Hydrol. Earth Syst. Sci., 10, 535–552, 2006
www.hydrol-earth-syst-sci.net/10/535/2006/

Hydrology and Earth System Sciences

A global evaluation of streamflow drought characteristics

A. K. Fleig[1], **L. M. Tallaksen**[1], **H. Hisdal**[1,2], **and S. Demuth**[3,4]

[1]Department of Geosciences, University of Oslo, Norway
[2]Norwegian Water Resources and Energy Directorate, Oslo, Norway
[3]IHP/HWRP Secretariat, Koblenz, Germany
[4]Institute of Hydrology, Albert-Ludwigs-Universität Freiburg, Germany

Received: 24 October 2005 – Published in Hydrol. Earth Syst. Sci. Discuss.: 18 November 2005
Revised: 6 June 2006 – Accepted: 29 June 2006 – Published: 21 July 2006

Abstract. How drought is characterised depends on the purpose and region of the study and the available data. In case of regional applications or global comparison a standardisation of the methodology to characterise drought is preferable. In this study the threshold level method in combination with three common pooling procedures is applied to daily streamflow series from a wide range of hydrological regimes. Drought deficit characteristics, such as drought duration and deficit volume, are derived, and the methods are evaluated for their applicability for regional studies. Three different pooling procedures are evaluated: the moving-average procedure (MA-procedure), the inter-event time method (IT-method), and the sequent peak algorithm (SPA). The MA-procedure proved to be a flexible approach for the different series, and its parameter, the averaging interval, can easily be optimised for each stream. However, it modifies the discharge series and might introduce dependency between drought events. For the IT-method it is more difficult to find an optimal value for its parameter, the length of the excess period, in particular for flashy streams. The SPA can only be recommended as pooling procedure for the selection of annual maximum series of deficit characteristics and for very low threshold levels to ensure that events occurring shortly after major events are recognized. Furthermore, a frequency analysis of deficit volume and duration is conducted based on partial duration series of drought events. According to extreme value theory, excesses over a certain limit are Generalized Pareto (GP) distributed. It was found that this model indeed performed better than or equally to other distribution models. In general, the GP-model could be used for streams of all regime types. However, for intermittent streams, zero-flow periods should be treated as censored data. For catchments with frost during the winter season, summer and winter droughts have to be analysed separately.

Correspondence to: A. K. Fleig
(a.k.fleig@geo.uio.no)

1 Introduction

Drought is a major natural hazard having severe consequences in regions all over the world. In Europe the drought of 2003 affected 19 countries, and the total costs were estimated to exceed 11.6 billion Euros (EurAqua, 2004). Currently (spring 2006) another severe drought is developing in western Europe. Portugal and Spain are already experiencing their worst drought in 60 and more than 100 years, respectively, and also France and the UK are fearing severe water shortages this summer (EUMETSAT, 2006). In 2003 many different sectors were affected, such as agriculture, forestry, water supply, energy and transport (navigation). The range of drought impacts is related to drought occurring in different stages of the hydrological cycle and usually different types of droughts are distinguished (e.g. Wilhite and Glantz, 1985; Tallaksen and van Lanen, 2004). The origin is a *meteorological drought*, which is defined as a deficit in precipitation. A meteorological drought can develop into a *soil moisture drought*, which may reduce agricultural production and increase the probability of forest fires. It can further develop into a *hydrological drought* defined as a deficit in surface water or groundwater, e.g. reducing water supply for drinking water, irrigation, industrial needs and hydropower production, causing death of fish and hampering navigation. Some authors also define drought in terms of its consequences, e.g. *socio-economical drought* (Wilhite and Glantz, 1985).

A general definition of drought is given by Tallaksen and van Lanen (2004), who define drought as "a sustained and regionally extensive occurrence of below average natural water availability". This definition relative to normal implies that droughts can occur in any hydroclimatological region and at any time of the year. In response to the different impacts of drought in different regions, a large number of quantitative drought characteristics have been developed. Recently published summaries can be found in, e.g. Heim (2002), Hisdal et al. (2004), Smakhtin and Hughes (2004) and Hayes

Published by Copernicus GmbH on behalf of the European Geosciences Union.

(2005). Expressed as a single number, drought characteristics are often referred to as drought indices or drought statistics.

The choice of a suitable drought characteristic for a specific study depends on the hydroclimatology of the region, the type of drought considered, the vulnerability of society and nature in that region, the purpose of the study and the available data. Due to the lack of an unique standard definition, this choice is subjective and a large number of different characteristics are used to describe and quantify droughts. In case of streamflow drought two main approaches of deriving drought characteristics can generally be distinguished (Hisdal et al., 2004). One is to analyse *low flow characteristics* such as a time series of the annual minimum n-day discharge, the mean annual minimum n-day discharge or a percentile from the flow duration curve (FDC). These characteristics describe the low flow part of the regime and characterise droughts according to their magnitude expressed through the discharge (Tallaksen et al., 1997). The development in time of a drought event is not considered. In the second approach, discharge series are viewed as a time dependent process, and the task is to identify the complete drought event from its first day to the last. In this way a series of drought events can be derived from the discharge series, and droughts can be described and quantified by several properties, such as drought duration or deficit volume. These so called *deficit characteristics* are commonly derived by the threshold level method (theory of runs).

As seen in Europe in 2003, however, a single drought event can cover a large region, spanning over different climate zones and affecting various human activities. In order to assess drought events covering large regions and to study the spatial aspects of drought, the definition and the identification of drought have to be consistent throughout large regions. Thus, a standard procedure has to be identified, which is capable to characterize droughts under different hydroclimatological and hydrogeological conditions. This will enhance comparative studies and be beneficial for the development of regional drought monitoring and forecasting systems. It is still advisable to describe the different kinds of drought (meteorological, soil moisture and hydrological drought) by separate drought characteristics, since they do not necessarily occur simultaneously nor exhibit the same severity. Keyantash and Dracup (2002), for example, evaluate the most commonly used drought indices for the different types of drought. For hydrological drought they found the total water deficit derived by the theory of runs superior as compared to the cumulative streamflow anomaly, the Palmer Hydrological Drought Severity Index and the Surface Water Supply Index. For streamflow drought, several methods to derive drought characteristics are compared by Cancelliere et al. (1995). Based on the application to five Italian streams they also suggest the use of the theory of runs, however, applied on a moving average series. Both evaluations are based on applications of the methods and indices to data with a monthly or longer time resolution. However, also droughts shorter than one month can be severe, for example when analyzed with respect to navigation (Cancelliere, 2005) and in a cold climate zone, where the frost free period may last only a few months. In these cases drought characteristics operating on shorter time resolutions are necessary to obtain more detailed information. Streamflow deficit characteristics derived from a daily discharge series are the focus of the present study.

In drought studies design events are used for the construction of water reservoirs, which are one of the most important measures to cope with drought. Hydrological design often requires extrapolation beyond the range of observations and design events can be determined based on an extreme value analysis (Coles, 2001). A number of authors have studied frequencies of extreme streamflow droughts as defined with the threshold level method, e.g. Zelenhasić and Salvai (1987), Clausen and Pearson (1995), Kjeldsen et al. (2000) and Shiau and Shen (2001). An extended review of drought frequency analysis for within-year droughts is given in Tallaksen (2000) and Tallaksen et al. (2004). Two common approaches to select extreme events from a time series are the block maxima (BM) and partial duration series (PDS) approach. In case of BM often the annual maxima is chosen. However, based on a case study in Norway, Engeland et al. (2004) found that for daily streamflow data a block size of at least two years was required for drought deficit volumes to avoid model bias. They therefore recommend the use of PDS as this allows to include a larger amount of events in the series and thus reduces the standard errors in the design event estimates. Still there are often large uncertainties related to the choice of extreme value distribution.

In this study the threshold level method in combination with three common pooling procedures is evaluated for its applicability to daily discharge series for streams in different climate zones and with different hydrological regimes. In addition, the application for regional drought studies including different types of streams is considered. The evaluated pooling procedures are: 1) the inter-event criterion (IT-method), 2) the moving-average filter of n days (MA-procedure) and 3) the sequent peak algorithm (SPA). The methods are applied to a global data set of 15 daily discharge series from a wide range of flow regimes including perennial as well as intermittent streams. Streamflow deficit characteristics, such as deficit volume and duration are derived for all series and the methods are evaluated based on the following criteria:

1. Objectivity: what kind of decisions have to be made by the user of the method?

2. Data requirements: can the method deal with missing data and do any other data considerations need to be made?

3. Reasonableness: can the selected drought events be considered "true" drought events for all series, e.g. are

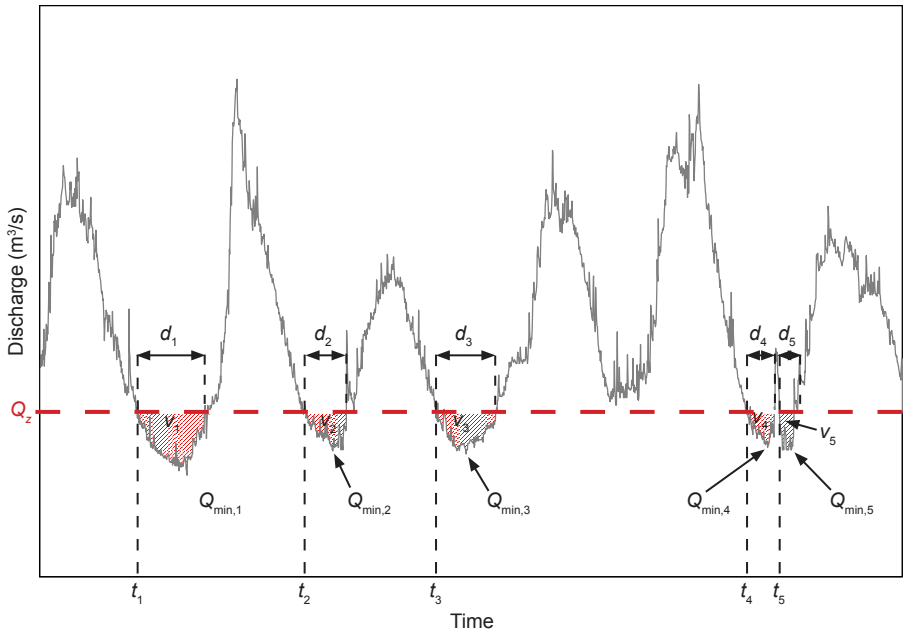

Fig. 1. Illustration of commonly used deficit characteristics as defined with the threshold level method: time of occurrence, t_i, duration, d_i, deficit volume or severity, v_i, and the minimum flow occurring during the drought event, $Q_{\min,i}$.

all major historical events represented, and do the selected events all represent droughts?

