On the objective identification of flood seasons

Juraj M. Cunderlik, Taha B. M. J. Ouarda, and Bernard Bobée

Institut National de la Recherche Scientifique, Eau, Terre, and Environnement, University of Québec, Sainte-Foy, Quebec, Canada

Received 30 April 2003; revised 21 October 2003; accepted 14 November 2003; published 31 January 2004.

[1] The determination of seasons of high and low probability of flood occurrence is a task with many practical applications in contemporary hydrology and water resources management. Flood seasons are generally identified subjectively by visually assessing the temporal distribution of flood occurrences and, then at a regional scale, verified by comparing the temporal distribution with distributions obtained at hydrologically similar neighboring sites. This approach is subjective, time consuming, and potentially unreliable. The main objective of this study is therefore to introduce a new, objective, and systematic method for the identification of flood seasons. The proposed method tests the significance of flood seasons by comparing the observed variability of flood occurrences with the theoretical flood variability in a nonseasonal model. The method also addresses the uncertainty resulting from sampling variability by quantifying the probability associated with the identified flood seasons. The performance of the method was tested on an extensive number of samples with different record lengths generated from several theoretical models of flood seasonality. The proposed approach was then applied on real data from a large set of sites with different flood regimes across Great Britain. The results show that the method can efficiently identify flood seasons from both theoretical and observed distributions of flood occurrence. The results were used for the determination of the main flood seasonality types in Great Britain. *INDEX TERMS:* 1821 Hydrology: Floods; 1860 Hydrology: Runoff and streamflow; 1869 Hydrology: Stochastic processes; *KEYWORDS:* flood seasonality, significance, sampling variability, nonparametric density, bootstrap resampling

Citation: Cunderlik, J. M., T. B. M. J. Ouarda, and B. Bobée (2004), On the objective identification of flood seasons, *Water Resour. Res.*, *40*, W01520, doi:10.1029/2003WR002295.

1. Introduction

[2] The proper determination of flood seasons is an important task with many practical applications in hydrology and water resources management. The information on flood seasonality is often used in seasonal flood frequency analysis for separating mixed-distribution floods generated by different atmospheric mechanisms [*GREHYS*, 1996; *Ouarda et al.*, 2001]. For instance, *Ouarda et al.* [2000] separated spring floods caused by snowmelt from summer/fall rainfall-generated floods in the province of Québec, Canada. In regional flood frequency analysis catchments are often grouped into regions according to the similarity in flood seasonality [*Ouarda et al.*, 1993; *Burn*, 1997; *Cunderlik and Burn*, 2002a]. Flood seasonality is also used for assessing hydrological homogeneity (or similarity) of a group of sites [*Cunderlik and Burn*, 2002c]. Seasonality of floods has important implications for the specification of flood-duration-frequency relationships [*Javelle et al.*, 2003]. Other applications include seasonal streamflow forecasting, watershed flood protection management, floodplain management, and reservoir operation.

[3] During the last decade, several studies emerged that were focused on various aspects of flood seasonality. *Hirschboeck* [1991] used information on hydro-climatic seasonality for the identification of mixed flood distributions. *Bayliss and Jones* [1993] described the seasonality of floods in Great Britain by measures based on directional statistics. *Ashkar et al.* [1993] and *Ouarda et al.* [1993] proposed a graphical procedure for determining flood seasons from peaks-over-threshold data. The procedure is based on plotting the mean annual number of exceedances against the time t for increasing threshold levels. Two different forms of the procedure were applied to gauging stations in the provinces of Québec and New Brunswick, Canada, and allowed partition of the two provinces into regions with similar flood seasons. *Magilligan and Graber* [1996] constructed contour maps of the mean day of flood occurrences and of a parameter that describes the variance of flood occurrences. The authors analyzed physiographic controls on flood timing using multiple regression models and suggested using the constructed contour maps for depicting the flood regime in New England. *Black and Werritty* [1997] identified geographical patterns of flood seasonality, using a database of events exceeding modest flood-flow thresholds at 156 gauging stations, and sought to explain them in terms of climatologic and catchment characteristics. *Krasovskaia* [1997] proposed a method for the hierarchical aggregation of monthly flow series into flow regime types by means of the minimization of an entropy-based objective function. *Burn* [1997] presented a regionalization approach that uses information related to the timing of flood events. The approach was based on the region of influence (ROI) pooling framework [*Burn*, 1990].

0043-1397/04/2003WR002295$09.00

[4] *Robson and Reed* [1999] described the effect of urbanization on flood seasonality. *Lecce* [2000] used cluster analysis on the data from 806 USGS gauging stations in the southeastern US to investigate spatial variations in the timing of the annual flood. *Whitfield and Cannon* [2000] proposed a polar plotting technique for seasonal hydrologic and climatic data. *Cunderlik and Burn* [2002a] derived and tested a detailed descriptor of flood seasonality. The authors explored the sensitivity of the seasonality descriptor to the record length and to the length of overlapping period. They have also shown that the information captured in flood seasonality is sufficient for the effective estimation of extreme flow quantiles from rural catchments in Great Britain. *Cunderlik and Burn* [2002b] explored the linkages between rain and flood seasonality and its applicability to regional flood frequency estimation at ungauged sites. *Archer* [2003] investigated broad characteristics of hydrological regimes in the upper Indus Basin using streamflow data from nineteen long-period stations in terms of annual and seasonal runoff.

[5] Generally, flood seasons are identified subjectively by visually assessing the temporal distribution of flood occurrences at the site of interest [see, e.g., *Ouarda et al.*, 1993; *Black and Werritty*, 1997; *Lecce*, 2000]. The plausibility of such locally identified seasons is then verified at a regional scale by comparing the at-site results with the results obtained from hydrologically similar neighboring sites. This approach is subjective and also highly time-consuming when applied to a large number of sites. Furthermore, if the seasons are not tested for significance, they may just be a product of sampling variability.

[6] The effect of sampling variability is indeed an important issue when assessing flood seasonality particularly from short records. Two types of errors can be made due to this effect. First, data from a short record can produce a pronounced seasonal flood distribution despite the true (but unknown) nonseasonal character of floods at the site of interest. Second, an apparently nonseasonal distribution of floods from a short sample may hide the true flood seasonality that is actually present at a given site.

[7] The main objective of this paper is to introduce a new, objective method for the identification of flood seasons. The proposed method tests the significance of seasons of high and low probability of flood occurrence by comparing the observed monthly variability of flood occurrences with the theoretical monthly flood variability in a nonseasonal model. The method also addresses the uncertainty resulting from sampling variability by quantifying the probability associated with the identified flood seasons. The performance of the proposed method is tested on many samples with different record lengths generated from several theoretical models of flood seasonality. The proposed approach is then applied to a large set of sites with a variety of flood regimes from different parts of Great Britain. The sites are then pooled into groups with similar temporal distributions of flood occurrences, representing the main flood seasonality types in Great Britain.

2. Method for the Identification of Flood Seasons

[8] Seasonality of flood occurrences can be tested by comparing the sampling variability of flood occurrences observed in a given record with the theoretical sampling variability of nonseasonal flood occurrences. A model of nonseasonal flood occurrences (floods with no seasonal preference) can be expressed by means of the circular uniform distribution as:

$$f(x) = P[X = x] = \frac{1}{360^\circ} \quad ; \quad 0^\circ \leq \mathrm{x} < 360^\circ \tag{1}$$

$$F(x) = P[X \leq x] = \frac{x}{360^\circ} \quad ; \quad 0^\circ \leq \mathrm{x} < 360^\circ \tag{2}$$

where x is the day of flood occurrence in degrees (by converting the 365 or 366 days of the year into 360 degrees), $f(x)$ is the probability density function and $F(x)$ is the cumulative density of flood occurrences. In this uniform, nonseasonal model a flood can occur, with the same probability, on any given day of the year. It is convenient to group dates of flood occurrence into months. Shorter intervals, such as one week, are inappropriate because the pattern of flood seasons diminishes even for long-record gauges, and therefore could not be discerned for gauges with typical length of records consisting of 20–30 flood observations [*Cunderlik and Burn*, 2002a]. Monthly temporal resolution of flood seasons is suitable for most practical applications. When the data are grouped into months, an adjustment must be applied when converting times into angles. According to *Mardia* [1972], the monthly frequencies can be adjusted so that they correspond to 360 days with all months having the same length. Then 1° will correspond to 1 day. The observed frequencies for 31-day months are multiplied by 30/31 and the frequency for February by 30/28 or by 30/29 respectively. The year is then reduced to 360 days but the sum (S) of the original frequencies f_i does not equal the sum (S') of the adjusted frequencies f_i'. To preserve the sum S, the final adjusted frequencies are obtained by multiplying f_i' by S/S'. The probability of the adjusted flood frequencies in a month, assuming the uniform distribution, will be then 1/12.

