Hydrological Drought Simulations: How Climate and Model Structure Control Parameter Sensitivity

L. A. Melsen and B. Guse

1Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, Netherlands, 2GFZ German Research Centre for Geosciences, Potsdam, Germany, 3Institute of Natural Resource Conservation, Department of Hydrology and Water Resources Management, Christian-Albrechts-University of Kiel, Kiel, Germany

Abstract Hydrological drought, defined as below-average streamflow conditions, can be triggered by different mechanisms, which are to a large extent dictated by the climate. Moreover, the simulation of hydrological droughts highly depends on the model structure and how drought-triggering mechanisms are parameterized. In this large-sample hydrological study, we investigate how climate and model structure control hydrological drought simulations. We conducted sensitivity analysis on parameters of three frequently used hydrological models (HBV, SAC, and VIC) for the simulation of drought duration and drought deficit over 605 basins covering more than 10 different Köppen-Geiger climates. The sensitivity analysis revealed that, as anticipated, different parameters are sensitive in different climates. However, not all expected drought mechanisms were reflected in the parameter sensitivity: Especially, the sensitivity of ET parameters does not align with the theory, and the role of snow parameters in snow-related droughts shows a distinction between degree-day-based models and energy-balance models. Besides parameter sensitivity being different over climates, we also found that parameter sensitivity differed over the different models. Where HBV and SAC did display fairly similar behavior, in VIC other model mechanisms were triggered. This implies that conclusions on driving mechanisms in hydrological drought cannot be based on hydrological models only, as different models would lead to different conclusions. Hydrological models can have heuristic value in drought research, to formulate new theories and identify research directions, but formulated theories on driving processes should always be backed up by observations.

1. Introduction

Hydrological droughts can be defined as phases of below-average streamflow (Tallaksen & Van Lanen, 2004). Anomalies in streamflow can negatively impact ecosystems and societies that adapted their water demand to (near) average water conditions, also if this anomaly occurs in the high-flow season (Tijdeman et al., 2018; Wilhite et al., 2007). Since precipitation and temperature patterns are changing (e.g., Fontrodona Bach et al., 2018) and expected to change further in the future (Knutti & Sedláček, 2013), insights into the simulation of hydrological droughts is of ample importance.

Drought is often characterized in terms of drought duration, the duration of below average streamflow (although often a lower threshold than “average” is defined), and drought deficit, the integrated difference between the defined threshold and the actual streamflow (Dracup et al., 1980). Van Lanen et al. (2013) and Van Loon and Laaha (2015) demonstrate that climate controls these two drought characteristics. Furthermore, different processes can trigger negative anomalies in streamflow, for instance extended periods without rainfall can propagate into a hydrological drought, or an anomaly in temperature can delay the snow melt process, subsequently causing a hydrological drought (Van Loon & Van Lanen, 2012). Different drought types prevail in different climates. Therefore, climate plays an important role both in drought-triggering mechanisms and drought characteristics.

For the simulation of drought characteristics, model structure comes in particular into play: The way in which the model structure is conceptualized controls the drought simulations (Melsen et al., 2019; Pfannerstill et al., 2014; Staudinger et al., 2011). There is a large range of hydrological models available, each with a different structure, which might influence the way in which hydrological droughts are simulated. However, representative observations of states and fluxes that control hydrological drought, such as...
Figure 1. Model structure of HBV(TUW), SAC, and VIC. The sensitivity to median drought duration and median drought deficit is shown as gray color intensity for the line around (duration) and background of (deficit) the different parameters. The displayed sensitivity is the average for all 605 basins covering a wide range of hydro-climatological conditions. Therefore, parameter sensitivity for individual basins will generally deviate from this average. This figure was inspired by Figure 3 of Herman et al. (2013).

The amount of water in storage and the release of water from storage, are hardly available (Henn et al., 2018). Therefore, model selection and model development for drought simulation are not straightforward.

A well-known diagnostic approach in hydrology is the use of sensitivity analysis to detect which model parameters are impacting a defined hydrological target variable or signature (Guse et al., 2016; Pianosi et al., 2016; Razavi & Gupta, 2015). Comparing parameter sensitivity among different models provides insights into similarities and differences in model structure functioning (Herman et al., 2013). As such, sensitivity analysis is a useful tool to investigate model functioning when simulating drought duration and drought deficit. Provided that these target variables are controlled by climate, sensitivity analysis should be conducted over a range of climates.

The goal of this study is to identify the hydrological model mechanisms that influence the simulation of drought duration and drought deficit and to relate this to climate and model structure. Given that different drought types prevail in different climates and that drought duration and drought deficit are controlled by climate, our hypothesis is that climate is reflected in the sensitivity of parameters representing different mechanisms. However, question is how the differences caused by climate relate to differences caused by model structure. Therefore, we will explore three hydrological models that are frequently used for hydrological drought simulation and make use of a large-sample data set covering over 10 different Köppen-Geiger climates.

2. Methods

To identify the model mechanisms that influence the simulation of drought duration and drought deficit, we conducted sensitivity analysis on a wide range of parameters in three hydrological models that are frequently used in drought research. The analysis was applied to 605 basins in the contiguous United States (CONUS), covering a wide range of hydro-climates. As such, we were able to identify which parameters mainly influence drought duration and drought deficit over a range of models and climates and investigated if this could be related to drought-triggering mechanisms as described by Van Loon and Van Lanen (2012).

2.1. Models

We consider three hydrological models: the Sacramento Soil Moisture Accounting model combined with SNOW-17 (SAC; Newman et al., 2015), the Variable Infiltration Capacity model Version 4.1.2h (VIC; Liang...
Table 1
Selected Parameters, Their Classification, and Their Boundaries for the HBV Model (See Section 2)

| Name | Unit | LB | UB | Description |
|------|------|----|----|-------------|
| 1    | Tm   | °C | -3.0 | 3.0 | Temperature where melt starts |
| 2    | Ts   | °C | Tr-0.01 | Tr-3 | Temp. below which precipitation is snow |
| 3    | Tr   | °C | 0.0 | 3.0 | Temp. above which precipitation is rain |
| 4    | DDF  | mm °C^-1 day^-1 | 0.04 | 12 | Degree-day factor |
| 5    | SCF  | — | 0.1 | 5.0 | Snow correction factor |
| 6    | LP   | — | 0.0 | 1.0 | Evaporation reduction threshold |
| 7    | P-T  | — | 1.0 | 1.74 | Priestley-Taylor coefficient |
| 8    | FC   | mm | 0.0 | 2000 | Max soil moisture storage |
| 9    | BETA | — | 0.0 | 20 | Nonlinear shape coefficient of groundwater partitioning function based on soil moisture level |
| 10   | K0   | day | 0.0 | 2.0 | Storage coefficient of very fast response |
| 11   | K1   | day | 2.0 | 30 | Storage coefficient of fast response |
| 12   | L    | mm | 0.0 | 100 | Reservoir threshold |
| 13   | PERC | mm day^-1 | 0.0 | 100 | Percolation rate |
| 14   | K2   | day | 30 | 250 | Storage coefficient of slow response |
| 15   | BMAX | day | 0.0 | 30 | Max baseflow of low flows |

Note. The selected parameters are based on Parajka et al. (2007); the parameter boundaries have been widened based on Uhlenbrook et al. (1999) and Abebe et al. (2010). The Priestley-Taylor parameter is based on Lhomme (1997). The parameters have been classified into five categories: snow (blue), evapotranspiration (green), soil moisture and shallow layer (purple), percolation (orange), and deep layer (red); colors and classification are consistent with Figure 1. LB = lower boundary; UB = upper boundary.

et al., 1994), and the TUWmodel (Parajka et al., 2007) following the structure of Hydrologiska Byråns Vattenbalansavdelning Model (HBV; Bergström, 1976, 1992; Uhlenbrook et al., 1999). Figure 1 provides a simplified overview of the structure of these models. All three models are frequently used for drought simulations; see for instance Van Werkhoven et al. (2008) and Wang et al. (2008) for application of SAC; Andreadis and Lettenmaier (2006), Sheffield and Wood (2007), Mo (2008), Shukla and Wood (2008), and Wang et al. (2008) for the application of VIC; and Van Loon & Van Lanen (2012, 2013) and Van Loon et al. (2014) for the application of HBV.

All models were run in a lumped fashion (the whole catchment represented by a single cell) for the period 1980–2008 using a sample of parameters. The first 5 years have been omitted as spin-up (1980–1985), to diminish the effect of initial conditions. The sampled parameters per model are provided in Tables 1, 2, and 3, for HBV, SAC, and VIC, respectively. Below, a short description of each model is given. A more thorough description of the models can be found in Melsen et al. (2018).

The HBV model (Bergström, 1976, 1992; Seibert & Vis, 2012) originates in Sweden. The model distinguishes three routines: snow accumulation and melt (the snow routine), soil moisture (the soil routine), and response and river routing (response function and routing routine). The snow routine is based on a degree-day approach, where elevation distribution was not accounted for in this study. In the soil moisture routine, soil storage and groundwater drainage are determined. Drained water enters the response routine of the model, which consists of an upper, nonlinear fast responding reservoir (the shallow groundwater layer), and a lower, slow responding linear reservoir (the deep groundwater layer), where the latter represents baseflow. The left panel in Figure 1 shows a schematic overview of the model structure (note that we merged the upper groundwater layer with the soil moisture layer for interpretation). We employed the TUWmodel (Parajka et al., 2007), which follows the structure of HBV, but is implemented in R, easing the use for high performance computing.

