Tracking Interannual Streamflow Variability with Drought Indices in the U.S. Pacific Northwest

JOHN T. ABATZOGLOU, RENAUD BARBERO, AND JACOB W. WOLF

Department of Geography, University of Idaho, Moscow, Idaho

ZACHARY A. HOLDEN

Northern Region, Forest Service, USDA, Missoula, Montana

(Manuscript received 3 October 2013, in final form 5 May 2014)

ABSTRACT

Drought indices are often used for monitoring interannual variability in macroscale hydrology. However, the diversity of drought indices raises several issues: 1) which indices perform best and where; 2) does the incorporation of potential evapotranspiration (PET) in indices strengthen relationships, and how sensitive is the choice of PET methods to such results; 3) what additional value is added by using higher-spatial-resolution gridded climate layers; and 4) how have observed relationships changed through time. Standardized precipitation index, standardized precipitation evapotranspiration index (SPEI), Palmer drought severity index, and water balance runoff (WBR) model output were correlated to water-year runoff for 21 unregulated drainage basins in the Pacific Northwest of the United States. SPEI and WBR with time scales encompassing the primary precipitation season maximized the explained variance in water-year runoff in most basins. Slightly stronger correlations were found using PET estimates from the Penman–Monteith method over the Thornthwaite method, particularly for time periods that incorporated the spring and summer months in basins that receive appreciable precipitation during the growing season. Indices computed using high-resolution climate surfaces explained over 10% more variability than metrics derived from coarser-resolution datasets. Increased correlation in the latter half of the study period was partially attributable to increased streamflow variability in recent decades as well as to improved climate data quality across the interior mountain watersheds.

1. Introduction

Water resources of the western United States depend upon winter snowpack as a natural reservoir and are sensitive to an array of atmospheric drivers (McCabe and Dettinger 2002; Clark 2010). Large interannual variability in winter precipitation across the western United States, where the majority of precipitation falls during the winter months, coupled with increasing water demand make the region susceptible to water scarcity (Wilhite et al. 2007). Widespread observations across the northwestern United States over the past 60 years find declines in annual streamflow of the bottom quartile of years (Luce and Holden 2009) and an advancement in the timing of snowmelt-dominated streamflow (e.g., Clark 2010; Stewart et al. 2005). These changes may be partially attributable to changes in precipitation (e.g., Luce et al. 2013); however, the influence of other climate factors, most notably temperature, has likely played a role in changes in volumetric runoff (e.g., Vano et al. 2012) and, in particular, runoff timing (Hidalgo et al. 2009). Consequentially, stationarity from the perspective of the influence of climatological drivers of streamflow as well as the management of water resources may be questioned in particularly sensitive natural or managed systems (e.g., Milly et al. 2008).

A variety of ways exist to better understand the response of a hydrologic system to climate forcings. Physically based hydrologic models provide a preferred tool for exploration; however, they contain their own limits in both computation ability and model assumptions. Drought indices provide an alternative means to integrate landscape-scale climatic forcing to the relativized difference

Corresponding author address: John Abatzoglou, Department of Geography, University of Idaho, 875 Perimeter Dr., MS 3021, Moscow, ID 83844-3021.
E-mail: jabatzoglou@uidaho.edu

DOI: 10.1175/JHM-D-13-0167.1

between water supply and demand across a variety of scales (Redmond 2002; Keyantash and Dracup 2002; Mishra and Singh 2010) and are used operationally to monitor and forecast drought and water resources. The primary climate variables used to compute drought indices and surface water supply are precipitation, potential evapotranspiration (PET), and temperature, with each drought index providing different weighting across these variables in addition to the sequence and duration associated with them (e.g., Heim 2002). Previous studies have evaluated the utility of drought indices to track measured hydrological, agricultural, and ecological indicators (e.g., Ellis et al. 2010; van der Schrier et al. 2011; McEvoy et al. 2012; Vicente-Serrano et al. 2012a). However, the utility of drought indices has not been fully vetted in regions of complex terrain or at spatial scales of individual watersheds. Redmond (2002) noted the need to better evaluate drought indices at smaller scales (e.g., individual watershed) where impacts are manifested and, in particular, across the western United States, which is characterized by complex energy and moisture gradients.

The scientific community faces a significant challenge to produce timely and more comprehensive assessments of the utility of drought indices given increasing vulnerability of water resources associated with multiple stressors (Wilhite et al. 2007). Changes in climate may alter the ability of various drought indices to track meaningful hydrologic metrics, particularly in regions that observe a significant shift in the phase of precipitation or change in the influence of PET on the surface water budget (Berghuijs et al. 2014). This is particularly true as many drought indices are calibrated to their historical record and often use overly simple PET approximations, although more complex PET methods have resulted in similar Palmer drought severity index (PDSI) values (Dai 2010, 2011). Likewise, Oudin et al. (2005) found that complex methods for estimating PET may not yield additional skill in rainfall–runoff modeling. Large differences in PET calculated using the Thornthwaite and Penman–Monteith methods have been noted across the western United States (e.g., van der Schrier et al. 2011) with repercussions for estimating climatic water balance (Crimmins et al. 2011). Likewise, drought indices that incorporate PET may yield divergent trajectories in a warming climate between temperature- and energy-based estimates of PET (Donohue et al. 2010; Sheffield et al. 2012), further emphasizing the need to quantify the effectiveness of different flavors of a single drought index.

This study examines water-year streamflow from 21 unregulated long-term stream gauges from 1948 to 2012 across the Pacific Northwest (PNW) of the United States given the importance of hydrologic drought on surface water availability in the region. These gauges are considered given their high-quality observations of natural flow and their distribution across heterogeneous watersheds with respect to climate and watershed hypsometry. Likewise, the PNW has observed significant changes in climate over the study period, including increases in temperature, growing season PET, and the portion of precipitation falling as rain (e.g., Mote 2003; Abatzoglou 2011; Abatzoglou et al. 2014). These changes have been further manifested through observed decreases in 1 April snow water equivalent (SWE; Mote et al. 2005), earlier snowmelt runoff (Stewart et al. 2005), and decreases in mean and the lower quartile of water-year streamflow (Luce and Holden 2009; Clark 2010).

Using a suite of drought indices and water-year streamflow at these 21 sites, we aim to understand: 1) how relationships between streamflow and different drought indices vary across watersheds; 2) whether there are differences in correlations between streamflow and drought indices that use only precipitation and those that incorporate estimates of PET, as well the degree to which PET estimated using the Thornthwaite and Penman–Monteith methods influences results; 3) what value is added by incorporating high-resolution climate surfaces; and 4) whether relationships have changed over the period of record.

2. Data and methodology

Daily streamflow records from 21 unregulated stream gauges across the PNW with high-quality records both in duration and completeness were acquired from the U.S. Geological Survey (USGS; Fig. 1 and Table 1). Gauges were chosen from reference gauges according to the Geospatial Attributes of Gauges for Evaluating Streamflow, version II (GAGES II; Falcone 2014) dataset, had complete data for 64 water years (1948/49–2011/12), and were used by Luce and Holden (2009) to examine long-term variability in streamflow records of the PNW. Watersheds within the PNW include snow-dominated, rain-dominated, and transient watersheds (Clark 2010) spanning various elevations, total annual precipitation, and precipitation seasonality. Water-year streamflow was defined by the summation of observed daily streamflow from 1 October to 30 September. A standardized runoff index (SRI; Shukla and Wood 2008; Elsner 2010) that represents a z score for cumulative water-year runoff was calculated for each stream gauge through a normal inverse cumulative distribution that applies the nonparametric kernel density bandwidth estimator of Botev et al. (2010). This nonparametric transformation overcomes some of the documented limitations of using a single prescribed probability

