The Potential Predictability of Precipitation Occurrence, Intensity, and Seasonal Totals over the Continental United States*

DANIEL J. SHORT GIANOTTI, BRUCE T. ANDERSON, AND GUIDO D. SALVUCCI

Department of Earth and Environment, Boston University, Boston, Massachusetts

(Manuscript received 8 November 2013, in final form 29 April 2014)

ABSTRACT

Using weather station data, the parameters of a stationary stochastic weather model (SSWM) for daily precipitation over the contiguous United States are estimated. By construct, the model exactly captures the variance component of seasonal precipitation characteristics (intensity, occurrence, and total amount) arising from high-frequency variance. By comparing the variance of the lower-frequency accumulations (on the order of months) between the SSWM and the original measurements, potential predictability (PP) is estimated. Decomposing the variability into contributions from occurrence and intensity allows one to establish two contributing sources of total PP. Aggregated occurrence is found to have higher PP than either intensity or the seasonal total precipitation, and occurrence and intensity are found to interfere destructively when convolved into seasonal totals. It is recommended that efforts aimed at forecasting seasonal precipitation or attributing climate variability to particular processes should analyze occurrence and intensity separately to maximize signal-to-noise ratios. Significant geographical and seasonal variations exist in all PP components.

1. Introduction

Precipitation data are very noisy as a result of many nonlinear processes and thresholds in both the physics of precipitation and our ability to measure it (Mandelbrot and Wallis 1968; Pruppacher and Klett 1997). At our common measurement time scales (e.g., daily), spatially based physical process models dominate our precipitation forecasting methodologies (Lynch 2008), whereas the chaotic behavior of the atmosphere beyond a few weeks renders these methods nearly useless for forecasts at seasonal scales or longer (Lorenz 1965). In time series analyses of precipitation point data, the perceived deterministic time scale switches; there is significant persistence in aggregated seasonal and annual totals (particularly when tied to slowly changing atmospheric indices; Katz and Parlange 1998), but daily data are generally considered random variables because of low

daily persistence (Katz and Parlange 1998) as well as the complex spatial behavior of precipitation (Olsson and Niemczynowicz 1996; Gong et al. 2003).

As we move from daily data to longer aggregated time scales, we see that precipitation data become less chaotic and less stationary (Mandelbrot and Wallis 1968)—that is, the probability distributions for precipitation random variables are not the same from one month, season, or year to the next. This is perhaps most easily envisioned as the superposition of multiple precipitation patterns at multiple time scales: diurnal evaporation and moisture convergence (Meisner and Arkin 1987), the annual cycle (Meisner and Arkin 1987), correlations with intraseasonal to multiannual climate oscillations [e.g. the Madden-Julian oscillation (MJO; Becker et al. 2011), El Niño–Southern Oscillation (ENSO; Ropelewski and Halpert 1986), Arctic Oscillation/North Atlantic Oscillation (AO/NAO; Hurrell 1995), and Pacific decadal oscillation (PDO; Mantua and Hare 2002)], and empirical trends over the course of our data record (Karl and Knight 1998). This nonstationarity is what fundamentally allows us to make successful predictions of precipitation despite the chaotic regime of atmospheric dynamics, as evidenced by the typically positive forecast skill scores on seasonal time scales (NOAA Climate Prediction Center 2013).

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JCLI-D-13-00695.s1>.

Corresponding author address: Daniel J. Short Gianotti, Boston University, 685 Commonwealth Ave., Boston, MA 02215.
E-mail: gianotti@bu.edu

We are interested in establishing the degree to which precipitation is potentially predictable at these climatic time scales [i.e. longer than the chaotic noise time scale on the order of days to weeks (Lorenz 1965; Leith 1978)]. Throughout this paper, we will use the terms high frequency and “weather noise” to refer to processes with a time scale on the order of days, and low frequency and “climate signal” to refer to processes—other than the annual cycle—with time scales longer than the weather noise. The characterization of high-frequency processes as noise rather than signal is simply a result of this study’s focus on climate-scale processes.

Potential predictability (PP) can be thought of as an upper bound on forecast skill or a partitioning of a time series into signal and noise. In the case of this study, the noise is due to short-persistence (on the order of days) processes in the Earth system, and the signal is due to a number of forcings, feedbacks, and slowly changing boundary conditions that lead to persistence beyond our noise time scale. This upper bound can provide us with useful information, including 1) the broad limits of our ability to confidently forecast longer-term climate behaviors, 2) indications of where and when our present forecasts are limited by an understanding of the system versus an absence of useful climate signals, and 3) a metric to test whether or not our current forecast models display signals and noise in the same manner as the real climate system.

The first point is particularly critical as unexpected droughts and floods damage infrastructure in U.S. towns, cities, and farmlands; and water resource reliability impacts human livelihoods through social, health, and economic means (NCADAC 2014). It is important to note that these impacts are as much an issue of the predictability of precipitation as its magnitude. For example, the ability of modern human infrastructure to provide for quite similar lives in the deserts of Arizona and the rainforests of Washington relies on a proper use of climatological data to predict future conditions. This prediction is obviously less effective if stationarity is improperly assumed, or if the data used to make predictions are insufficient for the time scale of infrastructure planning (Milly et al. 2008). Beyond the direct human impacts, increased precipitation variability has been found to affect ecosystem services—for example, through changes in plant community composition and ecosystem carbon cycling processes (Knapp et al. 2002)—and our ability to mitigate or adapt to these impacts are limited by our ability to predict them.

Since day-to-day precipitation is unpredictable on climatic time scales (i.e., beyond the weather noise time scale from days to weeks), we will think of it as a stochastic process with some amount of intrinsic variability.

From the forecasting perspective, we might equate precipitation variability with uncertainty and think that large variances must be related to low predictability. While it is certainly true that stochastic weather variability increases uncertainty, any climate signal—which we can think of as the variability of the weather noise distribution through time (e.g., Mandelbrot and Wallis 1968)—also increases precipitation variability. From a climate perspective, precipitation variance is a combination of (stochastic) weather noise and (potentially predictable) climate signal. If we think of separating our precipitation variance into a weather component and a climate component, the latter forms the basis for our climatic potential predictability, and large values of climate variability relative to weather variability suggest high PP. As an example, if the statistics of precipitation were stationary from year to year, there could be interannual variability caused by weather; however, there would be no climate-driven variability and thus no opportunity for prediction based on climate indicators (sea surface temperatures, atmospheric modes, land–atmosphere coupling). The degree of nonstationarity in the statistics of precipitation is thus one measure of potential predictability. In this paper, we will specifically assess how well (or poorly) stationary statistics can describe the interannual variability of precipitation.

Potentially predictable climate signals can manifest as deviations from stationary precipitation characteristics spanning from years to centuries. There has been considerable previous investigation into precipitation variability on these time scales, much of which focuses on partitioning natural variability and trend components. For instance, mean daily precipitation increased over the contiguous United States by roughly 10% over the twentieth century (Karl and Knight 1998). Precipitation occurrence in the contiguous United States has generally increased over the last century as well (Karl and Knight 1998; Groisman et al. 1999; Higgins and Kousky 2013; Pal et al. 2013), with some locations showing an additional 15 days per year of precipitation over the last half of the twentieth century (Higgins et al. 2007). At the same time, dry spell lengths have also increased, leading to more wet days in wet periods and longer droughts in dry periods (Groisman and Knight 2008; Pal et al. 2013), as well as possible shifts in the timing of the seasonal precipitation cycle (Pryor and Schoof 2008; Pal et al. 2013). Roughly half of the increase in the number of wet days has been attributed to the PDO and AO (Higgins et al. 2007); and modeling studies have linked up to 40% of the variability in annual precipitation in the U.S. Southeast, Southwest, and mid-Atlantic regions to sea surface temperatures (SSTs), and found significant amplifications of variance

in the central Great Plains due to land–atmosphere feedbacks (Koster et al. 2000).

