AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2017JD027577

Key Points:
- State-of-the art climate models systematically underestimate drought persistence
- A method for partitioning sources of uncertainty in model errors is developed
- Observation and model uncertainty dominate the total spread in drought persistence error

Supporting Information:
- Supporting Information S1

Correspondence to:
H. Moon and S. I. Seneviratne,
heewon.moon@env.ethz.ch;
sonia.seneviratne@ethz.ch

Citation:
Moon, H., Gudmundsson, L., & Seneviratne, S. I. (2018). Drought persistence errors in global climate models. *Journal of Geophysical Research: Atmospheres*, *123*, 3483–3496. https://doi.org/10.1002/2017JD027577

Received 10 AUG 2017
Accepted 3 MAR 2018
Accepted article online 7 MAR 2018
Published online 12 APR 2018

Drought Persistence Errors in Global Climate Models

H. Moon[1], L. Gudmundsson[1], and S. I. Seneviratne[1]

[1]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

Abstract The persistence of drought events largely determines the severity of socioeconomic and ecological impacts, but the capability of current global climate models (GCMs) to simulate such events is subject to large uncertainties. In this study, the representation of drought persistence in GCMs is assessed by comparing state-of-the-art GCM model simulations to observation-based data sets. For doing so, we consider dry-to-dry transition probabilities at monthly and annual scales as estimates for drought persistence, where a dry status is defined as negative precipitation anomaly. Though there is a substantial spread in the drought persistence bias, most of the simulations show systematic underestimation of drought persistence at global scale. Subsequently, we analyzed to which degree (i) inaccurate observations, (ii) differences among models, (iii) internal climate variability, and (iv) uncertainty of the employed statistical methods contribute to the spread in drought persistence errors using an analysis of variance approach. The results show that at monthly scale, model uncertainty and observational uncertainty dominate, while the contribution from internal variability is small in most cases. At annual scale, the spread of the drought persistence error is dominated by the statistical estimation error of drought persistence, indicating that the partitioning of the error is impaired by the limited number of considered time steps. These findings reveal systematic errors in the representation of drought persistence in current GCMs and suggest directions for further model improvement.

1. Introduction

The persistence of droughts is one of the main attributes that determines the level of their socioeconomic and ecological impacts (Bêche et al., 2009; Harou et al., 2010; Kelley et al., 2015). Lately, several studies have highlighted that climate models tend to underestimate drought persistence compared to the observational record (Ault et al., 2014; Wetter et al., 2014), which contributes to the uncertainty of future projections of drought risk. Furthermore, the characteristics of extremely prolonged droughts have been found to be similar over different time frames, including the twentieth century and longer paleoclimatic records (Ault et al., 2014). However, the driving mechanism of such droughts remains controversial and likely differs between individual events (Ault et al., 2014; Cook et al., 2015; Dai, 2011; Griffin & Anchukaitis, 2014; Shanahan et al., 2009; Stine, 1994). Thus, an in-depth assessment of the underestimation of drought persistence in current climate models in the recent past century may help to better understand similar issues found for longer time scales.

Measures of persistence (or memory) of drought events and related climate phenomena have been used to analyze drought variability in both models and observations. Commonly used measures include power spectra of the time series (Ault et al., 2014; Pelletier & Turcotte, 1997) or the Hurst exponent (Mandelbrot & Wallis, 1969) which have been used, for example, to measure long-term persistence in time series of drought indices (Tatli, 2015) or precipitation (Bunde et al., 2013; Kumar et al., 2013). By transforming continuous time series of drought indicators into categorical time series of dry and wet spells, the persistence of drought can be effectively estimated through the parameters of a binary first-order Markov chain model (Sericola, 2013). In this case the dry-to-dry transition probability (P_{dd}) is used as a measure of drought persistence (Jackson, 1975; Sharma & Panu, 2014; Wilby et al., 2015). The primary assumption for this measure is that the probability of drought and nondrought events is conditional only upon the previous time step. Interestingly, the dry spell length distribution can also be approximated by the geometric distribution given the assumption that the dry spell lengths that are estimated are binary sequences of independent trial results (in this case, drought and nondrought), and the probability of success (in this case, probability of drought) is equal for all the trials (Lee et al., 1986; Mathier et al., 1992).

Quantifications of drought persistence errors in a model ensemble are themselves subject to several sources of uncertainty, which include (i) uncertainty of the considered observations, (ii) uncertainty of the individual

models, (iii) internal variability, and (iv) errors related to the statistical estimation of drought persistence (i.e., finite sample effects). Assessing the relative contribution of each of these sources to the total uncertainty will aid interpretation in a model validation exercise. Uncertainty in the observation-based data sets generally refers to the combination of observational errors and the uncertainty related to data processing methods which depend on the interpolation scheme, station density, and quality control tools (Xie et al., 1996) as well as the subset of selected stations. Previous studies that have assessed global climate models (GCMs) globally usually relied on a single observational data set as a reference and sometimes accounted for observational uncertainty in the analysis by excluding the regions where observations were deemed unreliable. An alternative approach to account for observational uncertainty is to use independent observation-based data sets for the analysis and to investigate the associated spread. Model uncertainty can be characterized as the spread across different models, mainly generated from the discrepancy in their structure. Internal climate variability (or natural climate variability) corresponds to fluctuations generated in the climate system without changes in external forcing and can thus be approximated from the spread across initial condition ensembles of individual GCMs. Lastly, the fact that model evaluation is based on finite data sources leads to uncertainty in the statistical characterization of model performance. This statistical estimation error is often omitted in model validation studies, although it could affect the robustness of the analysis and hence also the conclusions.

In this study, we aim to investigate the ability of GCMs to simulate drought persistence, measured through the dry-to-dry transition probability (P_{dd}) in a Markov chain framework. Currently, a number of drought indicators are used in the scientific literature. Typically, these indices are tailored to specific types of analysis and are often categorized into meteorological (Keyantash & Dracup, 2002) hydrological (Tallaksen & van Lanen, 2004), or agricultural drought (Quiring & Papakryiakou, 2003). In addition, indicators like the Standardized Precipitation Index (McKee et al., 1993) or the Standardized Precipitation and Evapotranspiration Index (Vicente-Serrano et al., 2009) depend on the choice of a model distribution which in turn needs to be calibrated and is subject to uncertainty (Stagge et al., 2015). Finally, the choice of threshold to discriminate droughts from nondrought events is highly contextual and sometimes even subjective (Steinemann, 2014). Consequently, there is no single unifying drought definition, and the choice of drought indicator is usually dependent on the scope of individual assessments. Here we use precipitation anomalies as a drought indicator to separate dry (negative anomalies) from wet spells, which form the basis for estimating P_{dd}. The rationale underlying this decision is that precipitation is one of the climate variables that has been monitored with the best spatial and temporal coverage throughout the past century, and few additional assumptions are required (see section 3). In order to robustly assess the representation of P_{dd} in the models and to fully characterize the spread of the model error, we compare multiple observation-based data sets with a large ensemble of GCMs runs. Finally, we assess to which degree the spread in drought persistence error can be attributed to differences among models, observation uncertainty, internal variability, and the statistical estimation error of P_{dd} using an analysis of variance (ANOVA)-based approach. We close with a discussion on the link between the drought persistence error identified in this study to the issues found in precipitation variability and suggest potential implication of using the error partitioning method introduced in this study for the interpretation of climate model projections.

2. Data

2.1. Precipitation From the Fifth Phase of Coupled Model Intercomparison Projects Model Simulations

With the international consensus on establishing global Coupled Model Intercomparison Projects (CMIP), there have been continued phases of CMIPs since the early 2000s. The GCM simulations of the fifth phase of CMIP (CMIP5; K. E. Taylor, Stouffer, & Meehl, 2012) include historical simulations for the period 1850–2005 and Representative Concentration Pathways experiments built according to different greenhouse gas concentration scenarios from 2006 onward. Here we consider historical simulations (1901–2005) and Representative Concentration Pathway 8.5 simulations (2006–2010) to cover the 1901–2010 period. Only models with at least three initial condition ensemble members were included in the analysis to account for internal variability (Table 1).

