Comparison of different threshold level methods for drought propagation analysis in Germany

B. Heudorfer and K. Stahl

ABSTRACT

The Threshold Level Method is an approach that enables comparability across all hydrological levels. This advantage is used especially in studies on drought propagation. There are different calculation procedures for this method. The effect that the choice of a variable versus a constant threshold level method has on drought characteristics and drought propagation patterns has not been fully explored yet. Also, most drought propagation studies have analyzed modelled data, suggesting that applicability to observations be tested. We tested the Constant and the Variable Threshold Level Method for the 10th, 20th and 30th percentile on observed precipitation, streamflow, and groundwater data from Germany, and compared drought characteristics and drought propagation patterns by means of statistical analysis and synoptic assessment. The characteristic effects of choosing a variable versus a constant threshold are: (1) a substantial increase in short droughts, (2) a moderate decrease in intermediate droughts and (3) a minor increase in long droughts. Furthermore, in slow-reacting lowland catchments, theoretical propagation characteristics could mostly be confirmed. In faster-reacting upland catchments, this was not always the case and considerable differences arose. Sources of ambiguity were predominantly groundwater in lowlands and streamflow in the mountainous catchments. In conclusion, there is potential of diverging inference from the same data, depending on the chosen methodology.

Key words | catchment scale, constant threshold, drought duration, drought propagation, Germany, variable threshold

INTRODUCTION

Drought is a ‘creeping disaster’ (Mishra & Singh 2010) with complex and indirect mechanisms and impacts that are not trivial to study. Hydrological drought occurs when a meteorological drought starts to affect the terrestrial hydrosphere. When this happens, an abundance of processes and feedback mechanisms start to take effect that are yet to be fully understood. One way to deal with the complexity of the processes and feedbacks involved is the conceptualization of drought development as the propagation of negative hydro-climatic signals (i.e. meteorological droughts) through interconnected hydrological subsystems like the soil systems, surface water systems and groundwater bodies. Pinning down controls that slow down, speed up or deviate signal processing (i.e. catchment and climate controls) is one way to enhance process understanding as well as to determine an area’s vulnerability to hydro-climatic extremes, i.e. droughts.

This concept is referred to as ‘drought propagation’, introduced as a hypothetical framework as such by Changnon (1987), but first coined as a term by Eltahir & Yeh (1999). Examples of this approach can be found, for example, in Eltahir & Yeh (1999). They showed for the case of Illinois, USA, how seasonality and other atmospheric patterns govern the evolution of climatic anomalies and therefore droughts, while simultaneously groundwater is – unlike other hydrological variables – more vulnerable

doi: 10.2166/nh.2016.258
to drought events instead of flood events due to nonlinear aquifer drainage mechanisms. In a synthetic modelling study, Peters et al. (2003) showed the drought-delaying and -attenuating properties of groundwater systems, and Peters et al. (2006) revealed spatially diversified drought development mechanisms in the groundwater system of the Pang catchment in the UK. In the same catchment, Tallaksen et al. (2009) studied spatially aggregated drought characteristics in precipitation and simulated groundwater recharge, head and discharge and pointed out the differences in spatial patterns and vertical drought features. Also, the conceptual drought propagation approach allowed Vidal et al. (2010) to identify and explain the evolution of the pan-European benchmark drought events in France as well as local anomalies by means of a comparative analysis in different hydro-geological and hydro-climatic settings. Finally, Van Loon et al. (2012) and Van Loon & Van Lanen (2012) were able to distinguish six different seasonality-driven drought genesis processes across various climate zones by looking at drought propagation patterns. Their proposed drought typology, however, is yet to be validated in a wider range of catchments (Van Loon 2015) and by means of observation frameworks or other than previously used modelling frameworks. General drought propagation patterns are shown in Figure 1 and summed up in more detail by Van Loon (2013). It can be said that a lag, lengthening, attenuation and pooling occurs when droughts move from precipitation over soil moisture into streamflow and groundwater. These patterns are based upon the findings of the above-mentioned and other studies. In addition, occurrence and strength of these patterns underlie catchment and climate controls (Tijdeman et al. 2012; Van Lanen et al. 2013; Van Loon et al. 2014).

Within the drought propagation framework, one method to pin down patterns, processes and controls is to conduct comparative analysis on drought characteristics in the different levels of the hydrological cycle. Drought characteristics are features like drought duration, drought intensity (maximum deviation from threshold), deficit water volume (or severity) or the affected area. They can be obtained from the partial duration time series truncated by a threshold. This method is called the Threshold Level Method. Many studies on drought propagation use the Threshold Level Method in one way or another. This has to do with its ability to define drought characteristics in all possible hydrological variables in a uniform manner. This enables cross-variable frequency analysis of uniform statistical products (drought characteristics) which makes it a suitable tool for drought propagation studies (Van Loon 2015).