Much of the increase in twentieth-century precipitation is caused by increases at the extreme high end of the distribution (Karl and Knight 1998). Models suggest that heavy precipitation should increase with trends in increased atmospheric temperatures, perhaps with an increase in dry days as well (Groisman et al. 2005; Trenberth et al. 2003), and that probable maximum precipitation should increase following the Clausius–Clapeyron relation (Kunkel et al. 2013; Sun et al. 2007). Using maximum daily precipitation, this increase in heavy precipitation has been found to hold globally, but U.S. precipitation has shown a less coherent relationship (Westra et al. 2013). Changes in heavy precipitation have been tied to increased convective thunderstorm activity (except in the Great Lakes and mid-Atlantic; Changnon 2001), extratropical cyclones (Great Lakes and Northeast; Kunkel et al. 2012), frontal systems (central Great Plains; Kunkel et al. 2012), and the influence of tropical cyclones (Southeast and the “Cornbelt” region; Knight and Davis 2009; Kunkel et al. 2012). Much of the increase in heavy precipitation in the United States appears to have occurred after 1970 (Groisman et al. 2005), although intensities of similar magnitudes were observed at the same frequency around the turn of the twentieth century, suggesting a background of significant natural climate variability (Kunkel et al. 2003). As part of our analysis in this paper, we will assess the degree to which these observed changes in both intensity and occurrence processes can be said to contribute to precipitation’s variability and potential predictability.

Given its importance, significant research has addressed the issue of isolating predictable characteristics of climate-scale precipitation variance, using both process models (Koster et al. 2004; Moron et al. 2006; Jia and DelSole 2011) and observational approaches (Madden et al. 1999; Anderson et al. 2009; Robertson et al. 2009; Gianotti et al. 2013). PP is generally established from signal-to-noise ratios (Ferguson et al. 2011), variance partitioning (Koster et al. 2004; Jia and DelSole 2011; Boer 2009), or through the divergence of states from initial model conditions (Collins 2002). Precipitation is found to have lower typical PP than air temperatures (Collins 2002; Boer 2009; Feng et al. 2011)—often less than 30% of the total variance is deemed potentially predictable (Koster et al. 2000; Boer 2009; Westra and Sharma 2010; Feng et al. 2011)—and the processes driving PP in a certain location tend to be dominated by either ocean processes or land feedbacks (Koster et al. 2000). Prior work has separated the role of precipitation occurrence and intensity, and occurrence has been found to demonstrate more potential predictability than intensity in the tropics

(Moron et al. 2007; Robertson et al. 2009) and Chile (Verbist et al. 2010), perhaps due to the stronger spatial coherence of occurrence processes (Moron et al. 2006).

In this study, we assess PP for U.S. seasonal precipitation totals using purely observation-based data as opposed to modeling studies, which may have artifacts resulting from model specification. Studies of the United States have found higher PP in the U.S. West (Yilmaz and DelSole 2010; Feng et al. 2011), with possible feedbacks in the central Great Plains (Koster et al. 2004) using gridded precipitation data. Using station data, Gianotti et al. (2013) found the seasonal PP for rain gauge precipitation occurrence to be somewhat higher than other studies found for total precipitation, 30% on average across the United States, with some locations higher than 70% in certain seasons.

The primary components of this study and this paper are 1) a strategy for creating stationary stochastic weather models (SSWMs) to model precipitation under the assumption of stationary weather statistics (no climate variability), 2) the decomposition of precipitation variability into intensity and occurrence components, 3) simulation of precipitation data using the SSWMs as a null hypothesis of no potentially predictable climate signal against which to compare the observed precipitation data, 4) calculation of potential predictability from the observed data, and 5) discussion of spatial features of the calculated potential predictability values. Section 2 presents an overview of the data and statistical methods employed to quantify the potential predictability of climate-induced variations in seasonal-mean precipitation amounts, occurrences, and intensities across the United States. The temporal and spatial characteristics of potentially predictable signals are then presented in sections 3 and 4. Discussion of the results and implications for understanding the role climatic factors versus short-persistence processes play in generating these events follow in section 5.

2. Methods

a. Stochastic modeling

To establish potential predictability, we compare the interannual variability of observed data with the variability of annually stationary models designed to represent the observed data as well as possible within a stationary framework. We will refer to these models as stationary stochastic weather models to emphasize their use in simulating artificial weather data with no interannual variability aside from that created by weather processes with persistence on the order of days.

Daily precipitation has been modeled in the past, often as a chain-based process with some assumed parametric

or nonparametric intensity distribution (Stern and Coe 1984; Katz 1977; Wilks 1999; Wang et al. 2006; Robertson et al. 2006; Schoof and Pryor 2008; Hanson and Vogel 2008). Although generally fit to observational data, the simplest models are prone to underrepresentation of observed interannual variability (Gregory et al. 1993; Katz and Parlange 1998). To the degree that this is due to misrepresentation of high-frequency variance, it can be corrected by using higher chain order models and/or more complex intensity distributions.

When considering occurrence, we can envision a single probability distribution for intensity with a large density at zero to account for dry days (Salvucci and Song 2000; Robertson et al. 2009), or we can think of it as two processes: an occurrence process and an intensity process, which will not necessarily be statistically independent. This separation is potentially useful in cases where some physical processes may be tied more directly to either occurrence or intensity—for example, if soil moisture is connected to occurrence through convective triggering or SSTs are connected to intensity through synoptic convergence.

To construct our SSWMs, we use seasonally varying variable-order Markov chains to model daily precipitation occurrence and continuous parametric density functions (conditioned on the chain of occurrence patterns) to model intensity. To properly represent the high-frequency variability, we allow the occurrence models to vary their complexity via their chain order and use the corrected Akaike information criterion (AIC_c ; Hurvich and Tsai 1989) to select the appropriate order to optimize the goodness-of-fit/parsimony trade-off.

Our observational precipitation data is drawn from the U.S. Historical Climatology Network (USHCN) dataset beginning on 1 January 1900 and ending on 31 December 2009 (Williams et al. 2006). We focus on station-based data in order to ensure that results are not biased by interpolation algorithms, scale selection, changes in station locations, or coverage gaps as they may be with gridded data (Hamlet and Lettenmaier 2005). From the 1218 available weather USHCN stations, 774 are selected that have at least 80 yr of data with less than 5% of precipitation values missing—the longest period meeting these criteria is used.

Our occurrence models are variable-order Markov chains with daily varying transition probabilities and chain orders (see Gianotti et al. 2013). At each station, a data-pooling window is selected using the AIC_c to improve transition probability estimates. Pooling sizes for occurrence are typically between 15 and 35 days, and all days within that window (across all years) are used in estimating the transition probabilities for the central day's occurrence model. For each day of the year,

parameters are fit to the empirical data for chain orders from zero to five days, and the appropriate model for each day is selected, again using the AIC_c . Pooling sizes are fixed for a given station, while chain orders can vary from day to day. All combinations of pooling sizes from 1 to 91 days are tested using all chain orders from 0 to 5 days for each day of the year. The AIC_c is used over the similar Bayesian information criterion (Schwarz 1978) because the AIC_c has a weaker penalty for overfitting, potentially allowing more variability in the models, which leads to more conservative confidence intervals and PP estimates.

As an example, in Ajo, Arizona (pooling size 17 days), on 18 June the optimal chain order is zero. This means that the modeled probability of precipitation for 18 June is just the climatological frequency of occurrence in the pooling window (10–26 June) around 18 June. In contrast, in Selma, Alabama (pooling size 23 days), on 18 June the selected chain order is two, meaning that there are four different modeled probabilities of precipitation occurrence depending on whether it rained on both the 16th and 17th, neither the 16th nor 17th, just the 16th, or just the 17th. Each of those four probabilities is determined from the climatological frequency of occurrence following the matching two-day history for all days in the observed record within the pooling window of 7–29 June.