4. Robustness with respect to different streamflow regimes: are the results comparable for different kinds of streams?

5. Suitability for a frequency analysis: is the obtained drought series independent and identically distributed?

The evaluation focuses on within-year droughts. The first two criteria are evaluated based on theoretical considerations and criterion 3 on a visual inspection of the drought time series with respect to the discharge series. Criterion 4 is the evaluation of the outcome of criterion 3 for different flow regimes, and criterion 5 is evaluated based on statistical tests and exploratory data analysis. Finally, a frequency analysis of PDS of drought deficit characteristics is conducted, focusing on the choice of extreme value distributions.

The paper starts with a detailed description of the threshold level method and the three pooling procedures considered in this study. Subsequently, the Global Data Set used for the evaluation of the streamflow deficit characteristics is presented, followed by the application and evaluation of the threshold method in combination with the pooling procedures for daily streamflow series, including data considerations and requirements. The results of the frequency analysis of deficit characteristics are then presented. Finally, the main conclusions are summarised.

2 Threshold level method

The threshold level method originates from the theory of runs introduced by Yevjevich (1967), who originally defined droughts as periods during which the water supply does not meet the current water demand. Both the water supply, $S(t)$, as well as the water demand, $D(t)$, are expressed as time series, and a drought event is defined as an uninterrupted sequence of negative values in the supply-minus-demand series, $Y(t)=S(t)-D(t)$. Later, Yevjevich (1983) simplified the concept by applying a constant demand. The demand is represented by a threshold level, Q_z, and droughts are defined as periods during which the discharge is below the threshold level. Common deficit characteristics are the start of the drought, t_i, drought duration, d_i, deficit volume or severity, v_i, and the minimum flow occurring during the drought event, $Q_{\min,i}$, as illustrated in Fig. 1. Additional deficit characteristics can be defined, such as drought intensity, which is the ratio of deficit volume and duration, and recovery time. The latter is defined, e.g. by Correia et al. (1987) as the time it takes to compensate a certain fraction of the deficit volume by excesses of water above a certain recovery level.

In general, the threshold level can either represent a certain water demand, for example for power plants or water supply, or the boundary between normal and unusually low streamflow conditions. The threshold level might be fixed or varying over the year to reflect, e.g. seasonally different water demands. However, not all periods with relatively low flow

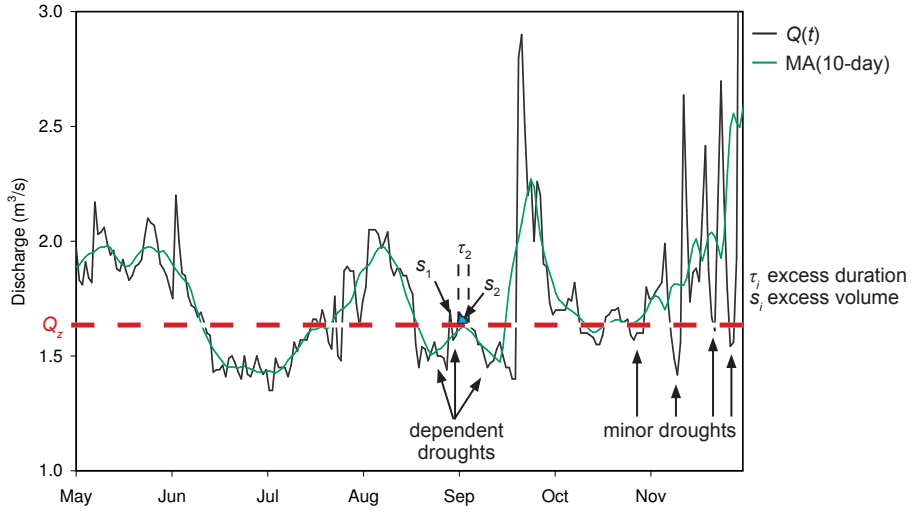

Fig. 2. Illustration of the pooling of mutually dependent droughts and the removal of minor droughts by a MA(10-day)-filter.

compared to a varying threshold are considered a drought, such as relative low flow periods due to a delayed onset of the snowmelt flood. Stahl (2001) and Hisdal et al. (2004) therefore used the terms streamflow deficiency or anomaly when defining deficit periods (periods with discharge below the threshold level) using a varying threshold level.

When the threshold level is set to represent the boundary between normal and unusually low streamflow, it is chosen based on the characteristics of the streamflow regime. In this case low flow indices, such as percentiles from the flow duration curve (FDC), are frequently applied for both perennial and intermittent streams. For regional studies these were found to give more consistent results than percentages of the mean (Tallaksen et al., 1997). Also linear combinations of the mean flow and the standard deviation have been applied for regional studies (e.g. Ben-Zvi, 1987). The choice of threshold level influences both the number of events and the presence of multi-year droughts in the derived drought series. When focus is, as in this study, on within-year droughts neither a large amount of multi-year droughts nor a large number of years without any droughts should be included in the series as these can complicate an extreme value analysis (Tallaksen et al., 1997). The threshold level has to be chosen as a compromise between these two features. For short data series the use of very low threshold levels can be problematic, as the derivation of statistical properties of droughts require a certain minimum number of events. These considerations do not reveal a single preferable threshold level, and its selection, and hence the definition of drought, remains a subjective decision. For perennial streams threshold levels between the 70-percentile flow (Q_{70}) and the 95-percentile flow (Q_{95}) from the FDC are frequently applied, which are the flows that are exceeded 70–95 percent of the time. For intermittent streams lower exceedance percentiles have to be chosen, depending on the percentage of zero flow. For example, Woo and Tarhule (1994) tested threshold levels between Q_5 and Q_{20} and Tate and Freeman (2000) threshold levels up to $Q_{12.5}$ for streams in Nigeria and southern Africa, respectively.

The threshold level method was developed for discharge series with a time resolution of one month or longer, but it has also been applied to daily discharge series, e.g. Zelenhasić and Salvai (1987) and Tallaksen et al. (1997). When the time resolution is short in comparison with the droughts to be studied two problems have to be considered in particular: the occurrence of minor droughts and mutually dependent droughts (Fig. 2). *Minor droughts* are events of short duration and small deficit volume. A high number of minor droughts in the sample may disturb an extreme value analysis and the number of minor droughts should thus be reduced. *Mutually dependent drought* events can occur during a prolonged period of low discharge when short excess periods with discharge above the threshold level divide the period of low discharge into several drought events. When the excess periods are of short duration, τ_i, and small excess volume, s_i, one would generally consider the whole period of low discharge to be one drought event. Short excess periods can be caused by short rainfall events or artificial influences. The split drought events are called *mutually dependent droughts*. They cannot be considered independent of one another, and e.g. for an extreme value analysis it is recommended to combine these into larger independent events. This can be done by so called *pooling procedures*, of which three common ones are described in details in the next section. In a regional study pooling is further recommended due to differences in catchment responses. For example in a slowly responding groundwater-fed catchment a short rainfall event during a prolonged dry period will lead to a much

smaller rise in streamflow as compared to a fast responding neighbouring catchment. As a result a drought might be split in one catchment but not in the other one.

2.1 Pooling of mutually dependent droughts

The *inter-event time method* (IT-method), introduced by Zelenhasić and Salvai (1987), pools drought events based on an inter-event time criterion (IT-criterion). Two mutually dependent droughts are pooled if they occur less than a predefined number of days, t_c, apart, i.e. $\tau_i \leq t_c$. The duration of the pooled drought event, the *full drought duration*, $d_{\text{pool f}}$, is defined to last from the first day of the first pooled event to the last day of the last pooled event, including the excess periods:

$$d_{\text{pool f}} = d_i + d_{i+1} + \tau_i \quad (1)$$

Furthermore, the pooled drought duration without excess periods, the *real drought duration*, $d_{\text{pool r}}$, can be of interest. The total pooled deficit volume, v_{pool}, is defined as the sum of the deficit volumes, v_i, of the pooled drought events:

$$v_{\text{pool}} = v_i + v_{i+1} \quad (2)$$

For studies focusing on, e.g. reservoir management, a more consistent definition would be to subtract the inter-event excess volume, s_i, from the sum of the deficit volumes. As parameter value Tallaksen et al. (1997) recommended t_c=5 days, based on the relationship between t_c and the mean deficit characteristics for two perennial streams in Denmark. However, they used an additional inter-event volume criterion (IV-criterion) for pooling. Zelenhasić and Salvai (1987) recommended t_c=6 days based on their experience with two perennial rivers in former Yugoslavia. This value has also been applied for intermittent streams (e.g. Woo and Tarhule, 1994).

In the *moving-average procedure* (MA-procedure; Tallaksen et al., 1997) a MA (n-day)-filter with a n-day averaging interval is employed. This smoothens the discharge series, and as a result short excess periods are filtered out and mutually dependent droughts are pooled (Fig. 2). In this way both the time period between two drought events as well as the magnitude of the discharge values below and above the threshold level control the pooling of events. The original discharge series is modified by the smoothing such that the pooled event often starts (ends) a few days after (before) the first (last) day with discharge below the threshold level. The excess volume is automatically subtracted from the total deficit volume of the pooled event. Alternatively, the MA-filter can be used only to identify the duration of pooled drought events, whereas the pooled deficit volumes are calculated from the original discharge values (Cancelliere et al., 1995). Since the daily values are calculated as a n-day average, there is a possibility of introducing dependency between drought events if one event occurs less than n days after the preceding one without being pooled to it (Hisdal et al., 2004). An averaging interval of n=10 days is suggested by Tallaksen et al. (1997) based on the study of two perennial rivers in Denmark.

