



Expected annual minima from an idealized moving-average drought index

James H. Stagge¹, Kyungmin Sung^{1,2}, Irene Felix Munyejuru¹, and Md Atif Ibne Haidar¹

¹Department of Civil, Environmental and Geodetic Engineering, The Ohio State University, Columbus, OH 43210, USA

²Korea Adaptation Center for Climate Change, Korea Environment Institute, Sejong, Republic of Korea

Correspondence: James H. Stagge (stagge.11@osu.edu)

Received: 14 May 2024 – Discussion started: 27 May 2024

Revised: 22 October 2024 – Accepted: 6 December 2024 – Published: 7 February 2025

Abstract. Numerous drought indices originate from the Standardized Precipitation Index (SPI) and use a moving-average structure to quantify drought severity by measuring normalized anomalies in hydroclimate variables. This study examines the theoretical probability of annual minima based on such a process. To accomplish this, we derive a stochastic model and use it to simulate 10×10^6 years of daily or monthly SPI values in order to determine the distribution of annual exceedance probabilities. We believe this is the first explicit quantification of annual extreme exceedances from a moving-average process where the moving-average window is proportionally large (5%–200%) relative to the year, as is the case for many moving-window drought indices. The resulting distribution of annual minima follows a generalized normal distribution rather than the generalized extreme-value (GEV) distribution, as would be expected from extreme-value theory. From a more applied perspective, this study provides the expected annual return periods for the SPI or related drought indices with common accumulation periods (moving-window length), ranging from 1 to 24 months. We show that the annual return period differs depending on both the accumulation period and the temporal resolution (daily or monthly). The likelihood of exceeding an SPI threshold in a given year decreases as the accumulation period increases. This study provides clarification and a caution for the use of annual return period terminology (e.g. the 100-year drought) with the SPI and a further caution for comparing annual exceedances across indices with different accumulation periods or resolutions. The study also distinguishes between theoretical values, as calculated here, and real-world exceedance probabilities, where there may be cli-

matological autocorrelation beyond that created by the moving average.

1 Introduction

The Standardized Precipitation Index (SPI) (Guttman, 1999; McKee et al., 1993) is used to measure meteorological drought operationally by many organizations, including the WMO and numerous drought monitors (Cammalleri et al., 2021; Hao et al., 2014; Heim and Brewer, 2012; Lawrimore et al., 2002; Sheffield et al., 2014; Svoboda et al., 2002; WMO and GWP, 2016). This index is particularly useful because it requires only precipitation data and mirrors commonly agreed-upon definitions of meteorological drought, a sustained and spatially extensive period of below-average water availability (Heim, 2002; Lloyd-Hughes, 2014; Tallaksen and Van Lanen, 2023). The SPI measures accumulated or mean precipitation during a moving window and normalizes this quantity relative to the historical climatology for that day of the year, thereby producing a normalized anomaly. The SPI is typically referred to using its accumulation period, the backwards-looking moving window, measured in months. The SPI-3 therefore represents precipitation anomalies based on the previous 3 months. The SPI can also be calculated at a daily temporal resolution, though the naming convention still typically refers to months, for example, using a 90 d moving window to calculate the SPI-3. Accumulated precipitation is bounded by zero and is typically positively skewed, commonly leading to the use of independently calibrated gamma distributions (Guttman, 1999; Lloyd-Hughes and Saunders, 2002; Stagge et al., 2015a; Stagge and Sung, 2022) to rep-

resent climatology and transform accumulated precipitation into percentiles. These percentiles are ultimately transformed to anomalies of the standard normal distribution, with a mean of 0 and a standard deviation of 1.

The fundamental SPI concept has since been expanded to a family of normalized drought indices, each quantifying anomalies from different portions of the hydrologic cycle. For example, normalized drought indices have measured anomalies in the climatic water balance, soil moisture, groundwater, and streamflow, referred to, respectively, as the Standardized Precipitation Evapotranspiration Index (SPEI, Beguería et al., 2014; Vicente-Serrano et al., 2010), the Standardized Soil Moisture Index (SSI, Sheffield et al., 2014), the Standardized Groundwater Level Index (SGI, Bloomfield and Marchant, 2013), and the Standardized Runoff Index (SRI, Shukla and Wood, 2008). The principles developed in this study are applicable to all normalized drought indices that employ a moving-average structure, but we will occasionally refer to the SPI as the simplest example of this broader class. It should be noted that these indices sometimes use instantaneous values, to which a moving window is not applied and the findings in this study are less relevant.

Values from normalized drought indices follow the standard normal distribution (mean = 0, standard deviation = 1) within the reference calibration period, resulting in relatively interpretable percentiles. These percentiles have then been used to develop thresholds. For example, the US Drought Monitor uses five categories classified from D0 to D4 based on SPI thresholds of -0.5 , -0.8 , -1.3 , -1.5 , and -2 (Xia et al., 2014). These thresholds correspond very roughly with percentiles of $\leq 30\%$, 20% , 10% , 5% , and 2% , which allows them to be compared with other percentile-based indices. In this study, we will use the USDM SPI thresholds as reference points, though using the exact corresponding percentiles. For example, an extreme drought (D3) is assumed to occur for SPI values of -1.5 to -2 or for values 1.5 to 2 SD (standard deviations) drier than what is typical for that time of year. The likelihood of a given day falling within the D3 category is therefore 4.4 %, the difference between exceedance probabilities for -1.5 (6.7 %) and -2 (2.3 %). These thresholds have statistical utility but are arbitrary. For example, different thresholds have been used previously (McKee et al., 1993), drought impacts are linked to a wide range of thresholds (Blauth et al., 2016; Stagge et al., 2015b), and the US Drought Monitor's blended Objective Drought Indicator uses additional logic to classify droughts (Anderson et al., 2013; Xia et al., 2014).

Confusion of interpretation can occur with regards to normalized drought indices because hydrologists often quantify extremes in terms of annual exceedance probabilities or their reciprocal, the return period. However, annual exceedance probabilities differ from the probability associated with a normalized drought index series because each day follows a standard normal distribution. For example, there is a 6.7 % chance that the SPI on 1 January will fall into the D3 cate-

gory or worse (see above), but there is also a 6.7 % chance of falling into this drought category on 2 January and all subsequent days of the year. Therefore, it is not correct to state that there is a 6.7 % chance of experiencing a D3 drought or worse (one SPI value less than -1.5) in a given year. The latter statement is what hydrologists typically define as return period or annual exceedance probability, which is clearly distinct from the daily or monthly SPI probabilities. If SPI values were independent and identically distributed (i.i.d.), the likelihood of a year with a single SPI value below a given threshold would be given by

$$1 - (1 - p)^n, \quad (1)$$

where p is the probability of exceedance for each time step, and n is the length of the sample period: 12 months or 365 d. The probability of at least one observation being below -1.5 in a given year then approaches 100 % for daily time steps and is 56.4 % for monthly time steps. However, SPI time series are fundamentally not i.i.d., instead being subject to a large amount of temporal autocorrelation due to the SPI's moving-average structure. The degree of temporal autocorrelation depends on the length of the moving window or of the accumulation period, which can be tuned to capture, alternatively, short droughts with smaller periods or to capture seasonal to multi-year droughts using longer periods. While temporal autocorrelation invalidates annual exceedance estimates from Eq. (1), the moving-average structure provides a predictable and pre-defined structure that can be leveraged to quantify the likelihood of annual exceedances.

Despite there being a robust field of research into moving-average models as part of the autoregressive moving-average (ARMA) family of time series models (Box and Jenkins, 1970; Wilks, 2011), we were unable to identify prior research quantifying the extreme behaviour of a moving-average sequence where the moving window is long relative to the time interval, as is the case for normalized drought indices. Solutions exist for the simple AR(1) case (Hirtzel, 1985a, b) and for an AR(n) case, although the latter requires high-dimension copulas and is unstable (Tsoukalas, 2022). Prior extreme-value theory for moving averages (Davis and Resnick, 1991; Rootzen, 1986) appears to break down under the conditions of the normalized drought indices, as we outline here. The purpose of this study is therefore to quantify the expected annual minima and their associated return periods from a theoretical moving-average series when simulated from daily or monthly sequences with moving windows that mimic common lengths for the SPI and other drought indices. Before simulating these annual minima, we first derive a stochastic model for an idealized moving-average series and show how, for several example sites, temporal autocorrelation from observations broadly follows this idealized model due to the underlying moving average. For the remainder of this study, we will focus on annual minima as relevant for droughts; however, all findings apply equally for positive

(maxima) extremes because the standard normal distribution is symmetrical.

