

Major Over- and Underestimation of Drought Found in NOAA's Climate Divisional SPI Dataset

DANIELA SPADE AND KIRSTEN DE BEURS

Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, Oklahoma

MARK SHAFER

Department of Geography and Environmental Sustainability, University of Oklahoma, and Oklahoma Climatological Survey, and Southern Climate Impacts Planning Program, Norman, Oklahoma

(Manuscript received 6 November 2019, in final form 17 July 2020)

ABSTRACT

Evaluation of the standardized precipitation index (SPI) dataset published monthly in the National Oceanic and Atmospheric Administration/National Centers for Environmental Information (NOAA/NCEI) climate divisional database revealed that drought frequency is being mischaracterized in climate divisions across the United States. The 3- and 6-month September SPI values were downloaded from the database for all years between 1931 and 2019; the SPI was also calculated for the same time scales and span of years following the SPI method laid out by NOAA/NCEI. Drought frequency is characterized as the total number of years that the SPI fell below -1 . SPI values across 1931–90, the calibration period cited by NOAA/NCEI, showed regional patterns in climate divisions that are biased toward or away from drought, according to the average values of the SPI. For both time scales examined, the majority of the climate divisions in the central, Midwest, and northeastern United States showed negative averages, indicating bias toward drought, whereas climate divisions in the western United States, the northern Midwest, and parts of the Southeast and Texas had positive averages, indicating bias away from drought. The standard deviation of the SPI also differed from the expected value of 1. These regional patterns in the NCEI's SPI values are the result of a different (sliding) calibration period, 1895–2019, instead of the cited standardized period of 1931–90. The authors recommend that the NCEI modify its SPI computational procedure to reflect the best practices identified in the benchmark papers, namely, a fixed baseline period.


1. Introduction

a. Overview

The standardized precipitation index (SPI) is a probability-based moisture index designed to measure whether moisture conditions are normal, abnormally dry, or abnormally wet, relative to a predefined calibration, or base, period. The SPI was developed by McKee et al. (1993) to address the need for a drought index that had few data requirements and that accommodated for the fact that our key usable water sources

(soil moisture, groundwater, snowpack, streamflow, and reservoir storage) respond to moisture deficits and the eventual arrival of precipitation on a distinct time scale.

The SPI is based on the probability of accumulating a given amount of precipitation in a specified period of time, ranging from 1 to 24 months, relative to a predefined base period. The acquired cumulative probability values associated with each precipitation value are converted to the standard normal random variable Z , which allows for the estimation of both dry and wet conditions. Because the SPI is based on the standard normal distribution, an SPI value of zero signifies the mean precipitation amount, relative to the base period, and the index becomes more negative or positive as dry or wet conditions, respectively, become more severe (Svoboda et al. 2012).

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Kirsten de Beurs, kdebeurs@ou.edu

DOI: 10.1175/JAMC-D-19-0272.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

Despite its computation, the SPI can be non-normally distributed. For example, the SPI becomes lower bounded and therefore nonnormally distributed when there is a high frequency of no precipitation (i.e., values of zero) because this reduces the size of the dataset used to construct the SPI (Wu et al. 2007). The size of the dataset is a key limitation in the ability of the SPI to accurately portray drought/wet conditions because it is a probability-related index; the parameters of the SPI's underlying probability distribution are sensitive to the length of record, particularly if the precipitation pattern changes between lengths of record (Guttman 1994; Wu et al. 2005). Hence, in arid climates or those climates with a distinct seasonality to the precipitation regime, the SPI is prone to erroneous results and must be used with caution.

Nonetheless, the SPI offers several advantages for operational use, including the fact that it requires only precipitation data to be computed, it can be calculated at various time scales, enabling it to capture both short- and long-term drought and abnormal wetness, and it is normalized with respect to location, allowing it to be comparable across regions with different climates (Keyantash and NCAR Staff 2018). For these reasons, the SPI is widely used operationally in the United States. For example, it is one of the key drought indicators used by the U.S. Drought Monitor (USDM), which is considered to be the standard for operational drought monitoring in the United States (Svoboda et al. 2002). Similarly, the National Centers for Environmental Information (NCEI), formerly called the National Climatic Data Center, releases monthly State of the Climate Reports that provide detailed drought discussions according to what is indicated by the climate divisional SPI and Palmer drought index, among other drought indicators, available through the NOAA/NCEI climate divisional database (nClimDiv; Vose et al. 2014a). Furthermore, in 2011 the SPI was recommended through the Lincoln Declaration on Drought as the internationally preferred index to be used by all national meteorological and hydrological services to characterize meteorological droughts (Hayes et al. 2011).

Modest data requirements and temporal flexibility of the SPI make it also popular in a research setting. Kangas and Brown (2007) used the SPI at various accumulation periods to investigate the effect of time scale on the spatial patterns of drought frequency and duration. Spinoni et al. (2014) used the 12-month SPI to construct a global drought frequency dataset, with the 12-month accumulation period accommodating the various precipitation regimes around the world. Using the 3-, 6-, and 12-month SPI, Livada and Assimakopoulos (2007) explored the spatiotemporal variability of drought intensity

and duration in Greece. Furthermore, the temporal flexibility of the SPI also helped Guttman (1999) clarify the time scale associated with the Palmer drought index.

The objectives of this study stem from an analysis of drought frequency according to the USDM dataset. During the analysis of the SPI, which is one of the drought indicators examined by the authors of the USDM to construct the weekly USDM maps, inconsistencies were identified between the theoretical drought frequency values for a normalized index and the drought frequency values obtained from the NCEI's SPI dataset, prompting a careful examination of the dataset across its 1931–90 calibration period (see the NCEI document found online at <ftp://ftp.ncdc.noaa.gov/pub/data/cirs/climdiv/divisional-readme.txt>). Specifically, the objectives of this study are to demonstrate that the NCEI's SPI dataset is not in agreement with the standard normal distribution across the calibration period according to the average and standard deviation of each climate division's SPI value across that period. This was found to impact the characterization of drought frequency throughout the contiguous United States across the base period. Consultation with the point of contact for the NCEI's drought datasets revealed that the NCEI's SPI values do not align with the standard normal distribution because the NCEI uses a sliding calibration period, 1895–2019, instead of the cited standardized period of 1931–90.