2.2. Interpolated Precipitation Observations and Reanalysis Products

To quantitatively assess the performance of the considered GCM simulations, five observation-based global precipitation data sets were selected, all covering at least the 1901–2010 time period (Table 2). The Climatic

Table 1
GCMs Used in This Study

GCM	Institution (country)	Ensemble members
CanESM2	Canadian Centre for Climate Modeling and Analysis (Canada)	5
CCSM4	National Center for Atmospheric Research (United States)	6
CESM1-CAM5	National Science Foundation, Department of Energy, National Center for Atmospheric Research (United States)	3
CNRM-CM5	Centre National de Recherches Météorologiques / Centre Européen de Recherche et Formation Avancées en Calcul Scientifique (France)	5
CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organisation in Collaboration with the Queensland Climate Change Centre of Excellence (Australia)	10
EC-EARTH	EC-EARTH Consortium	6
FIO-ESM	The First Institute of Oceanography, SOA (China)	3
GISS-E2-H	NASA Goddard Institute for Space Studies (United States)	5
GISS-E2-R		5
HadGEM2-ES	Met Office Hadley Centre (UK)	4
IPSL-CM5A-LR	Institut Pierre Simon Laplace (France)	4
MIROC5	Atmosphere and Ocean Research Institute, The University of Tokyo (Japan)	3
MPI-ESM-LR	Max Planck Institute for Meteorology (Germany)	3

Note. Names of GCMs, modeling institution, and number of ensemble simulations are given. GCM = global climate model.

Research Unit Timeseries version 3.1 (Harris et al., 2014), University of Delaware version 3.0.1 (Matsuura & Willmott, 2012), and Global Precipitation Climatology Centre (Schneider et al., 2014) data are generated from interpolated station data, and each data product uses a different subset of available stations as well as different interpolation and quality control methods. In addition, we use the 20CR (20th Century Reanalysis; Compo et al., 2006) and the European Centre for Medium-Range Weather Forecasts twentieth century reanalysis (Poli et al., 2016), which assimilate different variables and differ both with respect to the underlying models and assimilation schemes.

3. Methods

3.1. Quantifying Drought Persistence

To quantify drought persistence, we convert observed and modeled precipitation values into binary time series indicating dry (negative) and wet (positive) anomalies. For annual time series anomalies are computed by subtracting the linear trend (least squares estimate). For the monthly resolution, the series were first detrended using linear least squares regression. Subsequently, the long-term mean of each month was removed from the detrended time series. Drought persistence is then measured through the dry-to-dry transition probability P_{dd}, which is defined as the proportion of dry-to-dry transition out of all transitions from a dry status. The theoretical minimum of P_{dd} is 0.5 by construction as this value indicates a white noise system without memory (unless there is a tendency to systematically have wetter conditions after dry conditions). In a few cases P_{dd} values estimated from observed or simulated climate variables are slightly below the theoretical minimum, which can be explained by the sampling uncertainty. In the analyses, we consider both

Table 2
Observation-Based Data Set Used in This Study

Observation-based data set	Data features	Provided resolution
CRU TS3.1	3,500–9,000 weather station records from CLIMAT	Monthly, 0.5° × 0.5°
UDEL	4,100–18,000 weather station records mainly from GHCN v2	Monthly, 0.5° × 0.5°
GPCC	11,000–49,450 gauge stations	Monthly, 0.5° × 0.5°, 1.0° × 1.0°, and 2.5° × 2.5°
20CR	Assimilating variable: surface pressure	Subdaily, daily, and monthly, 2° × 2°
ERA-20C	Assimilating variable: surface pressure, surface winds	Subdaily, daily, and monthly, approximately 128 km × 128 km

Note. Names of data sets, number of minimum stations used (for interpolated data sets), assimilated variables (for atmospheric reanalysis), and temporal and spatial resolution are given. Number of stations taken from Schneider et al. (2014). CRU TS3.1 = Climatic Research Unit Timeseries version 3.1; UDEL = University of Delaware; GPCC = Global Precipitation Climatology Centre; 20CR = 20th Century Reanalysis; ERA-20C = ECMWF twentieth century reanalysis; CHCN v2 = Global Historical Climate Network version 2.

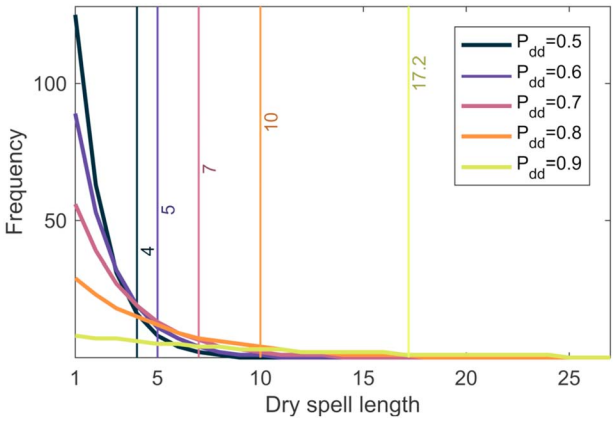

Figure 1. Dependence of the dry spell length frequency distribution on different P_{dd} values with constant P_{ww} (0.5). Vertical lines indicate the 90th percentile of the distributions for different P_{dd} values. The unit of *x* axis (dry spell lengths) is omitted as it varies depending on the temporal scale, for example, months for monthly time series and years for annual time series.

month-to-month and year-to-year transitions. In addition to the dry-to-dry transition probability P_{dd}, also dry-to-wet (P_{dw}), wet-to-wet (P_{ww}), and wet-to-dry (P_{wd}) transition probabilities can be estimated from the time series. The transition probabilities are related to each other such that $P_{dd} = 1 - P_{dw}$ and $P_{ww} = 1 - P_{wd}$, implying that the dynamics of the binary series can be captured by two parameters only. Under a first-order Markov chain assumption, these transition probabilities can then be used to simulate stochastic ensembles resembling the original binary time series, which in turn can be used to generate estimates of the dry spell length distribution. The statistical simulation of a Markov chain is conducted by first taking a random initial status (either dry or wet) of which subsequent statuses are determined based on the prior status and the corresponding transition probabilities derived from the original time series. The process is repeated until the length of constructed binary time series equals the length of the original time series. To facilitate the interpretation of the persistence metric (P_{dd}), Figure 1 shows how the drought length distribution depends on the value of P_{dd}. For constructing this figure, 1,000 time steps were considered for the statistical simulation of binary time series with a constant P_{ww} value of 0.5. For each P_{dd} value, the statistical simulation was repeated 1,000 times and median of the simulated samples was taken as the best estimates of dry spell length distributions. Dry spell lengths at 90th of each distribution are also presented, which increase with increasing P_{dd}.

Figure 2 shows an example of an annual precipitation anomaly time series at a grid cell in Eastern Europe from the Climatic Research Unit Timeseries version 3.1 data together with the dry spell length distribution derived from the time series with the uncertainty range obtained by statistical simulation. P_{dd} and P_{ww} estimated in this grid cell are also provided for better understanding of the conversion between Markov chain parameters and dry spell length distribution. More examples in different regions are presented in Figure S1 (supporting information).

To assess whether the Markov chain model and thus also the dry-to-dry transition probability (P_{dd}) later used for model validation are reasonable approximations for observed and modeled dry spell length distributions, a two sample Kolmogorov-Smirnov (KS) test was conducted to compare dry spell length distributions derived from the original time series to those derived from statistical simulation. For doing so, the previously described statistical simulation was repeated 1,000 times and the median of the bootstrapped distribution (black line in Figure 2b; gray shading corresponds to the range spanned by the individual simulations) compared with the observed or modeled dry spell length distributions using the KS test with a 0.05 significance level. Failure to reject the null hypothesis of equal distributions indicates that the Markov chain (and thus P_{dd}) is a reasonable approximation of the observed and modeled dry spell length distributions. Figure 3 shows the

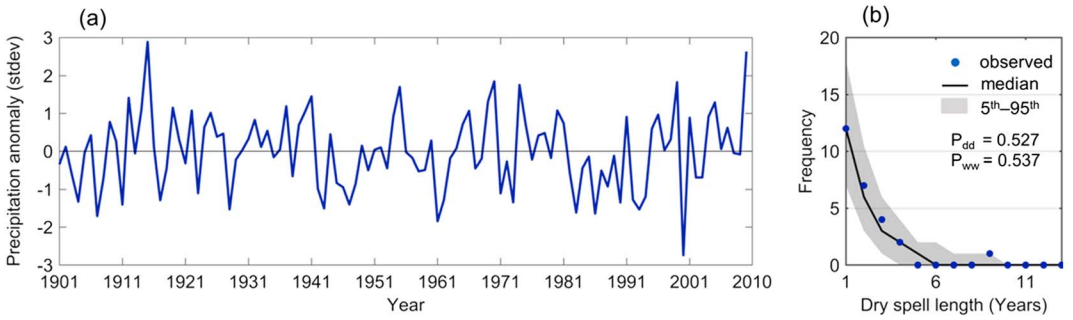

Figure 2. (a) Precipitation anomaly for 1901–2010 at a randomly chosen grid cell in Eastern Europe. The time series is taken from the Climatic Research Unit Timeseries version 3.1 data set. (b) Dry spell length distribution derived from the observed time series and median and 5th to 95th uncertainty range of 300 dry spell length distributions from statistical simulation of Markov chain.