[9] Once the flood occurrences have been grouped into months, significant flood seasons can be identified by comparing the variability of the observed monthly probabilities of flood occurrence (monthly relative frequencies) with the theoretical sampling variability of nonseasonal monthly probabilities of flood occurrence generated from the uniform distribution. On one hand, if the observed monthly relative frequencies of flood occurrence from a sample with a record length N are within the confidence intervals derived for the monthly relative frequencies from records with the same record length N simulated from the uniform distribution, then there is no evidence of significant flood seasonality. On the other hand, if a record could not originate from the uniform distribution, then it must have a nonuniform (seasonal) distribution of floods that are clustered into one, two or more modes. Statistical tests for modality, such as the nonparametric smoothed-bootstrap test given by *Efron and Tibshirani* [1993] can be used for determining the number of modes presented in the data [*Cunderlik and Burn*, 2002b]. *Fisher* [1993] describes several other tests of randomness against nonuniform alternatives developed for directional data. However, such tests do not provide any information on the temporal occurrence and duration of flood-rich seasons (seasons of high proba-

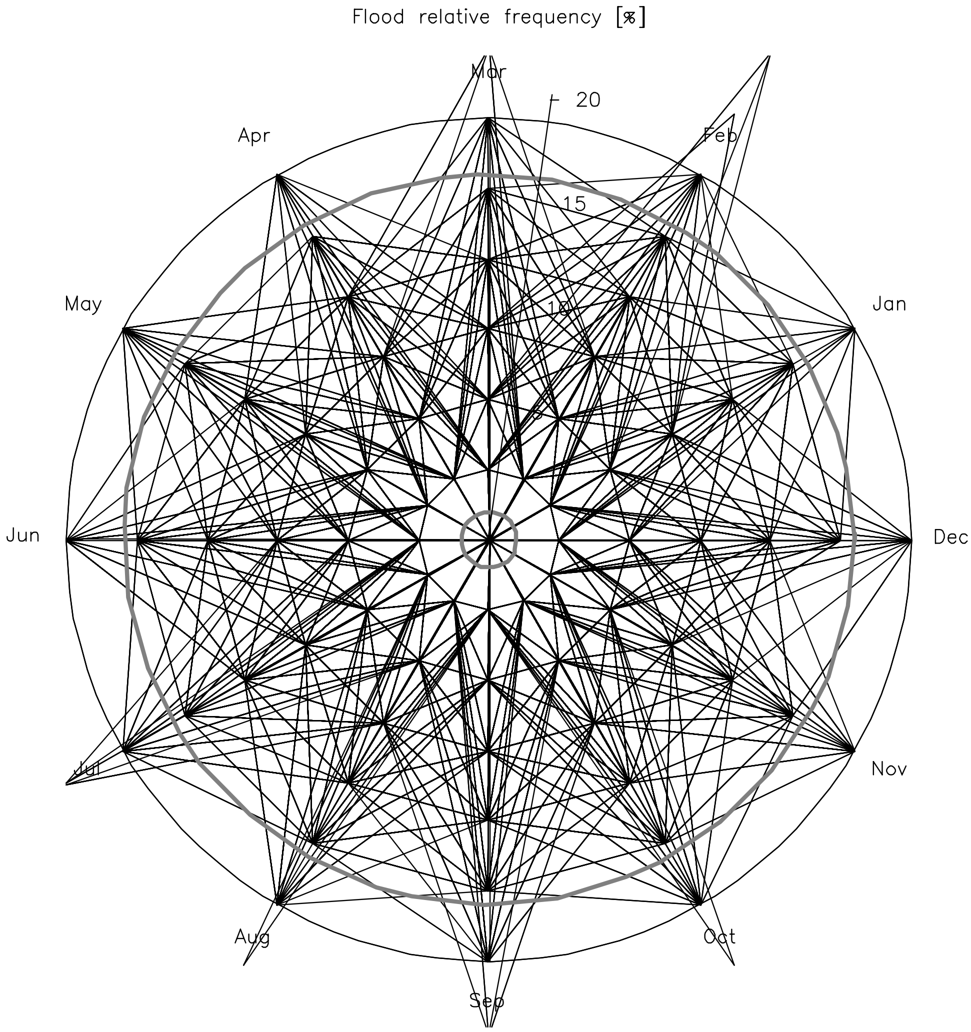

Figure 1. Sampling variability of relative flood frequencies calculated from 1000 samples with the record length of 30 observations generated from the uniform distribution. Thick shaded circles define the upper and lower one-sided confidence intervals.

bility of flood occurrence), and no information on the significance of flood-poor seasons (seasons of low probability of flood occurrence). The objective of the present work is therefore to develop a comprehensive technique for providing a complete picture of the distribution of flood seasons.

[10] The confidence intervals for the monthly relative frequencies generated from the circular uniform (nonseasonal) distribution of flood occurrences can be estimated according to the following procedure.

[11] 1. For a given record length of N observations, a large number, N_{Sim}, of simulated records with N flood occurrences are generated from the circular uniform distribution.

[12] 2. Generated dates of flood occurrence are grouped into months and monthly relative frequencies (probabilities of flood occurrence in a given month) adjusted according *Mardia* [1972] are calculated for each month. The sampling distribution of the adjusted relative frequencies is skewed, bounded ($0 \leq x \leq 100$), with the sampling mean $\bar{x} = 8.33\overline{3}$.

[13] 3. From the N_{Sim} sequences of monthly relative frequencies calculated from the simulated records of length N, one-sided (1-α)% confidence intervals are constructed as the αth (1-αth) empirical percentile intervals.

[14] 4. The (1-α)% one-sided confidence intervals are used for testing the hypothesis H_0, whether a given sequence of relative frequencies could originate from the uniform distribution at the α% significance level, against the alternative H_1 that the sequence originated from a nonuniform (seasonal) distribution of flood occurrences. The one-sided test is applied separately on relative frequencies above and below the uniform mean. If any monthly relative frequency is above the upper or below the lower confidence interval, then it is assumed at the significance level α that such sequence of relative frequencies could not originate from the uniform distribution. We suggest setting the significance level α to 5%. The α = 1% level was also considered but found too restrictive when applied on real data with more complex seasonality patterns.

[15] 5. The relative frequencies above and below the one-sided confidence intervals are considered to be significant flood-rich or flood-poor seasons.

[16] Figure 1 shows the sampling variability of relative frequencies obtained from 1000 samples with a record