The SAC model (Burnash et al., 1973; National Weather Service, 2002) originates in the United States, developed by the U.S. National Weather Service. Snow representation as included through SNOW-17 is a degree-day-based formulation. Comparable to the soil routine in HBV, SAC divides the soil into an upper and a lower zone for fast and slow response, respectively. In contrast to HBV, SAC distinguishes for both
### Table 2
**Selected Parameters and Their Boundaries for the SAC Model (See Section 2)**

| Name       | Unit       | LB   | UB   | Description                                      |
|------------|------------|------|------|-------------------------------------------------|
| MFMAX      | mm °C⁻¹ hr⁻¹ | 0.8  | 3.0  | Max melt factor                                 |
| MFMIN      | mm °C⁻¹ hr⁻¹ | 0.01 | 0.79 | Min melt factor                                 |
| UADJ       | km 6 hr⁻¹   | 0.01 | 0.40 | Wind adjustment factor for rain on snow         |
| SI         | mm         | 1.0  | 3500 | Snow water equivalent for 100% snow area        |
| SCF        | —          | 0.1  | 5.0  | Snow undercatch correction factor               |
| PXTEMP     | °C         | −3.0 | 3.0  | Temperature for rain/snow transition            |
| P-T        | —          | 1.0  | 1.74 | Priestley-Taylor coefficient                    |
| UZTWM      | mm         | 1.0  | 800  | Upper zone max storage of tension water         |
| UZFWM      | mm         | 1.0  | 800  | Upper zone max storage of free water            |
| UZK        | day⁻¹      | 0.1  | 0.7  | Upper zone free water lateral depletion rate    |
| ZPERC      | —          | 1.0  | 250  | Max percolation rate                            |
| REXP       | —          | 0.0  | 6.0  | Exponent of the percolation equation            |
| PFREE      | —          | 0.0  | 1.0  | Frac. percolating from upper to lower zone      |
| LZTWM      | mm         | 1.0  | 800  | Lower zone max storage of tension water         |
| LZFPM      | mm         | 1.0  | 800  | Lower zone max storage of free water            |
| LZFSM      | mm         | 1.0  | 1000 | Lower zone max storage of sec. free water       |
| LZPK       | day⁻¹      | 1⁻⁵ | 0.025| Lower zone prim. free water depletion rate       |
| LZSK       | day⁻¹      | 1⁻³ | 0.25 | Lower zone sec. free water depletion rate        |

*Note. The parameter boundaries are based on Newman et al. (2015); the Priestley-Taylor parameter has been adapted based on Lhomme (1997). The parameters have been classified into five categories: snow (**blue**), evapotranspiration (**green**), soil moisture and shallow layer (**purple**), percolation (**orange**), and deep layer (**red**); colors and classification are consistent with Figure 1. LB = lower boundary; UB = upper boundary.*

### Table 3
**Selected Parameters and Their Boundaries for the VIC Model (See Section 2)**

| Name       | Unit       | LB   | UB   | Description                                      |
|------------|------------|------|------|-------------------------------------------------|
| Tsmax      | °C         | 0.0  | 3.0  | Max temp. where snowfall can occur              |
| Tsmin      | °C         | Tmax-0.01 | Tmax-3.0  | Min temp. where rainfall can occur            |
| SR         | —          | 5 · 10⁻⁵ | 0.5  | Surface roughness of the snow pack              |
| NewAlb     | —          | 0.7  | 0.99 | New snow albedo                                 |
| albaa      | —          | 0.88 | 0.99 | Base in snow albedo function for accum.         |
| albeta     | —          | 0.66 | 0.98 | Base in snow albedo function for melt           |
| RZT1       | —          | 0.5  | 2    | Multipl. factor rootzone thickness layer 1      |
| RZT2       | —          | 0.5  | 2    | Multipl. factor rootzone thickness layer 2      |
| RZT3       | —          | 0.5  | 2    | Multipl. factor rootzone thickness layer 3      |
| Rmin       | —          | 0.1  | 10   | Multipl. factor min. stom. res. vegetation      |
| Bi         | —          | 10⁻⁵ | 0.4  | Infiltration shape parameter                    |
| Depth2     | m          | 0.1  | 3.0  | Depth of soil layer 2                          |
| Expt2      | —          | 4.0  | 30   | Exponent of the Brooks-Corey relation          |
| Ds         | —          | 10⁻⁴ | 1.0  | Frac. Dsmax nonlinear baseflow starts           |
| Dsmax      | mm day⁻¹   | 0.1  | 50   | Max velocity of the baseflow                   |
| Ws         | —          | 0.2  | 1.0  | Frac. Wsmax nonlinear baseflow starts           |
| Depth3     | m          | 0.1  | 3.0  | Depth of soil layer 3                          |

*Note. Based on Demaria et al. (2007), Chaney et al. (2015), and Melsen et al. (2016). Parameters 4–6 are usually hard-coded in the VIC model but were shown to be highly sensitive in Mendoza et al. (2015) and are therefore included in the sampling. The parameters have been classified into five categories: snow (**blue**), evapotranspiration (**green**), soil moisture and shallow layer (**purple**), percolation (**orange**), and deep layer (**red**); colors and classification are consistent with Figure 1. LB = lower boundary; UB = upper boundary.*
the upper and the lower zone two components: tension water that can only be removed from the soil by evapotranspiration and free water that can also drain to the lower zone or into baseflow.

The VIC model (Liang et al., 1994, 1996) is a land-surface model, which can solve both the water balance and the energy balance. Snow is represented through an energy-balance approach. The soil is divided into three layers. Evaporation only occurs in the top layer, while rooting depth determines which layers are available for transpiration. Percolation from the upper layer downward depends on the moisture level of the upper soil layer. The lowest layer delivers the baseflow.

2.2. Data and Catchment Classification

The 605 CONUS basins investigated in this study are a subset of the publicly available CAMELS data set of Newman et al. (2014, 2015) and Addor et al. (2017b, 2017a). This dataset contains daily U.S. Geological Survey (USGS) discharge observations for all basins (in this study only used to investigate the change in sensitivity over model performance) and basin-averaged daily forcing data based on Daymet (a product based on interpolated observations; see Thornton et al., 2012) for the period 1980–2008. The majority of the basins are small headwater catchments to minimize human impact. Catchment area ranges between 4 and 25,800 km² (median is 361 km²), elevations from 14 to 3,640 m (median is 454 m), and slopes from 0.05° to 14.3° (median is 1.5°). For more detailed information, we refer to the above mentioned studies.

To investigate how climate controls parameter sensitivity for hydrological drought simulations, all catchments have been classified into climate classes. Two different tracks were followed in classifying the catchments: the traditional climate classification as based on the Köppen-Geiger system (Köppen, 1936; Peel et al., 2007), resulting in clearly defined separate classes, and the Knoben classification (Knoben et al., 2018), which is based on hydrological regimes in response to climate, and results in gradual classes. Figure 2 provides an overview of the catchments for both classifications.

2.2.1. Köppen-Geiger Classification

The catchments were classified according to the Köppen-Geiger system (Köppen, 1936), based on Table 1 in Peel et al. (2007). In total, the 605 basins cover 14 different climate classes (Figure 2a). Only climate classes with six catchments or more were included in the analysis, which resulted in 10 different climate classes (4 climate classes, and the related 6 catchments, being excluded). A summary of the climate classes is provided in Table 4. Thirty percent of the catchments has a temperate climate with hot summers, meaning that the warmest month of the year has an average temperature above 22 °C and the coldest month a temperature between 0 and 18 °C. The second largest climate class, representing 26% of the catchments, has a cold climate with warm summers, implying the warmest month has an average temperature between 10 and 22 °C and the coldest month an average temperature below 0 °C. Furthermore, at least 4 months per year the monthly average temperature is above 10 °C (see Peel et al., 2007). The Köppen-Geiger classification was used to investigate changes in parameter sensitivity per climate class.

2.2.2. Knoben Classification

Knoben et al. (2018) developed a climate classification system specifically focussed on hydrological response. This classification of hydro-climates has a gradual scale and is based on three climate indices: the average aridity, seasonality of the average aridity, and the fraction precipitation falling as snow. Knoben et al. (2018)
Table 4

Short Description of the Climate Classes Considered in This Study, Based on Peel et al. (2007)

| Class | Subclass | Description                                      | [%] |
|-------|----------|--------------------------------------------------|-----|
| B     | Sk       | Arid cold steppe                                 | 1   |
| C     | sa       | Temperate with dry and hot summer                | 1   |
|       | sb       | Temperate with dry and warm summer               | 9   |
|       | fa       | Temperate without dry season, with hot summer    | 30  |
|       | fb       | Temperate without dry season, with warm summer   | 3   |
| D     | sb       | Cold, with dry and warm summer                   | 4   |
|       | sc       | Cold, with dry and cold summer                   | 5   |
|       | fa       | Cold, without dry season, with hot summer        | 15  |
|       | fb       | Cold, without dry season, with warm summer       | 26  |
|       | fc       | Cold, without dry season, with cold summer       | 5   |

Note: The last column indicates the percentage of catchments that fall within each climate class. This column adds up to 99% because the climate classes with less than six catchments were excluded from the analysis (1%).

demonstrated that, based on these three indices, hydrological response is better constrained compared to the Köppen-Geiger classification.