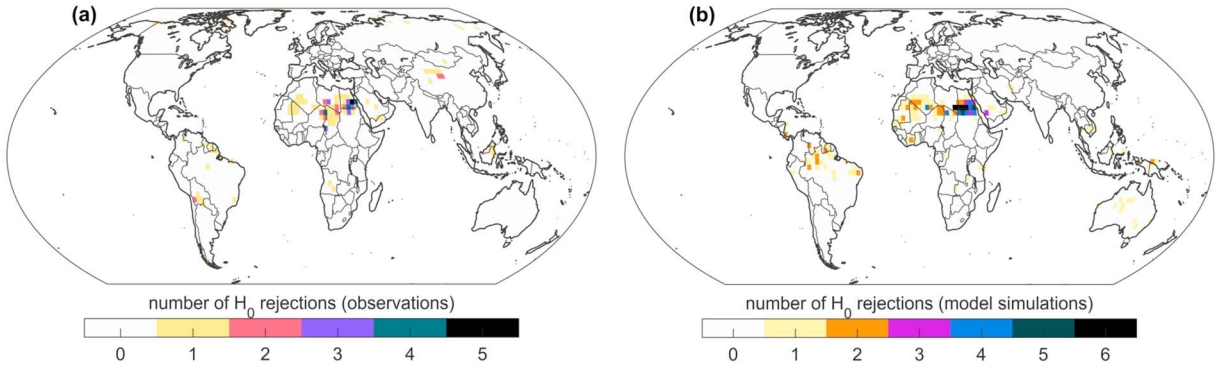

Figure 3. Assessment of the suitability of the Markov-chain assumption for approximating observed (a) and modeled (b) dry spell length distributions. Shown are the number of cases at which a two-sample Kolmogorov-Smirnov test rejects the null hypothesis that the observed (a) or modeled (b) dry spell length distribution can be approximated by a first-order Markov chain.

total number of cases at which the KS test rejects the null hypothesis that observed and modeled dry spell length distributions can be approximated with a Markov chain model at the monthly time scale. For observation-based data sets, there are no cases of rejection in most of the regions while a few grid cells located in Sahara, Northeast Brazil, and Southeast Asia have one to five cases of rejection out of five observation-based data sets indicating that the dry spell length distribution is not well presented with Markov chain model in the respective regions. For model simulations, the spatial distribution of rejections is similar to the observation-based estimates, with one to six cases out of 63 model simulations in the same region, which in this case confirms the reliability of the Markov chain assumption. At the annual time scale, there were no cases of rejection for model simulations and less than 10 grid cells had one to two rejections for observations (thus not shown in the figures).

Uncertainty in P_{dd} estimation is related to the finite number of samples in the observed and modeled time series at both time scales, and it is critical to assess whether this uncertainty substantially affects the model validation results. To quantitatively assess the uncertainty due to finite sample size, which also can be referred to as statistical estimation error, we generate the uncertainty distribution of P_{dd} using parametric bootstrapping (Basawa et al., 1990; Efron & Tibshirani, 1994). In each grid cell, 1,000 samples of binary time series are generated through statistical simulation under Markov chain assumption (described in the previous paragraph of this chapter) based on the P_{dd} and P_{wd} parameters estimated from the original time series. P_{dd} is then calculated for each statistical ensemble member, resulting in the uncertainty distribution for each P_{dd} estimate from the observational data and the ensemble runs of the GCMs. The *b*th bootstrapped sample of P_{dd} from observation *o* can be denoted as $P_{dd\ ob}$ and from ensemble *r* in model *m* as $P_{dd\ mrb}$. We compute the drought persistence error of ensemble *r* in model *m* against observation *o* as

$$E_{omrb} = P_{dd\ mrb} - P_{dd\ ob} \tag{1}$$

so that the uncertainty range of the estimated error can be spanned by the differences between bootstrapped samples from the observations and the models.

3.2. Partitioning the Spread of the P_{dd} Error

To investigate the respective roles of the sources of uncertainty on the drought persistence error, we aim at partitioning the spread of E_{omrb}. For doing so, we adapt a methodology that has been previously used to partition the spread in climate model projections (Hawkins & Sutton, 2009; Orlowsky & Seneviratne, 2013), with modifications that allow to account for observation uncertainty and the statistical estimation error of P_{dd}. Using an ANOVA-based approach, we aim to quantify the relative contribution of (i) differences among observational products, (ii) differences among climate models, (iii) internal climate variability, and (iv) statistical estimation uncertainty of P_{dd} to the total spread of E_{omrb}. As ensemble members are subordinate to the models and bootstrapped samples are subordinate to the ensemble members and observations (Figure 4), we apply a crossed and nested ANOVA (Krzywinski et al., 2014). A schematic of the partitioning is shown in

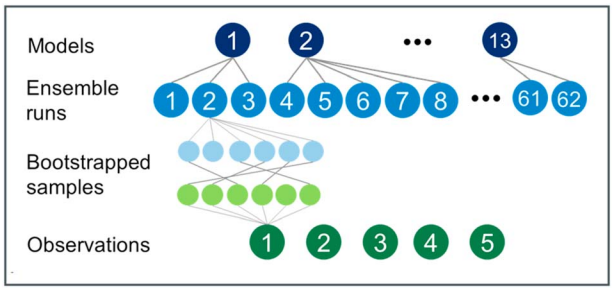

Figure 4. Hierarchical structure of drought persistence (P_{dd}) error.

Figure 5. This approach partitions the total sum of squares of E_{omrb} (SS_t) into model uncertainty (SS_m), observation uncertainty (SS_o), internal variability (SS_i), and sampling error (SS_s) as

$$SS_t = SS_o + SS_m + SS_i + SS_s = \sum_{o=1}^{N_o}\sum_{m=1}^{N_m}\sum_{r=1}^{R_m}\sum_{b=1}^{N_b}\left(E_{omrb}{}^2\right) - \text{cf} \quad (2)$$

where N_o is the number of observation data sets, 5, N_m is the number of models, 13, R_m is the number of ensemble members available for each model m, minimum 3 to maximum 10 (Table 1), and N_b is the number of bootstrapped samples per observation or model, 1000. We introduce the correction factor (cf) defined as

$$\text{cf} = \left(\sum_{o=1}^{N_o}\sum_{m=1}^{N_m}\sum_{r=1}^{R_m}\sum_{b=1}^{N_b} E_{omrb}\right)^2 / N_t \quad (3)$$

to correct the sum of squares of nonnested factors (Crawley, 2005) by subtracting it from the raw sum of squares of average error of each observation ($\overline{E_{o\cdots}{}^2}$) and model ($\overline{E_{\cdot m\cdot\cdot}{}^2}$) to adjust them as deviations from the total mean as

$$SS_o = N_s N_b \sum_{o=1}^{N_o} \overline{E_{o\cdots}{}^2} - \text{cf} \quad (4)$$

and

$$SS_{\text{mod}} = N_o N_b \sum_{m=1}^{N_m} R_m \overline{E_{\cdot m\cdot\cdot}{}^2} - \text{cf} \quad (5)$$