The following subsection provides background information on the SPI; sections 2 and 3 are dedicated to this study's data and methods, respectively; section 4 provides the results and section 5 documents studies that have utilized the NOAA/NCEI SPI dataset in recent years. Section 6 provides a brief summary of the results and conclusions.

b. Background on the SPI

The SPI is a Z-score-like measurement of accumulated precipitation, identifying the number of standard deviations above or below the mean (precipitation) of the base period. Positive SPI values indicate greater than average precipitation and negative values indicate less than average precipitation. Figure 1 illustrates the relationship of the SPI to the standard normal distribution and following the statistical theory of the normal distribution, it demonstrates that a location is expected to be on the dry end of the spectrum 16% of the time and on the wet end of the spectrum 16% of the time. Note that an SPI value between -1 and 1 is considered neutral; a value less than -1 is considered to be dry; a value greater than 1 is considered to be wet.

Because precipitation frequency distributions typically are not normally distributed, several statistical procedures are used to transform accumulated precipitation values to

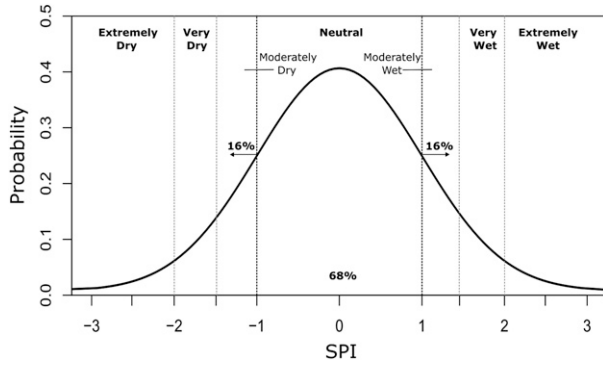


FIG. 1. Standard normal distribution with the SPI.

an SPI. The basic workflow first involves the selection of a precipitation accumulation period, which is directly relevant to the type of drought one is interested in identifying. For instance, the 3-month SPI is a viable accumulation period in the analysis of agricultural or soil moisture drought (Svoboda et al. 2012). The 3-month SPI calculated for the month of May utilizes the total accumulated precipitation for March, April, and May.

Next is the selection of a base period, which is a comparison period used to establish whether current conditions are normal, abnormally dry, or abnormally wet. The base period should ideally contain at least 30 years of continuous data that includes one long-term drought and one long-term wet period (Karl 1986; McKee et al. 1993; Edwards and McKee 1997). Guttman (1994) recommended that up to 80 years of data be used for reliable results in the estimation of extreme events.

Following the selection of the accumulation and base periods, a probability density function is chosen that best fits the long-term precipitation dataset. The cumulative probability of each precipitation value is calculated using the estimated parameters associated with the probability density function for each time scale of interest (1-, 3-, 6-month, etc.). Last, an equiprobability transformation is made in order to convert the cumulative probability to the standard normal random variable Z with a mean of zero and standard deviation of 1, yielding the SPI.

2. Data

NCEI SPI and precipitation

The contiguous United States consists of 344 climate divisions, with boundaries constructed through considerations of drainage basins, crop districts, climatic conditions, and county lines; each state contains between 3 and 10 climate divisions (Guttman and Quayle 1996). Divisional values of temperature and precipitation, from which the drought indices published in nClimDiv are derived, are estimated from area-weighted averages of gridpoint estimates that are interpolated from station data (Vose et al. 2014b). Climate divisional dataset files are available for download in “TXT,” “MAP,” and “KMZ” file formats.

SPI datasets were downloaded at 3- and 6-month time scales for the 48 contiguous United States from nClimDiv for all years from 1931 to 2019. The 3- and 6-month SPI time scales were selected for their relevance in capturing the short- and medium-term drought conditions (i.e., agricultural and meteorological drought) that impact the United States, and also because they are among the time scales used by NOAA/NCEI (hereinafter NCEI) to issue monthly State of the Climate Reports. The SPI at each time scale was downloaded using September as the ending period for the moving-total precipitation. September was selected to investigate end-of-water-year drought conditions.

Precipitation data were downloaded for each climate division to calculate the 3- and 6-month September SPI. The NCEI cites a base period of 1931–90 for all drought data posted in the nClimDiv database and this base period is also used in the calculated SPI dataset for consistency.

3. Methods

a. SPI computational procedure

The NCEI precipitation data are used to calculate the SPI. The NCEI fits monthly climate division precipitation data to the Pearson type-III (PE3) probability density function for all precipitation values $x > 0$:

$$f(x) = \frac{(x - \xi)^{\alpha-1} \exp - (x - \xi)/\beta}{\beta^\alpha \Gamma(\alpha)} \quad \text{for } \gamma > 0 \quad \text{and} \quad \xi \leq x < \infty \quad \text{and} \quad (1)$$

$$f(x) = \frac{(\xi - x)^{\alpha-1} \exp - (\xi - x)/\beta}{\beta^\alpha \Gamma(\alpha)} \quad \text{for } \gamma < 0 \quad \text{and} \quad -\infty < x \leq \xi, \quad (2)$$

where α , β , and ξ are the shape, scale, and location parameters, respectively, and are respectively given by

$$\alpha = 4/\gamma^2, \quad (3)$$

$$\beta = \frac{1}{2}\sigma|\gamma|, \quad \text{and} \quad (4)$$

$$\xi = \mu - 2\sigma/\gamma. \quad (5)$$

The $\Gamma(\alpha)$ in Eqs. (1) and (2) denotes the gamma function (GAM). The method of L-moments is used by the NCEI for estimation of the PE3's location, shape, and scale parameters. The Fortran computer program SPICOMPUTE by N. Guttman outlines the computational procedure used to obtain the NCEI SPI values and can be found online (<http://www1.ncdc.noaa.gov/pub/data/software/palmer/spi.f>). Here, the same probability density function and SPI computational procedure outlined in SPICOMPUTE are applied. See appendix B for a thorough description of the SPI computational procedure using the PE3 distribution and the method of L-moments for parameter estimation.

While the SPI is unbounded in theory, Guttman (1999) recommends truncating the cumulative probabilities from 0.001 to 0.999, which bounds the SPI between ± 3.09 . The estimation of extreme probabilities based on sample sizes < 100 may not be reliable. These bounds are enforced in the SPICOMPUTE FORTRAN code and are also applied in this study's method for consistency.

The ability of the PE3 probability distribution to accurately model the precipitation time series was evaluated using a two-sample Kolmogorov–Smirnov test and Pearson's chi-squared test, with a null hypothesis that the data comes from the normal probability distribution, at a significance level of 0.05.

b. Evaluation of the NCEI SPI against the derived PE3 SPI

To understand how the downloaded SPI values compared with the PE3-calculated SPI values based on the NCEI's precipitation data, several different experiments were designed.