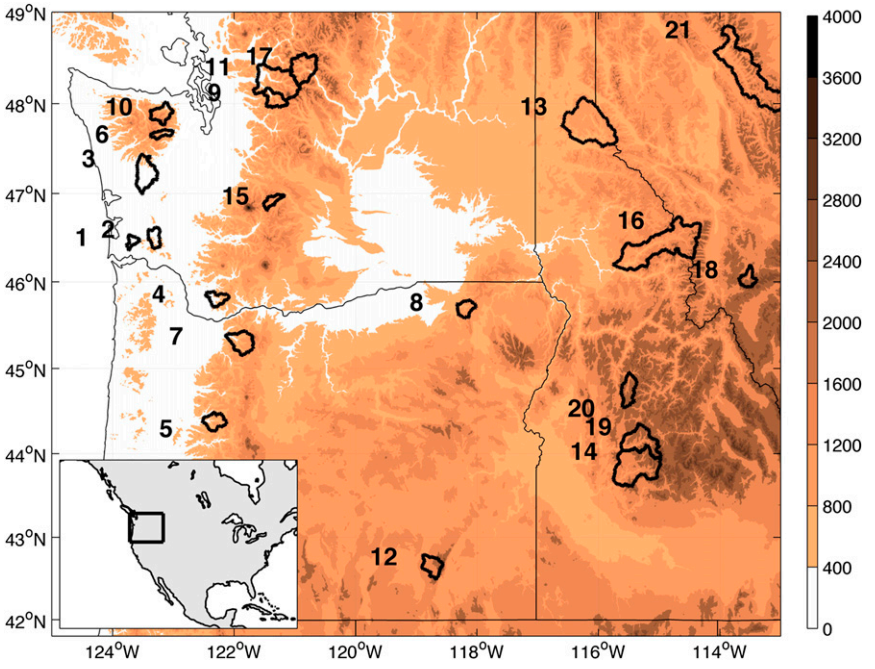

FIG. 1. Locations of the 21 watersheds and elevation (m). Numbers refer to the watershed ranking shown in Table 1. The inset shows the location of the study area.

distribution for hydrologic indices (e.g., Vicente-Serrano et al. 2012b).

A set of four established drought indices were selected: 1) PDSI (Palmer 1965), 2) standardized precipitation index (SPI; McKee et al. 1993), 3) standardized precipitation evapotranspiration index (SPEI; Vicente-Serrano et al. 2010), and 4) a normalized modified Thornthwaite water balance runoff (WBR) model (Willmott et al. 1985;

TABLE 1. List of stations ranked according to the ratio of 1 April SWE to October–March precipitation (SWE:*P*). Also shown are the annual-average precipitation and the percent of annual precipitation occurring in the months of April–June (AMJ). The last two columns indicate the metric that had the highest correlation with SRI.

Station rank	Station ID	SWE:*P* ratio	Annual-avg *P* (mm)	P_{AMJ} (%)	Best metric	r^2
1	12010000	0	3315	15	WBR-10, Jul	0.92
2	12020000	0	2077	16	WBR-6, Mar	0.88
3	12035000	0	3050	14	WBR-11, Jun	0.94
4	14222500	0.01	3551	16	Thorn-SPEI-11, Jun	0.95
5	14185000	0.01	2191	19	SPI-10, May	0.94
6	12054000	0.14	2369	14	SPI-8, Apr	0.84
7	14137000	0.14	2340	19	PM-SPEI-10, May	0.89
8	14020000	0.18	1004	23	WBR-8, May	0.88
9	12186000	0.30	3138	17	PM-SPEI-7, Apr	0.88
10	12048000	0.31	1817	14	WBR-7, Apr	0.84
11	12189500	0.34	2747	16	WBR-9, Apr	0.92
12	10396000	0.39	661	27	PM-SPEI-9, Apr	0.79
13	12413000	0.43	1204	23	SPI-8, Apr	0.88
14	13185000	0.53	903	21	WBR-7, Apr	0.91
15	12488500	0.53	1665	15	PM-SPEI-7, Mar	0.88
16	13337000	0.60	1303	24	PM-SPEI-9, May	0.88
17	12451000	0.63	1726	13	WBR-8, Apr	0.89
18	12332000	0.65	702	36	PM-SPEI-6, Mar	0.56
19	13235000	0.67	991	21	PM-SPEI-7, Apr	0.91
20	13313000	0.75	1042	21	PM-SPEI-7, Apr	0.81
21	12358500	0.80	1285	25	Thorn-SPEI-11, Jun	0.87
Pacific Northwest scale					WBR-12, Sep	0.86

TABLE 2. List of acronyms for indices and associated time scales used in this study. Asterisks denote a single time scale.

Acronyms	Meaning	Timescales used
SPI	Standardized precipitation index	1–12 months
CDD-SPI	Climate division data standardized precipitation index	1–12 months
PM-SPEI	Penman–Monteith standardized precipitation evapotranspiration index	1–12 months
Thorn-SPEI	Thornthwaite standardized precipitation evapotranspiration index	1–12 months
SRI	Standardized runoff index	Water year
Thorn-PDSI	Thornthwaite Palmer drought severity index	*
PM-PDSI	Penman–Monteith Palmer drought severity index	*
CDD-PDSI	Climate division data Palmer drought severity index	*
WBR	Water balance runoff	1–12 months

Dobrowski et al. 2013). PDSI is based on soil water balance equations by considering the magnitude and sequence of precipitation and PET in addition to soil available water holding capacity (AWC). However, PDSI's well-documented limitations include its calibration parameters, which can be adjusted for through the use of the self-calibrated PDSI (SC-PDSI; Wells et al. 2004), and its failure to discriminate precipitation phase, which potentially limits its applicability in snow-dominated and transient watersheds that define the hydrology of the western United States (e.g., Alley 1984; Vicente-Serrano et al. 2010). The use of SC-PDSI over PDSI failed to reflect significant differences, and hence, we constrained our focus to the more widely used PDSI. Both the SPI and SPEI are temporally flexible and applicable to different types of drought. The SPI assumes that precipitation variability is much greater than PET variability and only accounts for precipitation, whereas SPEI accounts for a simplified moisture balance by using precipitation minus PET. Similar to PDSI, neither account for snow dynamics as it pertains to moisture timing. Both the SPI and SPEI were calculated monthly by considering the cumulative precipitation or cumulative precipitation minus cumulative PET, respectively, over the past number of months (1–12 months were considered here) relative to historical conditions, which was then transformed into a near-Gaussian distribution using the nonparametric approached described previously for SRI. A modified Thornthwaite water balance model [Willmott et al. (1985); McCabe and Wolock (2011a); updated by Dobrowski et al. (2013)] that incorporates monthly temperature, precipitation, and PET was run at monthly time steps with AWC to model monthly runoff, defined as the excess precipitation or snowmelt not used by PET or to recharge soils. For compatibility with other drought indices, cumulative runoff from the WBR model of the previous 1–12 months was transformed to a near-Gaussian distribution using the aforementioned nonparametric method. While SPI and SPEI have prescribed time scales and do not incorporate information antecedent to the time period of interest, PDSI and WBR can entrain memory of conditions prior to the time period of interest.

Two of the drought indices and the WBR model consider evaporative demand; however, PET can be estimated several ways using climatological data, including a simple temperature-based approach via the Thornthwaite method (hereafter Thorn-PET; Thornthwaite 1948) and an energy-balance approach via the Penman–Monteith method (hereafter PM-PET; Allen et al. 1998). Thorn-PET is a widely used empirical transformation that only requires monthly-mean temperature and latitude. This contrasts with PM-PET, which is an energy-balance approach requiring temperature, latitude, elevation, wind speed, radiation, albedo, and vapor pressure deficit. We modified PM-PET to account for unrealistic variations in the surface energy budget when snow cover exists or prior to the onset of the growing season, when temperature is a limiting factor (Jarvis 1976). This was done using an empirical hyperbolic tangent function of Dai (2008) that accounts for precipitation phase where PET is set to zero for monthly-mean temperatures below −3°C and unmodified for temperatures exceeding 5°C.

Data required to calculate drought indices were derived from three primary sources. First, monthly precipitation and maximum, minimum, and dewpoint temperature are acquired from the Parameter–Elevation Regressions on Independent Slopes Model (PRISM) at 800-m resolution and aggregated to 4-km resolution (Daly et al. 2008) from 1895 to 2010. As the 800-m PRISM time series dataset was available through 2010, data for 2011–12 were estimated by applying monthly anomalies from phase 2 of the North American Land Data Assimilation System (NLDAS-2; Mitchell et al. 2004) to monthly averages from PRISM over a common time period (1981–2010). Because of the lack of long-term observations of downward solar radiation and 10-m wind speed, we used climatologically aided interpolation (e.g., Willmott and Robeson 1995), using monthly anomalies of wind speed and downward shortwave radiation from the National Centers for Environmental Prediction–National Center

for Atmospheric Research (NCEP–NCAR) reanalysis (1948–2012) superimposed with climatological monthly downward surface shortwave radiation and 10-m wind speed from NLDAS-2. The resultant monthly fields were bilinearly interpolated to the 4-km PRISM grid. Spatial AWC data for the top 250 cm of soil was retrieved from the State Soil Geographic (STATSGO; www.soilinfo.psu.edu/index.cgi) database and aggregated to match the 4-km scale of the climate data. Pixels fully contained within the contributing upstream drainage basin for each stream gauge site were aggregated to form a single time series for each basin. We compared correlations calculated using the aforementioned datasets to the coarser, but more readily available, PDSI and monthly precipitation data from the U.S. climate division data (CDD) from the National Climatic Data Center (NCDC). Monthly SPI was calculated identical as performed for higher-resolution data. We extracted SPI and PDSI data from the nearest climate division to each watershed.