2 Methods

This study is organized by first developing a simulation method to generate long sequences that meet two criteria: (1) values for each day or month follow the standard normal distribution, with a mean of 0 and a standard deviation of 1, and (2) values follow a uniformly weighted backwards-looking moving average. Following this derivation, we simulate extremely long sequences from this process and extract the annual minima for multiple thresholds. We also estimate the first four moments of the annual minima and test several extreme-value distributions to determine whether the annual minima can be adequately represented by a continuous probability distribution. Using these fitted distributions, we describe the annual exceedance probabilities for any threshold.

This approach is designed only to explore the theoretical behaviour of a simplified case, affected only by the structural persistence caused by the moving average. This approach does not consider additional climatological persistence caused by a region's climatology or by macroscale drivers like atmospheric teleconnections. We make this distinction between structural and climatological persistence throughout the study. Normalized drought indices in the real world are impacted by a combination of these and other factors (see Discussion), requiring site- and index-specific analyses. But, as we show for a variety of case study sites across climate regions, structural persistence typically represents the vast majority of temporal autocorrelation. Therefore, here, we present the solution to the limiting case of structural persistence only for clarification and as a benchmark for future comparisons.

2.1 Relating SPI to moving-average processes

The SPI calculates a moving average or moving sum of precipitation, which is then normalized for each day or month of the year relative to its historical climatology (Guttman, 1999; McKee et al., 1993). For the purposes of this study, we stipulate that the simplest SPI time series is also a moving-average series following the necessary standard normal distribution with a mean of 0 and a standard deviation of 1. The potential implications of this assumption are explored further in the Discussion. A random series of SPI- q values, where q is the accumulation period, can therefore be generated using random daily or monthly incremental changes, called innovations, Z_t , sampled from

$$Z_t \sim N(0, \sqrt{q}) \text{ i.i.d.} \quad (2)$$

These progressively enter and leave a moving average,

$$\text{SPI}(t) = \frac{1}{q} \sum_{j=0}^q Z_{t-j}, \quad (3)$$

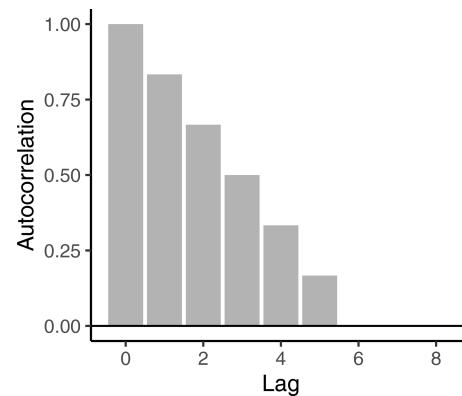


Figure 1. Autocorrelation function (ACF) for an SPI-6 generating process.

thereby generating an SPI(t) series. This produces the requisite standard normal distribution while maintaining a moving window of q time steps. The SPI has a backwards-looking "memory" of q time steps (days or months), where each is weighted equally. Within the autoregressive moving-average (ARMA) framework (Box and Jenkins, 1970; Wilks, 2011), such a model can be written as a moving-average process, MA($q-1$). In this notation, the moving window is written as $q-1$ rather than q because $q-1$ represents the number of time lags in addition to the innovation added in the current step, i.e. at a time lag of zero. Writing the model in standard ARMA notation allows the application of a deep body of literature regarding the properties of ARMA models. Using this standard notation, the MA($q-1$) process has $q-1$ MA coefficients of 1 and innovations with a standard deviation of \sqrt{q}/q . For example, it is possible to simulate an SPI-6 sequence using an MA(5) with MA coefficients of [1, 1, 1, 1, 1] and innovations randomly sampled from an i.i.d. Gaussian distribution $N\left(0, \frac{\sqrt{6}}{6}\right)$.

The autocorrelation function (ACF) for such a theoretical SPI- q series, represented by an MA($q-1$) process, has a linear decay of $1/q$ per time lag (Fig. 1). The ACF becomes zero past q lags because these innovations are no longer part of the moving average. Expanding on the previous SPI-6 example, temporal autocorrelation falls from 1 at lag 0 to 0.1666 at lag 5, followed by zero autocorrelation for lags of 6 months and beyond (Fig. 1). The same would occur for a 6-month SPI-6 series using a daily temporal resolution, but autocorrelation would decay linearly towards zero after 182 d (approximately 6 months).

Moving-average processes with discrete time lags can be converted to an infinite-order autoregressive process if the MA model is invertible (Granger and Andersen, 1978). For an ARMA process to be invertible, all roots of the characteristic polynomial must lie outside the complex unit circle (> 1) (Davidson, 1981; Granger and Andersen, 1978; Hallin, 1984). Roots of the MA($q-1$) process defined here

are exactly on the unit circle, making this theoretical model not invertible to an equivalent AR(∞) model.

2.2 Stochastic simulation from MA model

Normalized drought index time series were simulated using the moving-average model (Eqs. 2 and 3), with unique simulations being run using daily and monthly temporal resolutions. Monthly simulations were performed using accumulation periods of 1, 2, 3, 6, 9, 12, and 24 months to address the most commonly used moving windows for the SPI.

Daily simulations recreated these accumulation periods with an additional half-month window: 15, 30, 60, 90, 182, 274, 365, and 730 d. This permits a direct comparison of daily simulations with their monthly counterparts using the typical SPI- q naming scheme. This range is similar to the 1–60-month range considered by the US Drought Monitor (Svoboda et al., 2002; Xia et al., 2014), but we chose not to extend beyond 2 years (24 months). We also chose to include a 15 d window for completeness, both to show behaviour under extreme conditions and because some measures of flash drought rely on rapid decreases in water balance or soil moisture over 14 or 15 d periods (Christian et al., 2019; Lisonbee et al., 2022).

For daily and monthly experiments, we simulated a total of 10×10^6 years using 20 repeated simulations of 500 000 years. The result was 3.65×10^9 total days and 0.12×10^9 months. A total of 20 repeated simulations were used to test uncertainty and the impact of initial conditions while remaining conscientious of computational memory. Once it was determined that there were nearly imperceptible differences between statistical characteristics for the repeated simulations, the 20 simulations were combined to create the full 10×10^6 year dataset. While this creates 19 small discontinuities at the interface between the 20 simulations, this effect was minimal when viewed across the 0.12–3.65 billion individual time steps.

2.3 Temporal autocorrelation at case study sites

While the purpose of this study is primarily to determine the extreme behaviour from the moving-average structure underlying many drought indices, it is illustrative to understand how closely observed drought indices follow this idealized structure. To test this, we first derive SPI time series for the 1, 3, 6, 12, and 24 month moving windows using daily data from the Global Soil Wetness Project Phase 3 (GSWP3) (Kim, 2017; Dirmeyer et al., 2006). The GSWP3 forcing dataset is based on a the 20th Century Reanalysis dataset (Compo et al., 2011), with dynamical downscaling to maintain both high- and low-frequency signals and additional bias correction based on the GPCC (Meyer-Christoffer et al., 2015) for precipitation. The GSWP3 dataset used here includes daily precipitation for the period 1901–2010 at a global resolution of $0.5^\circ \times 0.5^\circ$. While a more comprehen-

sive investigation is needed in the future to examine multiple variables, e.g. SPI, SPEI, SSI, and SGI, at sites across the world, we have chosen to focus on 23 case study grid cells, aiming to span as many Köppen–Geiger climate zones (Peel et al., 2007) as feasible and to examine three to five sites in each of the six continents (Fig. S5). SPI was calculated following the stationary-spline approach of Stagge and Sung (2022) by fitting a seasonally cyclic spline to the parameters of the gamma distribution for positive precipitation and to those of the logistic distribution for the probability of zero precipitation.

