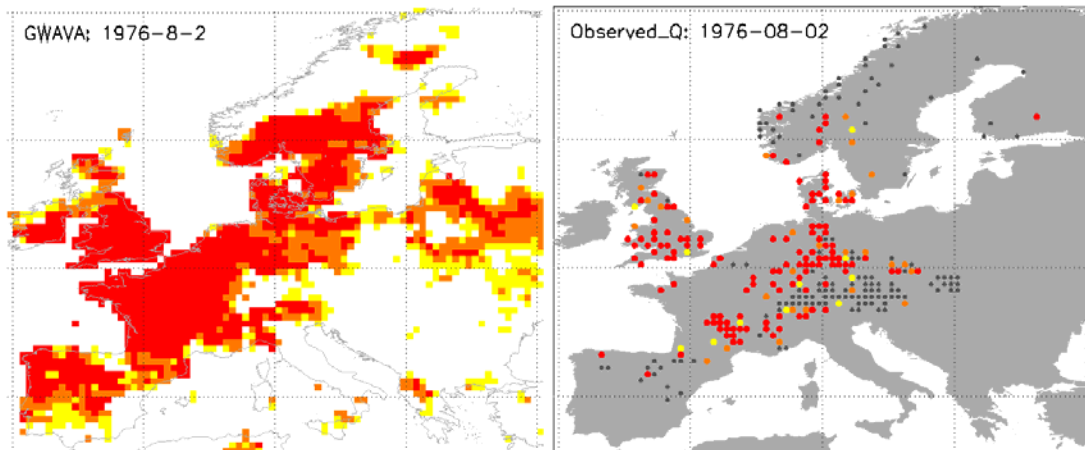




Technical Report No. 48

**SPACE-TIME CHARACTERISTICS OF LARGE-SCALE
DROUGHTS IN EUROPE DERIVED FROM
STREAMFLOW OBSERVATIONS AND WATCH
MULTI-MODEL SIMULATIONS**



Author names: Tallaksen, L.M., Stahl, K. & Wong, G.

Date: 31 July 2011



WATCH is an Integrated Project Funded by the European Commission under the Sixth Framework Programme, Global Change and Ecosystems Thematic Priority Area (contract number: 036946). The WACH project started 01/02/2007 and will continue for 4 years.

Title:	Space-time characteristics of large-scale droughts in Europe derived from streamflow observations and WATCH multi-model simulations
Authors:	Tallaksen, L.M., Stahl, K. & Wong, G.
Organisations:	University of Oslo
Submission date:	31 July 2011
Function:	This report is an output from Work Block 4; task 4.2
Deliverable	WATCH deliverable 4.2.3

Abstract

This report presents a detailed analysis of major historical droughts in Europe in the last part of the 20th century using novel methods for drought characterization that account for spatial extent and severity. The analyses were based on observed streamflow as well as runoff simulated by the WATCH multi-model ensemble comprising nine large-scale models. The ensemble median of the models was found to perform better than the ensemble mean when comparing two benchmark indices against observed streamflow for nine different anomaly levels. Based on the ensemble median and the 20% threshold from the empirical distribution (non-exceedance frequency), two extended drought periods (40% or more of the grid cells in drought) were identified, namely the autumn 1975 to late summer 1976, and spring and summer of 1990. A detailed spatial and temporal analysis revealed major differences in the drought development of the two events (build-up, consolidation and recovery phase), and in the models' ability to reproduce their behaviour. The occurrence of larger and longer drought events agreed largely among models, however, the spatial coverage and variability within the drought affected area varied considerably. In the last part of the report a comparison of various methods for deriving Severity-Area-Frequency (SAF) curves for different return periods is made, suggesting that the drought clustering method performs better than methods that do not account for spatial continuity.

List of content	page
1. Introduction	1
2. Data	3
2.1. Modelled runoff	3
2.2. Observed streamflow	3
3. Multi-model drought analysis for Europe	3
3.1. Background	3
3.2. Methods	4
3.2.1. Definition of drought days	4
3.2.2. Validation of drought days and years	4
3.2.3. Drought characteristics	4
3.3. Results	4
3.3.1. Anomaly validation	4
3.3.2. Drought characteristics	5
3.4. Discussion	9
4. Severity-Area-Frequency (SAF) curves	10
4.1. Background	10
4.2. Drought characteristics	10
4.3. Methods	11
4.3.1. Spatial and temporal analysis	11
4.3.2. SAF curves	11
4.3.3. Drought clustering method	11
4.4. Results and discussion	12
5. Conclusions	14
6. References	15

1. Introduction

Drought is an extreme, but temporary water shortage relative to the average (natural) condition of a region. It can be defined as a “a sustained and regional extensive occurrence of below average natural water availability” (Tallaksen & van Lanen, 2004), thus considering both the temporal and spatial aspect of the drought.

The primary cause of a drought is the lack of precipitation over a large area and for an extensive period of time, called a *meteorological drought*. The water deficit propagates through the hydrological cycle and gives rise to different types of drought. Combined with high evaporation rates a soil water deficiency may cause a *soil moisture drought* to develop. Subsequently, groundwater recharge and streamflow will be reduced and a *hydrological drought* may develop. A reduced recharge leads to lower groundwater heads and storage and finally to low river flows or even dried up river. Thus, drought has a wide range of impacts depending on the scale of the event and which components of the hydrological cycle are most severely affected. These include major social, economic and environmental impacts and Bradford (2000) and Wilhite (2000) give extensive reviews of the wide range of impacts that have been reported.

Drought frequently covers large areas, extending across several catchments or river basins, and manifest itself as a regional rather than a local event. Consequently, not only the severity and duration of a drought, but also the spatial extent and temporal aspects such as a drought’s persistence are considered important characteristics of the event (e.g. Andreadis *et al.*, 2005; Hisdal & Tallaksen, 2003). The generalized spatial extent of major drought events in Europe in the last decade is illustrated in Figure 1.1, and no systematic pattern in the area affected can be seen.

