JOURNAL OF APPLIED AND SERVICE CLIMATOLOGY

Drought Assessment in a Changing Climate: A Review of Climate Normals for Drought Indices

Joel Lisonbee[1], John Nielsen-Gammon[2], Blair Trewin[3], Gretel Follingstad[1], Britt Parker[4]

[1] *NOAA/National Integrated Drought Information System, and Cooperative Institute for Research in the Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, Colorado, USA*
[2] *Department of Atmospheric Sciences, Texas A&M University, College Station, TX, USA*
[3] *Australian Bureau of Meteorology, Melbourne, Victoria, Australia*
[4] *NOAA/National Integrated Drought Information System*

Corresponding Author: Joel Lisonbee, joel.lisonbee@noaa.gov

Should drought be considered an extreme dry period based on the entire record of available data? Or, should drought be considered a low in precipitation variability within the context of a present, contemporary climate? The two most common reference periods are the full period of record (all observed data or as much as possible) and a 30-year reference climatology. However, climate non-stationarity may render the "all-data" approach an inaccurate or obsolete comparison unless a trend is factored in. The aim of this review is to explore the literature for approaches to addressing these issues.

The World Meteorological Organization (WMO) has recommended a 30-year reference period for most climatological applications since 1935, but for drought assessments and drought indices the modus operandi has been to use as much data as possible. However, in the literature, the "all data" approach has been challenged by evident impacts from climate change-induced non-stationarity. Over the past several years, as potential errors in drought assessments became more apparent due to a stationarity assumption when applying drought indices, several studies have adopted shorter reference periods, with 30-years being the most common. Furthermore, several recent papers have recommended using short reference periods with more frequent data updates for drought assessments to be representative of a contemporary climate. Additionally, at least 18 non-stationary drought indices have been proposed in efforts to retain long datasets and account for non-stationarity in the climate system.

---

1. Introduction

In 1935 the World Meteorological Organization (WMO) instructed member nations to calculate climate normals using a 30-year period, from 1901 to 1930 (Guttman 1989). For most climatological applications, the 30-year reference period has been the standard. However, for drought, the modus operandi has typically been to include as much data as possible, usually the full period of record. The rationale for this is that a larger dataset reduces the sampling uncertainty and produces a more stable dataset. This is important because drought represents the extremes of climate variability at one tail of the statistical distribution. On the other hand, in regions that experience pronounced trends in the climate system, treating the distant past as "representative" means present or future drought assessments are compared to a very different climate of the past. This issue is especially significant when assessing long-term drought (over periods of multiple years, sometimes referred to as "megadroughts"), as internal variability will generally be lower over longer periods, and hence long-term trends will be larger relative to internal variability.

Several recent publications have helped frame the issue of drought assessment in a non-stationary climate (Paulo et al. 2016; Wang et al. 2021; Mi et al. 2022; Hoylman et al. 2022). For example, following their analysis of megadrought and pluvial events in climate projections over the 21st century, Stevenson et al. (2022) point out that "background trends are so large that if traditional stationary definitions are used to identify megadrought or pluvial, in those places where the trend emerges, the entire late 21st century is identified as a single large event" (Stevenson et al. 202, p.6).

These findings shed light on the broader question of how to define drought. The Glossary of Meteorology defines

Journal of Applied and Service Climatology, Volume 2024, Issue 001, DOI: doi.org/10.46275/JOASC.2024.05.001

drought as "a period of abnormally dry weather sufficiently long enough to cause a serious hydrological imbalance" (American Meteorological Society 2019). But what does "abnormally dry weather" mean in the context of a changing climate. Should drought be considered an extreme dry period based on the entire record of available data? Or, should drought be considered a dry period within the context of contemporary climate variability? Are different solutions better in different circumstances?

The aim of this review is to search the literature for established answers to the following research questions:

- What reference climatologies are used in drought assessments? What period of record does the literature support, 30-year climatology, a full period of record or another longer or shorter reference period? Are there studies on specific commonly referenced drought indices and indicators that provide a recommended reference period?

- Does the literature support adapting how drought should be monitored and assessed because of climate change*?*

2. Methods

This systematic quantitative literature review followed the methodology outlined in Pickering and Byrne (2014) along with an adaptation of the process outlined in Page et al. (2021). The topic and research questions have been outlined above. The keywords used and search results are summarized in Table 1 with the process of synthesizing these results shown in Figure 1, and detailed as follows:

As of January, 2023, A Web of Science search for the terms "Drought Assessment" and "Climate Change" produced 168 results that used these terms in the title, abstract or keywords. An additional search for the terms "Drought Assessment" and "non-stationarity" produced 6 results of which one was a duplicate from the previous search.

As described below, in Section 4.3, a portion of this review will focus on the use of the Standardized Precipitation Index (SPI) in drought assessment. A Web of Science search (title, abstract or keywords) for the terms "Standardized Precipitation Index" or "SPI" and "Climate Change" produced 3,015 results. This volume of search results indicates that this literature review will not be an exhaustive list of all papers on the topic, but rather a narrative of the history and current state of the research. To narrow down these results additional search terms were added. Adding the word "drought" in the search terms decreased this number to 802, and then filtering again by papers that mention "drought assessment" decreased the number of papers to 59. Within these results there were 31

TABLE 1: Literature review search terms and number of search results.

Search Term(s)	Search Engine	Date of search	Total
"Drought Assessment" and "Climate Change"	Web of Science (Title, Abstract and Keywords)	6 January 2023	168
"Drought Assessment" and "non-stationarity"	Web of Science (Title, Abstract and Keywords)	6 January 2023	6
["Standardized Precipitation Index" or "SPI"] and "Climate Change"	Web of Science (Title, Abstract and Keywords)	6 January 2023	3,015
["Standardized Precipitation Index" or "SPI"] and "Climate Change" and "drought"	Web of Science (Title, Abstract and Keywords)	6 January 2023	802
["Standardized Precipitation Index" or "SPI"] and "Climate Change" and "drought assessment"	Web of Science (Title, Abstract and Keywords)	6 January 2023	59
"drought"+ "reference period"+ "non-stationary"	Google Scholar (full text)	20 January 2023	884
"drought"+ "reference period"+ "non-stationary climate"	Google Scholar (full text)	20 January 2023	111

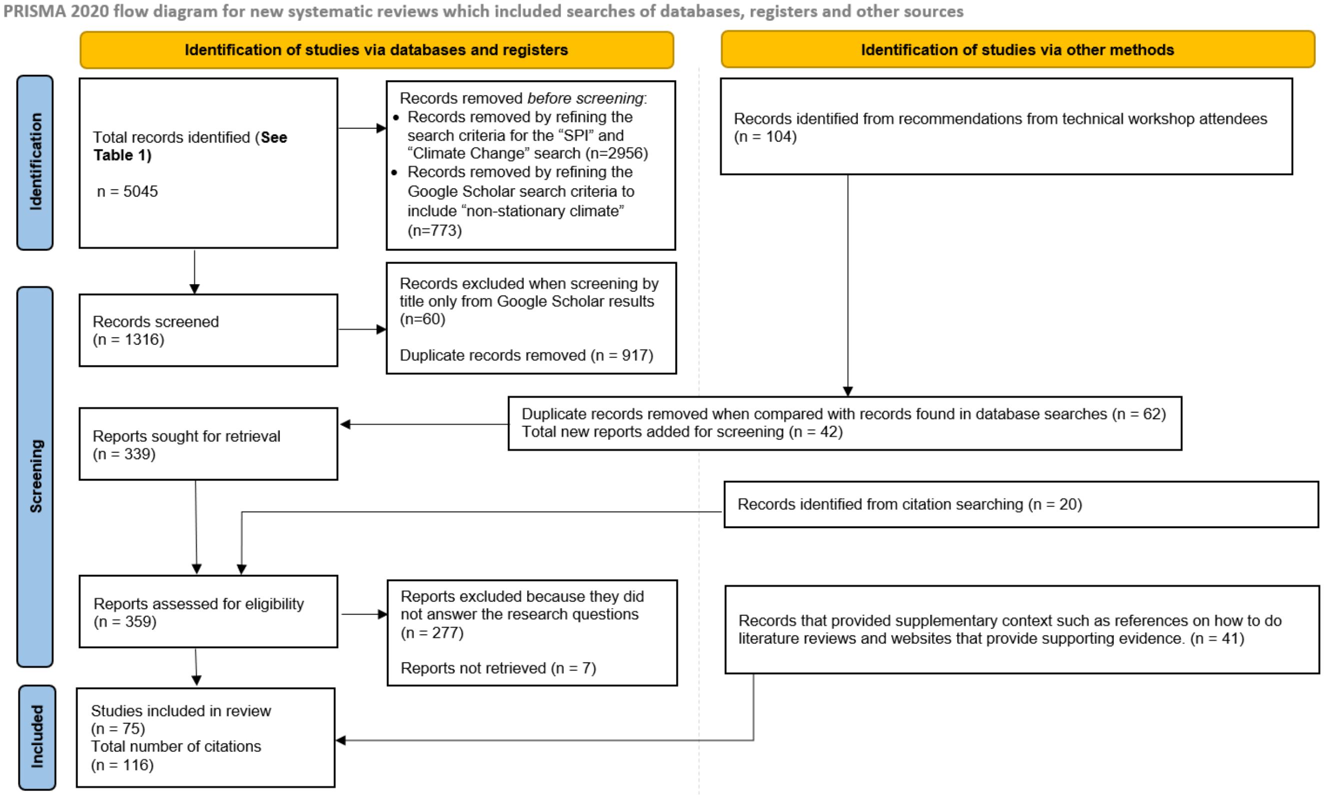

FIGURE 1: The process followed when synthesizing the literature search results (adapted from Page et al. 2021). For more information, visit: http://www.prisma-statement.org

papers that specifically mentioned "reference period".