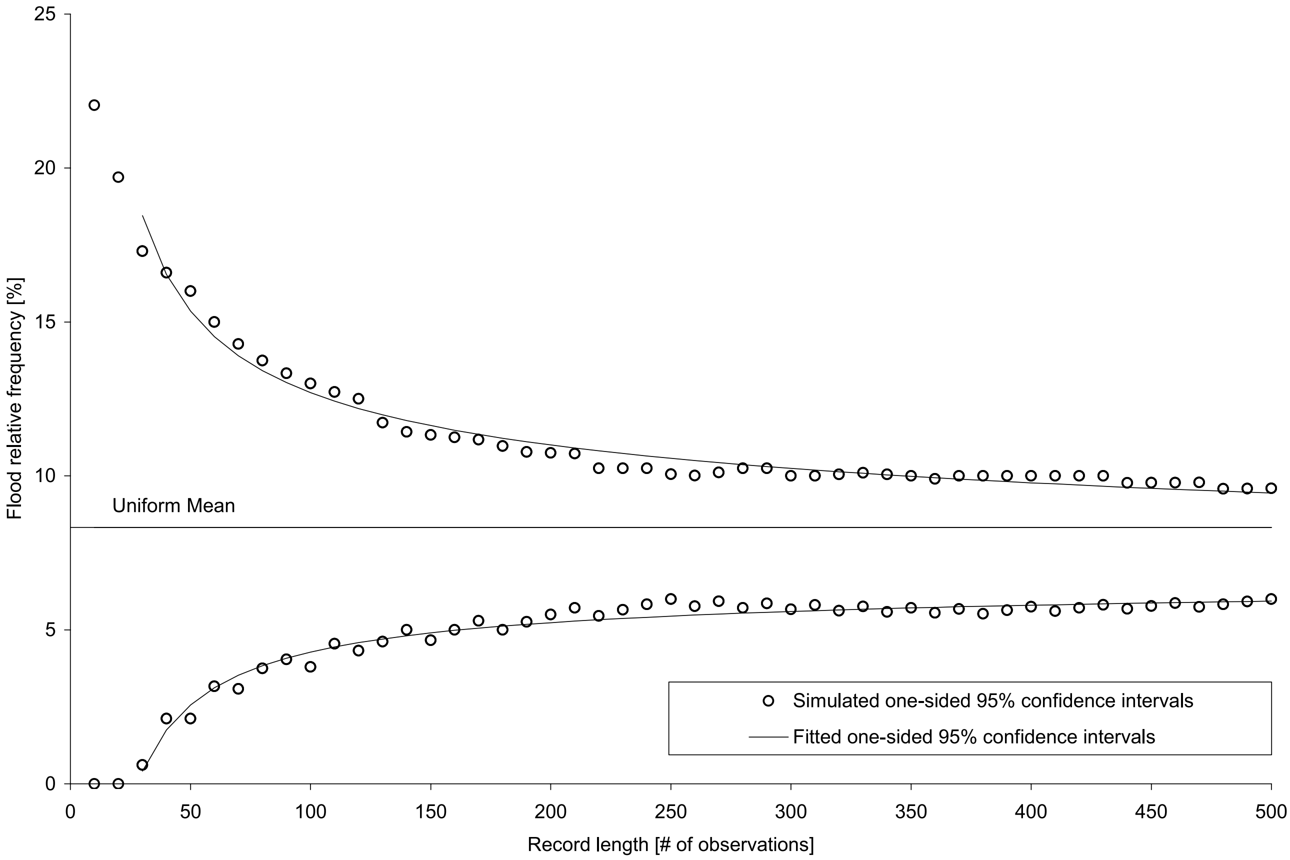

Figure 2. Simulated and fitted one-sided confidence intervals for monthly relative frequencies of flood occurrence as a function of record length N (based on 100,000 simulations from the uniform distribution).

length of 30 observations generated from the uniform distribution. Several peaks shown in Figure 1 are outside the upper and lower confidence intervals (shown by thick shaded circles) calculated for $N = 30$ observations. Figure 2 depicts the one-sided 95% upper and lower confidence intervals calculated from 100,000 simulated samples for record lengths ranging from 10 to 500 observations. The upper bound of 500 observations was set as a reasonable limit for long peaks-over-threshold samples. The confidence intervals for any record length from the interval $\langle 30, 500 \rangle$ can be approximated by

$$L_U^N = \frac{N + 11.491}{0.048\,N^{-1.131}} \quad ; \quad 30 \leq N \leq 500 \tag{3}$$

$$L_L^N = \frac{N - 27.832}{0.199\,N^{-0.964}} \quad ; \quad 30 \leq N \leq 500 \tag{4}$$

with $R^2 = 0.958$ for L_U^N and $R^2 = 0.960$ for L_L^N. Figure 2 suggests avoiding the assessment of flood seasonality from records with fewer than 30 observations because of large sampling variability. The oscillation of the data points about the fitted curves is caused by rounding the dates of flood occurrence to the nearest day.

[17] The one-sided confidence intervals defined above provide the bounds outside which all relative frequencies of flood occurrence could not, at the 5% significance level originate from the uniform, nonseasonal flood distribution model. By using the test outlined above, all significant nonuniform distributions of flood occurrences are classified as seasonal flood distributions. However, there may be cases of significant nonuniform distributions of flood occurrences, which do not necessary need to be "seasonal." Also, if the uncertainty resulting from the sampling variability of flood occurrences was ignored, two other errors could be possibly committed. The first error arises when the true temporal distribution of flood occurrences is seasonal, but because of a large sampling variability, no significant flood-rich or flood-poor seasons are found in a sample record generated from this distribution. The second error occurs when the flood seasonality from a record is found significant, but is actually not. Since the true distribution is not known, the sampling variability (uncertainty) associated with the probabilities obtained from the sample records must be taken into account and estimated.

[18] In order to address the uncertainty resulting from the sampling variability of flood occurrences a measure must be defined to estimate the probability of whether a given season is significant or nonsignificant. The observed data from a sample alone provides only a limited amount of information about the true temporal flood distribution. Therefore we suggest using a bootstrap resampling procedure to obtain a clearer picture about the sampling variability and the uncertainty associated with the estimates of flood frequencies. The idea is to generate N_{Bst} bootstrap samples from the available record and, then, for each sample, assess the significance of flood seasons using the method de-

scribed above. If more than $\alpha^*\%$ of the relative frequencies for a given season (month) are found significant, then it is assumed at the $\alpha^*\%$ probability that the sample originated from a parent distribution, which has a significant flood occurrence in that season. Thus it is possible to estimate the probability that the true (unknown) distribution really has a significant flood-rich or flood-poor season in any given month of the year.

[19] In order to classify the seasons of flood occurrence objectively we suggest defining "significant flood-rich seasons" comprising months having the seasonal probability $>\alpha^*\%$ with relative frequencies exceeding the upper confidence interval L_U^N for the uniform distribution. The category of "possibly significant flood-rich seasons" will then include months that do not exceed the upper confidence interval L_U^N, but more than $\alpha^*\%$ of the months obtained from the bootstrap samples did, and so there is a real chance that the true distribution actually has a significant flood-rich season in the corresponding month at the $\alpha^*\%$ significance level. In a similar fashion, months with the seasonal probability $>\alpha^*\%$ and relative frequencies below the lower confidence interval L_L^N for the nonseasonal distribution can be classified as "significant flood-poor seasons", and those inside the confidence intervals, but with more than $\alpha^*\%$ of the bootstrap months below the lower confidence interval L_L^N as "possibly significant flood-poor seasons" at the $\alpha^*\%$ significance level. We suggest setting α^* to 5 or 10%.

3. Performance Evaluation

[20] The performance of the proposed method for the identification of flood-rich and flood-poor seasons was assessed using a large number of records with different record lengths N simulated from several subjectively defined types of unimodal, bimodal and trimodal temporal distributions of flood occurrence. One scenario also involved samples generated from the uniform distribution. Distributions with four and more significant modes were not analyzed because such cases are very rare in the real world. Altogether, 33 different models of temporal flood distribution (1 uniform (nonseasonal model), 14 unimodal, 10 bimodal and 8 trimodal) were defined. They roughly cover the most basic flood regimes worldwide, including regimes with one main flood season (such as snowmelt or monsoon induced flood seasons), regimes with two flood seasons with equal or unequal probabilities of flood occurrence (such as dominant snowmelt induced season and secondary autumn rainfall induced flood season), or complex regimes with three equal or unequal flood seasons (such as spring snowmelt, summer storm, and autumn frontal induced flood seasons).