1) COMPARISON OF THE AVERAGE AND STANDARD DEVIATION OF SPI

At each time scale and for each climate division, the average and standard deviation of the downloaded and calculated SPI values are taken across the base period. These values are expected to come out to 0 and 1, respectively, in accordance with the standard normal distribution. These criteria are used to determine whether or not a climate division's SPI time series agrees with the

standard normal distribution. Average values below -0.01 are considered to be negatively biased, and values above 0.01 are considered to be positively biased. Likewise, a standard deviation below 0.99 represents a precipitation regime with a narrower distribution than the standard normal distribution, whereas a value above 1.01 is a wider distribution. This study compares the average and standard deviation of the NCEI SPI and the calculated SPI, based on the PE3 distribution, for all climate divisions. Note that the expectation is that all means are 0 and all standard deviations are 1. In addition, it is expected that the NCEI SPI and the PE3 SPI values be identical because of the reported NCEI SPI calculation method.

The impact on drought detection of a climate division with a negative average SPI value across the base period is that the climate division is biased toward indicating drought conditions. In contrast, a positive average SPI value indicates the climate division is biased away from drought, meaning it is more difficult for the index to detect genuine drought in that climate division (Figs. 2a,b, respectively). The impact on drought detection of a climate division with a wider distribution than the standard normal distribution (i.e., a standard deviation greater than 1) is that it is more difficult for the index to move between levels of drought than it would be if the distribution was truly normal (Fig. 2c). Likewise, the impact on drought detection of having a standard deviation below 1 is that it is overly easy for the index to move between levels of drought intensity (Fig. 2d). In statistical terminology, these adjustments to the distribution width represent changes to the kurtosis of the distribution.

2) COMPARISON OF NCEI SPI AND PE3 SPI BY CLIMATE DIVISION ACROSS TIME

Three scatterplots were created using the 6-month NCEI SPI dataset plotted against the calculated 6-month SPI dataset. A scatterplot was created for a climate division with a negative average, a climate division with a positive average and a climate division with an average of 0 across the base period. Each scatterplot consists of 120 points, 60 for each dataset. A regression analysis was performed with a significance level of 0.05 to determine a relationship between the two datasets. The purpose of this analysis was to determine how the NCEI dataset compares with the calculated dataset across the entire range of SPI values comprising the base period. For brevity, these results focus on the 6-month SPI as the 3-month SPI results were effectively identical.

From the standard normal distribution (Fig. 1), it is expected that approximately 9 drought years occur (16% of the time) across the 60-yr base period for each climate division. The same is expected for abnormally wet conditions.

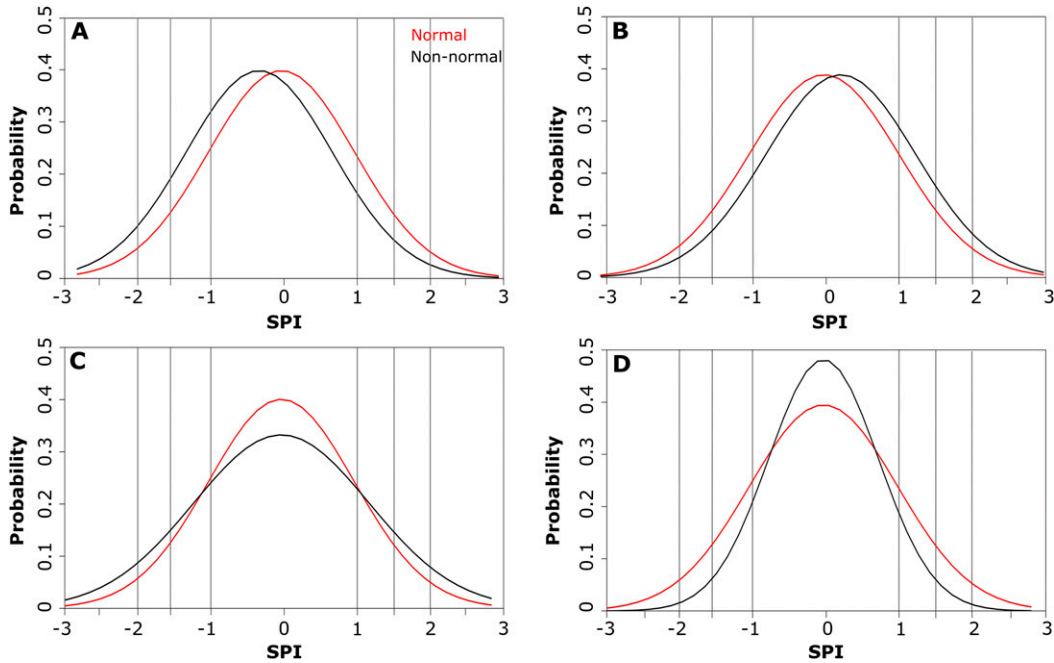


FIG. 2. Standard normal distribution (red curve) overlaid (black curves) with (a) a negatively biased distribution, (b) a positively biased distribution, (c) a distribution with a standard deviation that is greater than 1 (leptokurtic), and (d) a distribution with a standard deviation that is less than 1 (platykurtic).

To quantify the differences between the SPI values downloaded from the nClimDiv database (hereinafter referred to as NCEI3 for the 3-month SPI and NCEI6 for the 6-month SPI) and the calculated SPI values (abbreviated Calc3 and Calc6), the difference is taken between the total number of dry or wet years among the two datasets:

$$\text{NCEI3} - \text{Calc3} \quad \text{and} \quad (6)$$

$$\text{NCEI6} - \text{Calc6}. \quad (7)$$

A negative or positive value respectively indicates that more dry or wet years were estimated by the Calc SPI dataset, and a value of zero indicates that there is no difference between the two datasets.

4. Results

a. Goodness of fit

Goodness-of-fit tests were performed to confirm the normality of the Calc3 and Calc6 datasets. The goodness-of-fit tests were the Kolmogorov–Smirnov and Pearson’s chi-squared tests. Hypothesis tests for these two approaches produced *p* values of 1 for every climate division and in every time scale analyzed, indicating the PE3 probability distribution successfully transforms the original precipitation time series into a normal distribution.

b. 1931–90 average and standard deviation of the SPI

If the NCEI3 data behave as expected, average SPI values should equal 0 with a standard deviation of 1 for each climate division. Instead, this study found that 114 (33.1%) of the 344 climate divisions across the conterminous United States had a positive average SPI across the base period. The bulk of these climate divisions are concentrated across the western United States, with smaller groups of positive average values appearing along the Gulf Coast and Midwest (Fig. 3a). The same pattern is also shown by the NCEI6 data, although to a lesser extent, with 75 (21.8%) climate divisions showing a positive average (Fig. 3e). More climate divisions with negative averages were found across the base period [200 (56.1%) climate divisions in NCEI3 and 232 (67.4%) climate divisions in NCEI6]. The spatial pattern of negative averages is the same in both datasets—the majority of the climate divisions in the central and northeastern United States showed negative averages to varying extents.