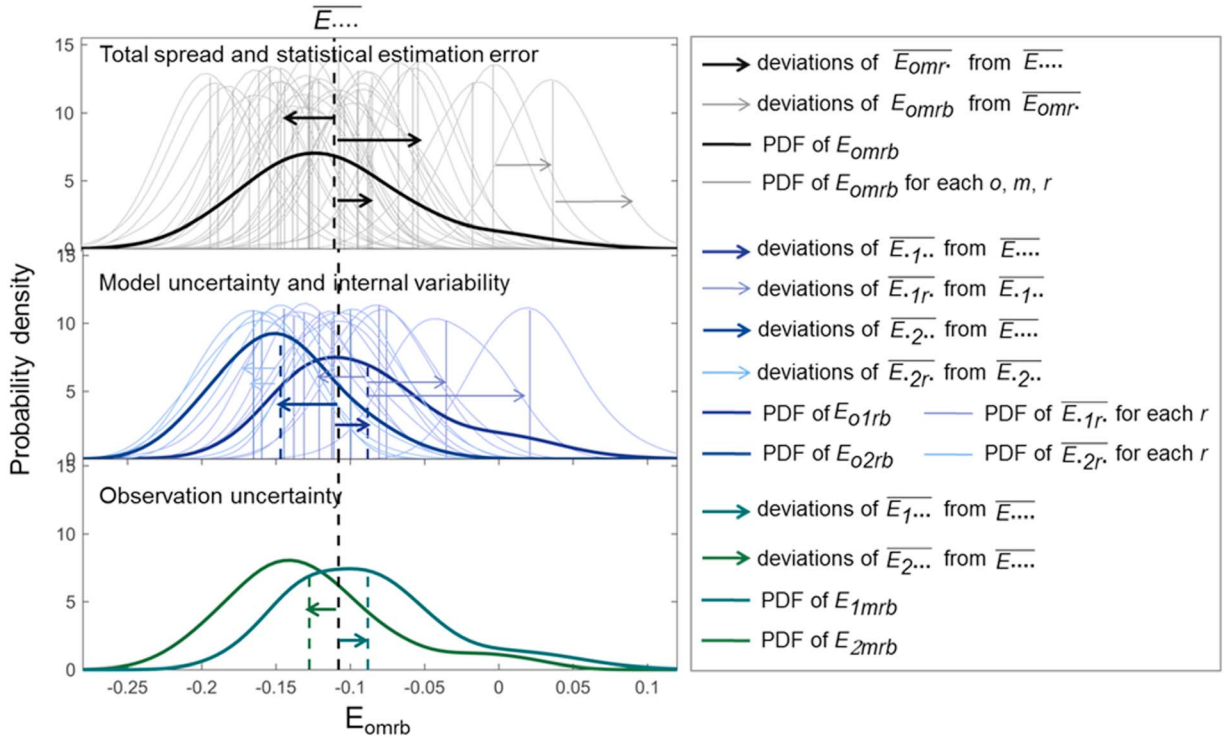

Figure 5. Schematic of partitioning the spread in the error, for the case $N_m = 2$, $N_o = 2$, $R_1 = 5$, $R_2 = 6$, and $N_b = 30$. See text for details and definitions.

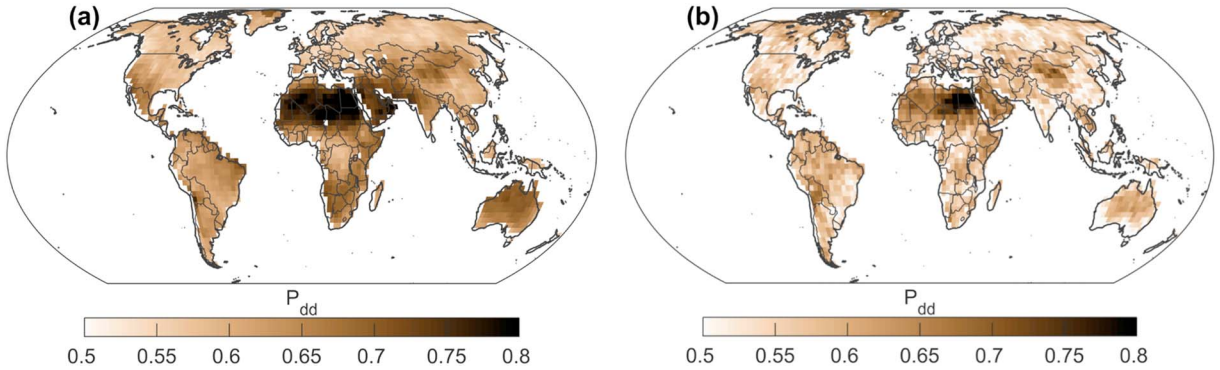

Figure 6. Mean drought persistence (P_{dd}) estimated from five observation-based data sets for 1901–2010 at (a) monthly and (b) annual scales.

where N_s is total number of model simulations, $\sum R_m$, and N_t is the total number of E_{omrb} calculated as $N_o N_b N_s$. Note that for SS_i, additional terms are subtracted from the raw sum of squares of average error of each ensemble simulation to account for the hierarchical structure of the problem.

$$\begin{aligned} SS_i &= N_o N_b \sum_{m=1}^{N_m} \sum_{r=1}^{R_m} \overline{E_{\cdot mr\cdot}}^2 - (SS_m + cf) \\ &= N_o N_b \sum_{m=1}^{N_m} \sum_{r=1}^{R_m} \overline{E_{\cdot mr\cdot}}^2 - N_o N_b \sum_{m=1}^{13} \overline{E_{\cdot m\cdot\cdot}}^2 \end{aligned} \tag{6}$$

SS_s is calculated as the sum of squared deviations of bootstrapped samples from average error of each ensemble simulation against each observation or as a residual sum of squares

$$SS_s = SS_t - N_b \sum_{o=1}^{N_o} \sum_{m=1}^{N_m} \sum_{r=1}^{R_m} \overline{E_{omr\cdot}}^2 = SS_t - SS_m - SS_o - SS_i \tag{7}$$

4. Results

4.1. Observed Global P_{dd} Patterns

Figure 5 shows the mean dry-to-dry transition probability (P_{dd}) over all observational data sets at each grid cell around the world for both monthly and annual time scales. At the monthly scale, global mean P_{dd} is 0.62 (with [0.58, 0.65] interquartile range) and at annual scale global mean P_{dd} is 0.57 (with [0.54, 0.60] interquartile range). The global means of observed maximum dry spell lengths are 19.1 months and 8.9 years for each time scale.

Monthly and annual P_{dd} values display clear different spatial patterns (Figure 6), indicating that monthly persistence does not necessarily propagate linearly to the yearly persistence. Some of the desert areas, for example, Sahara, Gobi, and Atacama, are commonly recognizable in both temporal scale with larger P_{dd} values. The results of KS test (Figure 3) also suggest higher uncertainty in P_{dd} estimation in Sahara.

4.2. Drought Persistence Error of CMIP5 Model Simulations

Figures 7a and 7c show the drought persistence error (E_{omrb}) averaged over all observational data sets and all model simulations. Overall, the models tend to underestimate P_{dd} at both monthly and annual scales, a feature that is particularly pronounced in western North America, western South America, Africa, and Northeast Asia. In around 40% of the land grid cells (38.9% at monthly scale and 37.9% at annual scale) at least 80% of the model simulations agree on the negative sign of E_{omrb}. For the positive sign, the same percentage of agreement was only found in 2.7% of the land grid cells at monthly scale and 1.5% at annual scale. Figures 6b and 6d show the global distribution of E_{omrb} and for each observational data set. For all considered observational data sets the tendency of the models to underestimate drought persistence is visible, but it should be noted that 20CR has stronger drought persistence at annual scale and suggests even larger negative drought persistence biases in the climate models than when using other reference data sets. It should be