Our modeled intensities are drawn from parameterized probability distributions fit to the first two moments of the observed data. To allow for bimodal precipitation distributions with always-finite variances, the intensity models take the form of a five-parameter gamma–gamma probability distribution mixture model:

$$f(x) = \lambda \frac{\beta_1^{\alpha_1} x^{\alpha_1-1} e^{-\alpha_1 \beta_1}}{\Gamma(\alpha_1)} + (1 - \lambda) \frac{\beta_2^{\alpha_2} x^{\alpha_2-1} e^{-\alpha_2 \beta_2}}{\Gamma(\alpha_2)}, \quad (1)$$

where $0 \leq \lambda \leq 1$, α_1 and α_2 are shape parameters, β_1 and β_2 are rate parameters, and $\Gamma(\cdot)$ is the gamma function. The gamma–gamma distribution was selected from among a number of 1–5 parameter models (exponential, gamma, Dagum, Weibull, kappa, lognormal, and mixtures of each) by the AIC_c as the best trade-off between parsimony and goodness of fit. As with the occurrence process, a model is constructed for each combination of station, day of year, pooling size, and chain order (conditioning on the previous m days' occurrence, rather than intensity), the data pooling size is selected using AIC_c , and then the chain order is selected also using AIC_c . This process is fairly computationally intensive, requiring over 10^8 nonlinear five-dimensional constrained optimization solutions, and yields an intensity model for each day of the year at each station. The 365

daily models for occurrence and intensity, each of which is interannually stationary, combine to make our SSWM for a given location. Selected chain orders for occurrence are generally zero to three days, while they are almost exclusively zero (no memory) for intensity. This suggests that, given 100 yr of observational data—even allowing data pooling (effectively increasing observations by an order of magnitude) for more robust parameterization—intensity estimates are not substantially improved by conditioning on the past few days' occurrence patterns.

We then create stochastically generated Monte Carlo weather-only simulations of the historical record using the SSWMs for each location. Simulations are run a single day at a time to make use of the short-term memory of the chain models—for each day, the probability of occurrence and the probability distribution for intensity are determined by the current model state, a function of day of year and the last 0–5 days of simulated data. For each station, 1000 simulations of the 80–110-yr record are produced. These simulations then form a distribution of precipitation records that match the observed record in length, and—for each day of the year—match the probability of precipitation occurrence, the likelihood of precipitation given the past few days of occurrence, and the intensity distribution. The important way in which they do not match the observed record is that—despite being stochastically generated and therefore displaying interannual variability—the SSWM data-generating mechanism is stationary from year to year, and therefore fails to incorporate low-frequency processes that might affect the occurrence/intensity distributions like those driven by SSTs, atmospheric modes, trends in greenhouse gas concentrations, and land cover changes. Any long-term aggregation of the SSWM simulations, then, should generally display lower variability than the observed records upon which they are based.

More importantly for our purposes, variability in the observed record that extends beyond the window of likely variabilities in the SSWM simulations suggests the existence of low-frequency (climate) variability. While not all low-frequency variability is deterministically predictable, this difference between the variability of the observed record and the variability of the distribution of simulated SSWM records is the basis of our estimate for potential predictability. For example, if in a given location the interannual variance of the observed autumn precipitation was equal to the interannual variance of autumn precipitation in the SSWM, then the observed variance can be explained entirely by a high-frequency noise process (our SSWM), and there is no separable low-frequency signal, hence no PP. To account for missing data, we simulate missing days using

the SSWM for that location. This is more conservative than either imputing missing values with resampled observations or omitting the data entirely. [The SSWMs have no persistence beyond five days, and thus any precipitation sequences (longer than five days) that incorporate SSWM data are underdispersed relative to the actual conditions, and will therefore tend to moderately underestimate the PP signal.]

Since the observations and simulations are at daily resolution and we recreate the seasonal cycle by using 365 occurrence/intensity models to comprise the SSWM for each station, we can look at interannual variability for any period greater than or equal to a single day. For this analysis, however, we will focus on the variance and PP of seasonal precipitation totals in an 89-day window, as well as the portion of that variance due to occurrence versus intensity. The variance in total seasonal precipitation in the observed station data ($\sigma_{\text{obs,tot}}^2$) is shown for 89-day windows centered on 15 January, 15 April, 15 July, and 15 October in Figs. 1a–d, respectively. Variability is generally highest in winter on the leeward sides of the southern Appalachians and the Pacific coastal ranges; in spring along the Gulf Coast and reaching into the Ozarks; in summer in the central Great Plains, Gulf Coast, and the Atlantic Coastal Plain; and in fall the broader Southeast and Pacific Northwest coast.

b. Variance decomposition: Occurrence and intensity

Previous research in the tropics has suggested that precipitation occurrence may be more predictable than intensity (Moron et al. 2006; Robertson et al. 2009). To separate the portion of the total variance that is due to occurrence versus intensity, we can use the law of conditional variance:

$$\text{Var}(S_{89}) = E_N(\text{Var}[S_{89}|N_{89}]) + \text{Var}_N(E[S_{89}|N_{89}]), \quad (2)$$

where S_{89} is the total precipitation at a given location in an 89-day window and N_{89} is the number of days with measured precipitation (≥ 0.254 mm) in that window. The first term on the right-hand side is referred to as the “expectation of conditional variances,” and the second as the “variance of conditional means;” for simplicity's sake we will denote them as $\sigma_{\text{obs,int}}^2$ and $\sigma_{\text{obs,occ}}^2$, respectively so that Eq. (2) can be rewritten as

$$\sigma_{\text{obs,tot}}^2 = \sigma_{\text{obs,int}}^2 + \sigma_{\text{obs,occ}}^2. \quad (3)$$

It is important to note that unless occurrence and intensity are independent, then $\sigma_{\text{obs,occ}}^2$ is different (by more than just a linear scaling factor) from the variance in the number of wet days— $\text{Var}(N_{89})$, as used in Gianotti et al. (2013)—and $\sigma_{\text{obs,int}}^2$ is not the same as the variance of the

FIG. 1. The variance of seasonal precipitation totals in the United States from at least 1930 to 2009 (when available, data goes back as far as 1900) for (a) December–February (DJF), (b) March–May (MAM), (c) June–August (JJA), and (d) September–November (SON). Daily data are taken from 774 USHCN stations.

intensity on wet days. They are the portions of total precipitation due to occurrence or intensity, rather than the variances of occurrence or intensity. For further information on the partitioning of occurrence and intensity see the [appendix](#).

Figure 2, the fraction of the total variability that can be attributed to occurrence $\sigma_{\text{obs,occ}}^2 / \sigma_{\text{obs,tot}}^2$, shows that occurrence contributes substantially to the total observed interannual variance—56% when averaged across all

stations and seasons. The portion of variance due to occurrence is generally higher in the west than the east and is generally highest in the fall.

3. Simulation results

The mean variance (mean of all 1000 simulations) in total precipitation from the SSWM simulations $\sigma_{\text{sim,tot}}^2$ is shown in Fig. 3, broken down again by 89-day season.

FIG. 2. The fraction of variance in the observed seasonal total precipitation record attributable to occurrence $\sigma_{\text{obs,occ}}^2 / \sigma_{\text{obs,tot}}^2$ for (a) DJF, (b) MAM, (c) JJA, and (d) SON. Subtract plotted values from one to obtain the fraction of variance attributable to intensity.

