The effect of climate on timescales of drought propagation in an ensemble of global hydrological models

Anouk I. Gevaert[1], Ted I. E. Veldkamp[2], Philip J. Ward[2]

[1] Faculty of Science, Vrije Universiteit Amsterdam, the Netherlands
5 [2] Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, the Netherlands

Correspondence to: Ted I. E. Veldkamp (ted.veldkamp@vu.nl)

Abstract. Drought is a natural hazard that occurs at many temporal and spatial scales and has severe environmental and socio-economic impacts across the globe. The impacts of drought change as drought evolves from precipitation deficits to deficits in soil moisture or streamflow. Here, we quantified the time taken for drought to propagate from meteorological

10 drought to soil moisture drought, and from meteorological drought to hydrological drought. We did this by cross-correlating the Standardized Precipitation Index (SPI) against standardized indices of soil moisture, runoff, and streamflow from an ensemble of global hydrological models forced by a consistent meteorological dataset. Drought propagation is strongly related to climate, occurring at sub-seasonal timescales in tropical climates and at up to multi-annual timescales in continental and arid climates. Winter droughts are usually related to longer SPI accumulation periods than summer droughts,

15 especially in continental and tropical savanna climates. The difference between the seasons is likely due to winter snow cover in the former and distinct wet and dry seasons in the latter. Model structure appears to play an important role in model variability, as drought propagation to soil moisture drought is slower in land surface models than in global hydrological models, but propagation to hydrological drought is faster in land surface models than in global hydrological models. The propagation time from SPI to hydrological drought in the models was evaluated against observed data at 297 in-situ

20 streamflow stations. On average, errors between observed and modeled drought propagation timescales are small and the model ensemble mean is preferred over the use of a single model. Nevertheless, there is ample opportunity for improvement as substantial differences in drought propagation are found at 20 % of the study sites. A better understanding and representation of drought propagation in models may help improve seasonal drought forecasting as well as constrain drought variability under future climate scenarios.

25 # 1 Introduction

Drought is a complex global phenomenon with severe environmental (Bond et al., 2005; Lewis and Sjöstrom, 2010; Reichstein et al., 2013; Turco et al., 2017; Vicente-Serrano et al., 2013) and socio-economic (Horridge et al., 2005; Stanke et al., 2013; Wegren, 2011) impacts. Data from global and regional models are commonly used to study droughts (e.g. Andreadis and Lettenmaier, 2006; Sheffield et al., 2004; Sheffield and Wood, 2008a) and water scarcity (Kummu et al.,

2016; Veldkamp et al., 2015; Wada et al., 2011), and how these will change in the future (Dai, 2013; Huang et al., 2017; Sheffield and Wood, 2008b; Trenberth et al., 2014). However, the evolution of drought from meteorological anomalies to deficits in soil moisture and streamflow in these models is still poorly understood (Van Lanen et al., 2013; Van Loon, 2015), despite the fact that drought impacts are more closely related to these latter components of the hydrological cycle (Van Loon,

5   2015).

Several types of drought can be distinguished (i.e. Van Loon, 2015; Mishra and Singh, 2010). The first stage or type of drought is called meteorological drought, and is caused by precipitation deficits that may or may not be combined with above-normal potential evaporation rates. If the precipitation deficits and/or increased evaporation rates are sustained for a sufficient period, they can result in lower than average soil moisture availability, which may result in a soil moisture drought.

10   In the same way, meteorological drought can propagate into lower streamflow and groundwater levels, which are both forms of hydrological drought. Another type of drought, socio-economic drought, is not based on a hydrological variable alone, but occurs when water availability is lower than water demand. As drought propagates from meteorological drought to soil moisture or hydrological drought, the characteristics of droughts change. As drought moves through components of the hydrological cycle, the onset of droughts tends to be lagged, there tend to be fewer but longer drought events, and the

15   droughts are attenuated (Van Lanen et al., 2013; Van Loon et al., 2012; Peters et al., 2003). The degree to which these drought propagation characteristics are observed depends on climate and catchment properties (Van Lanen et al., 2013; Van Loon and Laaha, 2015).

Several drought propagation studies have been carried out based on observational data at catchment to national scales, mainly focusing on meteorological and streamflow droughts (e.g. Barker et al., 2016; Haslinger et al., 2014; Van Loon and

20   Laaha, 2015; Lorenzo-Lacruz et al., 2013). However, the geographical extent covered by these studies is limited, mainly due to the lack of observational data with sufficiently long timescales needed to identify drought in many regions. Therefore, data from regional or global models are generally used to study drought at larger spatial scales. In some cases, studies have used data from a single large-scale model (e.g. Van Lanen et al., 2013; Lehner et al., 2006; Lloyd-Hughes et al., 2013; Sheffield et al., 2004). However, there can be considerable differences between hydrological outputs of different large-scale

25   models (Burke and Brown, 2008; Gudmundsson et al., 2012a, 2012b; Haddeland et al., 2011; Prudhomme et al., 2014), suggesting that multi-model approaches (e.g. Gudmundsson et al., 2012a; Stahl et al., 2012; Tallaksen and Stahl, 2014; Wang et al., 2009) are more appropriate. This is supported by the fact that studies have found that a model ensemble performs better than individual models (Gudmundsson et al., 2012a; Stahl et al., 2012). However, most regional or global drought studies have focused on drought frequency and severity (van Huijgevoort et al., 2013; Prudhomme et al., 2011,

30   2014) rather than on drought propagation, and those that have studied drought propagation have been limited to a single model (Van Lanen et al., 2013) or to a small selection of contrasting catchments (Van Loon et al., 2012).

The aim of this study is therefore to investigate drought propagation times in an ensemble of global hydrological models from meteorological drought to soil moisture drought, and from meteorological drought to hydrological drought. We focus on the effect of climate on drought propagation in particular, and distinguish between summer and winter droughts. In

Sect. 2, we describe the identification of drought, the method we use to quantify drought propagation timescales, and the validation analysis. The global models and the observational data are presented in Sect. 3. In Sect. 4, we describe and discuss the timescales of drought propagation and their relationship with climate in the multi-model ensemble mean. In addition, we briefly describe the model variability and perform a validation exercise. We wrap up with a brief summary and conclusion in

5   Sect. 5.

2 Methods

The timescales of drought propagation were determined by relating standardized indices of meteorological drought to indices of soil moisture and hydrological drought. These indices were derived from an ensemble of seven global models that were forced with a consistent meteorological dataset in the eartH2Observe project (Schellekens et al., 2016). For more details on

10   the model datasets, see Sect. 3. First, standardized indices of precipitation, soil moisture, runoff, and streamflow were calculated as a measure of drought. Second, we determined which timescales of meteorological drought were most strongly correlated to the other drought types. Finally, the model results were evaluated against observational data.

2.1 Drought indices

Meteorological drought was quantified by the widely used Standardized Precipitation Index (SPI; Mckee et al. 1993). The

15   SPI fits a pre-defined probability distribution to the frequency distribution of precipitation. Generally, the precipitation series have a monthly resolution, in which case the fitting is done for each month separately. Examples of possible distribution functions are the gamma, Pearson Type III, or log-normal (Guttman, 1999; Lloyd-Hughes and Saunders, 2002; Mckee et al., 1993) functions, though a non-parametric approach has also been developed (Hao and AghaKouchak, 2014). In this study, we use the gamma distribution as the pre-defined probability distribution, as this is commonly considered to be the most

20   suitable (Lloyd-Hughes and Saunders, 2002; Mckee et al., 1993; Stagge et al., 2015). Regions in which months with zero precipitation are common are problematic for the SPI (Wu et al., 2007), so the cumulative distribution function is corrected with the probability of zero precipitation (Naresh Kumar et al., 2009). The fitted probability distribution is then transformed to the normal distribution, resulting in negative (positive) values for dry (wet) conditions. Note that the SPI relates the precipitation amount in a certain month to average conditions at that particular location. In drier climates or seasons, water-

25   limited conditions may be the norm, while in wetter climates or seasons, water stress may not be relevant until severe drought thresholds are reached.

Advantages and disadvantages of the SPI are described in Hayes et al. (1999). A first advantage is that the SPI is relatively simple, requiring only precipitation data as an input. Secondly, the index is flexible. It can be used to compare different timescales of drought by aggregating the input precipitation time series over a number of months, known as the

30   accumulation period (typically 1, 3, 6, 12, or 24 months). Thirdly, the standardization facilitates comparisons of extreme conditions in different locations, but also over different timescales. A disadvantage is that the precipitation data used to

calculate the SPI may not be representative of surface conditions. In addition, the calculation of the index requires long time series (>30 years) of data (Guttman, 1999; Mckee et al., 1993; Zargar et al., 2011).

The standardization approach of the SPI is not only easily applied to different timescales, but can also be applied to other (hydrological) variables such as soil moisture or streamflow. In this way, we use the same methodology to identify soil

5 moisture and hydrological drought by defining the Standardized Soil Moisture Index (SSMI) (Hao and AghaKouchak, 2013), Standardized Runoff Index (SRI) (Shukla and Wood, 2008), and Standardized Streamflow Index (SSFI).

Soil moisture and runoff are controlled by pixel-scale precipitation and hydrological processes. Streamflow, on the other hand, is affected by upstream pixels. Therefore, we computed a catchment-aggregated SPI to quantify drought propagation from meteorological to streamflow droughts. The input for the aggregated SPI is the total precipitation falling within each

10 pixel and its upstream area, based on a 0.5° routing network (see Sect. 3). We did not include a travel time factor in this calculation, thus assuming that precipitation falling in the upper parts of each catchment will impact streamflow at the outlet within the same month. In the rest of the paper, meteorological drought is based on the pixel-based SPI when referring to soil moisture or runoff drought, and to the catchment-based SPI when referring to streamflow drought.

In this study, we quantified drought propagation from meteorological drought to soil moisture and hydrological droughts.