The Threshold Level Method was first introduced by Yevjevich (1967) and further elaborated by Zelenhasic & Salvai (1987) into an explicit drought definition approach, as it is used in drought analysis up to date. Two families of threshold level methods emerged: the Constant and the Variable Threshold Level Method. In the case of the constant threshold, as used for example by Tallaksen et al. (1997), Stahl (2001), Andreadis et al. (2005), Vidal et al. (2010), Sheffield et al. (2012) and Tallaksen & Stahl (2014), usually a single percentile of the long-term cumulative frequency distribution of the regarded hydrologic variable is used as a time invariant threshold. In the case of the variable threshold, which has been more widely used in recent years (Stahl 2001; Hirabayashi et al. 2008; Hannaford et al. 2011; Prudhomme et al. 2011; Tijdeman et al. 2012; Van Huijgenvoort et al. 2012; Van Loon & Van Lanen 2012; Van Loon et al. 2014), the threshold level varies over the year, following the seasonal amplitude. It can be calculated in numerous ways, for example by taking a smoothed monthly (Van Loon & Van Lanen 2012) or a daily percentile as the varying threshold level (Stahl 2001; Hannaford et al. 2011). Commonly used percentiles for the threshold level for both constant and variable threshold methods range from the 70% percentile up to the 95% percentile (Van Loon 2015), with the 80% percentile being the most common choice. This is true for perennial rivers common in the climate zone regarded in this study (warm temperate,
humid). Different percentiles might need to be chosen, for example for intermittent streams.

There are specific caveats for both methods. For example, provided there is a pronounced seasonality, as is the case in Germany, droughts will be defined almost exclusively during the dry season with the constant threshold. Now if droughts are viewed as below normal water availability – which is the most common definition (Mishra & Singh 2010) – then the constant threshold is not the correct way to define droughts, because deficits relative to the relative normal during the wet season can occur as well. From that perspective, a variable threshold is the better choice. On the other hand, drought can also be viewed as a social construct that defines specific periods of climate variability as a natural hazard in, from a physical perspective, a rather arbitrary manner. Which period this is, is defined by the impact it has on society. Therefore, we can technically only speak of drought when there is a real impact, which is during the dry season, mostly, when the absolute water deficit is highest. From that perspective, we actually could not speak of ‘droughts’ when we use the variable threshold, and should rather be calling it ‘relative water deficit’ (Stahl 2001), or use the constant threshold. In any case, a definite answer regarding to which drought definition scheme is the most suitable can only be given by ‘ground-truthing’ said methods by means of empirical impact data, by determining which framework can best predict impact occurrence. In this direction, recent efforts to build empirical impact databases are promising (e.g. Bachmair et al. 2015; Stahl et al. 2016).

As we can see, there are specific advantages and disadvantages of each method, depending on how we define drought. However, while there is some literature available that discusses severity-duration-frequency curves of streamflow for different Threshold Level Methods (Sung & Chung 2014; Sarailidis et al. 2015), as well as some recent non-peer-reviewed literature that discuss different Threshold Level Methods in some way (e.g. Boer et al. 2015; Beyene et al. 2015), to our knowledge no definite comparison of the constant and variable threshold has been published with regard to drought propagation. However, to choose the appropriate threshold method for one’s study or management application, more clarity regarding the implication of the choice is essential. The aim of this study is therefore: (1) to evaluate the difference in drought characteristics and drought propagation patterns when analysis is undertaken with the constant and the variable threshold, respectively, and (2) to test if the findings of previous studies – which were based on modelling frameworks, mostly – are reproducible in an observation framework. To do this, this study compares drought propagation patterns under one prominent constant and one prominent variable threshold level method calculation procedure, respectively. The study was conducted for precipitation, discharge and groundwater head in various catchments in Germany. Drought propagation patterns were evaluated by means of drought duration frequency distributions (hereafter referred to as DFDs) and synoptic assessment of occurrence and strength of drought propagation patterns such as lag, lengthening and pooling in the hydrographs.

**DATA**

Streamflow and groundwater data were taken from a dataset collected by Kohn et al. (2014) for a national study on the low flow conditions of 2011. The data pool contained, amongst others, about 280 groundwater gauges and 350 stream discharge gauges across Germany, provided by the administrative departments of the federal states. The data were taken from locations with near-natural conditions. From this dataset, four streamflow gauges with matching unconfined groundwater gauges within their catchment boundaries were selected. These are used for the case study investigations of this study. The groundwater gauges were chosen according to the key borehole approach (Peters et al. 2006), where data from few or only one borehole are taken as a representative for the aquifer.

Precipitation data were taken from the freely available E-OBS dataset (Haylock et al. 2008) as 0.25° grid cell averages. The grid cell precipitation was then interpolated to the catchment area with the Thiessen Polygon method. Precipitation data, as well as streamflow data, were available on a daily basis. However, a larger share of the groundwater records had weekly or two-weekly measurements, which were linearly interpolated to daily resolution. The length of all time series is 56 years (1975–2011), with groundwater heads in units of m (a.s.l.), and stream discharge and precipitation normalized to catchment size, in units of mm/d.
For confirmation/validation purposes, an alternative, largest possible sample of 167 precipitation gauges, 217 streamflow gauges and 117 groundwater gauges was compiled from the above-mentioned sources. Here, selection criteria were less strict than for the case study selection: the data were required to be near-natural and within the 1975–2011 measurement period. The spatial distribution was on a 0.25° grid for precipitation, and randomly distributed for groundwater and streamflow.

The assessment of catchment and aquifer settings was undertaken using various sources of hydrogeological (meta-)data. These data were taken mainly from generalized mapping products as well as primary borehole profiles and stratigraphic cross profiles. Used mapping products were the HUEK200 (BGR & SGD 2011a), the GUEK200 (BGR & SGD 2011b) and the Hydrological Atlas of Germany HAD (Bundesamt für Gewässerkunde 2005), which are compilations of hydrogeological maps depicting, for example, geological setting, hydrological conductivities, soil properties, river network densities etc. Borehole profiles and stratigraphic cross sections were available for the Lachte and Oertze aquifers. Topographic data were taken from the HydroSHEDS DEM, a 100 × 100 m grid (Lehner et al. 2008).