[21] For each model a record with $N = 10{,}000$ observations was constructed in such way that its empirical density function exactly corresponded to the subjectively predefined temporal flood distribution. The $f(x)$ functions were then expressed nonparametrically using the formula presented by *Cunderlik and Burn* [2002b]:

$$
\begin{aligned}
f(x,h) &= \frac{1}{nh}\sum_{i=1}^{n}\phi\left(\frac{360-|x-x_i|}{h}\right) \qquad ;\forall|x-x_i| > 180^\circ \\
f(x,h) &= \frac{1}{nh}\sum_{i=1}^{n}\phi\left(\frac{x-x_i}{h}\right) \qquad ;\forall|x-x_i| \le 180^\circ
\end{aligned}
\tag{5}
$$

where h is a bandwidth or smoothing factor, which determines the amount of smoothing that is applied to the data, and $\phi(x)$ is the standard normal density. Figure 3 shows the probability density functions $f(x)$ for five selected temporal distributions of flood occurrence. The first character in the notation used in Figure 3 describes the type of modality (e.g., 1 for unimodal), and the last character in the case of unimodal distributions describes the shape of the distribution (L for light-tailed and H for heavy-tailed), and in the case of multimodal distributions the equality of the modes (seasons) (E for modes with equal probability and U for unequal probability modes). The density functions in Figure 3 are for better clarity depicted in Cartesian coordinates. Since the proposed method is invariant to rotation around the origin, the modes of the distributions can occur in any other days as those shown in Figure 3. From the fitted nonparametric density functions (equation (5)) a large number of samples N_{Sim} with different record lengths N were generated, and the seasonal significance assessed according to the proposed method. The number of simulated samples N_{Sim} was set to 1000 and the number of bootstrap samples N_{Bst} to 500. Two types of evaluation criteria were chosen. The first, "strict" evaluation involved inclusion of only those simulated samples for which all seasons that were found significant exactly matched all significant seasons in the parent distribution. The second, "relaxed" criterion took the effect of sampling variability into account and thus also included samples for which the true significant seasons were identified only as possibly significant at the $\alpha^* = 5\%$ significance level. Again, the complete pattern (all significant seasonal and nonseasonal months) had to match in order to classify these samples as correctly identified.

[22] The results for the five main density functions showed in Figure 3 are summarized in Figure 4. The results obtained from the strict evaluation are depicted by "S", and those from the relaxed evaluation by "R". The results confirm the expectation that the performance of the proposed method strongly depends on both the distribution type of flood occurrences and on the sample record length. The best results were naturally obtained from samples generated from the uniform distribution (not shown), where 40 observations were sufficient to achieve the 100% strict performance of the method. Also, in the case of the unimodal light-tailed distribution 1L (Figure 3) the proposed method correctly identified significant flood seasons in all samples with record lengths of 60 observations or more. In the case of the heavy-tailed unimodal distribution 1H, the required record length increased to 150 observations. The 100% strict performance was achieved in bimodal samples generated from the distribution 2E, when the record length was around 300 observations compared to the 370 observations needed for the more complex 2U type with one main and one secondary flood season. In the samples generated from the trimodal distribution 3E, 410 and more observations were needed to achieve the 100% performance. By comparing the results obtained from the 1L and 1H, 2E and 2U distributions we can see how the performance of the method decreases as the complexity of the distributions increases. Rather long records are especially needed for obtaining reliable results from trimodal records; however, such complex distributions of flood

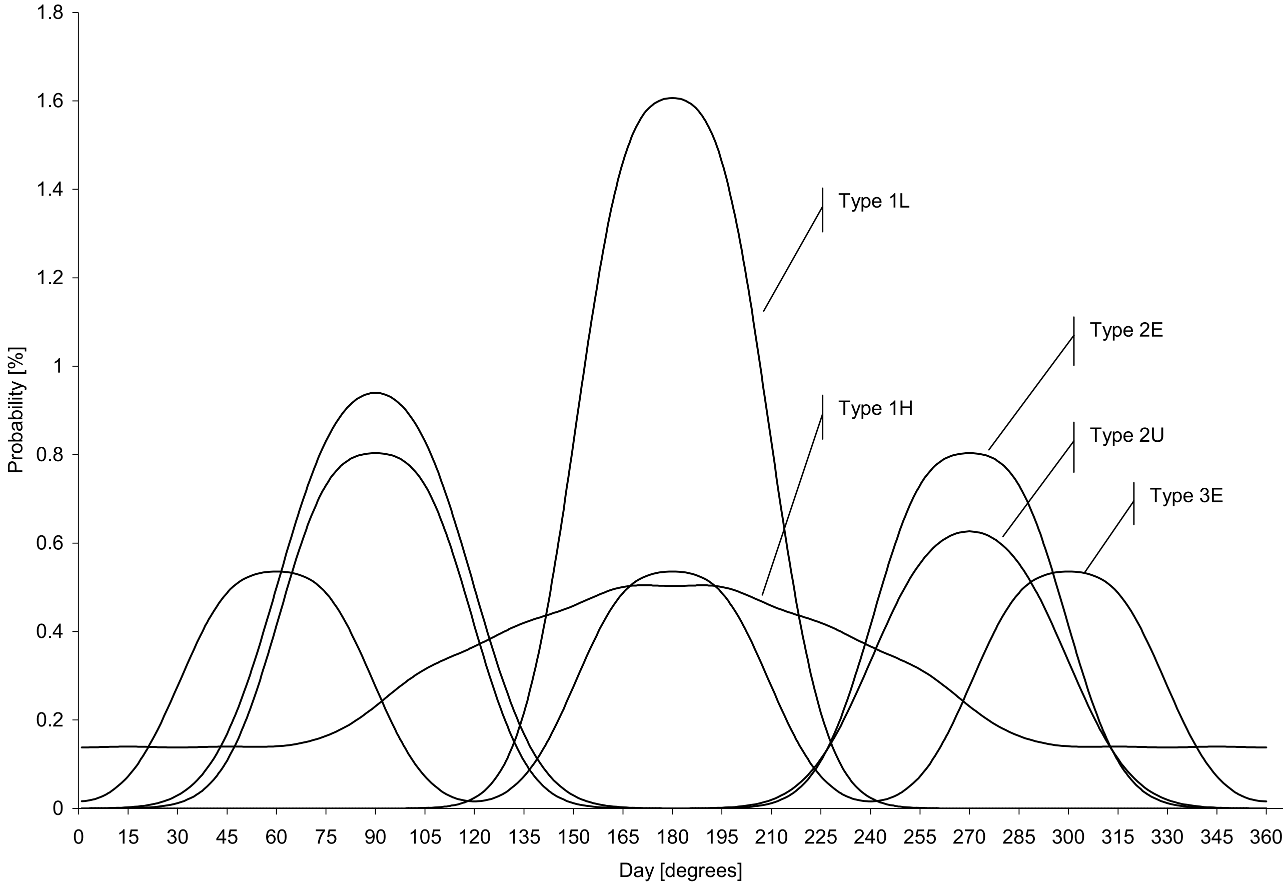

Figure 3. Probability density functions of five basic flood seasonality models (unimodal (types 1L and 1H), bimodal (types 2E and 2U), and trimodal (type 3E)).

occurrences are rather rare in the real world. Here the use of peaks-over-threshold samples might help to achieve better performance of the proposed method.

[23] If we also include the results where the true significant seasons were identified only as possibly significant (due to the effect of sampling variability), we notice a visible improvement in the performance of the method. Only about 40 observations are then needed to achieve the 100% relaxed performance in samples generated from the 1L distribution. In the case of the 1H distribution, 110 observations are needed for obtaining the 100% performance, 200 observations are needed for the bimodal 2E distribution, 260 observations for the bimodal 2U distribution, and 340 observations for the trimodal 3E distribution.

[24] The results obtained from the bimodal 2U distribution, which has one dominant and one secondary flood season (see Figure 3), show that for short record lengths, around 40 or more additional observations are needed for this distribution to achieve the performance obtained from the bimodal 2E distribution with two equal-probability seasons. The less defined secondary season is more difficult to capture, especially from samples with shorter record lengths. However, when the sampling uncertainty is taken into account, then the performance measure for the 2U type (2U-R) approaches the performance obtained from the 2E type based on the strict evaluation (2E-S in Figure 4).