15 Therefore, we calculated the SPI using accumulation periods of 1, 2, 3, 6, 9, 12, 24 and 36 months. These accumulation periods span sub-seasonal, (multi-)seasonal, and (multi-)annual timescales. For the other standardized indices (SIs), we used the 1-month accumulation period to identify short-term drought conditions.

2.2 Timescales of drought propagation

Drought propagation from meteorological to soil moisture or hydrological drought was based on correlations between the

20 SPI and target SI (SSMI, SRI or SSFI). The SI time series were prepared by applying two criteria to the target SI. First, we distinguished between drought conditions in summer and winter seasons. The summer (winter) season was defined as June, July, and August above (below) the equator, and December, January, and February below (above) the equator. Second, we focused on months corresponding to dry conditions, here defined as $SI \leq 0$. This threshold includes near-normal and mildly dry conditions, but leaves a larger sample size than when we limit the analysis to moderate or severe drought events.

25 Theoretically, moderate drought events ($SI \leq -1$) have an occurrence probability of about 16 % per year, which would correspond to about 5 events in a 30-year study period, for a total of 15 events considering a three-month season. Severe drought events ($SI \leq -1.5$) would result in six events during a 30-year period of three months based on the same reasoning, compared to 45 events using $SI \leq 0$ as a threshold.

After preparing the SI time series, we cross-correlated each of the SPI time series with the SSMI, SRI, and SSFI, without

30 considering lag between the time series. The drought propagation timescale was defined as the SPI accumulation period that is most closely related to the target drought index, which we call SPI-n. Pixels where the final correlations are not statistically significant (p = 0.05) were masked from the results. Autocorrelation is a potential issue when correlating time

series, as it reduces the degrees of freedom compared to a standard significance test. In this study, the effective degrees of freedom are based on the modified Chelton method (Pyper and Peterman, 1998) as in Barker et al. (2016).

The strength of the relationships between the Köppen-Geiger classification (Kottek et al., 2006) shown in Fig. 1, as well as certain model characteristics on SPI-n, were quantified using a variety of tests. We use the rank of SPI-n rather than the

5   duration of the accumulation period in months in the calculations. This means we assume that the difference between accumulation periods of 12 and 24 months (both in the order of annual timescales, differing by one rank SPI-n) to be equivalent to the difference between 1 and 2 months (both in the order of sub-seasonal timescales, differing by one rank SPI-n), but very unlike the difference between 1 and 12 months (sub-seasonal versus annual timescales, differing by 5 rank SPI-n). Statistical significance was based on (paired) t-tests and ANOVA tests. Since statistical significance does not reflect the

10   relevance of differences between groups, we use Cohen's d to quantify the effect size between two groups (Cohen, 1988). This metric is defined as:

$$d = \frac{\mu_1 - \mu_2}{\sigma_{pooled}} \tag{1}$$

where

$$\sigma_{pooled} = \sqrt{\frac{(n_1-1)\sigma_1^2 + (n_2-1)\sigma_2^2}{n_1+n_2-2}} \tag{2}$$

15   and where μ represents the group mean, σ the standard deviation, n the number of observations, and subscripts indicate the group. The outcome of the metric is thus the difference in group means relative to the standard deviation of the groups. A result of 1 therefore means that the group means differ by one standard deviation and those groups overlap by about 62 %. Previous studies tend to use fixed thresholds to interpret small (0.2), medium (0.5) and large (0.8) effect sizes, though these thresholds are considered to be rather arbitrary (Cohen, 1988; Lenth, 2001).
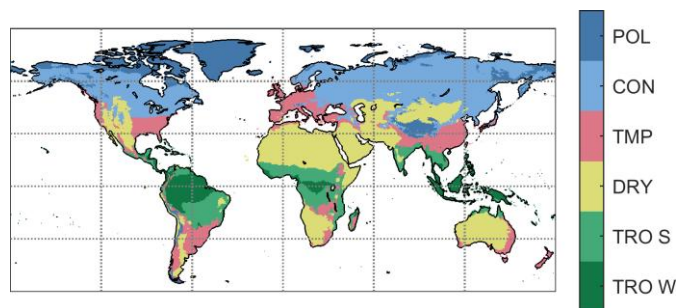
Figure 1: Map of the Köppen-Geiger climate classification used in this study. TRO W = tropical wet, TRO S = tropical savanna, DRY = dry, TMP = temperate, CON = continental, POL = polar.

2.3 Validation of drought propagation

Timescales of drought propagation from meteorological to streamflow drought in global hydrological models are validated against observational data. Sites with observational streamflow data were matched to model pixels by selecting the model pixel containing the in-situ site. Since there may be discrepancies between the model and actual river routing schemes, we

5   extracted the model streamflow data from the selected pixel as well as from the eight surrounding pixels. The in-situ site was then assigned to the model pixel with the lowest root mean square error (RMSE) between observed and modeled streamflow. Once the in-situ sites were matched to model pixels, the SPI-n were calculated as described in Sect. 2.2. For consistency, model results were also recalculated at the in-situ sites using a paired data approach.

The evaluation of SPI-n was based on the rank SPI-n rather than the length of the accumulation period in months, for the

10   same reason for which rank SPI-n were used in the statistical tests (Sect. 2.2). The performance metrics used in the evaluation were mean absolute error (MAE) and Spearman correlation coefficient.

3 Data

We assess drought propagation in seven global models from the eartH2Observe project (www.earth2observe.eu), as well as in the model ensemble mean. Three of the models are land surface models (LSMs): HTESSEL-CaMa (Balsamo et al., 2009;

15   Yamazaki et al., 2011), ORCHIDEE (D'Orgeval et al., 2008; Krinner et al., 2005; Ngo-Duc et al., 2007), and the SURFEX-TRIP modeling platform (Decharme et al., 2010, 2013). The other four models are Global Hydrological Models (GHMs): LISFLOOD (Van Der Knijff et al., 2010), PCR-GLOBWB (van Beek et al., 2011; Wada et al., 2014), W3RA (van Dijk, 2010; Van Dijk et al., 2014), and WaterGAP3 (Döll et al., 2009; Flörke et al., 2013). Other models in the eartH2Observe project were excluded because they did not provide all of the variables required for this study. The models are run with a

20   consistent 0.5° meteorological forcing dataset, the WATCH Forcing Data methodology applied to ERA-Interim reanalysis (WFDEI) data (Weedon et al., 2014). However, static fields such as land cover and soil physical properties were not prescribed, as these tend to be closely linked to the modeling system. Important characteristics of the models such as runoff mechanisms and representation of reservoirs or water use are presented in Table 1. For more information about the model datasets and project design, see Schellekens et al. (2016).

25   In this study, we used the precipitation, root-zone soil moisture, runoff, and streamflow datasets to calculate the SIs. We do not study groundwater droughts because HTESSEL-CaMa does not simulate this store, and two of the other models have not made the data available. In addition, groundwater is defined differently between models, complicating comparisons of this store. The common forcing dataset means that the SPI time series are identical for all models, while the SSMI, SRI, and SSFI are model-specific. The model ensemble mean was calculated as the average of the SIs. A consistent model dataset for

30   the eartH2Observe project is available from 1980–2012 , though we use the years 1983–2012 to avoid data gaps in the first years of the SI time series caused by the 36-month accumulation period for the SPI. In addition to the datasets used to

calculate the SIs, we used other model variables to relate these to the drought propagation results. These are the runoff coefficient, or the ratio of runoff to precipitation, and the ratio of surface runoff to total runoff.

Table 1: Overview of models and relevant characteristics

Model	Model type	Evaporation	Snow	Soil layers	Runoff	Reservoirs	Water use
HTESSEL-CaMa	LSM	Penman-Monteith	Energy balance	4	Saturation excess	No	No
LISFLOOD	GHM	Penman-Monteith	Degree-day	2	Saturation and infiltration excess	Yes	Yes
ORCHIDEE	LSM	Bulk method (Barella-Ortiz et al., 2013)	Energy balance	11	Green-Ampt infiltration	No	Irrigation only
PCR-GLOBWB	GHM	Hamon	Temperature based	2	Saturation excess	Yes (lakes only)	No
SURFEX-TRIP	LSM	Penman-Monteith	Energy and mass balance	14	Saturation and infiltration excess	No	No
W3RA	GHM	Penman-Monteith	Degree-day	3	Saturation and infiltration excess	No	No
WaterGaP3	GHM	Priestley-Taylor	Degree-day	1	Beta function	Yes	Yes

5 The validation of modeled drought propagation was based on observed precipitation and streamflow time series. We used gauge-based precipitation data from the Global Precipitation Climatology Centre's (GPCC) Full Data Reanalysis product version 7 (Schneider et al., 2015). The data are available as monthly precipitation totals at 0.5° resolution for the entire modeled period. Note that reanalysis precipitation data, such as used for the model forcing in this study, and the GPCC dataset are not truly independent. However, the dependence of reanalysis precipitation on both gauge and satellite
10 observations inhibits the selection of a completely independent dataset with global coverage that also has a record long enough for drought studies. Monthly streamflow data were obtained from the Global Runoff Data Centre (GRDC; Koblenz, Germany; http://grdc.bafg.de) database. We used only sites with a catchment area larger than 9000 km² and at least 15 years of data. In addition, we ensured that the sites were independent by searching for the most upstream stations that fit the previous requirements. All stations located downstream of these stations were excluded from the analysis. Finally, we only

report on sites where the correlation between SPI-n and the SSFI was statistically significant ($p < 0.05$). These criteria resulted in 297 sites, with an average data availability of 27 years.

4 Results and discussion

Here, we characterize and discuss the timescales of drought propagation from meteorological to soil moisture and hydrological drought at global scale. First, we describe drought propagation in the model ensemble mean and its relationship to climate. Second, we assess the variability between models and identify factors that may explain these differences. Finally, we compare the results of the model ensemble mean as well as the individual models to observational data.