Three static attributes were characterized for each basin: 1) the fraction of accumulated precipitation P remaining as SWE on 1 April, hereafter referred to as SWE:P ratio; 2) the ratio of water-year P during spring (April–June, using 1948–2012 climatology), hereafter referred to as P_{AMJ}; and 3) total water-year P (1948–2012 climatology; Table 1). The SWE:P ratio was computed as the ratio of 1 April SWE to October–March P using SWE and P from the Variable Infiltration Capacity (VIC) model (Liang et al. 1996) at $1/8^{\circ}$ spatial resolution from 1981 to 2010. Following Elsner et al. (2010), we classified basins as rain-dominated basins having SWE:P less than 0.3, transient basins as 0.3–0.6, and snow-dominated basins exceeding 0.6. These characteristics exhibit relationships with one another arising from climatological factors, with the wettest watersheds located in the lower elevations west of the Cascades having the lowest P_{AMJ} and lowest SWE:P.

Pearson's correlation coefficients were calculated between water-year SRI at each stream gauge site and 1) 1–12-month SPI ending January–September, 2) 1–12-month SPEI ending January–September using both Thorn-PET and PM-PET, 3) January–September PDSI using both Thorn-PET and PM-PET, and 4) normalized cumulative 1–12-month WBR ending January–September from PRISM data using the period of record 1948–2012 (Table 2). Correlations were also calculated for SPI and PDSI from climate division data over the same period. We also calculated Nash–Sutcliffe efficiency coefficients for the same relationships, but hereafter we show only correlation coefficients, as Nash–Sutcliffe statistics did not provide additional information beyond that obtained with correlation analysis given the normal distribution of the data. We further evaluated relationships within the period of record using a moving 31-yr correlation window. A stationary linear relationship assumes no significant change in the relationship through time. It is plausible that changes in correlations through time are strictly a function of changes in streamflow variability through time, but otherwise they have a stationary relationship. We evaluated the null model of stationary relationships between drought indices and streamflow by creating 10 000 time series y that were a linear function of SRI, plus random noise ε following the simple model (e.g., Neter et al. 1996):

$$y = \mathrm{SRI} + \varepsilon$$
$$\varepsilon = N(0, \sigma).$$

The error term is a random number from a normal distribution with mean 0 and standard deviation σ. For SPI, SPEI, and WBR, we vary σ to obtain correlations between drought indices and streamflow approximate to those derived in the observational record from 1948 to 2012. Moving correlation windows of 31-yr lengths were calculated from these synthetic datasets with the central 95% of values from these 10 000 simulations used to construct an envelope of potential values under the null model. Linear least squares trends of the observed 31-yr moving correlations and those resulting from the bootstrapped data were computed.

3. Results

Mean squared correlation coefficients $\overline{r^2}$ between SRI and drought indices of SPI, SPEI, PDSI, and WBR for the months of January–September averaged across all 21 basins are shown in Fig. 2. The results illustrate that $\overline{r^2}$ from SPI and SPEI were largely comparable with the strongest correlations obtained using 6–10-month time scales ending in April–June encompassing the vast majority of water-year P and explaining over 80% of the variance in SRI. A slight degradation of $\overline{r^2}$ for longer time scales of SPI and SPEI extending into July–September suggests that summer P and PET had negligible relationships to SRI across the study area, consistent with the nominal summer P in the region being lost either to evapotranspiration (ET) or soil moisture recharge. The PDSI $\overline{r^2}$ peaked in April (Figs. 2d,g) but explained only 60% of the variance in SRI. Normalized 9–12-month WBR ending in June–September exhibited the strongest $\overline{r^2}$ to streamflow at the regional scale, explaining 86% of the variance in SRI (Fig. 2c).

Correlations between drought indices and streamflow varied by watershed, with $r^2 > 0.9$ in some basins (Table 1). In general, correlations were higher in the wettest

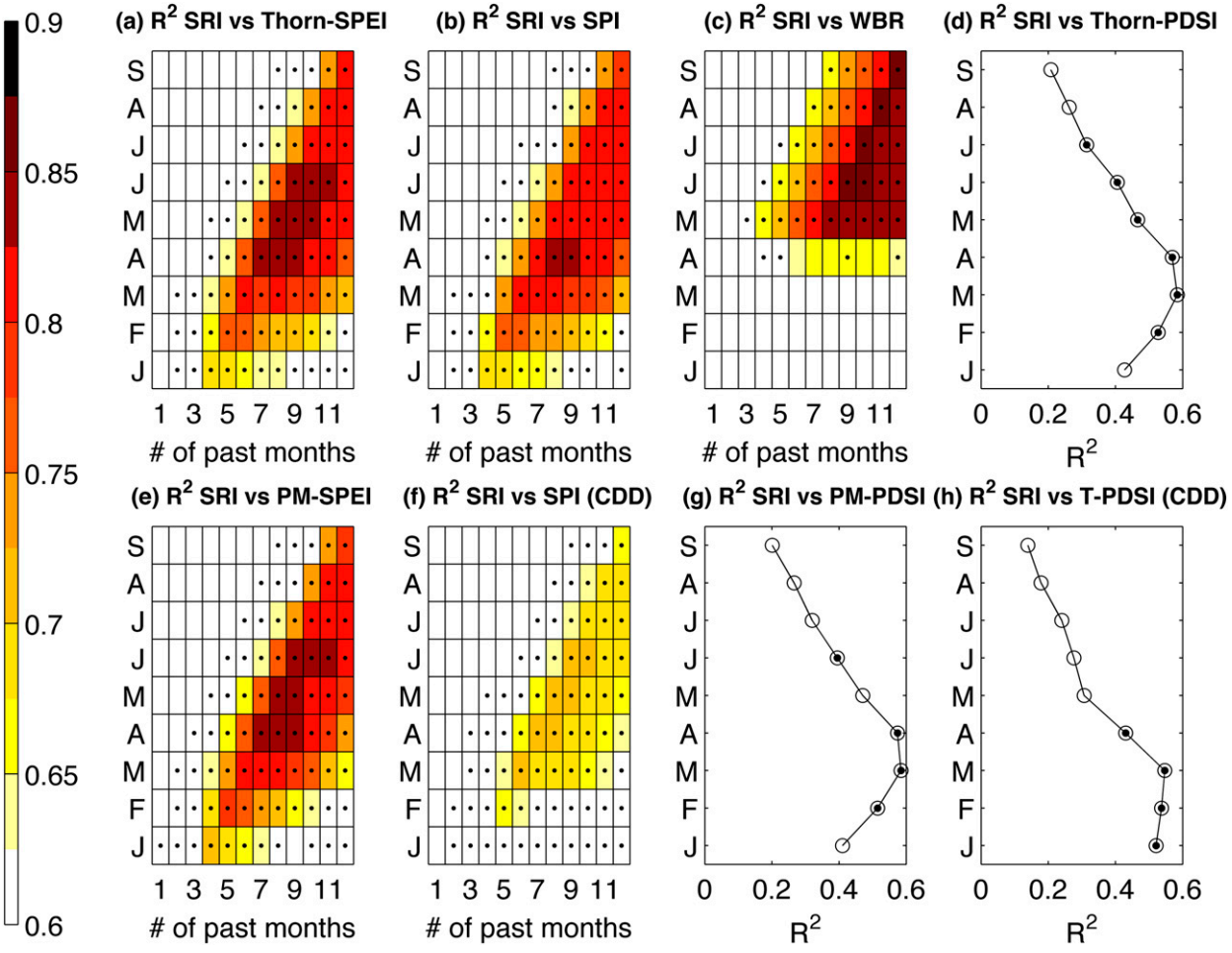

FIG. 2. Values of $\overline{r^2}$ averaged over the 21 stations between water-year standardized runoff and values of metric duration for (a) Thorn-SPEI, (b) SPI from PRISM data, (c) WBR, (d) Thorn-PDSI, (e) PM-SPEI, (f) SPI from CDD, (g) PM-PDSI, and (h) Thorn-PDSI from CDD. In (a),(b),(c),(e), and (f) squared correlation coefficients are shown in a matrix where the ending month is on the *y* axis (January–September) and the number of months is on the *x* axis. Black dots indicate that all watersheds display significant correlations at the two-sided 95% level according to a Monte Carlo random phase test.