Also in January 2023, a Google Scholar search for "drought"+"reference period"+"non stationary" produced 884 results, adding the term "climate" (i.e. "non stationary climate") and removing citations from the results reduced this to 111. These were sorted through by *title only* to remove papers that were very clearly irrelevant (i.e. material science, structural engineering, etc.). Based on titles alone there were 51 papers remaining.

On February 28th–March 1st, a technical workshop was held on the topic of drought assessment in a non-stationary climate. As attendees registered for this workshop they were asked if they were aware of any publications which addressed the specific research questions of this review. This resulted in 104 suggested papers, but nearly all were duplicates of previously identified searches.

Following these searches, duplicates were removed, and about 20 papers were added that were found as references in other papers. The final number of papers that were found on this topic is 359. These papers were screened by *title*, *abstract*, *results* and *conclusions* sections including a search through the text for keywords that would indicate if the paper answered the research questions. At the end of this process 76 papers remained to be explored at a deeper level and have been referenced in this literature review.

3. Brief Timeline of Climate Normals

When assessing if a weather pattern can be considered "abnormal" it needs to be compared to a reference. The first known use of the term "normal" to describe a comparable climate reference period in the meteorological literature was in an 1840 meteorological monograph by Heinrich Wilhelm Dove (1803-1879) (Guttman 1989). Dove's use of the term "normal" had several different contexts, but the context that survived into the late 19th century is that "normal" was equivalent to the average or mean of a long series of observations (Guttman, 1989). "Originally this designation had been used for zonal means of climate elements, but the [WMO] adopted it for temporal rather than zonal means" (Landsberg 197, p.2).

The idea gained momentum in an essay titled *Suggestions on a Uniform System of Meteorological Observations* by Professor C. H. D. Buys Ballot (1817-1890), then the director of The Royal Netherlands Meteorological Institute. What Buys Ballot initiated in this essay would eventually become the standard in meteorology and ultimately lead to the formation of an International Meteorological Organization (IMO) which was eventually succeeded by the WMO (Guttman, 1989; WMO, 2022). Guttman (1989) provides some background as to why thirty-years was chosen and some of the issues that arose right away from this decision. In 1872, the International

Meteorological Committee compiled mean values over a standard period to allow comparability between data from various stations.

> The doctrine gradually developed that climate is essentially constant during intervals that are long compared to human experience. It was assumed that long-term averages would converge to this stable value or normal. International agreements eventually led to the compromise that the appropriate interval for computing a normal would be 30 years. (Guttman 1989, p.602; see also references therein and Landsberg, 1975)

In 1935 the IMO instructed member nations to adopt a standard 30-year reference period which included the years 1901 to 1930, inclusive. In 1956 the WMO updated the reference period and established the idea of regular updates every 10 years with the 30-year reference period ending in the most recent year ending with zero (World Meteorological Organization 2007).

Historically, climatological normals have been used for two purposes. First, for comparison; "they form a benchmark or reference against which conditions (especially current or recent conditions) can be assessed" (World Meteorological Organization 2007, p.6). Second, they are used for prediction; "they are widely used for predictive purposes, as an indicator of the conditions likely to be experienced in a given location" (World Meteorological Organization 2007, p.6). While these applications of normals work well for most comparisons, they assume a stationary climate over time—essentially that the climate of today is sufficiently comparable to the climate of the past and the climate of the future. Early applications of climate normals warned about this assumption. Landsberg (1975) proposed that use of the concept of a climatological normal "did considerable harm" to the science of climatology. "One of the worst misinterpretations of the 'normal' concept was that a 'normal' value had, by itself, prognostic value for future events," (Landsberg, 1975, p.3). Guttman (1989) states, "the normals as they have been previously defined and published meet the needs of those making these kinds of comparisons. It is emphasized, however, that these comparisons imply very little about climatic change, non-random fluctuations, or extremes. They are simply an assessment of deviations from a reference" (Guttman, 1989, p.603).

Notwithstanding the WMO guidance for the use of a 30-year normal period, it has been shown that various climate normals ($N \neq 30$) can produce more accurate comparisons for various applications (Arguez and Vose 2011). While the focus of this review is on climate normals used specifically for drought assessment, a few non-drought applications of various climate normals include: a 50-year normal is ideal for the Atlantic Hurricane Season (Schreck et al. 2021); using an *optimal climate normals* technique showed the optimal normal for temperature is 10 years and for precipitation is 15 years (Huang et al. 1996; Livezey et al. 2007).

4. Drought Indices

This section reviews the introduction of climate indices specifically for drought assessment (hereafter "drought indices"; see Heim 2002; Quiring 2009; Mishra and Singh 2010; Dai 2011; Singh et al. 2022), and the period used to calculate the drought index. In this context the term "assessment" is used broadly to include drought monitoring and diagnosing—i.e. assessing when one is in a drought, and examining how extreme the individual droughts are in a historical context. It is not within the scope of this literature review to enumerate *all* the drought indices that have been produced (there are hundreds), but to look at commonly used indices and the amount of data used (reference periods, length of records, etc.) to assess drought (see Table 2).

Before enumerating some of these indices, it is worth pointing out that the selection of indices carries its own uncertainty in the overall drought assessment (Hoffmann et al. 2020; Satoh et al. 2021). Whether an index is selected for meteorological, agricultural, hydrological, socioeconomic, or ecological drought, can change the sign and magnitude of the drought assessment. This is why most studies focus on a single hydroclimate aspect. Satoh et al. (2021) and IPCC (2012) articulate this as *the issue of drought definition*; the drought definition selected for a study can be the dominant source of uncertainty within that drought assessment (Satoh et al. 2021). In fact, McColl et al. (2022) recommend moving away from drought indices altogether when interpreting climate models because (1) they are redundant, (2) many work on the assumption that they are consistent in space and time, i.e. a stationary climate, and (3) they introduce definitional ambiguity.

4.1 Some Key Drought Indices Developed Before 1990

Within this literature search, the earliest discovered reference to a drought index is by Foley (1957), which uses a cumulative precipitation anomaly based on all available data to quantitatively assess drought in Australia.

Palmer (1965) introduced what is now known as the Palmer Drought Severity Index (PDSI). The index is based on a water balance or hydrologic accounting approach to climatic analysis which allows for a calculation of the distribution of moisture excesses and deficiencies. This moisture supply-and-demand is estimated using a simple water balance model that uses temperature and precipitation as inputs and approximates the impact of potential evapotranspiration on soil moisture. This is generally accepted as the first attempt to objectively and numerically define drought. Palmer used the full record of data available; specifically, for western Kansas, USA, this was from January 1887–December 1957 (71 years); and for central Iowa, USA, January 1931–December 1957 (27 years). In addition to the PDSI in 1965, the same paper by Palmer also introduced the Palmer Hydrological Drought Index and the Palmer Moisture Anomaly Index (commonly known as the Z-Index).

Another, less well-known drought index proposed in 1965 was the Rainfall Anomaly Index (Van-rooy 1965) which measured the rainfall anomaly, as calculated using the full period of record, against a 9-member classification scheme ranging from extremely wet to extremely dry.

Gibbs and Maher (1967) introduced the use of rainfall deciles as drought indicators. Rainfall deciles are calculated using all available data. As of the writing of this review, rainfall deciles are still used operationally in Australia as a way to assess drought (Australian Bureau of Meteorology 2023).

In 1968, Palmer proposed another drought index based on the PDSI, this one specific to crops (Palmer 1968). The Crop Moisture Index is the sum of an evapotranspiration deficit (with respect to normal conditions) and soil water recharge. These terms are computed on a weekly basis using PDSI parameters, which consider the mean temperature, total precipitation, and soil moisture conditions from the previous week. This index also uses the full period of record to calculate a climate normal.

Another commonly used drought index proposed in 1968 is the Keetch-Byram Drought Index (Keetch and Byram 1968). Keetch–Byram drought index is a soil moisture deficit indicator usually used in fire risk assessment. It requires mean annual rainfall for the index calculation. When this index was first introduced, Keetch and Byram (1968) used all available data at that time.

The commonly used Aridity Index was introduced in 1977 (UNESCO 1977), more to assess which climates are considered arid and less for diagnosing the occasional lows in precipitation variability. The Aridity Index is calculated simply as the precipitation divided by the potential evapotranspiration over a given time period (usually annually) at a given location or broader region. The original proposal for the Aridity Index averaged the annual precipitation and potential evapotranspiration over all available data for a location, but the index has been applied to shorter periods to establish changes in aridity over time (Greve et al. 2019).

Two new drought indices were introduced in 1980, but with specific hydrologic applications. These were the (Hydrologic) Total Water Deficit (Dracup et al. 1980) and the Drought Area Index (DAI, Bhalme and Mooley 1980). The Total Water Deficit is calculated as the duration of drought multiplied by the average departure from "normal" within that duration. The DAI was developed as a method to improve understanding of monsoon rainfall in India, determining both flood and drought episodes using monthly precipitation (Bhalme and Mooley 1980). Both used a full period of record, but could reasonably be calculated with a shorter reference period.