4.1 Model ensemble mean

Drought propagation based on SPI-n for the model ensemble mean varies considerably, and all timescales from 1 to 36 months are represented in the results (Fig. 2). Summer soil moisture droughts (based on SSMI) are best represented by SPI-n of one or two months in wet regions such as the Amazon, but by much longer SPI-n in dry climates and some boreal regions. Results are more mixed when focusing on runoff droughts (based on SRI). Runoff droughts are most linked to precipitation deficits in the same month in dry climates such as the Sahel, southern Africa, and central Australia. Runoff droughts in other dry regions such as the Middle East, northern Africa, and the western USA, however, are related to much longer precipitation deficits up to several years. The patterns of drought propagation timescales from meteorological to streamflow drought are similar to the patterns from meteorological to runoff droughts, though SPI-n tends to be slightly longer. Longer SPI-n for streamflow compared to runoff can be expected as streamflow is simply routed runoff. In general, SPI-n are also longer, and drought propagation slower, for winter droughts than for summer droughts (Fig. 2). In this study, we focus on SPI-n rather than the strength of the relationship between SPI and other SI. However, the correlations behind SPI-n are generally high, with median values ranging from 0.67 to 0.74, depending on climate and season (Fig. S1). The strength of the relationship is highest in tropical climates (medians around 0.8) and lowest in polar climates (medians around 0.6).
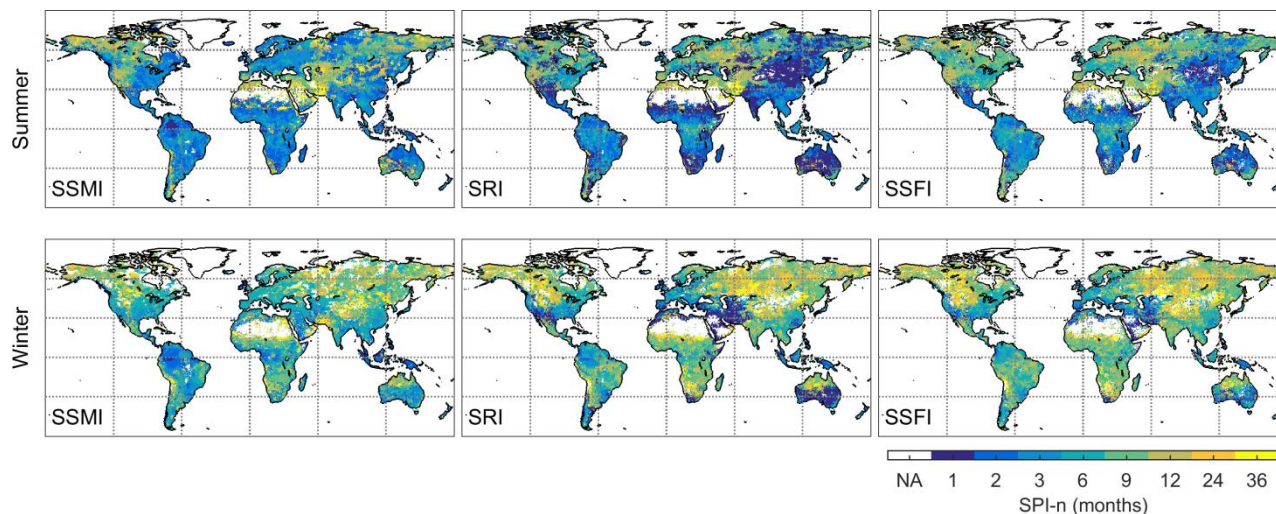
Figure 2: The SPI accumulation period (SPI-n) resulting in the highest correlations with model ensemble mean SSMI, SRI, and SSFI, for summer and winter droughts. Pixels where those correlations are not statistically significant (p < 0.05) are masked.

5   The relationship between drought propagation timescales and climate is further examined using the Köppen-Geiger classification (Kottek et al., 2006). We use six climate classes that reflect the five major climate types, with an additional distinction between tropical wet (i.e. tropical rainforest or monsoonal climates) and tropical savanna climates (Fig. 1). The results in Fig. 3 confirm that the climate type plays an important role in the timescale of drought propagation. As in Fig. 2, droughts in tropical climates tend to respond to short periods of precipitation deficits, while continental and polar climates

10   respond to longer periods of precipitation deficits. Overall, the variability within the tropical climate groups (both wet and savanna) is relatively low compared to dry climates, which are represented by the entire range of SPI accumulation periods studied. Despite the large variability in drought propagation timescales in dry climates, further distinctions between desert/savanna or hot/cold climates within the dry climate class did not have added value.
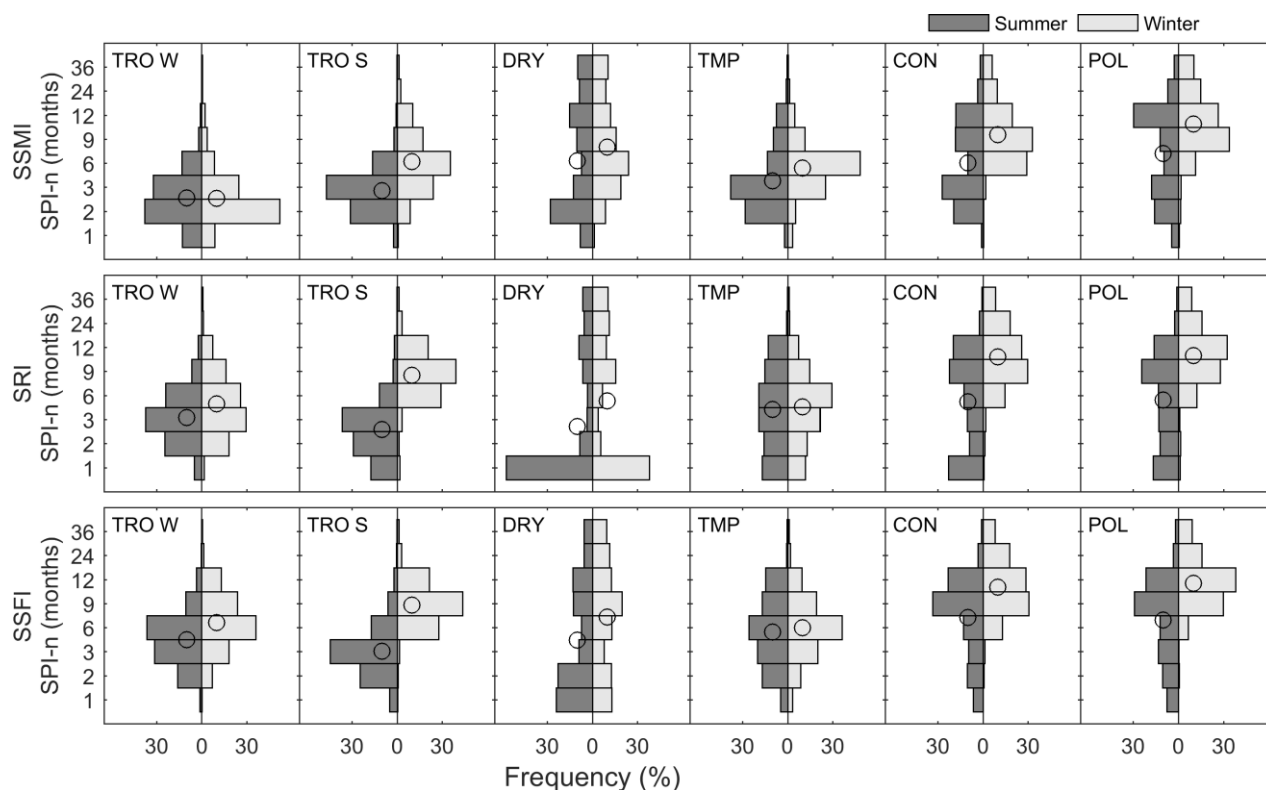
Figure 3: Histograms of the SPI-n in months by climate type, and for summer and winter droughts in SSMI, SRI, and SSFI. Circles represent the mean SPI-n per climate type and season. Abbreviations of climate types are the same as in Fig. 1.

5 Winter droughts are related to longer SPI-n than summer droughts in tropical savanna, continental and polar climates. Soil moisture droughts in tropical wet climates, and hydrological droughts in temperate climates, on the other hand, have similar propagation characteristics in summer and winter. The seasonality of SPI-n in tropical savanna climates may be related to the distinct wet (summer) and dry (winter) seasons in these regions: SPI-n are more similar to tropical wet climates in summer and to dry climates in winter. In continental climates, longer SPI-n may be due to snow cover in the winter. Precipitation in

10 the form of snow during fall and winter will not replenish soil moisture or runoff until temperatures rise sufficiently to melt the snow. In this way, drought conditions in winter may be more related to precipitation deficits in the previous summer. Indeed, soil moisture droughts may be more related to the SPI accumulated over the three-month period before snowfall than to the longer period starting with those three months and extending until the defined winter season. Note that we used a rather simple definition of summer and winter seasons by using the equator. However, since the climate along the equator is

15 almost exclusively tropical wet, which does not show significant seasonality, we expect that this does not significantly impact the results.