In order to compute the average aridity, first, the Thorntwaite’s Moisture Index $MI$ (Knoben et al., 2018; Willmott & Fedema, 1992) has to be determined:

$$MI(t) = \begin{cases} 1 - \frac{E_p(t)}{P(t)}, & \text{if } P(t) > E_p(t) \\ 0, & \text{if } P(t) = E_p(t) \\ \frac{P(t)}{E_p(t)} - 1, & \text{if } P(t) < E_p(t). \end{cases}$$  \hspace{1cm} (1)$$

$P(t)$ and $E_p(t)$ are mean monthly observations of precipitation and potential evapotranspiration, respectively. Both were obtained from the CAMELS data set (Addor et al., 2017b; Newman et al., 2015).

Based on the Moisture Index, average aridity $I_m$ is computed as

$$I_m = \frac{1}{12} \sum_{t=1}^{12} MI(t).$$  \hspace{1cm} (2)$$

The average aridity varies between $-1$ and $1$, with $-1$ indicating most arid conditions and $1$ indicating most humid conditions. Figure 2b shows the average aridity $I_m$ per catchment.

The seasonality in $I_m$, $I_{m,r}$ is defined as follows:

$$I_{m,r} = \max(MI(1, 2, ...12)) - \min(MI(1, 2, ...12)).$$  \hspace{1cm} (3)$$

$I_{m,r}$ varies between 0, indicating no intra-annual changes, and 2, indicating that the climate switches from fully arid to fully humid. Figure 2c shows the seasonality in average aridity $I_{m,r}$ per catchment.

The last index defined by Knoben et al. (2018) is the fraction of precipitation falling as snow, $f_s$, defined as

$$f_s = \frac{\sum_{t=1}^{12} P(T(t) \leq T_0)}{\sum_{t=1}^{12} P(t)},$$  \hspace{1cm} (4)$$

where $T(t)$ is mean monthly temperature and $T_0$ is a threshold temperature below which precipitation is assumed to fall as snow. The threshold temperature was set to $0 \, ^\circ C$, in line with Knoben et al. (2018). $f_s$ ranges between 0, no snowfall, and 1, all precipitation falling as snow. Figure 2d shows the $f_s$ per catchment.

These three indicators were used in combination with Multinomial Logistic Regression (section 2.4) to determine the change in parameter sensitivity over a gradual hydro-climatic scale.
2.3. Drought Identification and Sensitivity Analysis

An overview of the conducted procedure for the analysis is provided in Figure 3. First, for all three models the parameters were sampled using a space filling sampling strategy. Subsequently, all models were run with the complete sample for all basins over a 28-year period, of which the first 5 years were used as spin-up and omitted from the analysis. Then, for each individual model run, the median drought duration and median drought deficit were determined. Finally, sensitivity analysis was conducted by relating median drought duration and median drought deficit to the sampled parameters.

2.3.1. Parameter Sampling

The parameter sampling strategy relates to the sensitivity analysis method that was employed. Since the goal of this study is to investigate model functioning, a global sensitivity analysis method is preferred above a local method (Gupta & Razavi, 2018). We used the Distributed Evaluation of Local Sensitivity Analysis method (DELSA; Rakovec et al., 2014), a derivative based approach. Given the high number of models, parameters, and catchments evaluated in this study, it was decisive that DELSA is computationally cheaper than variance-based global methods such as Sobol’.

Consider a model $f$, a vector with parameter values $\theta$ with a length $k$, $k$ being the number of parameters, and model output $\psi$. In local sensitivity, the optimal parameter values $\theta_{opt}$ (where “optimal” can refer to any kind of measure that was defined, based on either model performance or another metric) are taken as base run, leading to the base model output $\psi_{base}$:

$$\psi_{base} = f(\theta_{opt}) = f(\theta_{opt, 1}, \theta_{opt, 2}, \theta_{opt, 3}, \ldots, \theta_{opt, k}).$$  \hspace{1cm} (5)

Subsequently, one parameter is perturbed, for example, with 1%, and the change in model output compared to the base run is evaluated:

$$\psi_{perturbed} = f(\theta_{opt, 1} \cdot 1.01, \theta_{opt, 2}, \theta_{opt, 3}, \ldots, \theta_{opt, k}).$$  \hspace{1cm} (6)

The change in $\psi$ between both runs compared to the change in the parameter represents the local partial derivative and is an indicator for sensitivity:

$$\frac{\partial \psi}{\partial \theta_1} \bigg|_{\theta_{opt}} = \frac{\psi_{base} - \psi_{perturbed}}{\theta_{opt, 1} - \theta_{opt, 1} \cdot 1.01}.$$  \hspace{1cm} (7)

DELSA applies local sensitivity analysis at several places throughout the parameter space, not only at $\theta_{opt}$, thereby mimicking global sensitivity analysis methods that also consider the full parameter space. Rakovec et al. (2014) compared DELSA results to the global variance-based Sobol’ results and demonstrated that DELSA can mimic the results of the computationally more demanding Sobol’ method with substantially fewer parameter samples.

Tables 1, 2, and 3 provide an overview of the parameters of each model that were selected for the sensitivity analysis in this study. The parameters were selected such that all the essential hydro-climatological processes are represented. The parameter boundaries have been determined to be consistent with literature, as mentioned in the table captions. First, using a quasi-random Sobol’ sequence (Sobol’, 1967), 100 base runs were sampled based on the provided parameter boundaries. For each of these base runs, local sensitivity
was determined: Each parameter was sampled 100 times by perturbation of the parameter value with 1% compared to the 100 base runs.

To test if 100 runs per parameter was sufficient to sample the parameter space, a Kolmogorov-Smirnov (KS) test was applied to the distribution of parameter sensitivities for 1,000 bootstrapped samples with a size of \( n = 50 \). With the KS test, it was evaluated whether the distribution in sensitivity, with respect to median drought duration and median drought deficit, significantly changed when 50 parameter samples were taken, compared to 100 (see also Figure SA.1 in the supporting information). The test revealed that only very rarely (i.e., in less than 0.00006% of the tested cases) the distribution significantly changed. From this we conclude that with 100 samples per parameter, we sufficiently capture parameter space.

Using 100 samples per parameter \( k \) results in a total of \( k \cdot 100 \) samples, plus 100 base-run samples per model—1,600 samples for HBV, 1,900 for SAC, and 1,800 for VIC. All models were run with the complete sample for all 605 basins, leading to a total of more than 3 million model runs.

### 2.3.2. Variable Threshold Method to Identify Drought Duration and Deficit

In our adopted definition, hydrological drought is a negative deviation from “normal” flow conditions. There are two frequently used methods to identify hydrological drought; the fixed threshold method (e.g., Apurv et al., 2017; Laaha et al., 2017) and the variable threshold method (e.g., Sheffield et al., 2009; Van Huijgevoort et al., 2012; Yevjevich, 1967). Both methods are based on a threshold level below which a hydrological drought is defined. In the fixed threshold method the same threshold is applied over the year, generally leading to drought identification in the low-flow season. The fixed threshold level is appropriate when focusing on human water consumption (Apurv et al., 2017) or to identify the absolute dry state of the system (Laaha et al., 2017). The variable threshold method accounts for seasonality and as such shows deviations compared to normal conditions for that time of the year. This is for instance relevant for hydropower production (Killingtveit & Engen, 2003). The variable threshold can be determined and applied based on different temporal intervals, for instance daily or monthly (Heudorfer & Stahl, 2016).

We decided to use the variable threshold method with a daily time step, in line with Hannaford et al. (2011), because our catchments cover a wide range of climates with differing seasonality. In catchments with high seasonality, the fixed threshold level would always indicate the low-flow season as a hydrological drought, while for catchments with low seasonality real deviations compared to normal would be identified, which hampers the comparison (see for instance also Sheffield et al., 2009). Furthermore, the variable threshold method indicates deviations from normal, which is consistent with the adopted drought definition (Tallaksen & Van Lanen, 2004) and drought typologies as defined by Van Loon et al. (2012). It should be stressed that the variable threshold method as drought definition should be regarded as conceptual rather than physical (Sheffield et al., 2009).

For each model run a variable threshold level was identified. For each day of the year the discharge at the 20th percentile was determined (i.e., this discharge is exceeded 80% of the time for that particular day of the year, therefore also often referred to as the Q80), based on 23 years of daily data. Subsequently, the variable threshold was determined by using a moving average of 30 days of the daily determined 20th percentiles, resulting in a smoothed variable threshold. Every time the simulated discharge dropped below the threshold, this period was classified as a drought. For each individual drought event, the duration (number of consecutive days that the discharge was below the threshold) and the deficit (integral of the difference between discharge and threshold level per drought event, in units mm) were determined. Finally, over the 23-year simulation period, the median drought duration and median drought deficit were determined and stored.

### 2.3.3. Sensitivity Analysis

Model output \( \psi \) as evaluated for the sensitivity analysis was median drought duration and median drought deficit. The local partial derivative (equation (7)) thus represents, for instance, the change in simulated median drought duration divided by the change in the parameter value. Finally, the relative contribution of each parameter to change in median drought duration was determined, leading to a value per parameter between 0 (not sensitive at all) and 1 (the only parameter influencing \( \psi \)). This means that for each parameter set \( \theta \), the total sensitivity adds up to 1. DELSA was applied to 100 different \( \theta \), leading to 100 different values between 0 and 1 for each parameter. We used the average sensitivity from the complete sample per basin as a measure of parameter sensitivity.