For each observed SPI time series, we then calculated the lagged correlation at 1 d increments using Pearson correlation. To compare observed autocorrelation with that expected from an idealized moving-average process, we generated 1000 independent replicates of 110-year daily time series following Eqs. (2) and (3), calculating the lagged correlation for each using an identical approach to that used for the observed time series. The inner 95 % percentile of all replicates was used to approximate the 95 % confidence interval, while the mean showed the expected autocorrelation. Lagged autocorrelation from the observed SPI was contrasted with that from the idealized moving-average replicates to show how typical SPI series mimic the idealized series and to illustrate how much autocorrelation is caused by the moving-average structure relative to climatological persistence for these example sites. We present four sites in North America in the text but show all other sites as figures in the Supplement.

2.4 Annual minima analysis

Block annual minima were extracted from the simulated daily and monthly time series using an annual period. Calendal years, rather than water years, were used for ease of interpretation and because the synthetic SPI series does not distinguish seasonality.

To find a univariate probability distribution that reasonably approximates the simulated annual minima, we first compared the sample L-moment ratios with theoretical values for common univariate distributions (Hosking and Wallis, 1995; Peel et al., 2001). For the purposes of this analysis, the sign of annual minima was changed to better fit distributions typically designed for maxima extremes. This sign change is reasonable because, while seasonal precipitation or other measured variables may be skewed, the transformed SPI distribution is symmetrical, normally distributed about zero, and the simulated SPI did not distinguish seasonality. L-moment plots were developed for the entire 10×10^6 year simulated series and for each of the 20 subset simulations to estimate uncertainty around the L-moment estimate. In addition, we evaluated quantile–quantile (q–q) plots, showing the empirical quantiles from the sample compared with the theoretical quantiles determined from the candidate distribution.

Once an appropriate univariate distribution was found, we fit this distribution using both maximum-likelihood es-

timation (MLE) and L moments, which have previously been shown to produce equivalent fits, particularly for large datasets (Beguería et al., 2014). Estimates across both methods were nearly equivalent. For all subsequent analyses, we report L-moment estimates. Goodness of fit was verified using q-q plots and the Akaike information criterion (AIC) (Akaike, 1998; Cavanaugh and Neath, 2019). From the fitted distributions, exceedance probabilities were estimated, along with their corresponding annual return periods for a range of SPI values using the L-moment estimates. Empirical estimates for several key exceedance probabilities were calculated as validation.

3 Results

3.1 Annual minima best fit distribution

To enable the development of continuous exceedance probability and return period curves, we sought to find a univariate probability distribution that provided a good fit for the annual minima. L-moment ratios closely match the theoretical moments of the three-parameter generalized normal distribution (Peel et al., 2001), also known as the three-parameter lognormal distribution (Fig. 2), indicating that this distribution best fits the data. This holds true across all accumulation periods, regardless of whether one is using monthly (Fig. 2) or daily underlying data (Fig. S1 in the Supplement). As the accumulation period increases towards 24 months, the annual minima appear to be increasingly like the normal, with L skewness decreasing towards zero and L kurtosis approaching the theoretical value of 0.1226 for normally distributed data. The generalized normal distribution is capable of representing the more skewed distributions for short accumulation periods and the more symmetrical distributions for longer accumulation periods, all while closely matching the simulated annual minima (Fig. 2). This extremely close fit supports the use of the generalized normal distribution as opposed to other potential distributions to capture the magnitude of annual minima. This extremely close fit for 10 million simulated observations, provides strong evidence. L-moment ratios were extremely stable, with nearly imperceptible differences between the 20 simulations (Fig. S2). The interquartile ranges (IQRs) for L-skewness and L-kurtosis estimates from daily simulations were 0.0007–0.0015 and 0.0003–0.0007, respectively, amounting to uncertainties of 0.8 %–6 % and 0.3 %–0.5 %, measured by the IQR divided by the median.

The cumulative distribution function for the generalized normal distribution can be described by the standard normal distribution $\Phi(Y)$ with an additional transformation, where Y is

$$Y = -\kappa^{-1} \log \left(1 - \frac{\kappa(x-\xi)}{\alpha} \right) \quad \text{where } \kappa \neq 0 \quad (4)$$

$$Y = \frac{x-\xi}{\alpha} \quad \text{where } \kappa = 0$$

In the above, ξ is the location parameter, α is a scale parameter, and κ is a shape parameter (Das, 2018; Hosking and Wallis, 1997). When $\kappa = 0$, the distribution reverts back to the normal distribution, with a mean of ξ and a standard deviation of α . This distribution is equivalent to the three-parameter lognormal distribution or a normal distribution fit in natural log space with parameters μ_{\log} , σ_{\log} , and ζ , representing the location in log space, the scale in log space, and a lower distribution bound (Das, 2018). For conversion between the parameters of the two, one can use the following relationships: $\kappa = -\sigma_{\log}$, $\alpha = \sigma_{\log} e^{\mu_{\log}}$, and $\xi = \zeta + e^{\mu_{\log}}$. The generalized normal distribution has been used in hydrology and meteorology studies (Basu and Srinivas, 2013; Das, 2018; Sangal and Biswas, 1970) for extreme-value analysis.

Notably, the annual minima values do not converge towards the generalized extreme-value (GEV) distribution (Fig. 2), as might be expected for extreme values with temporal autocorrelation (Berman, 1964; Davis and Resnick, 1991; Hirtzel, 1985b; Leadbetter et al., 1983; Rootzen, 1986). The Berman theorem (Berman, 1964; Coles, 2001) states that the maxima statistics of stationary Gaussian sequences with autocorrelations should converge towards a Gumbel distribution (GEV type I). We believe that this deviation from expectation is due to the clustering of extremes, which violates Berman's theorem. This proposed explanation is expanded upon in the Discussion.

Deviations from the GEV distribution are most noticeable for longer accumulation periods, like the SPI-24, which has sample L-skewness and L-kurtosis values of 0.0035 and 0.123, respectively (Fig. 2). These are nearly identical to theoretical values for the normal distribution (0 and 0.1226) and, thus, for the generalized normal distribution, whereas the GEV distribution cannot produce distributions with zero skewness (Fig. 2). While the deviation from the GEV distribution becomes smaller for shorter accumulation periods, it is notable how closely the empirical L moments from the simulations follow the generalized normal distribution (Fig. 2).

Quantile–quantile plots were used to further verify fitting skill, with all simulated extremes falling neatly along the distribution, with only slight deviations at the most extreme values (return periods $> 1 \times 10^6$) and no consistent patterns of bias (Fig. S3). The generalized normal distribution therefore accurately reproduces empirical quantiles, with little noticeable bias even at the extremes. These strong fits appear to be similarly accurate for short and long accumulation periods, though AIC values increase slightly (become worse) for longer accumulation periods (Fig. S4a). The AIC confirms that the generalized normal distribution produces a better fit than the GEV across all moving-average lengths, illustrated by a lower value (negative difference) in terms of AIC (Fig. S4b). Based on all of this evidence, all subsequent analyses are therefore based on the generalized normal distribution, except where empirical estimates are used as a validation.

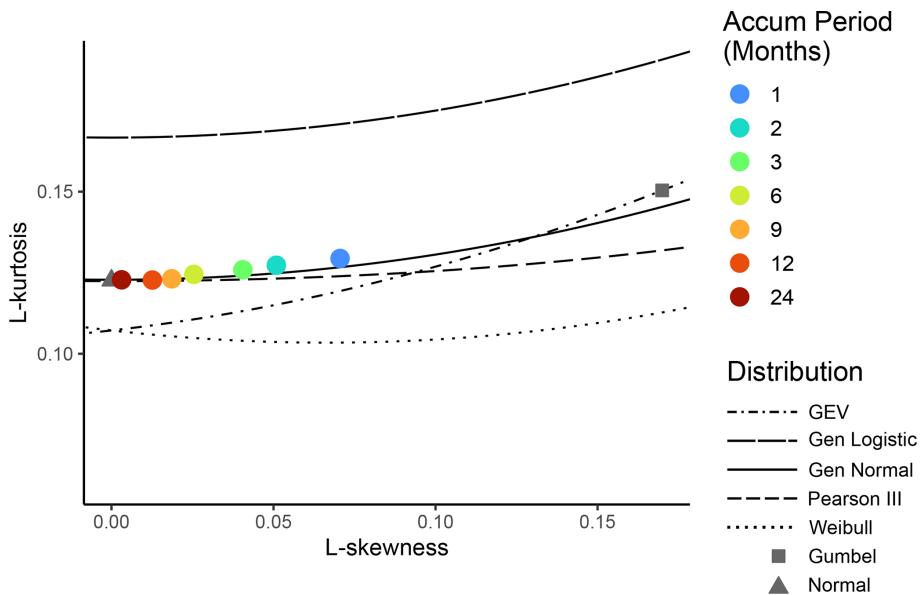


Figure 2. L-moment ratios for annual extremes from monthly simulated series. Coloured points refer to fitted moments across varying accumulation periods, while lines correspond to theoretical distributions. Note that this figure shows distributions with a flipped sign. True skewness for annual minima is negative. An equivalent figure for daily simulations is shown in the Supplement (Fig. S1).