In terms of statistical significance, ANOVA tests show that the mean SPI-n for each climate type are significantly different ($p < 0.05$) for both soil moisture and hydrological droughts, with just one exception. The means of SPI-n for winter

hydrological droughts in continental and polar climates are not significantly different. Similarly, the difference in propagation time between summer and winter droughts for all drought types are significantly different based on paired t-tests, except for soil moisture droughts in tropical wet climates. Despite the fact that most group means differ significantly, they are not always substantially different in magnitude. For example, the difference between mean summer and winter SPI-n for runoff droughts in temperate climates is very small.
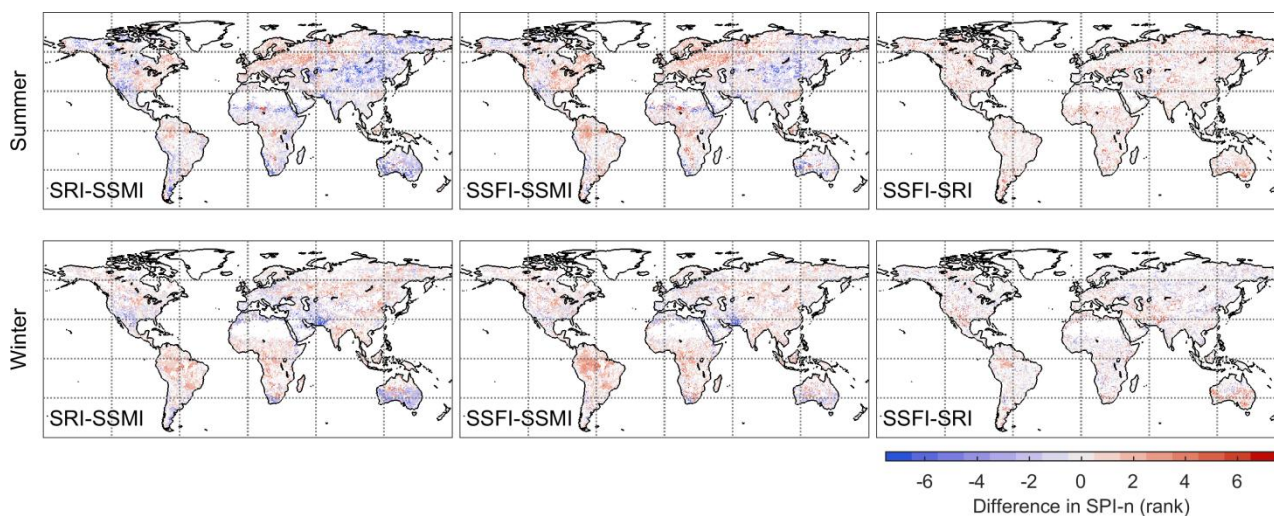
Figure 4: The difference in the rank of SPI-n for SRI and SSMI, SSFI and SSMI, SSFI and SRI. Pixels where the difference between accumulation periods are not statistically significant (p < 0.05) are masked.

The difference between the timescales of drought propagation to soil moisture and hydrological droughts can provide additional insights into the mechanisms of drought propagation in the models. The differences in the rank of SPI-n are shown in Fig. 4. As explained in Sect. 2.2, we use rank SPI-n rather than the duration in months because these are more useful in interpreting differences between (sub-)seasonal and annual timescales than the SPI-n in months. A difference of 1–2 rank SPI-n indicates that drought propagation occurs at similar timescales (i.e. sub-seasonal, seasonal or yearly time scales). Differences of more than four rank SPI-n represent large differences in drought propagation timescales, such as between sub-seasonal and yearly timescales. In summer, SPI-n for runoff are higher than for soil moisture in the Amazon, eastern North America, central Africa, and parts of Europe. This means that drought propagation to soil moisture is quicker than drought propagation to runoff, which suggests that subsurface runoff or baseflow is more important than surface runoff in these locations. The opposite is true for most parts of Australia, large parts of central and eastern Asia, and parts of western North America. In these locations, drought propagation to soil moisture drought is slower than drought propagation to runoff, which implies that surface runoff is an important component of total runoff. The differences can be substantial, with rank differences of five and more, which roughly represent the difference between sub-seasonal and annual timescales. In winter,

differences tend to be smaller, and longer drought propagation timescales for runoff than soil moisture (i.e. positive values in Fig. 4) are more common. Drought propagation timescales for streamflow droughts tend to be longer than for runoff drought, which is consistent with streamflow being routed runoff.

 Spearman correlation coefficients show that the difference in rank SPI-n of summer soil moisture and runoff droughts is related to the amount of surface runoff relative to total runoff ($\rho = -0.53$). The relationships with average annual precipitation ($\rho = 0.44$) and the runoff coefficient ($\rho = 0.36$) are slightly weaker. The amount of surface runoff relative to total runoff, annual average precipitation, and the runoff coefficient are also related to the difference between soil moisture and streamflow drought, though the relationships are slightly weaker (correlations are up to 0.1 lower). The difference in SPI-n between drought types in winter, as well as between runoff and streamflow droughts in summer, cannot be explained by these variables. Each of the correlation coefficients noted here is highly significant ($p < 0.01$).

In an additional analysis, we calculated SPI-n for the model ensemble mean focusing on mild droughts (SI ≤ -0.5) and moderate droughts (SI ≤ -1). When using mild droughts as a threshold value, the results are largely similar to those including near-normal conditions (SI ≤ 0), as shown in Fig. S2. Both the global patterns as well as the differences between the climate types are similar to the results shown in this section. However, the number of pixels masked due to insignificant correlations between SPI-n and the SSMI, SRI or SSFI increases. When moderate droughts are used as a threshold (Fig. S3), the proportion of pixels that are masked increases further, resulting in 40–50 % less data than when including near-normal conditions. In addition, the maps of SPI-n become noisier and the distributions of SPI-n by climate type (as shown in Fig. 3) flatten, especially in continental and polar climates. However, the relationships between the climate and seasonal group means remain similar.

The patterns of runoff drought propagation found in this study are similar to a previous study that calculated correlations between ensemble median runoff and precipitation percentiles (van Huijgevoort et al., 2013). Specifically, runoff is more closely related to shorter SPI (or precipitation percentile) accumulation periods in tropical regions, and to longer accumulation periods in continental and polar climates. However, more specific regional comparisons between these studies are hindered by differences in the approach of the two studies. In van Huijgevoort et al. (2013), correlations are based on the full precipitation and runoff time series, while in our study they are limited to below-average runoff conditions in either summer or winter months only. The results shown here are also in line with results from an observational study comparing SPI and SSFI in the United Kingdom. That study reported that SPI-n of 1–4 months were most closely related to SSFI, except in the southeast where some major aquifers are located and where longer SPI-n were found (Barker et al., 2016). That is just slightly shorter than the 2–6 months found in this study. Where drought propagation occurs at very short timescales, drought is mainly driven by precipitation deficits, possibly in combination with temperature anomalies (though these are not reflected in the SPI). Where drought propagation occurs at long timescales, attenuation by hydrological stores likely plays a more important role.

4.2 Model variability

The variability of SPI-n in the models underlying the model ensemble mean is shown in Fig. 5. Again, the standard deviation is not shown in months, but in the rank of SPI-n timescales studied. In summer, the standard deviation ranges from 1–4 rank SPI-n. Note that differences of 1–2 rank SPI-n indicate that models tend to agree on whether drought propagation occurs at

5   sub-seasonal, seasonal, or annual timescales. Model variability is low in temperate regions such as Europe and eastern North America, as well as in tropical regions such as the Amazon and Southeast Asia. The model variability is high in (semi-)arid regions such as the Sahel and central Australia. The patterns of model variability in hydrological drought propagation timescales are largely similar to those of soil moisture drought. However, the variability is patchier, and there are some regional differences. For example, model variability in SPI-n in tropical and temperate climates is slightly higher for runoff

10   than for soil moisture. In addition, model variability is lower in central Asia for runoff droughts than for soil moisture droughts.
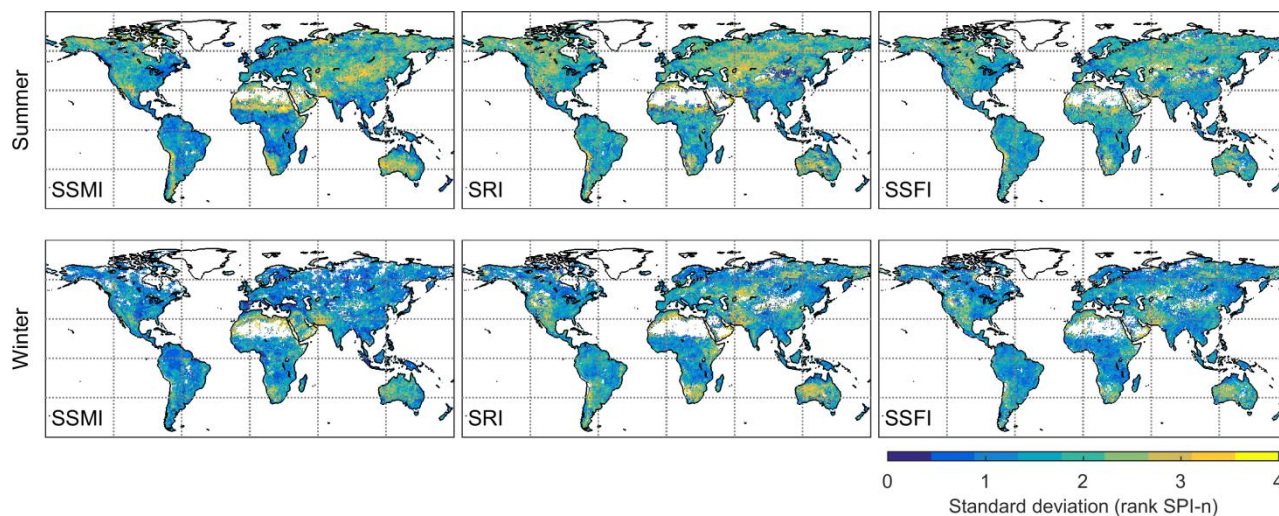
Figure 5: The standard deviation of the rank of SPI-n between the different hydrological models for summer and winter droughts
15   **in SSMI, SRI, and SSFI.**

Attributing observed differences between models to model characteristics is not straightforward, despite the common meteorological forcing, because there are considerable differences in model structures and parameterizations (Beck et al., 2016, 2017; Döll et al., 2016). However, we examine the relationship between drought propagation and specific model

20   characteristics to identify areas for further study. First, we examine the relationship between SPI-n and average soil moisture storage, as previous work has suggested that water storage plays an important role in drought propagation (Barker et al., 2016; Van Loon and Laaha, 2015). In addition, we examine the effect of model structural choices on drought propagation in an exploratory analysis.