The western United States shows the greatest grouping of positive averages whereas spatial grouping of negative averages is prevalent in the south-central, Midwest, and northeastern United States according to the NCEI3 and NCEI6 datasets.

Although the Calc3 and Calc6 datasets also contain climate divisions with averages that are different from 0 (Figs. 3c and 3g, respectively), the spatial

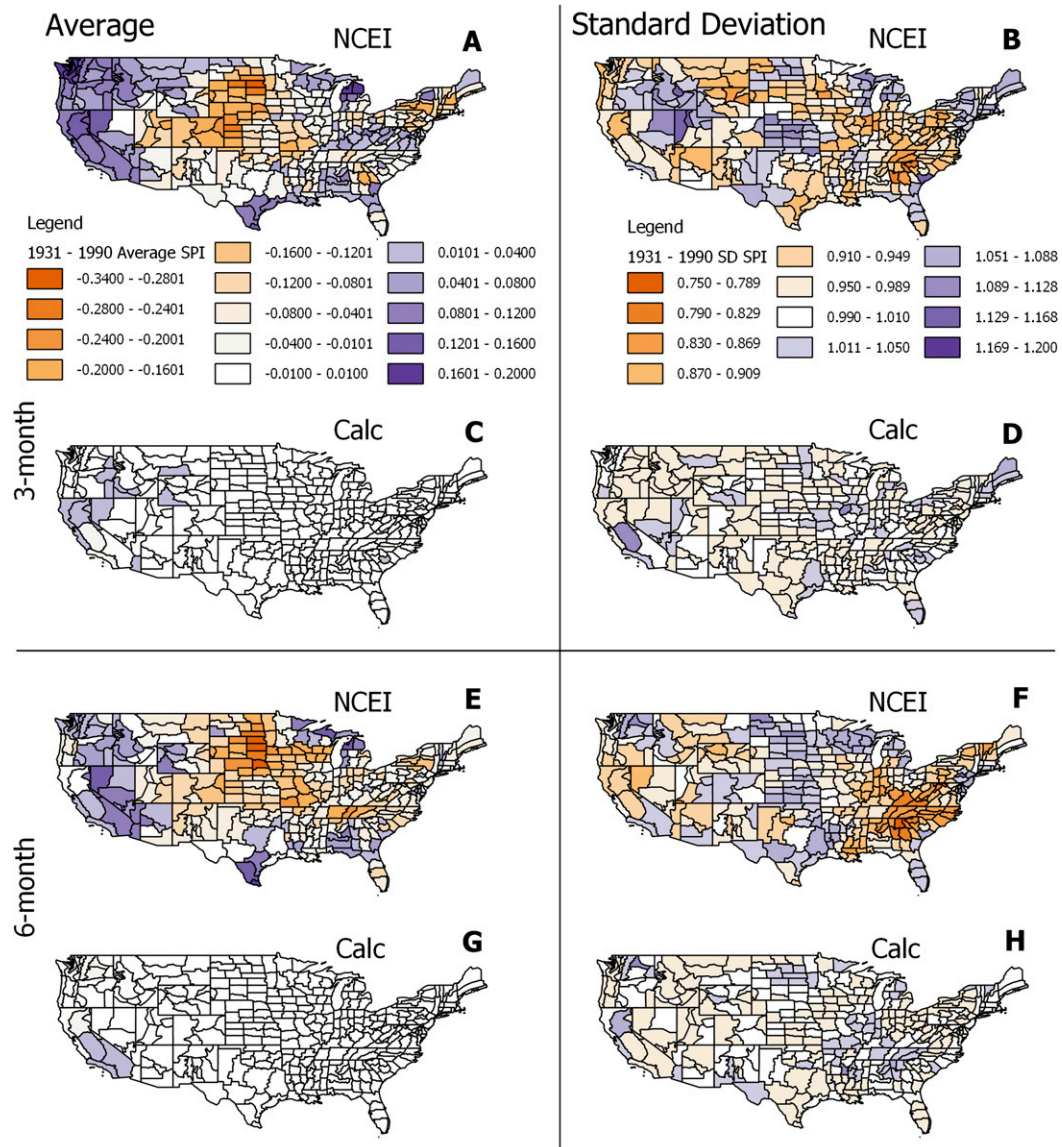


FIG. 3. The 1931–90 (left) average and (right) standard deviation of September SPI values according to the (a),(b) NCEI3, (c),(d) Calc3, (e),(f) NCEI6, and (g),(h) Calc6.

patterns seen in NCEI3 and NCEI6 could not be replicated at either time scale. Rather, only 16 (4.7%) climate divisions with positive and 3 (0.9%) climate divisions with negative averages appear in the Calc3 dataset with 3 (0.9%) and 7 (2.0%) in the Calc6 dataset, when following the PE3 methodology. The spatial distribution of the climate divisions with averages different from zero according to the Calc3 and Calc6 datasets appears to be random, with some grouping of positive averages for climate divisions in Northern California and eastern Oregon.

Spatial grouping of standard deviations of the SPI values different from 1 across the base period is also

evident in the NCEI3 and NCEI6 datasets (Figs. 3b and 3f, respectively). In total, 107 (31.1%) climate divisions have distributions wider than the standard normal distribution and 196 (57.0%) have distributions narrower than the standard normal distribution according to the NCEI3 dataset. Similarly, 108 (31.4%) have distributions wider than the standard normal distribution and 201 (58.4%) climate divisions have distributions narrower than the standard normal distribution in the NCEI6 dataset. So, the two datasets exhibit very similar biases in the shapes of the SPI distributions, seemingly independent of the time scale. In both datasets, the standard deviation is above 1 in climate divisions in the

western and midwestern United States. Interestingly, the northeastern United States switches between a standard deviation above 1 in the NCEI3 dataset to a value below 1 in the NCEI6 dataset, suggesting a seasonal component to the distributions. This is also seen in a group of climate divisions in Nevada, Utah, Idaho, and Oregon. Although not shown here, examination of the NCEI3 dataset for March and June also revealed a seasonal component in the spatial patterns of averages and standard deviations. This suggests that, according to the NCEI dataset, whether a climate division will be biased toward or away from drought is dependent on the ending period for the moving-total precipitation.