3.2 Observed vs. theoretical autocorrelation

The pattern of temporal autocorrelation for real-world SPI time series follows the pattern from the idealized moving-average time series, with a mostly linear decrease towards zero at the moving-window length, followed by fluctuations around zero (Figs. 3 and S6–S10). This implies that the structural persistence, occurring due to the moving average, is more important than climatological persistence in many locations and climate regions. The first row of Fig. 3 shows the first 20 of the 1000 replicates used to generate the 95 % interval. The red line (observed) generally remains within this 95 % interval for climates ranging from cold, hot summers (Winnipeg, group Dfb) to temperate (Columbus, group Cfa) to tropical monsoon (Miami, group Am) and hot deserts (Tucson, group BSh). This pattern holds for all other case study sites analysed outside of North America (Figs. S6–S10). Temporal autocorrelation beyond the limits shown here continues to fluctuate around zero but generally remains within the 95 % interval. Where there are deviations from the theoretical persistence, this could be evidence of randomness or of some climatological persistence. This can be investigated more thoroughly in future studies (Sect. 4.1).

3.3 Annual extreme values

Using the fitted generalized normal distribution, we explored the distribution and return periods for annual minima of the idealized moving-average time series described by Eqs. (2) and (3). For longer accumulation periods, the distribution of annual minima becomes less skewed, with a shift in the mean

towards zero (Figs. 4 and S2). Conversely, short accumulation periods have more skewness and are shifted towards the negative (more extreme). For daily data, the distribution mean increases from -2.53 for a 1-month accumulation period (SPI-1) to -0.77 for a 24-month period (SPI-24) (Figs. 5 and S2). For monthly data, this shift in the mean value for the annual minima increases from -1.63 to -0.61 for the SPI-1 and SPI-24. Differences between the daily and monthly resolutions are discussed in the next section.

Concurrently with an increase in the distribution mean for the SPI annual minima, distributions transition from skewed left for short accumulation periods towards more normally distributed (negligible skew) for long accumulation periods (Figs. 5 and S4). Variance also increases with an increased accumulation period. All distribution parameters and moments are presented in the Supplement (Figs. S12 and S13, Table S1).

The aforementioned distribution changes due to the accumulation period produce differences in the probability of annual threshold exceedances and their associated return period (Figs. 4 and 5). In Fig. 5, lines represent the probability of threshold exceedances derived from the fitted distribution, while vertical lines correspond to USDM thresholds. If one was to focus on the D4 exceptional drought threshold ($SPI < -2.0$) for a daily time series, the annual probability of a single SPI-3 exceedance of this threshold is 37.7 %, a return period of 2.65 years, while this probability decreases to 9.27 % for the SPI-24, corresponding to once every 10.79 years (Fig. 5a, Table 1). The discrepancy becomes even greater for extremely short accumulation periods. For example, the probability of at least one value being

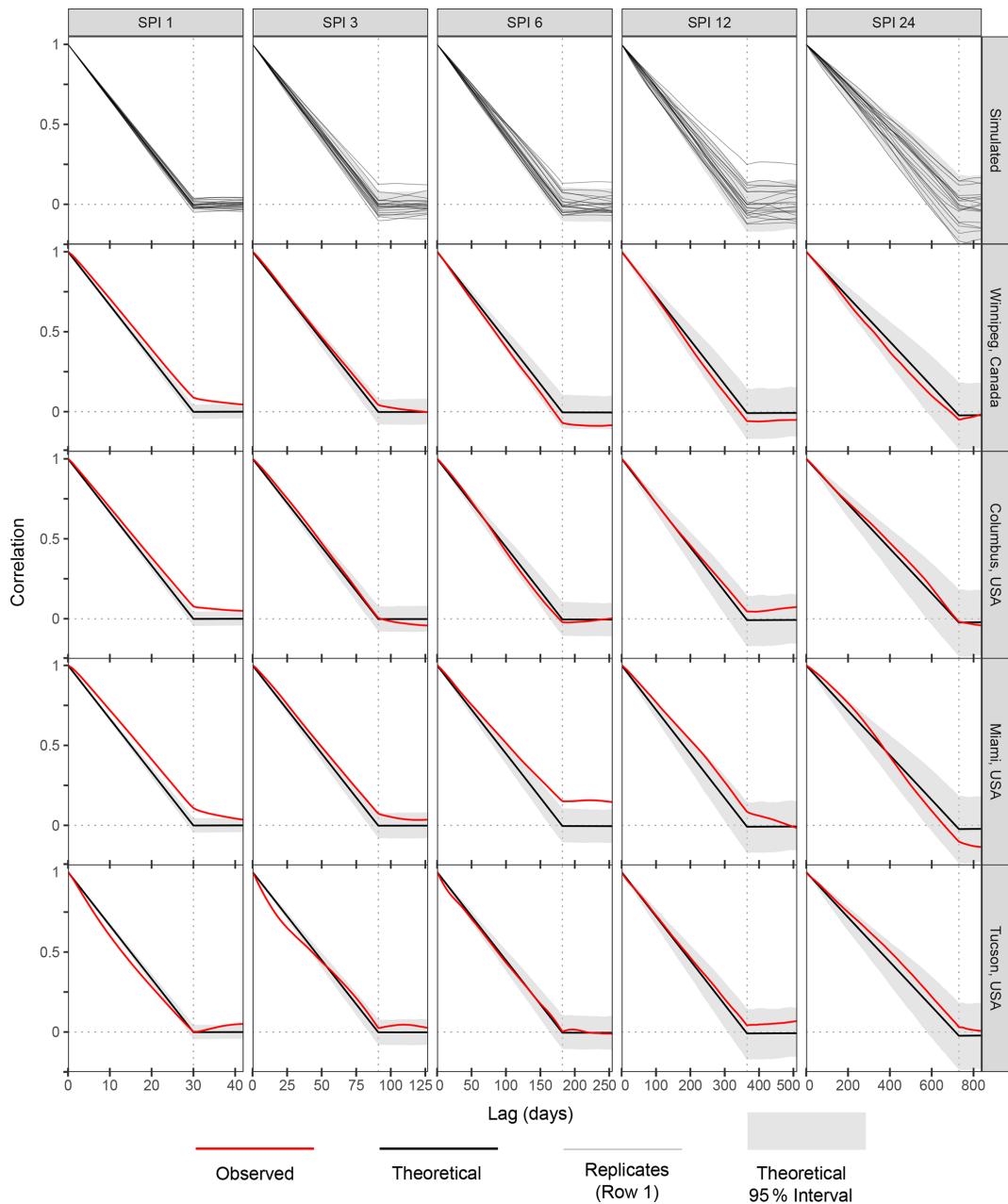


Figure 3. Lagged correlation for the SPI-1, 3, 6, 12, and 24 moving windows. The first row shows 20 replicates from random simulation, following Eqs. (2) and (3), in light grey. Subsequent rows show four North American grid cells representing climate zones ranging from coldest (Winnipeg) to warmest (Tucson). Red lines show observed temporal autocorrelation. Black lines show the expected (mean) temporal autocorrelation from 1000 replicate simulations, while the grey regions show the 95 % interval from these replicates. Dotted lines show the moving-window length. Note that each subfigure has a different x axis to better focus on the moving window.

below -2 for the SPI-0.5 (15 d period) is 87.6 % (return period of 1.14 years). In other words, a drought agency would declare a D4 drought every year if monitoring the SPI-0.5 but only once a decade if monitoring the SPI-24.