The *sequent peak algorithm* (SPA; Vogel and Stedinger, 1987) was developed for the design of water reservoirs. It derives the largest deficit volume of a discharge series for the whole period of record with respect to a threshold level, Q_z. It has also been used as a pooling procedure (e.g. Tallaksen et al., 1997). A time series of the deficit volume $w(t)$ is derived by summing up the daily deficits between the discharge at day t, $Q(t)$, and the threshold level and subtracting the excess volumes until $w(t)$ returns to zero (Fig. 3). When $w(t)$ equals zero, excess volumes are not subtracted, thus, $w(t)$ never turns negative:

$$w(t) = \begin{cases} w(t-1)+Q_z-Q(t) & \text{if} \quad w(t-1)+Q_z-Q(t)>0 \\ 0 & \text{if} \quad w(t-1)+Q_z-Q(t)\leq 0 \end{cases} \quad (3)$$

The largest deficit volume, $w_{\max}$, is then selected for each uninterrupted period of deficit, i.e. $w(t)>0$. A pooled drought event is considered to start on the first day with $w(t)>0$ and to end when $w_{\max,i}$ is reached. The total pooled deficit volume of the event is $w_{\max,i}$. Drought events are thus pooled until $w_{\max,i}$ is reached. In the time period following $w_{\max,i}$ and until the deficit volume, $w(t)$, is back to zero, the stream is not considered to be in a drought situation, since the average discharge of this period exceeds the threshold level.

2.2 Exclusion of minor droughts

When the MA-procedure is applied as a pooling procedure, minor droughts are automatically filtered out. When the IT-method or SPA is applied minor droughts have to be excluded in an additional step. For example Zelenhasić and Salvai (1987) exclude droughts with a deficit volume smaller than a certain percentage α of the maximum observed deficit volume, and Madsen and Rosbjerg (1995) exclude droughts with deficit volume or duration smaller than predefined percentages of the mean deficit volume or duration, respectively. Another possibility is to exclude droughts with a real drought duration shorter than a given minimum value, $d_{\min}$ (Jakubowski and Radczuk, 2004).

3 The Global Data Set

A global data set of 16 daily discharge series from around the world is used as a basis for the study (Fig. 4). The data set was assembled by the ASTHyDA project (ASTHyDA, 2006) in order to demonstrate the global variability of natural hydrological regimes. As such, the data set includes streams from most of the major climate zones as well as catchments of the same climate zone, but with different catchment characteristics (Rees et al., 2004). The catchment and discharge characteristics of the 16 streams are summarised in Table 1. The catchments are grouped according to their climate zones

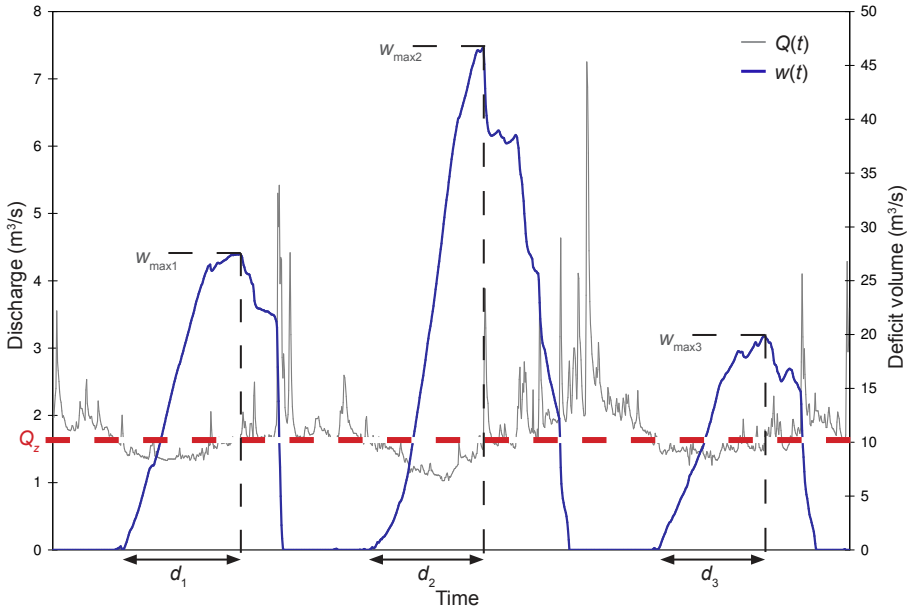

Fig. 3. Illustration of the derivation of the deficit characteristics duration, d_i, and deficit volume, $w_{\max,i}$, by the SPA.

Fig. 4. Catchments of the Global Data Set (modified from Rees et al., 2004).

following the Köppen climate classification (Köppen, 1930). In this study streams are classified as "perennial", when a stream is continuously flowing, "intermittent", when parts of a stream fall dry during dryer times, or "ephemeral", where precipitation is rare and water flows only directly after rainfall. Streams experiencing a frost season are additionally labelled "summer", since only droughts of the frost free season are considered. The hydrological regimes of the 16 stations are presented in Fig. 5.

Table 1. Catchment and discharge characteristics of the streams of the Global Data Set (with *AAR*: average annual precipitation; q: specific discharge; c_{zero}: percentage of zero-flow; *CV*: coefficient of variation).

Stream	Site	Country	Köppen Climate Zone	Streamflow type
Honokohau Stream	Honokohau	Hawaii, USA	Af: Tropical	perennial
Dawib	Dawib	Namibia	Bw: Dry – desert	ephemeral
Pecos River	Pecos	New Mexico, USA	BS: Dry – steppe	perennial, summer
Elands River	Elands River Drift	South Africa	Cw: Temperate – winter dry	intermittent
Bagmati River	Sundarijal	Nepal	Cw: Temperate – winter dry	perennial
Sabar	Alfarnatejo	Spain	Cs: Temperate – summer dry	intermittent
Arroyo Seco	Soledad	California, USA	Cs: Temperate – summer dry	intermittent
Ray	Grendon Underwood	United Kingdom	Cf: Temperate – no dry season	intermittent
Lambourn	Shaw	United Kingdom	Cf: Temperate – no dry season	perennial
Lindenborg	Lindenborg Bro	Denmark	Cf: Temperate – no dry season	perennial
Ngarururo	Kuripapango	New Zealand	Cf: Temperate – no dry season	perennial
Hurunui	Mandamus	New Zealand	Cf: Temperate – no dry season	perennial, summer
Lågen	Rosten	Norway	Df: Cold – no dry season	perennial, summer
Inva	Kudymkar	Russia	Df: Cold – no dry season	perennial, summer
Rhine	Lobith	The Netherlands	Df, Cf: Cold, Temperate	perennial, (summer)
Ostri	Liavatn	Norway	Df, ET: Cold, Polar	perennial, summer

Table 1. Continued.

Stream	Area (km^2)	Station Altitude (m a.m.s.l.)	Maximum Altitude (m a.m.s.l.)	*AAR* (mm)	q ($l/(s \cdot km^2)$)	c_{zero} (%)	*CV*	Number of used years
Honokohau Stream	11	256	ca. 1765[1]		98.36	0.0	1.23	53
Dawib	560	>200[1]	<2000[1]		0.02	98.2	14.17	4
Pecos River	490	2287	ca. 3993[1]	474–610	5.92 (7.34)[6]		1.30	68
Elands River	690	1000–1500[1]	poss. >3000[1]	500	3.53	3.0	2.44	13
Bagmati River	17	1600			62.88		1.02	22
Sabar	39	ca. 900[2]	1671[1]		4.54	50.9	3.62	29
Arroyo Seco	632	103		802–864	7.66	12.5	3.35	68
Ray	19	66	187	660	5.11	26.4	2.76	26
Lambourn	234	76	261	805	7.25	0.0	0.48	36
Lindenborg	214	5	113	741[3,4]	10.90	0.0	0.35	37
Ngarururo	370	500	1617	2000–2150[5]	46.97	0.0	1.06	34
Hurunui	1060	300	1987	1919	49.79 (45.16)[6]	0.0	0.86[6]	40
Lågen	1755	737	2200	700	52.24 (31.26)[6]	0.0	0.83[6]	84
Inva	2050	0–100[1]	200–500[1]	700–800	6.06 (6.91)[6]	0.0	1.87[6]	56
Rhine	160 800	10	4275	716	13.74 (13.00)[6]	0.0	0.51 (0.46)[6]	92
Ostri	235	733	2088	1560	44.69 (98.39)[6]	0.0	0.61[6]	34

[1] from The Times (1994)
[2] from SUR in English (2005)
[3] from Ovesen et al. (2000)
[4] average for the period 1971–1998
[5] Clausen (2003), personal communication
[6] summer

The warm and humid *tropical* climate is represented by Honokuhau Stream on Hawaii, which has a flashy streamflow behaviour due to strong and frequent convective rainfall events. The *dry* climate encompasses the *dry desert* climate, represented by the ephemeral river Dawib in Namibia, and the *dry steppe* climate, represented by the perennial Pecos River in New Mexico, USA. Here potential annual evaporation exceeds precipitation. The *temperate* climate is characterized by a high seasonal variation in temperature and in some regions also in precipitation. It is therefore further