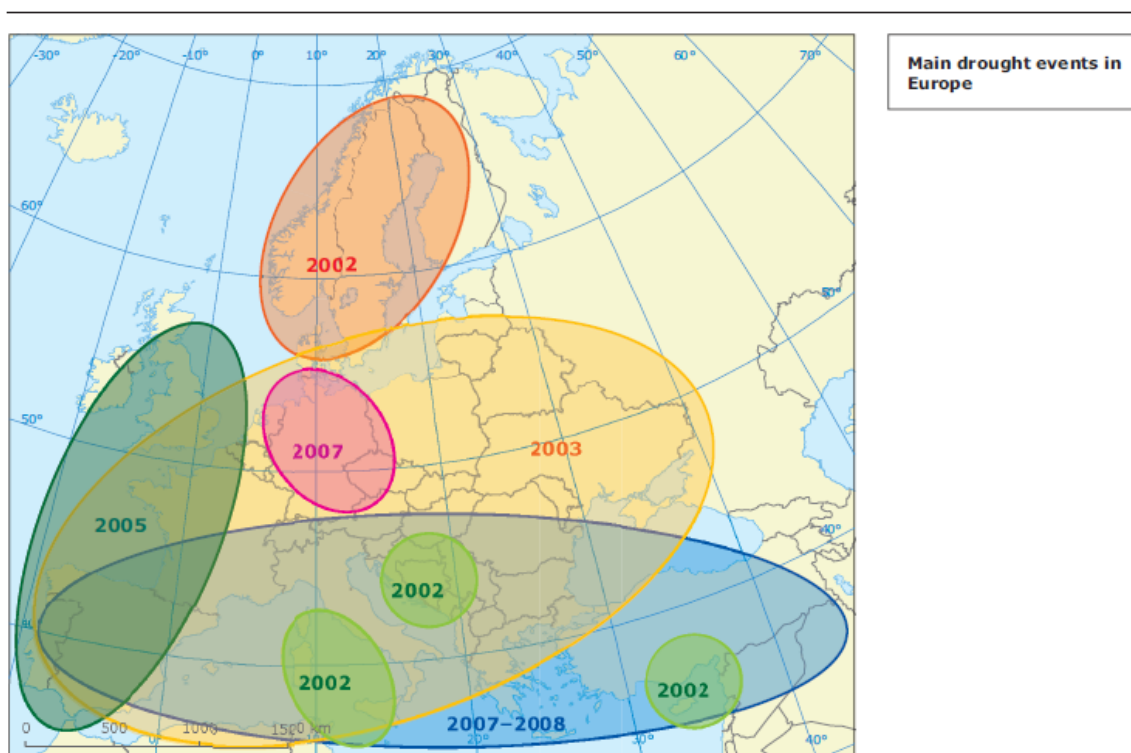


Figure 1.1 Recent extreme drought events in Europe and their approximately area coverage.

Recent publications have addressed the need to investigate the spatial and temporal characteristics of drought across regional and continental scales and different methodologies have been suggested as reviewed by Mishra & Singh (2011) and Sheffield & Wood (2011). The studies are motivated by the need to develop methodologies suited for validation of large-scale model (climate and hydrological) output in terms of drought, and further to predict based on these model output, useful and conveyable information about future changes in extremes. A net overall global drying is projected in recent studies as reviewed by Dai (2011), and more intense droughts affecting an increasing number of people have been observed globally since the 1970s in Africa, the Mediterranean, southeast Asia and eastern Australia (IPCC, 2007). Hence, quantifying the impact of drought through determining its associated characteristics, such as frequency, severity and spatial extent is essential in mitigating the effects of extreme drought.

Overall, droughts are caused by regional- or large-scale spatial and temporal anomalies in the climatic system. Regional drought-causing atmospheric situations are characterized by: i) an anomalous timing of a seasonal phenomenon; ii) an anomalous location of pressure centers and the track of cyclones; and iii) an anomalous persistence or persistent recurrence of dry weather patterns (Stahl & Hisdal, 2004). In a temperate climate like Europe, a meteorological drought is caused by persistent atmospheric circulation anomalies. In addition, its slow development allows for land-atmosphere feedbacks that may exacerbate and further expand the drought area (Seneviratne *et al.*, 2006). In the Mediterranean region, with its seasonal climate, severe droughts can for instance be caused by longer than usual influence of the subtropical high-pressure belt. Meteorological droughts can accordingly last several weeks or even months. In the mid-latitudes of western and northern Europe, "*blocking action*" is the major atmospheric anomaly to cause extended dry weather periods (Tallaksen & Fleig, 2009).

Extreme hydrological droughts are likely to be caused by a combination of climate anomalies and already low antecedent catchment storage (e.g. soil moisture, groundwater and lake storage). In a fast responding catchment a drought in streamflow during summer is normally caused by the weather conditions prevailing during the same summer, whereas in catchments with large storages, the preceding winter or even preceding years, can be as important (Tallaksen *et al.*, 2009). Dry winters can, for instance, be the cause of significant water resources stress in the following summer even when followed by above average summer rainfall. The wetness status of the land surface is also an important memory component of the climate system through important land-atmosphere interactions.

The derivation of drought events from a time series commonly refers to a certain threshold that distinguishes a drought event from a non-drought situation, and the event has thus a beginning and an end. A threshold might be imposed to derive drought events for each area unit (e.g. gridded model output). In some cases also an areal threshold is introduced, which implies that a certain fraction of the area should be in drought for the drought to start or continue (e.g. Santos, 1983; Tallaksen *et al.*, 2009). The threshold is frequently chosen to be a quantile (or equivalent percentile) of the empirical distribution, either fixed, seasonal dependent or monthly (daily) varying. In the latter case, the threshold can take on different values for each month (day) or a continuous moving average procedure can be adopted and the events referred to as anomalies rather than droughts.

In this study, focus is on hydrological drought, i.e. streamflow drought. Following a description of the data used, two specific drought studies are presented. The first presents a multi-model drought analysis for Europe focusing on the models ability to simulate major drought events in terms of their spatial and temporal dynamics, whereas the second study is concerned with quantifying the spatial and temporal characteristics of hydrological droughts in Europe through the use of severity-area-frequency (SAF) curves.

2. Data

Space-time characteristics of drought can be analysed based on either observations or model output. Observation would typically be point measurements (e.g. rainfall) or areal aggregations of hydrological variables (e.g. streamflow). Model output would normally be average values over a given area unit like a grid cell. In this study, modelled runoff at the grid cell scale is compared to observed streamflow in small river basins.

2.1 Modelled runoff

This study used output from nine of the WATCH models, i.e. GWAVA, HO8, HTESEL, JULES, LPJml, MATSIRO, MPI-HM, Orchidee and WaterGAP. Details of the models are given in Haddeland *et al.* (2011) and Gudmundsson *et al.* (2011a), including an overview of the schemes used for simulation of evapotranspiration, runoff generation and snowmelt. All nine models were run for the period 1963-2001 on a global 0.5 degree grid and forced by the WATCH Forcing Data (WFD). WFD were derived from the ERA-40 reanalysis product, interpolated to half-degree resolution and bias-corrected based on CRU and GPCC (Weedon *et al.*, 2011). The variable used here is daily total runoff (sum of the fast and slow runoff component) simulated for each grid cell in Europe (4425 land grids).

2.2 Observed streamflow

Streamflow observations from across Europe were available from the combined dataset of the European Water Archive of the UNESCO IHP FRIEND programme and the WATCH project (Stahl *et al.*, 2010). This dataset contains over 400 near-natural streamflow records for the period 1962-2004, hence covering the modelled period 1963-2000. As the basins are subscale to the model grid, each gauged river basin was first assigned to the 0.5 degree model grid cell in which its centroid lies. For model grids with more than one river basin, only the record from the largest basin within a grid cell was kept for the analysis, resulting in a dataset of 293 daily streamflow records that could be paired with model simulations of the corresponding grid cells. All daily streamflow data were converted to mm of runoff per unit area.

3. Multi-model drought analysis for Europe

3.1 Background

Hydrological drought occurrence, areal extent and temporal evolution are climate driven primarily, whereas variability within the affected area and persistence are influenced by land properties. This study explores how different large-scale hydrology models represent the latter aspects for major historical droughts in Europe in the period 1963 and 2000.

Model performance, i.e. the simulation of drought for different thresholds (or anomaly levels), were first validated with streamflow drought observations from a large data set of streamflow records (Section 2.2). The validation followed the anomaly and benchmark approach by Stahl *et al.* (2011, in press) with the aim of i) determining an appropriate threshold for drought definition and ii) testing whether the multi-model mean or the median adequately represents drought situations. Further, major historic droughts in Europe in terms of area coverage were selected and the dynamics in time and space were visualized to explore how the model ensemble simulates these events. The overall objective is to develop a better understanding of the robustness of drought characteristics, including variability, location and persistence, as simulated by large-scale hydrological models.

