On the spatio-temporal analysis of hydrological droughts from global hydrological models

G. A. Corzo Perez¹, M. H. J. van Huijgevoort¹, F. Voß², and H. A. J. van Lanen¹

¹Hydrology and Quantitative Water Management Group, Centre for Water and Climate, Wageningen University, Droevendaalsesteeg 4, 6708 PB Wageningen, The Netherlands
²Center for Environmental Systems Research, University of Kassel, Kassel, Germany

Received: 23 November 2010 – Published in Hydrol. Earth Syst. Sci. Discuss.: 19 January 2011
Revised: 24 August 2011 – Accepted: 24 August 2011 – Published: 19 September 2011

Abstract. The recent concerns for world-wide extreme events related to climate change have motivated the development of large scale models that simulate the global water cycle. In this context, analysis of hydrological extremes is important and requires the adaptation of identification methods used for river basin models. This paper presents two methodologies that extend the tools to analyze spatio-temporal drought development and characteristics using large scale gridded time series of hydrometeorological data. The methodologies are classified as non-contiguous and contiguous drought area analyses (i.e. NCDA and CDA). The NCDA presents time series of percentages of areas in drought at the global scale and for pre-defined regions of known hydroclimatology. The CDA is introduced as a complementary method that generates information on the spatial coherence of drought events at the global scale. Spatial drought events are found through CDA by clustering patterns (contiguous areas). In this study the global hydrological model WaterGAP was used to illustrate the methodology development. Global gridded time series of subsurface runoff (resolution 0.5°) simulated with the WaterGAP model from land points were used. The NCDA and CDA were developed to identify drought events in runoff. The percentages of area in drought calculated with both methods show complementary information on the spatial and temporal events for the last decades of the 20th century. The NCDA provides relevant information on the average number of droughts, duration and severity (deficit volume) for pre-defined regions (globe, 2 selected hydroclimatic regions). Additionally, the CDA provides information on the number of spatially linked areas in drought, maximum spatial event and their geographic location on the globe. Some results capture the overall spatio-temporal drought extremes over the last decades of the 20th century. Events like the El Niño Southern Oscillation (ENSO) in South America and the pan-European drought in 1976 appeared clearly in both analyses. The methodologies introduced provide an important basis for the global characterization of droughts, model intercomparison of drought identified from global hydrological models and spatial event analyses.

1 Introduction

Drought is defined as a “sustained and regionally extensive occurrence of below average water availability” (Tallaksen and van Lanen, 2004). It is triggered by low or lack of rainfall, often in combination with high evaporation rates. In regions with a cold climate, deviations from normal temperatures can also give rise to drought due to early snow accumulation or late snow melt (winter drought, van Loon et al., 2010). Different types of drought can be distinguished, i.e. meteorological drought, soil moisture drought and hydrological drought (e.g. groundwater storage, river flow, lake storage) (e.g. Tallaksen and van Lanen, 2004; Wilhite, 2000; Heim Jr, 2002). Droughts have large socio-economic and environmental impacts affecting many sectors. Between 1991 and 2005, 950 million people were affected by droughts worldwide and an economic damage of 100 billion US dollars was reported (UN-ISDR, 2009). Data for Europe from 2000–2006 show that each year on average 15 % of the EU total area and 17 % of the EU total population have suffered from the impact of droughts. The estimated total costs
of droughts for Europe over the past 30 years amount to 100 billion Euros (EC, 2007). Climate change projections indicate continental drying and a likely associated increase in dry spell length and frequency of drought over, for instance, many mid-latitude continental interiors that will lead to significant impacts (Bates et al., 2008). There is an urgent need to adapt to the negative impacts of drought, but knowledge is still limited, in particular on how large scale weather phenomena can be linked to hydrological drought that typically has a higher spatial variability (e.g. Fleig et al., 2010). This requires an adequate description of the spatio-temporal analysis of the different drought types, including how meteorological drought converts into hydrological drought (e.g. Peters et al., 2006; Tallaksen et al., 2009) for rather large areas (e.g. regions, continents or globe).

Droughts are derived preferably from observations (e.g. Lins and Slack, 1999; Douglas et al., 2000; Hisdal et al., 2001; Zhang et al., 2001; Stahl et al., 2010), but data availability is limited (e.g. Hannah et al., 2010). Combined observational-modeling frameworks are implemented for that reason. Global models that integrate the interaction of atmosphere-land (General Circulation Models, GCMs) are preferred, but their spatial scale is still too coarse to simulate sufficiently reliable land surface processes, including hydrological extremes. They are limited to an assessment of the average annual runoff (e.g. Milly et al., 2005). As an alternative, off-line approaches have been developed over the last decades, which include Global Hydrological Models (GHMs) and Land Surface Models (LSMs) that simulate the global and continental terrestrial water cycle (e.g. Liang et al., 1994; Döll et al., 1999). The GHMs and LSMS operate at a more detailed scale than the GCMs, which allows a better representation of the hydrological processes at the land surface, which likely will lead to a better identification of hydrological extremes. These models are forced with global reanalysis meteorological datasets (e.g. Sheffield and Wood, 2008) to simulate the past water cycle or with downscaled, bias-corrected future meteorological data derived from GCMs to generate possible future water cycles.

