**Hydrology and Earth System Sciences**

Open Access

# Hydrologic benchmarking of meteorological drought indices at interannual to climate change timescales: a case study over the Amazon and Mississippi river basins

E. Joetzjer[1], H. Douville[1], C. Delire[1], P. Ciais[2], B. Decharme[1], and S. Tyteca[1]

[1]CNRM-GAME – UMR3589, Groupe d'étude de l'atmosphère météorologique, 42 avenue G. Coriolis,
31057 Toulouse, France
[2]Laboratory of Climate Sciences and the Environment, L'Orme des Merisiers, 91191 Gif-sur-Yvette, France

*Correspondence to:* E. Joetzjer (emilie.joetzjer@meteo.fr)

**Abstract.** Widely used metrics of drought are still derived solely from analyses of meteorological variables such as precipitation and temperature. While drought is generally a consequence of atmospheric anomalies, the impacts to society are more directly related to hydrologic conditions. The present study uses a standardized runoff index (SRI) as a proxy for river discharge and as a benchmark for various meteorological drought indices (scPDSI, SPI, SPEI_th, and SPEI_hg respectively). Only 12-month duration droughts are considered in order to allow a direct (no river routing) comparison between meteorological anomalies and their hydrological counterpart. The analysis is conducted over the Mississippi and Amazon river basins, which provide two contrasted test beds for evaluating drought indices at both interannual (using detrended data) and climate change (using raw data) timescales. Looking first at observations over the second half of the 20th century, the simple SPI based solely on precipitation is no less suitable than more sophisticated meteorological drought indices at detecting interannual SRI variations. Using the detrended runoff and meteorological outputs of a five-member single model ensemble of historical and 21th century climate simulations leads to the same conclusion. Looking at the 21st century projections, the response of the areal fraction in drought to global warming is shown to be strongly metric dependent and potentially overestimated by the drought indices which account for temperature variations. These results suggest that empirical meteorological drought indices should be considered with great caution in a warming climate and that more physical water balance models are needed to account for the impact of the anthropogenic radiative forcings on hydrological droughts.

## 1 Introduction

Droughts are recurrent natural manifestations of climate variability that have plagued civilizations throughout history. They are often commonly classified into three types – meteorological, agricultural and hydrological – depending on which variable – respectively precipitation, soil moisture and river flow – is below normal conditions (Dai, 2011a). Meteorological drought often precedes and causes other types of droughts. Meteorological indices are therefore used not only for monitoring drought at regional to global scales, but also for anticipating their potential impacts on agriculture and water resources.

Several empirical meteorological drought indices have been proposed and applied at regional to global scales over the second half of the 20th century (e.g. Heim Jr., 2002). Nevertheless, evidence is building that human-induced climate change is perturbing the global hydrological cycle (e.g. Trenberth, 2011), making it necessary to analyse the validity of such indices in a warmer climate. While most 21st century climate scenarios project a global increase in the frequency, intensity and duration of droughts (Sheffield and Wood, 2008; Orlowsky and Seneviratne, 2012), the response

is still very uncertain at the regional scale and is not necessarily consistent from one metric to the other (e.g. Burke and Brown, 2008).

In the IPCC Fourth Assessment Report, the 20th century multi-decadal variations of drought were mainly discussed on the basis of the Palmer Drought Severity Index (PDSI, Palmer, 1965). This standardized index measures the departure of soil moisture using a simplified surface water balance model. It requires globally available precipitation ($P$) and temperature data as input for the calculation of potential evapotranspiration (PET) with Thornthwaite's (1948) equation, as well as the soil water field capacity. Analysis of global PDSI maps indicates that drought has generally increased throughout the 20th century (Dai et al., 2004). The PDSI was however criticized in several respects (e.g. Guttman, 1998; Vicente-Serrano et al., 2011). The underlying water balance model is quite empirical and was tuned using a limited number of instrumented sites in the US. This limitation was addressed by the development of the "self-calibrated" scPDSI (Wells et al., 2004). The empirically derived climate parameters and duration factors of this index are automatically calculated using the historical climatic data of the selected location. The Thornthwaite's approximation for the computation of PET was also criticized and recently replaced by a more physical but still empirical Penman–Monteith approach (Van der Schrier et al., 2011; Sheffield et al., 2012). Finally, it was argued that the PDSI cannot reflect the different timescales which characterize the impact of drought on different systems, including the surface hydrology (Vicente-Serrano et al., 2010).

In contrast, the Standardized Precipitation Index (SPI) of McKee et al. (1995) is a multi-scale index, computed as a standardized transform of cumulative precipitation over a given period (ranging typically between 1 and 48 months), but does not account for possible variations in the atmospheric demand. More recently, Vicente-Serrano et al. (2010) developed the Standardized Precipitation Evapotranspiration Index (SPEI) by applying a similar transform on cumulated $P$ minus PET. The aim was to combine the simplified water balance approach of the PDSI and the multi-scale nature of the SPI.

The superiority of the SPEI is however a matter of debate (Dai, 2011b). In spite of the criticisms of Guttman (1998) or Vicente-Serrano et al. (2010), the PDSI has been evaluated successfully at the regional or basin scale against both soil moisture and river discharge (Dai et al., 2004). Moreover, it compares relatively well with the 12-month SPEI (Vicente Serrano et al., 2011). In the recent IPCC SREX report on managing the risks of extreme events and disasters to advance climate change adaptation (Seneviratne et al., 2012), the PDSI was still used as a reference drought index, but the metric sensitivity of drought projections was highlighted as well as the need for more comprehensive comparisons of the various globally available drought indices.

The aim of the present study is to use a hydrologic drought index as a benchmark for assessing the variability of several meteorological drought indices at both interannual and climate change timescales. Given the limited instrumental record, the comparison will be conducted with both observations and an ensemble of global climate simulations spanning the 1850–2100 period. The simulations will allow us to test the robustness not only of the comparison made at the interannual timescale, but also of drought projections based on different meteorological indices.