Soil moisture storage in the models varies considerably between models, mainly due to differences in the definitions of root-zone soil moisture used to calculate the SSMI. The reported depths of the root zone range from 0.2 to 8 m, which may be a fixed value for all pixels or vary by pixel and/or land use type. Although we use standardized indices (SSMI) and not absolute values of soil moisture, the response time to changes in precipitation and/or evaporation will differ between soils

5 with large and small storage volumes. We examine the relationship between average root-zone soil moisture storage and average SPI-n for soil moisture droughts in Fig. 6. Soil moisture storage is averaged over space and time, and SPI-n in space, resulting in a single point for each model. The figure shows that drought propagation from meteorological to soil moisture drought is strongly related to average soil moisture storage, with correlation coefficients between 0.56 and 0.91, depending on the season and climate type. The impact of changes in storage on drought propagation is especially high in dry climates,

10 where relatively small differences in average soil moisture correspond to large differences in SPI-n. In comparison, the impact of storage on SPI-n is low in tropical wet climates. In general, changes in SPI-n with storage are smaller in winter than in summer. For tropical wet, continental, and polar climates the impact of storage is less than one rank SPI-n over the full range of storage volumes in winter, suggesting that storage is not an important driver of model variability in this season.
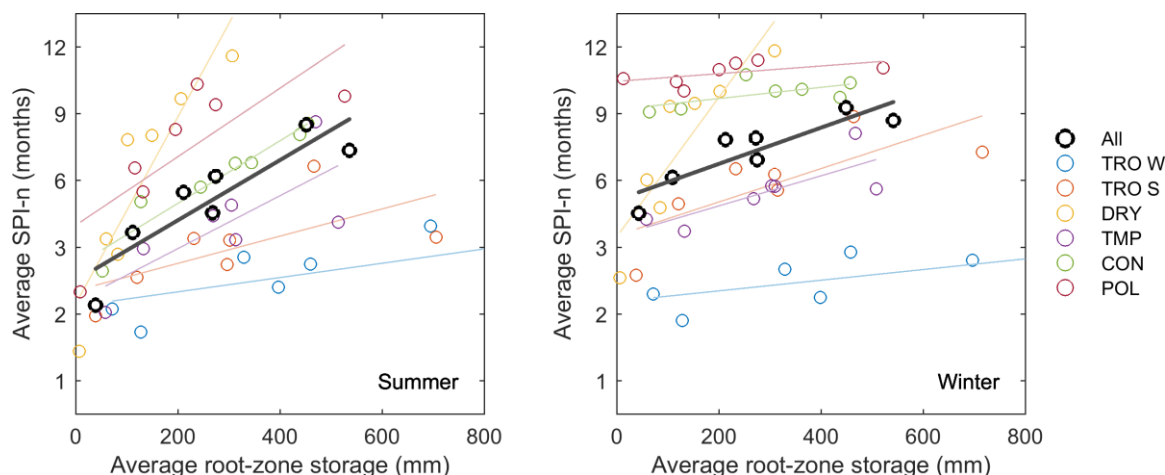
Figure 6: SPI-n for the SSMI averaged over all pixels and each climate type separately plotted against mean annual root-zone soil moisture storage averaged over the same region. Each point represents a model. Lines of best fit have been added for reference.

The relationship between soil moisture storage and drought propagation to runoff and streamflow is not included here

20 because the link is not as clear-cut. Runoff consists of surface and subsurface components. The subsurface component of runoff is related to water stored in groundwater and/or in the lowest layer of the soil, and not to the upper soil layers included in the root zone. Furthermore, groundwater data are not available for three out of seven models. The surface component, on the other hand, is only affected by soil moisture in as much that saturated conditions near the surface will result in a larger surface flow component. Therefore, it is impacted by the relative saturation of the soil rather than the storage itself.

In an exploratory analysis, we also examine the relationship between four qualitative factors related to model structure and the timescales of drought propagation. First, we compare SPI-n in LSMs and GHMs. Second, we study the effect of different evaporation schemes, specifically comparing Penman-Monteith evaporation schemes to more empirical temperature-based approaches. Third, we compare models that represent reservoirs with those that do not. Fourth, we group models by runoff
5 generation mechanisms, comparing models that include infiltration excess runoff generation to those that only represent saturation excess. In this last analysis, we exclude ORCHIDEE and WaterGAP3 because they use alternative methods of runoff generation; a Green-Ampt infiltration and a beta function, respectively.

The relative importance of these model characteristics based on Cohen's d (see Sect. 2.2) varies considerably over the different climates and drought types (Fig. 7). Overall, the tested model characteristics have a larger effect on soil moisture
10 droughts than on runoff and streamflow droughts. Grouping models by model type has the largest effect on mean soil moisture drought SPI-n for most climates, where drought propagation is slower in LSMs than in GHMs. For runoff droughts, on the other hand, GHMs tend to have higher SPI-n for runoff droughts than LSMs. However, it is unclear whether the representation of the energy balance is the primary source of the difference between the groups, or whether underlying or related model characteristics are key.
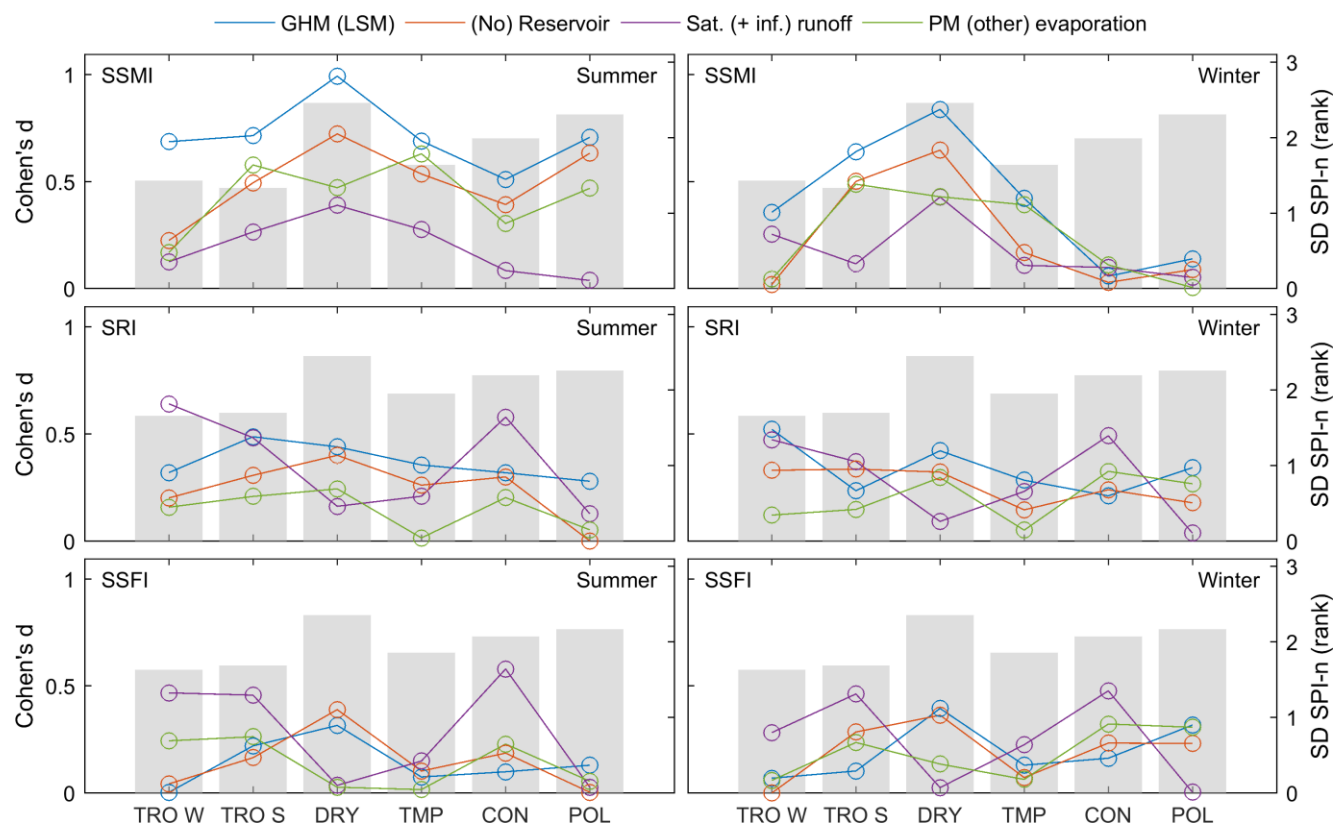
Figure 7: Effect sizes based on Cohen's d for model structural choices by climate for summer and winter droughts in SSMI, SRI, and SSFI. Bars represent the standard deviation of model SPI-n for each group. Abbreviations of climate types are the same as in Fig. 1.

5 Grouping models by runoff generation mechanisms has a smaller effect on average SPI-n for soil moisture droughts than the other studied factors. However, it can be more relevant for runoff and streamflow droughts, especially in tropical and continental climates. In these climates, including infiltration excess runoff leads to lower SPI-n and faster drought propagation. In tropical climates, this can be explained by the fact that high-intensity rainfall events exceeding the infiltration capacity of the soil are more common.

10 The simulation of reservoirs has a substantial impact on drought propagation, especially for soil moisture and runoff. This is consistent with previous work (Lorenzo-Lacruz et al., 2013), and also with increasing recognition that human adaptations should be taken into account in drought studies (Van Loon et al., 2016; Veldkamp et al., 2015). However, it is unclear what mechanism could be responsible for the effect of reservoirs on soil moisture droughts. The way in which reservoirs are implemented in global models affects runoff and streamflow, but has no direct link to soil moisture. Therefore, it is likely

15 that another mechanism is truly responsible for the observed differences.