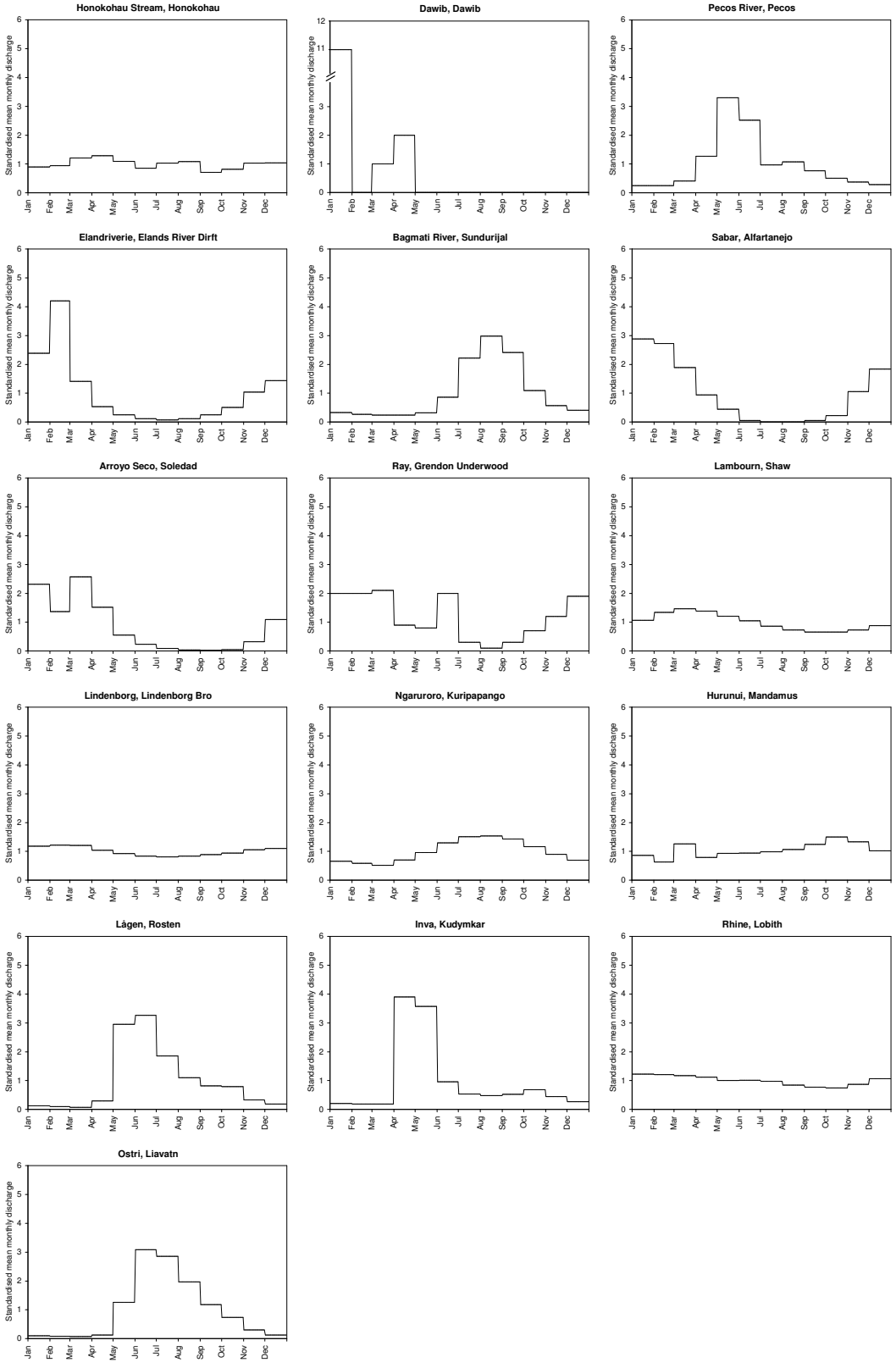

Fig. 5. Mean monthly discharges standardised by the mean discharge for the 16 stations of the Global Data Set.

classified by the timing of the dry season. The *temperate winter dry* climate is represented by Elands River in South Africa, which has an average annual precipitation (*AAR*) of 500 mm and may run dry for shorter periods, and Bagmati River in Nepal, which experiences a monsoon climate and never runs dry. The fast responding river Sabar in Spain with long seasonal zero-flow periods and the river Arroyo Seco in California, USA, experience a *temperate summer dry* climate. In the catchment of Arroyo Seco precipitation shows high inter-annual variability, and the river may run dry for several months in one year and not at all in another. Belonging to the climate zone of a *temperate climate without a dry season*, the three rivers Ray, Lambourn and Lindenborg are exposed to similar climate conditions (*AAR* 660–800 mm) but different hydrogeological characteristics. The catchment of Ray in the UK has impermeable soils resulting in a flashy streamflow behaviour with frequent zero-flow periods. The river Lambourn, also in the UK, has a high base flow contribution and a delayed low flow period from August till November. The river Lindenborg in Denmark shows a mixed flow regime. Also the two rivers from New Zealand, Ngaruroro and Hurunui, experience a temperate climate without a dry season, however with higher AAR ($\sim$2000 mm). Ngaruroro has a mixed response, but the baseflow contribution is lower than for Lindenborg. The catchment of Hurunui is partly snow covered during winter, but no pronounced snowmelt flood can be identified. Within the *cold* climate the length of the freezing period varies. When it is sufficiently long, low flows may also occur in winter as for the rivers Lågen in Norway and Inva in Russia. The catchment of Inva is located in a low lying area, which implies that the snowmelt flood is more distinct than for Lågen, whose catchment has an altitude range of 1500 m. The catchment of Ostri in Norway experiences a mixed *cold* and *polar* climate, and following the snowmelt flood melting water from glaciers contributes considerably to its summer flow. As an example of a large catchment covering two climate regions the river Rhine is included in the data set. At Lobith the catchment area is 160 800 km^2 and the catchment includes both temperate and cold climate regions.

All data series are quality controlled and periods of up to 15 days with incorrect values or missing data are filled in by interpolation. Years with more than 15 days of missing values are disregarded and the numbers of years with complete records are given in Table 1. The hydrological year is defined to start in the high flow season. For the only ephemeral river in the dataset, Dawib, only four consecutive years with a complete record exists. Therefore only a qualitative evaluation of the threshold level method can be provided for this type of stream.

4 Evaluation of the threshold level method for daily streamflow series

The threshold level method combined with the three pooling procedures as outlined in Sect. 2 is applied to the perennial and intermittent streams from the Global Data Set. Threshold levels are chosen to represent the range of commonly used threshold levels. Accordingly, Q_{90} and Q_{70}, are used for perennial streams, and percentiles between Q_{20} and Q_{70} for intermittent streams depending on their percentages of zero flow. Based on the experience form previous studies (Tallaksen et al., 1997; Hisdal et al., 2002; Engeland et al., 2004) threshold levels in the range between Q_{90} and Q_{70} for perennial streams are considered reasonable also for an extreme value analysis of droughts. In the following, the results obtained from the application of the threshold level method are summarised by discussing the necessary considerations for the different streamflow regimes, the determination of the pooling parameters, advantages and limitations of the different pooling procedures, and a comparison of the deficit characteristics derived with the three procedures.

4.1 Characteristics of streamflow regimes

The application of the threshold level method to *perennial* streams is normally straightforward, with the exception of streams experiencing pronounced seasonal differences as discussed below. *Intermittent* streams dry out, implying that the deficit volumes during zero-flow periods do not increase with increasing drought duration in the same way as during flow periods. The deficit volumes of intermittent streams can thus not be interpreted in the same way as those of perennial streams, and in case of a frequency analysis the deficit volumes of zero-flow periods should be treated as censored data. For *ephemeral* streams with rare and short flow events the severity of a drought is reflected by the duration of zero-flow periods and the total flow volume rather than by deficit periods during a flow event. The application of the threshold level method for daily data is therefore not favourable. Alternative drought characteristics are the duration of zero-flow periods, total volume of flow events or total annual discharge, as well as characteristics derived from groundwater or reservoir data.

Streamflow droughts can be of different origin caused by seasonally different hydroclimatological processes, for instance in regions experiencing a wet and a dry season or a warm and a cold season. In regions with a cold winter season two different types of streamflow droughts have to be distinguished: *summer droughts* caused by low precipitation and often accompanied by high evapotranspiration losses, and *winter droughts* occurring when the temperature is below the freezing point and water in the catchment is stored as snow and ice. If droughts are of different origin, it has to be decided whether deficit characteristics ought to be calculated for each type separately, e.g. in case of a frequency analysis,

or whether it is acceptable to derive a mixed series of drought events. If the droughts are to be separated, the seasons and the procedure to identify the season-specific droughts have to be defined. In addition, it is necessary to decide whether the threshold level should be based on the data of the whole year or of the season of interest.

Calculations focusing on only one season are recommended for streams with a cold winter season, since they often experience two annual low flow periods caused by different processes. Streams with one wet and one dry season usually experience only one low flow period, and in case of an all-year study only dry-season droughts are derived. However, in order to study wet-season droughts, separate calculations for the wet-season are necessary as suggested by Tate and Freeman (2000) for daily streamflow series in Southern Africa. Alternatively, a seasonally, monthly or daily varying threshold level could be used as suggested by Stahl (2001).

For frost influenced catchments Hisdal et al. (2001) specified fixed seasons and defined the summer as the period with mean monthly temperature above the freezing point. In case of an annual snowmelt flood, the start of the summer season is defined to be at the end of the flood period to avoid that very high flow values caused by winter precipitation influence the threshold level for the summer season. The use of fixed seasons does not account for the fact that the frost period actually varies from year to year and as a result drought events might be cut off or incorrectly classified into summer or winter droughts. However, defining the seasons for each year separately would require daily temperature data. Alternatively, droughts can be derived from the complete time series and classified as, e.g. summer droughts if their major part, in terms of deficit or duration, belongs to a predefined fixed summer season. This means that neither late-ending summer droughts are cut off, nor are parts of early winter droughts included in the sample of summer droughts. The events should in this case be classified prior to pooling. However, in catchments with a long winter season, as for the two Norwegian catchments Lågen and Ostri, it can happen that a severe summer drought develops into an even longer winter drought and is thus misclassified as a winter drought. For catchments with a short winter season, as for Pecos River in New Mexico, a whole winter season might be included in a drought classified as summer drought. There are no simple solutions to these difficulties encountered when classifying summer droughts that continue into the winter season. Should they be considered to end at some fixed date (censored data sample) or continue into the winter season (non-homogeneous data sample)? The use of a censored data sample has been suggested by Tate and Freeman (2000) for wet-season droughts and could also be tested for summer droughts.

A similar problem concerns catchments that cover a large altitude range or have a large areal extension in which case the seasonal behaviour can vary within the catchment as for the river Rhine at Lobith, the Netherlands. On average the lower parts of the catchment experience continuous frost periods only for a few days, while the frost periods in the mountainous areas last for several months. In such cases the specification of a summer season must depend on the hydrological regime at the site of interest, baring in mind the complexity of the catchment characteristics in the evaluation of the results.

If two distinct seasons are present it is recommended to derive separate threshold levels for each season, since a FDC of the summer period may differ considerably from the FDC of the complete data record. For example for the river Lågen, Norway, the fixed summer season, defined according to Hisdal et al. (2001), lasts from 15 June until 30 September. The 90-percentile flow for the whole year, $Q_{90\,\mathrm{Year}}$=2.97 m^3/s, is approximately five times smaller than for the summer, $Q_{90\,\mathrm{Summer}}$, and even lower than the observed lowest summer discharge of 6.30 m^3/s. Thus, with $Q_{90\,\mathrm{Year}}$ as threshold level no summer droughts would be selected.