**STUDY AREAS**

The four case catchments are situated in mid- and northern Germany. Lachte and Oertze are directly adjacent catchments located in the northern lowlands (Figure 2). They are catchments without distinct topographic features and comparable elevations, however the Oertze is about twice the size of the Lachte (Table 1). The Salz and Spree catchments are two ‘mountainous’ catchments located in the mountain ranges of Erzgebirge and Vogelsberg, respectively (Figure 2). Hereby the term ‘mountainous’ refers to the intermediate mountain ranges that are typical for central and southern Germany, with relatively steep slopes and elevation ranges of several hundred metres (so-called ‘Mittelgebirge’, sometimes called ‘intermediate mountains’ for distinction from the Alps). Furthermore, while the Spree catchment has a round shape, stretching over a hilly landscape, the Salz catchment is elongated with a more deeply incised valley. Elevation ranges of Spree and Salz catchments are similar, but with a larger share of high altitudes in the Salz (Table 1). The Spree catchment is about three times larger than the Salz catchment. There are no significant surface water storages (lakes, wetlands etc.) in these catchments.

Groundwater levels at the groundwater gauges in the two lowland catchments Oertze and Lachte are close to the surface, ranging from 2 to 4.5 m below ground level (Table 2).
The difference in altitude of the groundwater gauge relative to the closest stream branch is small, while groundwater levels are some 3–7 m above stream level, resulting in very small hydraulic gradients (Table 2), indicating high groundwater–surface water connectedness. In the two mountainous catchments, Salz and Spree, groundwater levels are deeper beneath the surface, with water levels ranging from 4.5 to 12 m below ground level (Table 2). Also, groundwater levels fluctuate on a magnitude twice as high as in the lowland gauges. As mountainous catchments, the height difference relative to the closest stream branch is much higher and groundwater levels are between 50 and 115 m above stream level. The distance to stream is not significantly shorter than in the two lowland catchments, resulting in relatively high hydraulic gradients (Table 2), indicating a lower groundwater–surface water connectedness.

Concerning aquifer characteristics, the Lachte and Örtze catchments stretch across the same large-scale, rather homogeneous aquifer system. It consists mostly of sand with some fraction of gravel. The aquifer depth of the top unconfined layer is about 20 m and conductivities range in the magnitude of about 10^{-3} to 10^{-4} m/s, according to the hydrological map HUEK200 (BGR & SGD 2011b). In the Salz and Spree catchments, aquifer characteristics are more heterogeneous. The Salz groundwater gauge is located on a small plateau above the Salz canyon about 120 m above the stream. According to HUEK200, conductivities in location of the gauge are about 10^{-3} to 10^{-5} m/s. The aquifer in the location of the Spree groundwater gauge is underneath a saddle in between three hills. The aquifer’s material is slope debris with conductivities of about 10^{-4} to 10^{-6} m/s according to HUEK200, and the aquifer is significantly sloped, according to GUEK200 (BGR & SGD 2011a).

### Methods

The overall approach was to study drought propagation on the basis of observed precipitation, streamflow and groundwater head records. Only observed data were used in contrast to many previous studies on drought propagation. The data were taken from four study catchments across Germany in order to account for catchment controls, following the comparative approach (Sivapalan 2009). The first step was to calculate drought durations according to the two threshold level methods. The second step was to calculate the drought DFDs. In addition to simple synoptic hydrograph analysis, the DFDs were used to assess the (change in) occurrence and strength of drought propagation patterns such as lag, lengthening and pooling.

To determine droughts with the Threshold Level Method, a variable (precipitation, streamflow or groundwater head) is defined as being in drought when it drops below a predefined threshold. Subsequent exceedance of the threshold terminates the drought.

\[
d(t) = \begin{cases} 
1 & \text{if } x(t) < r(t) \\
0 & \text{if } x(t) \geq r(t) 
\end{cases}
\]

with \( \delta(t) \) being the binary variable defining whether the hydrological variable \( x \) is below the threshold \( r \) on day \( t \), and therefore in drought.

This way, drought characteristics like timing, duration, intensity and deficit volume (or severity) can be quantified in an operative manner. The deficit volume is often argued to be the most suitable characteristic to study drought propagation. However, the deficit volume can only be calculated for flow variables. Groundwater head is a state variable. To convert it into a flow variable we need to calculate the

### Table 2 | Space-and-height relations of the groundwater gauges

<table>
<thead>
<tr>
<th></th>
<th>( h_{\text{gauge}} ) (m)</th>
<th>( h_{\text{gw,min}} ) (m)</th>
<th>( h_{\text{gw,max}} ) (m)</th>
<th>( h_{\text{gw,range}} ) (m)</th>
<th>( h_{\text{stream}} ) (m)</th>
<th>( L ) (m)</th>
<th>( I_{\text{min}} ) (–)</th>
<th>( I_{\text{max}} ) (–)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Örtze</td>
<td>50.8</td>
<td>46.28</td>
<td>47.79</td>
<td>1.51</td>
<td>41.0</td>
<td>876.0</td>
<td>0.006</td>
<td>0.008</td>
</tr>
<tr>
<td>Lachte</td>
<td>84.5</td>
<td>81.10</td>
<td>82.45</td>
<td>1.35</td>
<td>78.0</td>
<td>1259.0</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>Spree</td>
<td>326.4</td>
<td>318.09</td>
<td>321.81</td>
<td>3.72</td>
<td>268.0</td>
<td>790.0</td>
<td>0.063</td>
<td>0.068</td>
</tr>
<tr>
<td>Salz</td>
<td>306.9</td>
<td>294.64</td>
<td>297.84</td>
<td>3.20</td>
<td>182.0</td>
<td>1165.0</td>
<td>0.097</td>
<td>0.099</td>
</tr>
</tbody>
</table>