basins that received more of their precipitation during the winter months, particularly the wettest rain-dominated basins. While total precipitation amount, P seasonality, and SWE:P ratios are inherently coupled across the basins of interest, partial correlation analyses show that only P_{AMJ} was statistically significant. The maximum variance explained using SPI, SPEI, and WBR all were strongly negatively correlated to P_{AMJ} (r from -0.75 to -0.7, $p < 0.01$; Fig. 3a), indicating that some of the heterogeneity in correlations across the 21 basins is due to the seasonality of precipitation. Conversely, the maximum explained variance for Penman–Monteith PDSI (PM-PDSI) and Thornthwaite PDSI (Thorn-PDSI) was strongly negatively correlated with total precipitation ($r = -0.62$, $p < 0.01$; Fig. 3b), with PDSI having less utility in the wettest basins. Results from Table 1 suggest that 12-month WBR ending in September was the most highly correlated metric at the regional scale and at the watershed-type scale. At the individual station level, SPEI explained the most variance in 10 of 21 stations (8 of the 10 using PM method), with WBR and SPI explaining the most variance at 8 and 3 stations, respectively.

Comparable calculations of $\overline{r^2}$ using SPI and PDSI using data from NCDC climate divisions are shown in Figs. 2f and 2h. The relationships obtained using climate division data were qualitatively similar to those obtained using the higher-resolution PRISM climate surfaces. Generally, SPI calculated from PRISM explained 10%–15% more variance in streamflow than SPI calculated from climate division data, although the differences varied widely across the study area (Fig. 4a). Correlations between SRI and PDSI were also stronger using PRISM data compared to divisional data (Fig. 4b).

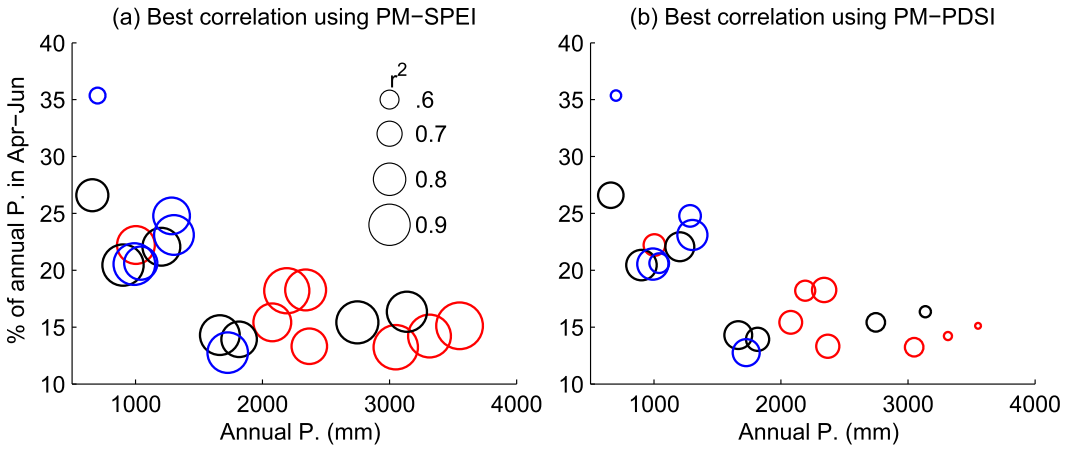

FIG. 3. Relationships between maximum squared correlation at each of the 21 watersheds for (a) PM-SPEI and (b) PM-PDSI for the three basin characteristics examined, including total annual precipitation (x axis), fraction of precipitation that falls in April–June (y axis), and the SWE:P ratio (colors). Red, black, and blue indicate rain, transient, and snow basins, respectively, using the criteria of Elsner et al. (2010). The sizes of the circles correspond to r^2.

Subtle differences between Penman–Monteith SPEI (PM-SPEI) and SPI ($\overline{r^2_{\text{PM-SPEI}}} - \overline{r^2_{\text{SPI}}}$) were identified for time scales that explained the maximum amount of interannual variability in streamflow (Fig. 5a). The SPEI provided slightly more explanatory power over the SPI from March to September for time scales of 9 months and less. The additional explained variance ($\overline{r^2_{\text{PM-SPEI}}} - \overline{r^2_{\text{SPI}}}$) for 6–9-month time scales ending in May–September was strongly correlated to P_{AMJ} ($r = 0.8, p < 0.01$) and total precipitation ($r = -0.78, p < 0.01$). The maximum explained variance for PM-SPEI was on average 1.5% more than the maximum explained variance by SPI (Fig. 4c). By contrast, 9–12-month SPI ending in late winter and early spring exhibited a stronger correlation to SRI than SPEI.

Differences in correlations to SRI using PM-SPEI and Thornthwaite SPEI (Thorn-SPEI) ($\overline{r^2_{\text{PM-SPEI}}} - \overline{r^2_{\text{Thorn-SPEI}}}$) shown in Fig. 5b were relatively small, largely confirming the results of Dai (2011). Our results suggest that PM-SPEI outperformed Thorn-SPEI for shorter time scales (1–6 months) encompassing March–September in basins with $P_{\text{AMJ}} > 20\%$. The difference in maximum explained variance between PM-SPEI and Thorn-SPEI was rather small for most basins and qualitatively similar to differences between PM-SPEI and SPI (Fig. 4d). PM-SPEI explained 4%–5% more variance than Thorn-SPEI for the two driest watersheds; otherwise, no relationships were seen to basin characteristics examined. Similar results were found between PM-PDSI and Thorn-PDSI (not shown).

Time-varying $\overline{r^2}$ between SRI and 9-month SPI (SPI-9), SPEI (SPEI-9), and WBR (WBR-9), as well as PDSI, are shown in Fig. 6. Our results suggest correlations increased over time. Averaged across all stations, a linear trend exceeding 8% explained variance over the three decades was observed for 9-month SPI and SPEI, with 6% and 12% increases observed for May 9-month WBR and PDSI, respectively. Changes were also present in time-varying $\overline{r^2}$ using climate division data, although again with lesser correlations compared to those using PRISM data.

Simulated changes in $\overline{r^2}$ using the null model depict a similar increase over the period of record, although peaking around a 31-yr period centered during the late 1980s and returning to correlations seen earlier in the record for more recent time periods. Figure 7a shows $\overline{r^2_{\text{SPI-9,null}}}$ and the 95% confidence interval of the 21-station mean along with the observed $\overline{r^2_{\text{SPI-9}}}$ estimated from PRISM. Increases in $\overline{r^2_{\text{SPI-9}}}$ using a moving 31-yr period from the mid-1960s to mid-1980s are generally consistent with the null model, indicating that such changes may occur with an otherwise stationary relationship and more variable streamflow record. However, a significant divergence for SPI following the late 1980s is inconsistent with stationary relationships. These discrepancies were also observed for SPEI and WBR when compared to the null model but were not apparent for PDSI. Conversely, $\overline{r^2}$ computed with climate division data appears to be fully explained by changes in streamflow variability and within the 95% confidence bounds of the null model (Fig. 7b).