We chose two among the world's largest river basins – Amazon and Mississippi – as a test bed for our analysis. While it would be interesting to extend the study to a larger number of basins, we believe that this subset is sufficient to illustrate our main findings (which are moreover confirmed by a global comparison of meteorological vs. hydrological time series simulated at each continental grid cell). Both Amazon and Mississippi are well documented in terms of climate and river discharge observations and are not too much influenced by human activities (dams and irrigation). Both show a substantial year-to-year variability (including during the dry season) of river discharge and a potential vulnerability to climate change. Nevertheless, they show contrasted climatological features. Precipitation in the Amazon Basin has a stronger annual cycle and a larger interannual variability than in the Mississippi Basin. The opposite is true for temperature and therefore for the atmospheric water demand (PET). These features are representative of the contrast between tropical and mid-latitude areas and might have consequences on the behaviour of the analysed meteorological drought indices.

For such large river basins, meteorological droughts generally precede their hydrological counterpart by a few weeks or months. In order to guarantee the relevance of our hydrological benchmark and to avoid the use of a river routing model, the focus will be only on 12-month droughts. Short-term droughts are therefore beyond the scope of the present study although they can be detected on a 12-month timescale if they show a sufficient magnitude and if the rest of the year is close to normal conditions. In other words, the 12-month deficit can be concentrated on a particular season, but we do not distinguish between wet-season versus dry-season droughts, which might have contrasted impacts on natural ecosystems and human societies.

Section 2 describes the input data (derived from either observations or climate simulations) and the methodology used for the calculation of both meteorological drought indices and the hydrologic benchmark. Section 3 first compares the ability of the different meteorological indices to capture the interannual variability of hydrological drought, as well as their skill to detect major hydrological droughts. Indices derived from the CNRM-CM5 climate scenarios are also analysed to highlight the contrasted sensitivity of the different drought indices to climate change. Section 4 discusses the results and draws the main conclusions of the study.

**Table 1.** Summary of the data sets used in the present study.

|  | 20th century | 1850–2100 |
|---|---|---|
| Data | 1° monthly precipitation: GPCC version 5 (Rudolf et al., 2011) | 1.4° CNRM-CM5 |
|  | 0.5° monthly surface air temperature: CRU TS.3 (Mitchell and Jones, 2005) | historical simulation and RCP8.5 climate |
|  | Monthly river discharge: GRDC (http://www.bafg.de/GRDC) | change scenario |
|  | 1° runoff and river discharge derived from off-line simulations (1951–2006): SURFEX hydrological system based on the ISBA land surface model and the TRIP river routing model (Decharme and Douville, 2007) | (five members) |

**Table 2.** Summary of the drought indices used in the present study.

| Type | Index | Based on | Multiscalar | References |
|---|---|---|---|---|
| Meteorological drought | sc-PDSI | simplified water balance | no | Palmer (1965) (PDSI) Wells et al. (2004) (scPDSI) |
|  | SPI | standardized transform of cumulative precipitation | yes | McKee et al. (1995) |
|  | SPEI_th | standardized transform of cumulative precipitation minus PET calculated through Thornthwaite's formula | yes | Vicente-Serrano et al. (2010) |
|  | SPEI_hg | standardized transform of cumulative precipitation minus PET calculated through Hargreaves's formula | yes | Vicente-Serrano et al. (2010) |
| Hydrological drought | SRI | standardized transform of cumulative runoff | yes | Shukla and Wood (2008) |

## 2 Data sets and methodology

### 2.1 Observed and simulated drought indices

All meteorological drought indices (SPI, SPEI_th, SPEI_hg, cf. summary in Table 2) are derived solely from monthly precipitation and surface air temperature. As far as observations are concerned, the selected global 20th century data sets are summarized in Table 1. Model outputs (monthly precipitation and temperature, but also monthly runoff for the hydrologic benchmark) have been derived from a five-member ensemble of 1850–2100 simulations obtained with the CNRM-CM5 global climate model (Voldoire et al., 2013). Each realization is the concatenation of a historical (i.e. 1850–2005) simulation driven by observed concentrations of greenhouse gases and sulfate aerosols (as well as realistic volcanic and solar forcings) and of a 21st century (i.e. 2006–2100) projection based on the RCP8.5 concentration scenario (corresponding to a $8.5\,\mathrm{W\,m^{-2}}$ radiative forcing at the end of the 21st century) proposed by the phase 5 of the Coupled Model Intercomparison Project (CMIP5, http://cmip-pcmdi. llnl.gov/cmip5/).

Although the aim of the study is not to compare simulated versus observed drought indices, but meteorological indices versus the hydrologic benchmark in both model and observations, precipitation and temperature observations (see Table 1) were first interpolated onto the model horizontal grid (about 1.4°) to ensure the same spatial resolution for all indices. On each grid cell, the scPDSI and the 12-month SPI and SPEI (hereafter SPI12 and SPEI12 respectively) were computed following the original algorithms proposed by Wells et al. (2004), McKee et al. (1995) and Vicente-Serrano et al. (2009) respectively. Cumulated $P$ was fitted with a gamma function, while a log-logistic function was preferred for $P$ minus PET (Vicente-Serrano et al., 2009) for the SPEI. While the simple Thornthwaite equation was used to compute PET from temperature and latitude for SPEI (hereafter SPEI_th) and scPDSI, another empirical formulation (Hargreaves and Samani, 1982) accounting more accurately for the role of solar radiation was tested for SPEI (hereafter SPEI_hg). Unlike in Van der Schrier et al. (2011) or Sheffield et al. (2012), more sophisticated formulations such as Penman–Monteith have not been tested given the lack of reliable (satellite) global observations of solar radiation before the 1980s.