Recently, six LSMs and five GHMs participated in a model intercomparison project (WaterMIP) that compares simulation results of these models in a consistent way. All models were run at 0.5° spatial resolution for the global land areas for a 15 year period (1985–1999) using a newly-developed global meteorological dataset (Weedon et al., 2010, 2011). Haddeland et al. (2011) describe the intercomparison setup and the first results of the multi-model global water balance. Most of these models also provide simulated gridded time series of daily hydrological variables for the period 1963–2001. A detailed analysis of the model outcome allows to improve our understanding of how hydrological droughts evolve at large scales. For this, in addition to the normal statistical analysis to obtain hydrological regimes, it is required to elaborate methodologies to be able to inter-compare the spatio-temporal development of global droughts among the models.

The methodology presented in this paper is a first step in the intercomparison of a suite of GHMs and LSMS (Haddeland et al., 2011) among each other and against other evidence to test their ability to identify large scale, space time patterns in historical hydrological droughts. Various approaches have been proposed to describe the spatio-temporal development of drought. Peters et al. (2006) and Tallaksen et al. (2009) propose methodologies, which they apply on the river basin scale to explore the area in drought and drought propagation in different gridded hydrometeorological variables simulated with a river basin model forced with local weather data (1960–1997). They did not consider spatially connected grid cells in drought, which here is further elaborated as a similar method, i.e. non-contiguous drought areas (NCDA) approach. On the continental scale Andreadis et al. (2005) studied drought in soil moisture and runoff simulated with the VIC land surface model and forced with a continental meteorological dataset (1920–2003) to assess area in drought, but additionally to identify spatially-connected grid cells in drought and next to obtain severity-area-duration curves for the United States. Sheffield et al. (2009) applied the VIC model using a global meteorological forcing dataset (1950–2000) to simulate soil moisture globally. Drought characteristics (e.g. frequency, duration, severity) were derived from time series of simulated gridded data. Spatial extent of soil moisture drought was calculated for the whole globe and for 20 areas covering the world. Sheffield et al. (2009) also investigated spatially connected regions of soil moisture drought for each continent and globally to calculate severity-area-duration curves. This type of spatial drought analysis cannot clearly identify and characterize specific spatial events for a particular time step. The method proposed here is based on clustering of patterns at each time step, named contiguous drought area (CDA). Spatio-temporal analysis of drought at large scales, in particular for hydrological drought, is a relatively new concept that is still under development.

This paper aims at further development of methodologies describing the spatio-temporal development of hydrological drought (i.e. runoff) on a global scale (non-contiguous and contiguous approaches). The methodologies are illustrated, as an example, with the outcome from the global hydrological model WaterGAP (Water – Global Analysis and Prognosis, Alcamo et al., 2003), which has been widely tested in research studies (Döll and Lehner, 2002; Döll et al., 2003). The results presented in this paper are the basis for further studies on the field of global hydrological drought analysis and multi-model comparisons.

2 Spatio-temporal drought methodology

Spatial and temporal information in a drought analysis can be studied in different ways (e.g. Tallaksen et al., 2009; Hisdal et al., 2004; Hannaford et al., 2010). The particular definition
of drought in this study differs from space to time, therefore an integrated drought concept is proposed. From a time series analytical point of view the definition of drought is interpreted as an anomaly of low values in the time series of the hydrological variable in a cell. On the other hand, in space anomalies of low values of a hydrological variable can be interpreted as a drought region. Therefore, the possible approaches can be represented as in Fig. 1.

- Option 1: Analysis of drought is performed by identifying anomalies determined and characterized by time series statistics (e.g. percentiles). Overall statistical information of the events are grouped and then the spatial analysis is done. This can be done by investigating the overall statistical information of a group of cells, or by spatially smoothing the statistical information.

- Option 2: Analysis of drought regions in one time frame (e.g. particular day, month or year) is performed (no knowledge from past or future time frames is assumed). Regions that have anomalies of water flow or levels are clustered (grouped) and labelled (classified) as drought. This cluster procedure is evaluated at different time frames and matching patterns of the anomalies of the regions are compared.

- Option 3: First the time dimension is considered as in Option 1, however, a second analysis is done per each time step. In evaluating the spatial regions of the drought at each time step and identifying contiguous areas it is possible to visualize drought development.

The main objective of these three options is to cover possible approaches to develop drought area analysis. These three options can be reflected in (i) overall cluster of statistics, (ii) Non-Contiguous Drought Area (NCDA) analyses, and (iii) the Contiguous Drought Area (CDA) analyses. The latter two methodologies are presented here to investigate hydrological drought at large scales. The NCDA analysis in this study focuses on the globe as a whole and on some selected large hydro-climatic regions. Additionally, the CDA analysis focuses on linked cells that have neighbours with the same conditions (e.g. drought state) and therefore can be clustered and labelled as spatial drought event.

The detailed procedure of the analysis can be divided into four steps as shown in Fig. 2. Although step 3 is not required for the CDA method, it is an important analysis that helps with the overall identification and location of drought. Hereafter we present the temporal drought analysis (steps 1 and 2) followed by the NCDA and CDA methods (steps 3 and 4).

