Forced and unforced variability of twentieth century North American droughts and pluvials

Benjamin I. Cook · Edward R. Cook ·
Kevin J. Anchukaitis · Richard Seager ·
Ron L. Miller

Abstract Research on the forcing of drought and pluvial events over North America is dominated by general circulation model experiments that often have operational limitations (e.g., computational expense, ability to simulate relevant processes, etc). We use a statistically based modeling approach to investigate sea surface temperature (SST) forcing of the twentieth century pluvial (1905–1917) and drought (1932–1939, 1948–1957, 1998–2002) events. A principal component (PC) analysis of Palmer Drought Severity Index (PDSI) from the North American Drought Atlas separates the drought variability into five leading modes accounting for 62% of the underlying variance. Over the full period spanning these events (1900–2005), the first three PCs significantly correlate with SSTs in the equatorial Pacific (PC 1), North Pacific (PC 2), and North Atlantic (PC 3), with spatial patterns (as defined by the empirical orthogonal functions) consistent with our understanding of North American drought responses to SST forcing. We use a large ensemble statistical modeling approach to determine how successfully we can reproduce these drought/pluvial events using these three modes of variability. Using Pacific forcing only (PCs 1–2), we are able to reproduce the 1948–1957 drought and 1905–1917 pluvial above a 95% random noise threshold in over 90% of the ensemble members; the addition of Atlantic forcing (PCs 1–2–3) provides only marginal improvement. For the 1998–2002 drought, Pacific forcing reproduces the drought above noise in over 65% of the ensemble members, with the addition of Atlantic forcing increasing the number passing to over 80%. The severity of the drought, however, is underestimated in the ensemble median, suggesting this drought intensity can only be achieved through internal variability or other processes. Pacific only forcing does a poor job of reproducing the 1932–1939 drought pattern in the ensemble median, and less than one third of ensemble members exceed the noise threshold (28%). Inclusion of Atlantic forcing improves the ensemble median drought pattern and nearly doubles the number of ensemble members passing the noise threshold (52%). Even with the inclusion of Atlantic forcing, the intensity of the simulated 1932–1939 drought is muted, and the drought itself extends too far into the southwest and southern Great Plains. To an even greater extent than the 1998–2002 drought, these results suggest much of the variance in the 1932–1939 drought is dependent on processes other than SST forcing. This study highlights the importance of internal noise and non SST processes for hydroclimatic variability over North America, complementing existing research using general circulation models.

Keywords Drought · Forcing · Predictability · North America · Dust Bowl

1 Introduction

Persistent, multi-year drought and pluvial events have been a recurrent feature of North American hydroclimate since at least the time of the Medieval Climate Anomaly (Cook et al. 1999, 2004, 2007; Fye et al. 2003). For the United States (US), the monetary impact of droughts far exceeds the costs of other natural disasters, including fires,

B. I. Cook (✉) · R. L. Miller
NASA Goddard Institute for Space Studies,
2880 Broadway, New York, NY 10025, USA
e-mail: bc9z@ldeo.columbia.edu

E. R. Cook · K. J. Anchukaitis · R. Seager
Lamont-Doherty Earth Observatory,
Palisades, NY 10964, USA

earthquakes, and hurricanes (Cook et al. 2007; Herweijer et al. 2007). Pluvials (anomalous wet periods) can have deleterious impacts as well (Trenberth and Guillemot 1996), and have also been instrumental in shaping resource use policies in the western US (Fye et al. 2003; Woodhouse et al. 2005). The societal importance of hydroclimatic events, coupled with evidence for recent trends towards increased desiccation (e.g., Seager et al. 2007; Touchan et al. 2010), has motivated a high level of interest in understanding the forcing and predictability of droughts and pluvials over North America (NA) (Herweijer et al. 2006; Seager et al. 2005b; Seager 2007; Schubert et al. 2009).

Experiments using general circulation models (GCM) have played a critical role in advancing our understanding of the dynamical mechanisms governing hydroclimatic variability (Hoerling and Kumar 2003; Schubert et al. 2009; Seager et al. 2003, 2005a; Seager 2007). When forced with either observed (Herweijer et al. 2006; Seager et al. 2005b) or idealized (Schubert et al. 2009) sea surface temperature (SST) patterns, GCMs are generally able to reproduce observed or expected patterns of drought with high fidelity (Herweijer et al. 2006; Seager et al. 2005b), albeit with some notable exceptions over NA (Cook et al. 2008; 2009; Seager et al. 2008a,b). Specifically, these experiments have helped illuminate the important role played by SSTs in the Pacific (Herweijer et al. 2006, 2007; Seager et al. 2005b; Seager 2007) and North Atlantic (Enfield et al. 2001; Kushnir et al. 2010; Mo et al. 2009; Wang et al. 2006) ocean basins.

Using GCMs to study drought and pluvial dynamics provides several major advantages over empirically based approaches. Based upon first principals, results from GCMs can be independently compared and verified against observational data sets and empirical studies, providing direct and independent insight into the physical mechanisms. GCMs are also informative for historical studies during time periods when observations are sparse or unavailable, including droughts during the Medieval Climate Anomaly (Feng et al. 2008; Seager et al. 2008a) and the nineteenth century (Herweijer et al. 2006). In addition to being forced by realistic boundary conditions and forcings, these models can also be driven with idealized forcings that are often larger or more persistent than observed (Schubert et al. 2009), maximizing the signal to noise ratio and allowing for clear dissection of causal mechanisms. These advantages are balanced by a number of significant shortcomings. GCMs are computationally expensive, limiting the number of ensembles or experiments that can be conducted. The models themselves are imperfect, often missing important processes or simulating existing processes incorrectly. Finally, it can be difficult to satisfyingly diagnose causal mechanisms in the models themselves

when dealing with competing sources of variance, especially in cases where the signal to noise ratio is low.

Most statistically based studies of drought and pluvial dynamics over NA have focused on describing simple statistical associations between drought variability and various forcing factors (e.g., Andreadis et al. 2005; Cook et al. 1997; Mo and Schemm 2008). Here we investigate the variability and forcing of drought and pluvial events over NA using a statistically based modeling approach that allows for large ensemble simulations and empirical separation of modes representing distinct SST related sources of drought variability. Our study offers a statistically based complement to existing research using GCM experiments, and we use previous insights into the dynamics and physics of drought forcing to guide the interpretation of results from our analyses. We focus on the four persistent events during the twentieth century, events that, for the most part, have been extensively investigated using GCM modeling experiments and are typical of events that have occurred in the more distant past. Our analysis is organized around two research questions: (1) what is the relative importance of various regions of SST forcing for simulating these drought and pluvial events? and (2) how much of the variability during these events can be ascribed to the SST forcing?

2 Data and methods

2.1 North American Drought Atlas

The North American Drought Atlas (NADA) (http://www.ncdc.noaa.gov/paleo/pdsi.html) version 2a (Cook et al. 2007) is a tree ring proxy based reconstruction of the Palmer Drought Severity Index (PDSI) (Palmer 1965). PDSI is an index of drought, incorporating supply (precipitation) and demand (evaporation as a function of temperature) information in it's calculation. It is locally normalized around a mean of zero, so that the standardized anomalies between different regions are directly comparable. Positive values of PDSI indicate wetter than normal (pluvial) conditions and negative values indicate drier (drought) conditions, with a typical range in values from -5 to $+5$. Tree ring chronologies and PDSI are effective at integrating over large spatial and temporal scales, making them ideal for use in studies of regional drought variability at seasonal timescales. The NADA product reconstructs PDSI for the summer (June–July–August) season using 1,821 tree ring chronologies, over a maximum of 286 $2.5° \times 2.5°$ gridpoints. This product is well validated and versions of the NADA have been used extensively in the study of North American drought variability (Cook et al. 1999; 2004; 2007; Fye et al. 2003; Herweijer et al. 2007). To develop our statistical model, we use data from the

drought atlas for years 1400 C.E.–2005 C.E., representing 251 continuous in time grid point locations. The time span chosen represents a tradeoff between the number of years available in the NADA and the spatial coverage of the dataset, which progressively decreases further back in time. While our resulting analysis is focused on the four twentieth century events, using over 600 years in our model development helps us maximize the representativeness of our modes of drought variability in our analysis and model (described next section). Our region of interest is western North America ($125°W - 95°W, 25°N - 50°N$), a region of high drought variability encompassing the primary features of all the major twentieth century events. This region is outlined (dashed box) in Figs. 6, 7, 8 and 9.