**Table 3.** Summary of the periods and methods used to compare meteorological vs. hydrological drought indices. Raw time series are used to analyse simulated climate change while detrended time series are preferred to discuss observed or simulated interannual variability.

| Data sets | Period | Detrending | Figures and tables |
|---|---|---|---|
| Gridded observations | 1951–1999 | Yes | Fig. 1 (meteorological indices) |
| Off-line hydrological simulation | 1951–1999 | Yes | Fig. 1 (SRI) |
| Coupled climate simulations | 1850–2100 | Yes | Fig. 2, Table 5 |
| | 1850–2100 | No | Figs. 3 and 4 |

**Table 4.** Schematic of the $2 \times 2$ contingency table used to assess the ability of the meteorological indices to detect a hydrological drought event: $A$ denotes the number of "hits", $B$ the number of "false alarms", $C$ the number of "misses", and $D$ the number of "correct non-events".

| | | Hydrological drought index: SRI12 | |
|---|---|---|---|
| | | $\leq x$th percentile | $> x$th percentile |
| Meteorological drought indices: SPI12, SPEI12_th, SPEI12_hg, or scPDSI | $\leq x$th percentile | $A$ | $B$ |
| | $> x$th percentile | $C$ | $D$ |

For all indices and in order to focus on interannual and longer timescales, annual mean values have been obtained by averaging monthly indices from January to December. Finally, basin average indices have been calculated, as well as the area of the basin in drought based on a common threshold (only for the simulated indices).

It must be here emphasized that the SPI and SPEI normalization was made in each grid cell before spatial averaging. While such a choice is somewhat arbitrary, it allows us to compute the areal fraction in drought (cf. Sect. 3.2) and to have a fair comparison with the PDSI, which is by definition a distributed index given the spatial variability of the soil water capacity (which is a key input parameter used in the simplified water balance model). Therefore, we have considered all drought indices as global gridded and monthly data sets that can be averaged in both space and time.

Hydrological drought has been defined using the SRI proposed by Shukla and Wood (2008), i.e. applying the same algorithm as for SPI12 but on the 12-month cumulated runoff. Runoff has been chosen rather than river discharge given the selected timescale (no need of a river routing model) and the possibility to compute the basin-averaged index and the areal fraction in drought exactly in the same way as for the meteorological indices. While runoff is a standard output of the CNRM-CM5 climate model, there is no observational counterpart so that we have used an off-line simulation of the ISBA land surface model (included in the CNRM-CM5 model) to produce a "pseudo-observed" gridded runoff. This was done by driving the ISBA land surface model with bias-corrected atmospheric reanalyses available over the 1951–2006 period (Alkama et al., 2011). In line with the comprehensive evaluation of Alkama et al. (2011), this "pseudo-observed" SRI12 (Fig. 1) is highly correlated with in situ river discharge observations over both Amazon

and Mississippi. This result makes us relatively confident about the relevance of our hydrologic SRI benchmark, which can be used to assess the behaviour of both observed and simulated meteorological drought indices. Moreover, it also means that the off-line ISBA simulation of land surface evapotranspiration is reasonable, at least at the annual timescale. This is the reason why we will also introduce a "Standardized Precipitation Actual Evapotranspiration Index" (SPAEI) by replacing PET by actual evapotranspiration in the SPEI algorithm. Note that the aim here is not to propose an alternative meteorological drought index given the difficulty to compute actual evapotranspiration from monthly observations, but just to highlight the consequences of the PET approximation in the SPEI algorithm.

## 2.2 Methodology

Before using the raw time series of the projected drought indices to assess the behaviour of the meteorological indices at the climate change timescale, the first step is to evaluate their interannual variability using both observations and simulations (see Table 3 for a summary of the selected periods and methods). For this purpose, and in order to get rid of the recent warming trend in each region, all basin-averaged indices have been detrended using cubic spline functions (Whaba, 1990; Ribes et al., 2010) with 2 and 4 degrees of freedom for detrending over a 49 and 251 yr time span respectively, before computing their correlation with the SRI12 benchmark. The Clayton skill score (CSS; Wilks, 2004), based on the probability for each index to be either above or below a given percentile of the distribution, has also been used to assess the ability of different indexes to detect major hydrological droughts. Given the contingency table given in Table 4, this skill score is simply computed as the difference between two conditional probabilities:
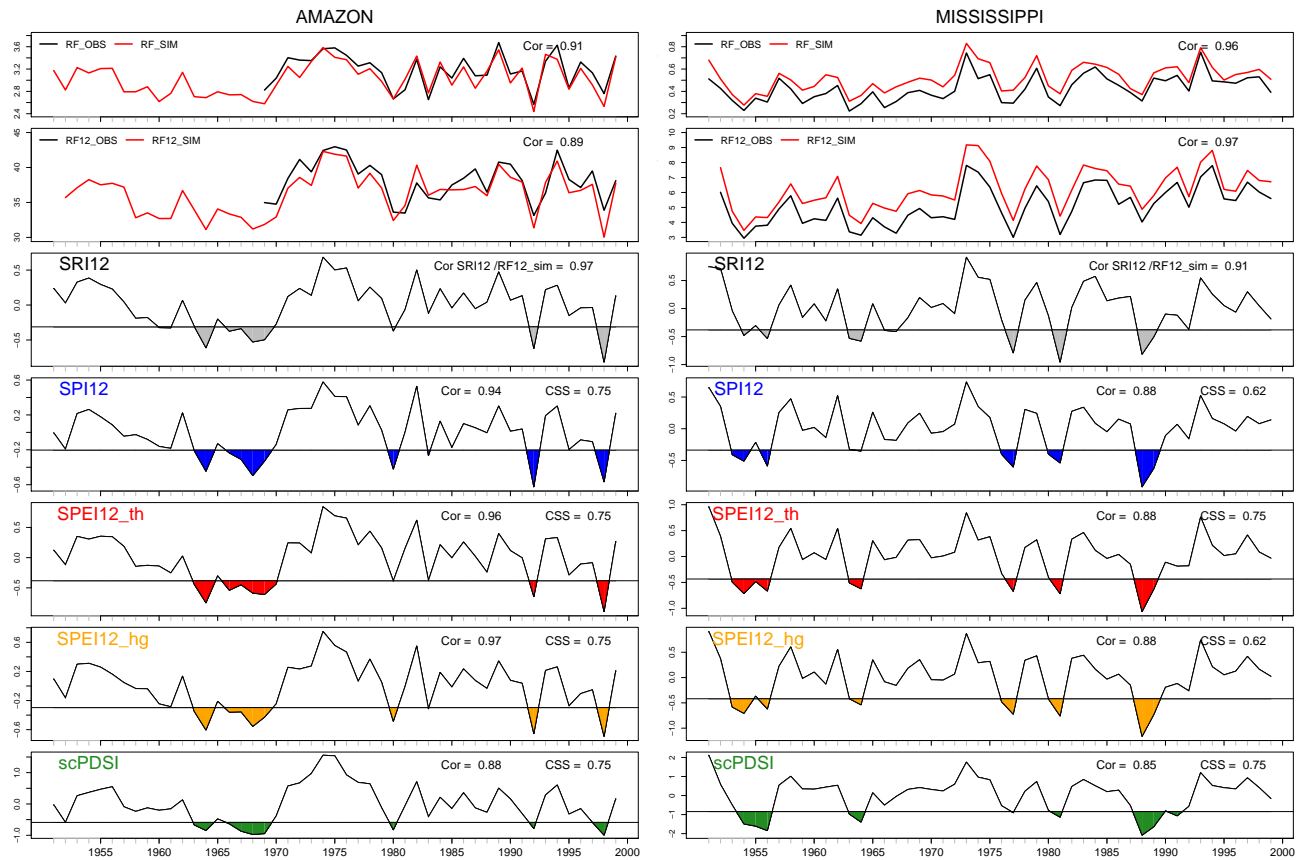
**Fig. 1.** Annual mean time series from 1951 to 1999 of river flow (RF mm day$^{-1}$), cumulated river flow over 12 months (RF12 mm day$^{-1}$), detrended SRI12, and detrended meteorological drought indices (SPI12, SPEI12_th, SPEI12_hg and scPDSI). The correlation and the CSSs between SRI12 and each meteorological index are indicated in the top right corner of each plot. For each index, droughts are defined under the 20th percentile and are shaded.