2.1 Drought analysis in the time series domain (steps 1 and 2)

To determine droughts from the modelling results the threshold level method (Yevjevich, 1967; Hisdal et al., 2004) is applied. With this method, a drought occurs when the variable of interest (e.g. precipitation, soil moisture, groundwater storage, or discharge) is below a predefined threshold (Fig. 3). The start of a drought event is indicated by the point in time when the variable falls below the threshold and the event continues until the threshold is exceeded again. Hence, each drought event can be characterised by its beginning, end and duration. Other commonly used drought characteristics are: deficit volume, calculated by summing up the differences between actual flow and the threshold level over the drought period, and minimum flow during an event (Hisdal et al., 2004). Both a fixed and a variable (seasonal, monthly, or daily) threshold can be used. In this study, a monthly threshold derived from the frequency duration curves of daily data are taken. For cells with a perennial runoff relatively low thresholds in the range from the 70 to 95-percentile can be considered reasonable (e.g. Hisdal et al., 2004; Fleig et al., 2006). In this study the 80-percentile is selected (e.g. Hisdal et al., 2001; Andreadis et al., 2005; Tallaksen et al., 2009; Wong et al., 2011), meaning that subsurface runoff values are exceeded 80% of the time. Although it is possible to think on more fuzzy bounds for the drought identification of a cell, this is not contemplated in this study. For cells with an intermittent or ephemeral runoff having a majority of zero flow, the 80-percentile could easily be zero in one or more months, and hence no drought events would be selected for these months. For these dry regions a higher threshold (low percentile) can be chosen as proposed by Fleig et al. (2006), or the region can be excluded because drought in streamflow is not so meaningful. The latter has been done in this study, although the proposed methodology can handle higher thresholds. Dry regions were excluded from the drought analysis when in more than 80% of the time the simulated runoff is zero. If the percentage of zero flow is lower than 80%, but the time series presents one single event, in the period 1963–2001, which appears in arid regions, it is also excluded. The discrete monthly threshold values are smoothly by applying a moving average of 30 days (e.g. van Loon et al., 2010). The smoothed threshold (T) and the hydrological variable are shown in Fig. 3. It is important to highlight that minor droughts are removed. Several methods have been proposed in the literature depending on dynamics of the hydrological variable (e.g. Fleig et al., 2006). The proposed methodology is flexible and different approaches and values can be used. In this study minor droughts with a duration smaller than 3 days (10% of a month) are excluded.

Fig. 1. Diagram of the steps in a temporal and spatial analysis.
The detailed procedure of the analysis can be divided into four steps as shown in Fig. 2. Although step 3 is not required for the CDA method it is an important analysis that helps with the overall representation of the drought region. Hereafter we present the temporal drought analysis (steps 1 and 2) followed by the NCDA and CDA methods (steps 3 and 4).

![Diagram of temporal and spatial analysis interaction](image)

**Fig. 2.** Diagram of temporal and spatial analysis interaction.

From the identification of drought events using the variable threshold method we extract the drought characteristics of each cell. The drought characteristics that fit our purpose are the average drought duration, average drought deficit volume and number of events. These drought characteristics are calculated as follows (similar to Sheffield et al., 2009; Tallaksen et al., 2009).

The drought duration (DD) is determined by the period of time the variable is under the drought threshold. DD is the total drought duration at cell \(c\) of a particular event.

\[
ADD = \frac{\sum_{i=1}^{N} DD}{N}
\]  

(1)

where ADD is the average drought duration at cell \(c\) (\(N\) is the number of droughts per cell).

The deficit volume (DV) is calculated by accumulating the daily deficit \((X - T)\) as visualized in Fig. 3. \(X\) represents the value of the simulated variable (in this study the subsurface runoff \(Q_{sb}\)).

\[
DV_i = \sum_{t=t_s}^{t_e} (X_t - T_t)
\]  

(2)

where \(DV_i\) is the deficit volume of the event \(i\). The initial and final time steps are represented by \(t_s\) and \(t_e\) respectively. \(T_t\) is the threshold per time step.

\[
RADV = \frac{\sum_{i=1}^{N} DV_i}{N} / \text{std}(X)
\]  

(3)

where RADV is the relative average deficit volume at cell \(c\). The division by the standard deviation (\(\text{std}(X)\)) is implemented since in the spatial analyses it is required to compare relative (standardized) and not absolute values.

\[
ADI = \frac{\sum_{i=1}^{N} DV_i}{N} / \text{ADD}(X)
\]  

(4)

where ADI is the Average Drought Intensity per cell.

**Fig. 3.** Drought characteristics calculated with variable threshold: (a) Hydrological variable and variable threshold, (b) daily intensity, (c) event deficit volume. Start and end of the events are named as 1 and 2.
2.2 Spatial methodology

The two proposed methodologies, NCDA and CDA can be explained as follows.

2.2.1 Non-contiguous drought areas (NCDA, step 3)

The spatial analysis presented here is based on the occurrence of an event (binary representation). Therefore, a discrete drought state \( D_s \) per cell was used. This can be expressed by a function (Eq. 5).

\[
D_s(t) = \begin{cases} 
1 & \text{if } X_t < T_t \\
0 & \text{if } X_t \geq T_t 
\end{cases}
\]  

(5)

where the drought state \( D_s \) per cell at time step \( t \) is determined by the hydrological simulated variable \( X \) at each time \( t \).

\[
PDA(t) = \frac{100}{A_{tot}} \cdot \sum_{c=1}^{N} (D_s(t) \cdot A)
\]

(6)

where PDA is the percentage of area in drought, at time \( t \) and relative to \( A_{tot} \) (total land area considered, e.g. the globe, climatic region, continent). The area at each cell \( A \) in this study was calculated according to its projection, \( N \) is number of cells. These equations were applied to global and selected hydro-climatic regions as follows.