2.2 Principal component analysis and modeling

In the drought modeling literature, it is common to conduct GCM experiments using idealized SST patterns (e.g., Schubert et al. 2009; Seager et al. 2005b). In a typical experiment, observed SSTs might be prescribed over the tropical Pacific or North Atlantic, with the rest of the ocean held to climatological SSTs or modeled as a thermodynamic mixed layer. Idealized SST patterns representing phases of ENSO or other SST modes are also often used. The point of such experiments is not to necessarily generate the most realistic patterns of observed drought variability; rather they are used to inform the underlying dynamics and the importance of various sources of forcing for specific drought and pluvial events. Our modeling approach and philosophy is quite similar. But instead of using a GCM with idealized SST patterns, we use a statistical model with empirically based modes of drought variability that represent distinct regions of SST forcing.

Principal component analysis (PCA) can be used to decompose a large, multivariate dataset into fewer variables that maximize the fraction of variability contained in the original data (Wilks 2006):

$$\mathbf{u} = [E]^T \mathbf{x}'$$

where $\mathbf{x}'$ are the anomalies (mean of zero) of the original data, $[E]$ is the matrix of empirical orthogonal function (EOF) loadings, and $\mathbf{u}$ is the matrix of principal component (PC) scores. When PCA is applied to spatially explicit data (as is common in climate analyses), the spatial expression of these variables is represented by the EOF loadings and associated time variability is contained in the PC score time series. Typically, the number of meaningful or interpretable PC scores is determined using one or more of a variety of selection criteria (Wilks 2006). The original data anomalies ($\mathbf{x}'$) can be recovered using the synthesis formula:

$$\mathbf{x}' = [E]\mathbf{u}$$

When the full set of variables (all PC/EOF pairs) is used in the synthesis formula, the full original dataset is reconstructed. If fewer than the full set of variables is used, the reconstruction is approximate and the reconstructed variance is less than that of the original data. Alternatively, the original data can be partially reconstructed from a subset of variables, and the remaining variables and residual component can be stochastically modeled. The subsequent reconstruction is a new dataset with a forced (deterministic) and unforced (stochastic) component. This latter approach serves as the foundation of our statistical model, allowing us to simulate PDSI with various levels of forcing (prescribed PC scores) and noise (modeled PC scores and residual variance) to determine to what extent we can reproduce drought and pluvial events given various assumptions about the underlying forcing.

Our conceptual PDSI model is summarized in Fig. 1. Using the scree test selection criteria (Cattell 1966), we identify five leading ('meaningful') modes of variability. The first three PCs show significant correlations over the period 1900–2005 with SSTs in various ocean basins, and are used to represent the potential SST forced variability in the PDSI data: PC 1 (tropical Pacific), PC 2 (North Pacific), and PC 3 (Atlantic). Results from the PCA are further discussed in Sect. 3.2.

For all our modeling scenarios, we either prescribe (i.e., use the original PC scores without modification) the three SST related PCs (1, 2, and 3), or model them by randomizing their phase information, preserving the spectral
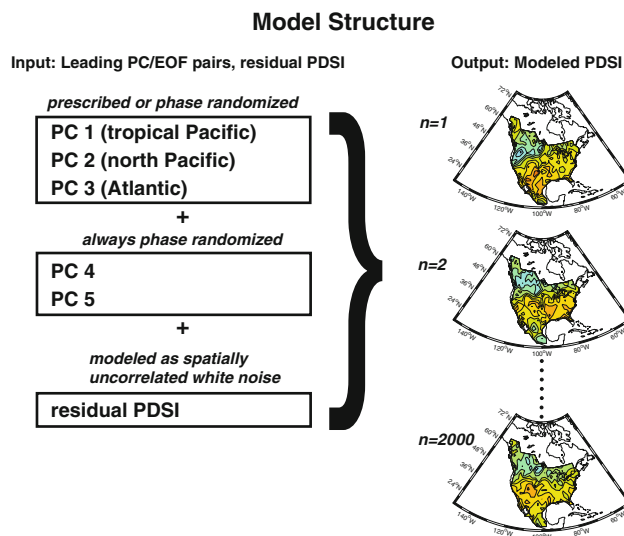
Model Structure

Fig. 1 Conceptual workflow for the PDSI modeling using prescribed and modeled PCs from the principal component analysis. PCs 1, 2, and 3 are used to represent the SST forced component of drought variability and are alternately prescribed or modeled via phase randomization. The remaining PCs and residual PDSI are always modeled

characteristics and serial autocorrelation of the underlying PC time series (Ebisuzaki 1997). PCs 4 and 5 are always phase randomized to preserve any underlying autocorrelation or frequency structure, and within our study are always considered to represent spatially coherent, but unforced, variability. For these five PC time series, the associated EOF loadings are used to constrain the spatial expression of the PC scores and are not modified. The residual variance, beyond the leading five PC scores, is modeled as spatially uncorrelated white noise, resampled using Gaussian distributions derived at each grid cell separately. Regardless of whether PCs 1, 2, or 3 are modeled or prescribed, all five leading PCs are represented and the total variance of the original PDSI dataset is preserved. For each experiment, we conduct a 2,000 member ensemble in order to sufficiently sample the noise space, although the results generally converge after about 100–200 resamples. To quantify the match between our modeled PDSI fields and the PDSI from the NADA, we use the anomaly correlation statistic (AC) (Wilks 2006), the product–moment coefficient of linear correlation between variables corresponding to the same location. The AC is commonly used in the verification of forecast fields and can be interpreted identically as a Pearson correlation.

For our null hypothesis (that the occurrence of any drought or pluvial event is random), we simulated a full random ensemble, modeling all five leading PC scores and residual PDSI. We then calculated the AC between the NADA PDSI and the modeled PDSI for each null case ensemble member and each drought and pluvial event. From the full histogram of ensemble ACs, we use the AC associated with the 95th percentile to establish our noise confidence limit. In other words, in the forced simulations (described next), any ensemble member with a higher AC than this null limit is considered to be skillfull at reproducing the drought/pluvial event of interest beyond simply random noise.

For the forced ensemble experiments, we alternately prescribe the first three PC scores (representing the forced component of PDSI variability). By examining the ensemble median and comparing these ensembles against the null case, we can determine to what extent different droughts and the pluvial are dominated by different forcing, and how well (and often) that forcing can reproduce the various drought and pluvial events above the random noise. Our forcing cases (described later) are PC 1 (tropical Pacific forcing), PCs 1–2 (tropical + North Pacific forcing), and PCs 1–2–3 (Pacific + Atlantic forcing).