[25] The effect of sampling variability can be also illustrated on the true parent distributions. The problem with short samples drawn from certain distribution types (particularly heavy-tailed and multimodal with main and secondary peaks) is that the frequency of flood occurrences is divided into more peaks, which may eventually fall within the confidence intervals for the uniform distribution. For example, less than 10 observations are needed to identify the single mode in type 1L as significant at the 5% significance level, but at least 55 observations are necessary for the single mode of the heavy-tailed type 1H (see Figure 3). Similarly, only 12 or more observations identify the two modes of the bimodal distribution type 2E as significant, but as the distribution tail becomes heavier, more data are needed to identify the modes (170 or more for the type 2E2 (not shown in Figure 3), whose probability is equal to half of the probability of the type 2E modes). The trimodal distribution of the type 3E in Figure 3 becomes significant when the record contains 35 or more observations, as opposed to the type 3E2 (not shown in Figure 3), whose probability is equal to half the probability of the type 3E modes, of which at least 230 observations are needed to identify these modes as significant. Again, if the bootstrap resampling is included in the testing, the performance of the method can be significantly improved (modes of the 2E2 distribution have a significant 16% probability of occurrence if generated from samples with 30 observations, and modes of the 3E2 distribution have a significant 12% probability of occurrence from the same samples with 30 observations).

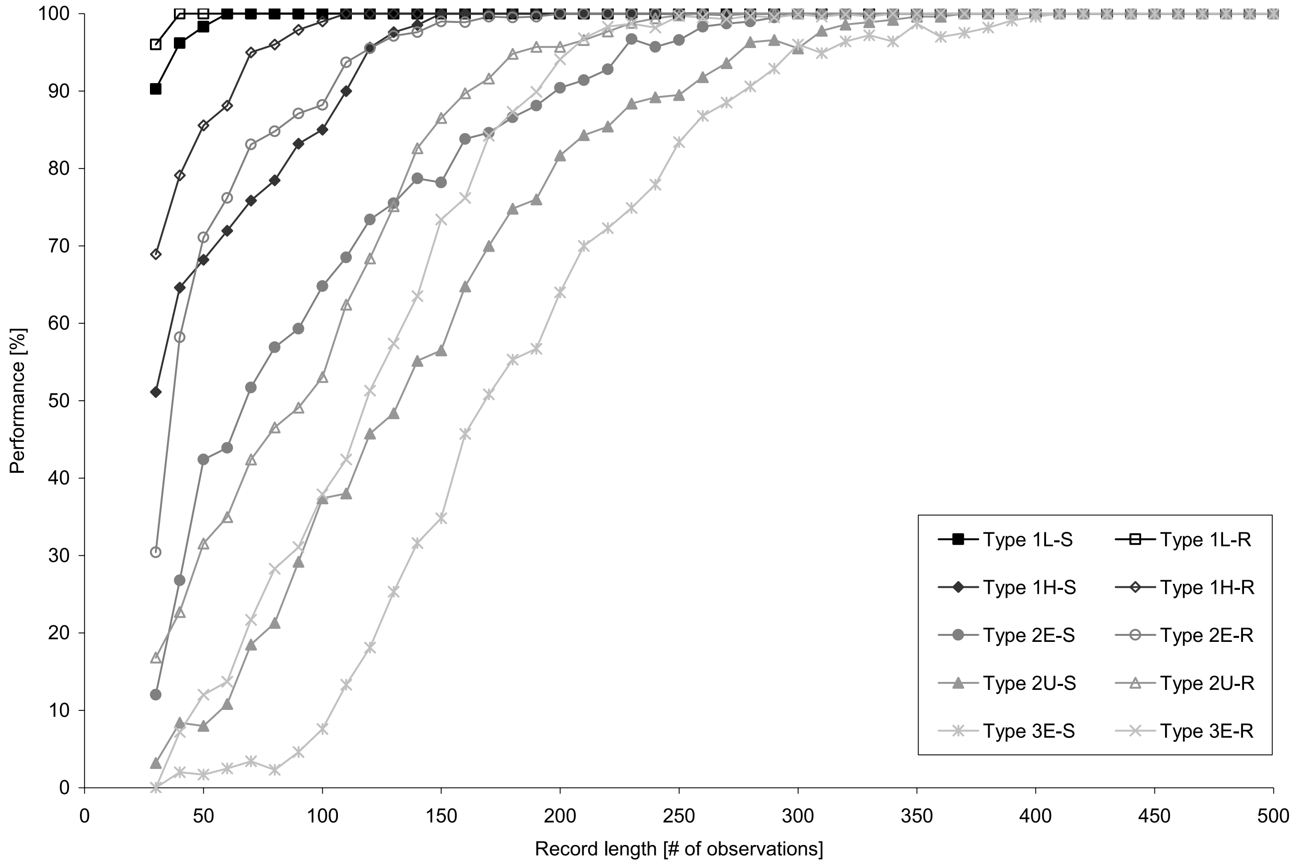

Figure 4. Relative numbers of correctly identified flood seasonality types for different flood distribution models and different sample record lengths (based on 1000 simulations). "S" marks strict evaluation results, and "R" marks relaxed evaluation results.

[26] The presented results indicate that for more complex distributions of flood occurrence, a rather large number of observations is needed in order to achieve acceptable performance of the proposed method. Annual maximum records with 100 or more observations are rather rare in most parts of the world; however, a 50-yearlong peaks-over-threshold (POT) record can easily provide 250 observations. Therefore, to achieve higher performance of the method and to obtain more reliable results, we strongly recommend using the POT data records.

4. Application

4.1. Study Area

[27] The proposed method was applied to a large number of POT records from Great Britain (GB). The POT records were chosen because they contain more information about flood seasonality than the annual maximum data. A comprehensive review of peaks-over-threshold modeling is presented by *Lang et al.* [1999]. The territory of GB was chosen because of its variable flood seasonality, and because of the high density of its network of gauging stations. A number of previously published studies [e.g., *Bayliss and Jones*, 1993; *Black and Werritty*, 1997; *Robson and Reed*, 1999; *Cunderlik and Burn*, 2002b] describing flood regimes in GB could be also used for the comparison of results.

[28] The POT data were obtained from the Flood Estimation Handbook (FEH) flood peak data CD-ROM [*Robson and Reed*, 1999]. The abstraction threshold was chosen in the FEH to yield an average of four or more flood peaks per year. More information about the POT extraction is given by *Bayliss and Jones* [1993] and in the FEH [*Robson and Reed*, 1999]. All catchments included in the database had to be essentially rural with minimal flood attenuation by reservoirs and lakes. From the subset of catchments that met these criteria a common 20-yearlong high-data density observation period from 1966 to 1985 was identified. Sites with at least 60 POT peaks during this period were then included in the database. There were 268 sites that fulfilled this criterion. In order to have the same record lengths for all 268 included sites, only the highest 60 peaks from each site were included in the database. The main reason to restrict the database to a common observation period was to eliminate the effect of climate variability and/or trends in the timing of floods resulting from different observation periods on the results [see, e.g., *Arnell and Reynard*, 1996; *Prudhomme et al.*, 2003]. The same number of observations for each analyzed record was chosen to eliminate the effect of different record lengths on the results, in order to apply the proposed method on every record with the same power. Both issues are considered important when results from individual gauges will be intercompared and used for the identification of main flood seasonality types in the study area.

4.2. Results

[29] To illustrate the effect of sampling variability on flood seasonality estimated from short samples, relative