In this study the following stepwise procedure has been adopted for season-specific calculations in frost influenced regions:

1. specification of the start and end date for the season of interest (fixed seasons);
2. determination of the threshold level based on data from the season of interest;
3. selection of drought events for the whole year;
4. classification of each drought event according to which season its longest part belongs to;
5. pooling of the drought events belonging to the season of interest.

4.2 The IT-method

For the IT-method the parameter value, t_c, has to be determined, and a sensitivity analysis of t_c is performed to judge whether the values of t_c=5 and 6 days as recommended by Tallaksen et al. (1997) and Zelenhasić and Salvai (1987), respectively, also apply to streams from other flow regimes. Five perennial and three intermittent streams representing the various regimes are chosen for the sensitivity study. In case of the perennial streams (Lindenborg, Ngaruroro, Bagmati River and Honokohau Stream) Q_{90} is applied as threshold level. For the intermittent streams the threshold is selected depending on the occurrence of zero flows, Q_{70} for Arroyo Seco, Q_{50} for Ray and Q_{20} for Sabar. The mean deficit characteristics of the annual maximum series (AMS) of non-zero values are analysed, rather than of the PDS, to reduce the influence of the number of minor droughts and pooled events. The mean values of deficit volume and real drought duration are calculated for t_c=0, 2, 5, 7, 10, 15, 20 and 30 days, where t_c=0 days represents the drought series without pooling. The

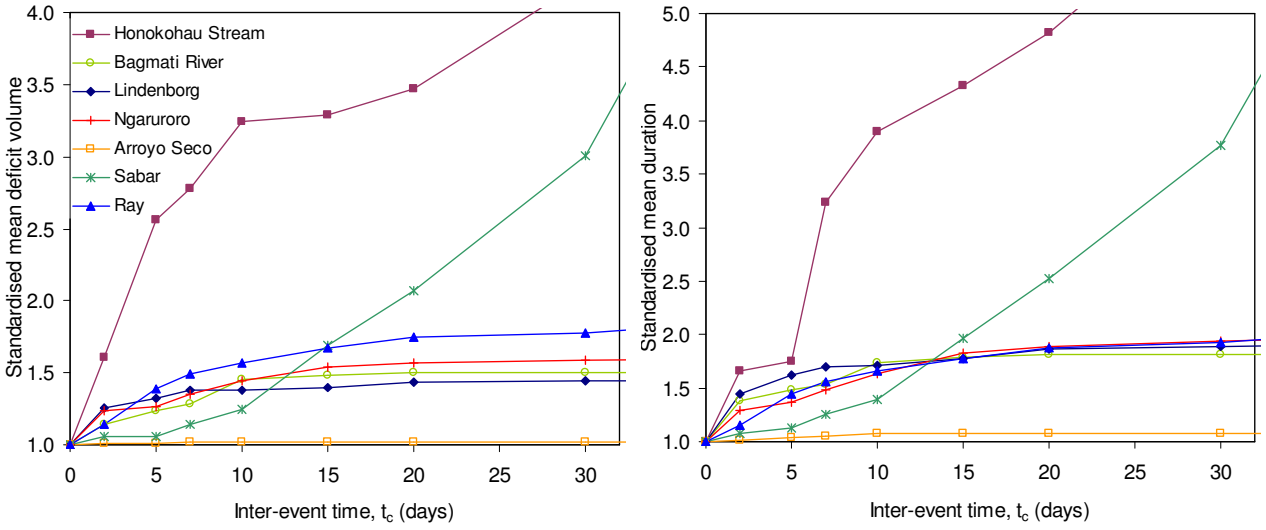

Fig. 6. Relationship between the inter-event time and the standardised mean deficit volume (left) and the standardised mean duration (right).

mean deficit volume and duration are standardised by the respective mean of t_c=0 day. The relationships between t_c and the deficit characteristics are shown in Fig. 6.

The sensitivity curves generally start to level out around t_c=5 days, and for most streams the deficit characteristics do not change substantially after t_c equals 10 to 15 days implying that a maximum of pooling is obtained. For Honokohau Stream and Sabar, however, the standardised mean deficit characteristics continue to increase. For Honokohau Stream this is due to the flashy discharge behaviour and the frequent occurrence of deficit periods. Sabar, on the other hand, is an intermittent stream with a distinct dry season. Its standardised mean deficit characteristics increase due to an increasing number of multi-year droughts with increasing t_c. This is mainly a result of the high threshold level chosen for Sabar (Q_{20}). Contrarily, for the other intermittent stream with a distinct dry season, Arroyo Seco, the choice of t_c is of little importance as the threshold level in this case is sufficiently low (Q_{70}) to avoid identifying wet-season and multi-year droughts. The choice of t_c is in this case not important, since the dry season is usually not split into several events. Hence, in case of a distinct dry season, pooling generally has no effect as long as the threshold level is selected sufficiently low. Further, the influence of pooling on the most extreme events is analysed. Sensitivity curves of both the maximum as well as the mean of the ten largest observations in the PDS of deficit volume and duration are plotted (not shown), again standardized by the respective values of the non-pooled series. The curves show a similar behaviour as the sensitivity curves for the mean of the AMS.

The sensitivity analysis suggests that the value recommended by Tallaksen et al. (1997), t_c=5 days, can be applied for perennial as well as intermittent streams with the exception of very flashy streams. The applicability of the IT-method for flashy streams is further discussed in Sect. 4.5.

4.3 The MA-procedure

For the MA-procedure Tallaksen et al. (1997) suggested an averaging interval of n=10 days for two perennial rivers in Denmark including the river Lindenborg. It is here tested whether the value of n=10 days can also be recommended for flashy (Honokohau Stream) and intermittent streams (Ray and Arroyo Seco). The mean deficit volume and mean real drought duration of the AMS of non-zero values are calculated and standardised by the respective mean of the non-pooled AMS. The MA-procedure has two effects which can both result in an increase of the mean deficit characteristics: pooling of events and exclusion of minor droughts. The latter only has an effect if it results in an increase in zero-drought years. The modification of the discharge series by using a MA (n-day)-filter implies, however, that the mean deficit characteristics are not strictly increasing with increasing n. An appropriate value for n can be selected when the mean deficit characteristics reach a maximum or when they level out caused by the pooling of events. The averaging interval n should be chosen as small as possible, since a MA (n-day)-filter modifies the discharge series.

AMS of drought events are obtained for n=5, 10, 15 and 20 days with a threshold level of Q_{90} for the perennial streams and Q_{70} and Q_{50} for Arroyo Seco and Ray, respectively. The relationships between n and the standardised mean deficit characteristics are displayed in Fig. 7. For Arroyo Seco, Lindenborg and Ray the curves flatten at n=5, 7 and 10 days, without being influenced by an increase in zero-drought years. These values for n can thus be considered

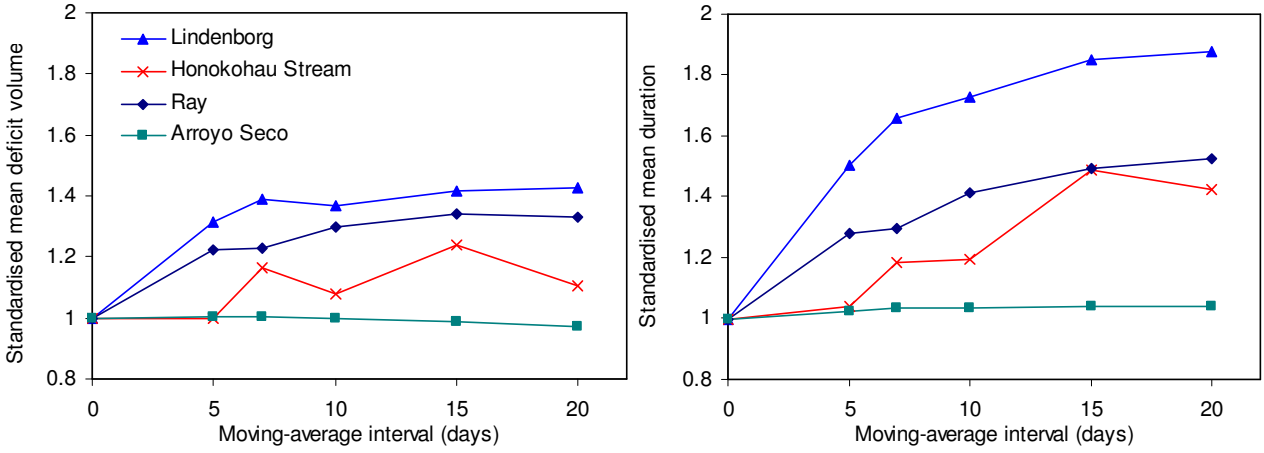

Fig. 7. Relationship between the moving-average interval and standardised mean deficit volume (left) and standardised mean duration (right) for Lindenborg, Honokohau Stream, Ray and Arroyo Seco.

optimal to select drought events. For Honokohau Stream the mean deficit characteristics reach a maximum at n=15 days and a local maximum at n=7 days. The maximum at 15 days is mainly caused by an increase in zero-drought years due to the exclusion of minor droughts, whereas the maximum at 7 days corresponds to the maximum deficit characteristics. A moving-average interval of n=7 days is therefore considered optimal for Honokohau Stream. The effect of pooling on the most extreme events is studied analogue to the analysis for the IT-method. For the MA-procedure the effect of pooling on the most extreme events in the PDS is found to be much smaller than on the mean values of the AMS. The only exception is Honokohau Stream for which the standardized maximum duration increases to about 2 for moving-average intervals of n=7 days and higher, whereas the standardized deficit volume remains nearly 1. This is a result of the duration of inter-event excess periods being included in the pooled drought duration, whereas inter-event excess volumes are automatically subtracted from the pooled deficit volume. In general it is concluded that an averaging interval of the order of 7 days is appropriate for these streams.