- Global spatial analysis of droughts: the spatial analysis is performed by applying Eqs. (5) and (6) to all cells across the globe at each time step. The result provides an overall assessment of the area in drought (% of global land area) as well as the critical periods (months or years) where this percentage is higher than normal.

- Regional drought spatial analysis (continents or climatic regions): in this study we use both continents and climatic regions. The spatial analysis of climatic regions is based on a division of the study space (globe) into regions that are defined by climate criteria. This is achieved by grouping cells that have similar climate properties. In this study we used the Köppen-Geiger classification (Köppe et al., 2006; Peel et al., 2007; Wanders et al., 2010). Wanders et al. (2010) applied the Köppen-Geiger classification rules that define each climate to the WATCH Forcing Data (Weedon et al., 2010, 2011) to obtain for each cell a classification. Cells that belong to the same Köppen-Geiger sub-climate were clustered. The two regions selected (Af, Cfb) in this study are highlighted in Fig. 4. The “Af” is a tropical rainforest climate that is characterized by constant high temperatures all the year and all months have average precipitation of at least 60 mm. The clusters are mainly located in the Amazonian and Indonesian regions. The “Cfb” is a humid temperate climate, which can be found in major parts of Europe, Eastern USA, south of South America, Eastern China, Eastern Australia and in some smaller regions.

2.2.2 Contiguous drought areas (CDA, step 4)

The NCDA provides the overall temporal evolution of the area in drought in a predefined region (e.g. globe, continent, hydro-climatic region). However, any of these predefined regions are arbitrary reference regions, they might separate spatially-continuous drought events (except for the case of the globe as reference region). Contiguous areas of drought (CDA) reflect spatial drought patterns composed by linked neighbours in a state of drought, classified here as spatial drought event. The binary representation created with the drought states (Eq. 5), from the time series analysis, allows for the application of pattern recognition techniques that group different cells based on the characteristics from their neighbours. This in general can be described as a clustering based on its neighbour conditions (e.g. connectivity).

For more reference on similar type of approaches please refer to Sheffield et al. (2009) who identified drought events in time and space using another clustering algorithm, and Andreas et al. (2005) and Zaidman et al. (2002) who presented
severity-area-duration curves with a partitioning, smoothing and filtering process for the clustering in the spatial dimension. The method used within this study can be described as follows.

1. Scan all cells of the data first by column then by row (Fig. 5).

2. Evaluate the state of the 8 (or 4 neighbours) and assign a preliminary class (spatial drought label) to the cell. Assignment is done as follows.

   - If actual position of the cell is in drought then look for a neighbour in drought.
   - If a neighbour is in drought and has a label (class) then change the label of the actual cell to the label of the neighbour.
   - If there is more than one label between the neighbours assign the lowest one and store the equivalence in the equivalence table.
   - Else assign a new label number.

3. Resolve the table of equivalence classes leaving linked classes as one class.

4. Make a second iteration from column cells then row and relabel based on the resolved equivalence classes.

The algorithm can be applied using different numbers of neighbours, for the present study we used 8 neighbours (DillenCourt et al., 1992).

The number of spatial events is reduced applying similar criteria as other previous studies (Sheffield et al., 2009; Talaksen et al., 2009). An areal threshold is defined with a minimum area required to have a spatial event (in this study the areal threshold was 2 cells, approx. a minimum of 5000 km²). Therefore in this study we excluded spatial events with an area below or equal to 2 connected cells in drought. This rather low areal threshold still allows for identifying spatial droughts in small units of the selected Köppen-Geiger regions.

3 WaterGAP model simulation

The gridded outcome from the global hydrological model WaterGAP was used in this study to further develop NCDA and CDA approaches for hydrological drought and to illustrate these across large scales. The global hydrological modeling system WaterGAP has been widely used on exploring the distribution and availability of water resources. It has been tested against river flow data and has shown to be a powerful tool to simulate the global water cycle. WaterGAP combines a global hydrological model (Alcamo et al., 2003; Döll et al., 2003) with several global water use models (Flörke et al., 2010; Flörke and Alcamo, 2004; Alcamo et al., 2003; Döll and Siebert, 2002). The WaterGAP Global Hydrology Model calculates surface and subsurface runoff, groundwater recharge and river discharge at a spatial resolution of 0.5° and is well applicable for global assessments related to water security, food security and freshwater ecosystems. To simulate the terrestrial water cycle spatially distributed physiographic information about elevation, slope, hydrogeology, land cover and soil properties, as well as location and extent of lakes, wetlands, and reservoirs is used. In addition climate data like precipitation, temperature and different radiation terms need to be specified (see Table 1). Optionally, water required for consumptive water use can be subtracted from surface waters. In this study only naturalized model simulations (i.e. without taking water management and water uses into account) were utilized (e.g. Haddeland et al., 2011). As a hydrological variable the simulated daily, gridded subsurface runoff was selected, which represents the groundwater component of the hydrological drought. So far, WaterGAP has not been calibrated at a daily timescale, nevertheless we expect this to have no significant effect on the drought analysis, because daily subsurface runoff appears to have no high intra-monthly flow variability for most of the cells. The
Table 1. WaterGAP model setup information.