Another way to interpret Fig. 5 is to make a horizontal comparison. The 2-year return period should be exceeded once every other year when measuring a sufficiently long

record. The threshold associated with this relatively commonplace occurrence varies from -2.5 , considered to be an extreme (D3) drought for the SPI-0.5, to -0.764 , considered to be only a moderate (D1) drought for the SPI-24 (Fig. 5a). The idea of experiencing an extreme short drought (0.5-month accumulation period) at least once every other year may be challenging for interpretation by the public. Again,

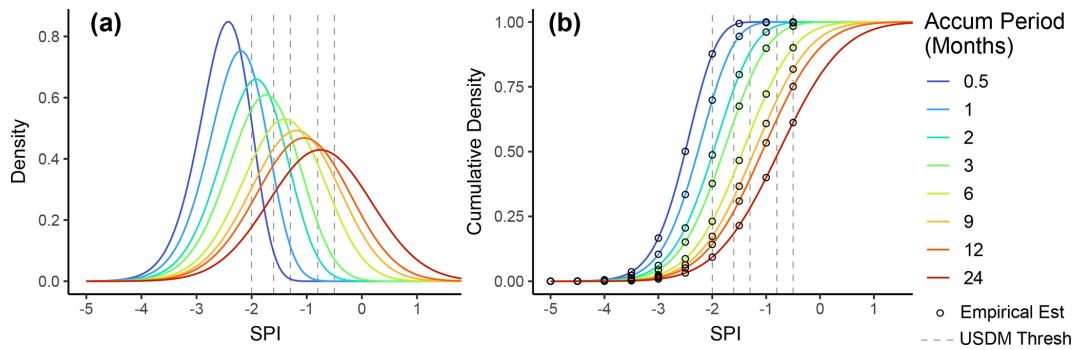


Figure 4. Annual minima (a) distribution and (b) cumulative probability density for daily sequences of varied accumulation periods, indicated by colour. Colours are identical to those in Fig. 2. Vertical grey lines correspond to US Drought Monitor thresholds for D0–D4 (-0.5 , -0.8 , -1.3 , -1.6 , and -2.0). Open points represent empirical estimates directly from simulations.

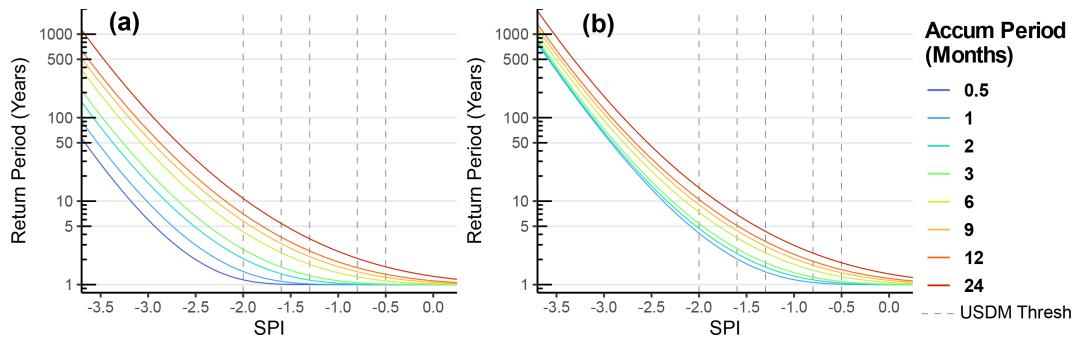


Figure 5. Return periods for (a) daily and (b) monthly sequences, with accumulation periods indicated by colour. Colours are identical to those in Figs. 2 and 3. Vertical grey lines correspond to US Drought Monitor thresholds, identically to Fig. 4.

Table 1. Annual return periods (in years) using daily simulation for various SPI thresholds and accumulation periods.

Accumulation period	Months	Days	Return period (years) for SPI threshold								
			-4	-3.5	-3	-2.5	-2	-1.5	-1	-0.5	0
0.5		15	185	28.2	5.97	1.99	1.14	1.01	1.00	1.00	1.00
1		30	289	45.4	9.61	2.97	1.43	1.06	1.00	1.00	1.00
2		60	473	76.8	16.4	4.83	2.05	1.26	1.04	1.00	1.00
3		90	659	107	22.9	6.60	2.65	1.48	1.11	1.02	1.00
6		182	1060	182	40.1	11.5	4.32	2.15	1.39	1.11	1.02
9		274	1440	251	55.6	15.8	5.78	2.73	1.65	1.22	1.06
12		365	1820	315	69.4	19.6	7.04	3.24	1.88	1.33	1.11
24		730	3520	569	118	31.6	10.8	4.65	2.50	1.63	1.26

this difference is solely due to the structural behaviour of the SPI's moving average and is important to understand when comparing SPI extreme occurrences from different accumulation periods.

Return periods for extreme SPI values rapidly increase beyond -3.5 (Fig. S14). Return periods from daily data for an SPI of -4 range from 185 to 3520 years, depending on the accumulation period, and from 14 600 to 291 000 years for an SPI equal to -5 . These values differ from those generated by Stagge et al. (2016), who focused on recurrence within a

given day of the year rather than on annual minima. However, both our study and that of Stagge et al. (2016) convey the same concern that using the SPI or other normalized drought indices to quantify tail behaviour at such extremely low probabilities is dubious given common record lengths of 100 years or less.

The difference in return period for annual SPI minima from a theoretical time series as noted here is solely due to the structural persistence caused by the size of the moving window. In case study sites using the SPI, structural

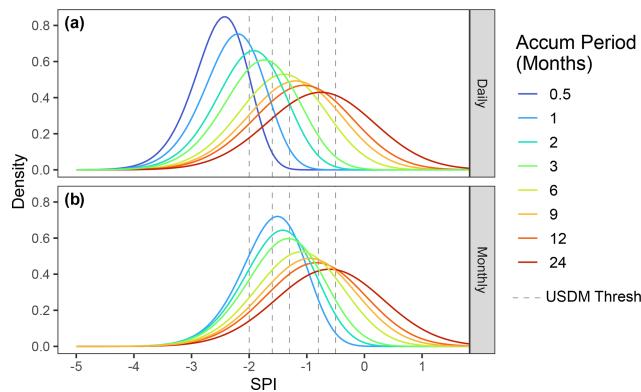


Figure 6. Fitted distribution of annual minima from daily (a) and monthly (b) time series. Colours and vertical lines are identical to Figs. 4 and 5.

persistence appears to play the predominant role in temporal autocorrelation (Fig. 3). The effect of structural persistence on annual minima can be partially explained because shorter accumulation periods necessarily have larger innovations (\sqrt{q}/q), leading to more erratic behaviour while maintaining the same overall standard normal distribution for each day of the year. As accumulation periods become larger, the moving average incorporates more individual days or months, requiring more sustained anomalies to produce extreme values. For the longest accumulation periods, where the moving-average window becomes longer than a year, the resulting time series slowly transitions from positive to negative or vice versa over the course of multiple years, thereby even producing occasional years in which the annual minima are greater than zero (Figs. 4 and 5).

3.4 Effect of temporal resolution

The temporal resolution of the underlying data (monthly or daily) has a strong impact on the annual minima of the simulated SPI. Using daily data shifts the distribution of the annual minima to become more extreme (more negative) across all accumulation periods, though the effect is strongest for short accumulation periods (Figs. 6 and 7). In turn, this makes return periods for monthly resolution data longer, even when considering the same threshold (Fig. 7). For example, the SPI-3 is likely to exceed -2 at least once every 2.65 years ($p = 0.378$) when using daily data but will likely only exceed this threshold once every 5.40 years ($p = 0.185$) when using monthly data (Tables 1 and 2). For the SPI-12, return periods for daily and monthly series become 7.04 and 10.44 years, respectively, for the -2 threshold. This is tied to the number of random samples (12 vs. 365) and the increased likelihood of a random extreme outlier despite moving windows of equal lengths.

There is little difference between the higher moments (variance, skewness, kurtosis) when comparing the distribu-

tion of annual extremes generated from daily or monthly data (Figs. S12 and S13) despite the moving window representing the same proportion of the year. Seemingly, the only difference is in the distribution mean, which is shifted to the more extreme (lower) when using daily data (Figs. S12 and S13). Because monthly and daily SPI data are often used interchangeably, the effect of the temporal resolution on annual minima is important for practitioners and drought monitoring agencies to acknowledge as it is purely an artefact of the calculation procedure and not of the climate.