For many climate divisions across the United States in the Calc3 and Calc6 datasets, the standard deviation of SPI values also differs from 1, although to a greater degree in the NCEI datasets (Figs. 3d and 3h, respectively). In total this study found 54 (15.7%) climate divisions with a standard deviation wider and 164 (47.7%) with a standard deviation narrower than the standard normal distribution in Calc3; this study found 53 (163) climate divisions with a standard deviation wider (narrower) than the standard normal distribution in Calc6. As was the case for the mean SPI computations, the proportion of standard deviations above and below the reference value (i.e., 1) were very similar across the 3- and 6-month SPI periods. Thus, it does not appear that the averaging period is responsible for the fluctuations. This is addressed further in the appendix. Nonetheless, regional patterns in the standard deviations above or below 1 are not seen in the Calc3 or Calc6 results.

c. Relationship between NCEI6 and Calc6, 1931–90

Figure 4 plots NCEI6 against the Calc6 dataset for climate divisions that have a negative average (Fig. 4a), positive average (Fig. 4b) and an average of zero (Fig. 4c) according to the NCEI6 dataset shown in Fig. 3a. The selected climate divisions are shown in bold in the map insets.

Figures 4a and 4b shows disagreements between the NCEI6 and Calc6 datasets along the entire range of SPI values. There is a nearly 1:1 relationship between the two dataset values (correlation coefficient squared $R^2 = 0.99$), but the intercept is below (Fig. 4a) and above (Fig. 4b) zero with a slope less than 1. The fact that there is not a lot of variation along the regression lines indicates that the two SPI datasets were derived from the same underlying precipitation dataset, otherwise there would be more variation along the line. For the climate division in Nebraska (Fig. 4a), the intercept is -0.29 with a slope of 0.96 , which is notably lower than 1. In terms of drought and wet events, this indicates that NCEI6 is underestimating the magnitude of

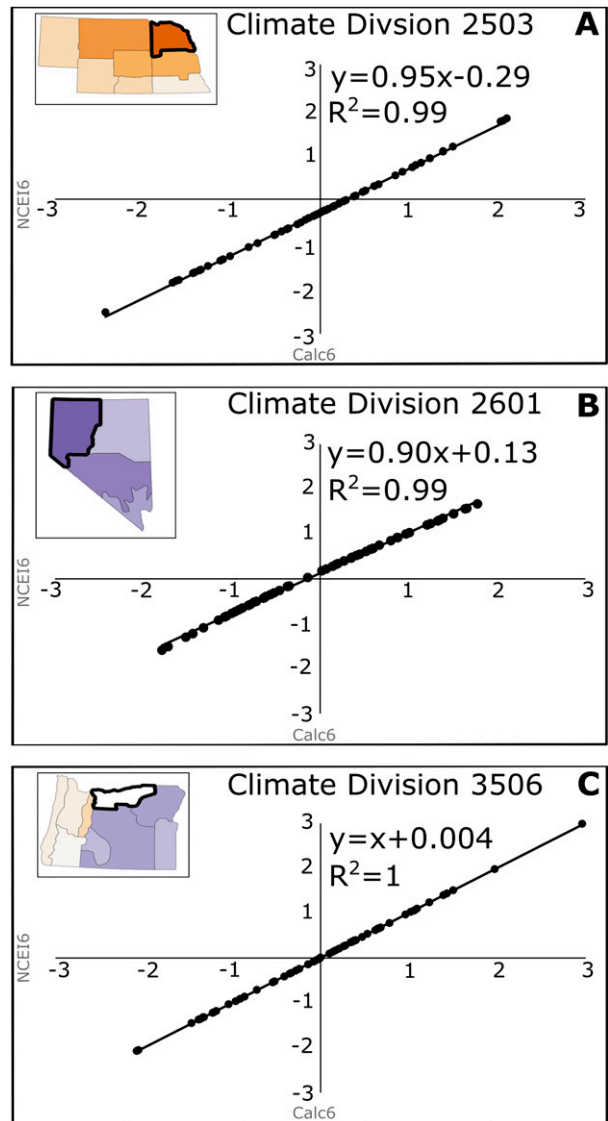


FIG. 4. The 1931–90 September NCEI6 values plotted against the Calc6 values for (a) a climate division with a negative average value across that time period according to the NCEI6, (b) a climate division with a positive average value, and (c) a climate division with an average value of 0.

precipitation events and overestimating the magnitude of drought events. For the climate division in Nevada (Fig. 4b), the intercept is 0.13 with a slope of 0.90 , which is significantly lower than 1. In terms of drought and wet events, this indicates that NCEI6 is underestimating the magnitude of dry events and overestimating the magnitude of wet events. In contrast, the climate division in Oregon (Fig. 4c) shows a 1:1 relationship ($R^2 = 1$) with a regression line centered over zero and a slope equal to 1, indicating that NCEI6 and Calc6 are estimating dry and wet events at the same magnitude.

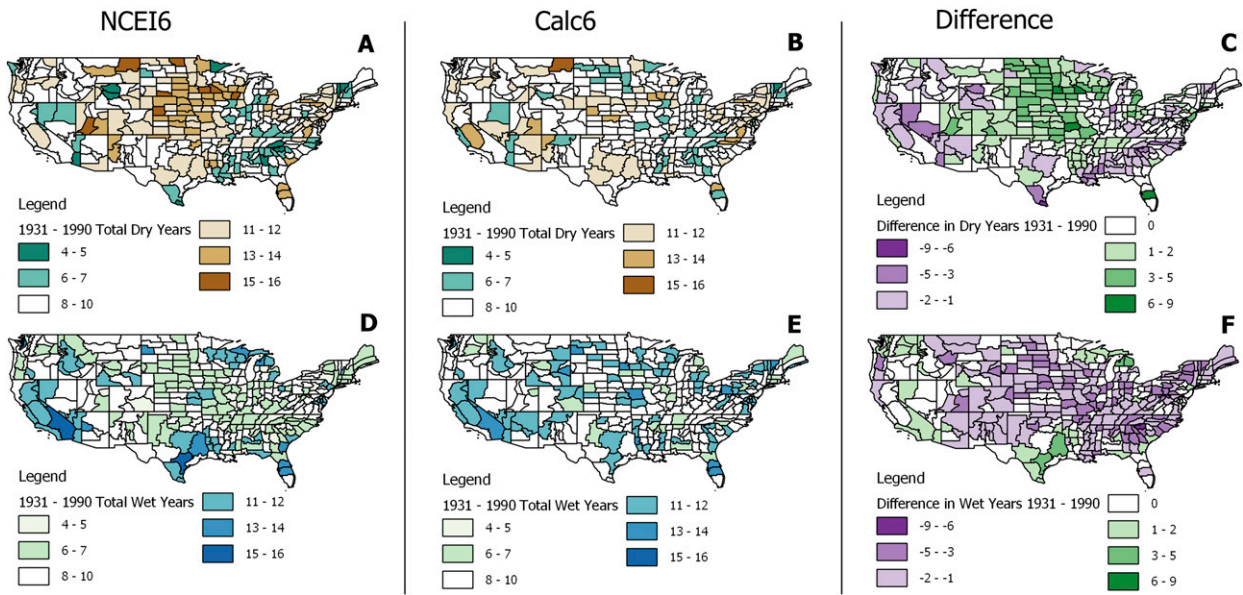


FIG. 5. Total number of (a),(b) dry and (d),(e) wet years across the 1931–90 time period according to the (left) NCEI6 and (center) Calc6 datasets, along with the difference in the total number of (c) dry and (f) wet years between the two datasets (i.e., NCEI6 – Calc6).