<table>
<thead>
<tr>
<th>Forcing</th>
<th>Veg balance</th>
<th>Soil data</th>
<th>Soil information used</th>
<th>Lakes/wetlands</th>
<th>Routing network</th>
</tr>
</thead>
<tbody>
<tr>
<td>precipitation, temperature,</td>
<td>global land cover</td>
<td>world dataset</td>
<td>land cover type dependent field</td>
<td>GLWD</td>
<td>DDM30</td>
</tr>
<tr>
<td>shortwave down (incoming short-wave)</td>
<td>characterization (Loveland et al.,</td>
<td>soil properties</td>
<td>capacity; permafrost included</td>
<td>(Lehner and</td>
<td>(Döll and</td>
</tr>
</tbody>
</table>

spatial resolution of the model results for each hydrological variable is 0.5° latitude by longitude, covering land areas defined by the CRU (Climate Research Unit of the University of East Anglia) land mask. Although WaterGAP was run from 1958 to 2001, only outcome from 1963 onwards is used since the model requires a warming period (e.g. 1958–1962). Only land point cells from the CRU mask were considered in this analysis (total land area is 146.7 million km²; distributed over 67 420 cells). The total land area above the equator represents around 78 % of the total land area.

We used the WATCH forcing data (Weedon et al., 2010, 2011) to force the WaterGAP model. The WATCH forcing variables are taken from the ERA-40 reanalysis product of the European Centre for Medium Range Weather Forecasting (ECMWF) as described by Uppala et al. (2005), and are interpolated to 0.5° spatial resolution, including elevation corrections as well as different methods for bias and/or under catch corrections. For detailed information on the forcing variables see Weedon et al. (2010, 2011).

4 Results

The development of the CDA and NCDA methodologies was evaluated for two different time periods. A long record of 38 years of daily time steps starting in 1963 was used for the NCDA. Part of this record, starting from 1976 (25 years) was used for the CDA. The period for the CDA was selected based on the highest number of droughts found in the run of the NCDA analysis. The CDA computational time is rather high and therefore this shorter period was beneficial. Both for the NCDA and the CDA, the same monthly variable thresholds were used for each cell, which were derived from the time series that cover the whole period of 38 years.

4.1 Time series results

Figure 6a–c shows global maps with drought characteristics. Some regions show up as white areas (Fig. 6a) meaning that in 80 % of the time the simulated subsurface is zero (e.g. Sahara desert in Africa). On the other hand it is possible to see regions with a low number of spatial events that are located on non-polar arid lands like, Sonora and Chihuahua, in Mexico and Atacama, in Peru and Chile. This drought behaviour of semi-arid and arid land is a clear outcome of the threshold method, which evaluates the events as a measure of the time when the flow is below a predefined low value. Climatic regions, where rainfall is erratic and low runoff is present for long periods of time, will have less number of events. Figure 6b shows a contrasting picture to what we saw in Fig. 6a. According to the definition low numbers of droughts go along with high average durations and the other way around. The Sahara desert and other extremely arid lands have drought durations of more than 2 years (500–1000 days), and arid lands (see Meigs, 1973, for a detailed definition) durations of around 0.5–1 year (100–300 days).

The distribution of the relative average deficit volume (RADV) is mixed and only spotted values can be seen in north west and middle USA, where high RADV seems to have some relation with average duration of 100 to 300 days (Fig. 6c). This pattern also occurs in the north eastern region of Brazil near the state of Maranhao. Figure 6d shows, as an example, the drought pattern at the beginning of 1976. The drought characteristics presented in Fig. 6 show the importance of further exploring global distribution of drought characteristics with the NCDA and CDA methodology presented in this study. The NCDA global results can be visualized at each time step (Fig. 6d) to further support the analysis of the global drought characteristics (Fig. 6a–c).

4.2 Global results (NCDA)

The Percentage of Drought Area (PDA) for each time step is explored using Eq. (6). Figure 7a shows the evolution of the PDA of the whole globe for the period between 1963 till 2000. From the yearly time series it is possible to derive the bound of maximum and minimum PDA, 30 % and 12 % respectively. The maximum PDA appears at the time series of Brazil near the state of Maranhao.

The color-coded map in Fig. 7b is built using the time series of PDA and a color representation for each percentage. The color-coded table shows some persistence among the years at the global scale (e.g. 1980s). Furthermore, Fig. 7b
Fig. 6. Hydrological drought characteristics derived from the subsurface runoff simulated with the WaterGAP model. (a) Total number of drought events, (b) average drought duration, (c) relative average deficit volume and (d) spatial distribution of drought events for 13 January 1976.
shows a low percentage of area in drought between the beginning of 1974 to the middle of the year 1976, as well as from the beginning of 1978 to first months of 1980. On the other hand, higher percentages can be seen from April till July 1982 and from February till August 1992. The highest percentage in 1992, seems to be not clearly preceded by other dry years. Other high PDAs can be seen in 1998 where the beginning of the high percentage appears at the end of 1997. Most of the historical high PDAs occur between calendar day 75 to 150 (April–May). This seems to happen in spring and the early summer for the Northern Hemisphere and transition summer–autumn in the Southern Hemisphere. The last 130 days of most years, the time series seems to be stable and no important variations can be seen from the beginning (day 230) to the end (day 360).