FIG. 3. The variance of seasonal precipitation totals in the United States, as simulated by interannually stationary models fit to the historical record for (a) DJF, (b) MAM, (c) JJA, and (d) SON. This variance serves as a best estimate for the degree of weather noise, that is, the variance due to only short-memory processes with time scales on the order of days. While the models used to generate the simulated data are designed to recreate the observed day-to-day variance exactly, they lack low-frequency, annually varying climate signals found in the observed system, and hence have lower overall variance for seasonal and longer accumulations (cf. Fig. 1).

This is the simulated analog of $\sigma_{\text{obs,tot}}^2$, shown in Fig. 1, but represents only the amount of variability that can be well modeled by stationary processes. Unsurprisingly, the spatial patterns are similar, with the major difference being that the simulated (i.e., resulting from high-frequency processes) variance of aggregated totals is lower in magnitude than the observed (high and low frequency) variance upon which it is based. To determine the contributions to $\sigma_{\text{sim,tot}}^2$ variability from intensity $\sigma_{\text{sim,int}}^2$ and occurrence $\sigma_{\text{sim,occ}}^2$, we decompose the variability of each simulated dataset in the same manner as for the observational data using Eq. (2). The fraction of the mean total simulated variability due to occurrence $\overline{\sigma_{\text{sim,occ}}^2} / \overline{\sigma_{\text{sim,tot}}^2}$ is shown in Fig. 4 as an analog to Fig. 2 (with overbars used to represent the mean of the distribution of simulations).

Focusing first on winter [December–February (DJF)], we notice a difference between the fraction of total variance due to occurrence in the observations (Fig. 2a) and simulations (Fig. 4a), specifically that the simulated data have a less pronounced west–east split in the fraction of variance due to occurrence (or intensity), as well as a smaller range than the observed data. This suggests that much of the low-frequency variability in the observations (absent in the simulations) is caused by processes that affect occurrence or intensity preferentially, with low-frequency occurrence patterns predominant in the western United States (and along the southern ENSO track) and low-frequency intensity variability

stronger in the eastern United States. For the other seasons, the same pattern of stronger observed occurrence/intensity separation holds, although the spatial patterns change (particularly in the Northwest in the summer and fall, the Mississippi/Ohio/Tennessee River valleys in the summer and fall, the central Great Plains in the spring through fall, and northern Appalachians in the winter through summer).

4. Results of potential predictability analysis

Using this variability data, we can now determine the seasonal potential predictability qualitatively suggested by the variance maps in the previous sections. The potential predictability is defined as the fraction of the observed variance that can be ascribed to climatic (low frequency) variability:

$$\text{PP}_{89,\text{tot}} = \frac{\sigma_{\text{obs,tot}}^2 - \overline{\sigma_{\text{sim,tot}}^2}}{\sigma_{\text{obs,tot}}^2}. \quad (4)$$

The subscript 89 in $\text{PP}_{89,\text{tot}}$ is a reminder that this is the potential predictability of 89-day seasonal precipitation, and the subscript “tot” is to clarify that this is defined for the variability of the total accumulated precipitation, rather than the portions of variability due to occurrence or intensity. The values of $\text{PP}_{89,\text{tot}}$ fall in the range from zero to one, although in some cases when there is little nonstationary signal and $\sigma_{\text{obs,tot}}^2$ lies in the middle of the

FIG. 4. The fraction of variance in the simulated seasonal total precipitation records attributable to occurrence $\sigma_{\text{sim,occ}}^2 / \sigma_{\text{sim,tot}}^2$ serving as an estimate of the fraction of high-frequency weather noise caused by occurrence. Subtract plotted values from one to obtain the fraction of variance attributable to intensity for (a) DJF, (b) MAM, (c) JJA, and (d) SON.

$\sigma_{\text{sim,tot}}^2$ distribution, $\sigma_{\text{sim,tot}}^2$ may exceed $\sigma_{\text{obs,tot}}^2$, leading to negative $\text{PP}_{89,\text{tot}}$ values, which imply no ability to extract low-frequency variability from the envelop of high-frequency noise (and hence no PP). As a reminder, if this same analysis was performed using two-day rather than 89-day accumulations (for a location with a chain order of, say, two or three days), the model's variance would exactly match the observed variance, the average of the simulated variances across all 1000 simulations would asymptotically equal the observed variance, and the potential predictability would be zero. It is only when accumulating beyond the chain order that the persistence of climatic boundary conditions (not represented by our models) can lead to precipitation totals that are unusual when compared to the stationary assumptions of an SSWM.

Using the decomposed simulated and observed variances, we can also establish the potential predictability of climate variability as driven by occurrence and intensity. The potential predictability of total 89-day precipitation due to intensity ($\text{PP}_{89,\text{int}}$) is defined as $\text{PP}_{89,\text{int}} = (\sigma_{\text{obs,int}}^2 - \sigma_{\text{sim,int}}^2) / \sigma_{\text{obs,int}}^2$. Similarly, $\text{PP}_{89,\text{occ}} = (\sigma_{\text{obs,occ}}^2 - \sigma_{\text{sim,occ}}^2) / \sigma_{\text{obs,occ}}^2$ is the potential predictability of 89-day precipitation totals due to occurrence.

To determine whether a PP value is statistically significantly different from zero or indistinguishable from zero, we calculate $\text{PP}_{89,\text{tot}}$, $\text{PP}_{89,\text{int}}$, and $\text{PP}_{89,\text{occ}}$ values using each of the simulations in place of the observed record. If the PP value (using the observed record) exceeds $1 - \alpha$ of the simulated values, then we can reject

the null hypothesis that our SSWM adequately reproduces the variability of the observed record (and therefore we can confidently say that $\text{PP} > 0$). If the PP value does not exceed $1 - \alpha$ of the simulated values, then we cannot reject the null hypothesis (equivalently our SSWM statistically reproduces the full variability of the observed record).

Seasonal maps of $\text{PP}_{89,\text{tot}}$, $\text{PP}_{89,\text{int}}$, and $\text{PP}_{89,\text{occ}}$ are shown in Figs. 5–8. Markers show station locations, colored by their PP values. Stations shown as circles are significantly different from $\text{PP} = 0$ at the $\alpha = 0.1$ level, and stations shown as triangles are not significantly different from $\text{PP} = 0$ at the $\alpha = 0.1$ level. Higher PP values are (perhaps not surprisingly) more likely to be statistically significantly different from zero, and thus represent conditions that cannot be represented by SSWMs with 0–5 days of memory. Lower PP values are typically not significantly different from zero, and thus the observed precipitation statistics at these locations can easily be recreated by our weather models.

Gridded values shown in the background of Figs. 5–8 are determined using a spatial analysis of the station-based PP values. As is evident in both the station values (symbols) and the interpolated background, there is a large degree of geographical coherence in the estimated PP values. This coherence is not a simple result of overall precipitation climatology (Pearson correlation between all 365×774 $\text{PP}_{89,\text{tot}}$ values and daily mean precipitation is -0.03), and so any spatial coherence that emerges from the analysis lends further confidence to

FIG. 5. Potential predictability of precipitation for a seasonal (89 day) window centered on 15 January, shown as the fraction of seasonal precipitation variance attributable to low-frequency (climate) processes for (a) PP of total precipitation ($PP_{89,tot}$, unitless); (b) PP of intensity processes ($PP_{89,int}$, unitless); and (c) PP of occurrence processes ($PP_{89,occ}$, unitless). Circles denote locations with PP significantly different from 0, and triangles denote locations where PP is not significantly different from 0 ($\alpha = 0.1$). Gridded underlying values are modeled based on spatial coherence. DJF displays the highest $PP_{89,tot}$, $PP_{89,int}$, and $PP_{89,occ}$ values of all seasons when averaged across stations. The DJF station-averaged PP values for all three metrics are nearly equal at roughly 0.3. Note that $PP_{89,tot}$ and $PP_{89,occ}$ have a very similar spatial pattern, with large potential climate signals along the U.S. West Coast, Rockies, Southwest, Florida, and the Michigan/Ohio River valley area. Intensity signals are notably lower in the Southwest and are slightly higher along the Piedmont Plateau/Atlantic Coastal Plain.

our estimates. Gridded estimates are determined using a generalized cross validation (GCV) scheme for fitting thin plate splines (ANUSPLIN; Wahba 1979; Hutchinson 1998). The 90% confidence intervals from the GCV analysis around the gridded values are fairly spatially homogeneous, suggesting a spatial uncertainty in our PP estimates of roughly ± 0.05 (maps of 90% confidence intervals are shown in the supplementary material).