\( h_{\text{gauge}} \): Elevation of groundwater gauges. \( h_{\text{gw,min}}/h_{\text{gw,max}}/h_{\text{gw,range}} \): Minimum, maximum and range of water levels in borehole. \( h_{\text{stream}} \): Elevation of the stream branch that is nearest to the groundwater gauge. \( L \): Distance between groundwater gauge and stream. \( I_{\text{min}}/I_{\text{max}} \): Gradient from gauge to stream.
storage coefficient, which requires information about aquifer properties, which was not available. Therefore deficit volume cannot be calculated for groundwater, and only drought duration was analysed as a proxy:

\[ \Delta_i = \sum_{t=1}^{T} \delta_i(t) \]  

(2)

where \( \Delta_i \) is the duration of drought \( i \), lasting from day \( t = 1 \) until day \( T \).

Two different widely used Threshold Level Methods are applied in this study, the constant and the variable threshold. The threshold value \( \tau \) is set to be the 80th percentile of the flow duration curve in exceedance terms (i.e. the 20th percentile expressed as non-exceedance) as argued for and used by, for example, Fleig et al. (2006), Tallaksen et al. (2009) and Van Loon (2013). As long as no ground-truthing with respect to impact data is possible, the choice of the threshold level method remains subjective. Therefore, we repeated parts of the analysis for the 70th and 90th percentile, in order to evaluate the influence of the subjectivity on the results. Then, for the constant threshold, to select the quantile, the duration curve of the complete period 1976–2011 is used, resulting in one single, time–invariant threshold level for a specific percentile (Figure 3). For the variable threshold, the 12-monthly duration curves are used, resulting in a threshold level that changes its height in the course of the year. To avoid the ‘staircase pattern’ of the threshold that occurs with this calculation procedure, the threshold is subsequently smoothed with a 30-day centred moving average (MA), following Van Loon & Van Lanen (2012).

In drought propagation studies, it is important to apply the same threshold height (the percentile) in order to maintain comparability of the derived drought characteristics (Van Loon 2013). However, when using a relatively low threshold level like the 80th percentile (or the 70th or 90th), the method produces many insignificant droughts in precipitation because precipitation is zero most days of the year. To avoid these insignificant droughts, a 30-day backwards MA is applied on the precipitation hydrograph. Because applying a MA leads to a significant shift in peaks and troughs, resulting in a systematic distortion of drought timing, the same 30-day backwards MA has to be applied on streamflow and groundwater hydrographs (Figure 4). Although this might lead to distortions in drought characteristics, this does not affect the comparative analysis in this study as long as the MA is uniformly applied on all data, which is ensured under the presented framework.

Minor peaks can exceed the threshold for a short period of time, cutting droughts in half, thus producing two superficially independent, but actually dependent droughts (Zelenhasic & Salvai 1987). To minimize this effect, pooling procedures are needed (Fleig et al. 2006). First, minor droughts of less than 3 days duration are excluded. Second, an inter-event time criterion, whose sensitivity is disputed (Fleig et al. 2006; Van Loon 2013; Sarailidis et al. 2015), of 5 days was applied on groundwater, streamflow and precipitation. With these methods we follow Van Loon (2013), who discussed this choice in great detail.

Next to statistical drought duration analysis, a synoptic assessment of the hydrographs is undertaken in order to analyse drought propagation patterns. For this purpose a standardized hydrograph \( \bar{x} \) was established:

\[ \bar{x}(t) = x(t) - r(x) \]  

(3)

After standardizing precipitation, streamflow and groundwater, \( r \) is set to zero, which enables the simultaneous display of all three variables within one plot (Figures 8–11),
enhancing comparability of cross-variable drought characteristics and drought propagation patterns.

RESULTS

Drought DFDs

In all catchments, droughts in precipitation have generally both the largest number and shortest duration, streamflow experiences fewer but longer droughts and groundwater has the fewest droughts with the longest durations (Figure 5, Table 3). The one exception is the Spree, where the total number of droughts in groundwater is higher than in streamflow (Table 3). All DFDs can be characterised as being right-skewed, with skewness increasing from precipitation via streamflow to groundwater (Figure 6). The Lachte and Salz groundwater variables experience rather extreme tailing. Groundwater in the Oertze catchment appears to be left-skewed based on Figure 5, but this might be due to
misleading bin width in the histogram since Figure 6 clearly shows that all variables are right-skewed, including groundwater in the Oertze.