On an individual station basis, the discrepancy between observed moving 31-yr $\overline{r^2_{\text{SPI-9}}}$ and $\overline{r^2_{\text{SPI-9,null}}}$ using PRISM data for the last 10 years of record was most pronounced over the interior PNW, whereas many of the watersheds in western Washington exhibited insignificant differences (Fig. 7c). Statistically significant differences for the last decade of observed moving 31-yr correlations

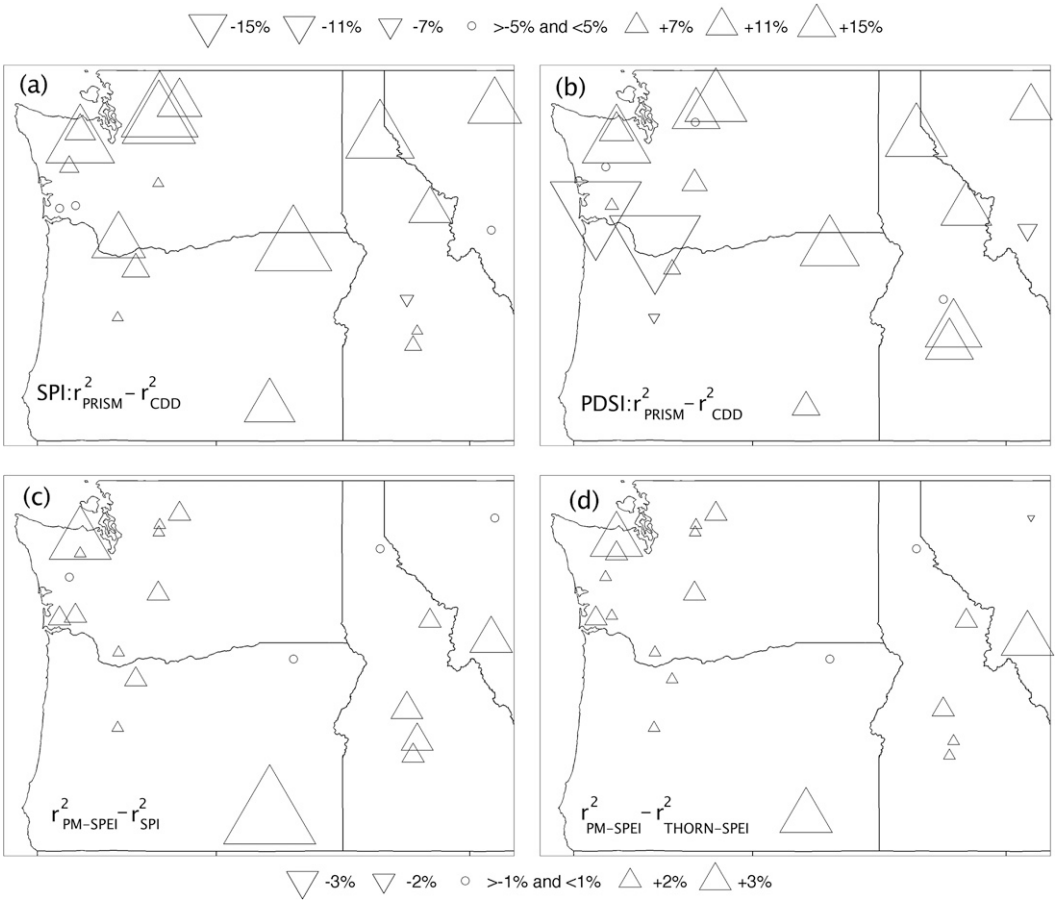

FIG. 4. Differences of max variance explained in standardized runoff (%) between (a) SPI from PRISM minus SPI from CDD, (b) PM-PDSI from PRISM minus PDSI from CDD, (c) PM-SPEI minus SPI (both from PRISM), and (d) PM-SPEI minus Thorn-SPEI (both from PRISM). The legend at the top refers to (a) and (b), while the legend at the bottom refers to (c) and (d).

versus the null model (estimated from the resampled null model data) were observed for seven of the nine basins east of the Cascades, with some basins having an addition 5% of explained variance relative to the null model. Conversely, observed changes in $\overline{r^2_{SPI\text{-}9}}$ computed using climate division data were consistent with the null model, with only one station having statistically significant differences (Fig. 7d). A possible cause of strengthened correlations in recent decades is the assimilation of climate data from high-elevation observations such as Snowpack Telemetry (SNOTEL) in PRISM, particularly in data-sparse regions with complex *P* patterns. Prior to the 1980s, the majority of information used in PRISM came from primarily lower-elevation National Weather Service Cooperative Observer Program (COOP) stations. Conversely, climate division data are solely derived from COOP data and may be less representative of interannual *P* variability in mountainous regions and in regions with greater spatial heterogeneity in *P* variability. The lack of comparable increases in observed $\overline{r^2_{SPI\text{-}9}}$ relative to the null model over watersheds in western Washington and Oregon is hypothesized to be a consequence of stronger spatial coherence between numerous COOP stations located in the region and precipitation falling in nearby watersheds (e.g., Luce et al. 2013).

4. Discussion and conclusions

Following Redmond (2002), we find that no single metric was universally optimal for tracking streamflow at the watershed scale in the PNW, but PM-SPEI and WBR generally had the highest predictive power while PDSI had the lowest predictive power. WBR was the best metric at the regional scale, which might not be surprising given its intended purpose rather than more generalized water supply–demand relationships of SPEI. Correlations of streamflow to SPEI and WBR

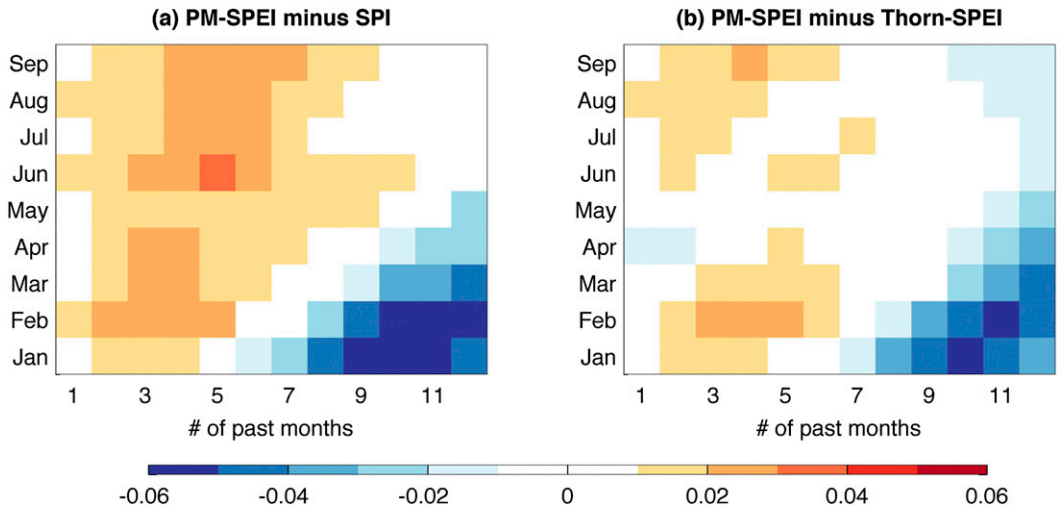

FIG. 5. Differences in 21-station $\overline{r^2}$ between (a) PM-SPEI and SPI and (b) PM-SPEI and Thorn-SPEI.

increased from winter through spring, similar to increasing skill of operational streamflow forecasts that incorporate late winter precipitation (e.g., Pagano et al. 2009). The generally lower correlations using PDSI are also similar to those seen in prior drought index comparisons (e.g., Vicente-Serrano et al. 2012a) and likely reflect PDSI's original design for monitoring soil moisture rather than runoff (Palmer 1965).

Drought indices that included atmospheric demand performed better than simpler indices, reinforcing results by McEvoy et al. (2012) and Vicente-Serrano et al. (2012a). The improved correlations with streamflow for indices having more sophisticated physical parameterization schemes that account for atmospheric demand (SPEI, WBR), while small, nonetheless suggest that such approaches are advantageous and may become more important in a changing climate (Barnett et al. 2005). An average of 1.5% additional variance was explained using SPEI over SPI, with more significant increases in drier regions that receive appreciable spring *P*. The influence of PET should be most pronounced in these watersheds given that the growing season *P* can be utilized by vegetation through ET and may contribute less to runoff, thereby contributing to interannual variability. Conversely, asynchronous seasonality of *P* and PET across much of the PNW results in interannual variations in volumetric streamflow being nearly entirely driven by interannual *P* variability (e.g., McCabe and Wolock 2011b). A broader analysis by Vicente-Serrano et al. (2012a) found that SPEI provided more explained variance than SPI in regions that receive more *P* during the growing season, thereby making atmospheric demand a more important contributor to the water balance.