FIG. 6. As in Fig. 5, but for a seasonal (89 day) window centered on 15 April. All of the PP metrics are generally lower in MAM than in DJF throughout the contiguous 48 states, with $PP_{89,int}$ particularly low and incoherent (station-based average of roughly 0.2). Springtime occurrence processes show higher potential predictability in the central Great Plains and New England as compared to winter, but lower in the Northwest and Great Lakes regions. Note that $PP_{89,tot}$ and $PP_{89,occ}$ have a similar pattern, although on average $PP_{89,occ}$ exceeds $PP_{89,tot}$ by 25%.

Looking at $PP_{89,tot}$ (Figs. 5a, 6a, 7a, and 8a), winter displays the most potential predictability, including broadly across the U.S. West, western Texas, Florida, and in the western Appalachians/Great Lakes regions. The U.S. West and Florida are strongly associated with ENSO precipitation (Ropelewski and Halpert 1986), and the area south of the Great Lakes may have connections to variability in lake effect snow, as driven by variability in winter storm tracks (Niziol et al. 1995). The Pacific Northwest shows high $PP_{89,tot}$ in the fall, also likely connected to ENSO patterns (Ropelewski and Halpert 1986). Also significantly, the Gulf and southern Atlantic coast, which displayed some of the highest observed precipitation variability during the June–November Atlantic hurricane season, shows little $PP_{89,tot}$ at these

FIG. 7. As in Fig. 5, but for a seasonal (89 day) window centered on 15 July. Summer is the potential predictability minimum for each of $PP_{89,tot}$, $PP_{89,int}$, and $PP_{89,occ}$, with station-averaged total and occurrence values in a similar range of roughly 0.15 and intensity PP at half that. The most predictable areas are the Rockies and central Great Plains for total precipitation and occurrence, and the North American monsoon region of Arizona/New Mexico for intensity.

times, suggesting that the low-frequency total precipitation signal in this area is small compared with the weather noise, a discouraging sign for seasonal forecasting of tropical storm-induced precipitation locally.

Note that $PP_{89,int}$ is shown in Figs. 5b, 6b, 7b, and 8b. Like $PP_{89,tot}$, $PP_{89,int}$ has a station-averaged maximum in the winter, with a similar PP “hot spot” in the western Appalachians/Great Lakes region and high values throughout much of the broader U.S. West, but without the low values across the Southwest. PP due to intensity is low and spotty in the summer and fall (with the highest summertime values in the Southwest monsoon region) and is much more spatially correlated with the intensity climate signal ($\sigma_{obs,int}^2 - \sigma_{sim,int}^2$, Pearson correlation $r = 0.573$) than it is anticorrelated with the weather noise ($\sigma_{sim,int}^2$, $r = -0.064$). The relatively low values of

$PP_{89,int}$ [DJF: 0.257; March–May (MAM): 0.181; June–August (JJA): 0.123; and September–November (SON): 0.156, when averaged across all stations] indicate that our SSWMs capture intensity variability well, suggesting that the low-frequency processes that drive the intensity of precipitation variance on seasonal scales are typically small relative to the high-frequency noise.

Results for $PP_{89,occ}$ are shown in Figs. 5c, 6c, 7c and 8c. The $PP_{89,occ}$ values (DJF: 0.237; MAM: 0.243; JJA: 0.203; and SON: 0.237) are generally higher than $PP_{89,tot}$ values; and, like $PP_{89,int}$, $PP_{89,occ}$ is much more highly correlated with its climate signal ($\sigma_{obs,occ}^2 - \sigma_{sim,occ}^2$, $r = 0.551$) than anticorrelated with its weather noise ($\sigma_{sim,occ}^2$, $r = 0.023$). Note that $PP_{89,occ}$ is also much more spatially correlated with $PP_{89,tot}$ than is $PP_{89,int}$ ($PP_{89,occ} \propto PP_{89,tot}$, $r = 0.821$ and $PP_{89,int} \propto PP_{89,tot}$, $r = 0.448$). The highest values of $PP_{89,occ}$ (greater than 0.7) occur in the U.S. West, through Texas, and in the Ohio River valley region in winter (Fig. 5c); in California, the Southwest, and the Cornbelt in the spring (Fig. 6c); in the Rockies and the Cornbelt in summer (Fig. 7c); and in the Pacific Northwest and along a swath between Milwaukee and Kansas City in the fall (Fig. 8c). From September through May (Figs. 5, 6, and 8), the patterns are similar to slight amplifications of the total PP, but the summer shows $PP_{89,occ}$ hot spots in the central Great Plains and the Rockies that are much higher than in the $PP_{89,tot}$ map (Figs. 7a,c).

This highlights an important point, namely that $PP_{89,tot} \neq PP_{89,occ} + PP_{89,int}$, and in fact the separated occurrence and intensity PP values are often higher than the total PP signal. We can think of this as a destructive interference of sorts, where the weather noise in the intensity process masks an otherwise persistent occurrence signal (or vice versa) when the two are examined together as $PP_{89,tot}$. The summer central Great Plains $PP_{89,occ}$ hot spot is a good example of this; most of the variability in the region is tied to occurrence processes (Fig. 2c), but roughly half of the weather noise is due to intensity processes (Fig. 4c), and so the climate signal stands out against the weather noise for occurrence (high values in Fig. 7c) much more so than for intensity (low values in Fig. 7b). When combined, however, the low signal-to-noise ratio from intensity weakens the overall signal-to-noise ratio for total precipitation, and so $PP_{89,tot}$ (Fig. 7a) is lower than $PP_{89,occ}$. This area is also noteworthy as the site of high modeled potential predictability due to land–atmosphere coupling (Koster et al. 2004). We might then imagine that this strong signal from occurrence is due to a precipitation recycling amplification in the number of rainy days in years with sufficient moisture in the area [as has been previously found by Findell et al. (2011)]. The difference between years in which

FIG. 8. As in Fig. 5, but for a seasonal (89 day) window centered on 15 October. Autumn PP is highest in the Northwest and the Cornbelt region, although intensity processes do not seem to demonstrate much coherent potential predictability. The Cornbelt area demonstrates the destructive interference of convolving intensity and occurrence signals— $PP_{89,tot}$ is lower than $PP_{89,occ}$ because high-frequency noise in the intensity processes mask climate signals in the occurrence processes when combined as total precipitation.

sufficient moisture is present to trigger recycling (and thereby extra rainy days) and those with insufficient moisture for this feedback leads to interannual variability not present in our annually stationary models, yielding potential predictability (i.e., $\sigma_{obs,occ}^2 > \sigma_{sim,occ}^2$, meaning that $PP_{89,occ} > 0$). In general, processes driving PP in these hot spot areas will need a more thorough investigation to establish regional modes of predictability.