To ease the interpretation, all sampled parameters have been classified based on the mechanism they represent: evapotranspiration (ET) parameters, snow parameters, soil moisture and shallow layer parameters,
percolation parameters, and deep layer parameters (see Figure 1 and Tables 1, 2, and 3). HBV has two parameters that represent a shallow groundwater layer (K0 and K1), while a shallow groundwater layer is not explicitly represented in VIC and SAC. Because we wanted the parameter classes to be consistent among the models, we decided to merge the shallow groundwater layer with the soil moisture group, because both relate to relatively fast response. This keeps the deeper layer parameters comparable among the models.

To investigate to what extent the results changed when another threshold level for drought identification was used, also, parameter sensitivity based on the Q90 and the Q70 for a random subsample of 20 catchments was determined. Indeed, a different threshold level lead to slightly different parameter sensitivities. The results changed more for drought duration than for drought deficit and changed more for VIC than for SAC and HBV. However, overall, the differences were (very) small.

Note that the model runs were not constrained based on discharge observations; the parameter sensitivity was determined based on 100 runs per parameter, regardless of the performance of this run compared to observations. The reason we chose this approach is because the goal of our study is to investigate model functioning, rather than model identifiability. Our approach is in line with Gupta and Razavi (2018), who demonstrate that global sensitivity analysis using a \( \psi \) based on a performance metric is inherently incomplete and distorts the information of relative parameter importance in the model.

### 2.4. Multinomial Logistic Regression for the Gradual Climate Scale

Finally, we tested if parameter sensitivity varied over hydro-climate using Multinomial Logistic Regression (MLR). MLR is a regression model that is used to predict probabilities of possible outcomes of dependent categorical variables. In our case, the dependent categorical variable is parameter class (consisting of five categories). These classes are categorical rather than ordinal because there is no meaningful way of ordering them. The independent variable is the climate, represented with the three gradual hydro-climate indicators from the Knoben classification. With MLR, the probability for each parameter class to be most sensitive given the independent variable (climate) is determined. The advantage of using MLR is that it not only interpolates the probabilities over the climate indicator values represented in the data (see the histograms in Figure 2) but is also able to extrapolate beyond the values covered in the dataset. Therefore, with MLR we can predict the probability of parameter class sensitivity in regions where all precipitation is falling as snow, even though this does not occur in any of the catchments included in the data set.

To improve the power of the test, MLR was applied to the parameter classes (ET, snow, soil moisture, percolation, and deep layer) rather than to individual parameters. The probability was determined that one of the parameter classes was most sensitive for a specific climate indicator, given the 605 basins that covered a range of the three climate indicators. For instance, if in all catchments with an average aridity of \(-1\) the class of ET parameters was most sensitive, this class has a probability of 1 for that aridity value. This probability can decrease with increasing average aridity values, thereby, for example, increasing the probability of the deep layer parameters.

Furthermore, using MLR, it was tested if any of the parameter classes could significantly be distinguished from the reference category, that is, if the class significantly contributed to improve the prediction \( (p < 0.05) \). The deep layer class was defined as reference category, since stream flow in drought conditions often relies on the release of water from the deeper storage layers. A reference category is required in order to apply MRL, and changing the reference class, for example, to the shallow layer, does not impact the probability distributions. It does, however, slightly change the significance results because the reference class is by definition “significant.”

### 3. Results

In this section, we first briefly discuss the variability in drought indicators for the three models as a result of changes in parameters. Subsequently, we discuss which parameters were most sensitive for the three models, and second, we discuss parameter sensitivity averaged over the different Köppen-Geiger climate classes. Finally, we link parameter sensitivity to the gradual hydro-climate indicators.

#### 3.1. Variability in Drought Indicators

To put the results into perspective, we first briefly discuss the variability in drought duration and drought deficit in the three models over the parameter sample, as shown in Figure 4. When evaluating the mean
Figure 4. The distribution of mean (a and c) and Coefficient of Variation (COV; b and d) of median drought duration (a and b) and median drought deficit (c and d) over all 605 basins. The mean median drought duration and deficit are determined per basin: The median duration was obtained per parameter sample, and the mean was taken over the complete parameter sample.

median drought duration over the 605 basins (i.e., the mean of the median drought durations that have been obtained over all simulations per model per basin), HBV displays a much higher variability in drought duration among the 605 basins than VIC and SAC. For mean median drought deficit, the differences among the models are smaller. More interesting from a sensitivity analysis point of view is the Coefficient of Variation (COV), defined as the standard deviation of the median drought duration and deficit (based on at least 1,600 model runs) divided by the mean median drought duration and deficit, respectively. As such, the COV shows the variation in drought indices that can be obtained by adapting the parameters. For both drought duration and drought deficit, HBV displays the highest COV, even though for this model the lowest number of parameters were sampled. For drought duration, SAC and VIC have a comparable COV, while for drought deficit, the COV of SAC is higher than the COV of VIC. In summary, with HBV both drought duration and drought deficit vary highly over different parameter values, while VIC shows the lowest variation in drought indices as a result of adapted parameters. We find for our tested instances the highest impact of model parameters on drought indicators for the model with the lowest number of sampled parameters.

3.2. The Most Sensitive Parameters Over Different Models
In this section, we discuss the parameter sensitivity for the three models and briefly discuss the differences among the models. The structure of the three investigated models is shown in Figure 1, with mean parameter sensitivity indicated. We refer to the sensitivity over different climates in the next section.

3.2.1. HBV
In HBV, median drought duration is mainly determined by the max soil moisture storage (FC), followed by the snow correction factor (SCF) and the storage coefficient of the deeper layer (K2). Whereas the role of subsurface storage properties on drought duration can be understood—larger storage means that it takes longer for a drought to develop—the role of the snow correction factor (which corrects for uncertainty in observations, which can particularly be high for snow because of wind) is not directly evident. Our hypothesis is that SCF can augment individual precipitation events such that when the snow is melted, the drought is
ended. The most sensitive parameter controlling drought deficit in HBV is again the max soil moisture storage (FC), followed by the shape coefficient that determines the fraction of water that drains to the shallow groundwater layer as compared to what remains in the soil moisture layer (BETA) and the snow correction factor (SCF). There is a large overlap in sensitivity for both drought deficit and drought duration, which can be a result of the role that these parameters have in controlling water input and storage.

3.2.2. SAC
The two most sensitive parameters in SAC for drought duration are the snow undercatch correction factor (SCF) and the Priestley-Taylor coefficient (P-T) controlling potential evaporation. Also, the parameter that defines the maximum storage of tension water in the upper zone (UZTWM) and the lower zone primary depletion rate (LZPK) displayed high sensitivity. Comparable to HBV, there are parameters controlling fluxes in and out of the system (here, SCF and P-T) and a parameter that controls the storage of water (here, UZTWM). Like for HBV, also for SAC there is large overlap in most sensitive parameters for drought duration and drought deficit. The main sensitive parameter for drought deficit is the snow undercatch correction factor (SCF), followed by the maximum storage of tension water in the upper zone (UZTWM) and the lower zone primary depletion rate (LZPK).

3.2.3. VIC
In the VIC model, most sensitive for both drought duration and drought deficit are the parameter that describes the maximum velocity of baseflow (Dsmax) and the parameter that determines the fraction of DSmax where nonlinear baseflow starts (Ds). Both parameters control the amount of water leaving the system through the deeper layer (baseflow). Other sensitive parameters are the depth of layer 2 (Depth2), the stomatal resistance (Rmin), the exponent of the Brooks-Corey relation (Expt2), and the depth of layer 3 (Depth3). Both depth parameters control the amount of water that can be stored in the (deeper) soil, while the Expt2 parameter controls the distribution of water over these two layers.

For all three models, large overlap is found in sensitive parameters for drought duration and drought deficit. Furthermore, all three models show a combination of flux parameters, controlling the amount of water available in the system, storage parameters, controlling the release of water from the soil, and parameters controlling the distribution of water over the upper and the lower layers. However, some differences in triggered mechanisms among the models can already be identified; when visually inspecting Figure 1, it can be recognized that in HBV and SAC snow parameters and soil moisture and shallow layer parameters tend to be more sensitive, while in VIC especially the deeper layer parameters show high sensitivity. In the next section, we investigate parameter sensitivity over different climate classes.

3.3. The Most Sensitive Parameters Over Different Climates
Figure 5 shows a heatmap with parameter sensitivity for median drought deficit and median drought duration, for different models and climate classes.

3.3.1. HBV
For HBV, we see that for the colder climates, the median drought deficit and median drought duration are particularly influenced by snow parameters DDF and SCF. Both parameters control the amount of water entering the system: the degree-day factor DDF in terms of snow melt rate and the snow correction factor SCF in terms of how much water is available to enter the system. The sensitivity of the snow parameters decreases when moving to the temperate and arid climates.

An interesting exception within the colder climates is the Dfa climate, which is characterized by a hot summer (average T of the warmest month >22 °C). The parameter sensitivity for this climate mimics the sensitivity of the temperate climates with a hot summer, Csa and Cfa. In the climates with a hot summer, soil moisture and shallow layer parameters FC (max soil moisture storage) and BETA (nonlinear shape coefficient) are most sensitive. The role of soil moisture storage (FC) on drought deficit and drought duration can physically be understood; when there is lack of input, flow depends on water that is released from storage. When the storage reservoir is smaller, less water is available for release during periods of low input, which can eventually lead to a streamflow drought, until new water is added to the system. Larger storage, obtained by varying soil storage parameters, means that it takes longer before a drought develops (decreasing the median drought duration) and sustained outflow from storage can decrease median drought deficit.