Note that the model characteristics tested here are not fully independent. For example, LSMs are less likely to model reservoirs than GHMs. In addition, the number of models is small compared to the number of differences in model structures and parameterizations. Therefore, grouping models by different characteristics can result in identical groups, making it impossible to untangle the two in the current study setup. One example is the snow scheme. Previous studies have suggested

20 that using energy-based or temperature-based snow schemes results in different model behavior, both based on the same models used in this study (Beck et al., 2017) as well as in other models (Haddeland et al., 2011). In this study, however, it is impossible to distinguish between model type and snow scheme since all LSMs use an energy-based approach, while GHMs use temperature-based approaches.

Insight into which factors are truly responsible for model differences can only be gained through exhaustive experiments

25 testing different model parameterizations and structures (i.e. Medlyn et al., 2015). Nevertheless, while the effect sizes cannot be used to confirm that a certain model characteristic is the true factor underlying observed differences, they can be used to identify directions for further study in more comprehensive analyses.

4.3 Evaluation against observations

The timescales of drought propagation from meteorological to streamflow drought in the models and model ensemble mean

30 have been evaluated against data from 297 in-situ streamflow stations (Fig. 8). Of the individual models, W3RA performs best for summer droughts and WaterGAP3 for winter droughts based on MAE and Spearman correlations. The performance of the model ensemble mean is similar to that of the best-performing model in both seasons, with the lowest mean absolute errors and the second-highest correlations with observed SPI-n. WaterGAP3 is the only model that was calibrated against

streamflow observations, which could be a reason for the performance in winter, even though it is outperformed by other models in summer. While the mean absolute errors of the ensemble mean and individual models are within one or two rank SPI-n, correlations are low for all models.

In addition to the overall performance metrics, we compare the number of models for which the difference with observed

5   SPI-n is small (absolute error ≤ 1) and the number of models for which this difference is large (absolute error ≥ 4) at each site (Fig. 8). The thresholds are based on the models' (in)ability to capture the overall timescales of drought in terms of sub-seasonal, seasonal or yearly timescales. At least six out of seven models are within one step of the observed SPI-n at 18 % of the study sites, and at least four models are within one step at nearly half of the study sites (47 and 49 % for summer and winter, respectively). This suggests that models are well able to capture observed drought propagation timescales at these

10 sites. However, the majority of models differ by at least 4 steps of SPI-n at approximately 20 % of sites. At these sites, the models show substantially different drought propagation timescales. These differences in SPI-n correspond to differences between sub-seasonal and annual timescales of drought propagation. Models tend to do well in eastern North America and Europe, as well as western North America in winter, but poorly in central North America and Australia. However, the errors between models and observations are not related to climate, catchment size, or SPI-n.
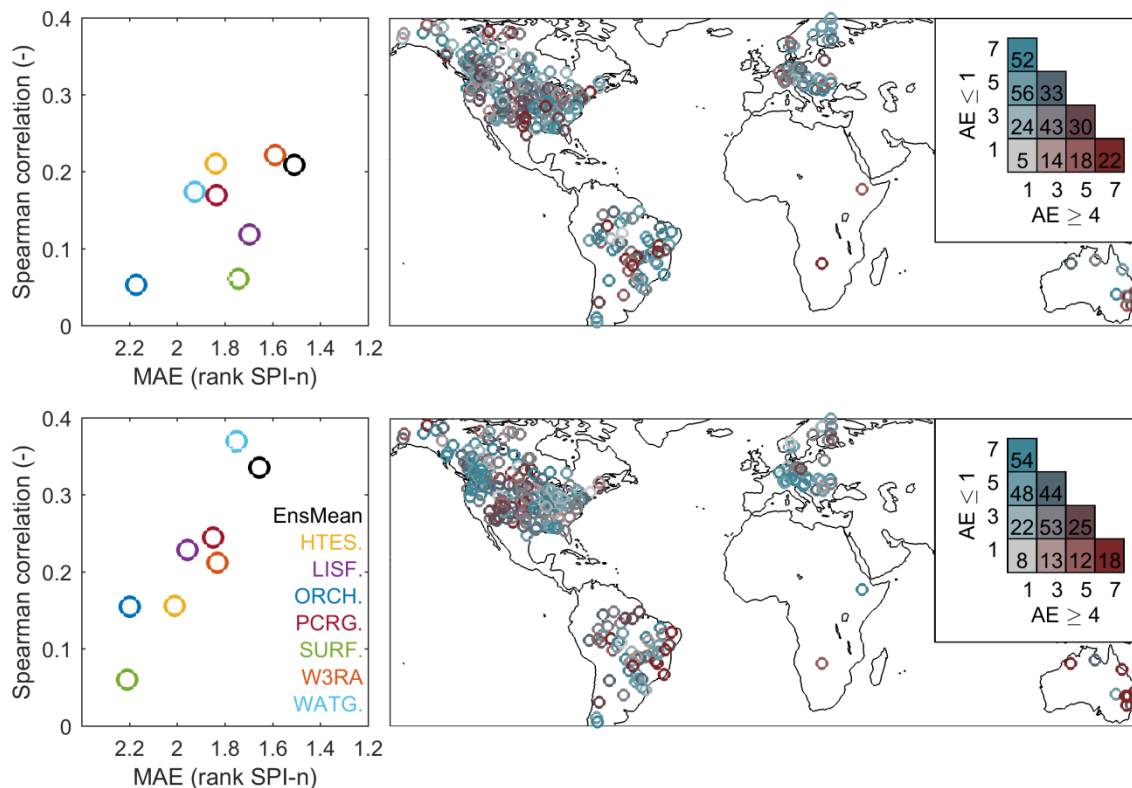
Figure 8: Comparison of model and observed summer (top) and winter (bottom) streamflow drought propagation timescales. The Spearman correlation and mean absolute error based on rank are compared, where better performing models fall in the upper

right corner. A map shows the number of models where the absolute error is ≤ 1 or ≥ 4 (right). The number of represented sites is indicated in the symbols of the legend.

It is important to note that some of the streamflow time series span as little as 15 years, which is shorter than the 30 years of data recommended for the calculation of SI. This means that the observational time series at some sites are too short to

5   capture the climatology of their locations. However, the average time series length (27 years) is close to the required 30 years. Furthermore, even where time series are relatively short we can evaluate whether the models capture the relationship between observed SPI and SSFI during the available time period. Another limitation of the evaluation is that the sites with observed data are not spread evenly across the globe, as most sites are located in the United States or Europe, with scarce data in Africa and Asia. Differences between the models and observations can be attributed to several types of errors (Van

10   Loon, 2015). The first type of error concerns errors in the model meteorological forcing data, but also in the GPCC precipitation data used to create the validation dataset. Then there are the errors in model structure and parameterizations of hydrologic processes, including the representation of anthropogenic influence on streamflow. Finally, there are errors in the (discretization of) the routing schemes employed by the models.

Unfortunately, evaluation of soil moisture drought propagation timescales is inhibited by the lack of root-zone soil moisture

15   data at global scale. While satellite soil moisture products are available, these are limited to the upper few centimeters of the soil (Owe et al., 2008), which is not representative of root-zone soil moisture. Field-measured soil moisture is also available, for example through the International Soil Moisture Network (Dorigo et al., 2011). However, only a handful of sites remain after applying the same site selection procedure as for streamflow drought validation (i.e. sufficiently long time series and a statistically significant relationship between SPI and SSMI).

20   **5 Conclusion**

This study evaluates timescales of drought propagation from meteorological to soil moisture and hydrological drought in an ensemble of seven land surface and global hydrological models. Drought propagation was quantified by cross-correlating standardized indices of hydrological variables. Here, we focus on soil moisture, runoff and streamflow droughts in summer and winter. However, the simple and flexible approach used here can be applied to other drought types, such as groundwater

25   droughts, and to other months or seasons.

Drought propagation is closely related to climate, with slower drought propagation in dry and continental climates and quicker drought propagation in tropical climates. Winter season drought propagation tends to be slower than in the summer, especially in tropical savanna and continental climates. This may be a result of the distinct wet and dry seasons in the former, and snow cover in the latter. Faster propagation of meteorological drought to runoff drought than to soil moisture drought

30   has been linked to a higher proportion of surface runoff, thereby causing a larger portion of total runoff to bypass the soil moisture store.

Model variability can be quite high, especially for summer droughts and in dry climates, where the socio-economic impacts of drought can be severe. Since the models were run with consistent forcing datasets, differences can be attributed to model parameterization and structure. For example, drought propagation from meteorological to soil moisture drought was generally slower in models with higher average soil moisture storage, and vice versa. Although the differences cannot be definitively attributed to specific model characteristics in the current experiment, we identified several directions for further study. In addition, grouping models by model type, runoff generation mechanisms, and representation of reservoirs all resulted in significantly different average drought propagation timescales.

The relationship between meteorological and streamflow drought in the global models was evaluated against observational data. Overall, the models were able to capture the timescales of drought propagation, as errors were relatively low on average. However, considerable model advancements can be made since there were large discrepancies between model and observed drought propagation at 20 % of the study sites.

A better understanding and representation of drought propagation in global models may improve drought forecasting (Cancelliere et al., 2007), especially when combined with the availability of accurate seasonal forecasts (Luo et al., 2007; Luo and Wood, 2007). Drought forecasting potential is expected to be higher in regions with relatively slow drought propagation, such as the dry and continental climates in this study, as drought forecasting for longer SPI-n tends to be more accurate than for shorter SPI-n (Mishra and Desai, 2005). Additional research using lagged SPI-n could assess the potential for forecasting different types of drought based on meteorological data. Improved representation of drought propagation in models is also crucial to constrain the impact of climate change on drought frequency and severity, and thereby improve the reliability of projected changes.

Data availability

The meteorological forcing, model outputs, and standardized indices from the eartH2Observe project are openly available from the project portal (www.earth2observe.eu). GPCC daily precipitation data can be obtained via https://www.esrl.noaa.gov/psd/data/gridded/data.gpcc.html and GRDC monthly streamflow data are available at http://grdc.bafg.de.