4.4 The SPA

An advantage of the SPA is that it requires no parameters in addition to the threshold level. However, the application of the SPA as pooling procedure reveals a major problem related to the original purpose of deriving the largest observed deficit volume. For example, within the 35 year long data record of Lindenborg the five most severe drought events occur between 1974 and 1978 using Q_{90} as threshold level. The fourth most severe event occurs in 1977 (Fig. 8, upper graph). A threshold of Q_{80} implies that the events from 1975, 76 and 77 are pooled into one multi-year drought and the maximum deficit volume is reached in October 1976 (Fig. 8, lower graph). The drought is thus considered to end in October 1976 and the year 1977 is considered to be drought-free. In general, drought events that occur shortly after a major event get pooled to this event, but are not accounted for in any way, neither in the deficit volume nor in the duration of the major event. This applies also to within-year droughts. Hence, the SPA is not suited as pooling procedure for the selection of a PDS. For an AMS its use should be limited to very low threshold levels to avoid that events following a major drought are not recognized.

4.5 Comparison of the pooling procedures

The degree of pooling of the three procedures is compared, employing n=7 days for the MA-procedure and t_c=5 days for the IT-method. The chosen threshold levels are the same as before. The deficit characteristics of the ten largest events in the PDS are compared, since the criteria to exclude minor droughts differ between the three procedures. It is found that for all types of streams the events derived with the MA-procedure show somewhat smaller deficit volumes and durations. This lower degree of pooling is mostly due to the choice of averaging interval. The IT-method and SPA pool drought events in a comparable manner for low threshold levels, not showing any streamflow-type specific differences. For higher threshold levels, the degree of pooling is much larger with the SPA and the problem of not recognising events following a major drought can occur. The largest differences in pooling are observed for fast responding catchments, such as Honokohau Stream and Ray. Honokohau Stream shows a particular flashy discharge behaviour due to the high frequency of rain events. For this stream the degree of pooling is much higher with the IT-method as compared to the MA-procedure and SPA as illustrated in Fig. 9. The shaded areas in Fig. 9 indicate the deficit volumes of

a non-pooled drought series and the lines below show the drought periods as they are pooled by the 5-day IT-method (upper lines), the MA (7-day)-filter (middle lines) and the SPA (lower lines). The events selected by the MA-procedure or the SPA last much shorter compared to the events derived with the IT-method. For example from mid August 1945 to mid September 1945 a series of seven minor drought events are pooled with a subsequent major drought (IT-method). Thus considering the pooled event to start on 13 August as opposed to the start of the major drought more than one month later on 19 September. With the MA-procedure and SPA these minor droughts are not pooled, since pooling is also determined by the inter-event excess volume. It is anticipated that the use of an additional IV-criterion would improve the IT-method.

It can be concluded that for the studied types of streams (perennial with and without cold winter and intermittent with and without dry season) the three tested pooling procedures can be applied with the following limitations:

- The IT-method is not recommended for flashy streams. An additional IV-criterion could possibly improve the method.
- The MA-procedure modifies the discharge series and thus deficit volume and duration. It might also introduce dependency between pooled drought events.
- As pooling procedure the SPA is only advisable for the study of AMS rather than PDS. In addition, it is limited to very low threshold levels.

General recommendations for season-specific drought studies are given in Sect. 4.1. For comparative studies the MA-procedure is considered the most flexible approach, keeping in mind the above mentioned limitations.

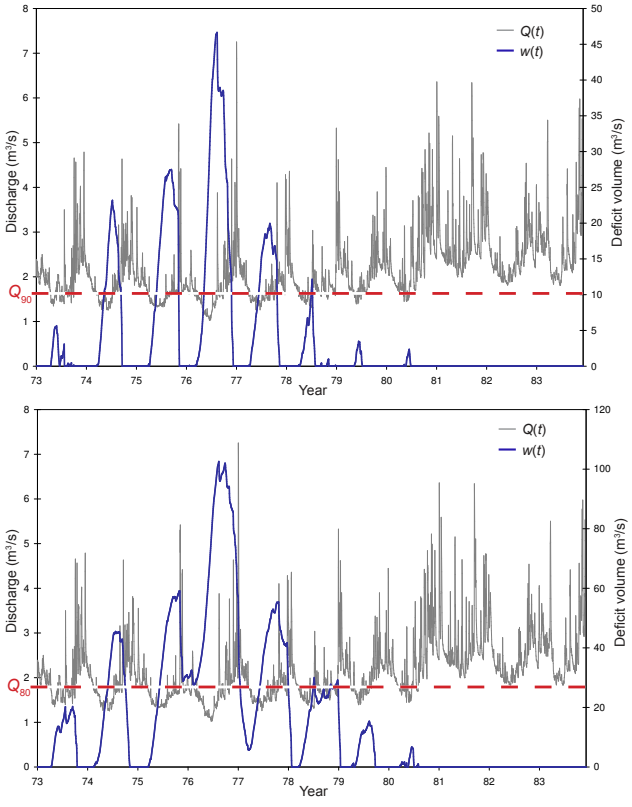

Fig. 8. Deficit volume, $w(t)$, in m^3/s derived by the SPA for the period 1973 to 1983 for the river Lindenborg, for two different threshold levels, Q_z; Upper: Q_z=Q_{90} and lower: Q_z=Q_{80}.

5 Frequency analysis

Partial duration series (PDS) of drought events are derived from time series of daily discharge using the threshold level method as outlined in the previous sections. The PDS model includes two stochastic model components, the number of extreme events occurring in a given time interval and the magnitude of the events, which is here expressed as deficit volume or duration. The Generalized Pareto (GP) distribution can be shown to be the limit distribution of scaled excesses over a limit, u, and is thus suited to model PDS of magnitudes, when the events included in the PDS exceed the limit u (Pickands, 1975). In practice u has to be determined based on subjective judgement (see below). When the number of extreme events in the PDS is assumed to follow the Poisson distribution and the magnitudes the GP distribution, the Generalized Extreme Value (GEV; Fisher and Tippett, 1928) results for the AMS. Annual exceedance probabilities can be estimated from the PDS provided the average number of events per year larger than the limit is known. Following Zelenhasić and Salvai (1987) the distribution of the largest streamflow drought (expressed as deficit volume or duration) in a given time interval, e.g. one year, $H(x)$, is derived based on $F(x)$, the distribution function of the magnitudes of all events which exceed the limit u within the time interval, combined with the distribution function of the number of droughts occurring in the time interval:

$$H(x) = Pr(E = 0) + \sum_{k=1}^{\infty} F^k(x) Pr(E = k) \quad (4)$$

where Pr(E=k) is the probability that k events occur during the time interval (E is the number of events). It is assumed that the drought deficit volume (duration) is an independent, identically distributed (*iid*) random variable. Only drought events lasting less than one year are included. For low threshold levels, such as Q_{90} and lower, Zelenhasić and Salvai (1987) considered all events included in the PDS after excluding minor droughts to be extreme events, in the sense that they are larger than u and the PDS can be modelled by the GP distribution. They excluded minor droughts with a deficit volume smaller than α=0.5 or 1% of the maximum

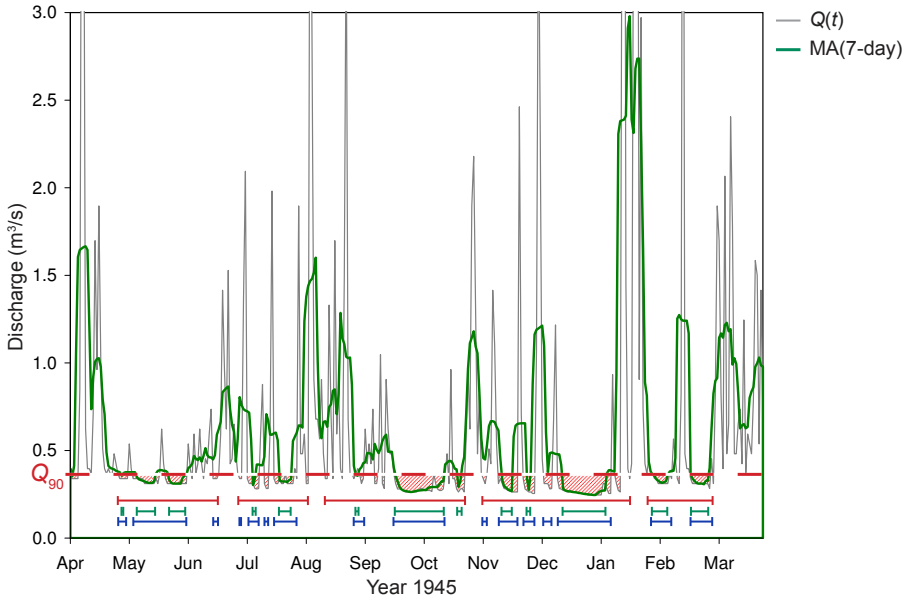

Fig. 9. Difference in pooling by the 5-day IT-method (upper lines), a MA(7-day)-filter (middle lines) and the SPA (lower lines) for Honokohau Stream for 1945. The non-pooled deficit periods are marked as shaded areas.

observed deficit volume, and found that an empirical check of the results for two streams in former Yugoslavia proved the method statistically correct for threshold levels Q_{90} and lower.