3.2 Methods

3.2.1 Definition of drought days

Hydrological drought days were defined as days when the total simulated runoff or observed streamflow was below a given threshold. Similar to the methodology used in many drought studies based on monthly data, the threshold z is defined by a given non-exceedance frequency (or quantile) of the empirical runoff distribution on the respective day of the year. Drought days were considered for non-exceedance frequency thresholds of $z=10, 20$ or 30% (here expressed as percentiles equivalent to Q90, Q80, Q70 in the exceedance notation common in hydrology). Time series of daily runoff in each grid cell, from each of the nine models, were first converted individually into their daily percentiles (values from 0 to 100). In addition, for each grid cell two summary time series were created: one describing the multi-model ensemble mean and one the ensemble median.

3.2.2 Validation of drought days and years

Two indices for benchmarking simulations of hydrological event dynamics introduced by Stahl *et al.* (2011) were used in this study. The indices are based on the agreement of the two corresponding series of simulated runoff and observed streamflow below or above a particular anomaly level z . Thus, they allow different properties related to dry and wet spells (events) to be evaluated.

The agreement index describes the overall agreement for each anomaly level z of the daily runoff, whereas the correlation coefficient is determined based on observed and simulated annual sums of days below or above z and thus, is a measure of how well the interannual variability of wet and dry years agree. Details about the indices can be found in Stahl *et al.* (2011). For thresholds such as 10, 20 or 30%, the approach effectively validates the simulated occurrence of drought days and the year-to-year variability of drought years with observations. It further provides a basis for choosing a suitable threshold for the derivation of other drought characteristics.

3.2.3 Drought characteristics

For the European domain the following daily characteristics were derived from the gridded time series of ensemble members and the ensemble median:

- **Spatial distribution of drought:** daily maps of the spatial distribution of grid cells in drought for different models as well as the ensemble median (colour coded grid cells for different thresholds)
- **Total drought area:** daily time series of the sum of grid cells in drought (in % of all land grid cells of the domain)
- **Spatial clustering:** daily time series of the mean area of contiguous clusters (minimum 2 grid cells) of drought (in % of the total drought area)
- **Location of drought centre:** daily time series of the centroid of the drought deficit, i.e. the volume below the threshold in all drought affected grid cells is used to derive a 2D centre of gravity (in lat/long coordinates)

The two largest historical events were then selected based on the daily series of total drought area covering $>40\%$, and their characteristics explored in detail. In addition, a ranking of the top ten drought years was established based on annual averages of total drought area.

3.3 Results

3.3.1 Anomaly validation

The agreement index and the correlation coefficient for the ensemble mean and median are shown in Figure 3.1 for nine anomaly levels. The agreement of drought and high flow days is better for more

moderate thresholds, as can be expected. Correlation of the inter-annual sums of drought and high flow days, however, is relatively insensitive to the threshold for many basins, particularly for the ensemble median. Only the most extreme threshold causes a rather low correlation for about 25% of the basins.

The thresholds considered for drought analysis are 10, 20 or 30%. The agreement index is particularly low for the ensemble mean and the days below the 10 and 20% thresholds. Agreement of observed and simulated drought days for the ensemble median below an anomaly level of 20% is considerably higher. The ensemble median appears to be a better representation than the ensemble mean. Thus, the 20% threshold was chosen for further analysis as the gain in the agreement index and correlation coefficient from a 20 to a 30% threshold is moderate.

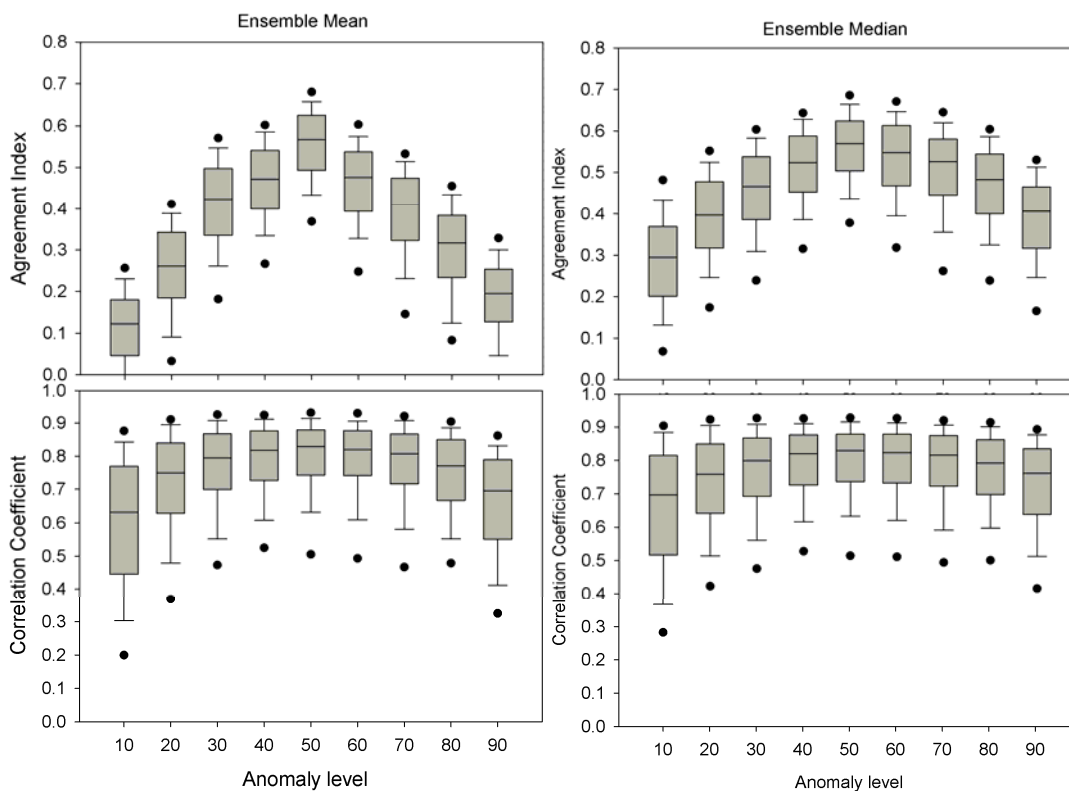


Figure 3.1 Agreement index of daily runoff and correlation of annual drought sums of days below and above particularly anomaly levels (thresholds) as defined in Stahl *et al.* (2011). Box plots show the distribution of the indices for 293 pairs of observations and simulations across Europe.

3.3.2 Drought characteristics

Based on the ensemble median and 20% threshold, two extended periods were identified with 40% or more of the grid cells in drought. These are autumn 1975 to late summer 1976 and spring and summer of 1990. Two shorter periods with an areal coverage greater than 40% of the domain occurred in January 1964 and in November 1983.

The evolution of the area in drought and spatial clustering of the 1975/76 drought are shown in Figure 3.2. While most models show a similar relative development, the total drought area and the level of clustering vary substantially. The drought can be described in three phases. During the first phase (build-up phase) throughout the autumn 1975, the total drought area increased and peaked in the

beginning of November 1975. The variation among models during this build-up phase is high with a spread of about 20% in drought area.