5. Summary and discussion

We have used observational data directly to obtain estimates of both the “weather noise” and nonstationary “climate signal” contributions to precipitation variability,

and have additionally decomposed these signals into contributions from occurrence and intensity processes. This is significant in that precipitation occurrence displays more potential predictability than intensity, suggesting that those processes responsible for triggering occurrence may be particularly important when attributing variability. This finding is in agreement with previous work in the tropics and Southern Hemisphere (Moron et al. 2007; Robertson et al. 2009; Verbist et al. 2010), suggesting that precipitation occurrence may generally be a more fruitful avenue for improving seasonal forecasts than intensity-related processes. We find that occurrence also displays higher PP in most cases than total precipitation, suggesting that modeling and forecasting efforts might be well served by focusing on occurrence first, rather than seasonal totals. Intensity is also occasionally more potentially predictable than total precipitation, and the two components should be investigated separately to avoid the loss of information that comes with convolving the two signals. Some care will be necessary in approaching occurrence using either global climate models (GCMs) or gridded datasets, as the spatially amalgamated discrete data will almost definitely yield different results at finite spatial scales rather than point measurements (Hughes et al. 1999).

Generally, the potential predictability results shown in Figs. 5–8 should help guide those attempting to attribute precipitation variability to processes. PP “hot spot” regions with high potential climate signals should guide researchers to both investigate processes and to work toward improved local forecasts if possible. Equally importantly, PP “cold spots” highlight regions in which climate signals may be difficult to disentangle from chaotic, high-frequency weather noise. Where PP values are low—and particularly where PP is not significantly different from zero—our SSWMs represent the interannual variability of 89-day precipitation totals well, despite having only at most five days of persistence and no representation of interannual processes. In some cases, the relative role of weather noise may not matter so much as the absolute magnitude of the potential climate signal $\sigma_{obs}^2 - \sigma_{sim}^2$. (Maps of this absolute potential signal strength for total precipitation, occurrence, and intensity are available as supplementary Figs. S2–S5.)

In the introduction, we suggested applying PP to 1) estimate the limits of our ability to confidently forecast longer-term climate behaviors in analyses of precipitation time series, 2) determine where and when our present forecasts are limited by an understanding of the system versus an absence of useful climate signals, and 3) test whether or not our current forecast models display signals and noise in the same manner as the real climate system. This study provides information on the

FIG. A1. An example of the decomposition of total precipitation into occurrence and intensity contributions following the law of conditional variance for USHCN weather station 335041 in McConnelsville, Ohio. (a) Observed 89-day precipitation vs number of precipitation events. (b) Histogram of seasons with a given number of precipitation events. (c) Mean precipitation values conditioned on number of events (green lines) and the standard deviation of the conditional means (red bar), the square root of the variance of conditional means σ_{occ}^2 . (d) Variance of precipitation conditioned on number of events (blue bars) and the mean of conditional variances σ_{int}^2 (dashed line). (e) Partitioning of variance, as from (c) and (d), as per the law of conditional variance.

first of these applications. Precipitation variability at locations with low PP is difficult to distinguish from the noise of a simple stationary weather model with only a few days of persistence, and so climate signals in those regions cannot be distinguished from what we think of as weather noise. Locations with higher PP values show significant divergences from stationary processes, and so may be predictable, but the ability to predict precipitation from climate indices in these locations is limited by both our knowledge of the proper climate indices and the coherence of their relationship with precipitation. Determining the degree to which each of

these limits our ability to make optimal forecasts of precipitation (i.e., the second proposed application for PP) is a substantial area of research, one that weather and climate forecast centers address each day. The third suggested application of PP information is to see whether global climate models follow patterns of potential predictability similar to the observational data upon which they are based [see previous work by Moron et al. (2006)]. These latter applications are important areas of future research.

Additional open questions remain regarding the drivers of potential predictability—specifically, how

much of PP is attributable to climate oscillations versus trends in the historic precipitation record? Also, the prediction of extremes is important both for planning and because increases in extreme precipitation events are one of the most consistent findings from past decades. Assessment of PP values for extreme events can be performed using our dataset and methods, but will require different metrics than those applied in this study.

Acknowledgments. The authors thank Angela Rigden for her help with the spatial analysis. We thank the U.S. Department of Energy and National Science Foundation for providing funding support for this work. We would also like to thank our reviewers for their time and their suggestions.

APPENDIX

Variance Decomposition into Occurrence and Intensity

The law of conditional variance states that

$$\text{Var}(\Pi) = E_{\Theta}(\text{Var}[\Pi | \Theta]) + \text{Var}_{\Theta}(E[\Pi | \Theta]). \quad (\text{A1})$$

We will use this to decompose total precipitation in a seasonal window into components due to the number of days of precipitation occurrence in that window, and the intensity on those days. Let π be the total precipitation observed in an 89-day window with θ days of precipitation occurrence in those 89 days. Here Π is a vector $(\pi_1, \pi_2, \dots, \pi_m)$ and Θ is similarly the vector $(\theta_1, \theta_2, \dots, \theta_m)$, where each element of the vectors is the seasonal observation for one year (i.e., in the first year there were θ_1 rainy days that sum to a total of π_1 millimeters of rain).

The green diamonds in Fig. A1a each represent a single winter's (θ_i, π_i) pair for a rain gauge at McConnelville Lock, Ohio, and Fig. A1b is a histogram of the number of precipitation events over all observed years. The second term on the right-hand side of Eq. (A1) is referred to as the variance of conditional means and is displayed visually in Fig. A1c. The blue circles in Fig. A1a are the mean seasonal precipitation conditioned on the number of wet days, as are the green lines in Fig. A1c. The red dot in Fig. A1c is the mean of the conditional means, and the error bars show one standard deviation—the square root of σ_{occ}^2 , as shown in Fig. A1e.

The first term on the right-hand side of Eq. (A1) is referred to as the expectation of conditional variances and is calculated by first taking the variances of the seasonal total precipitation as conditioned on the number of

wet days (blue error bars in Fig. A1a show one standard deviation, blue bars below in Fig. A1d show the variance). The mean across all observed θ values is shown with the dashed blue bar in Fig. A1d and as the blue region in Fig. A1e, also called σ_{int}^2 .

In this example, σ_{occ}^2 is greater than σ_{int}^2 as shown by the comparatively large spread between the blue dots in Fig. A1a as compared to the within- θ deviations around those means. This partitioning is equivalent to the notion of variance explained by a model (in this case the model is the mean precipitation per event dashed line in Fig. A1a) in a typical regression setting. If it always rained precisely the same amount on each wet day, our occurrence-based model would fit the data perfectly (no residual sum of squares), all of the blue dots would fall on the dashed red line in Fig. A1a, and there would be no σ_{int}^2 . Similarly, if it always rained the same amount in each season no matter how many wet days that season had, occurrence would explain none of the variance in the data, there would be no dispersion about the mean in Fig. A1c, and hence no σ_{occ}^2 .

The decomposition could potentially show less stability for seasons with very few events, as the estimated precipitation per event would be calculated using fewer observed data points (see supplementary material for discussion). In this analysis, this potential source of uncertainty is negligible for two reasons: 1) the coefficients of variation for the $\sigma_{\text{occ}}^2/\sigma_{\text{tot}}^2$ ratio are generally low and not a function of the number of wet days except for in one autumn and nine summertime locations (see supplementary Fig. S11), and 2) those locations with higher coefficients of variation (>0.25) are in the driest regions, which display consistently low (and not significantly different from zero) PP values for all of PP_{89,tot}, PP_{89,int}, and PP_{89,occ}.

