear models may be regarded as simple approaches, unable to capture the nonlinear dynamics driving atmospheric processes, the NIPA framework performs remarkably well overall. The combination of distinct ENSO phases and PCR isolates and highlights key climate signals and relevant teleconnections; the MC analysis provides confidence that the signals are not simply spurious correlations. These signals have the potential to significantly bolster season-ahead prediction skill, especially in extreme years when management is most difficult for decision makers, demonstrated by the high RPSS and HSS skill scores in the driest and wettest years. For the case study illustrated, NIPA provides good prospects for better capturing inter-annual variability compared with climatological forecasts, typical one-phase ("All Years") forecasts, and the dynamical NMME.

The LCRB, when viewed through the lens of composite analysis, is understood to be drier during La Nina and wetter during El Nino—and indeed the mean spring precipitation is statistically significantly lower for the La Nina years than the El Nino years. However, violin plots of observed precipitation (Figure 3) illustrate that there is significant variability in total observed precipitation within both phases. This is in agreement with *Larkin and Harrison* [2005], who find that across most locations in the U.S., upper quintile precipitation extremes with the same sign as a composite El Nino average occur in, at best, only half of the El Nino years identified. The NIPA methodology goes a step beyond using composite analysis to analyze ENSO events; it is intended to diagnose and predict the within-phase variance that is essentially masked when using compositing techniques.

Figure 5 (top left) indicates that within La Nina events (which are characterized by cooler than average tropical Pacific SSTs), the strength of the La Nina event is a good predictor for spring precipitation in the LCRB; the SST anomalies within the entire spatial domain typical of ENSO (excepting the cold tongue) are highly positively correlated with spring precipitation during La Nina events. Positive correlation implies that "warmer"

Table 4. RPSS Scores for NMME and NIPA, 1982–2010

	10 Driest	10 Wettest	All
NIPA	0.47	0.62	−0.15
NMME	−0.32	0.23	−0.35

SST anomalies (a weaker La Nina event) accurately predict higher precipitation, whereas "cooler" SST anomalies (a stronger La Nina event) predict dry conditions; the hindcasts for this phase correlate at $r = 0.61$, and are not driven by any single outlier. This stands in contrast to conventional thought that La Nina brings dry conditions to Texas. As an alternative to a static composite analyses, this methodology provides a dynamic way to predict precipitation in the LCRB conditioned on the phase of ENSO.

In light of these findings, the authors would like to propose that ENSO, as the dominant mode of variability in the A/O system, sets the stage through the large perturbation it causes in convective heating over the equatorial Pacific [*Kushnir et al*., 2006], disrupting atmospheric and oceanic circulations in the tropical Pacific and elsewhere [*Chiodi and Harrison*, 2013]. NIPA is meant to be utilized as a data-driven tool to help diagnose the climate drivers once the stage is set. Though it does not provide exact physical insight into the complex dynamics at play, it does highlight that intraphase ENSO variability may be large, yet predictable, for a given location.

The design of the NIPA framework allows it to easily be applied to hydroclimatic variables other than precipitation, such as wind speed, temperature, streamflow, etc., making it potentially appealing for a variety of applications. The authors acknowledge that forecasting one variable for one season in one location does not guarantee broad applicability; however, the MC methods implemented here provide a way to determine when significantly correlated SST grid cells are likely spurious or plausibly driven by cohesive areas of A/O variability. Ongoing work will explore the limits of applicability. Further testing and refinement of NIPA will continue to aid in better understanding its range of predictive capabilities, nuances, and shortcomings, ultimately in order to provide decision makers with reliable, useable, hydroclimatic forecast information on seasonal time scales.

Acknowledgments

Two of the data sets used in this study (NOAA's ERSSTV3b [*Smith et al*., 2008] and PRISM precipitation [*Daly et al*., 2008]) were obtained from the free to access IRI Data Library (iridl.ldeo.columbia.edu). The Hadley Centre's HADSLP2r [*Allan and Ansell*, 2006] was obtained from their website at www.metoffice.gov.uk. The Python code to run NIPA is available from Brian Zimmerman's Bitbucket repository at www.bitbucket.org/bgzimmerman. The authors acknowledge NOAA support through the SARP program (project NA130AR4310153) and to Ron Anderson and Bob Rose of the LCRA for their assistance and comments throughout the project.

References

Abbot, J., and J. Marohasy (2012), Application of artificial neural networks to rainfall forecasting in Queensland, Australia, *Adv. Atmos. Sci.*, *29*(4), 717–730.

Abbot, J., and J. Marohasy (2014), Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks, *Atmos. Res.*, *138*, 166–178.

Allan, R., and T. Ansell (2006), A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850-2004, *J. Clim.*, *19*(22), 5816–5842.