The Surface Water Supply Index, introduced by Shafer and Dezman (1982), is calculated by river basin based on snowpack, streamflow, precipitation, and reservoir storage using a Principle Component Analysis based on the full period of record at a location. This index classifies drought using normalized values in a scale similar to the PDSI. This is one example of how ingrained the use of the PDSI had become in the early 1980s.

The Soil Moisture Anomaly Index (Bergman et al. 1988) is another common drought index which requires "normal" precipitation and soil moisture values to assess drought. Bergman et al. (1988) did not define "normal".

4.2 Criticism of the Palmer Drought Severity Index

By the early 1980's the PDSI had become widely used in drought assessments and generally applied as a baseline or standard for comparing other drought indices and the practice of using the PDSI as a standard for drought index comparisons has continued with more recent papers (e.g. Bhalme and Mooley 1980; Cook et al. 1999; Dai et al. 2004; Heim 2002; Vicente-Serrano et al. 2010; Dai 2011; Ma et al. 2014; Gamelin et al. 2022). But this index was not without its problems.

Alley (1984) is very critical of the PDSI, suggesting most of the assumptions made are not physically sound and most of the thresholds set are "arbitrary", but did not address non-stationarity as a specific weakness of this index.

Karl (1986) investigated the PDSI and, using a comparison of 1931-1960 and 1951-1980 and a full record of 1895 -1983, discovered that the PDSI is highly sensitive to the base period used to calculate it. Presumably the inclusion/exclusion of the very dry 1930s in much of the United States is relevant to the findings. "By changing the base period used to calibrate the coefficients, the magnitude and sign of the PDSI change significantly in many areas of the United States" (Karl, 1986, abstract, p. 77). Karl

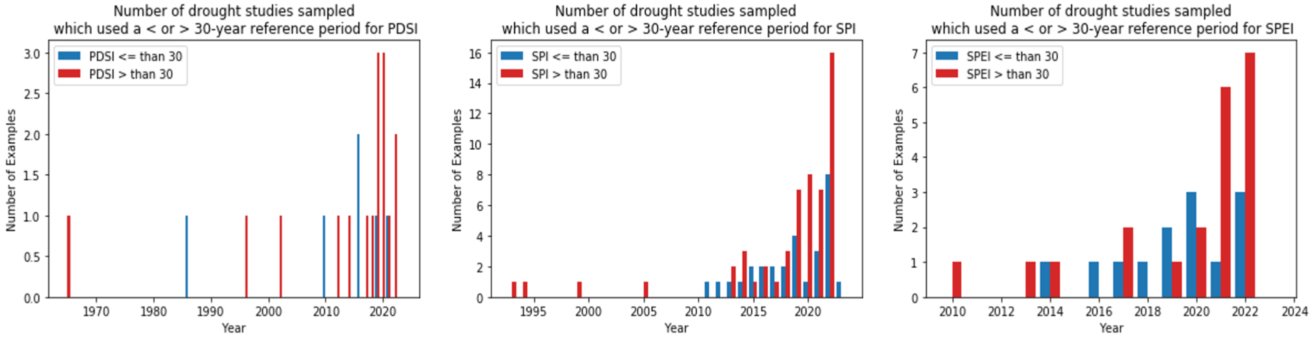

FIGURE 2: Of the subset of papers identified for this review, these graphs show the count of papers by year for (a) the PDSI, (b) the SPI and (c) the SPEI that used a reference period of 30 years or less compared with those that used a reference period of more than 30 years or used the full period of record.

(1986) provided a similar comparison with other Palmer (1965) indices and suggested that the Palmer Moisture Anomaly Index is less sensitive to reference periods, while the Palmer-Z index is the least sensitive. Within the scope of this literature review, this is the earliest reference that challenged the stationarity assumption, and the first time a shorter reference period is tested for drought assessment.

Heddinghaus and Sabol (1991) came to a similar conclusion as Alley. They further criticized the index for being discontinuous from one dry spell to the next. The PDSI as formulated by Palmer is measured from the beginning of a dry period to the end. However, it is very difficult to know whilst in a drought if a rainfall event marks the end of the drought or if it is simply a short reprieve within the longer dry period. “Problems thus arise in using the PDSI as an operational index since it may not be known until a later date which spell [wet or dry] the PDSI is really in” (Heddinghaus and Sabol, 1991, p. 243).

Figure 2a shows that of the small subset of papers examined in this review, the PDSI was calculated using a full period of record 16 times compared to a 30-year period only 6 times.

4.3 Standardized Precipitation Index (SPI)

The SPI was first introduced by McKee et al. (1993), and it quickly became a very popular drought indicator that is widely used and studied. A few reasons for its popularity are that it is easy to calculate, relatively easy to understand and interpret, and input data (precipitation, either in situ, remote sensed or modeled) is easily available for most regions of the world. Special attention will be given to the SPI in this literature review for two reasons. First, in 2009 at the *Interregional Workshop on Indices and Early Warning Systems for Drought* held in Lincoln, Nebraska, USA, the SPI was recommended as the drought index to be used globally among national meteorological and hydrological services. This became known as “the Lincoln Declaration” (Hayes et al. 2011). Second, most of the studies that challenge the stationarity assumption in drought assessment use the SPI as a starting place for their argument.

A specific advantage the SPI has for multi-national drought assessments is that the only data it requires is total monthly precipitation, which is a variable routinely exchanged by WMO members through monthly CLIMAT messages (although coverage is still limited in many parts of the world, particularly Africa). Other drought indices often require additional variables and/or data for shorter time periods, for which historical and current data are often more difficult to obtain across international borders.

“The SPI is, conceptually, simply the precipitation anomaly divided by the Standard deviation where the mean and standard deviation are determined from past records” (McKee et al., 1993, Section 2.0).

In practice, however, there are additional steps that must be taken. Precipitation doesn’t fit a normal (or gaussian) distribution—partly because precipitation cannot have a negative value and partly because precipitation is skewed: below-average precipitation tends to be more common than above-average precipitation. This is most pronounced at shorter timescales and in arid or semi-arid climates. Precipitation distributions are most often estimated using a gamma distribution—although Guttman (1999) suggests a Pearson-III distribution is ideal, Lana et al. (2001) used the Poisson-gamma distribution, and Lloyd-Hughes and Saunders (2002) explored a log-normal distribution for precipitation when calculating the SPI. The SPI is calculated over summary periods such as 1, 3, 6, 12, 24, or 48 months (sometimes described as chunks of daily precipitation, e.g. 30-day SPI, 90-day SPI and so forth). For the chosen time period, the data are fitted to the Gamma function to define the relationship of probability to precipitation. Next, an estimate of the inverse normal can be used "to calculate the precipitation deviation for a normally distributed probability density with a mean of zero and standard deviation of unity" (McKee et al., 1993, Section 2.0).

4.3.1 What reference period should be used for the SPI?

When McKee et al. (1993) first proposed the SPI, they recommended using a monthly precipitation dataset with "ideally a continuous period of at least 30 years" (McKee et al., 1993, Section 2.0). They further discussed how decision makers use hydrologic data "as a percent of average using recent climatic history (the last 30 to 100 years)", hence the goal of developing a drought index within that comparative reference period (McKee et al., 1993, Section 1.0). In the case study used within McKee et al. (1993) they calculated the SPI for Fort Collins, Colorado, USA, using 1889-1991 for the reference period, which represented the whole period of record to that time.

Guttman (1994) evaluated the statistics of a stable gamma distribution and showed that 60-70 years of data is ideal to estimate the tails of the distribution and smaller sample sizes create more unreliable probabilities in the cumulative probability function. "Mean departures decrease as the sample size increases with values near zero generally occurring with about 30 to 40 or more observations for the central tendency measure, about 40 to 50 or more for the dispersion measure, and about 60 to 70 for the skewness and kurtosis measures" (Guttman, 1994, Abstract; see also Guttman, 1999).

Wu et al. (2005) also tested the statistics of the skewness and kurtosis of the gamma distribution and confirmed that longer datasets are required for robust statistics. "The longer the length of record used in the SPI calculation, the more reliable the SPI values will be, especially for long-time-scale SPI values" (Wu et al., 2005, p. 518). However, Wu et al. (2005) also acknowledge within the conclusions of the paper that this result assumes a stationary climate. A shorter record length would be more appropriate where strong climate trends exist, although it would also lead to instability of the parameter estimates (Wu et al. 2005).

In 2013 the WMO published an SPI user guide (World Meteorological Organization 2012) which quotes from Guttman (1994) stating that 20-30 years of monthly data is the minimum data requirement, but 50-60 years, or more, are "optimal and preferred" (WMO, 2012, p.3). Longer SPI averaging periods may require even longer datasets. "Unless one has 80–100 years of data, the sample size is too small and the statistical confidence of the probability estimates on the tails (both wet and dry extremes) becomes weak beyond 24 months" (WMO, 2012, pp.6-7).

By the time of the WMO report, there was inconsistency among publications in the record length to calculate the SPI, and no standard reference period for the calculation of standardized drought indices had been promoted (Núñez et al. 2014). Some studies used the full record while others used a 30-year reference period, and others used something different altogether, although this was usually justified in the methods. Faergemann (2012) strongly recommends the use of a 40-year reference period for the calculation of the SPI. Stagge et al. (2015) and Stagge et al. (2017) recommended the use of a 30-year reference period. Carbone et al. (2018) found that extreme events have a large influence on the SPI and recommends that "record lengths of 60–70 years typically result in stable parameters and representative SPI values" (Carbone et al., 2018, p.615). They also found that "in general, these parameters remain relatively consistent through time when the reference period exceeds 60 years" (Carbone et al., 2018, p.611). From these examples it is concluded that there has not been a consistent length of record applied to SPI calculations.