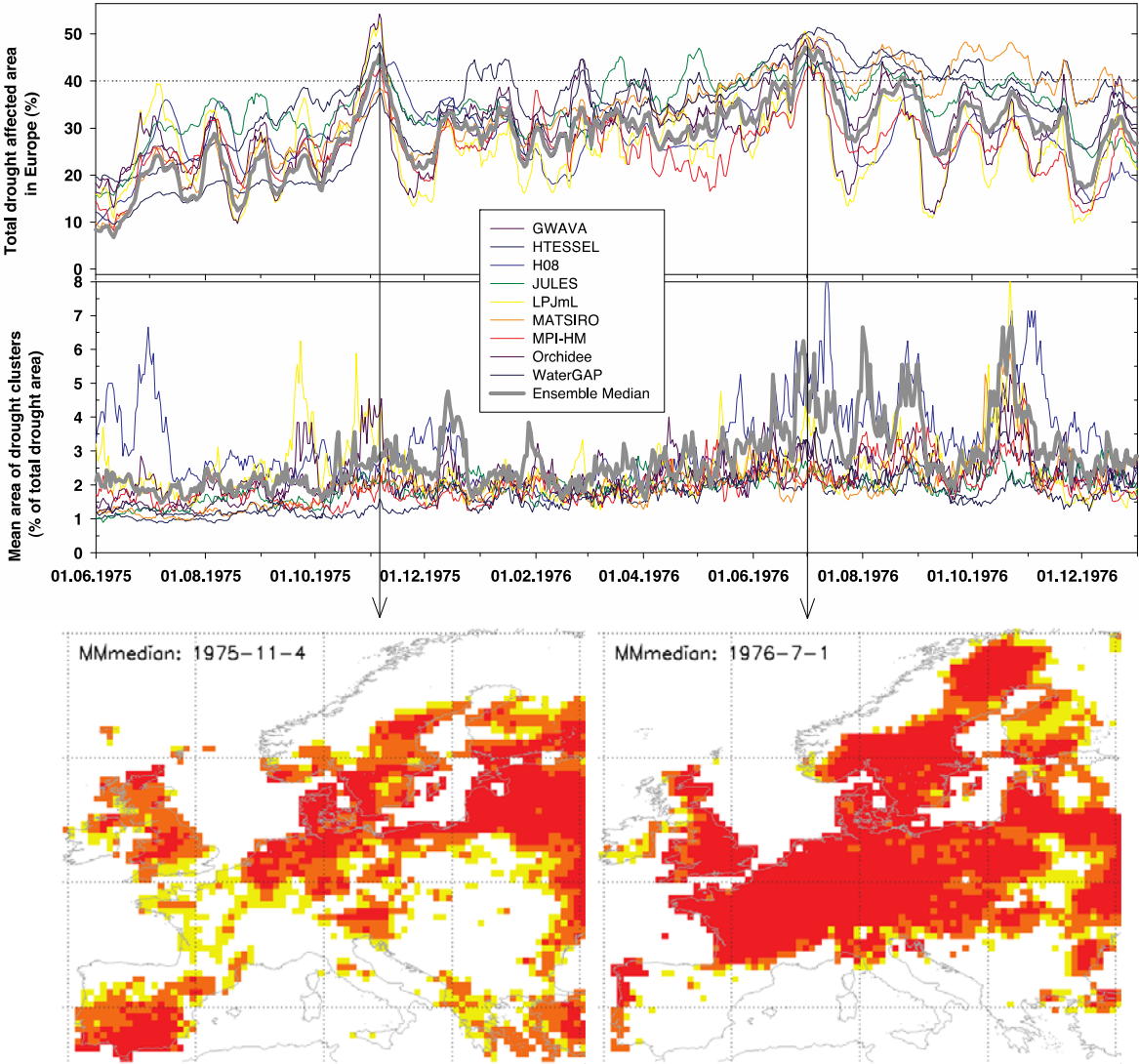


Figure 3.2 Evolution of total drought area and spatial drought clustering in 1975 and 1976. The two daily snapshots in the lower graphs illustrate the spatial distribution of drought. Grid cells with drought defined as runoff below the 10% threshold (red), 20% threshold (orange), and 30% threshold (yellow).

The build-up phase is followed by a period with constantly high drought areas of around 30% of the domain in the beginning of the year 1976. During this consolidation phase, models show less spread. Spatial clustering (mean area of drought clusters) remains relatively low during these first two phases of the drought, except for shorter periods of a few days. However, in the summer of 1976 both total drought area as well as spatial clustering increase, peaking at the end of June/beginning of July (peak phase). The spatial distribution of anomalies shown in the snapshots (maps in Figure 3.2) illustrates the differences in spatial clustering for the two peaks. The November peak has several disconnected areas that are affected during the build-up phase, whereas the summer 1976 peak covers a contiguous area across Europe.

Throughout the third phase (recovery phase) of the 1976 drought from July to the end of the year, models again vary strongly. The total drought area shows in some models a high degree of persistence around 35% and only a slow decline over time with small fluctuations. In other models, the total drought area fluctuates strongly between 10 and 40%. The spread in simulated area can be as large as 30% during this time. The variation among models in the spatial clustering increases over this recovery period with decreasing area in drought. The ensemble median generally shows a high spatial clustering within the range covered by the individual models.

The differences among models in location and area covered can also be visualized using the derived drought centroid. Figure 3.3 shows the development of the peak phase of the 1976 drought (15-day time steps). It displays the drought centroid and indicates the drought area by a circle scaled to the total area covered on the particular day. This allows plotting all models and the ensemble median (grey colour) in one map for comparison. The figure shows that the drought centre in northern central Europe is relative stable. Only after September it starts to move north and eastward (not shown). At times models are very similar in their representation of the drought, however, some models may 'break out' for a few days, but eventually the common pattern re-establish.

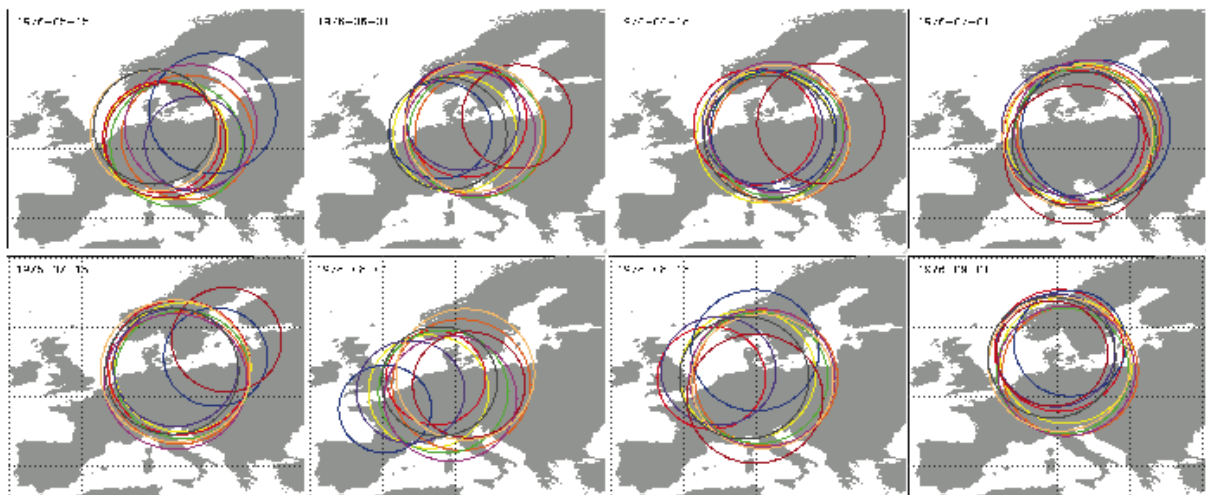


Figure 3.3 Drought centroid and relative area covered for all models (colours) and the ensemble median (grey circle) during the peak phase of the 1976 drought in Europe.