3 Results

3.1 Observed twentieth century drought and pluvial events

There are four persistent multi-year hydroclimatic events over NA during the twentieth century (Fig. 2). These are the target events for our statistical model, which we use to assess how well our defined SST forcing can reproduce the PDSI patterns. Negative (drought) anomalies are in orange/red and positive (pluvial) anomalies are in blue. The early twentieth century pluvial (1905–1917) was likely the largest pluvial event in the last thousand years (Woodhouse et al. 2005) and served as the foundation for overly generous water use

Fig. 2 PDSI from the NADA v2a, for the four persistent twentieth century hydroclimatic events over North America. The four events were the 1905–1917 pluvial, the 1932–1939 'Dust Bowl' drought, the 1948–1957 drought, and the 1998–2002 drought
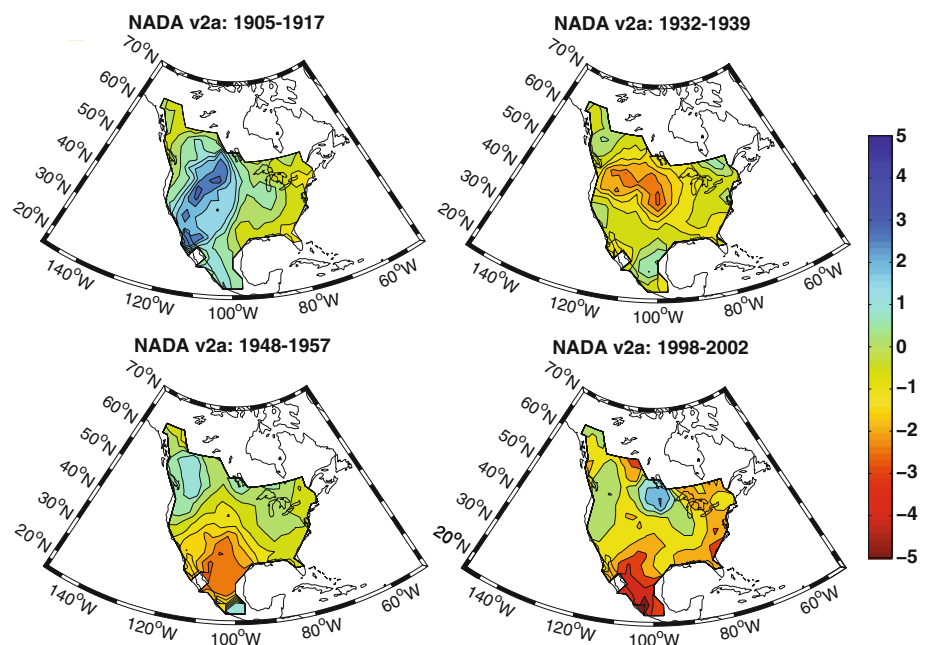
policies that drove much of the settlement and development in the western United States (Fye et al. 2003; MacDonnell et al. 1995). We were unable to find any published studies that explicitly discussed GCM simulations of the 1905–1917 pluvial event, although recent work suggests that pluvials should be much more predictable than droughts, based on reduced intra-ensemble spread in GCM experiments (Schubert et al. 2008). The first extended drought of the twentieth century occurred during the 1930s (1932–1939), named the 'Dust Bowl' in reference to the historically unprecedented dust storm activity that characterized this drought (Hansen and Libecap 2004). This drought is notable in that the spatial pattern diverged significantly from the canonical pattern associated with tropical Pacific SST forcing (Cook et al. 2009). Model simulations forced by observed SSTs alone incompletely simulate the Dust Bowl drought, with precipitation anomalies incorrectly centered in the southwest and southern Great Plains and without the expansive continental wide warming during the time period (Schubert et al. 2004; Seager et al. 2008b). Recent research has suggested that the dust storm activity arising from widespread land degradation at the time may have played an important role in forcing the unique drought pattern (Cook et al. 2008, 2009), possibly in concert with forcing from warm Atlantic SSTs and unforced atmospheric variability (Brönnimann et al. 2009; Hoerling et al. 2009). The latter two droughts (1948–1957 and 1998–2002) have spatial patterns fairly typical of La Niña forced droughts, centered in the southwest, Mexico, and southern Great Plains. Model experiments using GCMs are typically able to reproduce precipitation anomalies during these two droughts with high fidelity when forced with observed SSTs from the two time periods (Seager et al. 2005b). The 1998–2002 drought actually extended to 2004, but one study (Seager 2007) concluded that these latter two years of drought were unrelated to SST forcing and were likely a result of unforced atmospheric variability. Because we are most interested in the SST forced component, we focus on the earlier (1998–2002) part of this drought, when the potential for predictability is strongest.

3.2 PCA results

The eigenvalue trace (standardized to fractional variance) for the first 15 modes from our PCA is shown in Fig. 3. For our scree test, we fit a two-phase linear regression to the full eigenvalue trace. Our regression model identified eigenvalue five as the optimal breakpoint; we therefore retained the first five modes as our meaningful variables. These first five leading modes account for a cumulative 62% of the underlying variance in the original PDSI dataset. The associated EOF loadings are shown in Fig. 4. Negative values are in red and orange, indicating drought when the
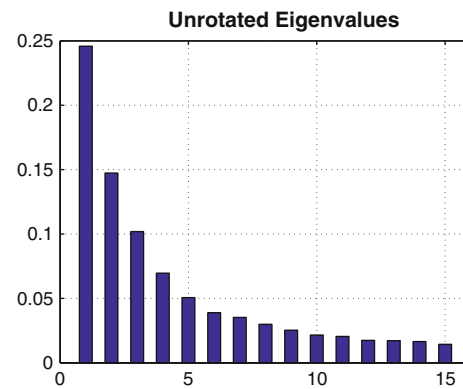

Fig. 3 First fifteen eigenvalues from the principal component analysis, standardized to fractional variance. The scree test selects the first five, accounting for a cumulative 61.53% of the variance in the drought atlas

accompanying PC scores are positive. EOF loadings 1 and 2 both have major centers of action in the southwest and southern Great Plains but, while EOF 1 shows continuous negative loadings throughout the west, EOF 2 is characterized by a dipole pattern with out of phase loadings in the Pacific northwest and the southwest. EOF 3 shows a strong east/west dipole with centers of action along the west coast and the Mississippi Valley. The spatial expression of these EOFs is remarkably similar to the first three rotated EOFs from a previous analysis of the NADA spanning the years 1000 C.E.–2003 C.E. (Herweijer et al. 2007). This helps support the idea that these patterns represent robust spatial expressions of PDSI variability over NA, rather than simply artifacts arising from the orthogonality constraints of the PCA.