Historically, two approaches have been taken for normalizing precipitation data for the SPI calculation: one is to normalize the dataset using a subset of the larger dataset that is more representative of a contemporary climate and the other is to use the full available record. Stagge and Sung (2022) point out there are inherent problems with both approaches. Using a shorter reference period, generally 30-years, precludes comparison with reference periods different from the one used to standardize the data, and "using a 30-yr subset…decreases the record length, thereby increasing parameter uncertainty" (Stagge and Sung, 2022, p.762). The second approach assumes all data over time fit the same distribution and ignores any long-term trends. "Further, this approach centers the SPI on the middle year of the record. If two studies were to use different record length, for example as new data became available, each would be centered on a different year, making comparisons across studies more challenging" (Stagge and Sung, 2022, p.762).

A few recent studies have tested the impact of reference period lengths on the SPI. Most notable are Paulo et al. (2016), Wang et al. (2021), Mi et al. (2022), and Hoylman et al. (2022).

Paulo et al. (2016) divided the full-record (1863-2007) into 30-year chunks to calculate the SPI and test different reference periods. They showed that the choice of reference period can have a substantial effect on SPI values and a longer record can "mask", or obscure, periods that would be considered precipitation deficits/surplus when using a more contemporary reference period. They recommend using a shorter reference period that is regularly updated when there are persistent or cyclic changes in the precipitation.

Wang et al. (2021) varied the size and scale parameters within the gamma distribution and observed the change in confidence interval width with varying data sizes. They conclude that "from the perspective of reducing uncertainty, the optimal record length for calculating SPI is about 70 years" (Wang et al., 2021, pp.1381-1382). Based on these results, Wang et al. (2021) further note that there are tradeoffs between reducing uncertainty by lengthening the reference period and assessing droughts in the future based on a climate normal period far away (temporally) from its reference climate (Wang et al., 2021).

Mi et al. (2022) present a new Ensemble Drought Index (EDI) based on integrating the common components of seven single drought indices including the SPI. They conclude that reference period is one aspect that leads to "variability in the statistical characteristics of the drought indices, which was mainly caused by non-stationary observation data series" (Mi et al., 2022, Section 5). They recommend that drought assessments should be done using long climate records without a significant trend.

Hoylman et al. (2022), point out that drought assessment has been outpaced by climate change. Looking specifically at SPI error (deviations from a modeled, stationary climate) the non-stationary simulations using observed 30-year moving windows confirm that contemporary SPI error "does not necessarily decrease given a longer climatology in locations where climate change velocities are high" (Hoylman et al., 2022 p.5).

From the small subset of papers examined in this literature review, the SPI was calculated using either a full period or a 60- 70-year reference period (Guttman 1994) 54 times and a 30-year SPI was used 29 times (Figure 2b). The push for shorter reference periods and alternative approaches to calculating the SPI began in the early 2010's.

4.4 Other Standardized Indices

Following the SPI there was a rise in standardized drought indices. These are usually indices that map the input data onto a normal distribution and divide the anomaly by the standard deviation to provide indices that can be compared with each other and over space and time (with caveats). Some of these include: the Standardized Runoff Index from Shukla and Wood (2008); the SPEI (Vicente-Serrano et al. 2010); the Standardized Soil Moisture Index (AghaKouchak 2014); a Standardized Palmer Index for hydro-meteorological use (Ma et al. 2014); a Standardized Precipitation Anomaly Index (Chanda and Maity 2015); and a standardized vapor pressure deficit drought index (SVDI; Gamelin et al. 2022). All of these indices categorize drought using the same or similar scale proposed by McKee et al. (1993). Namely, drought is considered mild when the index falls between 0 and -0.99, -1 to -1.49 is considered moderate drought, -1.5 to -1.99 is severe drought and anytime the index shows drought less than -2—or more than 2 standard deviations below the mean—this is considered an extreme drought (the SVDI is inverted such that positive values indicate drought conditions). The US Drought Monitor has used these values in conjunction with other drought indices to establish a drought severity index as follows: SPI values from -0.50 to -0.79 indicate abnormally dry (or a drought category of D0), -0.80 to -1.29 indicate moderate drought (D1), -1.3 to -1.59 indicate severe drought (D2), -1.60 to -1.99 indicate extreme drought (D3) and -2.00 or less indicates exceptional drought (D4).

All of these standardized indices assume a stationary climate and all but the SVDI (which used 1990-2012 reference period) use the full period of record to calculate the index. Many of these standardized indices have received similar criticism as the PDSI and the SPI, for example, Bartholomeus et al. (2014) noted that errors are introduced into the 6-month SPEI calculation based on the calibration period used.

Unlike the SPI, which considers only precipitation, many of these other indices attempted to incorporate a temperature component either implicitly (e.g. through evapotranspiration) or explicitly. One would expect indices which include temperature to show a stronger long-term trend in most climates, when compared to indices that use precipitation alone. A notable reference on this topic is IPCC (2012), their Section 3.5.1 and inset Box 3-3, "The Definition of Drought". This section points to what it calls "the issue of drought definition" and discusses the implications of using a drought index based on precipitation only compared to an index utilizing a temperature component (IPCC, 2012; Satoh et al., 2021). The IPCC's Fourth Assessment Report (IPCC 2007) included drought assessments for a changing climate that mostly drew from multivariate drought indices (primarily the PDSI), which incorporated a temperature signal. The subsequent Fifth Assessment Report (IPCC 2014) used a broader range of literature, which included the PDSI (and similar indices) but also included other indicators which use precipitation only. While using a broader range of drought indices would strengthen the assessment of drought in a stationary climate, the inclusion of precipitation-only drought indices for locations experiencing strong temperature trends may weaken the accuracy of the drought assessment. Another example of this is Vicente-Serrano et al. (2010), which proposed the SPEI. They showed fairly good agreement between the SPI and the

SPEI where a strong temperature trend was not evident but there were significant differences where temperatures increased over the analysis period (Vicente-Serrano et al., 2010, see their Figure 12, see also Vicente-Serrano et al. 2012). Stagge et al. (2015) applied a 30-year reference period to test various precipitation distribution differences between the SPEI and the SPI. Stagge et al. (2017) used the divergence between SPI and SPEI to show that climate change is affecting drought analysis in Europe. These represent a few examples that specifically pointed to the inclusion of temperature (or temperature derived) variables within drought assessment and the divergence from precipitation-only based drought indices.

It could be assumed that drought indices that include a temperature component (e.g. the SPEI) would be more susceptible to non-stationarity due to the underlying upward trend in global temperatures, and that research using these indices would therefore tend to use a shorter reference period to represent a contemporary climate. Figure 2c shows that this is not necessarily the case. Of the small sample of papers examined in this literature review the SPEI was used with a full period of record 21 times compared to a 30-year reference period 13 times.

Table 2 includes a small subset of common drought indices and notes which environmental variables they are derived from and the reference period originally used or recommended. From this small subset of drought indices there is not a consistent pattern among drought indices that use temperature-based variables and those that use a shorter reference period. However, many of these indices are several decades old and may have encouraged the use of a full record when originally proposed, but have been adapted to shorter reference periods, for use in a warmer climate (a few examples of papers that applied a 30-year reference period to the SPI and/or SPEI include: Meresa et al. 2016; Stagge et al. 2017; Mitra et al. 2018; Leng et al. 2020).

To summarize this section, the examination of common drought indices showed that when the majority of these indices (all indices prior to 1990) were introduced they were established using the full period of record. The rationale for this approach was that longer time series would provide more accuracy in representing the tails of the probability distribution, and more stable and reliable statistics. As climate change became more evident and mainstream within the scientific literature, the stationarity assumption of some well-established climate indices began to be questioned. Perhaps the most popular of these is the SPI, which is highly sensitive to the reference period length in a non-stationary climate (Russo et al. 2013; Núñez et al. 2014; Stagge et al. 2015; Paulo et al. 2016; Rashid and Beecham 2019a; Park et al. 2019; Shiau 2020; Song et al. 2020; Cammalleri et al. 2022; Hoylman et al. 2022).

TABLE 3: References and methodologies for calculating a non-stationary SPI

Reference	Novel non-stationary SPI methodology
Rashid and Beecham (2019)	Describes a matrix of covariates as input to the gamma distribution which considered large-scale climate variability to be sources of non-stationarity in addition to climate change
Park et al. (2019)	Considers not only the existing probability distribution parameter but also the return period
Shiau (2020)	Showed how gamma-distribution variations impact the SPI-based stationary and non-stationary drought analyses
Das et al. (2021)	Incorporated large-scale climatic oscillations as covariates in the location parameter of the gamma distributions

5. Statistical Methods to Account for Non-stationarity in Drought Indices

Many papers point to the climate stationarity assumption as a problem, and support adapting how drought is assessed because of climate change (the second research question of this review). The works of Landsberg in the mid-20th century questioned the use of climatological normals because of known trends and very-long time scale oscillations in the climate system (Landsberg 1975). Matalas (1997) questioned hydrologic flood models in a changing climate, speculating that increased temperatures would likely create hydrologic trends which are unaccounted for by the models. Koutsoyiannis (2006) showed that measurable uncertainty arises when assuming trends in runoff will always persist. As cited above, Hoylman et al. (2022) and Stevenson et al. (2022) provide more recent examples of how the assumption of climate stationarity produces errors in drought indicators and indices used in drought assessments.