The second drought with a total drought area reaching 40% and more, occurred in 1989-1990. Its evolution and characteristics show many similarities, but also distinct differences, to the 1976 event. Similar to 1976, the main event in summer 1990 is also preceded by considerable anomalies in the year before. The build-up and consolidation phase over the winter are not as clear as in 1976, but a high degree of spatial clustering indicates severe droughts in particular areas (Figure 3.4). Visualization of the snapshots show that after the autumn drought in Western Europe, the Mediterranean was affected over the winter, before the main drought area spreads again to the north (not shown). The peak phase in the summer of 1990 occurs already in May, but with lower spatial clustering (Figure 3.4) than in 1976.

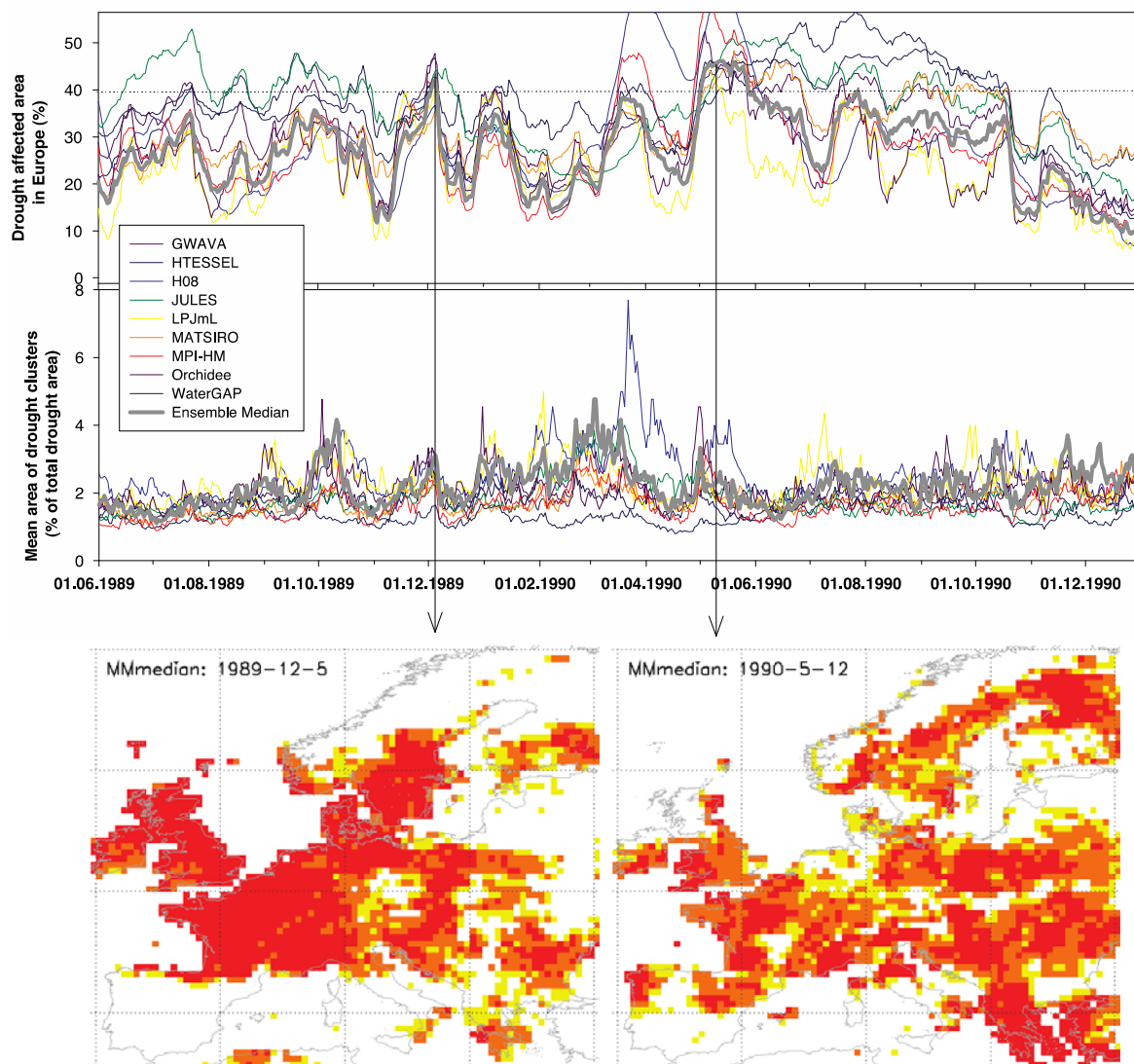


Figure 3.4 Evolution of total drought area and spatial drought clustering for the 1989/90 drought. The two daily snapshots in the lower graphs illustrate the spatial distribution of drought. Grid cells with drought defined as runoff below the 10% threshold (red), 20% threshold (orange) and 30% threshold (yellow).

In spring 1990, two drought centres in France and south-eastern Europe combined into one large European drought. Figure 3.5 shows the location development over the summer. Differences among the models are larger than for the 1976 event, but the central location is similar. However, the snapshots reveal that central European countries (e.g. Germany, Switzerland, etc.) were not directly affected, but as strong droughts persisted in Greece, France as well as the northeast, the centroid of the combined event lies in the centre.

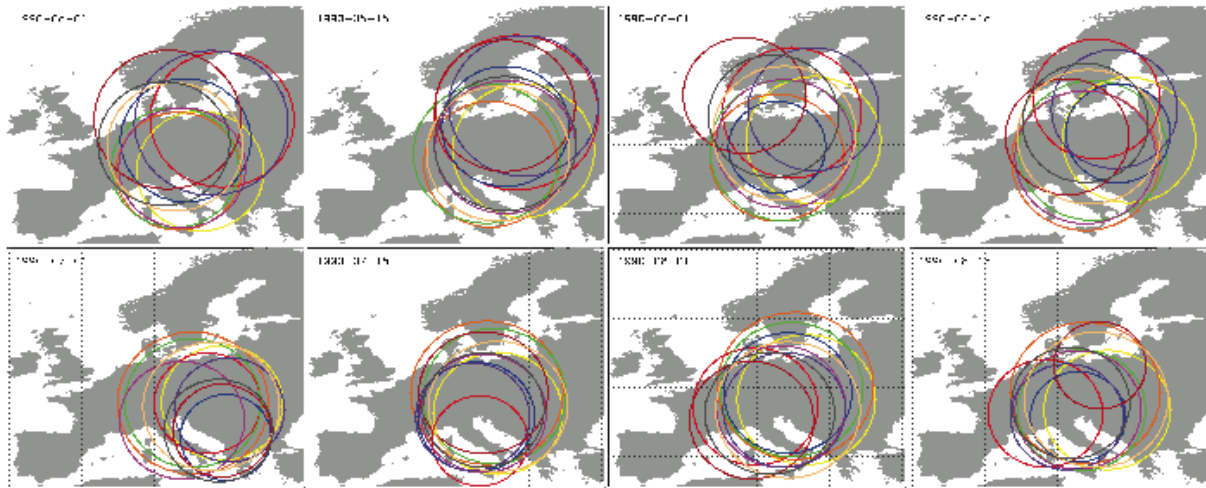


Figure 3.5 Drought centroid and relative area covered for all models (colours) and the ensemble median (grey circle) during the peak phase of the 1990 drought in Europe.