Correlations between the PCs and SSTs from the Hadley Centre (Rayner et al. 2003) show significant teleconnections with variability in the Pacific and Atlantic ocean basins (Fig. 5). PC 1 is inversely correlated with SSTs in the tropical Pacific, indicating drought conditions in the west and southwest when the eastern tropical Pacific is cooler than normal (i.e., 'La Niña' or 'La Niña-like' conditions). PC 2 also shows some significant correlations in ENSO regions, but also strongly correlates with SSTs in the central North Pacific and in a band along the northwest coast of North America. This SST correlation pattern is indicative of a negative phase Pacific Decadal Oscillation (PDO), typically associated with drought in the southwest and wet conditions in the northwest (McCabe et al. 2004, 2008). The underlying dynamics and the degree to which the PDO and Pacific decadal variability is independent from ENSO is poorly understood (McCabe et al. 2008), and it may be that the PCA artificially separates the Pacific sourced variability into extratropical and tropical modes. The first two PCs also show some significant correlations with SSTs in the North Atlantic, although this may simply
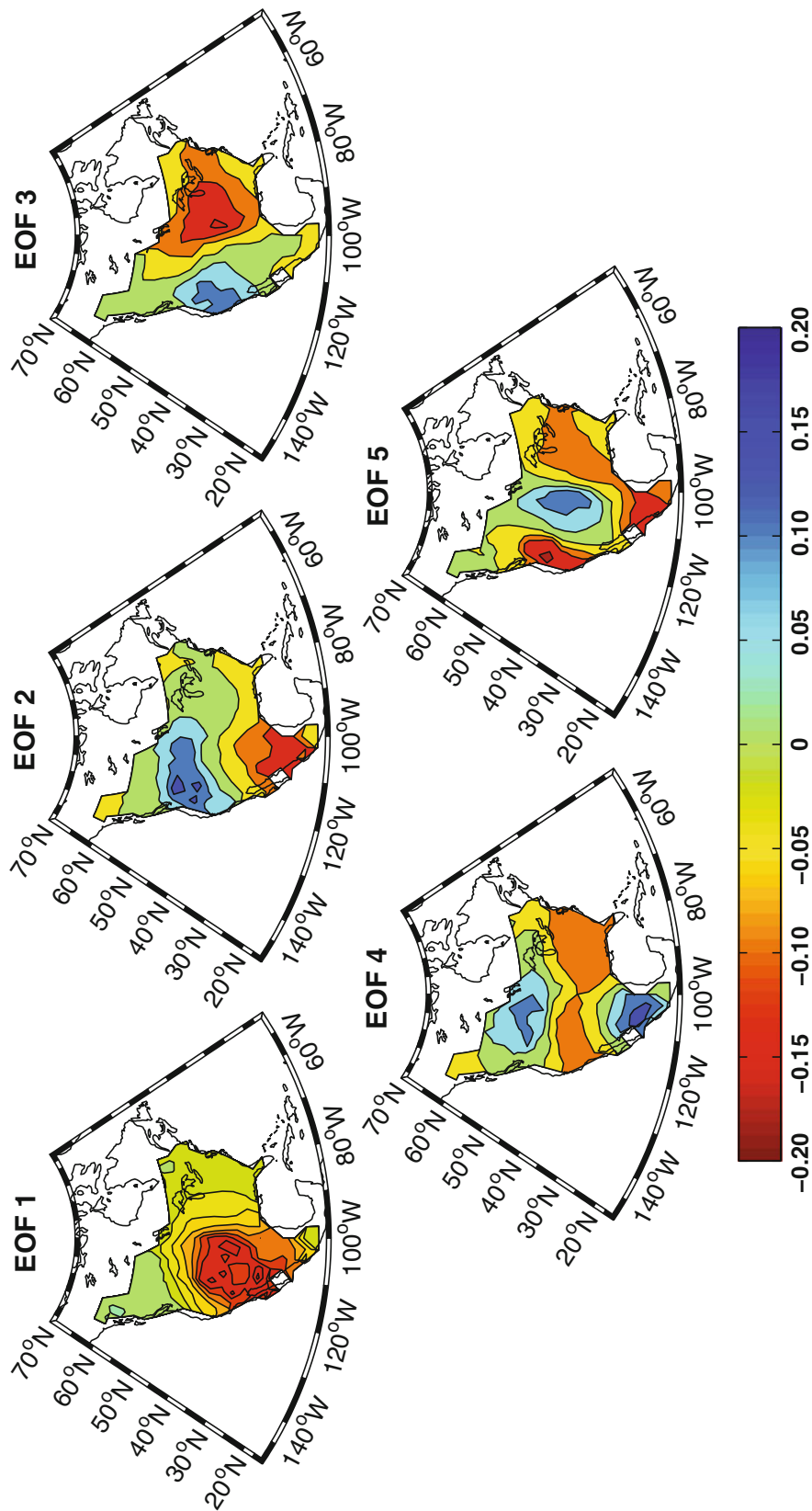
Fig. 4 Empirical orthogonal function (EOF) loadings for the first five modes. Note that negative values are in *red*, indicating drought when the accompanying principal component (PC) score time series are positive
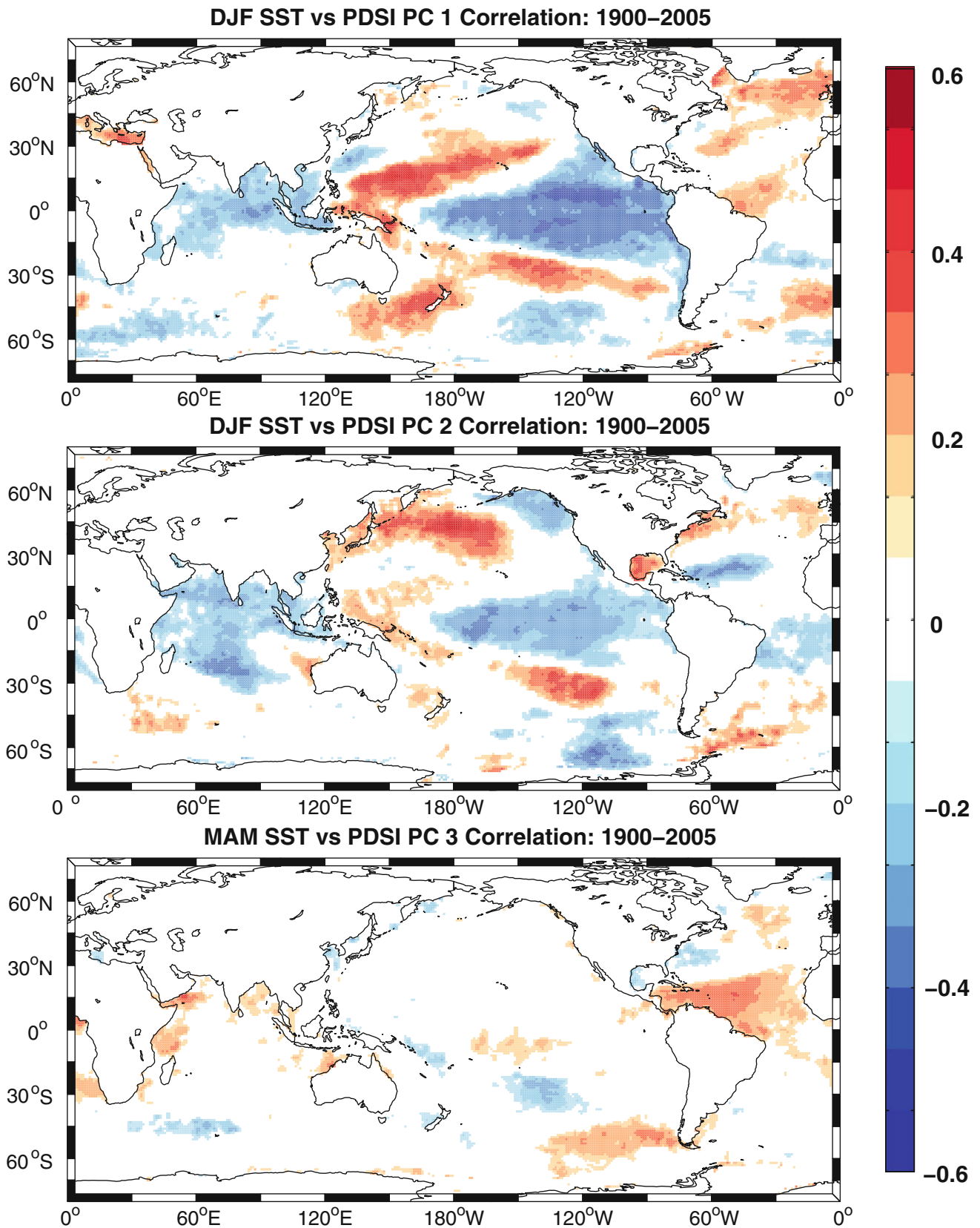
Fig. 5 Correlations (Spearman rank) between first three detrended principal component (PC) score time series and detrended sea surface temperatures from the Had ISST version 1 data set, 1900–2005. Insignificant correlations ($p > 0.10$) have been masked out
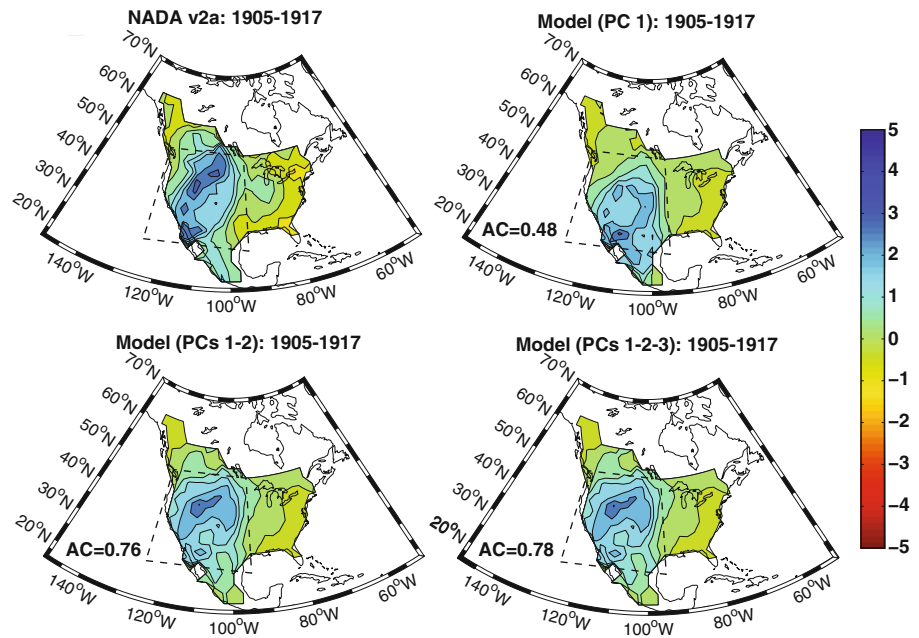
Fig. 6 PDSI from the NADA (*upper left*), and modeled with forcing from PC 1 (*upper right*), PCs 1–2 (*lower left*) and PCs 1–2–3 (*lower right*) for the 1905–1917 pluvial
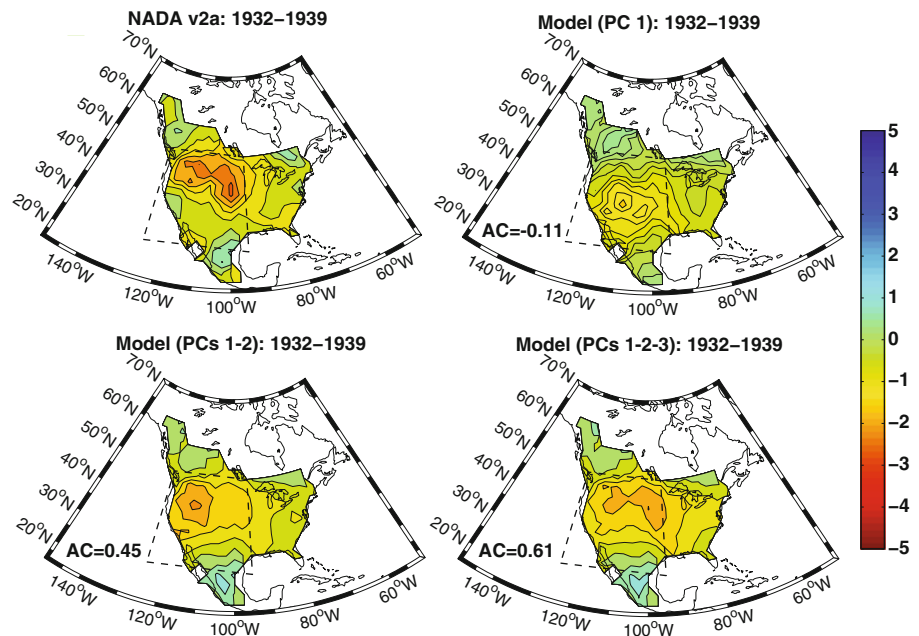