Many studies such as those discussed above have accounted for non-stationarity by adjusting the reference period used. The following studies have proposed some alternative approaches that maintain the full period of record.

Several papers focused on alternative ways to calculate the SPI to account for non-stationarity. Türkeş and Tatlı (2009) Introduced a time-varying SPI that considers the local-time means of the index and fits an upper and a lower envelope to precipitation. Dubrovsky et al. (2009) developed a relative SPI (rSPI) and relative PDSI (rPDSI)

TABLE 4: References and methodologies for calculating non-stationary drought indices

Index	Reference	Short description of methodology
Non-stationary Meteorological and Hydrological Drought Index	Zhang et al. (2021)	Uses climatic and anthropogenic indices as covariates, estimated by a non-stationary joint distribution model. The proposed index, entitled the Non-stationary Meteorological and Hydrological Drought Index considers the non-stationarity of precipitation and runoff, and their dependence structure" (Zhang et al., 2021, p.12)
Applied to the Standardized Streamflow Anomaly Index	Dutta and Maity (2021)	Propose a time-varying approach to drought assessment using a Bayesian Model Averaging based temporal network approach
Non-stationary Standardized Runoff Index	Wang et al. (2022)	Comprehensively considers the effects of climate change and human activities on runoff variability" (Wang et al., 2022, p.2435)
Applied to precipitation and temperature deciles.	Hughes et al. (2022)	Calculates deciles with a moving reference period and applies these to a farm profit model
Non-stationary Reconnaissance Drought Index	Bazrafshan and Hejabi (2018)	GAMLSS
Applied to the Fixed Runoff Threshold Level method and the Standardized Runoff Index	Jiang et al. (2019)	Established a separation framework including the variable runoff threshold level method and standardized runoff index based on a parameter transplantation method to quantify the impacts of climate change and human activities on hydrological drought

for climate change impact assessments. Russo et al. (2013) adapted the relative SPI approach from Dubrovsky et al. (2009) and expanded on the application. In these cases, the term "relative" is in comparison to a reference weather series in which the gamma distribution for precipitation was calculated at a location for two reference periods representing the present and a future climate. These curves were then compared to each other to calculate the probability differences between the two climate periods. Russo et al. (2013) also developed the Standardized Non-stationary Precipitation Index, which "is defined like the relative SPI but using a non-stationary Gamma distribution to transform the precipitation time series into corresponding time series of probability values" (Russo et al., 2013, p.7631). According to Russo et al. (2013) the advantage of this approach over the traditional SPI "is that it is able to model the entire time series without splitting the data into shorter periods. Fitting precipitation data to a non-stationary model is done by linearly varying the scale parameter of the Gamma distribution with time" (Russo et al., 2013, p.7631). They found this approach to be "more robust" at describing precipitation changes in a non-stationary climate when compared to the SPI.

Following Russo et al. (2013), several studies also took a non-stationary approach to calculating the SPI, some following the Russo et al. (2013) method (e.g., Chanda and Maity 2015; Salvi and Ghosh 2016; Song et al. 2020, others using adaptations. Novel adaptations used to calculate a non-stationary SPI that were found in this literature review are listed in Table 3 and the reader is directed to these references for details about the methodology of these approaches.

Villarini et al. (2010) applied a Generalized Additive Model in Location, Scale and Shape (or GAMLSS) approach to model rainfall and temperature, where trends were known to exist but assumed to change. GAMLSS is a statistical modeling and learning technique that can describe data that don't fit the simple linear regression model. It is a tool for modeling and predicting data that are complex, with multiple factors influencing the outcome, and where more insight from each variable is needed than just the central tendency (e.g., the spread and shape of the distribution). A GAMLSS approach, or a variation of it, has been applied to drought analysis, usually to fit precipitation data to a non-stationary gamma distribution (Wang et al. 2015; Bazrafshan and Hejabi 2018; Zou et al. 2018; Rashid and Beecham 2019b; Das et al. 2020, 2021; Wang et al. 2022). Several studies have specifically applied the GAMLSS methodology toward calculating a non-stationary SPI (Shiau 2020; Jehanzaib et al. 2021; Stagge and Sung 2022; Blain et al. 2022). These studies have shown that series in excess of 30 years can introduce errors in the statistics when stationarity is assumed in a non-stationary dataset (Jehanzaib et al. 2021). These compare and contrast the Generalized Linear Model for a non-stationary SPI proposed by Russo et al. (2013) to the Generalized Additive Model (Shiau 2020; Stagge and Sung 2022). These studies also show the relative weighting of covariates in implementing a non-stationary gamma distribution, including the use of time as a covariate (Blain et al. 2022).

Cammalleri et al. (2022) is critical of current non-stationary SPI approaches, citing that these methods are too complex to be applied operationally. They further find that when using a 30-year window, an update frequency of 10 years may not be enough to keep pace with climate change and they recommend an update cycle of every 5 years (Cammalleri et al. 2022; see also Stevenson et al. 2022).

In addition to the SPI, non-stationary adaptations have been made to other drought indices (Zhang et al. 2021; Dutta and Maity 2021; Wang et al. 2022; Hughes et al. 2022; Bazrafshan and Hejabi 2018; Jiang et al. 2019). The creation of non-stationary drought indices is to allow for the inclusion of complete data records while accounting for trends within the data. The proposed index, references and short description of the methodologies for new or adapted drought indices that were found in this literature search are listed in Table 4. The reader is directed to these references for details about the methodology of these approaches.

Other methods, although not yet applied in the literature to drought, have been used to account for temperature trends in the assessment of other climate extremes and have the potential to be adapted to drought. While shifting reference periods have been used as one tool to account for temperature trends in assessing the probability of events such as heatwaves, short reference periods may produce biased estimates (Zeder et al. 2023).

Finally, a few studies have considered an analogous problem, non-stationary aspects of extreme heavy precipitation. A variety of approaches have been developed in recent years and have been reviewed by Salas et al. (2018), Martel et al. (2021), Wasko et al. (2021), Yan et al. (2021), and Schlef et al. (2023). The articles by Yan et al. (2021) and Schlef et al. (2023) identify several different types of strategies. One strategy is to perform a traditional stationary analysis using the full period of record, but with frequent updates. Another strategy, which Schlef et al. (2023) call “simulated precipitation”, involves stationary analysis on shorter records, usually 30 years, with trends in probability assessed by interpolating between the 30-year analysis blocks from historical data and climate model simulations. A third strategy is a non-stationary analysis of all or part of the data record with time or a climate-related variable as a covariate. All three of these types of approaches have been discussed above in the drought assessment context. A fourth strategy involves using basic physical principles to estimate trends in frequency, usually by applying the Clausius-Clapeyron equation to add trends to historical estimates (Allen and Ingram 2002). Since drought lacks a simple physical expectation for the magnitude of its rate of change in a warming climate, this approach is not directly transferable to the drought context.

To summarize this section, in addition to shorter reference periods, there has been a recent (since 2013) surge in research regarding non-stationary drought indices that would retain the full period of record, but build into the assessment some mechanism to account for climatological trends. As of January 2023, there have been at least 18 non-stationary drought indices proposed. However, these require more complex statistics and would be difficult to implement in an operation drought monitoring system (Cammalleri et al. 2022).

6. Discussion

This literature review articulates the balancing act between having a period which represents the current climate, versus having a sufficiently large sample that characterizes the tails of the distribution. Importantly, there are distinct differences in drought assessment based on the choice of reference period.

To illustrate this, suppose a certain hypothetical location had 380 mm of precipitation in its driest year during the period 1900-1990. The same location then had three years below 355 mm from 1991-2020, and other data confirms that average rainfall has also declined during that period. In 2021 this location received 355 mm. How severe was that drought?

When using the full period of record and assuming a stationary climate, 355 mm at this location would be the fourth-lowest annual precipitation total in 120 years, so just below the 3rd percentile. Based on the US Drought Monitor Categories this could be assessed as an extreme, or D3, drought (Svoboda et al. 200, see their Table 1).

From the point of view of a short (e.g. 30-year) stationary drought assessment, 355 mm would be the fourth-lowest in 30 years, near the 9th percentile or severe (D2) drought according to the US Drought Monitor (Svoboda et al., 2002, see their Table 1).

If the full record was considered, but a non-stationary climate was accounted for within the drought assessment (see Section 5), there might be evidence that the climate is trending drier, demonstrating that it is even drier now that it was 30 years ago. Depending on the tools used for assessing the non-stationarity it would be reasonable to suggest that 355 mm should not be considered unusual or unexpected anymore; perhaps around the 10th to 20th percentile, or moderate (D1) drought.

In this hypothetical scenario three different approaches arrive at three different drought severities. It is the opinion of the authors that the "correct" period of record used to assess drought should be tailored to the reason for the drought assessment. One should consider why the drought is being assessed in the first place. For example, to assess climatological extremes a full period of record, and even a paleo-data record, could be considered. If the objective was to assess the potential impact of the drought on some system, in this case the reference period should align with the reference period for which the system has been designed. For example, if farmers typically take major weather events of the last few decades into account when growing crops, the reference period for drought assessment ought to be the last few decades. Finally, to assess the actual risk of an impactful extreme event occurring in the context of a present-day climate, the practitioner may consider using non-stationary statistics to more fully sample the tails of the distribution, while still accounting for a shift in the frequency and severity of extreme events.