$$\mathrm{CSS} = \frac{A}{A+B} - \frac{C}{C+D},$$

where $A$ is the number of meteorological droughts detected by the index that correspond to hydrological droughts (number of "hits"), $B$ the number of meteorological droughts that do not correspond to hydrological droughts (number of "false alarms"), $C$ the number of no-drought forecasts corresponding to hydrological droughts (number of "misses"), and $D$ the number of "correct non-events"). For a perfect detection, $B = C = 0$, so that $\mathrm{CSS} = 1$.

The CSS allows us to focus on particular drought events. Unfortunately, the relatively short river discharge time series is a strong limitation to our study that is based upon the 20th percentile of the distribution rather than on extreme events. For the observed annual mean time series, correlation and CSS have been calculated on detrended data with 2° spline function over a 49 yr period (1951–1999) with available river discharge data. For the sake of comparison, similar scores have been computed over 49 yr sliding windows for each 1850–2100 CNRM-CM5 climate simulation (the 20th percentile being estimated over the same 1951–1999

period as in the observations). In addition, scores of simulated and detrended indices with 4° spline function have been also estimated over the whole 251 yr integrations, using the 20th but also lower (10th and 5th) percentiles.

## 3 Results

### 3.1 Evaluation of meteorological drought indices against hydrological benchmark index at interannual timescale

Besides observed and ISBA-simulated variations of annual mean discharge at Obidos (Amazon) and Vicksburg (Mississippi), Fig. 1 shows the detrended time series for the various meteorological indices, as well as the ISBA-derived SRI12 for further comparison over years without discharge observations (over the Amazon Basin). Both correlations and CSS are slightly higher over the Amazon than over the Mississippi. Such a difference could be partly related to the different seasonality of precipitation and the possible contribution of early winter snowfall to the following year annual
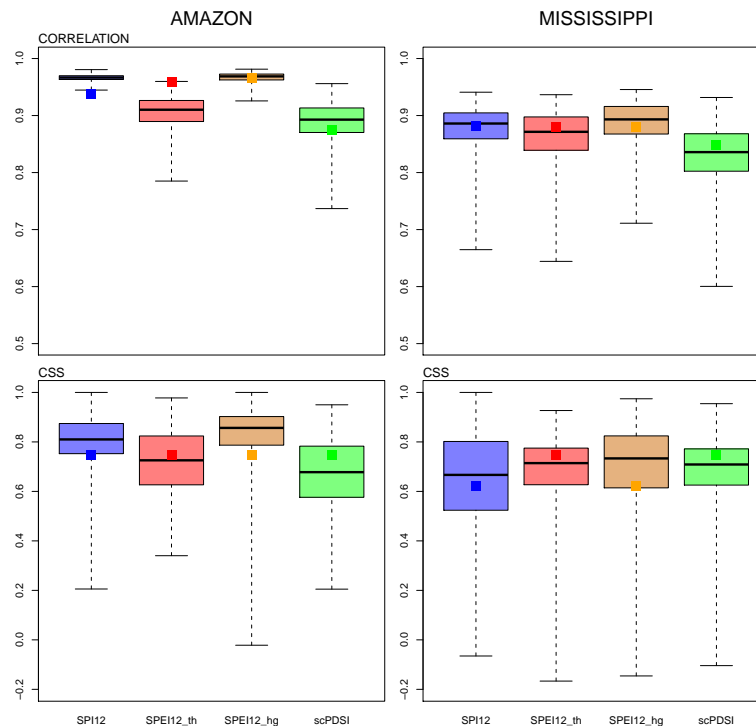
**Fig. 2.** Box-and-whisker plot of the sliding correlations and the CSS_20 calculated for the five members of the RCP8.5 scenario from 1850 to 2100 over a 49 yr time span for the Amazon and Mississippi watersheds. The boxes represent the 25th and the 75th percentile, the line the mean value, and the whiskers the minimum and the maximum values of the ensemble spread. The smaller squares indicate results obtained from observations (1951–1999).

mean runoff in the Mississippi Basin. Over the Amazon, the SPEI12_hg shows the best correlation with the SRI12 benchmark, closely followed by the SPEI12_th and SPI12. However, such differences are not significant and CSSs are the same for all three indices. Over the Mississippi, scores are also very close and longer time series would be useful to reach more robust conclusions about the relative skill of the different meteorological indices.