Fig. 7 PDSI from the NADA (*upper left*), and modeled with forcing from PC 1 (*upper right*), PCs 1–2 (*lower left*) and PCs 1–2–3 (*lower right*) for the 1932–1939 Dust Bowl drought

reflect the influence of the tropical Pacific on Atlantic SSTs. EOF loading patterns for PCs 1 and 2 are consistent with the typical spatial expression of ENSO and Pacific SST mediated drought teleconnections. In contrast, PC 3 correlates primarily with SSTs in the North Atlantic, with warm SSTs associated with drought in the central Great Plains and Mississippi valley. This SST correlation pattern and the associated EOF loadings are consistent with previous studies investigating the Atlantic influence on summer hydroclimate (e.g., Kushnir et al. 2010; Mo et al. 2009). PC 3 shows little correlation in the Pacific region, suggesting it is related to Atlantic SST variability that is

separate from Pacific forcing. The fact that this distinct Atlantic forcing is concentrated in PC 3 also makes sense in light of recent evidence suggesting that the direct influence of the Atlantic on drought is relatively weak, but may be most important for significantly amplifying forcing from the Pacific (e.g., Mo et al. 2009).

These sources of drought forcing (tropical Pacific, North Pacific, and Atlantic) have been identified in previous analyses (McCabe et al. 2004, 2008), providing further support that our selected PC/EOF pairs are dynamically meaningful in the context of drought variability over NA. The remaining two PCs (4 and 5) show few teleconnections

Fig. 8 PDSI from the NADA (*upper left*), and modeled with forcing from PC 1 (*upper right*), PCs 1–2 (*lower left*) and PCs 1–2–3 (*lower right*) for the 1948–1957 drought
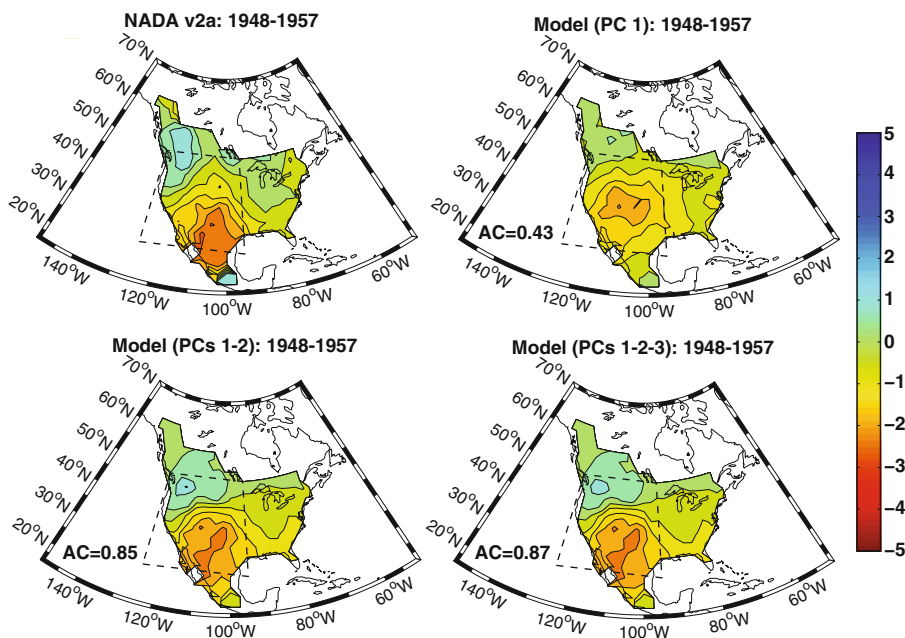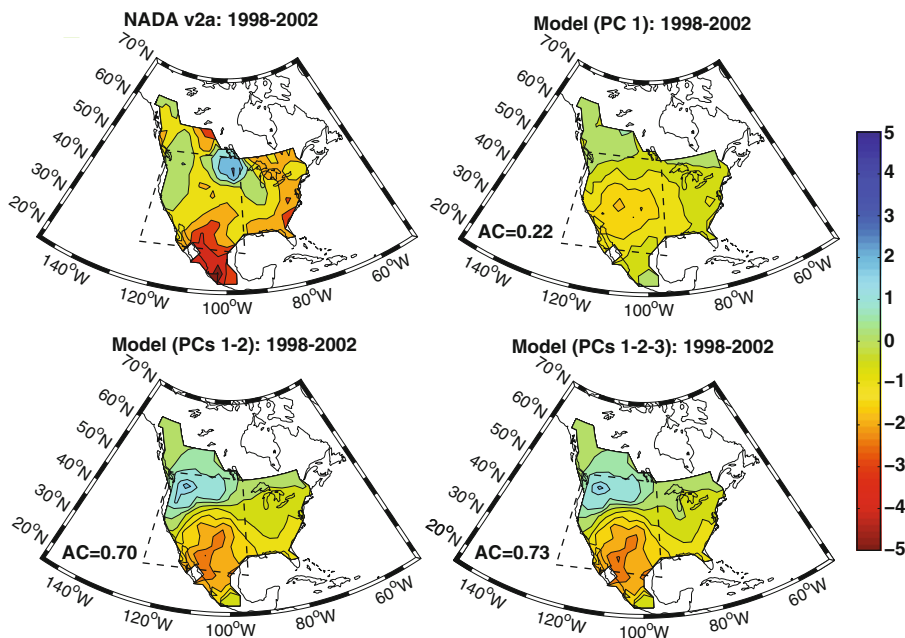

Fig. 9 PDSI from the NADA (*upper left*), and modeled with forcing from PC 1 (*upper right*), PCs 1–2 (*lower left*) and PCs 1-2-3 (*lower right*) for the 1998–2002 drought

to SSTs in any ocean basin (not shown), and their associated EOF loadings are difficult to interpret within the context of drought dynamics over NA. For our modeling exercise, we use these leading five PC/EOF pairs to represent the spatially coherent variability in the PDSI data, with the potential SST forced component of drought variability contained within the first three.