As mentioned above, the global PDAs show a seasonal behaviour, in particular between days 75 and 150, which seems to be associated with the transition of seasons, often together with high climate variability. To explore such phenomena, the PDAs for six major Köppen-Geiger hydro-climatic regions (Fig. 8) over the years 1991–1993 were plotted. The selected regions include a wide range of climates,
Fig. 9. Daily time series of percentage of area in drought for each year in the period 1963 to 2001 in Af and Cfb climates. (a) Areas with tropical rainforest climate (Af) and (b) Areas with maritime, temperate climate (Cfb).

i.e. warm equatorial with dry winter (Aw) or fully humid (Af), arid hot desert (BWh), temperate with warm summer (Cfb), snow-affected with a warm summer (Dfb) or cool summer with cold winter (Dfc). The selected hydro-climatic regions cover about 35% of the land surface of the globe, implying that these have a substantial influence on the global PDAs (Fig. 7a). Although, there is a clear variability in the lower frequency components of the time series the high frequency components dominate. Three climates have a clear seasonality (Aw, Dfb, Dfc). The Dfb climate has a periodic high PDA around day 90 of each year, when the cold season with snow enters the warm summer (Fig. 8e). The Dfc climate has two PDA peaks around days 90–120, which reflect the transition from the cold winter with snow into the cool summer (Fig. 8f). The Aw climate has less pronounced higher PDAs around day 150, which seem to be linked to the start of the monsoon (Fig. 8b). The climates with a clear seasonality have a sharp rise in the runoff either due to snowmelt or onset of the monsoon. Late snowmelt or arrival of the monsoon, which will not occur synchronously in the spatially-distributed units of a hydro-climatic region over the globe, leads to drought (anomaly), which in most year is short lived because the delayed snowmelt or monsoon will stop the drought (e.g. van Loon et al., 2011). Clearly, the definition of the drought through the daily smoothed monthly thresholds (Sect. 2.1) influences the occurrence of anomalies. Drought anomalies clearly linked to a particular time of the year (i.e. day 90–150) because of seasonal climates cause regular patterns in the global PDAs. These patterns need to be superimposed on the patterns of other climates. For example, the Cfb climate has irregular peaks at the end of the winter (e.g. 1992/1993), which likely coincide with low-rainfall winters (Fig. 8c). The fully humid equatorial climate (Af) and the hot desert climate (BWh) have in common that below average rainfall is not linked to seasons, which gives very unpredictable anomalies in runoff (Figs. 8a and d). The combination of more regular anomalies in seasonal climates (e.g. Aw, Dfb, Dfc) and rather unpredictable timing of anomalies in other climates (e.g. Af, BWh) results in the composite pattern shown in Fig. 7b. Furthermore, the unequal distribution of the snow-affected climates (e.g. Dfb, Dfc) over the globe, which predominantly occur in the Northern Hemisphere, cause patterns that show up only in the first part of the year (days: 75–150) and not in the second part.

4.3 Köppen-Geiger hydro-climatic regions (NCDA)

Based on the previous analysis the regions which we call unpredictable behaviour are the Af and the Cfb climates. Percentages of drought areas for the whole period (1963–2001) were calculated for these two hydro-climatic regions to investigate the spatial differences in the areal coverage across the world.

Figure 9a shows the 38 annual series of the PDA for the clusters with a Af climate (Fig. 4). The PDA values are higher than in the overall analysis for the globe (Fig. 7b), reaching a value of 62% at the beginning of 1998 and the end of 1997. This is a region that is well known for its richness in water resources on average, nevertheless, it appears to have a high number of anomalies in its low runoff, which can be explained by the rather large inter annual variability. Aside of this, in the year 1992 the most severe PDA over a whole year is shown.

For the second climate regime (Cfb, Fig. 4), the highest PDA is 36% (Fig. 9b). In this hydro-climatic region, 1992 seems not to have a high area in drought. The year 1976 has the highest PDA for the Cfb region. In the discussion section, the above-mentioned years with high PDAs will be compared with some documented events.
4.4 Spatial drought events (CDA)

In contrast to the NCDA presented in the previous sections, the contiguous approach (CDA) focuses on linking neighbouring drought states. Connected cells in drought \( D_s \) can be interpreted as a spatial drought event. Each spatial drought event (Sect. 2.2.2) is unique for a particular time step, and can be analysed further by exploring its spatial properties. Figure 10a shows, as an example, the results of a clustering analysis for 10 January 1976. For this date more than 800 spatial drought events were found across the globe. To explore the temporal evolution of the number of spatial drought events a color-coded table of 25 years with the number of clusters found on each day was plotted (Fig. 10b). The figure shows similar patterns as the NCDA. The minimum number of spatial events was around 400 and therefore the scale of the color-coded bar starts at 500, which means that events with a rather low number of spatial clusters are excluded. The year 1992 stands out as in the NCDA as the most critical year. Other extreme events, like 1982, 1987 and 1993 also show up in this graphical presentation. This illustrates that the number of spatial drought events seems to be congruent with the global area in drought.