7. Summary and Concluding Remarks

This literature review focused on addressing questions of climate non-stationarity and better understanding appropriate reference climatologies for drought indices, indicators and assessments, in the context of changing climate trends. The literature shows that drought assessments have typically used all available data, i.e. the full record, but occasionally a 30-year "normal" has been used in accordance with WMO guidelines. A few studies have used different reference periods, but these usually justified why they deviated from either of the more common approaches.

The topic of drought assessment in a non-stationary climate is broad in scope as evidenced by the large number of works on this topic that could not be included in this review. A few limitations to the current literature assessment should be considered. First, climate non-stationarity can describe changes in climate variability and extremes at a location even if the mean at a location remains unchanged. The broader literature search resulted in a variety of manuscripts about changing climate variability and drought assessment that were not included in this review. If the reader is interested in this aspect of the broader topic the following references may provide a good starting point: Gutzler and Robbins (2011), Bartholomeus et al. (2014), Cancelliere (2017), Lian et al. (2021), Marvel et al. (2021). Another area that was outside the scope of this literature review is the capacity of drought indices to produce sensible results in climates with large seasonal variations in precipitation, such as many tropical locations. While examining literature assessing drought indices in a non-stationary climate, only temporal changes were considered when assessing drought indices and not spatial changes. Finally, this review also highlights the need for further research in understanding the bounds of drought and effectively determining what constitutes the end of drought in an aridifying climate. Related to this is an investigation of compounding drought impacts (e.g. drought and wildfire) and how to separate the impacts of periodic drought from a multi decadal "megadrought" and from permanent climate change.

To conclude, drought assessment in a changing climate is difficult and prone to misinterpretation. What approaches should be taken to incorporate non-stationarity into drought assessment? This literature review has shown that several approaches have been tested to address the problem of non-stationarity, each with its strengths and weaknesses. The underlying questions are: When should drought be defined using all available data (treating the data as stationary), when should drought be defined using a shorter reference period (treating the shorter period as stationary), and when should drought be defined using all data but accounting for non-stationarity? Within the literature surveyed there is not yet consensus or a coalescence toward a singular approach, suggesting a need for continued research, discussion and collaboration which considers both the changing drought impacts and the application timeframe of decision makers. The references included herein should provide some guidance as to the history, options and science to support those decisions.

References

AghaKouchak, A., 2014: A baseline probabilistic drought forecasting framework using standardized soil moisture index: Application to the 2012 United States drought. *Hydrol. Earth Syst. Sci.*, **18**, 2485–2492, https://doi.org/10.5194/hess-18-2485-2014.

Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, https://doi.org/10.1038/nature01092.

Alley, W. M., 1984: The Palmer Drought Severity Index: Limitations and Assumptions. *J. Clim. Appl. Meteorol.*, **23**, 1100–1109, https://doi.org/10.1175/1520-0450(1984)023<1100:TPDSIL>2.0.CO;2.

American Meteorological Society, A., 2019: AMS Glossary: drought. https://glossary.ametsoc.org/wiki/Drought (Accessed March 31, 2023).

Arguez, A., and R. S. Vose, 2011: The definition of the standard WMO climate normal: The key to deriving alternative climate normals. *Bull. Am. Meteorol. Soc.*, **92**, 699–704, https://doi.org/10.1175/2010BAMS2955.1.

Australian Bureau of Meteorology, 2023: Drought: Rainfall Deficiencies and Water Availbility. http://www.bom.gov.au/climate/drought/ (Accessed February 27, 2023).

Bartholomeus, R. P., J. H. Stagge, L. M. Tallaksen, and J. P. M. Witte, 2014: How over 100 years of climate variability may affect estimates of potential evaporation. *Hydrol. Earth Syst. Sci. Discuss.*, **11**, 10787–10828, https://doi.org/10.5194/hessd-11-10787-2014.

Bazrafshan, J., and S. Hejabi, 2018: A Non-Stationary Reconnaissance Drought Index (NRDI) for Drought Monitoring in a Changing Climate. *Water Resour. Manag.*, **32**, 2611–2624, https://doi.org/10.1007/s11269-018-1947-z.

Bergman, K. H., P. Sabol, and D. Miskus, 1988: Experimental indices for monitoring global drought conditions. *Thirteen. Annu. Clim. Diagnostics Work.*, 190–197.

Bhalme, H. N., and D. A. Mooley, 1980: Large-Scale Droughts/Floods and Monsoon Circulation. *Mon. Weather Rev.*, **108**, 1197–1211.

Blain, G. C., G. da Rocha Sobierajski, E. Weight, L. L. Martins, and A. C. F. Xavier, 2022: Improving the interpretation of standardized precipitation index estimates to capture drought characteristics in changing climate conditions. *Int. J. Climatol.*, **42**, 5586–5608, https://doi.org/10.1002/joc.7550.

Cammalleri, C., J. Spinoni, P. Barbosa, A. Toreti, and J. V. Vogt, 2022: The effects of non-stationarity on SPI for operational drought monitoring in Europe. *Int. J. Climatol.*, **42**, 3418–3430, https://doi.org/10.1002/joc.7424.

Cancelliere, A., 2017: Non Stationary Analysis of Extreme Events. *Water Resour. Manag.*, **31**, 3097–3110, https://doi.org/10.1007/s11269-017-1724-4.

Carbone, G. J., J. Lu, and M. Brunetti, 2018: Estimating uncertainty associated with the standardized precipitation index. *Int. J. Climatol.*, **38**, e607–e616, https://doi.org/10.1002/joc.5393.

Chanda, K., and R. Maity, 2015: Meteorological Drought Quantification with Standardized Precipitation Anomaly Index for the Regions with Strongly Seasonal and Periodic Precipitation. *J. Hydrol. Eng.*, **20**, 1–8, https://doi.org/10.1061/(asce)he.1943-5584.0001236.

Cook, E. R., D. M. Meko, D. W. Stahle, and M. K. Cleaveland, 1999: Drought reconstructions for the continental United States. *J. Clim.*, **12**, 1145–1163, https://doi.org/10.1175/1520-0442(1999)012<1145:drftcu>2.0.co;2.

Dai, A., 2011: Drought under global warming: a review. *WIREs Clim. Chang.*, **2**, 45–65, https://doi.org/10.1002/wcc.81.

——, E. T. Kevin, and Q. Taotao, 2004: A Global Dataset of Palmer Drought Severity Index for 1870 – 2002: Relationship with Soil Moisture and Effects of Surface Warming. *J. Hydrometeorol.*, **5**, 1117–1130.

Das, J., S. Jha, and M. K. Goyal, 2020: Non-stationary and copula-based approach to assess the drought characteristics encompassing climate indices over the Himalayan states in India. *J. Hydrol.*, **580**, 124356, https://doi.org/10.1016/j.jhydrol.2019.124356.

Das, S., J. Das, and N. V. Umamahesh, 2021: Nonstationary Modeling of Meteorological Droughts: Application to a Region in India. *J. Hydrol. Eng.*, **26**, 1–14, https://doi.org/10.1061/(ASCE)HE.1943-5584.0002039.

Doesken, N., T. B. McKee, and K. John, 1991: *Development of a Surface Water Supply Index for the Western United States*. 1–69 pp.

Dracup, J. A., K. S. Lee, and E. G. Paulson, 1980: On the definition of droughts. *Water Resour. Res.*, **16**, 297–302, https://doi.org/10.1029/WR016i002p00297.

Dubrovsky, M., M. D. Svoboda, M. Trnka, M. J. Hayes, D. A. Wilhite, Z. Zalud, and P. Hlavinka, 2009: Application of relative drought indices in assessing climate-change impacts on drought conditions in Czechia. *Theor. Appl. Climatol.*, **96**, 155–171, https://doi.org/10.1007/s00704-008-0020-x.

Dutta, R., and R. Maity, 2021: Time-varying network-based approach for capturing hydrological extremes under climate change with application on drought. *J. Hydrol.*, **603**, 126958, https://doi.org/10.1016/j.jhydrol.2021.126958.

Faergemann, H., 2012: *Update on Water Scarcity and Drought indicator development. CIS Expert Group on Water Scarcity & Drought.* 1–23 pp.

Fischer, E. M., and Coauthors, 2023: Storylines for unprecedented heatwaves based on ensemble boosting. *Nat. Commun.*, **14**, 1–11, https://doi.org/10.1038/s41467-023-40112-4.

Foley, J. C., 1957: *Droughts in Australia. Review of Records From Earliest Years of Settlement to 1955*. 282 pp.

Gamelin, B. L., J. Feinstein, J. Wang, J. Bessac, E. Yan, and V. R. Kotamarthi, 2022: Projected U.S. drought extremes through the twenty-first century with vapor pressure deficit. *Sci. Rep.*, **12**, 8615, https://doi.org/10.1038/s41598-022-12516-7.

Gibbs, W. J., and J. V. Maher, 1967: Rainfall Deciles as Drought Indicators. *Aust. Bur. Meteorol. Bulletin*, **48**, 37.

Greve, P., M. L. Roderick, A. M. Ukkola, and Y. Wada, 2019: The aridity Index under global warming. *Environ. Res. Lett.*, **14**, https://doi.org/10.1088/1748-9326/ab5046.