In this study the program NIZOWKA (Jakubowski and Radczuk, 2004) is applied for the extreme value analysis including the selection of droughts. The magnitudes of the drought events comprising a PDS of drought deficit volume or duration, are derived for two threshold levels, Q_{90} and Q_{70}. As pooling procedure the 5-day IT-method is used, since the MA-procedure introduces dependency between pooled events for some of the streams. For Honokohau Stream, however, the MA(7-day)-filter is applied, as the IT-method was found to not be applicable for flashy streams. For the frequency analysis extreme events exceeding a limit u have to be selected. u should be high enough to exclude all non-extreme droughts to limit model bias and enable a good fit to the most extreme events. On the other hand, the number of remaining drought events has to be sufficiently high to avoid large uncertainties in the estimated design event. The optimum number of events is found as a compromise between selecting extremes following an extreme value distribution and a sample size needed for sufficiently precise estimations. Hisdal et al. (2002) tested two common approaches to determine the limit u for extreme value modelling of streamflow drought: the mean excess plot (or mean residual life plot) and the plot of the shape parameter of the GP distribution as a function of u (Coles, 2001). They found that the plot of the shape parameter is a very informative tool to determine u. The mean excess plot on the other hand is often difficult to interpret and thus seldom helpful. In this study the limit u is instead represented by two criteria to exclude minor droughts. Droughts are excluded when the deficit volume is smaller than a certain percentage, α, of the maximum observed deficit volume or the real drought duration is smaller than $d_{\min}$. It was decided to use both a minimum volume and duration criterion, as the two criteria affect the PDS of various types of streams differently. However, it is difficult to determine the best combination of the two limit criteria based on a shape parameter plot. The values for α and $d_{\min}$ are therefore rather determined based on a comparison of fitted GP/Poisson models for different values of α and $d_{\min}$ to the observations and a comparison of the drought estimates for a return period of 50 years. Nine combinations of α and $d_{\min}$ are tested with α=0, 0.5 or 1% and $d_{\min}$=1, 3 or 5 days. Requiring the use of the same combination of α and $d_{\min}$ for all streams and for the estimation of deficit volume and duration, the combination of α=0.5% and $d_{\min}$=3 days is found to be the best choice. However, for a drought study at an individual site another combination might be preferable. For Honokohau Stream, where the MA(7-day)-filter is applied, α=0% and $d_{\min}$=1 day was found to be the best choice, implying that in this case the exclusion of minor droughts by the MA-procedure is sufficient.

NIZOWKA allows several probability distributions to be fitted to the series of deficit volume or duration by the method of maximum likelihood. This encompasses the Poisson and Pascal distribution for the occurrence of the events and the Gamma, Weibull, Log-Normal, Johnson, Gumbel and Generalized Pareto distribution for the magnitudes of the PDS,

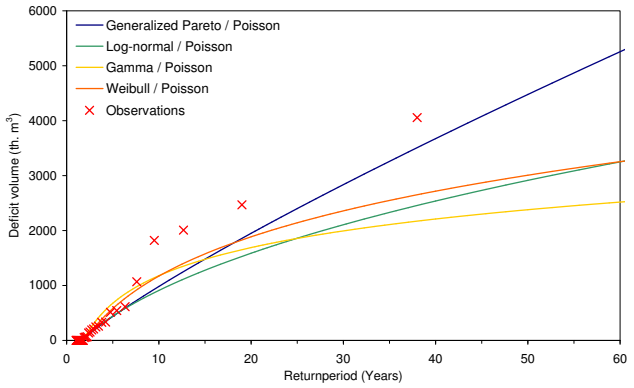

Fig. 10. Different distribution models for $H(x)$ compared to the AMS of observed deficit volumes for the river Lindenborg.

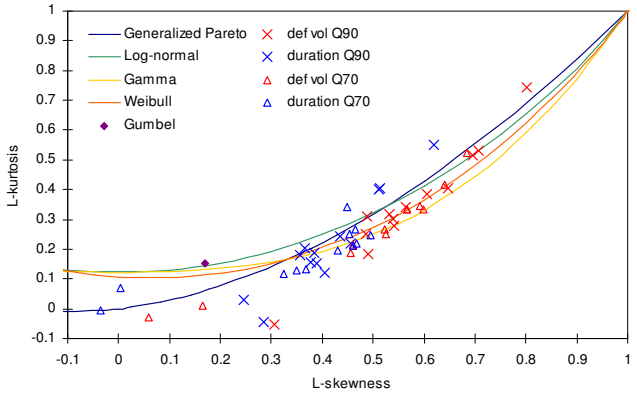

Fig. 11. L-moment ratios of the observed PDS of deficit volume and duration derived with threshold levels Q_{70} and Q_{90}, and for five theoretical distribution models.

$F(x)$. The model fits are tested by a χ^2-goodness-of-fit test (Haan, 1977). The test does not allow determination of the "best" or "true" distribution model (Stedinger et al., 1993), but gives an indication of which models perform reasonably well. The choice of the distribution for Pr(E=k) is observed to be of minor influence for $H(x)$, and the Poisson distribution is chosen. The fit of the obtained distribution functions for $H(x)$ is then visually compared with the observed AMS. In the visual comparison the overall fit to the complete AMS of observations is considered as well as the fit in the extreme range. The extreme range is considered to consist of the three to five largest observed events, depending on the length of the series. The models for $H(x)$ are in the following labelled according to the models used for $F(x)$ and Pr(E=k), e.g. a combined GP/Poisson model.

The described procedure is illustrated for deficit volumes for the river Lindenborg in which case the GP, Log-Normal, Weibull and Gamma distribution models all can be accepted at a significance level of 0.05 for the PDS. The highest attained significance level, ρ, in the χ^2-goodness-of-fit test is obtained for the Log-Normal model (ρ=0.71) and the second highest for the GP model (ρ=0.50). The visual comparison (Fig. 10) of the distributions for $H(x)$ shows that the GP/Poisson model gives the best fit to the extreme range, where the deviations between the different models is largest. Using the GP/Poisson model the relative deficit volume (deficit volume of an event relative to the daily mean discharge volume) of the drought event with a 50-year return period is estimated to 22.2. The estimates based on the other three models vary between 11.8 and 14.9, i.e. 33–47% lower than the GP/Poisson estimate. This demonstrates the high uncertainty in the estimate of return levels depending on the choice of distribution.

In total 38 estimations are successfully derived for the whole Global Data Set. This includes estimations for deficit volume, v, and duration, d, for ten perennial streams for two different threshold levels and duration for three intermittent streams (one threshold level). Duration is defined as the full drought duration (Eq. 1). No estimations are done for deficit volumes of intermittent streams. For Pecos River and Arroyo Seco no estimations were possible and for Shaw not for duration using Q_z=Q_{90}. For Hurunui estimations were possible only for deficit volumes with both threshold levels and for Lågen for both deficit characteristics but only when using Q_z=Q_{90}. In Table 2 the performance of the combined GP/Poisson model is compared to other models. Only models that achieved an attained significance level larger than 0.02 in the χ^2-goodness-of-fit test for $F(x)$ are considered. In 55% of the cases, the GP/Poisson gives the best fit to the most extreme events or to the whole range of drought events. In an additional 21.0% of the cases the GP/Poisson model performs equally well as the other models. Only in 24% of the cases other models perform better. The choice of the GP distribution is also supported by plotting L-skewness versus L-kurtosis for the PDS of deficit volume and duration (Fig. 11). All distributions, except the Gumble distribution, model the middle range of L-skewness and L-kurtosis values reasonably well, whereas for the lower and upper range, the GP distribution is preferable. The plot further shows that the PDS of deficit volume in general have a higher L-skewness and L-kurtosis than the PDS of duration. The selection of the GP/Poisson model, corresponding to a GEV model for the AMS, is in agreement with the extreme value theory, which states that excesses over a certain limit are Generalized Pareto (GP) distributed. The GP/Poisson model can thus be recommended, when estimating annual maxima from PDS of deficit characteristics in perennial and intermittent streams.

6 Conclusions

Droughts are natural hazards which can cover large regions and last for long periods of time. This implies that robust drought characteristics applicable in regions with different hydroclimatology and hydrogeology are needed. In this

Table 2. Performance of the combined GP/Poisson model compared to the performance of other models.

Estimation	Number of stations with estimations	Number of stations for which			
		GP/Poisson shows best overall fit	GP/Poisson fits best to extreme events	GP/Poisson fits equally well as others	Other models fit better
v: Q_z=Q_{90}	9	3	3	3	0
d: Q_z=Q_{90}	7	1	3	1	2
v: Q_z=Q_{70}	10	6	0	1	3
d: Q_z=Q_{70}	12	4	1	3	4
Total	38	14	7	8	9

study the threshold level method in combination with three common pooling procedures is evaluated to derive streamflow deficit characteristics from series with a daily time resolution. The pooling procedures are designed to overcome the problem of mutually dependent droughts. The method and procedures are evaluated based on their advantages and limitations as demonstrated in an application to 15 streams from around the world. The pooling procedures are judged with respect to their applicability to streams from different hydrological regimes and for comparability between streams. The threshold level method proved to be a suitable method for perennial and intermittent streams and useable both for all-year as well as seasonal series. It allows defining droughts depending on the purpose of the study as the threshold level can be chosen either to represent a specific water demand or the boundary to normal streamflow conditions. For ephemeral streams, other drought characteristics, like the duration of the zero-flow periods, are considered to provide more relevant information.

For regional applications, which include different types of streams, both the IT-method and the MA-procedure can be used as pooling procedure. The applicability of the SPA should, however, be limited to the study of AMS rather than PDS and to very low threshold levels to ensure that also events occurring shortly after major events are recognized. It is further recommended that the IT-method is extended to include both a time and volume based criterion as also suggested by Madsen and Rosbjerg (1995) and Tallaksen et al. (1997). For flashy streams with frequent crossing of the threshold level, the IT-method tends to pool too many events. In these fast responding catchments it is necessary to consider also the excess volume, which can be considerable also for short excess periods. The MA-procedure is applicable to both fast and slowly responding streams and its parameter, the averaging interval, can easily be optimised for each stream. A drawback is that it modifies the discharge series and may introduce dependency between the drought events.

A remaining challenge is how to define season-specific drought events, in particular severe summer droughts that continue as long winter droughts. A frequency analysis requires that the events are *iid*, which in this case is difficult to fulfil. A choice has to be made between working with censored data, i.e. summer droughts are cut off at the start of the winter season, or non-homogeneous data when combined summer and winter droughts are considered to belong to either the summer or winter season by a predefined rule, e.g. the longest duration. It remains to be tested what the best model is for the various flow regimes in cold regions.

Regional drought studies require a consistent set of drought characteristics that can be applied across the region. Deficit characteristics derived by the threshold level method proved to give comparable results for different kinds of streams provided that comparable threshold levels are chosen in accordance with the streamflow regimes. This is an advantage when estimates of design events are derived across a larger, often heterogeneous region. It should be emphasised that a methodology suitable for application in large regions, adapting to streams with widely differing flow regimes, would not necessarily imply the best choice for individual sites. In general, the choice of drought definition is a subjective choice that is made based on the purpose of the study, the hydrological regime, the type of drought considered, the demand and vulnerability of nature (and society) in that region and the available data. In addition, there are in most cases subjective elements inherited in the procedures themselves. For the threshold level method these include the choice of time resolution, threshold level, pooling procedure, criteria to exclude minor droughts, and parameters of the pooling criteria and criteria to exclude minor droughts.