Heterogeneous correlations across the region were partially explained by the set of basin characteristics analyzed, most notably *P* seasonality. Drought indices were more strongly correlated in rain-dominated basins west of the Cascades where the vast majority of precipitation falls from October to March. However, we note that climate data might be more representative across these watersheds because of the proximity of weather stations and the broader homogeneity in seasonal *P* on the windward side of the Cascades. Other factors that we did not consider, but that may be influential in watershed sensitivity to climate variability, include upstream geology (and base flow contributions) and vegetation that have been hypothesized to alter interbasin climate–streamflow relationships (e.g., López-Moreno et al. 2013). Confounding factors of representativeness of climate estimates and basin characteristics restrict us from fully resolving interbasin differences.

Minor differences in correlation to streamflow were found using the Penman–Monteith method compared to the Thornthwaite approach for estimating PET. These results appear to be in agreement with Dai (2011) and van der Schrier et al. (2011). We show that PM-SPEI explained significantly more streamflow variance when considering *P* and PET confined to the growing season when ET becomes more important to the water balance (i.e., excluding precipitation occurring prior to March). Furthermore, we found that PM-SPEI explained 4%–5% more variance than Thorn-SPEI in the driest two watersheds that received more than 25% of their annual *P* from April to June. Conceptually, the different approaches for estimating PET would be more important where ET plays a larger role on the local water balance, as opposed to

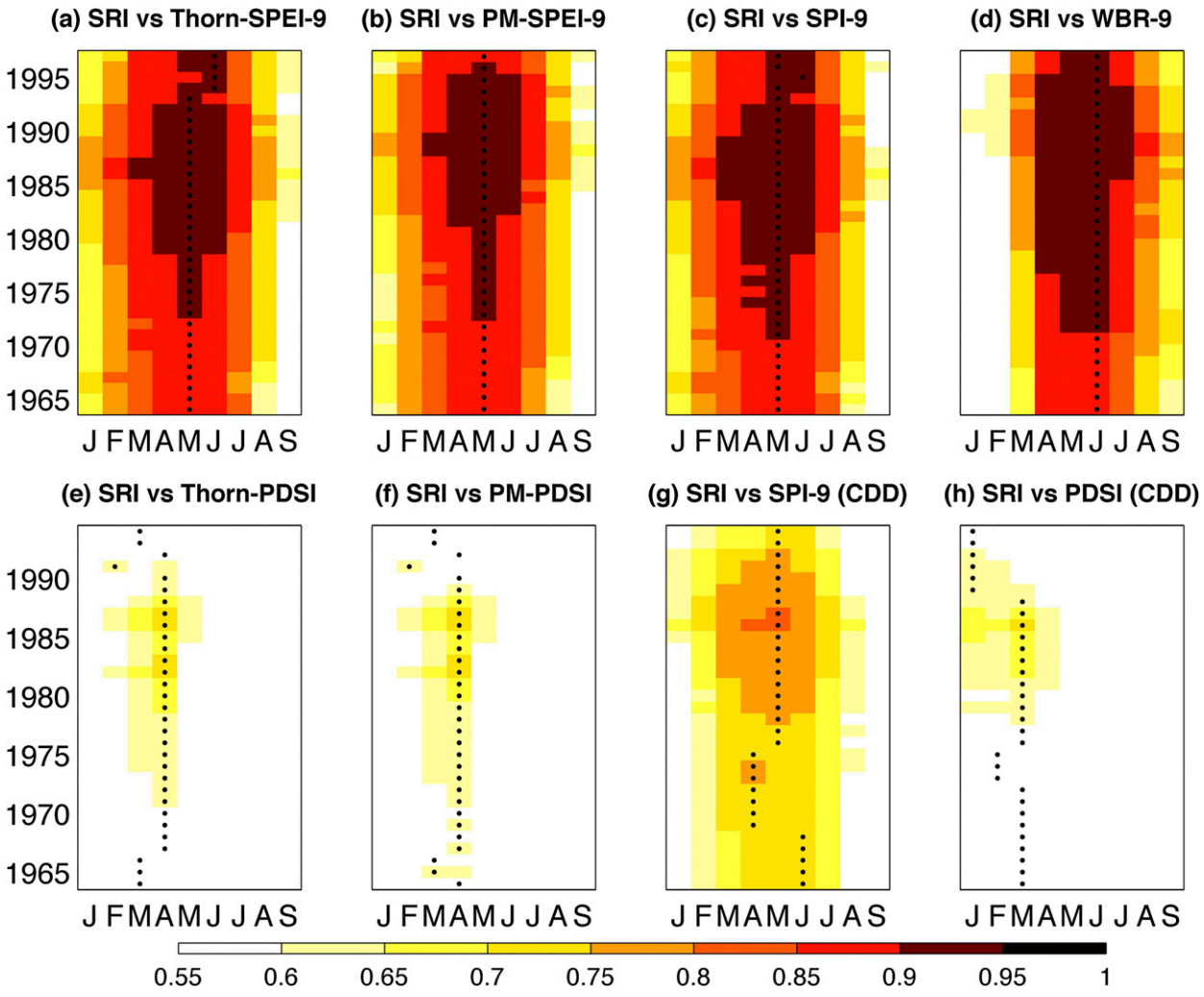

FIG. 6. Mean 31-yr moving-average $\overline{r^2}$ across all stations between SRI and (a) Thorn-SPEI-9, (b) PM-SPEI-9, (c) SPI-9, (d) WBR-9, (e) Thorn-PDSI, (f) PM-PDSI, (g) SPI-9 from CDD, and (h) PDSI from CDD. In (a)–(f), data are from PRISM. Years on the *y* axis are at the middle of the 31-yr average. Black dots indicate month of peak correlation for the given year.

watersheds with asymmetric seasonality of P and PET. Interannual variability in PET estimated using both approaches were strongly correlated for each station (r^2 =0.26–0.77), with the lowest correlations west of the Cascades. However, the range of interannual PET variability using the Penman–Monteith method was nearly 3 times that using the Thornthwaite method. A larger-scale analysis of PM-SPEI and Thorn-SPEI across watersheds where ET plays a larger role in the local water balance may further our cursory analysis. Likewise, limitations in the resolution and accuracy of forcing data restrict a definitive analysis of the merits of different means of approximating PET. Advances in topoclimatic modeling that account for finescale structure in radiation, wind speed, temperature, and vapor pressure deficit (e.g., Holden et al. 2011) may help better understand such relationships.

Increased correlation between streamflow and drought indices over the period of record was seen across most sites. Some of this increase is directly attributable to increased streamflow variability, as documented by Pagano and Garen (2005) and simulated by our null model. However, additional increases in correlation unexplained by changes in streamflow alone reveal nonstationarity in climate–streamflow relationships for certain watersheds. The discrepancy between the null model and observed correlations using the PRISM dataset were greatest at interior PNW gauges, where COOP observations are sparse and potentially less representative of precipitation received in mountain watersheds. We hypothesize that increasing data quality