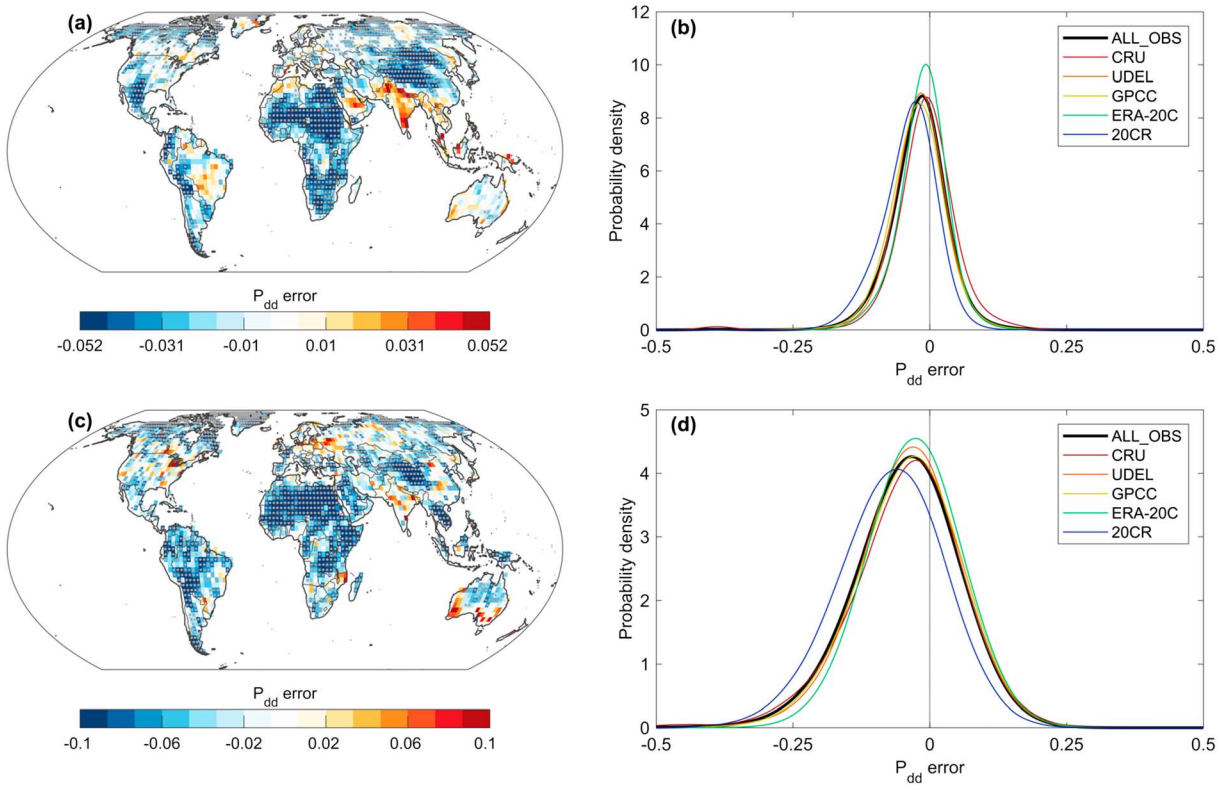

Figure 7. Spatial distribution of multi-simulation mean drought persistence error (E_{omrb}) for (a) monthly and (c) annual scales. Stippling indicates agreement in the sign of drought persistence error among more than 80% of the model simulations. Global distribution of drought persistence error for (b) monthly and (d) annual scales. CRU = Climatic Research Unit Timeseries version 3.1; UDEL = University of Delaware; GPCC = Global Precipitation Climatology Centre; ERA-20C = ECMWF twentieth century reanalysis; 20CR = twentieth century reanalysis.

noted for this result that 20CR relies on less direct observations than the other considered products (see Table 2). The "drizzle problem," where GCMs overestimate the number of days with light rainfall (Dai, 2006), might contribute to the larger biases found in desert regions.

Although the mean drought persistence bias indicates that the considered climate models tend to underestimate drought persistence, there are a few regions in which overestimation of P_{dd} is found. This is most pronounced at the monthly scale in the Arabian Peninsula and India, and there is a tendency for overestimation in the Mediterranean, southeastern South America, South Australia, and Southeast Asia. South Australia and India also show overestimation in annual P_{dd}.

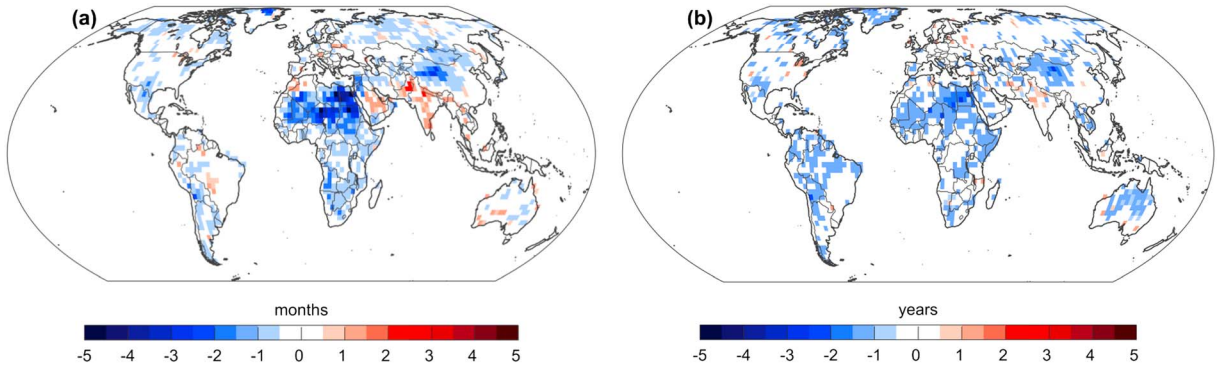

Figure 8. Spatial distribution of multi-simulation mean drought length bias against observed mean drought length at 95th percentile for (a) at monthly scale and (b) at annual scale.

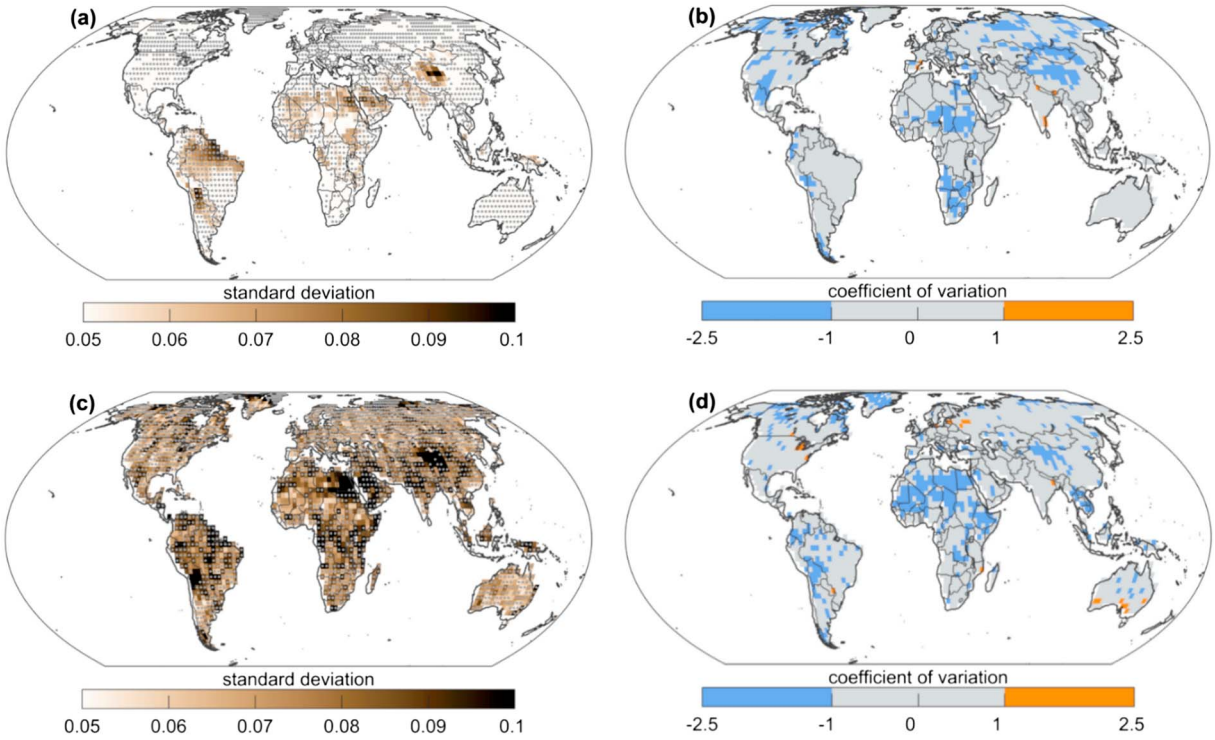

Figure 9. Standard deviation of drought persistence error at (a) monthly and (c) annual scales and grid cells at which the coefficient of variation E_{omrb} is larger or smaller than one for the (b) monthly and (d) annual scales. Stippling in (a) and (c) indicates regions with absolute value of coefficient of variation less than 1.

Additionally, most of these regions also show underestimation of wet persistence (P_{ww}) (supporting information Figure S2), which indicates that models are generally underestimating persistence of both dry and wet anomalies, at monthly and annual time scales.