4 Discussion

4.1 Structural persistence vs. real-world persistence

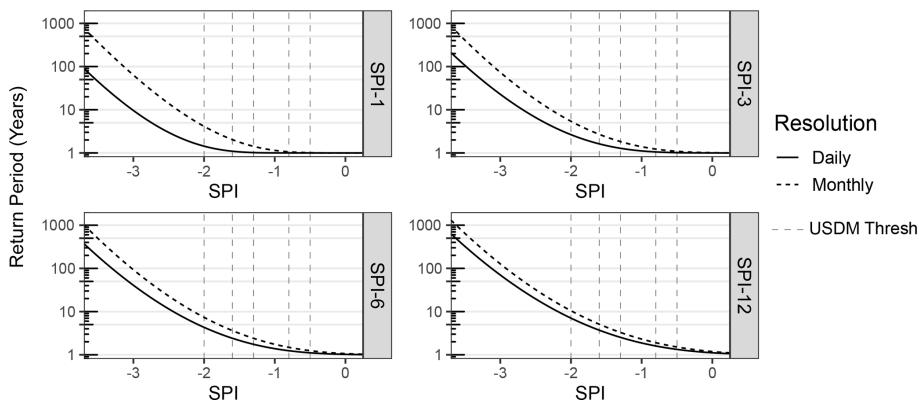
This study is, to our knowledge, the first attempt to quantify the return frequency of annual minima from a standardized drought index following a moving-average structure. To accomplish this, we defined the behaviour of a stochastic model that mimics the moving average of the SPI (Eqs. 2 and 3). In practice, the SPI and other drought indices can deviate from this general model in several ways, described next. Therefore, these results represent annual minima return periods for a highly idealized system as a bounding case, considering only structural persistence and Gaussian (symmetrical) innovations.

Normalized drought indices derived from gauge data can be subject to climatological persistence in addition to structural persistence. Structural persistence is caused by the moving-window structure, whereas climatological persistence is caused by climatological patterns, including frontal systems at shorter timescales and teleconnections like El Niño or the North Atlantic Oscillation at longer timescales. Another potential deviation from our theoretical SPI model is the seasonal regime, which may cause certain seasons to be more strongly correlated or to undergo rapid overturning of conditions. A final deviation from the real world is our assumption of symmetrical innovations (Eq. 2), which may not always hold true. This is particularly relevant for regions with low absolute precipitation, where individual large storm events may produce more extreme positive innovations than negative innovations.

Despite the potential for deviations from the idealized moving-average model that is the focus of this study, we found that most observed SPI series follow the expected structural persistence (Fig. 3). This finding suggests that these results are relevant for most SPI series, though one should check the persistence structure if they are to use this approach. Though the 23 global sites tested here do not show major climatological persistence, there may be regions with notable deviations from the assumed structural persistence created by the moving window. Testing of the persistence structure for SPI, SPEI, and other moving-average drought indices on a global, gridded scale would be needed to fully

Table 2. Annual return periods (in years) using monthly simulation for various SPI thresholds and accumulation periods.

Accumulation period	Return period (years) for SPI threshold								
	Months	-4	-3.5	-3	-2.5	-2	-1.5	-1	-0.5
1	2660	378	64.5	14	4.11	1.77	1.15	1.01	1.00
2	2830	391	67.6	15.3	4.7	2.05	1.27	1.05	1.00
3	2950	415	73.9	17.2	5.4	2.34	1.39	1.09	1.01
6	3290	489	92.7	22.8	7.34	3.13	1.75	1.24	1.06
9	3260	526	107	27.4	9.02	3.81	2.05	1.39	1.12
12	3840	613	124	31.8	10.4	4.37	2.31	1.51	1.19
24	6020	918	180	45.2	14.5	5.85	2.96	1.83	1.34

**Figure 7.** Return period comparison between daily (solid) and monthly (dotted) underlying data, where the SPI value is presented on the x axis and the return period is shown on the y axis in a log scale. USDM thresholds are shown, as in previous figures.

confirm this finding and should be addressed in future research.

4.2 Theoretical basis against GEV

Extreme-value theory, based on the work of Berman (1964), suggests that, for time series with high autocorrelation, including the moving average, the annual minima should asymptotically converge towards a member of the GEV distribution, namely the Gumbel distribution (Davis and Resnick, 1988; Eichner et al., 2006; Hirtzel, 1985b; Husler, 1990; Leadbetter et al., 1983; Rootzen, 1986). Notably, our simulations deviate from this, instead converging towards the generalized normal distribution (Figs. 2 and S1).

One potential explanation for this deviation is that the moving-average structure imparts extremal clustering (Coles, 2001; Moloney et al., 2019) to the time series, which violates the assumptions underlying the Berman theorem. Extremal clustering, the tendency for a time series to cluster at extreme levels, is quantified by the extremal index θ (Coles, 2001; Moloney et al., 2019). Clustering of extremes differs from temporal autocorrelation as it only considers clustering for events above a given extreme threshold (Auld and Papastathopoulos, 2021; Lindgren et al., 1983; Moloney

et al., 2019). The extremal index θ ranges from 0 to 1, representing completely independent extremes ($\theta = 1$), and progressively smaller values of θ , representing larger degrees of extremal clustering. The Berman theorem (Berman, 1964) and its subsequent derivations for moving-average sequences (Rootzen, 1986) are predicated on minimal extremal clustering. However, the moving averages used in normalized drought indices have increasing levels of extremal clustering for longer accumulation periods (Fig. S15, lower θ with higher accumulation). These levels of extremal clustering are still relatively minor but may partially explain the deviation from the GEV towards the generalized normal distribution, especially for longer accumulation periods (Fig. 2). More research would be required to confirm this hypothesized cause.

From a physical perspective, it is logical that moving-average drought indices have increased extremal clustering, particularly for long accumulation periods, because the longer moving-window approaches or surpasses the size of the annual block. Long moving-average windows produce relatively small incremental changes, which makes it unlikely to reach extreme minima in a given year without the preceding year already being quite low. This, in addition to smearing drought events across neighbouring calendar years, leads to greater levels of extremal clustering and lower val-

ues of θ (Fig. S15). Eichner et al. (2006) noted similar behaviour for annual block maxima derived from an autocorrelated series, finding that distributions become more normally distributed with higher degrees of correlation.

5 Conclusions

This study represents an advancement in the understanding of annual extremes derived from moving-average time series, of which normalized drought indices like the SPI are an important type. The SPI and other normalized drought indices are used by drought monitoring agencies throughout the world to quantify the relative severity of droughts and to classify conditions into discrete drought states. Because of the importance placed on these indices, this study explored the behaviour of a theoretical, idealized moving-window sequence with respect to annual minima. The major advances shown here can be viewed from two perspectives: an improved theoretical understanding of moving-window sequences and practical findings for the application of drought indices by drought monitoring agencies when setting decision thresholds or when communicating risk to the public.

From a theoretical perspective, this study presents a stochastic model to simulate a moving-average process where the moving-average window is proportionally large (5%–200%) relative to the year (Eqs. 2 and 3). This produced the first, to our knowledge, explicit quantification of annual extreme exceedances from such a sequence. We showed that the distribution of annual minima follows a generalized normal distribution rather than the GEV distribution, which was the initial expectation from extreme-value theory. This deviation is likely due to extremal clustering.

From an applied perspective, this study provides the expected annual return periods for the SPI or related drought indices, with common accumulation periods ranging from 1 to 24 months (Fig. 5, Tables 1 and 2). We show that the likelihood of exceeding an SPI threshold in a given year decreases (annual return period increases) as the accumulation period increases. The corollary of this finding is also true; the SPI threshold associated with a given annual return period becomes less extreme (closer to zero) for indices with longer accumulation periods. Practitioners have implicitly understood this relationship, even from the first definition of the SPI (McKee et al., 1993), which included a figure showing the number of unique droughts per 100 years decreasing with longer accumulation periods for a gauge in Fort Collins, CO, USA. Likewise, the European Drought Observatory's Combined Drought Index uses a more extreme threshold for short accumulation periods ($SPI1 < -2$) than for longer accumulation periods ($SPI3 < -1$) when classifying drought regions, which is in line with our findings, which suggest that these thresholds should have similar annual return periods (1–1.5 years) for daily data (Fig. 5, Table 1).

Despite an implicit understanding by drought practitioners, this study is the first to explicitly calculate the return period for a theoretical normalized drought index using a moving average. Drought managers can use this knowledge to better interpret exceedances of the SPI or other drought metrics.