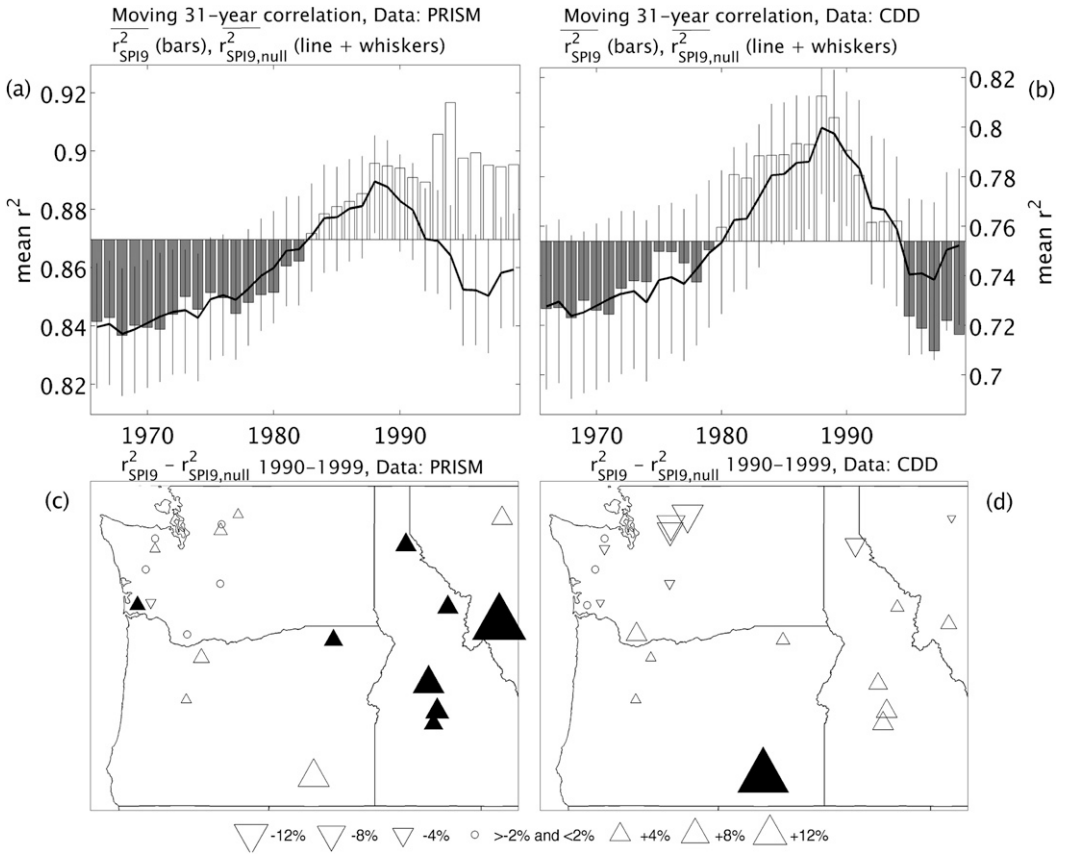

FIG. 7. Mean 31-yr moving $\overline{r^2}$ between SRI and 9-month SPI ending in May (bars) and the mean and 95% confidence interval of the null model (line and whiskers, respectively) averaged across 21 stations for (a) PRISM data and (b) CDD. Also shown is the difference of 31-yr moving $\overline{r^2}$ for the last 10 years of record (centered on 1990–99) between observations and (c) the null model for PRISM and (d) CDD, respectively. Filled symbols denote statistical significance at the 95% confidence level.

and spatial representation of climate datasets in complex terrain (e.g., inclusion of SNOTEL observations in PRISM) likely account for this disparity. Differences in intraseasonal-to-interannual *P* variability across complex terrain (e.g., Dettinger et al. 2004; Siler et al. 2013; Luce et al. 2013) during the cool season have been partially explained by variations in midlatitude flow and its influence on orographic precipitation enhancement. Prior to the incorporation of SNOTEL observations, *P* estimates in mountainous watersheds were estimated using observations from lower elevations that are unable to account for time-varying orographic enhancement. Whereas we postulate that improved data quality is associated with observed increases in correlation between drought indices and streamflow over the period of record, other mechanisms may contribute, including: 1) a decreasing fraction of *P* falling as snow, altering snow hydrology and increasing the utility of drought indices that only account for liquid precipitation; 2) changes in the relative influence of water demand on runoff, particularly with an earlier onset of spring across much of the region observed through advances in snowmelt timing and phenology (e.g., Cayan et al. 2001); 3) changes in *P* seasonality (e.g., Pagano and Garen 2005); and 4) changes in vegetative cover and ET of the upstream watershed.

Complex topography and its impacts on patterns of moisture and energy have emerged as important topics in mountain hydrology. The ability to monitor hydroclimatic variables at local and regional scales is urgently needed by resource managers, land owners, planners, and others across the western United States where decisions may often be made based on incomplete or insufficient data. We demonstrate that higher-spatial-resolution gridded climate surfaces from PRISM provide added value beyond more commonly used NCDC climate division data, explaining up to 80% of the interannual variability in water-year cumulative streamflow in the PNW. The monthly datasets used in this study are finely resolved relative to many global and regional datasets, but nonetheless cannot fully capture finescale characteristics in mountain

watersheds, partially because of an improved but still insufficient observational network. The divergence in observed correlation from PRISM data and that simulated by the null model demonstrates realized gains via assimilating precipitation data high-elevation observations. However, these results also suggest that analyses of longer-term hydroclimate variability and change in mountainous regions across the interior PNW using climate datasets may be less reliable.

Acknowledgments. We are appreciative of constructive feedback from Thomas Pagano and two anonymous reviewers who helped improve the quality of this manuscript. This research was funded by the U.S. Department of Agriculture National Institute for Food and Agriculture Award 2011-68002-30191, U.S. Department of the Interior via the Northwest Climate Science Center Awards G10AC00702 and G12AC20495, and NOAA Climate Impacts Research Consortium (RISA for the Northwest) under Award NA10OAR4310218.

REFERENCES

Abatzoglou, J. T., 2011: Influence of the PNA on declining mountain snowpack in the western United States. *Int. J. Climatol.,* **31,** 1135–1142, doi:10.1002/joc.2137.

——, D. E. Rupp, and P. W. Mote, 2014: Seasonal climate variability and change in the Pacific Northwest of the United States. *J. Climate,* **27,** 2125–2142, doi:10.1175/JCLI-D-13-00218.1.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith, 1998: Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, 300 pp. [Available online at www.fao.org/docrep/X0490E/X0490E00.htm.]

Alley, W. M., 1984: The Palmer drought severity index: Limitations and assumptions. *J. Climate Appl. Meteor.,* **23,** 1100–1109, doi:10.1175/1520-0450(1984)023<1100:TPDSIL>2.0.CO;2.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier, 2005: Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature,* **438,** 303–309, doi:10.1038/nature04141.

Berghuijs, W. R., R. A. Woods, and M. Hrachowitz, 2014: A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nat. Climate Change,* **4,** 583–586, doi:10.1038/nclimate2246.

Botev, Z. I., J. F. Grotowski, and D. P. Kroese, 2010: Kernel density estimation via diffusion. *Ann. Stat.,* **38,** 2916–2957, doi:10.1214/10-AOS799.

Cayan, D. R., M. D. Dettinger, S. Kammerdiener, J. M. Caprio, and D. H. Peterson, 2001: Changes in the onset of spring in the western United States. *Bull. Amer. Meteor. Soc.,* **82,** 399–415, doi:10.1175/1520-0477(2001)082<0399:CITOOS>2.3.CO;2.

Clark, G. M., 2010: Changes in patterns of streamflow from unregulated watersheds in Idaho, western Wyoming, and northern Nevada. *J. Amer. Water Resour. Assoc.,* **46,** 486–497, doi:10.1111/j.1752-1688.2009.00416.x.

Crimmins, S. M., S. Z. Dobrowski, J. A. Greenberg, J. T. Abatzoglou, and A. R. Mynsberge, 2011: Response to comments on "Changes in climatic water balance drive downhill shifts in plant species' optimum elevations." *Science,* **334,** 177, doi:10.1126/science.1205029.

Dai, A., 2008: Temperature and pressure dependence of the rain–snow phase transition over land and ocean. *Geophys. Res. Lett.,* **35,** L12802, doi:10.1029/2008GL033295.

——, 2010: Drought under global warming: A review. *Wiley Interdiscip. Rev.: Climate Change,* **2,** 45–65, doi:10.1002/wcc.81; Corrigendum, **3,** 617, doi:10.1002/wcc.190.

——, 2011: Characteristics and trends in various forms of the Palmer drought severity index during 1900–2008. *J. Geophys. Res.,* **116,** D12115, doi:10.1029/2010JD015541.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. A. Pasteris, 2008: Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *Int. J. Climatol.,* **28,** 2031–2064, doi:10.1002/joc.1688.

Dettinger, M., K. Redmond, and D. Cayan, 2004: Winter orographic precipitation ratios in the Sierra Nevada—Large-scale atmospheric circulations and hydrologic consequences. *J. Hydrometeor.,* **5,** 1102–1116, doi:10.1175/JHM-390.1.

Dobrowski, S. Z., J. T. Abatzoglou, A. K. Swanson, J. A. Greenberg, A. R. Mynsberge, Z. A. Holden, and M. K. Schwartz, 2013: The climate velocity of the contiguous United States during the 20th century. *Global Change Biol.,* **19,** 241–251, doi:10.1111/gcb.12026.

Donohue, R. J., T. R. McVicar, and M. L. Roderick, 2010: Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate. *J. Hydrol.,* **386,** 186–197, doi:10.1016/j.jhydrol.2010.03.020.