Guttman, N. B., 1989: Statistical Descriptors of Climate. *Bull. Am. Meteorol. Soc.*, **70**, 602–607, https://doi.org/10.1175/1520-0477(1989)070<0602:SDOC>2.0.CO;2.

——, 1994: On the Sensetivity of Sample L Moments to Sample Size. *J. Clim.*, **7**, 1026–1029, https://doi.org/10.1175/1520-0442(1994)007<1026:OTSOSL>2.0.CO;2.

——, 1999: ACCEPTING THE STANDARDIZED PRECIPITATION INDEX: A CALCULATION ALGORITHM 1. *JAWRA J. Am. Water Resour. Assoc.*, **35**, 311–322, https://doi.org/10.1111/j.1752-1688.1999.tb03592.x.

Gutzler, D. S., and T. O. Robbins, 2011: Climate variability and projected change in the western United States: regional downscaling and drought statistics. *Clim. Dyn.*, **37**, 835–849, https://doi.org/10.1007/s00382-010-0838-7.

Hayes, M., M. Svoboda, N. Wall, and M. Widhalm, 2011: The lincoln declaration on drought indices: Universal meteorological drought index recommended. *Bull. Am. Meteorol. Soc.*, **92**, 485–488, https://doi.org/10.1175/2010BAMS3103.1.

Heddinghaus, T. R., and P. Sabol, 1991: A Review of the Palmer Drought Severity Index ans Where Do We Go from Here? *7th Conference in Applied Climatology, 10-13 September 1991*, Boston, American Meteorological Society, 242–246.

Heim, R. R., 2002: A Review of Twentieth-Century Drought Indices Used in the United States. *Bull. - Am. Meteorol. Soc.*, 1149–1165.

Hoffmann, D., A. J. E. Gallant, and J. M. Arblaster, 2020: Uncertainties in Drought From Index and Data Selection. *J. Geophys. Res. Atmos.*, **125**, 1–21, https://doi.org/10.1029/2019JD031946.

Hoylman, Z. H., R. K. Bocinsky, and K. G. Jencso, 2022: Drought assessment has been outpaced by climate change: empirical arguments for a paradigm shift. *Nat. Commun.*, **13**, 1–8, https://doi.org/10.1038/s41467-022-30316-5.

Huang, J., H. M. van den Dool, and A. G. Barnston, 1996: Long-Lead Seasonal Temperature Prediction Using Optimal Climate Normals. *J. Clim.*, **9**, 809–817, https://doi.org/10.1175/1520-0442(1996)009<0809:LLSTPU>2.0.CO;2.

Hughes, N., W. Y. Soh, C. Boult, and K. Lawson, 2022: Defining drought from the perspective of Australian farmers. *Clim. Risk Manag.*, **35**, 100420, https://doi.org/10.1016/j.crm.2022.100420.

IPCC, 2007: *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. 104 pp. https://www.ipcc.ch/report/ar4/syr/.

——, 2012: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, C.B. Field et al., Eds., Cambridge University Press, p. 582.

——, 2014: *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 151 pp.

Jehanzaib, M., J. Yoo, H. H. Kwon, and T. W. Kim, 2021: Reassessing the frequency and severity of meteorological drought considering non-stationarity and copula-based bivariate probability. *J. Hydrol.*, **603**, 126948, https://doi.org/10.1016/j.jhydrol.2021.126948.

Jiang, S., M. Wang, L. Ren, C. Y. Xu, F. Yuan, Y. Liu, Y. Lu, and H. Shen, 2019: A framework for quantifying the impacts of climate change and human activities on hydrological drought in a semiarid basin of Northern China. *Hydrol. Process.*, **33**, 1075–1088, https://doi.org/10.1002/hyp.13386.

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. Clim. Appl. Meteorol.*, **25**, 77–86.

Keetch, J. J., and G. M. Byram, 1968: Drought Index for Forest Fire Control. *Res. Pap. South. Res. Stn.*, **38**, 35.

Keyantash, J., and J. A. Dracup, 2002: The An Evaluation of of Drought: Drought Indices. *Bull. Am. Meteorol. Soc.*.

Koutsoyiannis, D., 2006: Nonstationarity versus scaling in hydrology. *J. Hydrol.*, **324**, 239–254, https://doi.org/10.1016/j.jhydrol.2005.09.022.

Kraus, E. B., 1977: Subtropical Droughts and Cross-Equatorial Energy Transport. *Mon. Weather Rev.*, **105**, 1009–1018.

Lana, X., C. Serra, and A. Burgueño, 2001: Patterns of monthly rainfall shortage and excess in terms of the standardized precipitation index for Catalonia (NE

Spain). *Int. J. Climatol.*, **21**, 1669–1691, https://doi.org/10.1002/joc.697.

Landsberg, H. E., 1975: The Definition and Determination of Climatic Changes, Fluctuations and Outlooks. *2nd Carolina Geographical Symposium, Atmospheric Quality and Climatic Change.*, R.J. Kopec, Ed., Chapel Hill: U. North Carolina., 1–17.

Leng, X., X. Liu, Y. Gao, Y. Liu, Q. Yang, G. Sun, Y. Peng, and Y. Huang, 2020: Drought assessment of southwestern China based on HadGEM2-ES model under representative concentration pathway 4.5 scenario. *Nat. Hazards*, **102**, 307–334, https://doi.org/10.1007/s11069-020-03928-7.

Lian, X., and Coauthors, 2021: Multifaceted characteristics of dryland aridity changes in a warming world. *Nat. Rev. Earth Environ.*, **2**, 232–250, https://doi.org/10.1038/s43017-021-00144-0.

Livezey, R. E., K. Y. Vinnikov, M. M. Timofeyeva, R. Tinker, and H. M. van den Dool, 2007: Estimation and extrapolation of climate normals and climatic trends. *J. Appl. Meteorol. Climatol.*, **46**, 1759–1776, https://doi.org/10.1175/2007JAMC1666.1.

Lloyd-Hughes, B., and M. A. Saunders, 2002: A drought climatology for Europe. *Int. J. Climatol.*, **22**, 1571–1592, https://doi.org/10.1002/joc.846.

Ma, M., L. Ren, F. Yuan, S. Jiang, Y. Liu, H. Kong, and L. Gong, 2014: A new standardized Palmer drought index for hydro-meteorological use. *Hydrol. Process.*, **28**, 5645–5661, https://doi.org/10.1002/hyp.10063.

Martel, J.-L., F. P. Brissette, P. Lucas-Picher, M. Troin, and R. Arsenault, 2021: Climate Change and Rainfall Intensity–Duration–Frequency Curves: Overview of Science and Guidelines for Adaptation. *J. Hydrol. Eng.*, **26**, https://doi.org/10.1061/(asce)he.1943-5584.0002122.

Marvel, K., B. I. Cook, C. Bonfils, J. E. Smerdon, A. P. Williams, and H. Liu, 2021: Projected Changes to Hydroclimate Seasonality in the Continental United States. *Earth's Future*, **9**, 1–19, https://doi.org/10.1029/2021EF002019.

Matalas, N. C., 1997: Stochastic Hydrology In The Context Of Climate Change. *Clim. Change*, **37**, 89–101.

McColl, K. A., M. L. Roderick, A. Berg, and J. Scheff, 2022: The terrestrial water cycle in a warming world. *Nat. Clim. Chang.*, **12**, 604–606, https://doi.org/10.1038/s41558-022-01412-7.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The Relationship of Drought Frequency and Duration to Time Scales. *Eighth Conference on Applied Climatology*, Anaheim, California https://climate.colostate.edu/pdfs/relationshipofdroughtfrequency.pdf.

Meresa, H. K., M. Osuch, and R. Romanowicz, 2016: Hydro-meteorological drought projections into the 21-st century for selected polish catchments. *Water (Switzerland)*, **8**, https://doi.org/10.3390/w8050206.

Mi, Q., C. Ren, Y. Wang, X. Gao, L. Liu, and Y. Li, 2022: A robust ensemble drought index: construction and assessment. *Nat. Hazards*, https://doi.org/10.1007/s11069-022-05715-y.

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

Mitra, S., P. Srivastava, and J. Lamba, 2018: Probabilistic assessment of projected climatological drought characteristics over the Southeast USA. *Clim. Change*, **147**, 601–615, https://doi.org/10.1007/s10584-018-2161-y.

Núñez, J., D. Rivera, R. Oyarzún, and J. L. Arumí, 2014: On the use of Standardized Drought Indices under decadal climate variability: Critical assessment and drought policy implications. *J. Hydrol.*, **517**, 458–470, https://doi.org/10.1016/j.jhydrol.2014.05.038.

Page, M. J., and Coauthors, 2021: The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, **372**, https://doi.org/10.1136/bmj.n71.

Palmer, W. C., 1965: Meteorological Drought. *U.S. Weather Bur. Res. Pap. No. 45*, 58.

——, 1968: Keeping Track of Crop Moisture Conditions, Nationwide: The New Crop Moisture Index. *Weatherwise*, **21**, 156–161, https://doi.org/10.1080/00431672.1968.9932814.

Park, J., J. H. Sung, Y. J. Lim, and H. S. Kang, 2019: Introduction and application of non-stationary standardized precipitation index considering probability distribution function and return period. *Theor. Appl. Climatol.*, **136**, 529–542, https://doi.org/10.1007/s00704-018-2500-y.