Figure 10a illustrates that the distribution of the spatial drought events can follow regions that are not clearly bounded by hydro-climatic classifications. Spatial drought events can be identified inside climatic or regional (e.g. continents) boundaries or its areas can be in multiple hydro-climatic regions at the same time. This means that for detailed analysis of spatial events it is important to have unbounded regions to avoid separation of spatial events.

The main advantages of the CDA is the possibility to identify: (i) the area of a particular spatial event (number of cells or area), and (ii) location (geo-referenced). This can be done for a specific time step as demonstrated in Fig. 10a. Additionally, in a temporal study it also requires monitoring the evolution of spatial drought events for (long) time series (e.g. Andreadis et al., 2005; Lloyd-Hughes, 2010), as well as to

Fig. 10. Results of the CDA method applied for the analysis of number of drought clusters. (a) Drought clusters for 10 January 1976 (around 800 colors, each color represent a cluster and (b) color-coded table of the number of drought clusters on the whole Earth.
summarize statistically their information. In this study we monitor the maximum area of a spatial drought event and the duration (i.e. number of days in a calendar year) that it occurs somewhere in the world. This was done to keep track of the change in Maximum Spatial Drought Event (MSDE) in terms of its size and relative location at each time step. Figure 11a shows a color-coded table with the timing and the magnitude of the MSDEs for each day within the 25 year record contemplated in this part of the study. MSDEs of above 5000 cells (about 7.5% of the global land area) appear only around 6 times for a maximum of 30 days. Furthermore, we see, although weaker, similar patterns around days 75–150 as for the global area in drought (Fig. 7b). Because of the congruence of area in drought and number of spatial events it is possible that seasonal climates cause these patterns (Sect. 4.2).

The relative location of MSDE is an important indicator of changes in global drought. For this, the occurrence of the MSDE in a continent, is determined by the location of its geometrical centroid. This analysis for continents does not relate to a predefined region, and although a MSDE is counted for a particular continent, it may extend over the continents border. Figure 11b shows how many days the centroid of the MSDE occurs in the different continents. For instance, in 1999 the maximum event was found in Asia for over 300 days and for about 50 days in North America. The maximum area seems to be located predominantly in Asia. Africa has continuous records with events from 1983 till 1985. Exceptional cases in Europe can be highlighted like 1976 which has a high number of days that the maximum spatial drought event occur there (over 60 days).

5 Discussion

In this study the analysis of spatial and temporal droughts from a global hydrological model has been developed using overall statistics as well as the CDA and NCDA methodologies. A number of remarks can help on the understanding of the benefits and weakness of the methodologies. For the characterization of spatial drought events a unique formulation is required, since conventional averaging of spatial regions might not really represent relevant characteristics of the drought. The duration of a spatial event could be defined by the average, longest or shortest duration of the cluster of connected neighbourhood cells. However, to determine the deficit volume of a particular spatial event, it might not be accurate to add straightforwardly deficits. Instead, a weighting of normalized volumes can be calculated as an alternative. Moreover, the spatial extent of the events is changing in time (e.g. Andreadis et al., 2005; Lloyd-Hughes, 2010), as well as its centroid. Therefore a morphological and dynamic analysis needs to be included in the CDA method presented in this paper. In this sense, this paper did not cover the development of drought characteristics of these spatial events (e.g. dynamics of representative drought duration or deficit volume of a cluster).

The duration and deficit volume severity of a spatial drought event, at a particular time, is relevant information for drought management. This is commonly approached using severity-area-duration (SAD) curves (e.g. Sheffield et al., 2009), however, the approximation of this method is limited by the spatial interpolation used on the overall information. Therefore, the CDA could be a complementary methodology that provides a step by step spatial analysis.

Drought characteristics both for the NCDA and CDA methodologies are affected by the definition of drought, especially the choice of threshold, the approach to remove minor droughts (Hisdal et al., 2004; Fleig et al., 2006) and the
methods to deal with zero flows. Smoothing of the hydrograph by applying a 5-day moving average (Fleig et al., 2006; Wong et al., 2011) and implementing multiple threshold are commonly considered for the temporal analysis. A step forward in the analysis of spatial drought events is to explore the temporal and spatial sensitivity of applying different areal thresholds (in this study two cells). If higher areal thresholds are applied (e.g. Sheffield et al., 2009) only large events will be identified, which implicitly means exclusion of small (spotted) regions, i.e. small units of Köppen-Geiger climates. Another important feature is that the CDA algorithm used has a two dimensional neighbouring scheme (latitude and longitude). This scheme only includes regions with 8 contiguous cells in drought, but by increasing the range of neighbours to, for instance, 24 cells it will reduce the sensitivity of the areal threshold by only identifying larger spatial events. Another spatial criterion that requires investigation is the number of cells not in drought that should separate two adjacent spatial drought events. In this study, a single line of cells not in drought, may break one large event into two separate events.