In the temperate climates, especially for those with a warm rather than a hot summer (Csb, Cfb), we see an increased sensitivity for the reservoir threshold L, the percolation rate PERC, and the storage coefficient of
Figure 5. Parameter sensitivity versus Köppen-Geiger climates for the three different models, for median drought deficit (left) and median drought duration (right). The parameter labels have been colored according to their class, as shown in Figure 1.
the slow response, K2. This seems to indicate an increased role for the deeper layer storage (controlled by
the percolation toward the deeper layer) for drought duration and drought deficit in temperate climates.

3.3.2. SAC
In comparison to HBV, fewer parameters in SAC demonstrate high sensitivity. But there is also resemblance
with the results for HBV: Snow parameters dominate in colder climates, and soil moisture and shallow layer
parameters in the temperate and arid climates.

The most sensitive snow parameter in the colder climates is the snow undercatch correction factor SCF,
which has the same role as the SCF parameter in HBV: precipitation observation correction, thereby directly
controlling the amount of water available for snow storage. The role of melt rate, which was highly sensitive
in HBV, is less pronounced in SAC, perhaps because melt rate relates to two parameters in SAC: MFMAX
and MFMIN (which both do display a somewhat higher sensitivity).

Like for HBV, the parameter sensitivities in the cold climate with a hot summer (Dfa) resemble the param-
eter sensitivities in the temperate climates. Within the temperate climates, UZTWM, the upper zone max
storage of tension water, shows most sensitivity. This parameter controls storage in the upper zone, thereby
resembling the role of FC in HBV. Remarkable, however, is that only UZTWM shows high sensitivity, while
in theory the role of storage in drought duration and deficit could also be achieved by UZFWM, the upper
zone max storage of free water. Tension water can only be removed from the system by evaporation and
transpiration, while free water can drain to the deeper reservoir. The two temperate climates with a hot sum-
mer (Csa and Cfa) display higher sensitivity for the lower zone free water depletion rate LZPK indicating
an increased role for the deeper layer in temperate climates with hot summers—contrasting the results of
HBV, where this behavior was observed for temperate climates with warm summers.

The results for median drought deficit and median drought duration are very much alike in terms of pat-
tern, but sensitivities are generally lower for median drought duration, indicating a less clear role for all
parameters on variability in median drought duration. This is in line with the results presented in Figure 4,
showing a lower variability in drought duration compared to drought deficit for SAC.

3.3.3. VIC
Parameter sensitivity in the VIC model deviates from the sensitivities observed for SAC and HBV. Although
snow parameters show sensitivity in the colder climates, with declining sensitivity when moving toward
temperate and arid climates, the most sensitive parameters in almost all climates are Ds and Dsmax, both
related to the base flow from the deeper layer. Only in the arid climate (BSk), the evapotranspiration param-
eter related to the stomatal resistance, Rmin, displays more sensitivity. Also, in the temperate climates with
hot summers (Csa and Cfa), Rmin is more sensitive, and even the cold climate with a hot summer (Dfa)
does not show a slightly higher sensitivity to Rmin. Increased sensitivity to Rmin coincides with a higher sensitivity
in Depth2, the depth of soil layer 2, representing storage in the shallow layer.

Overall, VIC seems to show fewer differences in parameter sensitivity among the different climates. Subtle
differences exist, such as sensitivity of snow parameters in cold climates and the sensitivity of stomatal
resistance in arid and warm climate, but deep layer parameters dominate in almost all climates. Like for
SAC, for VIC the results for median drought deficit and median drought duration closely resemble each
other in terms of pattern, but generally lower sensitivities are found for median drought duration.

The results in this section indicate that climate controls parameter sensitivity: Different model mechanisms
are triggered over varying climates to simulate drought duration and drought deficit. The results, however,
also demonstrate that the triggered model mechanisms vary among model structures. All models display
higher sensitivity for snow parameters in the colder climates, but in some cases this is overwhelmed by the
sensitivity for other parameters (deep layer parameter in VIC). In the next section, we will further explore
climate and model structure controls.

3.4. Parameter Sensitivity Over Gradual Hydro-Climatic Indicators
In the following analysis, we focus on the average parameter sensitivity per parameter class, as indicated in
Tables 1, 2, and 3 and Figure 1: evapotranspiration parameters, snow parameters, soil moisture and shal-
low layer parameters, percolation parameters, and deep layer parameters. This provides a more complete
overview of the triggered model mechanisms than the single most sensitive parameter alone. Subsequently,
we applied Multinomial Logistic Regression (MLR) to obtain a predictive model of parameter sensitivity
over three gradual hydro-climatic indicators. Figure 6 and 7 depict an overview of the results.
Figure 6. Most sensitive parameter classes for drought duration, related to Knobens’ hydro-climate indicators (note that arid implies negative values in our definition of aridity; see equation (2)). Probability of a parameter class to be the most sensitive class for median drought duration (parameters classified according to Figure 1) based on all 605 catchments that were sorted for three different indicators: average aridity, seasonality of the average aridity, and fraction precipitation falling as snow. Probability is determined with Multinomial Logistic Regression (MLR), where deep layer parameters were taken as reference to which the other classes have been compared. Classes that have a significant ($p < 0.05$) contribution to the prediction are displayed in darker shades; the lighter shades indicate insignificance.

Figure 6 shows the results from the multinomial logistic regression for drought duration. In line with the results from the previous section, it is visible that parameter sensitivity varies over climate. Both HBV and SAC (although not significant in the latter case) depict increased sensitivity for soil moisture and shallow layer parameters with increasing aridity. This finding is in agreement with the results of Haslinger et al. (2014), which showed increasing importance of subsurface storage properties with drier conditions. For VIC the role of soil moisture and shallow layer parameters does not vary over aridity. With increasing wet conditions, both SAC and HBV display an increased role for ET, but for HBV that role is taken over by percolation parameters in the very wet regions. VIC shows deviant behavior; deep layer parameters display increasing sensitivity with increasing wetness. Also, for seasonality and fraction precipitation falling as snow, HBV and SAC display comparable behavior. ET parameters show high sensitivity for low seasonality and low fraction of precipitation falling as snow, and with increasing seasonality and fraction precipitation falling as snow, the sensitivity of snow parameters increases (although not always significant). In contrast, parameter sensitivity in VIC does not vary significantly over seasonality, while snow parameters become significantly more important in regions with a higher fraction precipitation falling as snow.

Figure 7 shows the multinomial logistic regression results for drought deficit. Also for drought deficit, parameter sensitivity varies over climate, although VIC shows less variation in parameter sensitivity over...
Figure 7. Most sensitive parameter classes for drought deficit, related to Knobens’ hydro-climate indicators (note that arid implies negative values in our definition of aridity; see equation (2)). Probability of a parameter class to be the most sensitive class for median drought deficit (parameters classified according to Figure 1) based on all 605 catchments that were sorted for three different indicators: average aridity, seasonality of the average aridity, and fraction precipitation falling as snow. Probability is determined with Multinomial Logistic Regression (MLR), where deep layer parameters were taken as reference to which the other classes have been compared. Classes that have a significant ($p < 0.05$) contribution to the prediction are displayed in darker shades; the lighter shades indicate insignificance.

The extent to which the different parameter classes can significantly be distinguished is an indicator of the effectiveness of the hydro-climatic indicator to predict sensitive parameter classes for hydrological drought simulations. In that perspective, aridity and seasonality seem more effective in predicting parameter sensitivity for drought duration and deficit than the fraction precipitation falling as snow. The three climate characteristics are most effective in the prediction of sensitive parameters for HBV, based on significance of the parameter classes. In contrast, parameter sensitivity in VIC is harder to predict given these climate characteristics. This can indicate that the model results are more constrained by the model structure in VIC compared to HBV. This is in line with the results of section 3.1, where HBV displayed the highest variability in drought duration and drought deficit over the parameter sample.
Overall, the results in this section demonstrated that both climate and model structure control parameter sensitivity. That parameter sensitivity depends on climate is in line with expectation, because different drought typologies will prevail in different climates. The variation in parameter sensitivity among the models, however, indicates that different models trigger different mechanisms to simulate drought duration and drought deficit.

4. Discussion

4.1. Methodological Implications

We presented an extensive sensitivity analysis study with drought duration and drought deficit as target variables. Some methodological choices were made that might influence the results of the study.

4.1.1. Sensitivity Analysis Methods

Ample sensitivity analysis methods are available (Pianosi et al., 2016). The goal of this study, investigating model functioning, indicates that global sensitivity analysis should be applied rather than local sensitivity analysis (Gupta & Razavi, 2018). Within the pool of global methods, we selected DELSA because it is computationally frugal. However, Pappenberger et al. (2008) compared several global sensitivity analysis methods and demonstrated that different sensitivity analysis methods will lead to different results. The presented parameter sensitivities are therefore specific for the setup and methods chosen in this study. Important, however, is that all models were treated in the same way, making the relative comparison among the models and over climates robust.