Allan, R. P., and B. J. Soden (2008), Atmospheric warming and the amplification of precipitation extremes, *Science*, *321*(5895), 1481–1484.

Anderson, J., H. van den Dool, A. Barnston, W. Chen, W. Stern, and J. Ploshay (1999), Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction, *Bull. Am. Meteorol. Soc.*, *80*(7), 1349–1361.

Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata (2007), El Niño Modoki and its possible teleconnection, *J. Geophys. Res.*, *112*, C11007, doi:10.1029/2006JC003798.

Barnston, A. G. (1994), Linear statistical short-term climate predictive skill in the Northern Hemisphere, *J. Clim.*, *7*(10), 1513–1564.

Bazo, J., M. D. L. N. Lorenzo, and R. Porfirio da Rocha (2013), Relationship between monthly rainfall in NW Peru and tropical sea surface temperature, *Adv. Meteorol.*, *2013*, 1–9.

Becker, E., H. van den Dool, and Q. Zhang (2014), Predictability and forecast skill in NMME, *J. Clim.*, *27*(15), 5891–5906.

Block, P., and L. Goddard (2012), Statistical and dynamical climate predictions to guide water resources in Ethiopia, *J. Water Resour. Plann. Manage.*, *138*(3), 287–298.

Block, P., and B. Rajagopalan (2007), Interannual variability and ensemble forecast of Upper Blue Nile Basin *Kiremt* season precipitation, *J. Hydrometeorol.*, *8*(3), 327–343.

Block, P. J., F. A. Souza Filho, L. Sun, and H.-H. Kwon (2009), A streamflow forecasting framework using multiple climate and hydrological models, *J. Am. Water Resour. Assoc.*, *45*, 828–843.

Board, T. W. D. (2012), Water for Texas: 2012 state water plan, technical report, Tex. Water Dev. Board, Austin, Tex. [Available at https://www.twdb.texas.gov/waterplanning/swp/2012/index.asp.]

Branković, Č., T. N. Palmer, and L. Ferranti (1994), Predictability of seasonal atmospheric variations, *J. Clim.*, *7*(2), 217–237.

Camberlin, P., S. Janicot, and I. Poccard (2001), Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic vs. ENSO, *Int. J. Climatol.*, *21*(8), 973–1005.

Capotondi, A., A. T. Wittenberg, M. Newman, E. Di Lorenzo, J.-Y. Yu, P. Braconnot, J. Cole, B. Dewitte, B. Giese, and E. Guilyardi (2014), Understanding ENSO diversity, *Bull. Am. Meteorol. Soc.*, *96*(6), 921–938.

Chen, C.-J., and A. P. Georgakakos (2014), Hydro-climatic forecasting using sea surface temperatures: Methodology and application for the southeast US, *Clim. Dyn.*, *42*(11-12), 2955–2982.

Chiodi, A. M., and D. E. Harrison (2013), El Niño impacts on seasonal US atmospheric circulation, temperature, and precipitation anomalies: The OLR-event perspective*, *J. Clim.*, *26*(3), 822–837.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *Int. J. Climatol.*, *28*(15), 2031–2064.

Enfield, D. B., A. M. Mestas-Nuñez, and P. J. Trimble (2001), The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US, *Geophys. Res. Lett.*, *28*(10), 2077–2080.

Epstein, E. S. (1969), A scoring system for probability forecasts of ranked categories, *8*(6), 985–987.

Frankignoul, C. (1985), Sea surface temperature anomalies, planetary waves, and air-sea feedback in the middle latitudes, *23*, 357–390.

Gershunov, A. (1998), ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Implications for long-range predictability, *J. Clim.*, *11*(12), 3192–3203.

Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona (2005), A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts, *41*, W10410, doi:10.1029/2004WR003467.
Hartmann, H., S. Becker, and L. King (2008), Predicting summer rainfall in the Yangtze River basin with neural networks, *Int. J. Climatol.*, *28*(7), 925–936.
Heidke, P. (1926), Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst, *Geografiska Annaler*, *8*, 301–349.
Infanti, J. M., and B. P. Kirtman (2014), Southeastern U.S. rainfall prediction in the North American multi-model ensemble, *J. Hydrometeorol.*, *15*(2), 529–550.
Jiménez Cisneros, B. E., T. Oki, N. W. Arnell, G. Benito, J. G. Cogley, P. Döll, T. Jiang, and S. S. Mwakalila (2014), Freshwater resources, in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change*, edited by C. B. Field et al., pp. 229–269, Cambridge Univ. Press, Cambridge, U. K.
Kirtman, B. P., et al. (2013), The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction, *Bull. Am. Meteorol. Soc.*, *95*(4), 585–601.
Kumar, A., and M. Hoerling (2011), Annual cycle of Pacific-North American seasonal predictability associated with different phases of ENSO, *J. Clim.*, *11*(10), 3295–3308.
Kushnir, Y., W. A. Robinson, P. Chang, and A. W. Robertson (2006), The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability*, *J. Clim.*, *19*(23), 5949–5970.
Larkin, N. K., and D. Harrison (2005), On the definition of El Niño and associated seasonal average U.S. weather anomalies, *Geophys. Res. Lett.*, *32*, L13705, doi:10.1029/2005GL022738.
Lloyd-Hughes, B., and M. A. Saunders (2002), Seasonal prediction of European spring precipitation from El Nio-Southern Oscillation and local sea-surface temperatures, *Int. J. Climatol.*, *22*(1), 1–14.
Markowski, G. R., and G. R. North (2003), Climatic influence of sea surface temperature: Evidence of substantial precipitation correlation and predictability, *J. Hydrometeorol.*, *4*(5), 856–877.
McCabe, G. J., M. A. Palecki, and J. L. Betancourt (2004), Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States, *Proc. Natl. Acad. Sci. U. S. A.*, *101*(12), 4136–4141.
Mo, K. C., and D. P. Lettenmaier (2014), Hydrologic prediction over the conterminous United States using the national multi-model ensemble, *J. Hydrometeorol.*, *15*(4), 1457–1472.
Neelin, J. D., A. Bracco, H. Luo, J. C. McWilliams, and J. E. Meyerson (2010), Considerations for parameter optimization and sensitivity in climate models, *Proc. Natl. Acad. Sci. U. S. A.*, *107*(50), 21,349–21,354.
North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng (1982), Sampling errors in the estimation of empirical orthogonal functions, *Mon. Weather Rev.*, *110*(7), 699–706.
Phillips, I. D., and J. Thorpe (2006), Icelandic precipitation—North Atlantic sea-surface temperature associations, *Int. J. Climatol.*, *26*(9), 1201–1221.
Quan, X., M. Hoerling, J. Whitaker, G. Bates, and T. Xu (2006), Diagnosing sources of US seasonal forecast skill, *J. Clim.*, *19*(13), 3279–3293.
Rajagopalan, B., E. Cook, U. Lall, and B. K. Ray (2000), Spatiotemporal variability of ENSO and SST teleconnections to summer drought over the United States during the twentieth century, *J. Clim.*, *13*(24), 4244–4255.
Regonda, S. K., B. Rajagopalan, M. Clark, and E. Zagona (2006), A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin, *42*, W09404, doi:10.1029/2005WR004653.
Ropelewski, C., and M. Halpert (1986), North American precipitation and temperature patterns associated with the El Nino/Southern Oscillation (ENSO), *Mon. Weather Rev.*, *114*, 2352–2362.
Ropelewski, C., and C. Halpert (1987), Global and regional scale precipitation patterns associated with the El Nino/Southern Oscillation, *Mon. Weather Rev.*, *115*(8), 1606–1626.
Sadoff, C., and M. Muller (2009), *Water Management, Water Security and Climate Change Adaptation: Early Impacts and Essential Responses*, Global Water Partnership, Stockholm.
Shabbar, A., and W. Skinner (2004), Summer drought patterns in Canada and the relationship to global sea surface temperatures, *J. Clim.*, *17*(14), 2866–2880.
Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore (2008), Improvements to NOAA's historical merged land–Ocean surface temperature analysis (1880–2006), *J. Clim.*, *21*(10), 2283–2296.
Takahashi, K., A. Montecinos, K. Goubanova, and B. Dewitte (2011), ENSO regimes: Reinterpreting the canonical and Modoki El Niño, *Geophys. Res. Lett.*, *38*, L10704, doi:10.1029/2011GL047364.
Wang, H., S. Schubert, M. Suarez, and R. Koster (2010), The physical mechanisms by which the leading patterns of SST variability impact U.S. precipitation, *J. Clim.*, *23*(7), 1815–1836.
Wang, L., W. Chen, and R. Huang (2008), Interdecadal modulation of PDO on the impact of ENSO on the east Asian winter monsoon, *Geophys. Res. Lett.*, *35*, L20702, doi:10.1029/2008GL035287.
White, W. B., A. Gershunov, and J. Annis (2008), Climatic influences on midwest drought during the twentieth century, *J. Clim.*, *21*(3), 517–531.
Wilks, D. S. (1995), *Statistical Methods in the Atmospheric Sciences: An Introduction*, Academic, Elsevier Science.
Wolter, K., and M. S. Timlin (1993), Monitoring ENSO in COADS with a seasonally adjusted principal component index, paper presented at the 17th Climate Diagnostics Workshop, CIMMS and the School of Meteorology, University of Oklahoma.
Wolter, K., and M. S. Timlin (2011), El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI. ext), *Int. J. Climatol.*, *31*(7), 1074–1087.