For this purpose, correlations and CSS have also been estimated over 49 yr sliding windows from our five-member ensemble of 1850 to 2100 climate simulations, with model-derived SRI12 taken as a reference. As explained in Sect. 2.2, all time series have been here detrended with 2-degree spline functions before computing correlation and CSS. Results are summarized in box-and-whisker plots (Fig. 2). In line with observations, all model-derived meteorological indices are relatively skillful over both river basins. Ranking them is particularly difficult over the Mississippi where differences in mean scores are very small. Results are more contrasted over the Amazon where SPI and SPEI_hg outperform other indices. This suggests that the details of the index computation (SPEI_hg versus SPEI_th) are as important as the choice of the index (SPEI vs. SPI or PDSI). The apparent superiority of SPEI_hg vs. SPEI_th (obvious over the Amazon, less clear over the Mississippi) did not show up in the observations. This might be due to the intrinsic uncertainty of scores based on 49 yr time series only, as well as to possible biases of the CNRM-CM5 model (for instance a dry bias over the Amazon; Joetzjer et al., 2013), which might increase the relative contribution of PET (vs. precipitation) in the SPEI calculation.

How sensitive are our CSSs to the quantile chosen as a threshold for drought definition? Considering now moderate ($q20$), severe ($q10$) and extreme ($q5$) droughts over the whole 1850–2100 period (Table 5), the simple SPI is the best proxy of 12-month hydrological droughts, closely followed by the SPEI_hg. Indeed, SPEI scores improve when PET is calculated with the Hargreaves in place of the Thornthwaite equation. Note that the scPDSI and the SPEI_th that estimate both PET through Thornthwaite show very similar skill.

In summary, precipitation remains the main driver of runoff at the interannual timescale, and accounting for PET (for SPEI) or even a simplified water balance (for sc-PDSI) does not improve the detection of 12-month hydrological droughts. Taking into account PET allows the SPEI to reach the same skill as the SPI when using the Hargreaves formula. As shown in Table 5, such a conclusion is not specific to the Amazon and Mississippi river basins, but also holds when averaging scores over all land grid points in the CNRM-CM5 model.

**Table 5.** Correlation and CSS between various meteorological drought indices and the reference standardized runoff index (SRI12) in the CNRM-CM5 model. Scores were calculated for average and detrended indices over the Amazon and Mississippi watersheds, as well as averaged over the globe on the basis of grid-cell rather than basin-scale indices (long: $-180°$ E, $+180°$ W; lat: $-60°$ S, $+60°$ N). The CSS was calculated using the 5th, 10th and 20th percentiles of the annual drought index distribution. Mean and standard deviation (sd) based on our five-member ensemble of 1850–2100 simulations are shown. Highest (bold) and lowest (italics) mean values are also shown.

| | | Amazon | | Mississippi | | Global | |
|---|---|---|---|---|---|---|---|
| | | mean | sd | mean | sd | mean | sd |
| | | **Correlation** | | | | | |
| | SPI12 | **0.97** | 0.001 | **0.88** | 0.023 | **0.89** | 0.014 |
| | SPEI12_th | 0.89 | 0.017 | 0.86 | 0.027 | 0.76 | 0.064 |
| | SPEI12_hg | 0.96 | 0.006 | 0.88 | 0.023 | 0.84 | 0.036 |
| | scPDSI | *0.88* | 0.004 | *0.84* | 0.030 | *0.75* | 0.033 |
| | | **CSS** | | | | | |
| | SPI12 | **0.84** | 0.071 | 0.68 | 0.082 | **0.69** | 0.052 |
| *q*20 | SPEI12_th | 0.70 | 0.050 | 0.64 | 0.129 | 0.56 | 0.081 |
| | SPEI12_hg | 0.82 | 0.059 | **0.69** | 0.124 | 0.64 | 0.065 |
| | scPDSI | *0.68* | 0.072 | *0.63* | 0.070 | *0.53* | 0.068 |
| | SPI12 | **0.79** | 0.053 | **0.61** | 0.101 | **0.65** | 0.073 |
| *q*10 | SPEI12_th | 0.65 | 0.077 | *0.55* | 0.084 | 0.52 | 0.095 |
| | SPEI12_hg | 0.77 | 0.038 | 0.59 | 0.065 | 0.59 | 0.084 |
| | scPDSI | *0.64* | 0.049 | 0.61 | 0.047 | *0.49* | 0.088 |
| | SPI12 | **0.77** | 0.089 | **0.56** | 0.092 | **0.59** | 0.105 |
| *q*5 | SPEI12_th | *0.66* | 0.068 | *0.53* | 0.068 | 0.47 | 0.118 |
| | SPEI12_hg | 0.72 | 0.092 | *0.53* | 0.068 | 0.54 | 0.112 |
| | scPDSI | *0.66* | 0.106 | *0.53* | 0.120 | *0.44* | 0.121 |

## 3.2 Climate change timescale

Moving to the raw model outputs, Fig. 3 shows the projection of the areal fraction of the Amazon and Mississippi basins in moderate, severe and extreme drought conditions (respectively defined under the 20th, 10th and 5th percentile estimated over the whole 1850–2100 period). Results obtained with the SRI12 benchmark are compared to the fractions derived from each meteorological index, as well as with the SPAEI to highlight the influence of the PET approximation on the simulated trends. Bold lines represent the ensemble mean value for each percentile. The envelope is defined by the minimum and maximum values among the five members for severe drought only (10th percentile), as an indication of the internal variability of the CNRM-CM5 climate model.