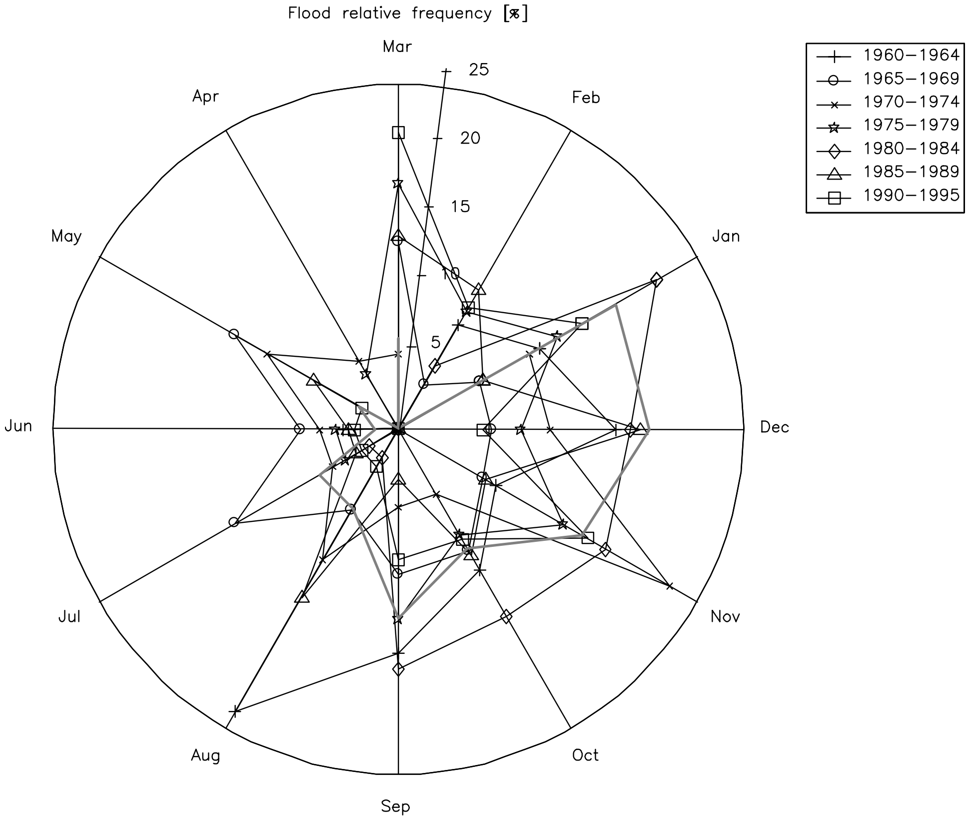

Figure 5. Relative frequencies of flood occurrence from Findhorn at Forres (ID 7002) from different 5-yearlong nonoverlapping observation periods. Thick shaded line represents relative frequencies from the common observation period 1966–1985.

frequencies of flood occurrence calculated from different 5-yearlong nonoverlapping observation periods (with a minimum of 20 POT peaks for each period) for the Findhorn station at Forres (ID 7002) were plotted in Figure 5. The thick shaded line represents the relative frequencies from the common 1966–1985 period. We can see that the individual temporal flood distributions change significantly from period to period. Also, the difference in relative frequencies of flood occurrence can reach 20% in some months. Figure 5 clearly demonstrates the importance of long records for reducing the effect of sampling variability, and the importance of a common observation period for all evaluated sites. The seasonality of floods is sensitive to the POT threshold of interest. The POT floods used in Figure 5 were abstracted above a threshold defined in FEH for the whole observation period. If individual thresholds defined for each 5-yearlong period were used instead (daily records were not available), then the sensitivity of flood seasonality to a given period of record would likely be even more pronounced than that seen in Figure 5.

[30] The percentage numbers of monthly relative frequencies found significant in the 500 bootstrap samples generated from the Findhorn at Forres POT record from the common period 1966–1985 are given in Table 1. Table 1 reports that among the 500 bootstrap samples, all February and April relative frequencies and 75% of all June frequencies were classified as significant flood-poor seasons. On the other hand 73% of December frequencies and 71% of January frequencies were found to be significant flood-rich seasons. The number of significant November relative frequencies were much lower, only 42%, but this result still highly exceeds the $\alpha^* = 5\%$ limit, thus being considered highly significant. The significant flood-rich and flood-poor seasons, identified by the proposed method in Findhorn at Forres from the common observation period, are also depicted in Figure 6. The method identified one unimodal significant flood-rich season from November to January and three discontinuous significant flood-poor seasons in February, April, and June. Because of the effect of sampling variability we can also expect potentially significant flood-rich season in September and a potentially significant flood-poor seasons in March, May and July, thus having a pronounced, contiguous February to July flood-poor season. The months of August and October are nonsignificant in terms of flood seasonality. The small inset in the upper part of Figure 6 gives a schematized view of the flood regime at this site, according to the significance of flood seasons.

[31] The proposed method was then applied in a similar fashion to the remaining 267 sites across Great Britain. The number of bootstrap samples N_{Bst} was again set to 500 and the significance level α^* to 5%. At each site the method identified at least one significant flood-rich and one significant flood-poor season. The majority of analyzed sites (78%) have unimodal distribution with only one significant flood-rich season. A total of 59 sites (22%) have significant bimodal structure, and most of them are concentrated along the west coast, particularly in Wales. Our criterion

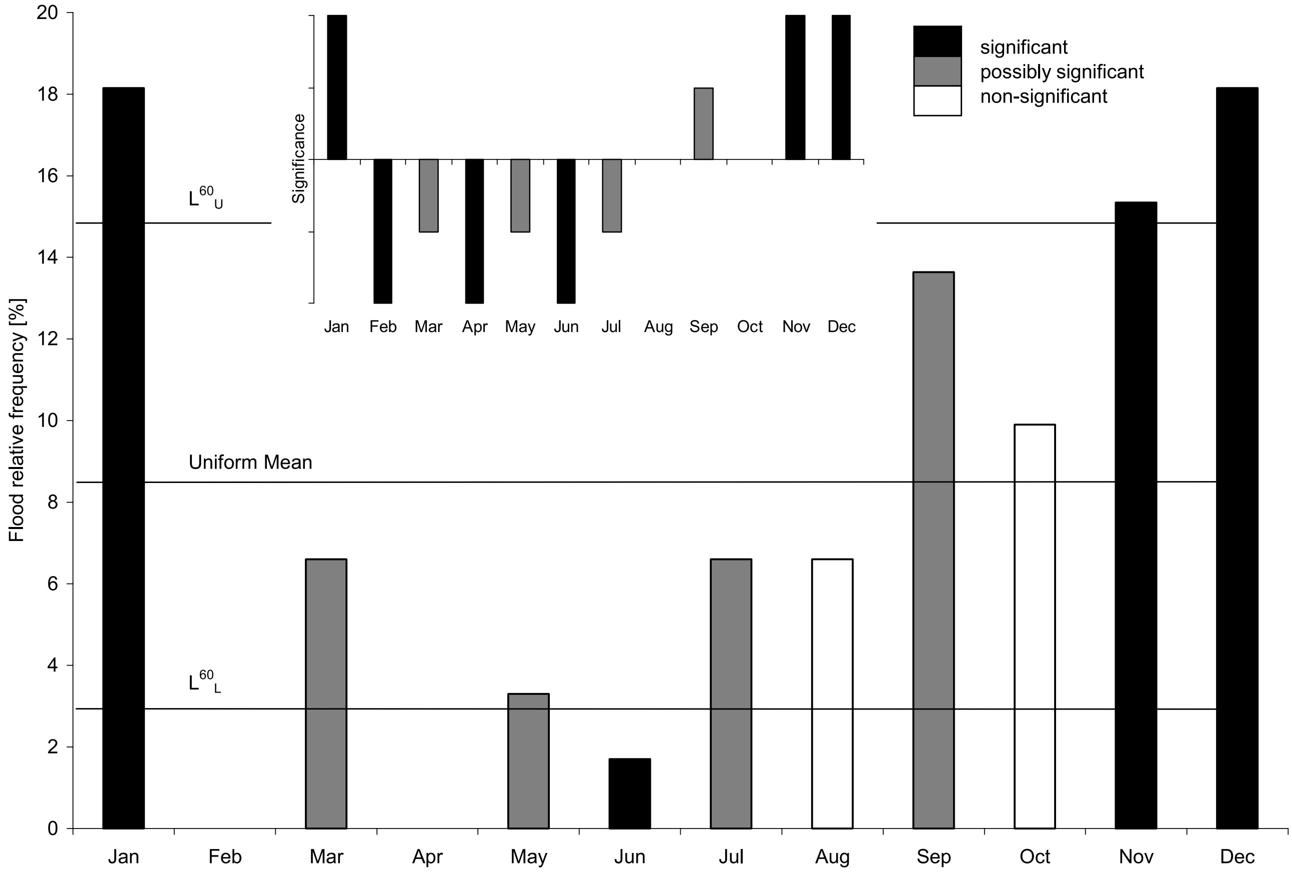

Figure 6. Identified flood seasons from Findhorn at Forres (ID 7002) from the common observation period 1966–1985. Horizontal lines depict the upper and lower one-sided confidence intervals. The smaller figure represents a schematized flood regime at this site according to the significance of monthly relative frequencies of flood occurrence.

for multimodality was that individual seasons in a multimodal distribution must be separated at least by one nonsignificant or possibly significant month. Almost half of all analyzed sites (47.0%) have two months that can be classified as a significant flood-rich season, 40.3% of the sites have three significant months, 7.5% only one month and 5.2% four significant flood-rich months. The results for significant flood-poor seasons revealed that 182 sites (67.9%) have unimodal distribution with only one significant flood-poor season and 86 sites (32.1%) have two significant flood-poor seasons. There is no visible geographical spatial pattern when the sites are plotted according to modality of flood-poor seasons. Significant flood-poor seasons are generally longer than flood-rich seasons. A total of 35.1% of all sites have 4 months that can be classified as significant flood-poor seasons, 21.6% have 3 months, 20.9% have 5 months, 11.2% have 2 months, 5.6% have 1 month, 4.5% have 6 months, and only 1.1% have 7 months of flood-poor season.