Ellis, A. W., G. B. Goodrich, and G. M. Garfin, 2010: A hydroclimatic index for examining patterns of drought in the Colorado River basin. *Int. J. Climatol.,* **30,** 236–255, doi:10.1002/joc.1882.

Elsner, M. M., and Coauthors, 2010: Implications of 21st century climate change for the hydrology of Washington State. *Climatic Change,* **102,** 225–260, doi:10.1007/s10584-010-9855-0.

Falcone, J. A., cited 2014: GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. U.S. Geological Survey. [Available online at http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml.]

Heim, R. R., Jr., 2002: A review of twentieth-century drought indices used in the United States. *Bull. Amer. Meteor. Soc.,* **83,** 1149–1165.

Hidalgo, H. G., and Coauthors, 2009: Detection and attribution of streamflow timing changes to climate change in the western United States. *J. Climate,* **22,** 3838–3855, doi:10.1175/2009JCLI2470.1.

Holden, Z. A., J. T. Abatzoglou, L. S. Baggett, and C. Luce, 2011: Empirical downscaling of daily minimum air temperature at very fine resolutions in complex terrain. *Agric. For. Meteor.,* **151,** 1066–1073, doi:10.1016/j.agrformet.2011.03.011.

Jarvis, P. G., 1976: The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philos. Trans. Roy. Soc. London,* **B273,** 593–610, doi:10.1098/rstb.1976.0035.

Keyantash, J., and J. A. Dracup, 2002: The quantification of drought: An analysis of drought indices. *Bull. Amer. Meteor. Soc.,* **83,** 1167–1180.

Liang, X., D. P. Lettenmaier, and E. F. Wood, 1996: One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model. *J. Geophys. Res.,* **101,** 21 403–21 422, doi:10.1029/96JD01448.

López-Moreno, J. I., S. M. Vicente-Serrano, J. Zabalza, S. Beguería, J. Lorenzo-Lacruz, C. Azorin-Molina, and E. Morán-Tejeda, 2013: Hydrological response to climate variability at different time scales: A study in the Ebro basin. *J. Hydrol.,* **477,** 175–188, doi:10.1016/j.jhydrol.2012.11.028.

Luce, C. H., and Z. A. Holden, 2009: Declining annual streamflow distributions in the Pacific Northwest United States, 1948–2006. *Geophys. Res. Lett.,* **36,** L16401, doi:10.1029/2009GL039407.

——, J. T. Abatzoglou, and Z. A. Holden, 2013: The missing mountain water: Slower westerlies decrease orographic enhancement in the Pacific Northwest USA. *Science,* **342,** 1360–1364, doi:10.1126/science.1242335.

McCabe, G. J., and M. D. Dettinger, 2002: Primary modes and predictability of year-to-year snowpack variations in the western United States from teleconnections with Pacific Ocean climate. *J. Hydrometeor.,* **3,** 13–25, doi:10.1175/1525-7541(2002)003<0013:PMAPOY>2.0.CO;2.

——, and D. M. Wolock, 2011a: Century-scale variability in global annual runoff examined using a water balance model. *Int. J. Climatol.,* **31,** 1739–1748, doi:10.1002/joc.2198.

——, and ——, 2011b: Independent effects of temperature and precipitation on modeled runoff in the conterminous United States. *Water Resour. Res.,* **47,** W11522, doi:10.1029/2011WR010630.

McEvoy, D. J., J. L. Huntington, J. Abatzoglou, and L. Edwards, 2012: An evaluation of multi-scalar drought indices in Nevada and eastern California. *Earth Interact.,* **16**, doi:10.1175/2012EI000447.1.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, *Eighth Conf. on Applied Climatology,* Anaheim, CA, Amer. Meteor. Soc., 179–184.

Milly, P. C., 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.

Mishra, A. K., and V. P. Singh, 2010: A review of drought concepts. *J. Hydrol.,* **391,** 202–216, doi:10.1016/j.jhydrol.2010.07.012.

Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.,* **109,** D07S90, doi:10.1029/2003JD003823.

Mote, P. W., 2003: Trends in temperature and precipitation in the Pacific Northwest. *Northwest Sci.,* **77,** 271–282.

——, A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier, 2005: Declining mountain snowpack in western North America. *Bull. Amer. Meteor. Soc.,* **86,** 39–49, doi:10.1175/BAMS-86-1-39.

Neter, J., M. H. Kuutner, C. J. Nachtsheim, and W. Wasserman, 1996: *Applied Linear Regression Models.* McGraw-Hill, 720 pp.

Oudin, L., C. Michel, and F. Anctil, 2005: Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 1—Can rainfall–runoff models effectively handle detailed potential evapotranspiration inputs? *J. Hydrol.,* **303,** 275–289, doi:10.1016/j.jhydrol.2004.08.025.

Pagano, T., and D. Garen, 2005: A recent increase in western U.S. streamflow variability and persistence. *J. Hydrometeor.,* **6,** 173–179, doi:10.1175/JHM410.1.

——, ——, T. R. Perkins, and P. A. Pasteris, 2009: Daily updating of operational statistical seasonal water supply forecasts for the western U.S. *J. Amer. Water Resour. Assoc.,* **45,** 767–778, doi:10.1111/j.1752-1688.2009.00321.x.

Palmer, W. C., 1965: Meteorological drought. Weather Bureau Research Paper 45, 65 pp.

Redmond, K., 2002: The depiction of drought. *Bull. Amer. Meteor. Soc.,* **83,** 1143–1147.

Sheffield, J., E. F. Wood, and M. L. Roderick, 2012: Little change in global drought over the past 60 years. *Nature,* **491,** 435–438, doi:10.1038/nature11575.

Shukla, S., and A. W. Wood, 2008: Use of a standardized runoff index for characterizing hydrologic drought. *Geophys. Res. Lett.,* **35,** L02405, doi:10.1029/2007GL032487.

Siler, N., G. Roe, and D. Durran, 2013: On the dynamical causes of variability in the rain-shadow effect: A case study of the Washington Cascades. *J. Hydrometeor.,* **14,** 122–139, doi:10.1175/JHM-D-12-045.1.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2005: Changes toward earlier streamflow timing across western North America. *J. Climate,* **18,** 1136–1155, doi:10.1175/JCLI3321.1.

Thornthwaite, C. W., 1948: An approach toward a rational classification of climate. *Geogr. Rev.,* **38,** 55–94, doi:10.2307/210739.

van der Schrier, G., P. D. Jones, and K. R. Briffa, 2011: The sensitivity of the PDSI to the Thornthwaite and Penman–Monteith parameterizations for potential evapotranspiration. *J. Geophys. Res.,* **116,** D03106, doi:10.1029/2010JD015001.

Vano, J. A., T. Das, and D. P. Lettenmaier, 2012: Hydrologic sensitivities of Colorado River runoff to changes in precipitation and temperature. *J. Hydrometeor.,* **13,** 932–949, doi:10.1175/JHM-D-11-069.1.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno, 2010: A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Climate,* **23,** 1696–1718, doi:10.1175/2009JCLI2909.1.

——, and Coauthors, 2012a: Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interact.,* **16**, doi:10.1175/2012EI000434.1.

——, J. I. López-Moreno, S. Beguería, J. Lorenzo-Lacruz, E. Morán, and C. Azorín-Molina, 2012b: Accurate computation of a streamflow drought index. *J. Hydrol. Eng.,* **17,** 318–332, doi:10.1061/(ASCE)HE.1943-5584.0000433.

Wells, N., S. Goddard, and M. J. Hayes, 2004: A self-calibrating Palmer drought severity index. *J. Climate,* **17,** 2335–2351, doi:10.1175/1520-0442(2004)017<2335:ASPDSI>2.0.CO;2.

Wilhite, D. A., M. D. Svoboda, and M. J. Hayes, 2007: Understanding the complex impacts of drought: A key to enhancing drought mitigation and preparedness. *Water Resour. Manage.,* **21,** 763–774, doi:10.1007/s11269-006-9076-5.

Willmott, C. J., and S. M. Robeson, 1995: Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatol.,* **15,** 221–229, doi:10.1002/joc.3370150207.

——, C. M. Rowe, and Y. Mintz, 1985: Climatology of the terrestrial seasonal water cycle. *J. Climatol.,* **5,** 589–606, doi:10.1002/joc.3370050602.