A second practical finding is that annual minima from a normalized drought index (e.g. the SPI or SPEI) depend on whether one uses daily or monthly data, even for the same accumulation period. Drought practitioners have largely understood that daily data are noisier and, thus, more subject to single-day deviations towards extremes, but this effect has not been quantified explicitly to date. Further, researchers tend to use daily or monthly resolution data interchangeably if the accumulation periods are equivalent, which, as we have shown here, produce different results when viewed as annual exceedances.

We therefore propose several recommendations. The first is a general recommendation for users of the SPI and related drought indices to be careful with language and thoughtful about interpretation when using a normalized drought index to determine whether a given event is particularly extreme. Our goal is to clarify the difference between the probability of the SPI exceeding -2 on a specific day and the probability of it exceeding -2 in a given year (commonly called the return period). The former is the definition of the SPI, whereas the latter is covered in this study. This distinction should be clear when considering the fact that the likelihood of $SPI < -2$ on any given day is 2.3 %, but the likelihood of experiencing $SPI < -2$ in a year ranges from 6.9 % to 87.7 %, depending on the accumulation period and the temporal resolution. This leads to the second recommendation, which is that practitioners should exercise caution when comparing the likelihood or severity of a particularly extreme SPI value using a short accumulation period with that of one using a long accumulation period. It is most appropriate to compare an index with itself, and, if one must make cross-index comparisons, there may be a need to use different thresholds, as is done by the European Drought Observatory's Combined Drought Index.

Our final recommendation is for more research to explore the findings shown here. For example, more research should explore the degree to which climatological persistence and seasonality affect temporal autocorrelation across a range of climate regions and drought indices, expanding on the example sites and SPI values tested in Sect. 3.2. The theoretical case developed here should act as a baseline under an idealized moving-average time series. However, deviations from this may be illustrative. While this result was meant as the simplest case, using the block annual minima for an idealized time series, future research could explore more drought-specific indices like duration.

Data availability. All data are available via an open-access repository (<https://doi.org/10.5281/zenodo.14814790>, Stagge et al., 2025).

Code availability. All code is available via an open-access repository (<https://doi.org/10.5281/zenodo.14814790>, Stagge et al., 2025). This code has been tested to generate all figures and tables presented in this paper.

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/hess-29-719-2025-supplement>.

Author contributions. JHS: conceptualization, methodology, formal analysis, writing (original draft). KS, IM, and MAIH: conceptualization, writing (review and editing).

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Special issue statement. This article is part of the special issue "Drought, society, and ecosystems (NHESS/BG/GC/HESS inter-journal SI)". It is not associated with a conference.

Acknowledgements. The authors would like to thank Peter Craigmille for discussions regarding this paper and its theoretical underpinnings. The authors would also like to acknowledge support from the Byrd Polar and Climate Research Center at the Ohio State University.

Financial support. This research has been supported by the US National Science Foundation Directorate for Geosciences (grant no. 2002539).

Review statement. This paper was edited by Floris van Ogtrop and reviewed by two anonymous referees.

References

Akaike, H.: Information Theory and an Extension of the Maximum Likelihood Principle, in: Selected Papers of Hirotugu Akaike, edited by: Parzen, E., Tanabe, K., and Kitagawa, G., Springer, New York, NY, 199–213, https://doi.org/10.1007/978-1-4612-1694-0_15, 1998.

Anderson, M. C., Hain, C., Otkin, J., Zhan, X., Mo, K., Svoboda, M., Wardlow, B., and Pimstein, A.: An Inter-comparison of Drought Indicators Based on Thermal Remote Sensing and NLDAS-2 Simulations with U.S. Drought Monitor Classifications, *J. Hydrometeorol.*, 14, 1035–1056, <https://doi.org/10.1175/JHM-D-12-0140.1>, 2013.

Auld, G. and Papastathopoulos, I.: Extremal clustering in non-stationary random sequences, *Extremes*, 24, 725–752, <https://doi.org/10.1007/s10687-021-00418-2>, 2021.

Basu, B. and Srinivas, V. V.: Formulation of a mathematical approach to regional frequency analysis, *Water Resour. Res.*, 49, 6810–6833, <https://doi.org/10.1002/wrcr.20540>, 2013.

Beguería, S., Vicente-Serrano, S. M., Reig, F., and Latorre, B.: Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring, *Int. J. Climatol.*, 34, 3001–3023, <https://doi.org/10.1002/joc.3887>, 2014.

Berman, S. M.: Limit Theorems for the Maximum Term in Stationary Sequences, *Ann. Math. Stat.*, 35, 502–516, <https://doi.org/10.1214/aoms/117703551>, 1964.

Blauthut, V., Stahl, K., Stagge, J. H., Tallaksen, L. M., De Stefanis, L., and Vogt, J.: Estimating drought risk across Europe from reported drought impacts, drought indices, and vulnerability factors, *Hydrol. Earth Syst. Sci.*, 20, 2779–2800, <https://doi.org/10.5194/hess-20-2779-2016>, 2016.

Bloomfield, J. P. and Marchant, B. P.: Analysis of groundwater drought building on the standardised precipitation index approach, *Hydrol. Earth Syst. Sci.*, 17, 4769–4787, <https://doi.org/10.5194/hess-17-4769-2013>, 2013.

Box, G. E. and Jenkins, G. M. (Eds.): *Time Series Analysis, Forecasting and Control*, Holden Day, San Francisco, ISBN 0816211043, 1970.

Cammalleri, C., Arias-Muñoz, C., Barbosa, P., de Jager, A., Magni, D., Masante, D., Mazzeschi, M., McCormick, N., Naumann, G., Spinoni, J., and Vogt, J.: A revision of the Combined Drought Indicator (CDI) used in the European Drought Observatory (EDO), *Nat. Hazards Earth Syst. Sci.*, 21, 481–495, <https://doi.org/10.5194/nhess-21-481-2021>, 2021.

Cavanaugh, J. E. and Neath, A. A.: The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements, *WIREs Comput. Stat.*, 11, e1460, <https://doi.org/10.1002/wics.1460>, 2019.

Christian, J. I., Basara, J. B., Otkin, J. A., Hunt, E. D., Wakefield, R. A., Flanagan, P. X., and Xiao, X.: A Methodology for Flash Drought Identification: Application of Flash Drought Frequency across the United States, *J. Hydrometeorol.*, 20, 833–846, <https://doi.org/10.1175/JHM-D-18-0198.1>, 2019.

Coles, S.: *An Introduction to Statistical Modeling of Extreme Values*, Springer, London, <https://doi.org/10.1007/978-1-4471-3675-0>, 2001.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I.,

Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, *Q. J. Roy. Meteorol. Soc.*, 137, 1–28, <https://doi.org/10.1002/qj.776>, 2011.

Das, S.: Goodness-of-Fit Tests for Generalized Normal Distribution for Use in Hydrological Frequency Analysis, *Pure Appl. Geophys.*, 175, 3605–3617, <https://doi.org/10.1007/s00024-018-1877-y>, 2018.

Davidson, J. E. H.: Problems with the estimation of moving average processes, *J. Econom.*, 16, 295–310, [https://doi.org/10.1016/0304-4076\(81\)90032-4](https://doi.org/10.1016/0304-4076(81)90032-4), 1981.

Davis, R. and Resnick, S.: Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution, *Stoch. Process. Their Appl.*, 30, 41–68, [https://doi.org/10.1016/0304-4149\(88\)90075-0](https://doi.org/10.1016/0304-4149(88)90075-0), 1988.

Davis, R. A. and Resnick, S. I.: Extremes of Moving Averages of Random Variables with Finite Endpoint, *Ann. Probabil.*, 19, 312–328, 1991.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface, *B. Am. Meteorol. Soc.*, 87, 1381–1398, doi:10.1175/BAMS-87-10-1381, 2006.

Eichner, J. F., Kantelhardt, J. W., Bunde, A., and Havlin, S.: Extreme value statistics in records with long-term persistence, *Phys. Rev. E*, 73, 016130, <https://doi.org/10.1103/PhysRevE.73.016130>, 2006.

Granger, C. W. J. and Andersen, A.: On the invertibility of time series models, *Stoch. Process. Their Appl.*, 8, 87–92, [https://doi.org/10.1016/0304-4149\(78\)90069-8](https://doi.org/10.1016/0304-4149(78)90069-8), 1978.