Figure 8 shows the mean drought length bias at the 95th percentile of the dry spell length distribution, estimated on the multisimulation and observation mean P_{dd}. For doing so the observed and modeled dry spell length distributions were estimated on the basis of the multiobservation and multisimulation mean P_{dd} and P_{ww} values (using statistical simulation; see section 3). Subsequently, the difference between the observed and the modeled dry spell length distribution at the 95th percentile was calculated. The spatial distribution of drought length bias generally follows the pattern of the drought persistence bias in Figure 7. At monthly scale, the strongest underestimation in drought length (up to 5 months) is found in Sahara. Substantial overestimation of drought length in Arabian Peninsula and India is also noticeable. The spatial pattern at annual scale is rather homogenous over all regions. Considering the scale difference, drought length bias at annual scale is considerably smaller than monthly, unlike the larger persistence bias (Figure 7) which is related to the generally smaller values of P_{dd} at annual scale. Drought length biases at other percentiles are presented in Figure S3 (supporting information).

Figures 9a and 9c show the standard deviation of model error at monthly and annual scales. While the spread of E_{omrb} at monthly scale is only large in a few regions (e.g., Amazon, central Andes, Tibetan Plateau, and Sahara), most of the regions show substantial spread in E_{omrb} at annual scale. To compare the relative magnitude of spread against the magnitude of multisimulation mean bias, the coefficient of variation is calculated (Figures 9b and 9d). In many regions, the absolute value of the coefficient exceeds 1, which indicates that the standard deviation is larger than the mean of E_{omrb}. Nevertheless, pronounced underestimation can be confirmed at monthly scale, in western North America, central, and South Africa, and at annual scale, in Greenland, western South America, Tibetan Plateau, and mainland Southeast Asia, while there are only few grid cells with mean overestimation larger than the spread.

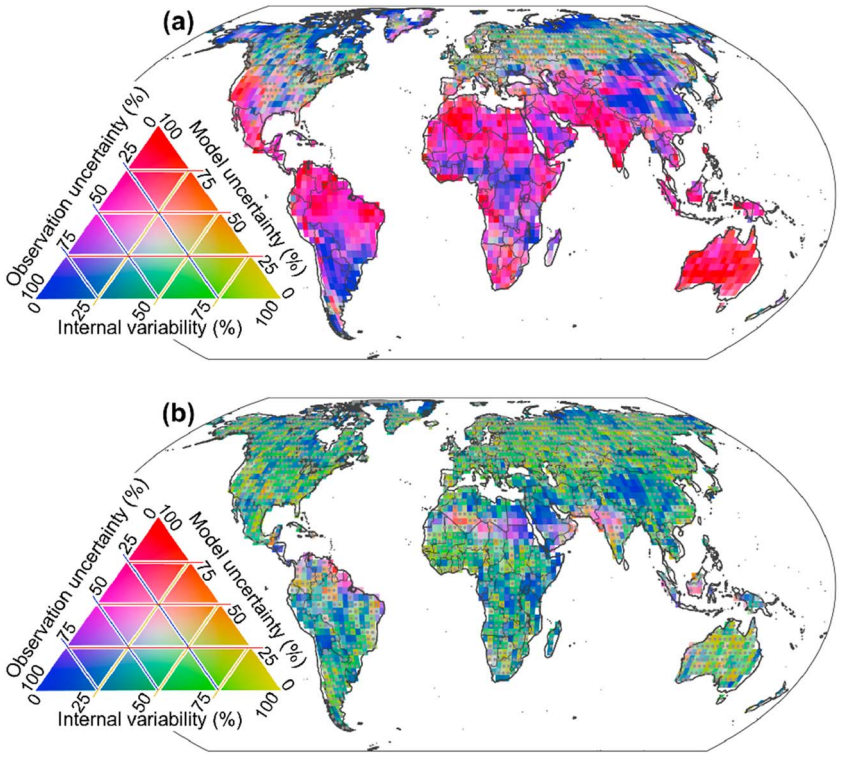

Figure 10. Partitioning the spread of the drought persistence error (E_{omrb}) into model uncertainty, observation uncertainty, and internal variability for (a) monthly and (b) annual scales. Stippled areas indicate that the summated contribution from the other sources is smaller than contribution from statistical estimation error.

4.3. ANOVA of Drought Persistence Error

As a substantial spread in the drought persistence error is found, we apply the ANOVA-based partitioning to quantitatively assess the relative contribution from the four possible sources of uncertainty to the total spread (see section 3.2). Figure 10 shows the spatial patterns determined by the relative contribution from, which are observation uncertainty, model uncertainty, and internal variability, where contribution from the statistical estimation error is only used for indicating reliability of the partitioning. Interestingly, the overall spatial patterns clearly differ depending on the considered time scale. The spread in monthly drought persistence error is mostly related to differences among the considered observational products and model uncertainty, except for high-latitude regions. Dominant contributions from model uncertainty are especially pronounced in western North America, Amazon, Northeast Brazil, Algeria, India, and Australia. Even some regions with scarce ground measurements, for example, Sahara, Amazon, and Australia, show similar or even higher contributions from model uncertainty to the spread in E_{omrb} at the monthly time scale, although a higher contribution of observation uncertainty might have been expected. In southern South America, central Africa, Arabian Peninsula, and East Asia, observation uncertainty is dominating to the total spread. Interestingly, a large spread among observation-based data sets for monthly precipitation in North Africa and Arabian Peninsula and other low-precipitation regions was reported in a previous study (Tanarhte et al., 2012). In contrast, the spread in annual drought persistence error is dominated by a combination from observation uncertainty and internal climate variability. In the Arabian

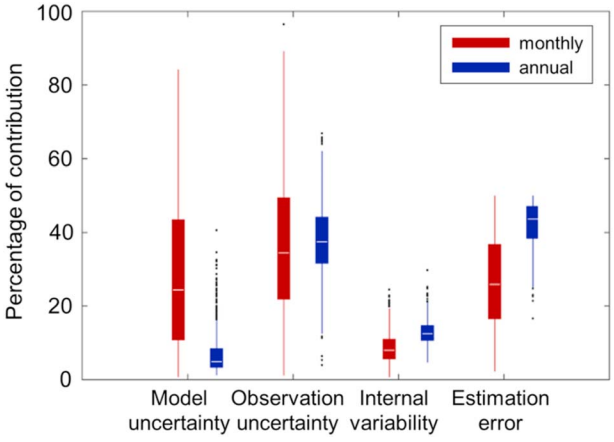

Figure 11. Box plots showing global distribution of contribution from each source of uncertainty to the total spread of drought persistence error for monthly (red) and annual (blue) scales. Each box represents the interquartile range, and the median is shown as white line. The whiskers show the range (25th quantile −1.5·IQR, 75th quantile +1.5·IQR), and values outside this range are indicated as dots.

Peninsula, Eastern Africa, and the Tibetan Plateau, observation uncertainty is more pronounced. Note, however, that for the annual scale, the statistical estimation error is larger than the sum of the relative contribution of all other sources of uncertainty at most locations, indicating that most of the spread (Figure 9c) is caused by finite sample effects.

Figure 11 shows the global distribution of the contributions from each source to the total spread in drought persistence error. At the monthly scale, overall contribution from observational and model uncertainty is clearly larger than that from internal variability. At the annual scale, the median of the contribution from statistical estimation error is larger than any other source, implying that the partitioning is unreliable in most of the regions (Figure 10b). In addition, a substantially lower contribution from model uncertainty is also noticeable at the annual scale.

5. Discussion

5.1. Drought Persistence Error

While previous studies have highlighted the underestimation of interannual to multidecadal precipitation variability in current climate models (Ault et al., 2012; Dai, 2006; Kumar et al., 2013), we investigated here whether state-of-the-art climate models capture meteorological drought persistence at shorter time scales. We find a systematic underestimation of the persistence of drought at both the monthly and annual scales in the considered CMIP5 models, when compared to five observational reference products over the 1901 to 2010 time frame. This underestimation is clear, despite the large spread in drought persistence error, which is in many regions attributable to observational uncertainty. Consequently, this study complements previous research (Ault et al., 2014; Wetter et al., 2014) by showing that the underestimation of drought persistence already occurs if year-to-year or month-to month variability is considered.