For SRI12, CNRM-CM5 under the RCP8.5 concentration scenario does not show any trend in the areal fraction of the Amazon Basin touched by hydrological drought, while a clear increase is projected over the Mississippi Basin. This response does not agree with the contrasted long-term variations derived from the various meteorological drought indices. The SPI12 behaves as a better proxy of SRI12 than scPDSI and SPEI12 over the Amazon Basin where precipitation change seems to dominate the long-term evolution of hydrological droughts and surface warming remains of marginal control. Conversely, the SPI12 evolution is in contradiction with the SRI12 evolution over the Mississippi Basin, where increased evapotranspiration seems to exceed increased precipitation and leads to more frequent and/or extended hydrological droughts at the end of the 21st century. This result highlights the SPI limitations, where and when temperature trends become strong enough to alter evapotranspiration without or despite changes in precipitation. Nevertheless, accounting for changes in PET does not necessarily solve the problem, as emphasized by Fig. 4. Indeed, the SPEI response to global warming is strongly dependent on the PET calculation. The strong sensitivity shown by SPEI12_th over both basins suggests that Thornthwaite's formula is not adequate for climate change studies and should be at least superseded by more robust approaches (e.g. Hargreaves or Penman–Monteith). The sensitivity of the PDSI to the PET calculation is controversial. For the 20th century Van der Schrier et al. (2011) showed weak sensitivity while Sheffield et al. (2012, Supplement) attribute this apparent weak sensitivity to inconsistencies in the forcing data sets and simulation configuration. Over the 21st century, and in line with Sheffield's results, it is likely that the large increase of the areal fraction in drought obtained with this index is also due to the simplistic PET calculation in the original algorithm.
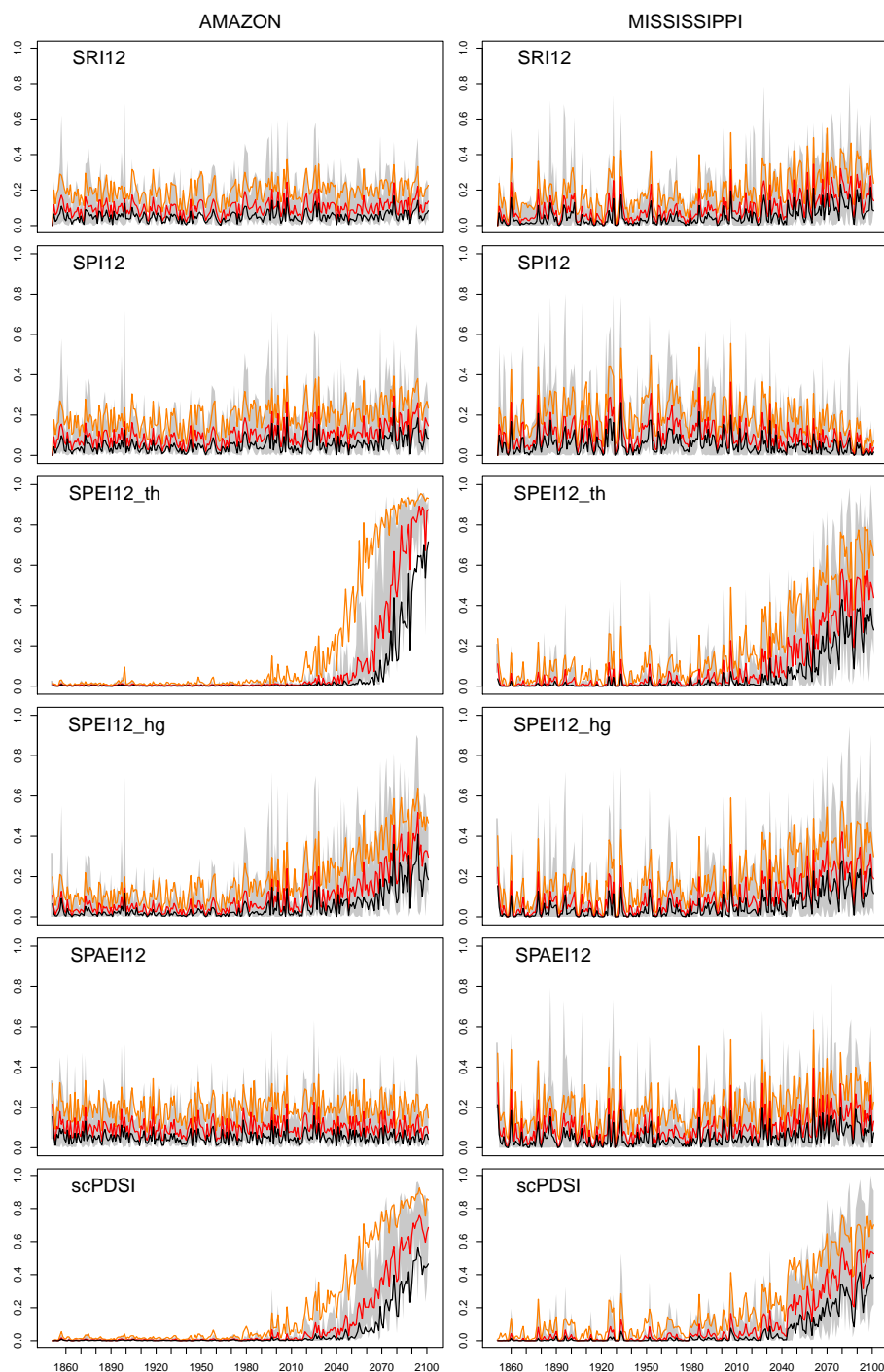
**Fig. 3.** Time series from 1850 to 2100 of the ensemble mean value of the areal fraction in drought condition in the Amazon and Mississippi basins. Moderate, severe and extreme droughts are defined locally as below the 20th (orange), the 10th (red) and the 5th (black) percentile. The envelop around the red line is defined by the minimum and maximum values among the five members.

Not surprisingly, the SPAEI12, accounting for actual rather than potential ET, shows more consistency with the "target" SRI12 than the other indices over both river basins. This confirms the limitation of the empirical meteorological indices for hydrological applications.