REFERENCES

- Anderson, B. T., J. Wang, S. Gopal, and G. D. Salvucci, 2009: Influence of daily rainfall characteristics on regional summertime precipitation over the southwestern United States. *J. Hydrometeorol.*, **10**, 1218–1230, doi:10.1175/2009JHM1104.1.
- Becker, E. J., E. H. Berbery, and R. W. Higgins, 2011: Modulation of cold-season U.S. daily precipitation by the Madden–Julian oscillation. *J. Climate*, **24**, 5157–5166, doi:10.1175/2011JCL4018.1.
- Boer, G., 2009: Changes in interannual variability and decadal potential predictability under global warming. *J. Climate*, **22**, 3098–3109, doi:10.1175/2008JCL12835.1.
- Changnon, S. A., 2001: Thunderstorm rainfall in the conterminous United States. *Bull. Amer. Meteor. Soc.*, **82**, 1925–1940, doi:10.1175/1520-0477(2001)082<1925:TRITCU>2.3.CO;2.
- Collins, M., 2002: Climate predictability on interannual to decadal time scales: The initial value problem. *Climate Dyn.*, **19**, 671–692, doi:10.1007/s00382-002-0254-8.
- Feng, X., T. DelSole, and P. Houser, 2011: Bootstrap estimated seasonal potential predictability of global temperature and

- precipitation. *Geophys. Res. Lett.*, **38**, L07702, doi:[10.1029/2010GL046511](#).
- Ferguson, I. M., P. B. Duffy, T. J. Phillips, X. Liang, J. A. Dracup, S. D. Schubert, and P. Pegion, 2011: Non-stationarity of the signal and noise characteristics of seasonal precipitation anomalies. *Climate Dyn.*, **36**, 739–752, doi:[10.1007/s00382-010-0850-y](#).
- Findell, K., P. Gentile, B. Lintner, and C. Kerr, 2011: Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation. *Nat. Geosci.*, **4**, 434–439, doi:[10.1038/ngeo1174](#).
- Gianotti, D. J., B. T. Anderson, and G. D. Salvucci, 2013: What do rain gauges tell us about the limits of precipitation predictability? *J. Climate*, **26**, 5682–5688, doi:[10.1175/JCLI-D-12-00718.1](#).
- Gong, X., A. G. Barnston, and M. N. Ward, 2003: The effect of spatial aggregation on the skill of seasonal precipitation forecasts. *J. Climate*, **16**, 3059–3071, doi:[10.1175/1520-0442\(2003\)016<3059:TEOSAO>2.0.CO;2](#).
- Gregory, J., T. Wigley, and P. Jones, 1993: Application of Markov models to area-averaged daily precipitation series and interannual variability in seasonal totals. *Climate Dyn.*, **8**, 299–310, doi:[10.1007/BF00209669](#).
- Groisman, P. Ya., and R. W. Knight, 2008: Prolonged dry episodes over the conterminous United States: New tendencies emerging during the last 40 years. *J. Climate*, **21**, 1850–1862, doi:[10.1175/2007JCLI2013.1](#).
- , and Coauthors, 1999: Changes in the probability of heavy precipitation: Important indicators of climatic change. *Climatic Change*, **42**, 243–283, doi:[10.1023/A:1005432803188](#).
- , R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. N. Razuvaev, 2005: Trends in intense precipitation in the climate record. *J. Climate*, **18**, 1326–1350, doi:[10.1175/JCLI3339.1](#).
- Hamlet, A. F., and D. P. Lettenmaier, 2005: Production of temporally consistent gridded precipitation and temperature fields for the continental United States. *J. Hydrometeorol.*, **6**, 330–336, doi:[10.1175/JHM420.1](#).
- Hanson, L. S., and R. Vogel, 2008: The probability distribution of daily rainfall in the United States. *Proc. World Environmental and Water Resources Congress 2008*, Honolulu, HI, ASCE, doi:[10.1061/40976\(316\)585](#).
- Higgins, R. W., and V. Kousky, 2013: Changes in observed daily precipitation over the United States between 1950–79 and 1980–2009. *J. Hydrometeorol.*, **14**, 105–121, doi:[10.1175/JHM-D-12-062.1](#).
- , V. B. S. Silva, W. Shi, and J. Larson, 2007: Relationships between climate variability and fluctuations in daily precipitation over the United States. *J. Climate*, **20**, 3561–3579, doi:[10.1175/JCLI4196.1](#).
- Hughes, J. P., P. Guttorp, and S. P. Charles, 1999: A non-homogeneous hidden Markov model for precipitation occurrence. *J. Roy. Stat. Soc.*, **48C**, 15–30, doi:[10.1111/1467-9876.00136](#).
- Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269**, 676–679, doi:[10.1126/science.269.5224.676](#).
- Hurvich, C. M., and C.-L. Tsai, 1989: Regression and time series model selection in small samples. *Biometrika*, **76**, 297–307, doi:[10.1093/biomet/76.2.297](#).
- Hutchinson, M. F., 1998: Interpolation of rainfall data with thin plate smoothing splines. Part I: Two-dimensional smoothing of data with short range correlation. *J. Geogr. Info. Decision Analysis*, **2**, 139–151.
- Jia, L., and T. DelSole, 2011: Diagnosis of multiyear predictability on continental scales. *J. Climate*, **24**, 5108–5124, doi:[10.1175/2011JCLI4098.1](#).
- Karl, T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull. Amer. Meteor. Soc.*, **79**, 231–241, doi:[10.1175/1520-0477\(1998\)079<0231:STOPAF>2.0.CO;2](#).
- Katz, R. W., 1977: Precipitation as a chain-dependent process. *J. Appl. Meteor.*, **16**, 671–676, doi:[10.1175/1520-0450\(1977\)016<0671:PAACDP>2.0.CO;2](#).
- , and M. Parlange, 1998: Overdispersion phenomenon in stochastic modeling of precipitation. *J. Climate*, **11**, 591–602, doi:[10.1175/1520-0442\(1998\)011<0591:OPISMO>2.0.CO;2](#).
- Knapp, A. K., and Coauthors, 2002: Rainfall variability, carbon cycling, and plant species diversity in a mesic grassland. *Science*, **298**, 2202–2205, doi:[10.1126/science.1076347](#).
- Knight, D. B., and R. E. Davis, 2009: Contribution of tropical cyclones to extreme rainfall events in the southeastern United States. *J. Geophys. Res.*, **114**, D23102, doi:[10.1029/2009JD012511](#).
- Koster, R. D., M. J. Suarez, and M. Heiser, 2000: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeorol.*, **1**, 26–46, doi:[10.1175/1525-7541\(2000\)001<0026:VAPOPA>2.0.CO;2](#).
- , and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, doi:[10.1126/science.1100217](#).
- Kunkel, K. E., D. R. Easterling, K. Redmond, and K. Hubbard, 2003: Temporal variations of extreme precipitation events in the United States: 1895–2000. *Geophys. Res. Lett.*, **30**, 1900, doi:[10.1029/2003GL018052](#).
- , —, D. R. Kristovich, B. Gleason, L. Stoecker, and R. Smith, 2012: Meteorological causes of the secular variations in observed extreme precipitation events for the conterminous United States. *J. Hydrometeorol.*, **13**, 1131–1141, doi:[10.1175/JHM-D-11-0108.1](#).
- , T. R. Karl, D. R. Easterling, K. Redmond, J. Young, X. Yin, and P. Hennon, 2013: Probable maximum precipitation and climate change. *Geophys. Res. Lett.*, **40**, 1402–1408, doi:[10.1002/grl.50334](#).
- Leith, C., 1978: Predictability of climate. *Nature*, **276**, 352–355, doi:[10.1038/276352a0](#).
- Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333, doi:[10.1111/j.2153-3490.1965.tb01424.x](#).
- Lynch, P., 2008: The origins of computer weather prediction and climate modeling. *J. Comput. Phys.*, **227**, 3431–3444, doi:[10.1016/j.jcp.2007.02.034](#).
- Madden, R. A., D. J. Shea, R. W. Katz, and J. W. Kidson, 1999: The potential long-range predictability of precipitation over New Zealand. *Int. J. Climatol.*, **19**, 405–421, doi:[10.1002/\(SICI\)1097-0088\(19990330\)19:4<405::AID-JOC355>3.0.CO;2-U](#).
- Mandelbrot, B. B., and J. Wallis, 1968: Noah, Joseph, and operational hydrology. *Water Resour. Res.*, **4**, 909–918, doi:[10.1029/WR004i005p00909](#).
- Mantua, N. J., and S. R. Hare, 2002: The Pacific decadal oscillation. *J. Oceanogr.*, **58**, 35–44, doi:[10.1023/A:1015820616384](#).
- Meisner, B. N., and P. A. Arkin, 1987: Spatial and annual variations in the diurnal cycle of large-scale tropical convective cloudiness and precipitation. *Mon. Wea. Rev.*, **115**, 2009–2032, doi:[10.1175/1520-0493\(1987\)115<2009:SAAVIT>2.0.CO;2](#).
- Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer, 2008: Stationarity is dead: Whither water management? *Science*, **319**, 573–574, doi:[10.1126/science.1151915](#).
- Moron, V., A. W. Robertson, and M. N. Ward, 2006: Seasonal predictability and spatial coherence of rainfall characteristics