Guttman, N. B.: Accepting the Standardized Precipitation Index: A Calculation Algorithm, *J. Am. Water Resour. Assoc.*, 35, 311–322, <https://doi.org/10.1111/j.1752-1688.1999.tb03592.x>, 1999.

Hallin, M.: Spectral factorization of nonstationary moving average processes, *Ann. Stat.*, 12, 172–192, 1984.

Hao, Z., AghaKouchak, A., Nakhjiri, N., and Farahmand, A.: Global integrated drought monitoring and prediction system, *Sci. Data*, 1, 140001, <https://doi.org/10.1038/sdata.2014.1>, 2014.

Heim, R. R.: A Review of Twentieth-Century Drought Indices Used in the United States, *B. Am. Meteorol. Soc.*, 83, 1149–1165, [https://doi.org/10.1175/1520-0477\(2002\)083<1149:AROTDI>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<1149:AROTDI>2.3.CO;2), 2002.

Heim, R. R. and Brewer, M. J.: The Global Drought Monitor Portal: The Foundation for a Global Drought Information System, *Earth Interact.*, 16, 1–28, <https://doi.org/10.1175/2012EI000446.1>, 2012.

Hirtzel, C. S.: Extreme values of autocorrelated sequences, *Appl. Math. Comput.*, 16, 327–345, [https://doi.org/10.1016/0096-3003\(85\)90014-1](https://doi.org/10.1016/0096-3003(85)90014-1), 1985a.

Hirtzel, C. S.: Statistics of extreme values of a first-order Markov normal process: An exact result, *Atmos. Environ.*, 19, 1207–1209, [https://doi.org/10.1016/0004-6981\(85\)90305-1](https://doi.org/10.1016/0004-6981(85)90305-1), 1985b.

Hosking, J. R. M. and Wallis, J. R.: A Comparison of Unbiased and Plotting-Position Estimators of L Moments, *Water Resour. Res.*, 31, 2019–2025, <https://doi.org/10.1029/95WR01230>, 1995.

Hosking, J. R. M. and Wallis, J. R.: Regional frequency analysis, Cambridge University Press, Cambridge, UK, 244 pp., ISBN 0521019400, 1997.

Husler, J.: Extreme Values and High Boundary Crossings of Locally Stationary Gaussian Processes, *Ann. Probabil.*, 18, 1141–1158, <https://doi.org/10.1214/aop/1176990739>, 1990.

Kim, H.: Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set], Data Integration and Analysis System (DIAS) [data set], <https://doi.org/10.20783/DIAS.501>, 2017.

Lawrimore, J., Heim, R. R., Svoboda, M., Swail, V., and Englehart, P. J.: Beginning a new era of drought monitoring across North America, *B. Am. Meteorol. Soc.*, 83, 1191–1192, <https://doi.org/10.1175/1520-0477-83.8.1191>, 2002.

Leadbetter, M. R., Lindgren, G., and Rootzén, H.: Nonstationary, and Strongly Dependent Normal Sequences, in: *Extremes and Related Properties of Random Sequences and Processes*, edited by: Leadbetter, M. R., Lindgren, G., and Rootzén, H., Springer, New York, NY, 123–141, https://doi.org/10.1007/978-1-4612-5449-2_6, 1983.

Lindgren, G., Leadbetter, M. R., and Rootzén, H.: Extremes and related properties of stationary sequences and processes, Springer, New York, <https://doi.org/10.1007/978-1-4612-5449-2>, 1983.

Lisonbee, J., Woloszyn, M., and Skumanich, M.: Making sense of flash drought: definitions, indicators, and where we go from here, *J. Appl. Serv. Climatol.*, 2021, 1–19, 10.46275/JOASC.2021.02.001, 2022.

Lloyd-Hughes, B.: The impracticality of a universal drought definition, *Theor. Appl. Climatol.*, 117, 607–611, <https://doi.org/10.1007/s00704-013-1025-7>, 2014.

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

McKee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, in: *Proceedings of the 8th Conference on Applied Climatology*, 17–22 January 1993, Anaheim, CA, 179–183, 1993.

Meyer-Christoffer, A., Becker, A., Finger, P., Rudolf, B., Schneider, U., and Ziese, M.: GPCC Climatology version 2015 at 0.25°: Monthly land-surface precipitation climatology for every month and the total year from rain-gauges built on GTS-based and historic data, Global Precipitation Climatology Centre at Deutscher Wetterdienst, 10 pp., https://doi.org/10.5676/DWD_GPCC/CLIM_M_V2015_025, 2015.

Moloney, N. R., Faranda, D., and Sato, Y.: An overview of the extremal index, *Chaos Interdisciplin. J. Nonlin. Sci.*, 29, 022101, <https://doi.org/10.1063/1.5079656>, 2019.

Peel, M. C., Wang, Q. J., Vogel, R. M., and McMahon, T. A.: The utility of L-moment ratio diagrams for selecting a regional probability distribution, *Hydrolog. Sci. J.*, 46, 147–155, <https://doi.org/10.1080/02626660109492806>, 2001.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, *Hydrol. Earth Syst. Sci.*, 11, 1633–1644, <https://doi.org/10.5194/hess-11-1633-2007>, 2007.

Rootzen, H.: Extreme Value Theory for Moving Average Processes, *Ann. Probabil.*, 14, 612–652, 1986.

Sangal, B. P. and Biswas, A. K.: The 3-Parameter Lognormal Distribution and Its Applications in Hydrology, *Water Resour. Res.*, 6, 505–515, <https://doi.org/10.1029/WR006i002p00505>, 1970.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., and Ogallo, L.: A Drought Monitoring and Forecasting System for Sub-Saharan African Water Resources and Food Security, *B. Am. Meteorol. Soc.*, 95, 861–882, <https://doi.org/10.1175/BAMS-D-12-00124.1>, 2014.

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

Stagge, J. H.: Data repository “jstagge/spi_returnperiod: Version associated with published manuscript”, v1.1.1, Zenodo [code and data set], <https://doi.org/10.5281/zenodo.14814790>, 2025.

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

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

Stagge, J. H., Kohn, I., Tallaksen, L. M., and Stahl, K.: Modeling drought impact occurrence based on meteorological drought indices in Europe, *J. of Hydrol.*, 530, 37–50, <https://doi.org/10.1016/j.jhydrol.2015.09.039>, 2015b.

Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F., and Stahl, K.: Response to comment on ‘Candidate Distributions for Climatological Drought Indices (SPI and SPEI)’, *Int. J. Climatol.*, 36, 2132–2138, <https://doi.org/10.1002/joc.4564>, 2016.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., and Stephens, S.: The Drought Monitor, *B. Am. Meteorol. Soc.*, 83, 1181–1190, <https://doi.org/10.1175/1520-0477-83.8.1181>, 2002.

Tallaksen, L. M. and Van Lanen, H. A.: Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater, in: 2nd Edn., Elsevier, <https://doi.org/10.1016/C2017-0-03464-X>, 2023.

Tsoukalas, I.: The tales that the distribution tails of non-Gaussian autocorrelated processes tell: efficient methods for the estimation of the k -length block-maxima distribution, *Hydrolog. Sci. J.*, 67, 898–924, <https://doi.org/10.1080/02626667.2021.2014056>, 2022.

Vicente-Serrano, S., Beguería, S., and López-Moreno, J.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, *J. Climate*, 23, 1696–1718, 2010.

Wilks, D. S.: Chapter 9 – Time Series, in: *International Geophysics*, Volume 100, edited by: Daniel, S. W., Academic Press, 395–456, <https://doi.org/10.1016/B978-0-12-385022-5.00009-9>, 2011.

WMO and GWP – World Meteorological Organization and Global Water Partnership: *Handbook of Drought Indicators and Indices, Integrated Drought Management Programme (IDMP)*, Geneva, ISBN 978-92-63-11173-9, 2016.

Xia, Y., Ek, M. B., Peters-Lidard, C. D., Mocko, D., Svoboda, M., Sheffield, J., and Wood, E. F.: Application of USDM statistics in NLDAS-2: Optimal blended NLDAS drought index over the continental United States, *J. Geophys. Res.-Atmos.*, 119, 2947–2965, <https://doi.org/10.1002/2013JD020994>, 2014.