Overall, the climate models' ability to simulate precipitation variability at long time scales is likely related to the representation of tropical sea surface temperature and El Niño–Southern Oscillation variability in the GCMs, as previously suggested (Meinke et al., 2005; Schubert et al., 2016). Consequently, it is not unlikely that drought persistence in current climate models also depends on how well large-scale ocean-atmosphere variability is represented. On shorter time scales, the general underestimation of drought persistence shown in this study could possibly be attributed to issues in the representation of land-atmosphere couplings, such as the soil moisture-precipitation coupling at the regional scale (Guillod et al., 2015; Koster et al., 2006; Orlowsky & Seneviratne, 2010; C. M. Taylor, de Jeu, et al., 2012). Consequently, the differences in the spatial distribution of the drought persistence error at monthly and annual time scales identified in this study might be related to differences in the governing processes. However, to which degree this is the case and how errors on short time scales (e.g., day to day) translate to errors on long time scales (e.g., year to year) remains to be investigated.

5.2. Partitioning the Spread in the Drought Persistence Error

While the considered GCMs display a systematic underestimation of drought persistence, the range of underestimation spanned by these models and by the observational data sets is substantial. To further investigate this issue, we developed an ANOVA-based approach that allows to partition the contribution of unprecise observations, differences among climate models, internal climate variability, and statistical uncertainty to the total spread in the drought persistence error. The partitioning method applied in this study is motivated by an approach that is used to partition the uncertainty in future climate projections into model uncertainty, scenario uncertainty, and internal variability (Addor et al., 2014; Hawkins & Sutton, 2009; Orlowsky & Seneviratne, 2013; I. H. Taylor, Burke, et al., 2012). In this study, we expand this framework for the model validation task. To this end, we incorporated observation uncertainty together with the statistical estimation error of E_{omrb} into the existing framework. The analysis showed that both observational uncertainty and the statistical estimation uncertainty can play a significant role in climate model validation.

The results of the partitioning of the spread in the drought persistence error have interesting implications both for future model validation studies and for model development. On the one hand, the analysis highlights that differences among the considered climate models are in many cases larger than the uncertainty of observations, highlighting the potential of future model development. On the other hand, the analysis also shows that issues with the considered observational data products (especially in data scarce regions) as well

as finite sample effects (annual time scale) can impair the model validation exercise. Consequently, the results highlight the importance of considering the latter uncertainties in future model validation studies. Furthermore, the partitioned spread in this study might be used to assess the efficiency of model selection based on observational constraint for reducing the model spread in future projections (Cox et al., 2013; Hall & Qu, 2006). The efficiency is expected to be higher in the regions where the range of observation uncertainty is substantially narrower than the model uncertainty.

6. Conclusions

Persistent meteorological drought triggers subsequent drying of other hydrological variables on land (Van Loon, 2015) such as soil moisture or water storage, which can, for example, have strong impacts on vegetation (Nicolai-Shaw et al., 2017) and regional temperature (Seneviratne et al., 2010). In this study, we investigated how well current generation GCMs represent drought persistence over the twentieth century and the recent past. For doing so we focused on meteorological drought and compared dry-to-dry transition probabilities (P_{dd}) at yearly and monthly time scales of an ensemble of simulations from the CMIP5 archive with observational data sets (interpolated observations and reanalyses). Overall, the results highlight that the considered models tend to underestimate drought persistence at monthly and annual time scales. This is an interesting addition to previous results (Ault et al., 2014; Kumar et al., 2013), which indicate that CMIP5 models underestimate the risk of prolonged droughts on multidecadal time scales.

To investigate the substantial spread in the drought persistence error, we develop a new methodology to effectively partition the spread into its main components, which include (i) observation uncertainty, (ii) model uncertainty, (iii) internal variability, and (iv) statistical estimation error of the considered validation metric. At the monthly time scale, observation uncertainty (which is often neglected in model validation studies) and model uncertainty are the main contributors to the total spread. At annual time scales, the statistical estimation error is dominant in most regions, followed by combined contributions from internal variability and observational uncertainty. This analysis reveals regions where improvements of GCMs or model selection can substantially reduce the spread of model uncertainty. Reducing uncertainty of drought simulations is important for improving drought projections in a changing climate.

Acknowledgments
This research was funded by the ERC DROUGHT-HEAT project (contract 617518). We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank Jan Sedlacek and Urs Beyerle for the retrieval and preparation of CMIP5 model data. We also acknowledge the providers of precipitation data of the following products: GPCC (http://gpcc.dwd.de), CRU TS (https://crudata.uea.ac.uk/cru/data/hrg/), UDEL (http://climate.geog.udel.edu/~climate/), 20CR (https://www.esrl.noaa.gov/psd/data/20thC_Rean/), and ERA-20C (http://apps.ecmwf.int/datasets/).

References

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., & Seibert, J. (2014). Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. *Water Resources Research*, *50*, 7541–7562. https://doi.org/10.1002/2014WR015549

Ault, T. R., Cole, J. E., & St. George, S. (2012). The amplitude of decadal to multidecadal variability in precipitation simulated by state-of-the-art climate models. *Geophysical Research Letters*, *39*, L21705. https://doi.org/10.1029/2012GL053424

Ault, T. R., Cole, J. E., Overpeck, J. T., Pederson, G. T., & Meko, D. M. (2014). Assessing the risk of persistent drought using climate model simulations and paleoclimate data. *Journal of Climate*, *27*(20), 7529–7549. https://doi.org/10.1175/JCLI-D-12-00282.1

Basawa, I. V., Green, T. A., McCormick, W. P., & Taylor, R. L. (1990). Asymptotic bootstrap validity for finite Markov chains. *Communications in Statistics - Theory and Methods*, *19*(4), 1493–1510. https://doi.org/10.1080/03610929008830275

Bêche, L. A., Connors, P. G., Resh, V. H., & Merenlender, A. M. (2009). Resilience of fishes and invertebrates to prolonged drought in two California streams. *Ecography*, *32*(5), 778–788. https://doi.org/10.1111/j.1600-0587.2009.05612.x

Bunde, A., Büntgen, U., Ludescher, J., Luterbacher, J., & von Storch, H. (2013). Is there memory in precipitation? *Nature Climate Change*, *3*(3), 174–175. https://doi.org/10.1038/nclimate1830

Compo, G. P., Whitaker, J. S., & Sardeshmukh, P. D. (2006). Feasibility of a 100-year reanalysis using only surface pressure data. *Bulletin of the American Meteorological Society*, *87*(2), 175–190. https://doi.org/10.1175/BAMS-87-2-175

Cook, B. I., Ault, T. R., & Smerdon, J. E. (2015). Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Science Advances*, *1*(1), e1400082–e1400082. https://doi.org/10.1126/sciadv.1400082

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., & Luke, C. M. (2013). Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, *494*(7437), 341–344. https://doi.org/10.1038/nature11882

Crawley, M. J. (2005). Analysis of variance. In *Statistics: An introduction using R* (pp. 155–185). London, UK: John Wiley. https://doi.org/10.1002/9781119941750

Dai, A. (2006). Precipitation characteristics in eighteen coupled climate models. *Journal of Climate*, *19*(18), 4605–4630. https://doi.org/10.1175/JCLI3884.1

Dai, A. (2011). Drought under global warming: A review. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(1), 45–65. https://doi.org/10.1002/wcc.81

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootsrap*. New York: Chapman & Hall/CRC.