Some of the NCDA and CDA results obtained in this study with a single global hydrological model (WaterGAP) have interesting similarities to reported historical events. The major drought area obtained for the year 1992 correspond to the El Niño Northern Oscillation (ENSO) year. The years 1997 and 1998, which correspond also to the ENSO phenomena (clearly present in Indonesia and South America, Aceituno, 1988; Chokkalingam et al., 2005), show to have a high percentage of drought area in the NCDA (Af region) and a high number of spatial drought events in the CDA.

The European drought in 1976 is clearly elucidated in the CDA and NCDA, and its durations seem to match with previous studies (Hisdal et al., 2004; Hannaford et al., 2010). The Cfb region has an important component in Europe (Fig. 4) and in many studies the 1976 European drought has been studied due to the high economic impact (Stahl, 2001). Although in our study, with the CDA, the maximum spatial event in 1976 is not as high as it was expected (Fig. 9a), that year had a long period of time with the maximum drought event in the globe. In 1991–1992 there was a prolonged drought over the Iberia peninsula (Hisdal et al., 2004). The CDA shows a maximum drought event centroid in a long period of time over US, having a link to the severe extreme drought in over half of the country during 1987–1989 (Fig. 11b). This drought was the subject of national headlines when it resulted in the extensive fires in Yellowstone National Park in 1988 (Sheffield et al., 2009; Andreadis et al., 2005; Shukla and Wood, 2008). It is documented that Africa suffered from droughts in the last quarter of the twentieth century (Prospero and Nees, 1986). The years 1984 and 1985 show to have the largest spatial drought event in Africa, which is highly associated to the drought and famine in East Africa. Ethiopia, usually considered the breadbasket of Eastern Africa, was hit by a extreme drought in the early 1980’s. A dry year in 1981 resulted in low crop yields. This period can be clearly identified in the Fig. 11b. Application of the NCDA to hydro-climatic regions that are affected by phenomena like the ENSO can be useful to test the approach (with single or multi-model), since with the global information (Fig. 6) it is not possible to capture these phenomena. Other regional distributions of the earth, can help on model inter-comparison, since models might perform better or worse according to different hydro-climatic regions.

6 Conclusions

In this study, two concepts of spatio-temporal analysis of large scale drought have been further developed. These concepts are the non-contiguous drought areas (NCDA) and the contiguous drought areas (CDA) approach. The NCDA analysis for the globe and pre-defined hydro-climatic regions builds upon principles used at the river basin scale as proposed by Tallaksen et al. (2009). In a similar way, temporal analysis with a variable threshold followed by spatial assessment was presented (Fig. 6). The CDA presented here is based on the use of an objective and automatic grouping of spatial drought events. This objective determination of spatial drought events uses pattern recognition on binary representations of drought and connected labelling algorithms. The NCDA and CDA analyses used subsurface flow values from the WaterGAP global hydrological model (second part 20th century).

The NCDA is useful for assessing overall spatial drought information at global and regional scales, while the CDA showed specific extreme areas that are unbounded by predefined regions. NCDA results seem to capture important events that might appear to be connected to relevant historical events. The ENSO phenomena and the droughts in Europe (1976 and 1992) show up in different ways in outcome from both methods. With the NCDA a high percentage of area in drought was identified and with the CDA a high area spatial events was found.

The centroid of the maximum spatial drought event in the CDA analysis was used as a measure of extreme events in the different continents (Fig. 11b). Its duration is assumed to reflect drought persistence in a continent. Both the NCDA and CDA investigate drought patterns (Figs. 9a and 11b), which suggest that global drought has its maximum extent in a particular season (March-May). Seasonal climates characterized by irregular snow melt (e.g. Dfb and Dfc climates) or Monsoon are the reason for these rather short anomalies. Dominant occurrence of these climates in the Northern Hemisphere contributes to these seasonal drought patterns.

The CDA shows to be able to identify spatial drought events and at the same time it is possible to follow critical (the most extended) events along time as illustrated for the continents. However, important characteristics, like deficit volume and duration of spatial events, require an analysis to be able to reach indices that characterize their drought
severity. Aside of this, morphological changes and tracking of spatial drought events seems to be the way forward to further explore spatial events and see how well global hydrological models capture historical events. Although the NCDA and CDA methodologies presented here used subsurface flow, these are flexible and can be applied to similar variables and other drought identification approaches, e.g. the Sequent Peak Algorithm (Hisdal et al., 2004). Further research steps are being taken towards the sensitivity of thresholds (e.g. introduction of multiple thresholds), the removal of minor droughts and the treatment of cells with high number of zero flow in a multi-model intercomparison context.

The two methodologies presented here can capture the main spatial temporal characteristics of large scale droughts. The results can help to: (i) understand drought generation as response to climate drivers and physical river basin structures, including possible world-wide synchronicity, and (ii) used to assess the impact of global change on drought (21st Century drought).

Acknowledgement. This research was undertaken as part of the European Union (FP6) funded Integrated Project Water and Global Change (WATCH, contract 036946). The research is part of the programme of the Wageningen Institute for Environment and Climate Research (WIMEK-SENSE). The study contributes to the UNESCO IHP-VII programme (Cross-cutting programme FRIEND and Theme 1 Adapting to the impacts of global changes). The research is part of the European Union (FP6) funded Integrated Project Water and Global Change (W A TCH, contract 036946). The research is part of the European Union (FP6) funded Integrated Project Water and Global Change (WATCH, contract 036946).

References