- in the tropical setting of Senegal. *Mon. Wea. Rev.*, **134**, 3248–3262, doi:[10.1175/MWR3252.1](https://doi.org/10.1175/MWR3252.1).
- , —, —, and P. Camberlin, 2007: Spatial coherence of tropical rainfall at the regional scale. *J. Climate*, **20**, 5244–5263, doi:[10.1175/2007JCLI1623.1](https://doi.org/10.1175/2007JCLI1623.1).
- NCADAC, cited 2014: Third National Climate Assessment Report. U.S. Global Change Research Program. [Available online at <http://nca2014.globalchange.gov/report/>.]
- Niziol, T., W. Snyder, and J. Waldstreicher, 1995: Winter weather forecasting throughout the eastern United States. Part IV: Lake effect snow. *Wea. Forecasting*, **10**, 61–77, doi:[10.1175/1520-0434\(1995\)010<0061:WWFTTE>2.0.CO;2](https://doi.org/10.1175/1520-0434(1995)010<0061:WWFTTE>2.0.CO;2).
- NOAA Climate Prediction Center, cited 2013: CPC verification summary. NOAA. [Available online at <http://www.cpc.ncep.noaa.gov/products/verification/summary/>.]
- Olsson, J., and J. Niemczynowicz, 1996: Multifractal analysis of daily spatial rainfall distributions. *J. Hydrol.*, **187**, 29–43, doi:[10.1016/S0022-1694\(96\)03085-5](https://doi.org/10.1016/S0022-1694(96)03085-5).
- Pal, I., B. T. Anderson, G. D. Salvucci, and D. J. Gianotti, 2013: Shifting seasonality and increasing frequency of precipitation in wet and dry seasons across the U.S. *Geophys. Res. Lett.*, **40**, 4030–4035, doi:[10.1002/grl.50760](https://doi.org/10.1002/grl.50760).
- Pruppacher, H. R., and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation*. 2nd ed. Kluwer Academic, 954 pp.
- Pryor, S. C., and J. T. Schoof, 2008: Changes in the seasonality of precipitation over the contiguous USA. *J. Geophys. Res.*, **113**, D21108, doi:[10.1029/2008JD010251](https://doi.org/10.1029/2008JD010251).
- Robertson, A. W., S. Kirshner, P. Smyth, S. P. Charles, and B. C. Bates, 2006: Subseasonal-to-interdecadal variability of the Australian monsoon over North Queensland. *Quart. J. Roy. Meteor. Soc.*, **132**, 519–542, doi:[10.1256/qj.05.75](https://doi.org/10.1256/qj.05.75).
- , V. Moron, and Y. Swarinoto, 2009: Seasonal predictability of daily rainfall statistics over Indramayu district, Indonesia. *Int. J. Climatol.*, **29**, 1449–1462, doi:[10.1002/joc.1816](https://doi.org/10.1002/joc.1816).
- Ropelewski, C. F., and M. S. Halpert, 1986: North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Wea. Rev.*, **114**, 2352–2362, doi:[10.1175/1520-0493\(1986\)114<2352:NAPATP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1986)114<2352:NAPATP>2.0.CO;2).
- Salvucci, G. D., and C. Song, 2000: Derived distributions of storm depth and frequency conditioned on monthly total precipitation: Adding value to historical and satellite-derived estimates of monthly precipitation. *J. Hydrometeorol.*, **1**, 113–120, doi:[10.1175/1525-7541\(2000\)001<0113:DDOSDA>2.0.CO;2](https://doi.org/10.1175/1525-7541(2000)001<0113:DDOSDA>2.0.CO;2).
- Schoof, J. T., and S. C. Pryor, 2008: On the proper order of Markov chain model for daily precipitation occurrence in the contiguous United States. *J. Appl. Meteor. Climatol.*, **47**, 2477–2486, doi:[10.1175/2008JAMC1840.1](https://doi.org/10.1175/2008JAMC1840.1).
- Schwarz, G., 1978: Estimating the dimension of a model. *Ann. Stat.*, **6**, 461–464, doi:[10.1214/aos/1176344136](https://doi.org/10.1214/aos/1176344136).
- Stern, R., and R. Coe, 1984: A model fitting analysis of daily rainfall data. *J. Roy. Stat. Soc.*, **147A**, 1–34, doi:[10.2307/2981736](https://doi.org/10.2307/2981736).
- Sun, Y., S. Solomon, A. Dai, and R. Portmann, 2007: How often will it rain? *J. Climate*, **20**, 4801–4818, doi:[10.1175/JCLI4263.1](https://doi.org/10.1175/JCLI4263.1).
- Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.*, **84**, 1205–1217, doi:[10.1175/BAMS-84-9-1205](https://doi.org/10.1175/BAMS-84-9-1205).
- Verbist, K., A. W. Robertson, W. M. Cornelis, and D. Gabriels, 2010: Seasonal predictability of daily rainfall characteristics in central northern Chile for dry-land management. *J. Appl. Meteor. Climatol.*, **49**, 1938–1955, doi:[10.1175/2010JAMC2372.1](https://doi.org/10.1175/2010JAMC2372.1).
- Wahba, G., 1979: How to smooth curves and surfaces with splines and cross-validation. Tech. Rep., University of Wisconsin–Madison, 16 pp.
- Wang, J., B. T. Anderson, and G. D. Salvucci, 2006: Stochastic modeling of daily summertime rainfall over the southwestern United States. Part I: Interannual variability. *J. Hydrometeorol.*, **7**, 739–754, doi:[10.1175/JHM518.1](https://doi.org/10.1175/JHM518.1).
- Westra, S., and A. Sharma, 2010: An upper limit to seasonal rainfall predictability? *J. Climate*, **23**, 3332–3351, doi:[10.1175/2010JCLI3212.1](https://doi.org/10.1175/2010JCLI3212.1).
- , L. V. Alexander, and F. W. Zwiers, 2013: Global increasing trends in annual maximum daily precipitation. *J. Climate*, **26**, 3904–3918, doi:[10.1175/JCLI-D-12-00502.1](https://doi.org/10.1175/JCLI-D-12-00502.1).
- Wilks, D. S., 1999: Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric. For. Meteorol.*, **93**, 153–169, doi:[10.1016/S0168-1923\(98\)00125-7](https://doi.org/10.1016/S0168-1923(98)00125-7).
- Williams, C. N., R. S. Vose, D. R. Easterling, and M. J. Menne, 2006: United States Historical Climatology Network daily temperature, precipitation, and snow data. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory Oak Ridge, TN, digital media. [Available online at http://gcmd.nasa.gov/records/GCMD_CDIAC_NDP070.html.]
- Yilmaz, M. T., and T. DelSole, 2010: Predictability of seasonal precipitation using joint probabilities. *J. Hydrometeorol.*, **11**, 533–541, doi:[10.1175/2009JHM1187.1](https://doi.org/10.1175/2009JHM1187.1).