3.3 Modeling: ensemble median patterns

Figures 6, 7, 8 and 9 show the median PDSI patterns for each drought and pluvial event from our forced ensemble simulations: prescribed PC 1 (tropical Pacific forcing), PCs 1–2 (tropical + North Pacific forcing) and PCs 1–2–3 (Pacific + Atlantic forcing). The AC statistics comparing the ensemble median PDSI pattern and the PDSI pattern from the NADA are in the lower left corners of the model panels, calculated over the region bounded by the dashed lines ($125°W - 95°W$, $25°N - 50°N$). As mentioned previously, the separation of PCs 1 and 2 may be somewhat artificial because of the orthogonality constraints imposed by the PCA. We therefore focus primarily on comparing the tropical+North Pacific case and the Pacific + Atlantic case.
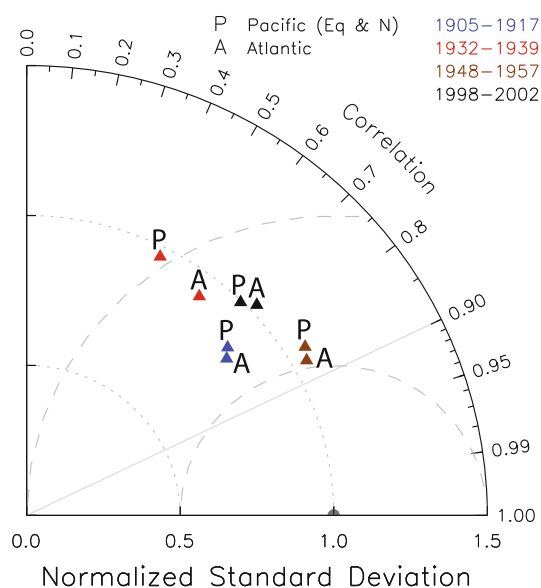
Fig. 10 Taylor diagram summarizing ensemble median results from all simulations for all events

For the pluvial event (1905–1917), the tropical + North Pacific forcing case reproduces the spatial pattern and intensity of the PDSI pattern from the NADA (Fig. 6). Almost the entire western region is wetter than normal in both the simulated and observed pattern, with the major wet anomaly extending along an axis from the southwest and into the northern Great Plains. The addition of Atlantic forcing in the Pacific + Atlantic case does little to change the actual median pluvial pattern, suggesting strongly that the Atlantic played little role in forcing the early twentieth century pluvial. For the 1932–1939 Dust Bowl drought, Pacific forcing does a poor job of reproducing the drought pattern (Fig. 7). The AC is actually negative in the tropical Pacific only case; this improves in the tropical +North Pacific case, but the drought is centered too far west, incorrectly spreads into the southwest, and misses the intense drying over the northern Plains. The inclusion of Atlantic (lower right panel) forcing markedly improves the general drought pattern, with the main drought now extending from the northwest across to the northern and central plains. However, the drought itself is still too diffuse in the west, too dry in the southwest and northern Mexico, and not dry enough in the northern Plains. The median AC for all forcing cases during the Dust Bowl is also lower when compared to other drought and pluvial events, implying a significant role for non-SST related processes.

The 1948–1957 (Fig. 8) and 1998–2002 (Fig. 9) drought simulations showcase the importance of Pacific forcing for persistent droughts over NA. For 1948–1957, the model does a remarkable job reproducing both the intensity and the spatial extent (Fig. 8) of the drought in the tropical + North

Pacific forcing case. The pattern with Atlantic forcing looks almost identical to the pattern for Pacific only forcing, suggesting that Atlantic SSTs played little (if any) role in forcing this drought. The modeled 1998–2002 drought pattern (Fig. 9) is quite similar to the observed and modeled 1948–1957 drought, also with primary forcing coming from the Pacific. Notably in this latest drought, the model cannot, in the ensemble median at least, reproduce the intensity of the drought in the southwest and Mexico, and also misses the pluvial conditions in the northern Plains. This suggests that a significant component of this most recent drought derives from non SST forced processes, although not to the same extent as the 1932–1939 Dust Bowl.

Results from all of our ensemble simulations can be compared and summarized by plotting the ensemble median AC statistics on a Taylor diagram (Fig. 10). By far, the Dust Bowl drought (red) is the biggest outlier, and the only event for which Atlantic forcing appears to significantly improve the ensemble median drought pattern (but see the next section regarding the 1998–2002 drought). Pacific forcing alone appears capable of reliably simulating the other events. Even with the addition of Atlantic forcing, however, the fidelity of the Dust Bowl simulation is weaker than any other event.

3.4 Modeling: ensemble statistics

In addition to the ensemble median, we can also examine the full ensemble distribution of simulated PDSI. The histograms of AC statistics between all ensemble members and the NADA are shown for the tropical + North Pacific forcing and Pacific + Atlantic forcing (Fig. 11). Vertical bars indicate the 95% confidence limit, calculated from the null hypothesis random ensemble described previously, and the percentage of ensemble members passing this threshold are indicated in the upper left corners.

Increasing the level of forcing from PCs 1–2 (tropical + North Pacific forcing, red lines) to PCs 1–2–3 (Pacific + Atlantic forcing, blue lines) results in both increased ensemble median anomaly correlations and also the percentage of individual ensemble members that exceed the 95% significance threshold. Tropical + North Pacific forcing alone results in over 90% of the individual ensemble members passing the noise threshold for the 1905–1917 pluvial (96.6%) and the 1948–1957 drought (94.5%). The addition of the Atlantic forcing increases the proportion of passing ensemble members only marginally (99.0 and 96.3% for this pluvial and drought, respectively). For the 1998–2002 drought, forcing from the Pacific dominates (66.7% passing in full Pacific forcing case), but the Atlantic plays an important secondary role (82.1% passing in the Pacific + Atlantic case). This improvement is not reflected in the ensemble median, which only

Fig. 11 Frequency histogram of anomaly correlations between observed drought and pluvial events and 2,000-member ensemble simulations from the tropical +North Pacific case (PCs 1–2, *red lines*) and Pacific + Atlantic case (PCs 1–2–3, *blue lines*). Ordinate is relative frequency and abscissa is the anomaly correlation; integrated area under the curves equals unity. *Vertical bars* indicate the 95% null confidence limit, calculated from the full random ensemble. Curves have been smoothed using a three-point running mean in order to facilitate easier comparisons
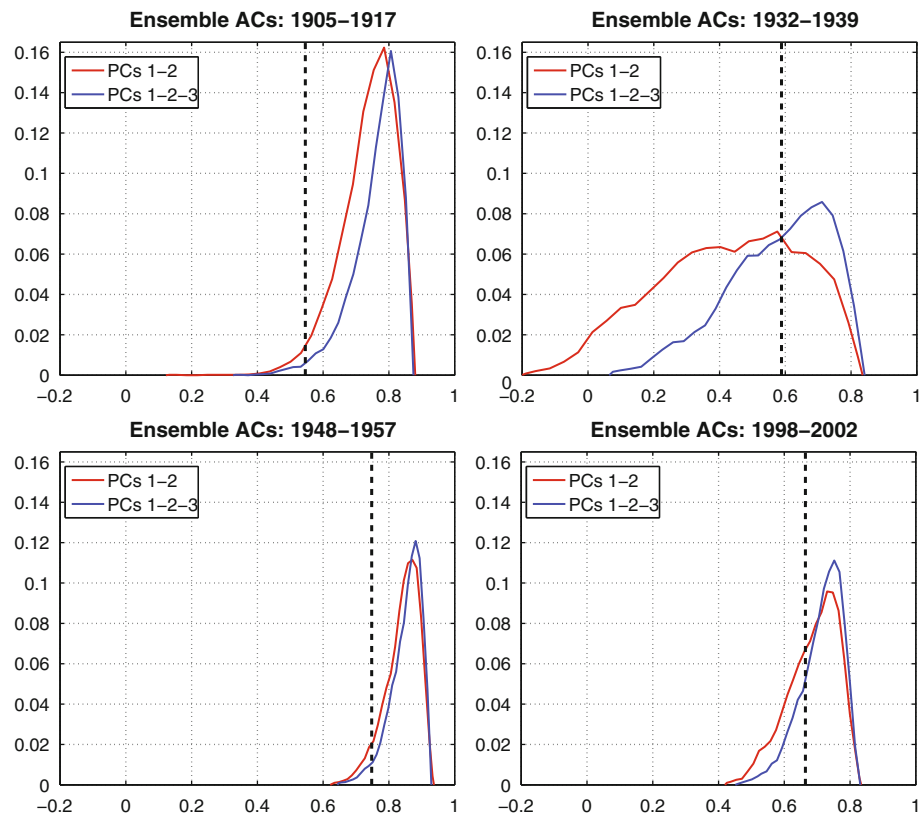
increases from 0.70 to 0.73 for the full Pacific and Pacific + Atlantic cases, respectively.