[32] The flood-rich and flood-poor seasons objectively identified at the 268 individual sites can be used for the identification of groups of sites with similar flood seasonality and for exploring any spatial patterns these groups may possess. A cluster analysis was applied for grouping the sites according to the similarity in their flood seasonality. Rather than using actual relative frequencies of flood occurrence (highly uncertain due to sampling variability), a schematized representation of flood seasonality was used. At each site the temporal flood distribution was described by a set of twelve monthly values: zeros represented nonsignificant months, the values of 0.5 (−0.5) months with possibly significant flood-rich (flood-poor) seasons and the values of 1 (−1) months with significant flood-rich

Table 1. Percentage Numbers of Significant Monthly Relative Frequencies Derived From 500 Bootstrap Samples Generated From the Findhorn at Forres POT Record (ID 7002) From the Common Observation Period 1966–1985[a]

Season	January	February	March	April	May	June	July	August	September	October	November	December
Flood-rich	**71.37**	0.00	0.62	0.00	0.00	0.00	0.44	0.11	**28.29**	3.79	**42.04**	**72.88**
Flood-poor	0.00	**100.00**	**6.99**	**100.00**	**36.93**	**75.03**	**7.02**	4.22	0.33	0.19	0.07	0.00

[a]Bold entries depict results significant at the α^* = 5% significance level.

(flood-poor) seasons. The twelve values describing the site's flood seasonality were used as the clustering attributes in a K-means cluster analysis. The K-means clustering technique is a nonhierarchical method that partitions the observations into K mutually exclusive clusters. The method defines groups by reallocating objects among clusters with the objective to minimize the within-group variance and to maximize the between-group variance. The optimal number of clusters was determined using the silhouette plot. The silhouette plot displays a measure of how close each point in one cluster is to points in the neighboring clusters [*The MathWorks*, 2000]. This measure ranges from +1, indicating points that are very distant from neighboring clusters, through 0, indicating points that can be equally assigned to one cluster or another, to −1, indicating points that are probably assigned to the wrong cluster. A quantitative way to compare different cluster outputs is to look at the average silhouette values for different numbers of clusters. In the case of our data the best result was obtained for three clusters that lead to well defined, separated, silhouette peaks with no negative values (not shown). The average silhouette value was 0.37.

[33] Figure 7 shows the identified three types of flood seasonality on a map of Great Britain. The types have a well-defined spatial pattern. The type 1 sites are dominant mainly in Scotland, northwest England and north Wales. The type 3 sites have prevailing occurrence in central England and eastern coast of GB. Finally, the type 2 sites are spatially concentrated in southern England, Wales, and make a spatial transition between the types 1 and 3 elsewhere. This spatial pattern is in a good correspondence with results published in other studies [*Bayliss and Jones*, 1993; *Black and Werritty*, 1997; *Robson and Reed*, 1999]. Figure 8 explains the three flood seasonality types in terms of the average relative frequencies of flood occurrence. The small inset in the upper part of Figure 8 summarizes the delineated types in terms of the average significance of flood seasons: +1 (−1) means all 268 sites have a significant flood-rich (flood-poor) season in the particular month. In the type 1, floods occur predominantly in the autumn, with the average maximum in November. The occurrence of floods at the end of the winter and spring seasons is rather low in this type. The type 1 also includes most of the 59 sites that were found bimodal in the previous step of the analysis. This can be seen from the inset in Figure 8, although the bimodal structure of the average significance index was partially smoothed-out by the averaging process. The type 3 includes sites where floods occur mainly during the winter and spring (with the average maximum in January) and with a low occurrence of flooding in the summer and autumn seasons. The type 2 represents again a transition between the types 1 and 3.

[34] Since the uncertainty related to the actual values of relative frequencies is high, it is better to adopt a schematized description of flood regime like the one shown in the upper part of Figure 8 (or for one site as shown in the upper part of Figure 6).

5. Conclusions

[35] In this paper we have proposed a new, objective approach to the identification of significant flood-rich and flood-poor seasons. The proposed method tests the significance of periods with high and low probability of flood occurrence by comparing the observed variability of flood occurrences with the theoretical variability in the uniform, nonseasonal model. The method also takes into account the uncertainty resulting from the sampling variability by quantifying the significance of the identified flood seasons. The performance of the method was tested on a large number of samples generated from different predefined distributions of flood seasonality assuming various sample record lengths. The results indicated that the performance of the method strongly depends on both the record length and the underlying flood distribution. The performance improves considerably when the sampling variability is taken into account. The sample record length plays a crucial role in flood seasonality studies. We recommend using POT data for assessing flood seasonality since they provide a greater amount of seasonal information than the annual maximum data records. Records with less than 30 observations should be used only in rare cases and their results evaluated with extreme care. Also, instead of using actual relative frequencies of flood occurrence for describing flood seasonality we suggest using a schematized alternative based on the significance of flood seasons.

[36] The proposed approach was applied on a large set of 268 gauging sites with flood regimes representing different parts of Great Britain. At every site the method identified at least one significant flood-rich and one significant flood-poor season. Using the K-means clustering algorithm, the set of sites was then grouped into three main types of flood seasonality in Great Britain. The delineated types are

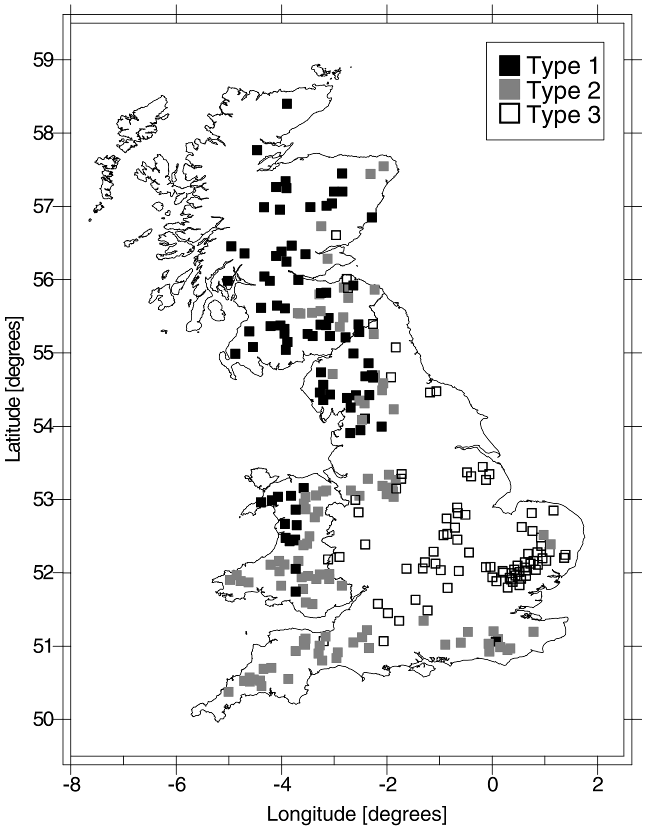

Figure 7. Identified main flood seasonality types in Great Britain.