When comparing the DFDs derived from the constant threshold to the DFDs derived from the variable threshold, significant differences become apparent. DFDs show higher right-skewness when derived from the variable threshold (Figure 6). At the same time some long droughts were identified, resulting in more extensive tailing of the DFD (Figure 5). This is generally the case for all DFDs, however precipitation DFDs show small differences from catchment to catchment (Figure 5) and experience only small differences in product moments for the constant and varying threshold applications (Figure 6). Figure 7(a) confirms these observations. While precipitation on average does not show major differences, streamflow and groundwater head experience a substantial increase in the number of very short droughts when the variable threshold instead of the constant threshold is applied. At the same time, droughts with medium lengths become less frequent. Tails show a minor increase in droughts again, especially in groundwater. Because Figure 7(a) is only based on four catchments, a larger sample was drawn as described in the data chapter. The resulting Figure 7(b) supports the inference based on Figure 7(a), supporting generalizability of the observed patterns. These patterns also sustain for the 10th and 30th percentile (Figure A1, available with the online version of this paper).

It is interesting that slow and fast reacting patterns cannot be grouped according to classical catchment controls, e.g. lowland and mountainous topography, which typically distinguish large, fine grained (thus slow reacting) aquifers from smaller, coarse grained (thus fast reacting) aquifers, respectively. In particular, for the higher 70th percentile (Figure A2, available with the online version of this paper), streamflow in the two mountainous catchments Salz and Spree experience more tailing, i.e. longer droughts – especially under the constant threshold – than the two low-lying, flat catchments Lachte and Oertze. This is unexpected, since streamflow of sloped catchments should react flashier to rainfall events, interrupting drought events more frequently, thus leading to a prevalence of short droughts, not

Table 3 | Drought statistics

<table>
<thead>
<tr>
<th></th>
<th>n_c</th>
<th>n_v</th>
<th>Δ(n_c/n_v)</th>
<th>d_c (d)</th>
<th>d_v (d)</th>
<th>Δ(d_c/d_v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lachte P</td>
<td>147</td>
<td>152</td>
<td>1.03</td>
<td>18</td>
<td>17</td>
<td>0.94</td>
</tr>
<tr>
<td>Q</td>
<td>38</td>
<td>61</td>
<td>1.61</td>
<td>69</td>
<td>44</td>
<td>0.64</td>
</tr>
<tr>
<td>G</td>
<td>16</td>
<td>12</td>
<td>0.75</td>
<td>165</td>
<td>217</td>
<td>1.32</td>
</tr>
<tr>
<td>Oertze P</td>
<td>133</td>
<td>139</td>
<td>1.05</td>
<td>20</td>
<td>19</td>
<td>0.95</td>
</tr>
<tr>
<td>Q</td>
<td>52</td>
<td>71</td>
<td>1.37</td>
<td>51</td>
<td>39</td>
<td>0.76</td>
</tr>
<tr>
<td>G</td>
<td>23</td>
<td>30</td>
<td>1.30</td>
<td>115</td>
<td>86</td>
<td>0.75</td>
</tr>
<tr>
<td>Salz   P</td>
<td>137</td>
<td>145</td>
<td>1.06</td>
<td>19</td>
<td>18</td>
<td>0.95</td>
</tr>
<tr>
<td>Q</td>
<td>51</td>
<td>94</td>
<td>1.84</td>
<td>52</td>
<td>28</td>
<td>0.54</td>
</tr>
<tr>
<td>G</td>
<td>11</td>
<td>12</td>
<td>1.09</td>
<td>240</td>
<td>222</td>
<td>0.93</td>
</tr>
<tr>
<td>Spree  P</td>
<td>137</td>
<td>137</td>
<td>1.00</td>
<td>19</td>
<td>19</td>
<td>1.00</td>
</tr>
<tr>
<td>Q</td>
<td>38</td>
<td>56</td>
<td>1.47</td>
<td>70</td>
<td>48</td>
<td>0.69</td>
</tr>
<tr>
<td>G</td>
<td>43</td>
<td>42</td>
<td>0.98</td>
<td>61</td>
<td>61</td>
<td>1.00</td>
</tr>
</tbody>
</table>

n_c, number of droughts with the constant threshold; n_v, number of droughts with the variable threshold; d_c, average duration of droughts with the constant threshold; d_v, average duration of droughts with the variable threshold.

Table 3 | Drought statistics

Figure 6 | (a) and (b): Product moments of the drought duration distributions in Figure 5. Positive skew equals right skewness.
longer tails. This might be due to unexpectedly high storage capacities, however, the long tails in the Salz and Spree streamflow DFDs also coincide with a long tail in their corresponding precipitation DFDs, indicating a drier, more continental climate in these two catchments (Figure A2).

Concerning groundwater, we expected alluvial lowland aquifers to be slow responding and therefore experiencing few short droughts (since they are typically large, fine grained aquifers), and small mountainous aquifers to be fast responding, experiencing few long droughts (since typically smaller, coarser grained aquifers are located in this setting). However, exceptions exist in the data presented here, where one lowland and mountainous gauge experiences droughts with few and long durations including some multi-year droughts (Lachte and Salz, Figure 5, Table 3, Figure A2), and one lowland and mountainous gauge experiences many short drought durations (Oertze and Spree, Figure 5, Table 3, Figure A2). This suggests the existence of a slow responding aquifer system in Lachte (lowland) and Salz (mountainous), and a fast aquifer system in Oertze (lowland) and Spree (mountainous).