Of all four events, Atlantic forcing appears to be most important during the 1932–1939 Dust Bowl drought. In the tropical + North Pacific simulation, the ensemble median AC = 0.45, and less than one third (28.3%) of ensemble members pass the 95% threshold. With the addition of Atlantic forcing, the median AC increases to 0.61 and the number of ensemble members passing the noise limit nearly doubles (53.8%). Despite these improvements, however, the fidelity of the Dust Bowl simulations is much weaker compared to the other events. This supports several recent studies (Brönnimann et al. 2009; Cook et al. 2008, 2009, 2010; Hoerling et al. 2009), highlighting the importance of non-SST related factors during the Dust Bowl.

3.5 Model uncertainties

We tested the sensitivity of our modeling framework to our PCA methodology. Conducting the PCA on the correlation matrix gave similar results, although with a lower fraction of the variance in the leading modes, diminished EOF loadings in the west, and higher EOF loadings in the eastern US. This arises because, in the original PDSI dataset, the variance is higher in the west and the southwest compared to the eastern portion of the domain. A

covariance matrix based PCA therefore weights the results towards the west and southwest; using the correlation matrix normalizes the variance, and more evenly distributes the loadings. Because we were most interested in drought in the western US, we chose to use results from the covariance matrix based PCA. We also experimented with varimax rotation of our EOFs, but found the results highly sensitive to our truncation criteria. Results from our unrotated analysis show good consistency with other North American drought studies that have found similar EOF loading patterns (Herweijer et al. 2007; McCabe et al. 2004, 2008). Given this similarity to other studies, and because we could find no compelling physically based reason to rotate, we decided to use results from the unrotated analysis.

4 Discussion and conclusions

Understanding the predictability of drought and pluvial events over NA depends on being able to separately quantify the forced, deterministic component of drought variability and the unforced, stochastic variance arising from background noise in the climate system. Many studies addressing these issues use GCM experiments, where the ensemble mean or median is used to indicate the common signal or forced component (e.g., Schubert et al. 2009) and

the unforced, internal variability is represented by the intra-ensemble variance (e.g., Schubert et al. 2008). This is a powerful approach but, because of the computational expense, many GCM experiments are limited in the size of ensembles that can be conducted. Seager et al. (2008b), for example, only used 16 member ensembles for their simulations of the nineteenth and twentieth centuries, while Schubert et al. (2004) used 14 in their simulation of the twentieth century. While the limited ensemble sizes are understandable given the time and expense required to run a GCM, it is difficult to determine whether the full noise space has been adequately sampled. There is also the implicit assumption that the forcings and teleconnections of interest are well represented in the model itself. The study presented here is able to avoid some of these issues, enabling large, computationally efficient ensemble simulations using a statistical modeling approach. Our major results are summarized:

- Forcing from the tropical + North Pacific is sufficient to reproduce both the early twentieth century pluvial (1905–1917) and the 1950s drought (1948–1957). The median drought pattern from the tropical +North Pacific forcing ensemble captures both the spatial pattern and intensity of these events. The addition of Atlantic forcing does little to improve either the ensemble median drought/pluvial patterns or the proportion of ensemble members passing the 95% noise threshold.
- For the 1998–2002 drought, the ensemble median from the tropical + North Pacific case generally captures the spatial distribution of the drought. The addition of Atlantic forcing has little impact on the median drought pattern, but substantially increases the proportion of ensemble members that exceed the noise threshold (66.7–82.1% passing). Even with Atlantic forcing, however, the median simulated drought pattern does not reach the intensity of the observed drought, implying other factors may be responsible.
- In the Dust Bowl (1932–1939), tropical + North Pacific forcing does a poor job of reproducing the drought, with a median drought pattern centered too far west and less than one third of all ensemble members passing the noise threshold. Addition of Atlantic forcing improves the median drought pattern, and nearly doubles the percent of ensemble members passing the 95% threshold (from 28.3 to 53.7%). Even with the Atlantic forcing, however, the fidelity of our simulation of the Dust Bowl is much weaker than for the other droughts, emphasizing the important role of non SST forcing or internal noise for shaping anomalies during the Dust Bowl drought.

Results from our statistical model largely confirm many of the insights gained from GCM experiments, and provide additional support for much of our understanding of drought dynamics using an independent methodology. For persistent hydroclimatic variability over NA, forcing from the Pacific dominates, with Atlantic forcing sometimes (but not always) playing an important secondary role. Even when the Atlantic is an important factor, however, variability and forcing coming from internal noise or non-SST sources can be significant (e.g., the 1932–1939 and 1998–2002 droughts). This implies there may be a greater limit on the predictability of persistent NA drought and pluvial events than GCM simulations would indicate. For example, the ensemble median from the Pacific + Atlantic forcing ensemble under predicts the severity of the 1998–2002 drought, even while reproducing the general spatial pattern in the ensemble median. In fact, the intensity and spatial pattern of the simulated 1998–2002 drought looks remarkably similar to the observed and modeled pattern for the 1948–1957 drought. The 1948–1957 drought, then, may be indicative of the upper limits of hydroclimatic predictability over NA, at least when SSTs are the only forcing considered. Also of interest, although beyond the scope of this paper, is the fact that our model avoids some of the 'false positive' droughts simulated by GCM studies. For example, in the GCM ensemble from Schubert et al. (2004), the ensemble mean (and most ensemble members) shows an intense drought over the Plains during the 1970s (Fig. 1, lower panel), at odds with the observational record, which showed near normal conditions. Our model produces drought conditions in the west, but with anomalies consistent with the NADA. This suggests there still may be some value added information to be gained through statistical modeling as a point of comparison with GCM simulations.

Of all events, Atlantic forcing appears to have been most important during the Dust Bowl drought, resulting in the biggest improvement in both the ensemble median and ensemble spread compared to the tropical +North Pacific forcing only case. Even then, only about half of the ensemble members from the Pacific + Atlantic forcing case exceed the 95% noise threshold; this (as well as the ensemble median AC) is much lower than comparable simulations of the other droughts and pluvial. This suggests a nearly equally important role for non SST sources of variability during the Dust Bowl. Whether this variability comes from internal noise in the climate system or is forced from other sources (e.g., the land surface) is currently widely debated (Brönnimann et al. 2009; Hoerling et al. 2009). Recent GCM based studies, however, support the idea that land surface and dust aerosols feedbacks during this period were important (Cook et al. 2008, 2009, 2010). Land degradation and subsequent dust storm activity during the Dust Bowl drought was widespread, intensive, and unprecedented over at least the instrumental period (Hansen and Libecap 2004). In the dust feedback

hypothesis, increased dust aerosol loadings in the atmosphere from these dust storms led to a 'Charney' type feedback, with increased shortwave reflection, reduced net radiation at the surface and top of the atmosphere, and compensatory subsidence that inhibited precipitation. To adequately address this within the current modeling framework, however, we would need a way to statistically diagnose this variability coming from the land surface and dust aerosols. This is beyond the capabilities of the relatively simple model we have developed here.