It was further found that the Generalized Pareto model is a good choice for the distribution of the magnitudes of drought events (PDS of deficit volume and duration) for most streams, thus supporting the theoretical base of extreme value modelling. There are large uncertainties related to fitting distributions based on observations only, in particular in the tail of the distributions. It is therefore recommended to let the choice of distribution function be guided by extreme value theory as this will likely give better predictions of the most extreme events.

A regional monitoring system commonly adapts to simple measures like relating the current streamflow to a value from the flow duration curve. A drought forecasting system on the other hand, depends on our ability to link large scale climate drivers to the frequency and occurrence of drought at the land surface. This requires that both the temporal and spatial development of drought causing processes in the climate and terrestrial system can be compared. For streamflow droughts the threshold level method is found to be a flexible approach for a wide range of flow regimes, capturing both the duration and the severity of a drought event. The possible link between drought and large-scale ocean and atmosphere patterns will be investigated in a further study, including deficit characteristics as presented here.

Acknowledgements. This study was conducted as contribution to the ASTHyDA project (Analysis, Synthesis and Transfer of Knowledge and Tools on Hydrological Drought Assessment through a European Network, http://www.geo.uio.no/drought), an accompanying measure project in the EC's 5th Framework Programme (EVK1-CT-2002-80023). The authors would like to thank the ASTHyDA consortium for providing the Global Data Set and all institutions contributing to the data set. Most series in the data set are taken from the Northern European FRIEND database, the European Water Archive (http://ne-friend.bafg.de/servlet/is/7397/?lang=en). The authors would further like to thank the three referees, S. M. Vicente Serrano, A. Cancelliere and one anonymous referee as well as the editor P. Molnar for their valuable comments.

Edited by: P. Molnar

References

ASTHyDA: Analysis, Synthesis and Transfer of Knowledge and Tools on Hydrological Drought Assessment through a European Network, http://www.geo.uio.no/drought, 2006.

Ben-Zvi, A.: Indices of hydrological drought in Israel, J. Hydrol., 92, 179–191, 1987.

Cancelliere, A.: Interactive comment on "A global evaluation of streamflow drought characteristics" by A. K. Fleig et al., Hydrol. Earth Syst. Sci. Discuss, 2, S1221–S1224, 2005.

Cancelliere, A., Ancarani, A., and Rossi, G.: Identification of drought periods on streamflow series at different time scales, in: Water Resources Management under Drought or Water Shortage Conditions, edited by: Tsiourtis, N. X., Balkema, Rotterdam, 49–57, 1995.

Clausen, B. and Pearson, C. P.: Regional frequency analysis of annual maximum streamflow drought, J. Hydrol., 173, 111–130, 1995.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer Series in Statistics, Springer-Verlag, London, 2001.

Correia, F. N., Santos, M. A., and Rodrigues, R.: Engineering risk in regional drought studies, in: Engineering, Reliability and Risk in Water Resources, edited by: Duckstein, L. and Plate, E. J., NATO ASI Series E: Applied Sciences No. 124, Martinus Ninjhoff Publ., Dordrecht, 61–86, 1987.

Engeland, K., Hisdal, H., and Frigessi, A.: Practical extreme value modeling of hydrological floods and droughts: a case study, Extremes, 7, 5–30, 2004.

EUMETSAT: Drought in Europe, available at: http://www.eumetsat.int/idcplg?IdcService=SS_GET_PAGE&ssDocName=005280&l=en&ssTargetNodeId=115, 2006.

EurAqua: Discussion Document: Towards a European Drought Policy, available at: http://www.euraqua.org, 2004.

Fisher, R. A. and Tippett, L. H. C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample, Proc. Cambridge Phil. Soc. 24(2), 180–190, 1928.

Haan, C. T.: Statistical Methods in Hydrology, The Iowa State University Press, Ames, Iowa, 1977.

Hayes, M. J.: Drought Indices, available at: http://www.drought.unl.edu/whatis/indices.htm, 2005.

Heim Jr., R. R.: A review of twentieth-century drought indices used in the United States, BAMS, 83(8), 1149–1165, 2002.

Hisdal, H., Stahl, K., Tallaksen, L. M., and Demuth, S.: Have streamflow droughts in Europe become more severe or frequent?, Int. J. Climatol., 21, 317–333, 2001.

Hisdal, H., Tallaksen, L. M., and Frigessi, A.: Handling non-extreme events in extreme value modelling of streamflow droughts, in: FRIEND 2002 – Regional Hydrology: Bridging the Gap between Research and Practice, IAHS Publ. no. 274, 281–288, 2002.

Hisdal, H., Clausen, B., Gustard, A., Peters, E., and Tallaksen, L. M.: Event Definitions and Indices, in: Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater, edited by: Tallaksen, L. M. and van Lanen, H. A. J., Developments in Water Science, 48, Elsevier Science B.V., Amsterdam, 139–198, 2004.

Jakubowski, W. and Radczuk, L.: Estimation of Hydrological Drought Characteristics NIZOWKA2003 – Software Manual, on accompanying CD to: Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater, edited by: Tallaksen, L. M. and van Lanen, H. A. J., Developments in Water Science, 48, Elsevier Science B.V., Amsterdam, 2004.

Keyantash, J. and Dracup, J. A.: The quantification of drought: An evaluation of drought indices, BAMS, 83(8), 1167–1180, 2002.

Kjeldsen, T. R., Lundorf, A., and Rosbjerg, D.: Use of a two-component exponential distribution in partial duration modellingo f hydrological droughts in Zimbabwean rivers, Hydrol. Sci. J., 45(2), 285–298, 2000.

Köppen, W.: Das geographische System der Klimate, in: Handbuch der Klimatologie, edited by: Köppen, W. and Geiger, R., Verlag von Gebrüder Borntraeger, Berlin, C1–C44, 1930.

Madsen, H. and Rosbjerg, D.: On the modeling of extreme droughts, in: Modelling and Management of Sustainable Basin-scale Water Resources Systems (Proc. Boulder Symp., July 1995), IAHS Publ. no 231, 377–385, 1995.

Ovesen, N. B., Iversen, H. L., Larsen, S. L., Müller-Wohlfeil, D.-I., Svendsen, L. M., Blicher, A. S., and Jensen, P. M.: Afstrømningsforhold i danske vandløb, Danmarks Miljøundersøgelser, Faglig rapport fra DMU, 340, 2000.

Pickands, J.: Statistical inference using extreme order statistics, Ann. Stat., 3, 119–131, 1975.

Rees, G., Marsh, T., Roald, L. Demuth, S., van Lanen, H. A. J, and Kaspárek, L.: Hydrological Data, in: Hydrological Drought – Processes and Estimation Methods for Streamflow and Ground-

water, edited by: Tallaksen, L. M. and van Lanen, H. A. J., Developments in Water Science, 48, Elsevier Science B.V., Amsterdam, 99–138, 2004.

Shiau, J. T. and Shen, H. W.: Recurrence analysis of hydrologic droughts of differing severity, J. Water Res. Planning and Management (ASCE), 127(1), 30–40, 2001.

Smakhtin, V. U. and Huges, D. A.: Review, automated estimation and analyses of drought indices in South Asia, Working Paper 83, International Water Management Institute, Colombo, Sri Lanka, 2004.

Stahl, K.: Hydrological Drought – a Study across Europe. PhD Thesis Albert-Ludwigs-Universität Freiburg, Freiburger Schriften zur Hydrologie no. 15, Freiburg, Germany, also available at: http://www.freidok.uni-freiburg.de/volltexte/202, 2001.

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency analysis of extreme events, in: Handbook of Hydrology, edited by: Maidment, D. R., McGraw-Hill, Inc., New York, NY, 18.1–18.66, 1993.

SUR in English: Alfarnatejo, available at http://canales.diariosur.es/fijas/pueblos/ingles/alfarnatejoi.htm, 2005.

Tallaksen, L. M.: Streamflow drought frequency analysis, in: Drought and Drought Mitigation, edited by: Vogt, J. V. and Somma, F., Advances in Natural and Technological Hazards Research, vol. 14, Kluwer Academic Publishers, Dordrecht, the Netherlands, 103–117, 2000.

Tallaksen, L. M. and van Lanen, H. A. J. (Eds.): Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater, Developments in Water Sciences 48, Elsevier B.V., Amsterdam, 2004.

Tallaksen, L. M., Madsen, H., and Clausen, B.: On the definition and modelling of streamflow drought duration and deficit volume, Hydrol. Sci. J., 42(1), 15–33, 1997.

Tallaksen, L. M., Madsen, H., and Hisdal, H: Frequency Analysis, in: Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater, edited by: Tallaksen, L. M. and van Lanen, H. A. J., Developments in Water Science, 48, Elsevier Science B.V., Amsterdam, 199–271, 2004.

Tate, E. L. and Freeman, S. N.: Three modelling approaches for seasonal streamflow droughts in southern Africa: the use of censored data, Hydrol. Sci. J., 45(1), 27–42, 2000.

The Times: Atlas of the World – Comprehensive Edition, Times Books, a division of Harper Collins Publishers, London, 1994.

Vogel, R. M. and Stedinger, J. R.: Generalized storage-reliability-yield relationships, J. Hydrol., 89, 303–327, 1987.

Wilhite, D. A. and Glantz, M. H.: Understanding the Drought Phenomenon: The Role of Definitions, Water Int., 10, 111–120, 1985.

Woo, M. and Tarhule, A.: Streamflow droughts of northern Nigerian rivers, Hydrol. Sci. J., 39(1), 19–34, 1994.

Yevjevich, V.: An objective approach to definitions and investigations of continental hydrologic droughts, Hydrology Papers 23, Colorado State University, Fort Collins, USA, 1967.

Yevjevich, V.: Methods for determining statistical properties of droughts, in: Coping with droughts, edited by: Yevjevich, V., da Cunha, L., and Vlachos, E., Water Resources Publications, Colorado, 22–43, 1983.

Zelenhasić, E. and Salvai, A.: A Method of Streamflow Drought Analysis, Water Resour. Res., 23(1), 156–168, 1987.