**Drought propagation patterns**

The Oertze catchment is one of the lowland catchments, where a typical drought propagation pattern under the constant threshold is visible (Figure 8(a)). The major 1991 precipitation drought starts on the 1st of February, translates with a delay of nearly 4 months into streamflow, and 20 days later in mid-June into groundwater. Pooling can be observed, as the peaks that cut precipitation and streamflow droughts in half are not present in groundwater. In effect, the groundwater drought prolongs, with 174 days in contrast to 103 and 106 days for precipitation and streamflow, respectively. When the variable threshold instead of the constant threshold is applied, changes in drought characteristics can be observed (Figure 8(a) and 8(b)). The streamflow and groundwater droughts start about 3 months earlier, thus streamflow and groundwater droughts are prolonged. Since the timing of the precipitation drought does not change, the delay between precipitation, streamflow and groundwater droughts is smaller now. However, while these differences are notable, the observed propagation patterns – pooling of
the two 91 streamflow droughts in groundwater as well as the delay and prolonging of droughts from precipitation to streamflow to groundwater – hold under both thresholds. Similar observations can be made in 1992.

The Lachte is the second lowland catchment. Herein, the precipitation winter drought at the turn of 2008/2009 does not translate into a hydrological drought, at least under the constant threshold (Figure 9(a)). However, a major streamflow summer drought develops on 24th of May after a period of intermediate precipitation surplus, indicating that storage conditions were affected during the wet period in winter, thus corresponding with the wet-to-dry season type drought after Van Loon & Van Lanen (2012). This 5-month (154 days) streamflow drought subsequently translates into a prolonged 6-month (182 days) groundwater drought with a delay of 2 months. Under the variable threshold (Figure 9(b)), the precipitation drought at the beginning of 2009 does produce a small streamflow drought (48 days) with little delay. The disturbed streamflow system recovers immediately before dropping into drought conditions again.
Compared to the constant threshold, the onset is about 1 month earlier. This streamflow drought with a pooled duration of 176 days propagates further into a groundwater drought some 2.5 months later. Here, under the variable threshold, the drought in groundwater ($d = 145$ days) does not prolong. Also, streamflow and groundwater droughts seem rather disconnected from precipitation droughts. They do not develop instantly from precipitation droughts, indicating the importance of large storage and thus memory in this catchment. In general, the appearance of drought propagation patterns (delay, prolonging, pooling) is more ambivalent than in the Oertze, with more variation when moving from the constant to the variable threshold.

In the Salz catchment, especially in streamflow, under the constant threshold, no streamflow drought is triggered by the four consecutive medium to major precipitation droughts in the course of the winter from late October 1995 to mid-May 1996. Only a 6-month (182 days) long groundwater drought occurs from the 27th of July onwards, 3 months after the onset of the last precipitation drought (Figure 10(a)). However, with application of the variable threshold (Figure 10(b)), the series of precipitation winter droughts from late October 1995 to mid-May 1996 instantly trigger a series of streamflow droughts with increasing intensity, followed by a delayed and prolonged groundwater drought. This pattern resembles classic drought propagation patterns, unlike under the constant threshold, where a streamflow drought does not occur and drought appears to propagate from precipitation directly into groundwater. This change in drought detection is owed almost exclusively to changes in streamflow, whereas groundwater experiences only a medium intensification, however a significant shift in timing. Both drought characteristics and drought propagation patterns are substantially different when another type of threshold method is applied.

In the Spree catchment, when applying the constant threshold, roles seem to be reversed. The major 1976 precipitation drought, starting on the 21st of February, propagates into groundwater first and into streamflow afterwards, with a delay of 1.5 and 3 months, respectively (Figure 11(a)). Also, the length of the groundwater drought (109 days) is less than that of streamflow (145 days), and pooling in groundwater does not occur. Groundwater experiences intermediate recoveries between the three consecutive droughts, reacting much faster than streamflow, where pooling occurs. When the variable threshold is applied, the situation changes. The onset of streamflow and groundwater droughts is simultaneous now on the 8th/9th of March, with a delay of only half a month after the precipitation drought (Figure 11(b)). Still, groundwater (136 days duration) recovers much faster than streamflow (226 days pooled duration), however both recover only intermediately and drop into drought condition.
at the turn of the years 1976/77, this time with a streamflow drought preceding the groundwater drought.

When different percentiles are taken as threshold level heights, a general pattern is observable. Moving from the 10th via the 20th to the 30th percentile, the drought onset generally shifts to an earlier date, while the drought termination shifts to a later date, resulting in longer drought durations (Figures A3–14, available with the online version of this paper), which is in accordance with the observed patterns in the DFDs (Figures A1–2). Also, additional droughts appear, and drought intensity increases, as the overall distance of hydrograph to threshold increases (Figures A3–14). This suggests that, while having a big impact on the unique drought characteristics themselves, the choice of the percentile as threshold level height is not a primary factor in influencing the overall drought propagation patterns, though it does influence their magnitude. Different behaviour in lowlands and highlands when applying a different percentile are not apparent.

**DISCUSSION**

**The effect of the choice of threshold level method**

The first objective of this study was to evaluate the difference that the choice of Threshold Level Method makes. For this purpose, drought duration was chosen as it is directly derivable from all hydrological variables in a uniform manner, including groundwater head. The overall best drought characteristic to study drought propagation would be deficit volume, which, however, is not trivial to calculate for groundwater (Van Loon 2013). Also, Ko & Tarhule (1994) and Oosterwijk et al. (2009) showed that different drought characteristics are highly correlated, so drought duration can be taken as representative for other characteristics. This study revealed characteristic effects on drought DFDs that can be summarized as follows: when the variable instead of the constant threshold is applied, for streamflow and groundwater there is on average:

- a substantial increase in short droughts;
- a moderate decrease in droughts of medium duration;
- a minor increase in long droughts.