The aggregation of the daily characteristics to a mean annual drought area allows the top ten drought years simulated with each model and by the ensemble median to be selected. Table 3.1 shows the rank of these events and the two previously discussed events rank first or second in most models. In addition, the year 1989 described earlier as leading up to the 1990 event, ranks high on its own. For the subsequent ranks, the order of the models varies somewhat, but still has a substantial overlap. It is also notable that drought years seem to cluster, particularly the early 1990s appear frequently in the list.

Table 3.1 The ten most extreme drought years in Europe in terms of the average annual drought area simulated by the multi-model ensemble and the ensemble median

Rank	Median	GWAVA	HTESSEL	H08	JULES	LPJml	Matsiro	MPI-HM	Orchidee	WaterGAP
1	1976	1976	1976	1990	1976	1976	1964	1990	1976	1976
2	1990	1989	1990	1976	1989	1964	1963	1976	1964	1990
3	1964	1990	1964	1989	1990	1989	1976	1993	1990	1989
4	1989	1973	1989	1964	1992	1996	1990	1994	1989	1964
5	1996	1993	1992	1993	1964	1993	1992	1989	1972	1973
6	1993	1964	1996	1983	1993	1990	1996	1969	1996	1996
7	1992	1996	1993	1994	1996	1992	1965	1983	1973	1972
8	1973	1992	1974	1996	1995	1969	1993	1992	1993	1992
9	1994	1991	1975	1992	1973	1963	1991	1996	1992	1993
10	1972	1972	1973	1973	1975	1994	1973	2000	1991	1974

3.4 Discussion

Due to the scarcity of observations a direct validation of the drought characteristic 'area' is not possible, however, the validation with individual river basins does give confidence, particularly in the use of the multi-model median anomalies as a basis for drought analysis. It was found that the occurrence of larger and longer drought events agree largely among models, which suggests a strong influence of the common forcing. However, the exact spatial dimensions and the variability within the drought affected area (implicitly expressed in the spatial clustering characteristic) varied considerably among the models. In particular, the building-up and recovery phase of a major drought event is modelled differently. This likely can be related to the conceptualization of hydrological processes, in particular water storage and release, in the different models.

4. Severity-Area-Frequency (SAF) curves

4.1 Background

This study aims at improving our understanding of hydrological extremes for different areal extents by establishing methods of identifying the associated frequency, severity and scale of these extremes. More specifically, it is concerned with quantifying the spatial and temporal characteristics of hydrological droughts in Europe and expressing this interaction between the spatial and temporal characteristics through the use of severity-area-frequency (SAF) curves. Different methodologies for evaluating the spatial and temporal analysis of hydrological drought in Europe are investigated and compared using the concept of SAF curves.

4.2 Drought characteristics

Runoff simulations from the large-scale hydrological model GWAVA (Meigh *et al.*, 1999), which has shown to perform satisfactory in previous studies (Haddeland *et al.*, 2011), was selected for this study. Details of the simulation are given in Section 2.1. The main aim was to demonstrate and evaluate different SAF approaches using data from the whole European domain. For each grid cell, the total runoff was derived (sum of subsurface and surface runoff), and for consistency and easy comparison of drought characteristics, monthly varying drought thresholds were adopted (derived from monthly quantile series). These were calculated for each grid cell relative to their corresponding month and each monthly runoff is replaced by their respective quantile, ranging from zero to 1. It is common to use an exceedance threshold of 0.2 (or 20% if express as a percentile) to define drought events. Hence, more extreme droughts are represented by lower-value quantiles. For each month, a drought map based on these quantiles is plotted. An example is given in Figure 4.1 for the European summer drought of 1976.

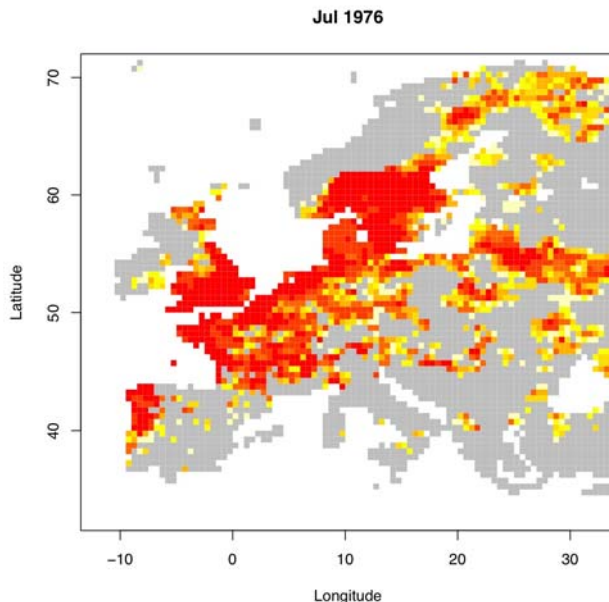


Figure 4.1 Drought map for July 1976, based on simulations from the GWAVA model; 10% threshold (red), 20% threshold (orange) and 30% threshold (yellow).

The spatial extent of a drought can be large without causing as much impact as a smaller scale, but more intense drought. Hence, it is essential that the overall drought deficit is considered. For each month, the total drought deficit were calculated relative to the drought threshold of each cell and divided by the number of grid cells in drought for that month, to give the average drought deficit.

4.3 Methods

4.3.1 Spatial and temporal analysis

Preliminary analysis of the drought spatial extent over the 38-year period indicated an evident wet and dry decadal cycle. The GWAVA model shows an overall good agreement with the observations, with the exception of a number of extreme drought events in the early 1970s and mid 1990s (Figure 4.2).

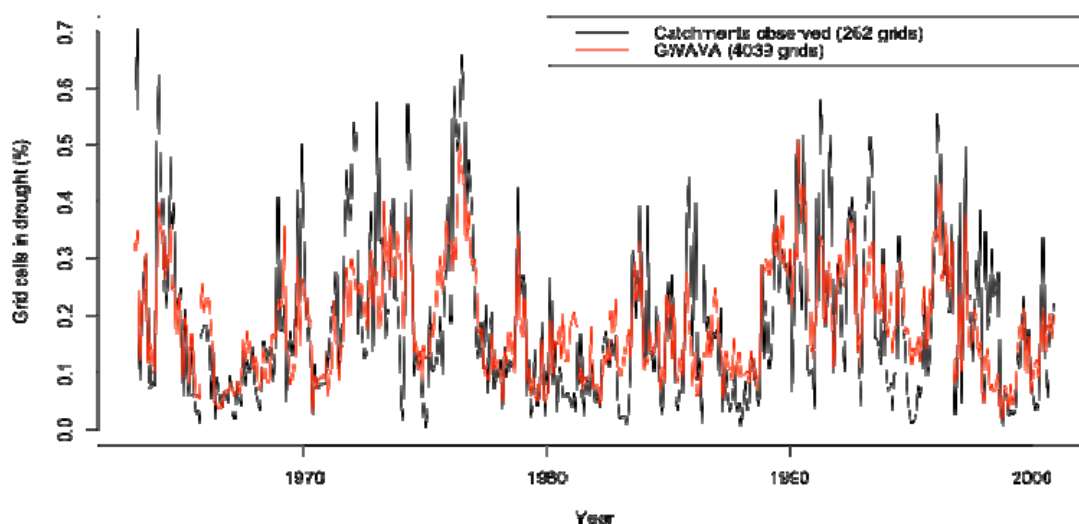


Figure 4.2 Monthly percentage of land area in drought for the 293 cells with paired data: observations (black) and GWAVA model (red), 1963-2000.