Another important point is that DELSA does not consider parameter interaction. Parameter interaction is likely to occur in the investigated models, given the high number of parameters involved. There are methods available where parameter interaction is accounted for, but they have a substantially higher computational demand. We evaluated the trade-off between accounting for interaction on the one hand and using a higher number of models, parameters, and catchments on the other hand and considered the latter more relevant for our research question.

4.1.2. Model Performance

Whereas local sensitivity analysis is generally applied around \( \theta_{opt} \), we applied local sensitivity analysis over the full parameters space. This is in line with global sensitivity analysis approaches, where the full range of parameters is evaluated. This does imply, however, that we also evaluated parameter sensitivity in regions with low model performance. An argument against this approach and the subsequent conclusions can be that the changes in sensitivities among the models might decline when the models are calibrated.

Because DELSA is a hybrid method, applying local sensitivity analysis globally, the local sensitivity determined for several points can be related to model performance of the run for that particular point. In this way, it can be tested if the sensitivities of the parameters change with model performance. Figures S2.1 to S2.6 show parameter sensitivity versus model performance. Model performance is expressed in the Nash-Sutcliffe Efficiency over the inverse discharge, NSE(1/Q), as suggested by Pushpalatha et al. (2012) to focus on the lower 20% of the discharge range. It can be seen that parameter sensitivity can vary over model performance. Sometimes, higher sensitivity is observed for the runs with the highest model performance, but this can be related to the relatively low number of runs that achieved this model performance. The opposite also occurs: lowest sensitivity for the runs with the highest model performance. In specific cases, increasing sensitivity with increasing model performance can be observed, for example, for an ET parameter in the Dfa climate for median drought deficit in the VIC model. Furthermore, it is visible that model performance is generally lower in the arid (Bsk) climates, as was also demonstrated by Melsen et al. (2018). Although these are valid arguments to put the presented results into perspective, we stress that our goal was to evaluate model functioning rather than model identification (Gupta & Razavi, 2018), and we therefore deliberately do not account for model performance in the analysis.

4.1.3. On the Choice for the Variable Threshold Method

We applied the variable threshold method to determine hydrological drought: first, because this is in line with the definition that hydrological drought is a streamflow anomaly, but also because it allows comparison among catchments with highly variable seasonalities. The use of the variable threshold, however, has an important caveat. An implication of this approach is that each day of the year is in drought about 20% of the time over the full simulation period. This challenges a process-based interpretation of the results, because different processes will trigger streamflow anomalies throughout the year. The results are thus,
as a consequence of the definition of the variable threshold, a mixed signal of all these anomaly-triggering processes. We evaluated median drought duration and median drought deficit, summary statistics of the drought time series. An important underlying hypothesis of our analysis is that each region has prevailing drought types, which have most influence on these summary statistics.

An alternative would be to use the fixed threshold method. With a fixed threshold, seasonal low flows are identified rather than streamflow anomalies. This means that droughts will mainly occur in the same period of the year throughout the simulation period, and probably, but not necessarily, fewer drought-triggering processes are involved. Problems with this approach in the context of our study is that this approach does not align with our adopted drought definition and the drought-triggering processes as described by Van Loon and Van Lanen (2012) and that it hampers comparison among catchments with variable climates and seasonalities (Sheffield et al., 2009). We argue that both methods have their advantages and disadvantages and that the most appropriate method depends on the adopted drought definition and the research question.

4.2. Parameter Sensitivity Versus Drought Type
Van Loon and Van Lanen (2012) identified five different hydrological drought types (besides composite droughts), based on the processes causing the drought. Probably, most catchments in this study experience several drought types and composite droughts (a combination of drought types) throughout the simulation period, but certain drought types will prevail in certain climates. In this section, we discuss the different drought types and if their mechanisms are reflected in the parameter sensitivity, based on Figure 5, 6, and 7. The three different snow-related drought types cannot be distinguished based on parameter sensitivity alone and are therefore discussed as a group.

**Classical rainfall deficit drought** is exclusively caused by an extended period with lack of rainfall. This drought type can occur in every climate. Subsurface storage properties determine how fast a classical rainfall deficit drought develops during a period of absent rainfall. In HBV, FC is the parameter that determines the amount of water in storage, and this parameter indeed shows high sensitivity over all climates. From all storage parameters in SAC, the UZTWM (upper zone max storage of tension water) also displays high sensitivity overall climates, but other storage parameters, such as the upper zone storage of free water, show less sensitivity. In VIC, the depth of both soil layers 2 and 3 displays sensitivity over the full climate range. Indeed, subsurface storage properties seem to play an important role in all climates.

**Wet-to-dry-season drought** starts with a period of rainfall deficit, followed by a dry season with a high potential evaporation. This drought type occurs in climates with clear precipitation seasonality (Van Loon & Van Lanen, 2012). The climates included in this study with a clear precipitation seasonality are Csa, Csb, Dsb, and Dsc. The mechanisms of this drought type are expected to be visible in evapotranspiration parameters. However, Figure 5 shows that the climates with a clear seasonality do not necessarily display (higher) sensitivity for evapotranspiration parameters. The seasonality displayed in Figures 6 and 7 is based on aridity rather than precipitation but also indicates yearly variation. These figures display the exact opposite of what would be expected based on the mechanisms of a wet-to-dry-season drought: ET parameters are mostly sensitive in climates with low seasonality (for HBV and SAC). Either this drought type does not occur frequently in the investigated catchments and therefore has no strong signal in the sensitivity or the mechanisms of this drought type are not mimicked in the simulation. Overall, it seems that in this case the model structure and thus the way in which evapotranspiration is regulated by the model parameters is a stronger factor. Hereby, it should be kept in mind that ET can be controlled in different ways in the models, such as with specific coefficients (HBV and SAC), by thresholds (HBV) or with specific factors for different layers (VIC).

**Rain-to-snow-season drought** starts with a period of rainfall deficit and continues into a period where precipitation does fall, but in the form of snow. Consequently, there is no replenishment of reservoirs until the snow starts to melt. The cold snow season drought is caused by below-average temperatures in cold and temperate climates, delaying snow melt thereby leading to a decrease in streamflow. The warm snow season drought is caused by high temperatures in the snow season, leading to a too early onset of snow melt and a lack of snow to melt in the "normal" snow melt period. All three drought types only occur in climates with a clear snow fall and melt season. Mechanisms of these droughts are expected to be visible in the parameter representing the threshold temperature above which snow starts to melt, or the threshold temperature that determines the transition from rain to snow. These parameters in all three models (Tm, Ts, and Tr for HBV, PXTEMP for SAC, and SR, Tmax, and Tsm for VIC) do display sensitivity in the colder climates, confirming that this mechanism is triggered. In VIC, the surface roughness of the snow
pack (SR) is the most sensitive snow parameter, which controls sensible and latent heat fluxes and as such can influence the onset of snow melt. Both in HBV and SAC, however, the snow correction factor (SCF) is the most sensitive snow parameter. This parameter simply controls the amount of water in storage in the system and therefore does not reflect the temperature-driven mechanisms of the snow-related droughts. It seems that this parameter has such a strong influence on drought simulation that it declines the role of the temperature-driven parameters.

Relating the parameter sensitivity to drought types, we recognize that some of the drought mechanisms as described in Van Loon and Van Lanen (2012) are visible in the parameter sensitivity, but certainly not all. Subsurface storage properties were found to be important in all models, but the importance of ET parameters differs per model and does not align with the theory, and the role of snow parameters clearly shows a distinction between the more conceptual snow modules in SAC and HBV, and the energy-balance-based snow implementation in VIC.

4.3. Implications of the Results

In their study on drought propagation, Apurv et al. (2017, p.3) state that “Simple conceptual models contain fewer numbers of parameters, which makes it easier to track the influence of each hydrologic process on the final outcome.” Although this statement in itself is true, the results of this study imply that no process-based conclusion can be drawn concerning hydrological drought based upon a single model. We showed that parameters representing comparable mechanisms (or processes) in different models had differing sensitivity to hydrological drought indicators; that is, it differs per model which parameters, representing different processes, are triggered to simulate hydrological drought. This implies that the choice of the model determines the conclusion of the study.

More often, simple conceptual models are used to identify driving processes. Van Loon et al. (2014), for example, demonstrate based on a study using a single simple conceptual model, which prolonged droughts in dry seasons occur due to evapotranspiration processes. With the results of this study in mind, the conclusions of Van Loon et al. (2014) could have been different when another model was employed: Based on HBV and SAC, it would be concluded that ET processes are most important when simulating hydrological drought in regions with low seasonality, while based on VIC, the conclusion would be that this is driven by deep layer processes.

The comparison of our results to the drought mechanisms as described by Van Loon and Van Lanen (2012) further demonstrates that not all drought mechanisms that theoretically could occur are well (enough) represented in the employed models. The way in which drought processes are simulated by a model can have profound effect on long term projections, thereby impacting the conclusions of the model study. Therefore, especially the way in which drought development mechanisms related to evapotranspiration and snow processes are implemented in models should be subject to closer scrutiny.