d. 1931–90 total number of dry and wet years

Climate divisions in the central, Midwest, and northeastern United States are dry biased with standard deviations greater than 1 according to the NCEI3 and NCEI6 datasets, making it more difficult for the SPI to detect wet conditions relative to the standard normal distribution. The effects of the dry bias in these regions are shown in Fig. 5a, which illustrates the total number of dry years that occurred across the 1931–90 base period according to the NCEI6 dataset. Unsurprisingly, each of the three regions exhibits a greater number of dry years than expected.

Climate divisions in the western, southeastern, and northern Midwest parts of the United States are biased away from drought (and toward wet conditions) with standard deviations less than 1 according to the NCEI6 dataset, making it more difficult for the SPI to detect dry conditions relative to the standard normal distribution. The effects of the wet bias in these regions are shown in Fig. 5d, which shows the total number of wet years that occurred across the 1931–90 base period according to the NCEI6 dataset. A greater number of wet years than would be expected is observed in many regions.

Because the SPI is a standardized index, each climate division should theoretically contain the same number of dry or wet years. According to the statistical theory of the standard normal distribution, each climate division should be on the dry end of the spectrum 16% of the time and on the wet end of the spectrum 16% of the

time. This was found to be the case for the majority of the climate divisions based on the Calc6 dataset across the 1931–90 time period (Figs. 5b,e), wherein 219 climate divisions across the United States experienced dry conditions for the expected total of 8–10 dry years and 221 climate divisions experienced wet conditions for the expected total of 8–10 wet years across the 60-yr period. Although not shown here, 199 and 220 climate divisions contained the expected total number of dry and wet years, respectively, according to the Calc3 dataset. Figures 5c and 5f elucidate spatial differences between the total number of dry and wet years estimated by the two datasets (NCEI6–Calc6), and these differences arise in climate divisions with averages different from zero shown in Fig. 3e. Climate divisions with negative average values, such as those in the Midwest and southeast United States, estimated far more dry years than the Calc6 dataset; the northern Great Plains in particular estimated between 7 and 9 additional dry years (Fig. 5c). The same regions estimated far fewer wet years than the Calc6 dataset (Fig. 5f).

These results demonstrate the cumulative effect that the biased average and standard deviations in the NCEI6 dataset had on drought detection across the base period.

5. Discussion

Consultation with the point of contact for the nClimDiv database revealed that the NCEI's SPI values are

dry/wet biased in this study's results because the dataset is in fact calibrated using the full period of record (here, 1895–2019) and not 1931–90. This moving-window approach to calibrating the data means that there is no fixed baseline associated with the dataset, the calibration period is updated every year. The lack of a baseline makes NCEI's SPI data unusable as a research dataset because the baseline is updated continuously as new data become available; consequently, it is impossible to accurately compare across studies as they each may be comparing against different historical records. Therefore, it becomes impossible to use the dataset to identify changes in frequency of occurrence or intensity of drought. For example, studies typically compare present anomalies with an established baseline to assess change, such as departure from 20th-century average (e.g., IPCC, National Climate Assessment, and NCEI State of the Climate Report).

Another concern is that calculating the SPI in such a fashion violates the formulation originally proposed by McKee et al. (1993), who recommended a recent climatic history as the basis for comparison to derive the SPI. Further, a fixed base period was used by Edwards and McKee (1997), who recommend a fixed calibration period containing at least one dry spell and one wet spell. The NCEI's approach does not follow this standard and users are not aware of the discrepancy because the online metadata associated with the product indicate that "all drought data are calibrated using the period 1931–1990." Consequently, the 1931–90 period was used to compare against the NCEI's SPI formulations in this study, which revealed the inconsistencies. However, other studies that did not conduct a validation of the NCEI dataset for the base period would be unaware that it lacks a fixed baseline.

A review of the literature identified several studies that have used the NCEI SPI dataset in recent years. Cumbie-Ward and Boyles (2016) used the climate divisional NCEI SPI dataset at various time scales and for all months across a 10-yr period ending in 2015 to make an objective analysis of the performance of an experimental high-resolution SPI dataset. Lu et al. (2019) used the NCEI SPI (1-, 2-, 3-, 6-, 9-, and 12-month time scales) as well as the NCEI Palmer drought indices in the derivation of an experimental monthly agricultural drought index across the 1895–2013 time period. In addition to product calibration and validation, the climate divisional SPI dataset has also been used in recent years to link groundwater levels with drought in the central United States (Whitemore et al. 2016; Leelaruban et al. 2017). Despite the fact that each of these studies downloaded the same datasets from the same data source, the authors are all analyzing a different set of

SPI values because the calibration period is not fixed. For example, in 1934 climate division 2503 in Nebraska has a 6-month September SPI value of -1.35 based on a 1931–90 base period, -1.62 based on an 1895–2009 base period, and -1.58 based on an 1895–2019 period. Thus, the only way for various research studies to be comparable would be if the respective authors happened to publish simultaneously.

6. Recommendations

The authors recommend that the NCEI modify its SPI computational procedure to reflect the best practices identified by McKee et al. (1993) and Edwards and McKee (1997), namely, the use of a fixed baseline period. Further, the NCEI should go back through its SPI database and recompute the entire SPI archive using a fixed baseline period, presumably 1931–90. This will allow the authors of previous studies that have incorporated the NCEI SPI dataset to reevaluate their previous conclusions using the revised NCEI SPI dataset.

7. Conclusions

Because the SPI is based on the standard normal distribution, it should have a mean of zero and a standard deviation of 1 across the base period. Our results indicate inconsistencies between the NCEI's divisional SPI datasets and the standard normal distribution. Examination of the NCEI's 3- and 6-month September SPI datasets across the cited 1931–90 base period indicates climate divisions in the western United States, parts of the southeastern United States, and the northern Midwest are biased away from drought; climate divisions throughout the central to northern plains, Midwest, Northeast, and parts of the Southwest were found to be biased toward drought (Fig. 3). Consultation with the NCEI dataset's point of contact revealed that these biases appeared in the dataset over the 1931–90 time period because the dataset is in fact calibrated using a moving window, in contrast to what is stated in the online metadata associated with the product.