Paulo, A., D. Martins, and L. S. Pereira, 2016: Influence of Precipitation Changes on the SPI and Related Drought Severity. An Analysis Using Long-Term Data Series. *Water Resour. Manag.*, **30**, 5737–5757, https://doi.org/10.1007/s11269-016-1388-5.

Phillips, I. D., and G. R. McGregor, 1998: The utility of a drought index for assessing the drought hazard in Devon and Cornwall, South West England. *Meteorol. Appl.*, **5**, 359–372, https://doi.org/10.1017/S1350482798000899.

Pickering, C., and J. Byrne, 2014: The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *High. Educ. Res. Dev.*, 534–548, https://doi.org/10.1080/07294360.2013.841651.

Quiring, S. M., 2009: Monitoring drought: An evaluation of meteorological drought indices. *Geogr. Compass*, **3**, 64–88, https://doi.org/10.1111/j.1749-8198.2008.00207.x.

Rashid, M. M., and S. Beecham, 2019a: Development of a non-stationary Standardized Precipitation Index and its application to a South Australian climate. *Sci. Total Environ.*, **657**, 882–892, https://doi.org/10.1016/j.scitotenv.2018.12.052.

——, and ——, 2019b: Development of a non-stationary Standardized Precipitation Index and its application to a South Australian climate. *Sci. Total Environ.*, **657**, 882–892, https://doi.org/10.1016/j.scitotenv.2018.12.052.

Russo, S., A. Dosio, A. Sterl, P. Barbosa, and J. Vogt, 2013: Projection of occurrence of extreme dry-wet years and seasons in Europe with stationary and non-stationary Standardized Precipitation Indices. *J. Geophys. Res. Atmos.*, **118**, 7628–7639, https://doi.org/10.1002/jgrd.50571.

Salas, J. D., J. Obeysekera, and R. M. Vogel, 2018: Techniques for assessing water infrastructure for nonstationary extreme events: a review. *Hydrol. Sci. J.*, **63**, 325–352, https://doi.org/10.1080/02626667.2018.1426858.

Salvi, K., and S. Ghosh, 2016: Projections of Extreme Dry and Wet Spells in the 21st Century India Using Stationary and Non-stationary Standardized Precipitation Indices. *Clim. Change*, **139**, 667–681, https://doi.org/10.1007/s10584-016-1824-9.

Satoh, Y., and Coauthors, 2021: A quantitative evaluation of the issue of drought definition: A source of disagreement in future drought assessments. *Environ. Res. Lett.*, **16**, https://doi.org/10.1088/1748-9326/ac2348.

Schlef, K. E., and Coauthors, 2023: Incorporating non-stationarity from climate change into rainfall frequency and intensity-duration-frequency (IDF) curves. *J. Hydrol.*, **616**, https://doi.org/10.1016/j.jhydrol.2022.128757.

Schreck, C. J., P. J. Klotzbach, and M. M. Bell, 2021: Optimal Climate Normals for the North Atlantic Hurricane Activity. *Geophys. Res. Lett.*, **48**, 1–9, https://doi.org/10.1029/2021GL092864.

Shafer, B. A., and L. E. Dezman, 1982: Development of a Surface Water Supply Index (SWSI) to assess the severity of drought conditions in snowpack runoff areas (Colorado) . *Proc. East. Snow Conf. 39th Annu. Meet.*, 164–175.

Shiau, J. T., 2020: Effects of Gamma-Distribution Variations on SPI-Based Stationary and Nonstationary Drought Analyses. *Water Resour. Manag.*, **34**, 2081–2095, https://doi.org/10.1007/s11269-020-02548-x.

Shukla, S., and A. W. Wood, 2008: Use of a standardized runoff index for characterizing hydrologic drought. *Geophys. Res. Lett.*, **35**, 1–7, https://doi.org/10.1029/2007GL032487.

Singh, R., M. Kumari, S. Bindal, and I. Gupta, 2022: Quantification of Drought Condition Using Drought Indices: A Review. 231–241.

Song, Z., J. Xia, D. She, L. Zhang, C. Hu, and L. Zhao, 2020: The development of a Nonstationary Standardized Precipitation Index using climate covariates: A case study in the middle and lower reaches of Yangtze River Basin, China. *J. Hydrol.*, **588**, 125115, https://doi.org/10.1016/j.jhydrol.2020.125115.

Stagge, J. H., and K. Sung, 2022: A Nonstationary Standardized Precipitation Index (NSPI) Using Bayesian Splines. *J. Appl. Meteorol. Climatol.*, **61**, 761–779, https://doi.org/10.1175/JAMC-D-21-0244.1.

——, L. M. Tallaksen, L. Gudmundsson, A. F. Van Loon, and K. Stahl, 2015: Candidate Distributions for Climatological Drought Indices (SPI and SPEI). *Int. J. Climatol.*, **35**, 4027–4040, https://doi.org/10.1002/joc.4267.

——, D. G. Kingston, L. M. Tallaksen, and D. M. Hannah, 2017: Observed drought indices show increasing divergence across Europe. *Sci. Rep.*, **7**, 1–10, https://doi.org/10.1038/s41598-017-14283-2.

Stevenson, S., S. Coats, D. Touma, J. Cole, F. Lehner, J. Fasullo, and B. Otto-Bliesner, 2022: Twenty-first century hydroclimate: A continually changing baseline, with more frequent extremes. *Proc. Natl. Acad. Sci. U. S. A.*, **119**, 1–9, https://doi.org/10.1073/pnas.2108124119.

Svoboda, M., and Coauthors, 2002: The Drought Monitor. *Bull. Am. Meteorol. Soc.*, **83**, 1181–1190, https://doi.org/10.1175/1520-0477(2002)083<1181:TDM>2.3.CO;2.

UNESCO, 1977: *Map of the World Distribution of Arid Regions; Explanatory Note*. 54 pp.

Van-rooy, M. P., 1965: A Rainfall Anomaly Index (RAI), Independent of the Time and Space. *Notos*, **14**, 43–48.

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. Clim.*, **23**, 1696–1718, https://doi.org/10.1175/2009JCLI2909.1.

——, and Coauthors, 2012: Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interact.*, **16**, 1–27, https://doi.org/10.1175/2012EI000434.1.

Villarini, G., J. A. Smith, and F. Napolitano, 2010: Nonstationary modeling of a long record of rainfall and temperature over Rome. *Adv. Water Resour.*, **33**, 1256–

1267, https://doi.org/10.1016/j.advwatres.2010.03.013.

Wang, W., J. Wang, and R. Romanowicz, 2021: Uncertainty in spi calculation and its impact on drought assessment in different climate regions over China. *J. Hydrometeorol.*, **22**, 1369–1383, https://doi.org/10.1175/JHM-D-20-0256.1.

Wang, Y., J. Li, P. Feng, and R. Hu, 2015: A Time-Dependent Drought Index for Non-Stationary Precipitation Series. *Water Resour. Manag.*, **29**, 5631–5647, https://doi.org/10.1007/s11269-015-1138-0.

Wang, Y., and Coauthors, 2022: A New Non-stationary Hydrological Drought Index Encompassing Climate Indices and Modified Reservoir Index as Covariates. *Water Resour. Manag.*, **36**, 2433–2454, https://doi.org/10.1007/s11269-022-03151-y.

Wasko, C., S. Westra, R. Nathan, H. G. Orr, G. Villarini, R. Villalobos Herrera, and H. J. Fowler, 2021: Incorporating climate change in flood estimation guidance. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, **379**, https://doi.org/10.1098/rsta.2019.0548.

World Meteorological Organization, 2007: *the Role of Climatological Normals in a Changing Climate*. https://library.wmo.int/index.php?lvl=notice_display&id=16659#.Y_yqgnbMKCg.

——, 2012: *Standardized Precipitation Index User Guide*.

Wu, H., M. J. Hayes, D. A. Wilhite, and M. D. Svoboda, 2005: The effect of the length of record on the standardized precipitation index calculation. *Int. J. Climatol.*, **25**, 505–520, https://doi.org/10.1002/joc.1142.

Yan, L., L. Xiong, C. Jiang, M. Zhang, D. Wang, and C. Y. Xu, 2021: Updating intensity–duration–frequency curves for urban infrastructure design under a changing environment. *Wiley Interdiscip. Rev. Water*, **8**, 1–24, https://doi.org/10.1002/wat2.1519.

Zeder, J., S. Sippel, O. C. Pasche, S. Engelke, and E. M. Fischer, 2023: The Effect of a Short Observational Record on the Statistics of Temperature Extremes. *Geophys. Res. Lett.*, **50**, 1–11, https://doi.org/10.1029/2023GL104090.

Zhang, T., X. Su, and K. Feng, 2021: The development of a novel nonstationary meteorological and hydrological drought index using the climatic and anthropogenic indices as covariates. *Sci. Total Environ.*, **786**, 147385, https://doi.org/10.1016/j.scitotenv.2021.147385.

Zhao, H., X. Pan, Z. Wang, S. Jiang, L. Liang, X. Wang, and X. Wang, 2019: What were the changing trends of the seasonal and annual aridity indexes in northwestern China during 1961–2015? *Atmos. Res.*, **222**, 154–162, https://doi.org/10.1016/j.atmosres.2019.02.012.

Zou, L., J. Xia, L. Ning, D. She, and C. Zhan, 2018: Identification of Hydrological Drought in Eastern China Using a Time-Dependent Drought Index. *Water*, **10**, 315, https://doi.org/10.3390/w10030315.