In contrast, precipitation droughts do not differ much between both threshold methods (Figure 7(a) and (b)). This is probably due to the generally flashy behaviour of precipitation, leading to sharp rises and descents in the hydrograph. This way, applying another type of threshold does not produce large differences in drought duration. In addition, precipitation DFDs are rather homogenous across all catchments (Figure 5). A possible explanation for this is the rather uniform climate conditions across...
Germany; presumably, there would be a higher variability if data were taken from a global dataset in contrasting climates, especially when considering climates with distinct seasonal precipitation patterns.

Furthermore, all DFDs are generally right-skewed. Skewness increases when the variable instead of the constant threshold is applied. This right-shift can be explained with the correspondence of the variable threshold with the overall seasonal amplitude. Any hydrological variable with inherent seasonality fluctuates around its seasonal amplitude, which corresponds to the variable threshold, thus more frequent threshold crossings produce shorter droughts in higher numbers, resulting in a higher right-skew in the associated DFD when the variable threshold is applied. All of the above-summarized patterns in the DFD statistics sustain for the threshold level height tested in this study (10th, 20th, 30th percentile), while overall drought numbers generally decreasing when moving from 50th via the 20th to the 10th percentile.

By means of the synoptic assessment of the normalised hydrographs (Figures 8–11), drought propagation patterns like pooling, lag and lengthening are comparatively stable in lowland catchments like the Lachte and Oertze (Figures 8 and 9) when another type of threshold is applied. In mountainous catchments like Salz and Spree, however, the differences between the two threshold level methods are more pronounced (Figures 10 and 11). As was shown, this was mainly due to streamflow, which crossed the respective threshold at completely different times and magnitudes when one or the other Threshold Level Method was applied. For example, in the Salz, we looked at a winter drought (Figure 10). Winter is the peak of the wet season in Germany, which obviously coincides with a peak in the variable threshold, thus the constant and variable threshold are far apart from each other during that time. This is true for the lowland catchments as well, but in the mountainous regions we additionally do not have reliable baseflow during that time, thus under precipitation deficits, the streamflow hydrograph can drop fast, which can result in very severe droughts with the variable threshold, though not with the constant threshold, which is much lower during that time. In contrast, in summer (see Figure 11), due to very low flows in general, we have the situation that quite small changes in the threshold can produce big differences in drought characteristics, simply because the range of values during this low flow season is narrow and often close to 0. Thus, in both cases the more ambivalent behaviour that we observed in mountainous catchments is probably due to a combination of a higher seasonal contrast and a lower baseflow.

Groundwater was ambivalent as well, up to the degree that slow and fast reacting patterns could not be grouped according to classical catchment controls, e.g. lowland and mountainous topography, which typically distinguish large, fine grained (thus slow reacting) aquifers from smaller, coarse grained (thus fast reacting) aquifers, respectively. Instead, small-scale, within-catchment (or within-aquifer) variability plays a role. A good case for this point can be made by the Lachte and Oertze catchments. Even though both have near-identical hydrogeological settings, the Lachte groundwater gauge, being located in a groundwater recharge area next to a small stream tributary, experiences few and long droughts (Table 3, Figure 5). On the contrary, the Oertze gauge, located within the riverine valley 870 metres from the main river branch, i.e. in a groundwater transition zone, experiences many short droughts (Table 3, Figure 5). This is in agreement with Peters et al. (2006), who found that for the Pang catchment in the UK, droughts in groundwater head near the water divide were much more persistent than those near the stream, and points to the problem concerning representativeness of key boreholes as discussed further below. However, contrary to their explanation of faster aquifer response close to the stream, we hypothesise that it is more likely an effect of convergence. The Oertze gauge is located near the stream and thus has a larger ‘catchment’, i.e. experiences a high degree of accumulation from converging recharge areas (or rather aquifer volumes). In consequence, it shows a high seasonal regularity (Figure A15, available with the online version of this paper) due to steady inflow from these areas. The seasonal amplitude reliably exceeds the threshold during the wet season, thus no multi-year droughts occur in this location. The seasonal amplitude in the Lachte gauge is much less reliable, instead supra-annual patterns with cycles on longer timescales dominate the overall dynamics (Figure A15), coinciding with a twice as high Coefficient of Variation (0.66 for Lachte, 0.32 for Oertze). This is, hypothetically, because it receives a smaller portion of steady water inflow from ‘upstream’ converging aquifer...
volumes, and has therefore a higher vulnerability to climate fluctuations, which ultimately results in longer (multi-year) drought durations. Therefore, it is hypothesised that the main control here is a groundwater gauge’s degree of convergence, which is in turn dependent on its location relative to areas of groundwater recharge, transition, and discharge within an aquifer.