The NCEI must make it clear that the dataset in its current form cannot be used to identify changes in frequency of occurrence or intensity of drought because the SPI values for any given climate division evolve over time as the calibration period continually expands. Further, the current dataset should not be used in research settings, as results become impossible to reproduce given that the dataset values change over time. The authors strongly recommend that the NCEI modify its SPI computational procedure from a moving window to a fixed baseline approach and republish the entire SPI archive using this fixed baseline period.

TABLE A1. Climate type associated with the climate divisions that had SPI values in disagreement with the standard normal distribution according to the Calc3 and Calc6 datasets.

Calc3		Calc6	
Climate division	Climate type	Climate division	Climate type
205	Warm deserts	309	Southeastern U.S. plains
401	Western cordillera	402	Western cordillera
402	Western cordillera	405	Mediterranean California
403	Western cordillera	406	Mediterranean California
404	Mediterranean California	407	Warm deserts
405	Mediterranean California	901	Ozark–Ouachita–Appalachian forests
406	Mediterranean California	2906	South-central semiarid prairies
1005	Cold deserts	4506	Western cordillera
1006	Cold deserts	4507	Cold deserts
1007	Cold deserts	4508	Cold deserts
1103	Southeastern U.S. plains		
2405	West-central semiarid prairies		
2601	Cold deserts		
3508	Western cordillera		
3509	Cold deserts		
3806	Southeastern U.S. plains		
4202	Warm deserts		
4510	Cold deserts		
4803	Cold deserts		

Acknowledgments. The authors greatly appreciate the three anonymous reviewers for providing their suggestions. We also thank the editor, Dr. Ellis.

APPENDIX A

Calc3 and Calc6 Average and Standard Deviation

On the basis of the average of the Calc3 SPI values across the base period, 19 climate divisions have an SPI that does not follow a standard normal distribution (i.e., nonzero average). Seventeen of these climate divisions are in the western United States and, based on the Level III Ecoregion data provided by the EPA, all but four of these climate divisions reside in either a Mediterranean climate or desert climate (western Idaho, eastern Oregon, and southwest Wyoming). The remaining two climate divisions reside in nonarid climates. Likewise, according to the Calc6 average SPI across the base period, 10 climate divisions have an SPI that does not follow a standard normal distribution across the base period. Eight of these climate divisions are in the western United States, with the majority residing in either a Mediterranean or desert climate. The remaining two climate divisions reside in nonarid climates; see Table A1. Wu et al. (2007) demonstrated that the SPI can become nonnormally distributed whenever there is a high probability of zero values in the dataset and advised caution when applying the SPI to study drought in arid climates or climates with a distinct seasonality to the precipitation. Therefore, the

nonzero averages in Calc3 and Calc6 in this study are attributed to the high probability of zero values in the datasets, because the climate divisions with SPI values not following the standard normal distribution reside in regions with a distinct seasonality in the precipitation regime. Thus, we should be cognizant that the nonzero averages of the SPI in these calculations may be an indicator that the SPI is not fully valid for these climate divisions.

The standard deviations differ from 1 in the Calc3 and Calc6 datasets. This study attributes this to the method of L-moments used to calculate the parameters of the PE3 distribution. L-moments are the expectations of linear combinations of order statistics, measuring the same aspects of a distribution as the central moments. Hosking (1990) gives a comprehensive description of the theory of L-moments, which are defined for continuous probability distributions, but in practice they often must be estimated from a finite sample (Hosking 1996). To verify the effect of L-moments on the standard deviation of SPI values in the Calc3 and Calc6 datasets, the climate division precipitation data was fit to the two-parameter GAM (Thom 1958), which was used by McKee et al. (1993) in the original development of the SPI. For this analysis, the parameters were calculated via maximum likelihood estimation (Venables and Ripley 2002) instead of the method of L-moments. Following this approach, only 12 (3.5%) of the climate divisions were found to have standard deviations in disagreement with the standard normal distribution for the GAM-derived Calc3 and Calc6 datasets, respectively.

APPENDIX B

SPI Computational Procedure via Method of L-Moments

Integrating Eqs. (1) and (2) from section 3a over the range of precipitation values yields the cumulative distribution function:

$$F(x) = G\left(\alpha, \frac{x - \xi}{\beta}\right) / \Gamma(\alpha) \quad \text{for } \gamma > 0 \quad \text{and} \quad (B1)$$

$$F(x) = 1 - G\left(\alpha, \frac{\xi - x}{\beta}\right) / \Gamma(\alpha) \quad \text{for } \gamma < 0, \quad (B2)$$

where $G(\alpha, x)$ denotes the incomplete gamma function:

$$G(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t} dt. \quad (B3)$$

The L-moments are the expectations of linear combinations of order statistics, measuring the same aspects of a distribution as the central moments. Hosking (1990) gives a comprehensive description of the theory of L-moments and L-moment ratios. The latter are computed for higher order L-moments ($r \geq 3$):

$$t_r = \ell_r / \ell_2, \quad (B4)$$

where

$$\ell_r = n^{-1} \sum_{j=1}^n w_{jn}^{(r)} x_{jn}, \quad (B5)$$

$$w_{jn}^{(1)} = 1, \quad (B6)$$

$$w_{jn}^{(2)} = 2(j-1)/(n-1) - 1, \quad \text{and} \quad (B7)$$

$$(r-1)(n-r+1)w_{jn}^{(r)} = (2r-3)(2j-n-1)w_{jn}^{(r-1)} - (r-2)(n+r-2)w_{jn}^{(r-2)} \quad \text{for} \quad r \geq 3. \quad (B8)$$

Given the L-moments above, the parameters of the PE3 distribution are calculated following Hosking and Wallis (1997). If $0 < |t_3| < 1/3$, then $z = 3\pi t_3^2$ and

$$\alpha \approx \frac{1 + 0.2906z}{z + 0.1882z^2 + 0.0442z^3}. \quad (B9)$$

If $1/3 \leq |t_3| < 1$, then $z = 1 - |t_3|$ and

$$\alpha \approx \frac{0.36067z - 0.59567z^2 + 0.25361z^3}{1 - 2.78861z + 2.56096z^2 - 0.77045z^3}. \quad (B10)$$

Following the estimation of α , the parameters of the PE3 distribution are calculated as

$$\gamma = 2\alpha^{-1/2} \text{sign}(t_3),$$

$$\sigma = t_2 \pi^{1/2} \alpha^{1/2} \Gamma(\alpha) / \Gamma\left(\alpha + \frac{1}{2}\right), \quad \text{and}$$

$$\mu = t_1. \quad (B11)$$

Cumulative probability values are obtained after plugging the values of Eq. (B11) into Eqs. (B1) and (B2).