Griffin, D., & Anchukaitis, K. J. (2014). How unusual is the 2012–2014 California drought? *Geophysical Research Letters*, *41*, 9017–9023. https://doi.org/10.1002/2014GL062433

Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J., & Seneviratne, S. I. (2015). Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nature Communications*, *6*(1), 6443. https://doi.org/10.1038/ncomms7443

Hall, A., & Qu, X. (2006). Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophysical Research Letters*, *33*, L03502. https://doi.org/10.1029/2005GL025127

Harou, J. J., Medellín-Azuara, J., Zhu, T., Tanaka, S. K., Lund, J. R., Stine, S., et al. (2010). Economic consequences of optimized water management for a prolonged, severe drought in California. *Water Resources Research*, *46*, W05522. https://doi.org/10.1029/2008WR007681

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *International Journal of Climatology*, *34*(3), 623–642. https://doi.org/10.1002/joc.3711

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108. https://doi.org/10.1175/2009BAMS2607.1

Jackson, B. B. (1975). Markov mixture models for drought lengths. *Water Resources Research*, *11*(1), 64–74. https://doi.org/10.1029/WR011i001p00064

Kelley, C. P., Mohtadi, S., Cane, M. A., Seager, R., & Kushnir, Y. (2015). Climate change in the fertile crescent and implications of the recent Syrian drought. *Proceedings of the National Academy of Sciences*, *112*(11), 3241–3246. https://doi.org/10.1073/pnas.1421533112

Keyantash, J., & Dracup, J. A. (2002). The quantification of drought: An evaluation of drought indices. *Bulletin of the American Meteorological Society*, *83*(8), 1167–1180. https://doi.org/10.1175/1520-0477-83.8.1167

Koster, R. D., Sud, Y. C., Guo, Z., Dirmeyer, P. A., Bonan, G., Oleson, K. W., et al. (2006). GLACE: The global land–atmosphere coupling experiment. Part I: Overview. *Journal of Hydrometeorology*, *7*(4), 590–610. https://doi.org/10.1175/JHM510.1

Krzywinski, M., Altman, N., & Blainey, P. (2014). Points of significance: Nested designs. *Nature Methods*, *11*(10), 977–978. https://doi.org/10.1038/nmeth.3137

Kumar, S., Merwade, V., Kinter, J. L., & Niyogi, D. (2013). Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations. *Journal of Climate*, *26*(12), 4168–4185. https://doi.org/10.1175/JCLI-D-12-00259.1

Lee, K. S., Sadeghipour, J., & Dracup, J. A. (1986). An approach for frequency analysis of multiyear drought durations. *Water Resources Research*, *22*(5), 655–662. https://doi.org/10.1029/WR022i005p00655

Mandelbrot, B. B., & Wallis, J. R. (1969). Some long-run properties of geophysical records. *Water Resources Research*, *5*(2), 321–340. https://doi.org/10.1029/WR005i002p00321

Mathier, L., Perreault, L., Bobée, B., & Ashkar, F. (1992). The use of geometric and gamma-related distributions for frequency analysis of water deficit. *Stochastic Hydrology and Hydraulics*, *6*(4), 239–254. https://doi.org/10.1007/BF01581619

Matsuura, K., & Willmott, C. J. (2012). Terrestrial precipitation: 1900–2010 gridded monthly time series. Retrieved from http://climate.geog.udel.edu/~climate/html_pages/Global2011/README.GlobalTsP2011.html

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th conference on applied climatology* (Vol. 17, pp. 179–183). MA: American Meteorological Society Boston.

Meinke, H., DeVoil, P., Hammer, G. L., Power, S., Allan, R., Stone, R. C., et al. (2005). Rainfall variability at decadal and longer time scales: Signal or noise? *Journal of Climate*, *18*(1), 89–96. https://doi.org/10.1175/JCLI-3263.1

Nicolai-Shaw, N., Zscheischler, J., Hirschi, M., Gudmundsson, L., & Seneviratne, S. I. (2017). A drought event composite analysis using satellite remote-sensing based soil moisture. *Remote Sensing of Environment*, *203*, 216–225. https://doi.org/10.1016/j.rse.2017.06.014

Orlowsky, B., & Seneviratne, S. I. (2010). Statistical analyses of land–atmosphere feedbacks and their possible pitfalls. *Journal of Climate*, *23*(14), 3918–3932. https://doi.org/10.1175/2010JCLI3366.1

Orlowsky, B., & Seneviratne, S. I. (2013). Elusive drought: Uncertainty in observed trends and short- and long-term CMIP5 projections. *Hydrology and Earth System Sciences*, *17*(5), 1765–1781. https://doi.org/10.5194/hess-17-1765-2013

Pelletier, J. D., & Turcotte, D. L. (1997). Long-range persistence in climatological and hydrological time series: Analysis, modeling and application to drought hazard assessment. *Journal of Hydrology*, *203*(1–4), 198–208. https://doi.org/10.1016/S0022-1694(97)00102-9

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., et al. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083–4097. https://doi.org/10.1175/JCLI-D-15-0556.1

Quiring, S. M., & Papakryiakou, T. N. (2003). An evaluation of agricultural drought indices for the Canadian prairies. *Agricultural and Forest Meteorology*, *118*(1-2), 49–62. https://doi.org/10.1016/S0168-1923(03)00072-8

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014). GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, *115*(1–2), 15–40. https://doi.org/10.1007/s00704-013-0860-x

Schubert, S. D., Stewart, R. E., Wang, H., Barlow, M., Berbery, E. H., Cai, W., et al. (2016). Global meteorological drought: A synthesis of current understanding with a focus on SST drivers of precipitation deficits. *Journal of Climate*, *29*(11), 3989–4019. https://doi.org/10.1175/JCLI-D-15-0452.1

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3–4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Sericola, B. (2013). Discrete-time Markov chains. In *Markov chains: Theory and applications* (pp. 76–87). London, Great Britain: ISTE Ltd and John Wiley. https://doi.org/10.1002/9781118731543.ch1

Shanahan, T. M., Overpeck, J. T., Anchukaitis, K. J., Beck, J. W., Cole, J. E., Dettman, D. L., et al. (2009). Atlantic forcing of persistent drought in West Africa. *Science*, *324*(5925), 377–380. https://doi.org/10.1126/science.1166352

Sharma, T. C., & Panu, U. S. (2014). Modeling of hydrological drought durations and magnitudes: Experiences on Canadian streamflows. *Journal of Hydrology: Regional Studies*, *1*, 92–106. https://doi.org/10.1016/j.ejrh.2014.06.006

Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F., & Stahl, K. (2015). Candidate distributions for climatological drought indices (SPI and SPEI). *International Journal of Climatology*, *35*(13), 4027–4040. https://doi.org/10.1002/joc.4267

Steinemann, A. (2014). Drought information for improving preparedness in the Western States. *Bulletin of the American Meteorological Society*, *95*(6), 843–847. https://doi.org/10.1175/BAMS-D-13-00067.1

Stine, S. (1994). Extreme and persistent drought in California and Patagonia during mediaeval time. *Nature*, *369*, 546–549. https://doi.org/10.1038/369546a0

Tallaksen, L. M., & van Lanen, H. A. J. (2004). *Hydrological drought: Processes and estimation methods for streamflow and groundwater*. Amsterdam, Netherlands: Elsevier.

Tanarhte, M., Hadjinicolaou, P., & Lelieveld, J. (2012). Intercomparison of temperature and precipitation data sets based on observations in the Mediterranean and the Middle East: Precipitation and temperature data sets. *Journal of Geophysical Research*, *117*, D12102. https://doi.org/10.1029/2011JD017293

Tatli, H. (2015). Detecting persistence of meteorological drought via the Hurst exponent. *Meteorological Applications*, *22*(4), 763–769. https://doi.org/10.1002/met.1519

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Taylor, C. M., de Jeu, R. A. M., Guichard, F., Harris, P. P., & Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *Nature*, *489*(7416), 423–426. https://doi.org/10.1038/nature11377

Taylor, I. H., Burke, E., McColl, L., Falloon, P., Harris, G. R., & McNeall, D. (2012). Contributions to uncertainty in projections of future drought under climate change scenarios. *Hydrology and Earth System Sciences Discussions*, *9*(11), 12,613–12,653. https://doi.org/10.5194/hessd-9-12613-2012

Van Loon, A. F. (2015). Hydrological drought explained: Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, *2*(4), 359–392. https://doi.org/10.1002/wat2.1085

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2009). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Wetter, O., Pfister, C., Werner, J. P., Zorita, E., Wagner, S., Seneviratne, S. I., et al. (2014). The year-long unprecedented European heat and drought of 1540—A worst case. *Climatic Change*, *125*(3–4), 349–363. https://doi.org/10.1007/s10584-014-1184-2

Wilby, R. L., Prudhomme, C., Parry, S., & Muchan, K. G. L. (2015). Persistence of hydrometeorological droughts in the United Kingdom: A regional analysis of multi-season rainfall and river flow anomalies. *Journal of Extreme Events*, *02*(02), 1550006. https://doi.org/10.1142/S2345737615500062

Xie, P., Rudolf, B., Schneider, U., & Arkin, P. A. (1996). Gauge-based monthly analysis of global land precipitation from 1971 to 1994. *Journal of Geophysical Research*, *101*(D14), 19,023–19,034. https://doi.org/10.1029/96JD01553