## 4   Discussion and conclusion

The present study aimed at comparing globally available empirical meteorological drought indices on one tropical (Amazon) and one mid-latitude (Mississippi) river basin, first in
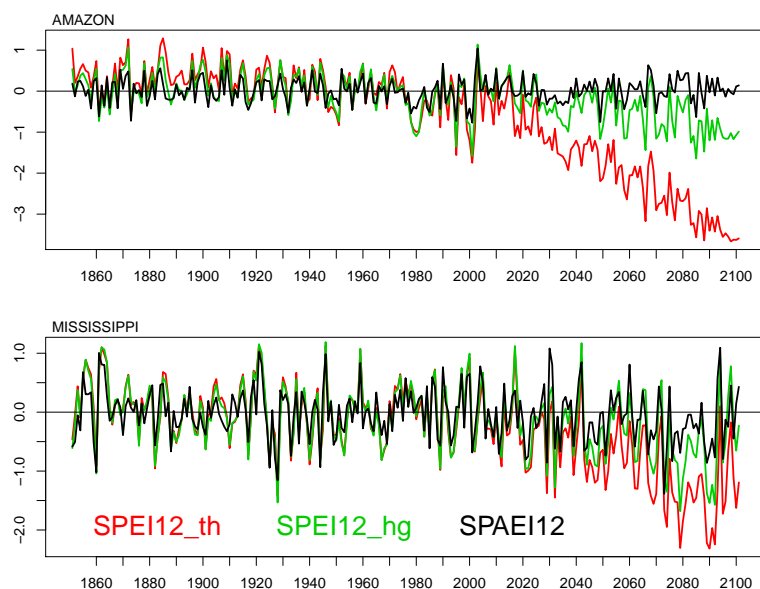
**Fig. 4.** Raw SPEI12 time series averaged over the Amazon (upper panel) and the Mississippi (lower panel) watersheds for one member of the CNRM-CM5 1850–2100 simulations.

their skill to detect interannual variations, then in their response to anthropogenic climate change. The focus is only on 12-month droughts, and a standardized runoff index (SRI), closely related to the river discharge, is used as a hydrologic benchmark.

At interannual timescales and over both basins, the simple SPI, based solely on precipitation, is no less suitable than more sophisticated empirical indices also using temperature inputs. This is true not only for observations, but also in the CNRM-CM5 climate simulations. When using the Hargreaves formula, the SPEI scores are however very close to the SPI scores. In contrast, the Thornthwaite formula systematically leads to lower scores. Such conclusions should be however tempered. First, there might be some regional heterogeneities in the ranking of the four indices given the weak spread between all indices, not only over the selected basins, but also when averaging the scores obtained over all land grid cells between $60°$ S and $60°$ N (cf. Table 5). Moreover, similar scores calculated on shorter timescale (3 and 6 months respectively, not shown) indices suggest a slight superiority of the SPEI_hg compared to the simple SPI.

Beyond the ability of the various meteorological indices to account for the interannual variability of annual streamflow, and in line with the conclusions of Burke and Brown (2008), Kingston et al. (2009) and Burke (2011), our study also emphasizes that drought projections are strongly index-dependent given the differing impact of temperature in their calculation. While the SPEI was recently proposed as a drought index sensitive to global warming (Vicente Serrano et al., 2010), it shows a stronger drying of the Amazon and Mississippi basins than indicated by our hydrologic benchmark. This discrepancy is less pronounced when estimating

PET with Hargreaves, especially for the Mississippi, showing that precipitation is not the only driver of the long-term drought variations. Such inconsistencies can lead to differences at the end of the 21st century, and are also discernible from the end of the 20th century as demonstrated by Sheffield et al. (2012) for the PDSI.

A caveat of the present study is the fact that we have neglected potential vegetation feedbacks in our climate projections. Under a higher atmospheric $CO_2$ concentration, the stomatal closure for instance might alter the relationship between meteorological and hydrological droughts as the stomatal closure partly regulates water exchange within the soil–plant–atmosphere continuum. The ISBA land surface model implemented in CNRM-CM5 model still uses a common Jarvis-type formulation (Jarvis, 1976) for the computation of the stomatal conductance and does not account for a direct $CO_2$ effect on plant transpiration. This effect is also neglected by the meteorological drought indices. This remark highlights again the fundamental limitations of such empirical indices, which can be relevant for present-day climate but less suitable for long-term projections. This caveat however does not change our main conclusion: besides the choice of a concentration scenario (here RCP8.5, i.e. the most severe scenario considered in CMIP5) and of a global climate model (here CNRM-CM5), the index definition and the associated PET calculation also represent a major source of uncertainty for drought projections (Taylor et al., 2013). Note that only one concentration scenario and one global climate model have been considered in this study, but that preliminary analyses of the different scenarios obtained with CNRM-CM5 as well as of the RCP8.5 scenario obtained with a subset of CMIP5 models suggest that CNRM-CM5 is not an outlier

among the CMIP5 models and that the index definition is as important as the choice of the scenario/model as a source of uncertainty for drought projections (details are given in the Supplement).

Finally, another limitation of the present study is the arbitrary choice of the SRI benchmark. Besides runoff and river discharge, other impact-oriented benchmarks could have been proposed such as soil moisture (e.g. the SMA – soil moisture anomaly; Orlowsky and Seneviratne, 2013) or photosynthesis activity, which can be derived from satellite observations. Nevertheless, such observations only cover a few decades (only since the early 1980s) and are sometimes still difficult to interpret given the limitations of remote sensing techniques (e.g. Anderson et al., 2011).

Therefore, the main alternative for drought monitoring and projections is probably the use of process-oriented land surface models, which can be either driven by observed atmospheric forcings (e.g. Sheffield and Wood, 2007) and bias-corrected climate scenarios or directly coupled to global climate models (e.g. Sheffield and Wood, 2008). Given the intrinsic uncertainties related to the various physical and biological processes represented in such land surface models (e.g. Betts et al., 2007), a multi-model approach is however strongly encouraged.

## References

Alkama, R., Decharme, B., Douville, H., and Ribes, A.: Trends in global and basin-scale runoff over the late 20th century: Methodological issues and sources of uncertainty, J. Climate, 24, 2983–2999, doi:10.1175/2010JCLI3921.1, 2011.