Agreement with previous modelling studies

The second objective of this study was to analyse observational data to evaluate possible agreement with those studies on the subject of drought propagation, on which the theoretical concepts were developed, and which are based almost exclusively on modelling experiments (Peters et al. 2003, 2006; Tallaksen et al. 2009; Di Domenico et al. 2010; Vidal et al. 2010; Van Loon & Van Lanen 2012; Van Loon et al. 2012). Only Eltahir & Yeh (1999) and Fendekova et al. (2011) took advantage of observational data, mostly. A further comparison with modelling results would require a model for each of the study catchments which was beyond the project’s scope. Instead, we discuss the reproducibility of previously found forms of drought characteristic distributions and theoretical drought propagation patterns in an observational framework as presented in this paper.

The DFDs that were analysed show, in general, classical drought propagation patterns with decreasing numbers and increasing durations when moving from precipitation via streamflow to groundwater. This is in line with previous drought propagation studies (e.g. Tallaksen et al. 2006, 2009). Also, as already pointed out, drought propagation patterns like pooling, lag and lengthening of droughts are comparatively stable in lowland catchments like the Lachte and Oertze (Figures 8 and 9) when the other type of threshold is applied. Results in these areas are therefore in good agreement with previous studies as well (e.g. Van Loon & Van Lanen 2012; Van Loon et al. 2012; Van Loon 2013), while in mountainous catchments like Salz and Spree, the differences between the two threshold level methods are more pronounced (Figures 10 and 11). As pointed out, this was mainly due to a more ambivalent behaviour of streamflow due to less reliable baseflow conditions.

Parallel to the explanation of a more ambivalent streamflow in mountainous settings, an additional problem is posed by the key borehole approach. The modelling frameworks generally reproduce a catchments storage as a whole, while with the key borehole approach, as used in this study, data from few or only one borehole are taken as a representative for the aquifer. This approach has disadvantages since drought characteristics are highly variable within an aquifer (Peters et al. 2006). Still, the key borehole approach was a necessity because the availability of groundwater gauges is scarce and groundwater modelling is time and cost intensive and is subject to many assumptions. Also, we did not have access to data from dense gauging grids. In any case, the conclusions made with data from key boreholes have limited explanatory power for aquifers in its entirety, since evidence exists that the ability to adequately characterise an aquifer’s behaviour in its entirety on grounds of only one or few gauges is disputable (Peters et al. 2006). Thereby, the explanatory power is lower in the mountainous gauges (Salz and Spree), which represent only a subset of multi-aquifer storage systems that consists of a variety of smaller storage units of different size, texture etc. The explanatory power is higher for the lowland gauges (Lachte and Oertze) due to flat and rather homogenous aquifer conditions, which is more representative concerning the whole groundwater storage in the respective catchment.

CONCLUSIONS

The study has shown the characteristic effects of the two investigated threshold level methods. Those are, under the variable relative to the constant threshold: (1) a substantial increase in short droughts, (2) a moderate decrease in intermediate droughts and (3) a minor increase in long droughts. This holds for streamflow and groundwater, but not precipitation. Also, all DFDs are right-skewed under the constant threshold and skewness increases even more under the variable threshold. These patterns generally sustain in DFD statistics for the different thresholds considered (10th, 20th and 30th percentile). Additionally it was shown how different hydrogeological settings can impact propagation patterns. Most prominently, it was found that the two lowland groundwater gauges showed significantly different response patterns, hypothetically due to different degrees
of convergence, i.e. the amount of aggregated water they receive.

Still, it is the lowland setting where the main results of previous studies on drought propagation can be generally confirmed, while in the mountainous catchments, theoretical drought propagation patterns were less often true, mainly due to lower baseflow. This study gives insight how, especially in the mountainous catchments, while different percentiles for the threshold did not change drought propagation patterns in a major way, the drought propagation patterns can be instable across threshold methods since drought characteristics like onset, intensity and duration of droughts change significantly when another type of threshold is applied. This happens up to the extent that individual droughts disappear completely or emerge at entirely different times. The effect is most pronounced in streamflow, less extreme in groundwater and least extreme in precipitation. The differences in the occurrence of actual events are to some extent averaged out in the statistics, such as the DFDs we studied, where drought propagation patterns were more stable, even when different percentiles were chosen as threshold levels. Still, this points to the risk of diverging inference that can be drawn from the same data, especially in hydrograph discussion, depending on which type of threshold, constant or varying, is applied. Thus, in any application, our results suggest that the distinct advantages and disadvantages of each method have to be weighed carefully because they can lead to diverging inferences from the results.

ACKNOWLEDGEMENTS

This study was supported through a scholarship from the Rosa Luxemburg Foundation to the first author. The authors would like to thank the three reviewers, who provided very helpful suggestions and added quality to the study. Streamflow and groundwater data were collected by Kohn et al. (2014) with special thanks to the following federal and regional state authorities of Germany, who provided the data: LfU, DWD, HLUG, LANUV, LFULG, LLUR, LUG, LHW, LKNM, NLWKN, StALU-MS, StALU-VP, StALU-WM and TLUG. For the precipitation data, we further acknowledge the E-OBs dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metaoffice.com) and the data providers in the ECA&D project (www.ecad.eu).

REFERENCES

BGR & SGD 201b HÜK200: Hydrologische Karte von Deutschland 1:200.000. Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, Germany.


Sivapalan, M. 2009 The secret to ‘doing better hydrological science’: change the question! Hydrol. Process. 23 (9), 1391–1396.

Stahl, K. 2001 Hydrological Drought – A Study across Europe. PhD, Albert-Ludwigs-Universität, Freiburg, Germany.


First received 10 December 2015; accepted in revised form 27 July 2016. Available online 30 September 2016