4.3.2 SAF curves

Different methods of constructing Severity-Area-Frequency (SAF) curves were explored. The spatial independence (deficit of all grid cells in drought) and spatial smoothing (average deficit over 2x2 and 3x3 grid cells in drought) methods do not take into account the spatial relationships of drought. These two methods demonstrated a higher sensitivity to change in mean drought severity for larger drought areas as compared to the drought clustering method, whereas a higher sensitivity to change in mean drought severity at smaller spatial extent was found using the drought clustering technique. In the following, the details and results of the drought clustering method are presented.

Spatial areas covering from 1-10% and 15% of Europe's land area, or approximately 100,000 km² to a maximum drought extent of 1482,000 km², are considered. These percentages were chosen to be reasonable since the number of months, which has at least 15% of land area in drought, is only 256 months out of a total of 456 months (here the interest is in studying drought behaviour for extreme scenarios). The number of pixels corresponding to the percentage of land area in drought is calculated based on Europe's total land area. In addition, the drought recurrence intervals of 10, 30, 50 and 100 years are used to examine the behaviour of the SAF curves.

4.3.3 Drought clustering method

In this method, spatial considerations are taken into account using a simple clustering algorithm (Andreadis *et al.*, 2005; Sheffield and Wood, 2008). Since droughts are regional phenomena, the spatial characteristics and distribution of a drought is essentially the contiguity of its extent. Based on the

calculation of the drought deficit, the severity of all grid cells are first calculated, that is $severity = 1 - drought\ deficit$. For each month, the severity of those grid cells that fall below the drought threshold, are then ranked in an ascending order according to their severity and the grid cell with the highest severity is chosen as cluster centre. Figure 4.3 provides an illustration of the procedure, which in the first time step identifies all grid cells in drought and the one with the highest severity is chosen as the cluster centre (Figure 4.3(a)).

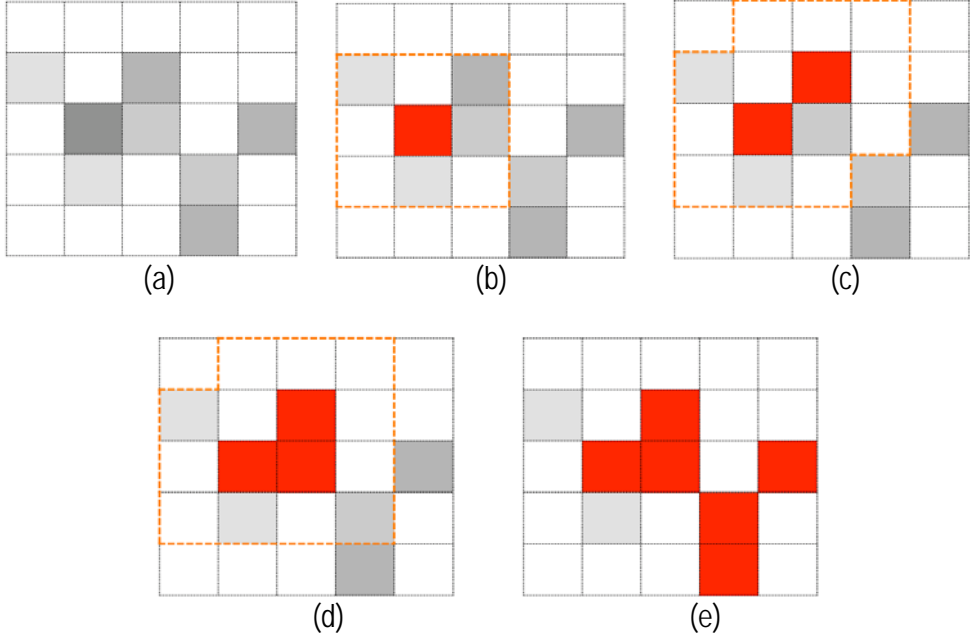


Figure 4.3 Procedure for the drought clustering method.

In instances where ties occur, the highest ranked severity is chosen, as shown in Figure 4.3(b), and its surrounding neighbourhood is searched for grid cells in drought with the second most severe drought quantile. The new neighbourhood is then extended, comprising the previous one and the current neighbourhood, and again searched for the next most severe grid cell (Figure 4.3(c)). This procedure is then repeated until it satisfies the number of grid cells required for 1-10% and 15% spatial extent (Figures 4.3(d) and (e)). The mean drought severity of this cluster is then calculated. If there is more than one cluster of similar size, then the cluster area with the maximum mean drought severity is chosen to represent that particular month. Similarly, if for a given month, there is no drought cluster that satisfies the specified spatial extent, then this is not recorded. This procedure is repeated for each month and is evaluated for the drought percentage area selected (1-10% and 15%).

4.4 Results and discussion

Table 4.1 shows the basic statistics of the mean drought severity obtained from the drought clustering method, and a general decrease in the number of months that satisfy the spatial extent given is seen, especially for larger drought areas. However, there is less variation between the mean drought severities of smaller percentage drought areas as compared to larger percentage drought areas. Figure 4.4 displays the corresponding SAF curves.

Table 4.1 Basic statistics of mean drought severity associated with 1-10% and 15% of drought area, using the drought clustering method

Percentage of drought area (No. of pixels)	Length (months)	Mean Severity	Standard deviation
1% (40)	442	0.953	0.025
2% (79)	396	0.945	0.025
3% (119)	354	0.941	0.025
4% (158)	301	0.939	0.025
5% (198)	257	0.938	0.024
6% (237)	233	0.937	0.023
7% (277)	210	0.935	0.021
8% (317)	184	0.934	0.022
9% (356)	163	0.933	0.021
10% (395)	145	0.932	0.021
15% (593)	77	0.931	0.019

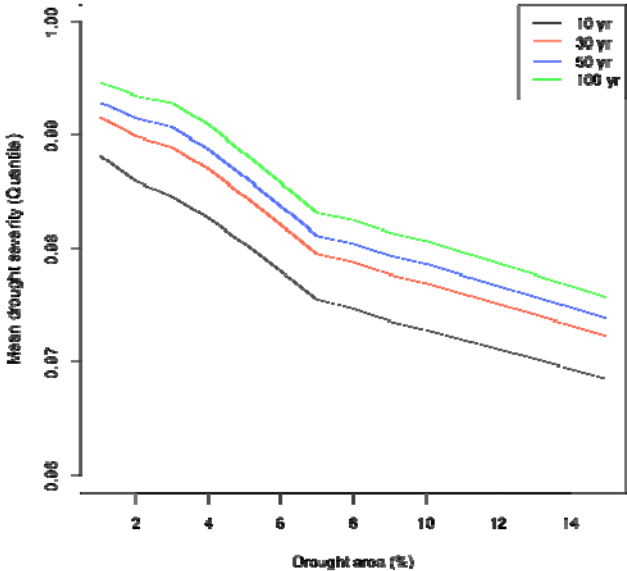


Figure 4.4 Severity-Area-Frequency (SAF) curves of 10, 30, 50 and 100 year return periods, using the drought clustering method.