Anderson, M. C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J. R., and Kustas, W.: Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States, J. Climate, 24, 2025–2044, 2011.

Betts, R. A., Boucher, O., Collins, M., Cox, P. M., Falloon, P., Gedney, N., Hemming, D. L., Huntingford, C., Jones, C. D., Sexton, D., and Webb, M.: Projected increase in continental runoff due to plant responses to increasing carbon dioxide, Nature, 448, 1037–1041, doi:10.1038/nature06045, 2007.

Burke, E. J.: Understanding the Sensitivity of Different Drought Metrics to the Drivers of Drought under Increased Atmospheric $CO_2$, J. Hydrometeorol., 12, 1318–1393, 2011.

Burke, E. J. and Brown, S. J.: Evaluating uncertainties in the projection of future drought, J. Hydrometeorol., 9, 292–299, 2008.

Dai, A., Trenberth, K. E., and Qian, T.: 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, doi:10.1175/JHM-386.1, 2004.

Dai, A.: Drought under global warming: A review, Climatic Change, 2, 45–65, 2011a.

Dai, A.: Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008, J. Geophys. Res., 116, D12115, doi:10.1029/2010JD015541, 2011b.

Decharme, B. and Douville, H.: Global validation of the ISBA Sub-Grid Hydrology, Clim. Dynam., 29, 21–37, doi:10.1007/s00382-006-0216-7, 2007.

Guttman, N. B.: Comparing the Palmer Drought Index and the Standardized Precipitation Index, J. Am. Water Resour. Assoc., 34, 113–121, 1998.

Hargreaves, G. H. and Samani, Z. A.: Estimating potential evapotranspiration, J. Irrig. Drain. Eng. ASCE, 188, 223–230, 1982.

Heim Jr., R. R.: A review of twentieth-century drought indices used in the United States, B. Am. Meteorol. Soc., 83, 1149–1165, 2002.

Jarvis, P. G.: The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field, Philos. T. Roy. Soc. Lond. B, 273, 593–610, doi:10.1098/rstb.1976.0035, 1976

Joetzjer, E., Douville, H., Delire, C., and Ciais, P.: Present-day and future Amazonian precipitation in global climate models: CMIP5 versus CMIP3, Clim. Dynam., 41, 2921–2936, doi:10.1007/s00382-012-1644-1, 2013.

Kingston, D. G., Tood, M. G., Taylor, R. G., Thompson, J. R., and Arnell, N. W.: Uncertainty in the estimation of potential evapotranspiration under climate change, Geophys. Res. Lett., 36, 3–8, doi:10.1029/2009GL040267, 2009.

McKee, T. B., Doesken, N. J., and Kleist, J.: Drought monitoring with multiple time scales, Proceedings of the 9th Conference on Applied Climatology, AMS, Boston, MA, 233–236, 1995.

Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, doi:10.1002/joc.1181, 2005.

Orlowsky, B. and Seneviratne, S. I.: Global changes in extreme events: Regional and seasonal dimension, Climatic Change, 110, 669–696, doi:10.1007/s10584-011-0122-9, 2012.

Orlowsky, B. and Seneviratne, S. I.: Elusive drought: uncertainty in observed trends and short- and long-term CMIP5 projections, Hydrol. Earth Syst. Sci., 17, 1765–1781, doi:10.5194/hess-17-1765-2013, 2013.

Palmer, W. C.: Meteorological drought. Research Paper No. 45, US Dept. of Commerce, available at: http://www.ncdc.noaa.gov/oa/climate/research/drought/palmer.pdf (last access: 6 December 2013), 1965.

Ribes, A., Azais, J. M., and Planton, S.: A method for regional climate change detection using smooth temporal patterns, Clim. Dynam., 35, 391–306, 2010.

Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A., and Ziese, M.: New GPCC full data reanalysis version 5 provides high-quality gridded monthly precipitation data, GEWEX News, 21, 4–5, 2011.

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in climate extremes and their impacts on the natural physical environment, in: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, edited by: Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G.-K., Allen, S. K., Tignor, M., and Midgley, P. M.: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC), Cambridge University Press, Cambridge, UK, and New York, NY, USA, 109–230, 2012.

Sheffield, J. and Wood, E. F.: Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle, J. Geophys. Res., 112, D17115, doi:10.1029/2006JD008288, 2007.

Sheffield, J. and Wood, E. F.: Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations, Clim. Dynam., 31, 79–105, 2008.

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

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

Taylor, I. H., Burke, E., McColl, L., Falloon, P. D., Harris, G. R., and McNeall, D.: The impact of climate mitigation on projections of future drought, Hydrol. Earth Syst. Sci., 17, 2339–2358, doi:10.5194/hess-17-2339-2013, 2013.

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

Trenberth, K. E.: Changes in precipitation with climate change, Climate Res., 47, 123–138, 2011.

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

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, J. Climate, 23, 1696–1718, 2009.

Vicente-Serrano, S. M., Beguería S., Lopez-Moreno J. I., and Angulo, M.: A New Global 0.5. Gridded Dataset (1901–2006) of a Multiscalar Drought Index: Comparison with Current Drought Index Datasets Based on the Palmer Drought Severity Index, J. Hydrometeorol., 11, 1033–1043, 2010.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: Comment on "Characteristics and trends in various forms of the Palmer Drought Severity Index (PDSI) during 1900–2008" by Aiguo Dai, J. Geophys. Res., 116, 1–9, 2011.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: description and basic evaluation, Clim. Dynam., 40, 2091–2121, 2013.

Wahba, G.: Spline models for observational data, Society for Industrial and Applied Mathematics (SIAM), CBMS-NSF Regional Conference Series in Applied Mathematics, doi:10.1137/1.9781611970128, 1990.

Wells, N., Goddard, S., and Hayes, M.: A self-calibrating Palmer Drought Severity Index, J. Climate, 17, 2335–2351, 2004.

Wilks, S.: Statistical Methods in the atmospheric sciences second edition, International Geophysics Series, Academic Press, 624 pp., 2004.