REFERENCES

Cumbie-Ward, R. V., and R. P. Boyles, 2016: Evaluation of a high-resolution SPI for monitoring local drought severity. *J. Appl. Meteor. Climatol.*, **55**, 2247–2262, <https://doi.org/10.1175/JAMC-D-16-0106.1>.

Edwards, D. C., and T. B. McKee, 1997: Characteristics of 20th century drought in the United States at multiple time scales. Colorado State University Department of Atmospheric Science Climatology Rep. 97-2, 155 pp., <http://hdl.handle.net/10217/170176>.

Guttman, N. B., 1994: On the sensitivity of sample L moments to sample size. *J. Climate*, **7**, 1026–1029, [https://doi.org/10.1175/1520-0442\(1994\)007<1026:OTSOSL>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1026:OTSOSL>2.0.CO;2).

—, 1999: Accepting the standardized precipitation index: A calculation algorithm. *J. Amer. Water Resour. Assoc.*, **35**, 311–322, <https://doi.org/10.1111/j.1752-1688.1999.tb03592.x>.

—, and R. G. Quayle, 1996: A historical perspective of U.S. climate divisions. *Bull. Amer. Meteor. Soc.*, **77**, 293–304, [https://doi.org/10.1175/1520-0477\(1996\)077%3C0293:AHPOUC%3E2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077%3C0293:AHPOUC%3E2.0.CO;2).

Hayes, M., M. Svoboda, N. Wall, and M. Widhalm, 2011: The Lincoln declaration on drought: Universal meteorological drought index recommended. *Bull. Amer. Meteor. Soc.*, **92**, 485–488, <https://doi.org/10.1175/2010BAMS3103.1>.

Hosking, J. R. M., 1990: L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Stat. Soc.*, **52B**, 105–124, <https://www.jstor.org/stable/2345653>.

—, 1996: FORTRAN routines for use with the method of L-moments version 3. IBM T. J. Watson Research Center Res. Rep. RC20525, 157 pp.

—, and J. R. Wallis, 1997: *Regional Frequency Analysis—An Approach Based on L-Moments*. Cambridge University Press, 224 pp.

Kangas, R., and T. Brown, 2007: Characteristics of US drought and pluvials from a high-resolution spatial dataset. *Int. J. Climatol.*, **27**, 1303–1325, <https://doi.org/10.1002/joc.1473>.

Karl, T. R., 1986: The sensitivity of the Palmer drought severity index and Palmer’s Z-index to their calibration coefficients including potential evapotranspiration. *J. Climate Appl. Meteor.*, **25**, 77–86, [https://doi.org/10.1175/1520-0450\(1986\)025<0077:TSOTPD>2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025<0077:TSOTPD>2.0.CO;2).

Keyantash, J., and NCAR Staff, Eds., 2018: The climate data guide: Standardized precipitation index (SPI). NCAR/UCAR, <https://climatedataguide.ucar.edu/climate-data/standardized-precipitation-index-spi>.

Leelarubhan, N., G. Padmanabhan, and P. Oduor, 2017: Examining the relationship between drought indices and groundwater levels. *Water*, **9**, 82, <https://doi.org/10.3390/w9020082>.

- Livada, I., and V. D. Assimakopoulos, 2007: Spatial and temporal analysis of drought in Greece using the standardized precipitation index (SPI). *Theor. Appl. Climatol.*, **89**, 143–153, <https://doi.org/10.1007/s00704-005-0227-z>.
- Lu, J., G. J. Carbone, and P. Gao, 2019: Mapping the agricultural drought based on the long-term AVHRR NDVI and North American Regional Reanalysis (NARR) in the United States, 1981–2013. *Appl. Geogr.*, **104**, 10–20, <https://doi.org/10.1016/j.apgeog.2019.01.005>.
- McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. *Eighth Conf. on Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–186.
- Spinoni, J., G. Naumann, H. Carrao, P. Barbosa, and J. Vogt, 2014: World drought frequency, duration, and severity for 1951–2010. *Int. J. Climatol.*, **34**, 2792–2804, <http://doi.org/10.1002/joc.3875>.
- Svoboda, M., and Coauthors, 2002: The Drought Monitor. *Bull. Amer. Meteor. Soc.*, **83**, 1181–1190, <https://doi.org/10.1175/1520-0477-83.8.1181>.
- , M. Hayes, and D. Wood, 2012: Standardized precipitation index user guide. World Meteorological Organization Doc. WMO-No. 1090, 24 pp., https://library.wmo.int/doc_num.php?explnum_id=7768.
- Thom, H. C., 1958: A note on the gamma distribution. *Mon. Wea. Rev.*, **86**, 117–122, [https://doi.org/10.1175/1520-0493\(1958\)086<0117:ANOTGD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1958)086<0117:ANOTGD>2.0.CO;2).
- Venables, W. N., and B. D. Ripley, 2002: *Modern Applied Statistics with S*. Springer, 498 pp.
- Vose, R. S., and Coauthors, 2014a: NOAA's Gridded Climate Divisional Dataset (CLIMDIV). NOAA National Climatic Data Center, accessed 4 June 2018, <https://doi.org/10.7289/V5M32STR>.
- , and Coauthors, 2014b: Improved historical temperature and precipitation time series for U.S. climate divisions. *J. Appl. Meteor. Climatol.*, **53**, 1232–1251, <https://doi.org/10.1175/JAMC-D-13-0248.1>.
- Whittemore, D., J. Butler, and B. Wilson, 2016: Assessing the major drivers of water-level declines: New insights into the future of heavily stressed aquifers. *Hydrol. Sci. J.*, **61**, 134–145, <https://doi.org/10.1080/02626667.2014.959958>.
- Wu, H., M. J. Hayes, D. A. Wilhite, and M. D. Svoboda, 2005: The effect of length of record on the standardized precipitation index. *Int. J. Climatol.*, **25**, 505–520, <https://doi.org/10.1002/joc.1142>.
- , M. D. Svoboda, M. J. Hayes, D. A. Wilhite, and F. Wen, 2007: Appropriate application of the standardized precipitation index in arid locations and dry seasons. *Int. J. Climatol.*, **27**, 65–79, <https://doi.org/10.1002/joc.1371>.