Acknowledgements

This research received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 603608, Global Earth Observation for integrated water resource assessment: eartH2Observe.

References

Andreadis, K. M. and Lettenmaier, D. P.: Trends in 20th century drought over the continental United States, Geophys. Res. Lett., 33(10), doi:10.1029/2006GL025711, 2006.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B. J. J. M., Hirschi, M. and Betts, A. K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, J. Hydrometeorol., 10(3), 623–643, doi:10.1175/2008JHM1068.1, 2009.

Barella-Ortiz, A., Polcher, J., Tuzet, A. and Laval, K.: Potential evaporation estimation through an unstressed surface-energy balance and its sensitivity to climate change, Hydrol. Earth Syst. Sci., 17(11), 4625–4639, doi:10.5194/hess-17-4625-2013, 2013.

Barker, L. J., Hannaford, J., Chiverton, A. and Svensson, C.: From meteorological to hydrological drought using standardised indicators, Hydrol. Earth Syst. Sci., 20, 2483–2505, doi:10.5194/hess-20-2483-2016, 2016.

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Miralles, D. G., McVicar, T. R., Schellekens, J. and Bruijnzeel, L. A.: Global-scale regionalization of hydrologic model parameters, Water Resour. Res., 52(5), 3599–3622, doi:10.1002/2015WR018247, 2016.

Beck, H. E., Van Dijk, A. I. J. M., De Roo, A., Dutra, E., Fink, G., Orth, R. and Schellekens, J.: Global evaluation of runoff from ten state - of - the - art hydrological models, Hydrol. Earth Syst. Sci., 21, 2881–2903, doi:https://doi.org/10.5194/hess-21-2881-2017, 2017.

van Beek, L. P. H., Wada, Y. and Bierkens, M. F. P.: Global monthly water stress: 1. Water balance and water availability, Water Resour. Res., 47(7), W07517, doi:10.1029/2010WR009791, 2011.

Bond, W. J., Woodward, F. I. and Midgley, G. F.: The global distribution of ecosystems in a world without fire., New Phytol., 165(2), 525–537 [online] Available from: http://www.ncbi.nlm.nih.gov/pubmed/15720663, 2005.

Burke, E. J. and Brown, S. J.: Evaluating Uncertainties in the Projection of Future Drought, J. Hydrometeorol., 9(2), 292–299, doi:10.1175/2007JHM929.1, 2008.

Cancelliere, A., Mauro, G. Di, Bonaccorso, B. and Rossi, G.: Drought forecasting using the standardized precipitation index, Water Resour. Manag., 21(5), 801–819, doi:10.1007/s11269-006-9062-y, 2007.

Cohen, J.: Statistical power analysis for the behavioral sciences, in Statistical Power Analysis for the Behavioral Sciences., 1988.

D'Orgeval, T., Polcher, J. and de Rosnay, P.: Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes, Hydrol. Earth Syst. Sci., 12(6), 1387–1401, doi:10.5194/hess-12-1387- 489 2008, 2008.

Dai, A. G.: Increasing drought under global warming in observations and models, Nat. Clim. Chang., 3(1), 52–58, doi:10.1038/nclimate1633, 2013.

Decharme, B., Alkama, R., Douville, H., Becker, M. and Cazenave, A.: Global Evaluation of the ISBA-TRIP Continental Hydrological System. Part II: Uncertainties in River Routing Simulation Related to Flow Velocity and Groundwater Storage,

J. Hydrometeorol., 11(3), 601–617, doi:10.1175/2010JHM1212.1, 2010.

Decharme, B., Martin, E. and Faroux, S.: Reconciling soil thermal and hydrological lower boundary conditions in land surface models, J. Geophys. Res. Atmos., 118(14), 7819–7834, doi:10.1002/jgrd.50631, 2013.

van Dijk, A. I. J. M.: The Australian Water Resources Assessment System; Technical Report 3, Landscape Model (version 0.5) technical description, Canberra, ACT, Australia., 2010.

Van Dijk, A. I. J. M., Renzullo, L. J., Wada, Y. and Tregoning, P.: A global water cycle reanalysis (2003-2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble, Hydrol. Earth Syst. Sci., doi:10.5194/hess-18-2955-2014, 2014.

Döll, P., Fiedler, K. and Zhang, J.: Global-scale analysis of river flow alterations due to water withdrawals and reservoirs, Hydrol. Earth Syst. Sci., 13(12), 2413–2432, doi:10.5194/hess-13-2413-2009, 2009.

Döll, P., Douville, H., Güntner, A., Müller Schmied, H. and Wada, Y.: Modelling Freshwater Resources at the Global Scale: Challenges and Prospects, Surv. Geophys., 37(2), 195–221, doi:10.1007/s10712-015-9343-1, 2016.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A. and Jackson, T.: The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci., doi:10.5194/hess-15-1675-2011, 2011.

Flörke, M., Kynast, E., Bärlund, I., Eisner, S., Wimmer, F. and Alcamo, J.: Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study, Glob. Environ. Chang., 23(1), 144–156, doi:10.1016/j.gloenvcha.2012.10.018, 2013.

Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N., Voss, F. and Koirala, S.: Comparing Large-Scale Hydrological Model Simulations to Observed Runoff Percentiles in Europe, J. Hydrometeorol., 13(2), 604–620, doi:10.1175/JHM-D-11-083.1, 2012a.

Gudmundsson, L., Wagener, T., Tallaksen, L. M. and Engeland, K.: Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe, Water Resour. Res., 48(11), 1–20, doi:10.1029/2011WR010911, 2012b.

Guttman, N. B.: Accepting the standardized precipitation index: A calculation algorithm, J. Am. Water Resour. Assoc., 35(2), 311–322, doi:10.1111/j.1752-1688.1999.tb03592.x, 1999.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P. and Yeh, P.: Multimodel estimate of the global terrestrial water balance: setup and first results, J. Hydrometeorol., 12, 869–884, doi:10.1175/2011JHM1324.1, 2011.

Hao, Z. and AghaKouchak, A.: Multivariate Standardized Drought Index: A parametric multi-index model, Adv. Water Resour., 57, 12–18, doi:10.1016/j.advwatres.2013.03.009, 2013.

Hao, Z. and AghaKouchak, A.: A Nonparametric Multivariate Multi-Index Drought Monitoring Framework, J. Hydrometeorol., 15(1), 89–101, doi:10.1175/JHM-D-12-0160.1, 2014.

Haslinger, K., Koffler, D., Schöner, W. and Laaha, G.: Exploring the link between meteorological drought and streamflow: Effects of climate-catchment interaction, Water Resour. Res., 50, 2468–2487, doi:10.1002/2013WR015051, 2014.

Hayes, M., Svoboda, M., Whilhite, D. A. and Wilhite, D. A.: Monitoring drought using the standardized precipitation index, Drought a Glob. Assess. Vol. 1, 80(3), 429–438, doi:http://dx.doi.org/10.1108/17506200710779521, 1999.

Horridge, M., Madden, J. and Wittwer, G.: The impact of the 2002-2003 drought on Australia, J. Policy Model., doi:10.1016/j.jpolmod.2005.01.008, 2005.

Huang, S., Leng, G., Huang, Q., Xie, Y., Liu, S., Meng, E. and Li, P.: The asymmetric impact of global warming on US drought types and distributions in a large ensemble of 97 hydro-climatic simulations, Sci. Rep., doi:10.1038/s41598-017-06302-z, 2017.

van Huijgevoort, M. H. J., Hazenberg, P., van Lanen, H. A. J., Teuling, A. J., Clark, D. B., Folwell, S., Gosling, S. N., Hanasaki, N., Heinke, J., Koirala, S., Stacke, T., Voss, F., Sheffield, J. and Uijlenhoet, R.: Global Multimodel Analysis of Drought in Runoff for the Second Half of the Twentieth Century, J. Hydrometeorol., 14(5), 1535–1552, doi:10.1175/JHM-D-12-0186.1, 2013.

Van Der Knijff, J. M., Younis, J. and De Roo, A. P. J.: LISFLOOD: a GIS‐ based distributed model for river basin scale water balance and flood simulation, Int. J. Geogr. Inf. Sci., 24(2), 189–212, doi:10.1080/13658810802549154, 2010.

Kottek, M., Grieser, J., Beck, C., Rudolf, B. and Rubel, F.: World map of the Köppen-Geiger climate classification updated, Meteorol. Zeitschrift, doi:10.1127/0941-2948/2006/0130, 2006.

Krinner, G., Viovy, N., de Noblet-ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochem. Cycles, doi:10.1029/2003GB002199, 2005.

Kummu, M., Guillaume, J. H. A., de Moel, H., Eisner, S., Flörke, M., Porkka, M., Siebert, S., Veldkamp, T. I. E. and Ward, P. J.: The world's road to water scarcity: shortage and stress in the 20th century and pathways towards sustainability, Sci. Rep., doi:10.1038/srep38495, 2016.

Van Lanen, H. A. J., Wanders, N., Tallaksen, L. M. and Van Loon, A. F.: Hydrological drought across the world: Impact of climate and physical catchment structure, Hydrol. Earth Syst. Sci., 17(5), 1715–1732, doi:10.5194/hess-17-1715-2013, 2013.

Lehner, B., Döll, P., Alcamo, J., Henrichs, T. and Kaspar, F.: Estimating the impact of global change on flood and drought risks in Europe: A continental, integrated analysis, Clim. Change, doi:10.1007/s10584-006-6338-4, 2006.

Lenth, R. V: Some Practical Guidelines for Effective Sample Size Determination Some Practical Guidelines for Effective Sample Size Determination, Am. Stat., doi:10.1198/000313001317098149, 2001.

Lewis, J. and Sjöstrom, J.: Optimizing the experimental design of soil columns in saturated and unsaturated transport experiments., J. Contam. Hydrol., 115(1–4), 1–13 [online] Available from: http://www.ncbi.nlm.nih.gov/pubmed/20452088, 2010.