4.4. So, Which Model Is Right?

We showed that climate controls parameter sensitivity in hydrological drought simulation but that model structure also influences parameter sensitivity. The question that remains is which model is correct. HBV and SAC are comparable in terms of number of processes explicitly parameterized and also demonstrated rather comparable behavior, while VIC has more processes explicitly parameterized and demonstrated fairly different behavior. Droughts are a complex interplay of precipitation and temperature anomalies, evapotranspiration processes, and soil moisture and groundwater storages (Seneviratne, 2012). To identify which physical processes are involved during hydrological drought, models only will not provide any answers. Seneviratne et al. (2012), for example, showed based on observations in a research catchment in Switzerland that runoff responded earlier to precipitation deficit in the 2003 central European drought than soil moisture, which was in contrast to what regional models simulated. Teuling et al. (2013) analyzed total subsurface storage at the catchment scale and showed that ET amplifies drought but also discuss that the role of ET changes over the course of a drought; the driving process of a hydrological drought can change over time. Henn et al. (2018) used a water balance approach in combination with soil moisture and snow observations to estimate ET in the Sierra Nevada mountain range during the 2013–2015 drought. Their results contain signals that indicate reduced ET as evidenced by large-scale drought-induced tree mortality: a process not represented in any of the investigated conceptual models. Models are, unavoidably, a simplified representation of reality, thereby not capturing all processes that influence the water cycle. Models have heuristic
value: to formulate new hypotheses and to provide directions (Oreskes et al., 1994). However, especially for complex phenomena such as droughts, direct physical interpretation of simulated processes can be deceitful.

5. Conclusion

In this large-sample hydrological study, we compared parameter sensitivity in three hydrological models (SAC, VIC, and HBV) over a range of climates. We investigated the controls of climate and model structure on model mechanisms that are triggered when simulating drought duration and drought deficit. Parameter sensitivity related to Köppen-Geiger climate classes and hydro-climatic indicators revealed that parameter sensitivity varies over climate. The relation between parameter sensitivity and climate was anticipated and allowed us to explore whether the drought-triggering mechanisms as described by Van Loon and Van Lanen (2012) were reflected in parameter sensitivity. We found that some of the drought mechanisms are visible in the parameter sensitivity, but certainly not all: Especially, the sensitivity of ET parameters does not align with the theory, and the role of snow parameters in snow-related droughts clearly shows a distinction between degree-day-based models and energy-balance models. Besides parameter sensitivity being different over climates, we also found that parameter sensitivity differed among the models. Whereas in VIC the deep layer parameters showed the highest sensitivity for drought duration and drought deficit, HBV and SAC displayed a larger role for snow parameters and evapotranspiration parameters. This implies that conclusions on driving mechanisms in hydrological drought cannot be based on hydrological models only, as different models would lead to different conclusions. Provided that not all drought-triggering mechanisms are captured well by the models and that different models would lead to different conclusions concerning relevant processes, hydrological models should mainly be used for their heuristic value in drought research: to formulate new hypotheses and to identify research directions. Formulated hypotheses on driving processes should always be backed up by observations.

Acknowledgments

We would like to thank A.J. (Ryan) Teoh for feedback on the first version of this manuscript and Stephan Theber for suggestions and discussions during EGU 2018. L. M. thanks colleagues from NCAR (Martyn Clark, Andrew Newman, Naoki Mizukami, and Nans Addor) for support with setting up the calculations. B. G. thanks the DFG for financial support (Project GU 1466/1-1 Hydrological Consistency in Modelling). All data for the analyses in this study are publicly available online (at https://www.hydroshare.org/resource/8aa8b1d840474e5c94749c3baf8bc5ce/).

References

Abebe, N., Ogden, F., & Pradhan, D. (2010). Sensitivity and uncertainty analysis of the conceptual HBV rainfall-runoff model: Implications for parameter estimation. *Journal of Hydrology*, 389, 301–310. https://doi.org/10.1016/j.jhydrol.2010.06.007
Addor, N., Newman, A., Mizukami, N., & Clark, M. (2017a). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 1, 5293–5313. https://doi.org/10.5194/hess-21-5293-2017
Addor, N., Newman, A., Mizukami, N., & Clark, M. (2017b). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. version 1.0. Boulder, CO: UCAR/NCAR. https://doi.org/10.5065/D6G73C3Q
Andreadis, K. M., & Lettenmaier, D. P. (2006). Trends in 20th century drought over the continental United States. *Geophysical Research Letters*, 33, L10403. https://doi.org/10.1029/2006GL025711
Aputv, T., Sivapalan, M., & Cai, X. (2017). Understanding the role of climate characteristics in drought propagation. *Water Resources Research*, 53, 9304–9329. https://doi.org/10.1002/2017WR021445
Bergström, S. (1976). Development and application of a conceptual runoff model for Scandinavian catchments (RJO 7). Norrköping: SMHI.
Bergström, S. (1992). The HBV model—Its structure and applications (4): SMHI reports hydrology.
Burnash, R., Ferral, R., & McGuire, R. (1973). A generalized streamflow simulation system—Conceptual modeling for digital computers: U.S. Department of Commerce, National Weather Service and State of California, Department of Water Resources.
Chaney, N., Herman, J., Reed, P., & Wood, E. (2015). Flood and drought hydrologic monitoring: the role of model parameter uncertainty. *Hydrology and Earth System Sciences*, 19, 3239–3251. https://doi.org/10.5194/hess-19-3239-2015
Demaria, E. M., Nijsen, B., & Wagener, T. (2007). Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model. *Journal of Geophysical Research*, 112, D11113. https://doi.org/10.1029/2006JD007534
Dracup, J., Lee, K., & Paulson Jr, E. (1980). On the definition of droughts. *Water Resources Research*, 16(2), 297–302.
Fontrudona Bach, A., van der Schrier, G., Melsen, L., Klein Tank, A., & Teuling, A. (2018). Widespread and accelerated decrease of observed mean and extreme snow depth over Europe. *Geophysical Research Letters*, 45, 12,312–12,319. https://doi.org/10.1029/2018GL079799
Gupta, H., & Razavi, S. (2018). Revisiting the basis of sensitivity analysis for dynamical Earth system models. *Water Resources Research*, 54, 8692–8717. https://doi.org/10.1029/2018WR022668
Guse, B., Pfannerstill, M., Gafurov, A., Fohrer, N., & Gupta, H. (2016). Demasking the integrated information of discharge: Advancing sensitivity analyses to consider different hydrological components and their rates of change. *Water Resources Research*, 52, 8724–8743. https://doi.org/10.1002/2016WR018894
Hannaford, J., Lloyd-Hughes, B., Keef, C., Parry, S., & Prudhomme, C. (2011). Examining the large-scale spatial coherence of European drought using regional indicators of precipitation and streamflow deficit. *Hydrological Processes*, 25, 1146–1162. https://doi.org/10.1002/hyp.7725
Haslinger, K., Koffler, D., Schönler, W., & Laaha, G. (2014). Exploring the link between meteorological drought and streamflow: Effects of climate-catchment interaction. *Water Resources Research*, 50, 2468–2487. https://doi.org/10.1002/2013WR015051
Henn, B., Painter, T., Bormann, K., McGirk, B., Flint, A., Flint, L., et al. (2018). High-elevation evapotranspiration estimates during drought: Using streamflow and NASA Airborne Snow Observatory SWE Observations to close the Upper Tuolumne River Basin water balance. *Water Resources Research*, 54, 746–766. https://doi.org/10.1002/2017WR020473
Herman, J., Reed, P., & Wagener, T. (2013). Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior. *Water Resources Research*, 49, 1400–1414. https://doi.org/10.1002/wrcr.20124
Heudorfer, B., & Stahl, K. (2016). Comparison of different threshold level methods for drought propagation analysis in Germany. Hydrology Research, 48(5), 1311–1326. https://doi.org/10.2166/nh.2016.258

Killingtveit, A., & Engen, I. (2003). Nedbør, temperatur og tilsig Aret 2002/2003 - en beskrivelse av klima og tilsig forhold i norge og sverige (precipitation, temperature and inflow year 2002/2003—A description of climate and inflow conditions in norway and sweden) (8), Oslo, Norway: Olje- og Energidepartementet.

Knochen, W., Woods, R., & Freer, J. (2018). A quantitative hydrological climate classification evaluated with independent streamflow data. Water Resources Research, 54, 5088–5109. https://doi.org/10.1002/2018WR022913

Knutti, R., & Sedlacek, J. (2013). Robustness and uncertainties in the new CMIP5 climate model projections. Nature Climate Change, 3, 369–373. https://doi.org/10.1038/NCLIMATE1716

Köppen, W. (1936). Das geographische System der Klimate. In W. Köppen, & G. Geiger (Eds.), Handbuch der klimatologie. Berlin: Gebr Borntraeger.

Lahm, G., Gauster, T., Tallaksen, L., Vidal, J., Stahl, K., Pudhommne, C., et al. (2017). The European 2015 drought from a hydrological perspective. Hydrology and Earth System Sciences, 21, 3001–3024. https://doi.org/10.5194/hess-21-3001-2017

Lhomme, J. (1997). An examination of the Priestley-Taylor equation using a convective boundary layer model. Water Resources Research, 33(11), 2571–2578. https://doi.org/10.1029/97WR01897

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research, 99(D7), 14,415–14,458. https://doi.org/10.1029/94JD00483

Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. Global and Planetary Change, 13, 195–206. https://doi.org/10.1016/0921-8181(95)00046-1

Melsen, L., Addor, N., Mizukami, N., Newman, A., Torfs, P., Clark, M., et al. (2018). Mapping (dis)agreement in hydrologic projections. Hydrology and Earth System Sciences, 22, 1775–1791. https://doi.org/10.5194/hess-22-1775-2018

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Clark, M., & Uijlenhoet, R. (2016). Representation of spatial and temporal variability in large-domain hydrological models: Case study for a mesoscale prealpine basin. Hydrology and Earth System Sciences, 20, 2207–2226. https://doi.org/10.5194/hess-20-2207-2016

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Mendoza, P., et al. (2019). Subjective modeling decisions can significantly impact the simulation of flood and drought events. Journal of Hydrology, 568, 1140–1155. https://doi.org/10.1016/j.jhydrology.2018.11.046

Mendoza, P. A., Clark, M., Barlage, M., Rajagopalan, B., Samaniego, L., Abramowitz, G., & Gupta, H. (2015). Are we unnecessarily constraining the ability of complex process-based models? Water Resources Research, 51, 716–728. https://doi.org/10.1002/2014WR015820

Mo, K. (2008). Model-based drought indices over the United States. Journal of Hydrometeorology, 9, 1212–1230. https://doi.org/10.1175/2008JHM1002.1

National Weather Service. (2002). III.5-SC-SC conceptualization of the sacramto soil moisture accounting model: National Oceanic and Atmospheric Administration (NOAA).