The SAF curves appear markedly different from those obtained from the spatial independence and spatial smoothing methods. This can be attributed to the spatial clustering effect. Firstly, there is more variation in the mean drought severity for different return periods for smaller drought spatial extent. Furthermore, there is a higher sensitivity to change in mean drought severity for smaller drought areas, than observed for the other two methods. This suggests that after a certain drought area threshold, the mean drought severity does not vary much for larger areas. The largest difference observed is between SAF curves associated with the 10-year and 30-year return periods. There is also a noticeable change in the behaviour of the curves at 3 and 7%. The observed change at 3% can be attributed to quantization at 1 and 2% spatial coverage. On the other hand, the change at 7% may indicate a different drought clustering behaviour for small and large-scale droughts.

5. Conclusions

This report has presented a detailed analysis of major historical drought in Europe in the last part of the 20th century using novel methods for drought characterization that account for spatial extent and severity. The analysis included both streamflow observations and runoff simulated from the WATCH multi-model ensemble of nine large-scale models. The occurrence of the most extreme drought events agreed largely among the models, however, major differences were found in the models ability to reproduce the different stages in drought development (build-up, consolidation and recovery phase). Consequently, the choice of model will influence the estimation of drought statistics, such as the frequency and severity of extreme events, commonly derived from past and future hydrological model simulations at the large scale.

Similar large spread between models was observed in a diagnostic evaluation of the same set of WATCH models in terms of their ability to simulate runoff variability in Europe (Gudmundsson *et al.*, 2011a; Gudmundsson *et al.*, 2011b). Therefore, it can be concluded that any single model should be applied with care in large-scale modelling exercises and rather an ensemble of models should be used. The ensemble median was here found to perform better than the ensemble mean based on comparing two benchmark indices against observed streamflow for overlapping grid cells.

Different methods for deriving Severity-Area-Frequency (SAF) curves were explored based on simulated runoff from one of the ensemble models, and it was found that the drought clustering method performs better than methods that do not account for spatial continuity. In a further study, the multi-model ensemble will be used to derive extreme statistics for additional drought variables, like the annual maximum drought area and severity. Moreover, the detection of the dynamics of the drought affected area as identified for Europe in this study, has the potential to reveal links with large scale climate drivers and thus, ultimately improve drought forecasting.

References

- Andreadis, K.M., Clark, E. A., Wood, A.W., Hamlet A.F. and Lettenmaier, D. P. (2005) Twentieth-Century Drought in the Conterminous United States, *Journal of Hydrometeorology* 6, 985-1001.
- Bradford, R.B. (2000) Drought events in Europe. In: *Drought and Drought Mitigation in Europe* (ed. by J. Vogt & F. Somma), Advances in Natural and Technological Hazards Research, Vol. 14. Kluwer Academic Publisher, Dordrecht, the Netherlands, 7–20.
- Dai, A. (2011) Drought under global warming: a review. *Advanced review*, Vol. 2, 45-65, John Wiley & Sons, Ltd.
- Gudmundsson, L., Wagner, T., Tallaksen, L.M. & Engeland, K. (2011a) Seasonal evaluation of nine large-scale hydrological models across Europe (Water Res. Res., submitted)
- Gudmundsson, L., L. M. Tallaksen, K. Stahl, D. B. Clark, E. Dumont, S. Hagemann, N. Bertrand, D. Gerten, J. Heinke, N. Hanasaki, F. Voß, S. Koirala (2011b), Comparing Large-scale Hydrological Models to Observed Runoff Percentiles in Europe. *J. Hydrometeorol.* (submitted).
- Haddeland, I., D. B. Clark, W. Franssen, F. *et al.*, (2011), Multi-model estimate of the global terrestrial water balance: Setup and first results, *Journal of Hydrometeorology*, doi:10.1175/2011JHM1324.1. (in press)
- Hisdal, H. & Tallaksen, L.M. (2003) Estimation of regional meteorological and hydrological drought characteristics. *J. Hydrology*, 281(3), 230-247.
- IPCC (2007) Working Group I Fourth assessment report 'The Physical Science Basis'. Geneva, Switzerland.
- Meigh, J.R., McKenzie, A.A. & Sene, K.J. (1999) A grid-based approach to water scarcity estimates for eastern and southern Africa. *Water Resources Management* 13, 85-115.
- Mishra, A.K. & Singh, V.P. (2011) A review of drought concepts. *J. Hydrology*, 391, 202-216.
- Santos, M.A. (1983) Regional droughts: a stochastic characterization. *J. Hydrology*. 66, 183–211.
- Seneviratne, S.I., D. Lüthi, M. Litschi, and C. Schär, 2006: Land-atmosphere coupling and climate change in Europe. *Nature*, 443, 205-209.
- Sheffield, J. & Wood, E.F. (2011) Drought – Past Problems and Future Scenarios. Earthscan, May 2011, 192 pg.
- Sheffield, J. & Wood, E.F. (2008) Global Trends and Variability in Soil Moisture and Drought Characteristics, 1950-2000, from observation-driven simulations of the terrestrial hydrologic cycle. *Journal of Climate* 21, 432-458.
- Stahl, K. & Hisdal, H. (2004) Hydroclimatology. In: Tallaksen, L. M. & Lanen, H. A. J. van (Eds) *Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater*. Developments in Water Sciences 48, Elsevier Science BV, The Netherlands, pg. 19-51.
- Stahl, K., H. Hisdal, J. Hannaford, L. M. Tallaksen, H. A. J. van Lanen, E. Sauquet, S. Demuth, M. Fendekova, and J. Jordar (2010), Streamflow trends in Europe: evidence from a dataset of near-natural catchments, *Hydrology and Earth System Sciences*, 14 (12), 2367-2382, doi:10.5194/hess-14-2367-2010.
- Stahl, K., Tallaksen, L.M., Gudmundsson, L. & Christensen, J.H. (2011) Streamflow data from small basins: a challenging test to high resolution regional climate modeling. *Journal of Hydrometeorology*. doi: 10.1175/2011JHM1356.1 (in press)

- Tallaksen, L.M., Hisdal, H. & van Lanen, H.A.J. (2009) Space-time modeling of catchment scale drought characteristics. *J. Hydrol.*, 375, 363-372.
- Tallaksen, L.M. & Fleig, A. (2009) Atmosphere: climatic drivers of drought, Chapter 3. In: Wipfler, L. & van Lanen, H.A.J. (Eds.) Extended Guidance Document on the Natural System & Drought, XEROCHORE Technical Report, Wageningen, pg. 7-16.
- Tallaksen, L.M. & van Lanen, H.A.J. (2004) (Eds) *Hydrological Drought – Processes and Estimation Methods for Streamflow and Groundwater*. Developments in Water Sciences 48. Elsevier B.V., the Netherlands, 580p.
- Weedon, G. P., S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Sterle, J. C. Adam, N. Bellouin, O. Boucher, and M. Best (2011), Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology*, doi:10.1175/998 2011JHM1369.1 (in press)
- Wilhite, D.A. ed. (2000) *DROUGHT, A Global Assessment*, Vol I & II. Routledge Hazards and Disasters Series, Routledge, London.