Lloyd-Hughes, B. and Saunders, M. A.: A drought climatology for Europe, Int. J. Climatol., doi:10.1002/joc.846, 2002.

Lloyd-Hughes, B., Shaffrey, L. C., Vidale, P. L. and Arnell, N. W.: An evaluation of the spatiotemporal structure of large-

scale European drought within the HiGEM climate model, Int. J. Climatol., doi:10.1002/joc.3570, 2013.

Van Loon, A. F.: Hydrological drought explained, Wiley Interdiscip. Rev. Water, 2(4), 359–392, doi:10.1002/wat2.1085, 2015.

Van Loon, A. F. and Laaha, G.: Hydrological drought severity explained by climate and catchment characteristics, J. Hydrol., 526, 3–14, doi:10.1016/j.jhydrol.2014.10.059, 2015.

Van Loon, A. F., Van Huijgevoort, M. H. J. and Van Lanen, H. A. J.: Evaluation of drought propagation in an ensemble mean of large-scale hydrological models, Hydrol. Earth Syst. Sci., 16(11), 4057–4078, doi:10.5194/hess-16-4057-2012, 2012.

Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N. and Van Lanen, H. A. J.: Drought in the Anthropocene, Nat. Geosci., 9(2), 89–91, doi:10.1038/ngeo2646, 2016.

Lorenzo-Lacruz, J., Vicente-Serrano, S. M., González-Hidalgo, J. C., López-Moreno, J. I. and Cortesi, N.: Hydrological drought response to meteorological drought in the Iberian Peninsula, Clim. Res., 58, 117–131, doi:10.3354/cr01177, 2013.

Luo, L. and Wood, E. F.: Monitoring and predicting the 2007 U.S. drought, Geophys. Res. Lett., 34(22), doi:10.1029/2007GL031673, 2007.

Luo, L., Wood, E. F. and Pan, M.: Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions, J. Geophys. Res. Atmos., 112(10), doi:10.1029/2006JD007655, 2007.

Mckee, T. B., Doesken, N. J. and Kleist, J.: The relationship of drought frequency and duration to time scales, AMS 8th Conf. Appl. Climatol., (January), 179–184, doi:citeulike-article-id:10490403, 1993.

Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., Hickler, T., Jain, A. K., Luo, Y., Parton, W., Prentice, I. C., Thornton, P. E., Wang, S., Wang, Y.-P., Weng, E., Iversen, C. M., McCarthy, H. R., Warren, J. M., Oren, R. and Norby, R. J.: Using ecosystem experiments to improve vegetation models, Nat. Clim. Chang., 5(6), 528–534, doi:10.1038/nclimate2621, 2015.

Mishra, A. K. and Desai, V. R.: Drought forecasting using stochastic models, Stoch. Environ. Res. Risk Assess., 19(5), 326–339, doi:10.1007/s00477-005-0238-4, 2005.

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

Naresh Kumar, M., Murthy, C. S., Sesha sai, M. V. R. and Roy, P. S.: On the use of Standardized Precipitation Index (SPI) for drought intensity assessment, Meteorol. Appl., 16, 381–389, doi:10.1002/met.136, 2009.

Ngo-Duc, T., Laval, K., Ramillien, G., Polcher, J. and Cazenave, A.: Validation of the land water storage simulated by Organising Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) with Gravity Recovery and Climate Experiment (GRACE) data, Water Resour. Res., doi:10.1029/2006WR004941, 2007.

Owe, M., de Jeu, R. A. M. and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, J. Geophys. Res., 113, F01002, doi:10.1029/2007JF000769, 2008.

Peters, E., Torfs, P. J. J. F., van Lanen, H. A. J. and Bier, G.: Propagation of drought through groundwater - A new approach using linear reservoir theory, Hydrol. Process., 17, 3023–3040, doi:10.1002/hyp.1274, 2003.

Prudhomme, C., Parry, S., Hannaford, J., Clark, D. B., Hagemann, S. and Voss, F.: How Well Do Large-Scale Models Reproduce Regional Hydrological Extremes in Europe?, J. Hydrometeorol., doi:10.1175/2011JHM1387.1, 2011.

5 Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D.: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment, Proc. Natl. Acad. Sci. U. S. A., 111(9), 3262–3267, doi:10.1073/pnas.1222473110, 2014.

Pyper, B. J. and Peterman, R. M.: Comparison of methods to account for autocorrelation in correlation analyses of fish data,
10 Can. J. Fish. Aquat. Sci., 55(9), 2127–2140, doi:10.1139/f98-104, 1998.

Reichstein, M., Bahn, M., Ciais, P., Frank, D., Mahecha, M. D., Seneviratne, S. I., Zscheischler, J., Beer, C., Buchmann, N., Frank, D. C., Papale, D., Rammig, A., Smith, P., Thonicke, K., van der Velde, M., Vicca, S., Walz, A. and Wattenbach, M.: Climate extremes and the carbon cycle, Nature, 500, 287–295, doi:10.1038/nature12350, 2013.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., van Dijk, A., Sperna Weiland, F., Minvielle, M., Calvet, J.-
15 C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W. and Weedon, G. P.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, Earth Syst. Sci. Data Discuss., 1–35, doi:10.5194/essd-2016-55, 2016.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B. and Ziese, M.: GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data, Glob. Precip.
20 Climatol. Cent., doi:10.5676/DWD_GPCC/FD_M_V6_050, 2015.

Sheffield, J. and Wood, E. F.: Global trends and variability in soil moisture and drought characteristics, 1950-2000, from observation-driven simulations of the terrestrial hydrologic cycle, J. Clim., 21(3), 432–458, doi:10.1175/2007JCLI1822.1, 2008a.

Sheffield, J. and Wood, E. F.: Projected changes in drought occurrence under future global warming from multi-model,
25 multi-scenario, IPCC AR4 simulations, Clim. Dyn., 31(1), 79–105, doi:10.1007/s00382-007-0340-z, 2008b.

Sheffield, J., Goteti, G., Wen, F. and Wood, E. F.: A simulated soil moisture based drought analysis for the United States, J. Geophys. Res. D Atmos., 109(24), 1–19, doi:10.1029/2004JD005182, 2004.

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.

30 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, doi:10.1002/joc.4267, 2015.

Stahl, K., Tallaksen, L. M., Hannaford, J. and Van Lanen, H. A. J.: Filling the white space on maps of European runoff trends: Estimates from a multi-model ensemble, Hydrol. Earth Syst. Sci., 16(7), 2035–2047, doi:10.5194/hess-16-2035-2012, 2012.

Stanke, C., Kerac, M., Prudhomme, C., Medlock, J. and Murray, V.: Health Effects of Drought: A Systematic Review of the Evidence, PLoS Curr., doi:10.1371/currents.dis.7a2cee9e980f91ad7697b570bcc4b004, 2013.

Tallaksen, L. M. and Stahl, K.: Spatial and temporal patterns of large-scale droughts in Europe: Model dispersion and performance, Geophys. Res. Lett., 41(2), 429–434, doi:10.1002/2013GL058573, 2014.

5   Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R. and Sheffield, J.: Global warming and changes in drought, Nat. Clim. Chang., 4, 17–22, doi:10.1038/nclimate2067, 2014.

Turco, M., von Hardenberg, J., AghaKouchak, A., Llasat, M. C., Provenzale, A. and Trigo, R. M.: On the key role of droughts in the dynamics of summer fires in Mediterranean Europe, Sci. Rep., doi:10.1038/s41598-017-00116-9, 2017.

Veldkamp, T. I. E., Wada, Y., de Moel, H., Kummu, M., Eisner, S., Aerts, J. C. J. H. and Ward, P. J.: Changing mechanism

10 of global water scarcity events: Impacts of socioeconomic changes and inter-annual hydro-climatic variability, Glob. Environ. Chang., 32, 18–29, doi:10.1016/j.gloenvcha.2015.02.011, 2015.

Vicente-Serrano, S. M., Gouveia, C., Camarero, J. J., Begueria, S., Trigo, R., Lopez-Moreno, J. I., Azorin-Molina, C., Pasho, E., Lorenzo-Lacruz, J., Revuelto, J., Moran-Tejeda, E. and Sanchez-Lorenzo, A.: Response of vegetation to drought time-scales across global land biomes, Proc. Natl. Acad. Sci., 110(1), 52–57, doi:10.1073/pnas.1207068110, 2013.

15 Wada, Y., Van Beek, L. P. H. and Bierkens, M. F. P.: Modelling global water stress of the recent past: On the relative importance of trends in water demand and climate variability, Hydrol. Earth Syst. Sci., 15(12), 3785–3808, doi:10.5194/hess-15-3785-2011, 2011.

Wada, Y., Wisser, D. and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dynam., 5(1), 15–40, doi:10.5194/esd-5-15-2014, 2014.

20 Wang, A., Bohn, T. J., Mahanama, S. P., Koster, R. D. and Lettenmaier, D. P.: Multimodel ensemble reconstruction of drought over the continental United States, J. Clim., 22(10), 2694–2712, doi:10.1175/2008JCLI2586.1, 2009.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J. and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing data methodology applied to ERA-Interim reanalysis data, Water Resour. Res., 50(9), 7505–7514, doi:10.1002/2014WR015638, 2014.

25 Wegren, S. K.: Food Security and Russia's 2010 Drought, Eurasian Geogr. Econ., doi:10.2747/1539-7216.52.1.140, 2011.

Wu, H., Svoboda, M. D., Hayes, M. J., Wilhite, D. A. and Wen, F.: Appropriate application of the Standardized Precipitation Index in arid locations and dry seasons, Int. J. Climatol., 27, 65–79, doi:10.1002/joc.1371, 2007.

Yamazaki, D., Kanae, S., Kim, H. and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resour. Res., doi:10.1029/2010WR009726, 2011.

30 Zargar, A., Sadiq, R., Naser, B. and Khan, F. I.: A review of drought indices, Environ. Rev., 19, 333–349, doi:10.1139/a11-013, 2011.