Newman, A., Clark, M., Sampson, K., Wood, A., Hay, L., Bock, A., et al. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance. Hydrology and Earth System Sciences, 19, 209–223. https://doi.org/10.5194/hess-19-209-2015

Newman, A., Sampson, K., Clark, M., Bock, A., Viger, R., & Blodgett, D. (2014). A large sample watershed-scale hydrometeorological dataset for the contiguous USA. Boulder, CO: UCAR/NCAR. https://doi.org/10.5065/DeMWF24D

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the Earth Sciences. Science, 263(5147), 641–646. https://doi.org/10.1126/science.263.5147.641

Pappenberger, F., Beven, K., Ratto, M., & Matgen, P. (2008). Multi-method global sensitivity analysis of flood inundation models. Advances in Water Resources, 31(1), 1–14. https://doi.org/10.1016/j.advwatres.2007.04.009

Parajka, J., Merz, R., & Bloschl, G. (2007). Uncertainty and multiple objective calibration in temporal regional water balance modelling: case study in 320 Austrian catchments. Hydrological Processes, 21, 435–446. https://doi.org/10.1002/hyp.6253

Peel, M., Finlayson, B., & McMahon, T. (2007). Updated world map of the Köppen-Geiger climate classification. Journal of Hydrology, 330, 1–28. https://doi.org/10.1016/j.jhydrol.2006.06.005

Piao, S., Hurtt, G. C., Cao, L., Chen, A., Shi, M., Liang, X., et al. (2014). Land-use responses and their impacts on future climate change. Science, 345(6197), 933–936. https://doi.org/10.1126/science.1253161

Prudhomme, C., Georgakakos, K., Girvetz, E. H., Lettenmaier, D. P., & Wood, E. F. (2017). A global terrestrial water balance model for large-scale hydrologic simulation. Journal of Geophysical Research: Atmospheres, 122(16), 8606–8629. https://doi.org/10.1002/2016JD026015

Rakovec, O., Hill, M. C., Clark, M. P., Weeris, A. H., Teuling, A. J., & Uijlenhoet, R. (2014). Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models. Water Resources Research, 50, 409–426. https://doi.org/10.1002/2013WR014063

Razavi, S., & Gupta, H. (2015). What do we mean by sensitivity analysis? The need for comprehensive characterization of global sensitivity in earth and environmental systems models. Water Resources Research, 51, 3070–3092. https://doi.org/10.1002/2014WR016527

Selbert, J., & Viss, M. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. Hydrology and Earth System Sciences, 16, 3315–3325. https://doi.org/10.5194/hess-16-3315-2012

Seneviratne, S. (2012). Historical drought trends revisited. Nature, 491, 338–339. https://doi.org/10.1038/491338a

Seneviratne, S., Lehner, I., Gurtz, J., Teuling, A. J., Lang, H., Moser, U., et al. (2012). Swiss prealpine Rhielohlbach research catchment and lysimeter: 32 year time series and 2003 drought event. Water Resources Research, 48, W06526. https://doi.org/10.1029/2011WR011749

Sheffield, J., Andreassian, K., Wood, E., & Lettenmaier, D. (2009). Global and continental drought in the second half of the twentieth century: Severity area duration analysis and temporal variability of large-scale events. Journal of Climate, 22, 1962–1981. https://doi.org/10.1175/2008JCLI2722.1

Sheffield, J., & Wood, E. (2007). Global trends and variability in soil moisture and drought characteristics, 1950-2000, from observation-driven simulations of the terrestrial hydrologic cycle. Journal of Climate, 21, 432–457. https://doi.org/10.1175/2007JCLI1822.1

Shukla, S., & Wood, A. (2008). Use of a standardized runoff index for characterizing hydrologic drought. Geophysical Research Letters, 35, L02405. https://doi.org/10.1029/2007GL032487

Sobol’, I. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 7(4), 784–802. https://doi.org/10.1016/0041-5553(67)90144-9
Staudinger, M., Stahl, K., Seibert, J., Clark, M. P., & Tallaksen, L. (2011). Comparison of hydrological model structures based on recession and low flow simulations. *Hydrology and Earth System Sciences*, 15, 3447–3459. https://doi.org/10.5194/hess-15-3447-2011

Tallaksen, L. M., & Van Lanen, H. A. J. (2004). *Hydrological drought: Processes and estimation methods for streamflow and groundwater developments in water science*. Amsterdam, The Netherlands: Elsevier Science.

Teuling, A., Van Loon, A., Seneviratne, S., Lehner, I., Aubinet, M., Heinesch, B., et al. (2013). Evapotranspiration amplifies European summer drought. *Geophysical Research Letters*, 40, 2071–2075. https://doi.org/10.1002/grl.50495

Thornton, P. E., Thornton, M., Mayer, B., Wilhelmi, N., Wei, Y., & Cook, R. (2012). Daymet: Daily surface weather on a 1 km grid for North America, pp. 1980–2012. Tennessee, USA: Oak Ridge National Laboratory. http://daymet.ornl.gov/

Tijdeman, E., Barker, L., Svoboda, M., & Stahl, K. (2018). Natural and human influences on the link between meteorological and hydrological drought indices for a large set of catchments in the contiguous United States. *Water Resources Research*, 54, 6005–6023. https://doi.org/10.1029/2017WR022412

Uhlenbrook, S., Seibert, J., Leibundgut, C., & Rodhe, A. (1999). Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal*, 44(5), 779–797.

Van Huijgevoort, M. H. J., Hazenberg, P., Van Lanen, H. A. J., & Uijlenhoet, R. (2012). A generic method for hydrological drought identification across different climate regions. *Hydrology and Earth System Sciences*, 16, 2437–2451. https://doi.org/10.5194/hess-16-2437-2012

Van Lanen, H., Wanders, N., Tallaksen, L., & Van Loon, A. (2013). Hydrological drought across the world: impact of climate and physical catchment structure. *Hydrology and Earth System Sciences*, 17, 1715–1732. https://doi.org/10.5194/hess-17-1715-2013

Van Loon, A. F., & Laaha, G. (2015). Hydrological drought severity explained by climate and catchment characteristics. *Journal of Hydrology*, 526, 3–14. https://doi.org/10.1016/j.jhydrol.2014.10.059

Van Loon, A. F., Tijdeman, E., Wanders, N., Van Lanen, H., Teuling, A., & Uijlenhoet, R. (2014). How climate seasonality modifies drought duration and deficit. *Journal of Geophysical Research: Atmospheres*, 119, 4640–4656. https://doi.org/10.1002/2013JD020383

Van Loon, A. F., Van Huijgevoort, M. H. J., & Van Lanen, H. A. J. (2012). Evaluation of drought propagation in an ensemble mean of large-scale hydrological models. *Hydrology and Earth System Sciences*, 16, 4057–4078. https://doi.org/10.5194/hess-16-4057-2012

Van Loon, A. F., & Van Lanen, H. (2012). A process-based typology of hydrological drought. *Hydrology and Earth System Sciences*, 16, 1915–1946. https://doi.org/10.5194/hess-16-1915-2012

Van Loon, A. F., & Van Lanen, H. (2013). Making the distinction between water scarcity and drought using an observation-modeling framework. *Water Resources Research*, 49, 1483–1502. https://doi.org/10.1002/wrcr.20147

Van Werkhoven, K., Wagener, T., Reed, P., & Tang, Y. (2008). Characterization of watershed model behavior across a hydroclimatic gradient. *Water Resources Research*, 44, W01429. https://doi.org/10.1029/2007WR006271

Wang, A., Bohn, T., Mahanama, S., Koster, R., & Lettenmaier, D. (2008). Multimodel ensemble reconstruction of Drought over the continental United States. *Journal of Climate*, 22, 2694–2712. https://doi.org/10.1175/2008JCLI2586.1

Wilhite, D., Svoboda, M., & Hayes, M. (2007). Understanding the complex impacts of drought: A key to enhancing drought mitigation and preparedness. *Water Resources Management*, 21, 763–774. https://doi.org/10.1007/s11269-006-9076-5

Willmott, C., & Fedema, J. (1992). A more rational climatic moisture index. *The Professional Geographer*, 44(1), 84–88. https://doi.org/10.1111/j.0033-0124.1992.tb00844.x

Yevjevich, V. (1967). An objective approach to definitions and investigations of continental hydrologic droughts. *Hydrology Papers, Colorado State University*, 23, 1–18.