Finally, results from the 1998–2002 ensemble suggest that, in some cases, the ensemble mean or median may not be the best metric for determining which factors are important during a given drought event. For this drought, the addition of Atlantic forcing did little to improve the ensemble median drought pattern, but did substantially increase the chance of any give ensemble member exceeding the noise threshold. This, in particular, highlights the potential problem with limited ensemble sizes in GCM experiments, which may not have a large enough sample size to resolve this shift in the distribution of the ensemble.

References

Andreadis K, Clark E, Wood A, Hamlet A, Lettenmaier D (2005) Twentieth-century drought in the conterminous United States. J Hydrometeorol 6(6):985–1001

Brönnimann S, Stickler A, Griesser T, Ewen T, Grant A, Fischer A, Schraner M, Peter T, Rozanov E, Ross T (2009) Exceptional atmospheric circulation during the "Dust Bowl". Geophys Res Lett 36(8):L08,802

Cattell R (1966) The scree test for the number of factors 1. Multivar Behav Res 1(2):245–276

Cook B, Miller R, Seager R (2008) Dust and sea surface temperature forcing of the 1930s "Dust Bowl" drought. Geophys Res Lett 35(8):L08,710

Cook B, Miller R, Seager R (2009) Amplification of the North American "Dust Bowl" drought through human-induced land degradation. Proc Natl Acad Sci 106(13):4997

Cook B, Seager R, Miller R (2010) Atmospheric circulation anomalies during two persistent North american droughts: 1932–1939 and 1948–1957. Clim Dyn (in press)

Cook E, Meko D, Stockton C (1997) A new assessment of possible solar and lunar forcing of the bidecadal drought rhythm in the western United States. J Clim 10(6):1343–1356

Cook E, Meko D, Stahle D, Cleaveland M (1999) Drought reconstructions for the continental United States*. J Clim 12(4):1145–1162

Cook E, Woodhouse C, Eakin C, Meko D, Stahle D (2004) Long-term aridity changes in the Western United States. Science 306(5698):1015–1018

Cook E, Seager R, Cane M, Stahle D (2007) North American drought: reconstructions, causes, and consequences. Earth Sci Rev 81(1–2):93–134

Ebisuzaki W (1997) A method to estimate the statistical significance of a correlation when the data are serially correlated. J Clim 10(9):2147–2153

Enfield D, Mestas-Nunez A, Trimble P (2001) The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the Continental US. Geophys Res Lett 28:2077–2080

Feng S, Oglesby R, Rowe C, Loope D, Hu Q (2008) Atlantic and Pacific SST influences on Medieval drought in North America simulated by the Community Atmospheric Model. J Geophys Res 113(D11101)

Fye F, Stahle D, Cook E (2003) Paleoclimatic analogs to twentieth-century moisture regimes across the United States. Bull Am Meteorol Soc 84(7):901–909

Hansen Z, Libecap G (2004) Small farms, externalities, and the Dust Bowl of the 1930s. J Political Econ 112(3):665–694

Herweijer C, Seager R, Cook E (2006) North American droughts of the mid to late nineteenth century: a history, simulation and implication for Mediaeval drought. Holocene 16(2):159

Herweijer C, Seager R, Cook E, Emile-Geay J (2007) North American droughts of the last millennium from a gridded network of tree-ring data. J Clim 20(7):1353–1376

Hoerling M, Kumar A (2003) The perfect ocean for drought. Science 299(5607):691

Hoerling M, Quan X, Eischeid J (2009) Distinct causes for two principal US droughts of the 20th century. Geophys Res Lett 36(19):L19,708

Kushnir Y, Seager R, Ting M, Naik N, Nakamura J (2010) Mechanisms of Tropical Atlantic SST Influence on North American Hydroclimate Variability. J Clim (in press)

MacDonnell L, Getches D, Hugenberg Jr W (1995) The law of the Colorado River: coping with severe sustained drought. Water Resour Bull 31(5):825–836

McCabe G, Palecki M, Betancourt J (2004) Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. Proc Natl Acad Sci 101(12):4136–4141

McCabe G, Betancourt J, Gray S, Palecki M, Hidalgo H (2008) Associations of multi-decadal sea-surface temperature variability with US drought. Quat Int 188(1):31–40

Mo K, Schemm J (2008) Droughts and Persistent Wet Spells over the United States and Mexico. J Clim 21(5):980–994

Mo K, Schemm J, Yoo S (2009) Influence of ENSO and the Atlantic Multidecadal Oscillation on drought over the United States. J Clim 22:5962–5982

Palmer W (1965) Meteorological drought. Res Pap 45:1–58

Rayner N, Parker D, Horton E, Folland C, Alexander L, Rowell D, Kent E, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys Res 108(D14):4407

Schubert S, Suarez M, Pegion P, Koster R, Bacmeister J (2004) On the cause of the 1930s Dust Bowl. Science 303(5665):1855–1859

Schubert S, Suarez M, Pegion P, Koster R, Bacmeister J (2008) Potential predictability of long-term drought and pluvial conditions in the US Great Plains. J Clim 21(4):802–816

Schubert S, Gutzler D, Wang H, Dai A, Delworth T, Deser C, Findell K, Fu R, Higgins W, Hoerling M et al. (2009) A US CLIVAR project to assess and compare the responses of global climate models to drought-related SST forcing patterns: overview and results. J Clim 22(19):5251–5272

Seager R (2007) The turn of the century North American drought: global context, dynamics, and past analogs. J Clim 20(22):5527–5552

Seager R, Harnik N, Kushnir Y, Robinson W, Miller J (2003) Mechanisms of hemispherically symmetric climate variability. J Clim 16(18):2960–2978

Seager R, Harnik N, Robinson W, Kushnir Y, Ting M, Huang H, Velez J (2005a) Mechanisms of ENSO-forcing of hemispherically symmetric precipitation variability. Q J R Meteorol Soc 131(608):1501–1527

Seager R, Kushnir Y, Herweijer C, Naik N, Velez J (2005b) Modeling of tropical forcing of persistent droughts and pluvials over western North America: 1856–2000. J Clim 18(19):4065–4088

Seager R, Ting M, Held I, Kushnir Y, Lu J, Vecchi G, Huang H, Harnik N, Leetmaa A, Lau N et al. (2007) Model projections of an imminent transition to a more arid climate in southwestern North America. Science 316(5828):1181

Seager R, Burgman R, Kushnir Y, Clement A, Cook E, Naik N, Miller J (2008a) Tropical Pacific forcing of North American medieval megadroughts: testing the concept with an atmosphere model forced by coral-reconstructed SSTs. J Clim 21(23):6175–6190

Seager R, Kushnir Y, Ting M, Cane M, Naik N, Miller J (2008b) Would advance knowledge of 1930s SSTs have allowed prediction of the Dust Bowl drought? J Clim 21(13):3261–3281

Touchan R, Anchukaitis K, Meko D, Sabir M, Attalah S, Aloui A (2010) Spatiotemporal drought variability in northwestern Africa over the last nine centuries. Clim Dyn 1–16

Trenberth K, Guillemot C (1996) Physical processes involved in the 1988 drought and 1993 floods in North America. J Clim 9(6):1288–1298

Wang C, Enfield D, Lee S, Landsea C (2006) Influences of the Atlantic warm pool on Western Hemisphere summer rainfall and Atlantic hurricanes. J Clim 19(12):3011–3028

Wilks D (2006) Statistical methods in the atmospheric sciences. In: International geophysics series, vol 91. Academic Press, New York

Woodhouse C, Kunkel K, Easterling D, Cook E (2005) The twentieth-century pluvial in the western United States. Geophys Res Lett 32(7):L07,701