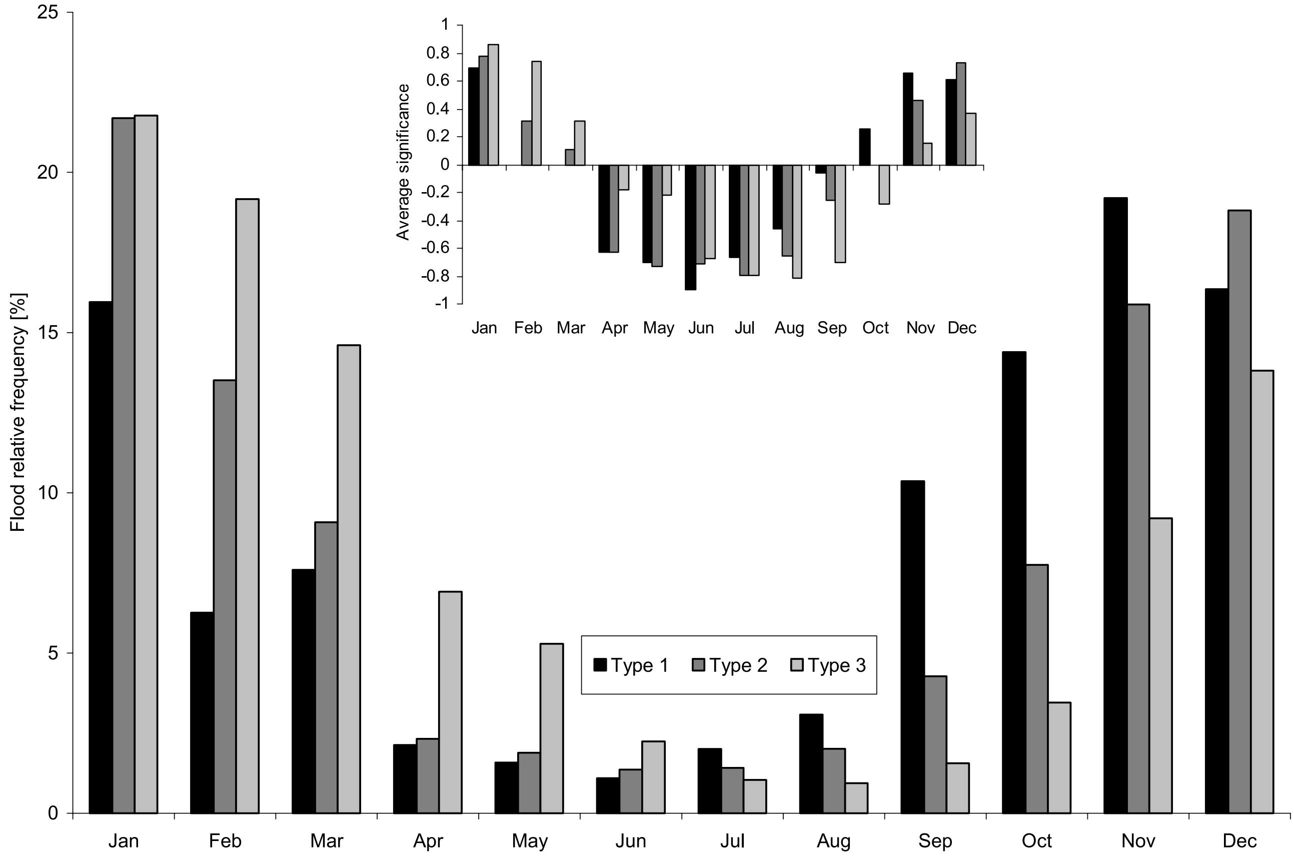

Figure 8. Average relative frequencies of flood occurrence and average significance of flood-rich and flood-poor seasons for the three identified flood seasonality types in Great Britain.

spatially well separated leading to three natural regions of flood regime. The regions are in good correspondence with flood regions identified in other studies. The results show that the proposed method can efficiently identify flood seasons from both theoretical and observed distributions of flood occurrence. The technique can be a useful tool for the investigation of climate and land use change effects on flood regime.

[37] **Acknowledgments.** The authors would like to thank two anonymous reviewers for their valuable comments, which significantly improved the quality of this article.

References

Archer, D. (2003), Contrasting hydrological regimes in the upper Indus Basin, *J. Hydrol.*, *274*, 198–210.

Arnell, N. W., and N. S. Reynard (1996), The effect of climate change due to global warming on river flows in Great Britain, *J. Hydrol.*, *183*, 397–424.

Ashkar, F., N. El-Jabi, and T. B. M. J. Ouarda (1993), Etude des variations saisonnières des crues par le modèle de dépassement, *Rev. Sci. Eau*, *6*, 131–152.

Bayliss, A. C., and R. C. Jones (1993), Peaks-over-threshold flood database: Summary statistics and seasonality, *Rep. 121*, 61 pp., Inst. of Hydrol., Wallingford, U. K.

Black, A. R., and A. Werritty (1997), Seasonality of flooding: A case study of north Britain, *J. Hydrol.*, *195*, 1–25.

Burn, D. H. (1990), Evaluation of regional flood frequency analysis with a region of influence approach, *Water Resour. Res.*, *26*(10), 2257–2265.

Burn, D. H. (1997), Catchment similarity for regional flood frequency analysis using seasonality measures, *J. Hydrol.*, *202*, 212–230.

Cunderlik, J. M., and D. H. Burn (2002a), The use of flood regime information in regional flood frequency analysis, *Hydrol. Sci. J.*, *47*(1), 77–92.

Cunderlik, J. M., and D. H. Burn (2002b), Analysis of the linkage between rain and flood regime and its application to regional flood frequency estimation, *J. Hydrol.*, *261*, 115–131.

Cunderlik, J. M., and D. H. Burn (2002c), Local and regional trends in monthly maximum flows in southern British Columbia, *Can. Water Resour. J.*, *27*(2), 191–212.

Efron, B., and R. J. Tibshirani (1993), *An Introduction to the Bootstrap*, Chapman and Hall, New York.

Fisher, N. I. (1993), *Statistical Analysis of Circular Data*, Cambridge Univ. Press, New York.

GREHYS (1996), Intercomparison of flood frequency procedures for Canadian rivers, *J. Hydrol.*, *186*, 85–103.

Hirschboeck, K. K. (1991), Flood hydroclimatology, in *Flood Geomorphology*, edited by V. R. Baker, R. C. Kochel, and P. C. Patton, pp. 27–49, John Wiley, Hoboken, N. J.

Javelle, P., T. B. M. J. Ouarda, and B. Bobée (2003), Spring flood analysis using the flood-duration-frequency approach: Application to the provinces of Quebec and Ontario, Canada, *Hydrol. Processes*, *17*(18), 3717–3736.

Krasovskaia, I. (1997), Entropy-based grouping of river flow regimes, *J. Hydrol.*, *202*, 173–191.

Lang, M., T. B. M. J. Ouarda, and B. Bobée (1999), Towards operational guidelines for over-threshold modeling, *J. Hydrol.*, *225*, 103–117.

Lecce, S. A. (2000), Seasonality of flooding in North Carolina, *Southeast. Geogr.*, *41*(2), 168–175.

Magilligan, F. J., and B. E. Graber (1996), Hydroclimatological and geomorphic controls on the timing and spatial variability of floods in New England, USA, *J. Hydrol.*, *178*, 159–180.

Mardia, K. V. (1972), *Statistics of Directional Data*, Academic, San Diego, Calif.

Ouarda, T. B. M. J., F. Ashkar, and N. El-Jabi (1993), Peaks over threshold model for seasonal flood variations, in *Engineering Hydrology: Proceed-*

ings of the Symposium, edited by C. Y. Kuo, pp. 341–346, Am. Soc. of Civ. Eng., Reston, Va.

Ouarda, T. B. M. J., M. Haché, P. Bruneau, and B. Bobée (2000), Regional flood peak and volume estimation in northern Canadian basins, *J. Cold Regions Eng.*, *14*(4), 176–191.

Ouarda, T. B. M. J., C. Girard, G. S. Cavadias, and B. Bobée (2001), Regional flood frequency estimation with canonical correlation analysis, *J. Hydrol.*, *254*, 157–173.

Prudhomme, C., D. Jakob, and C. Svensson (2003), Uncertainty and climate change impact on the flood regime of small UK catchments, *J. Hydrol.*, *277*, 1–23.

Robson, A., and D. Reed (1999), *Flood Estimation Handbook*, vol. 3, *Statistical Procedures for Flood Frequency Estimation*, Inst. of Hydrol., Wallingford, U. K.

The MathWorks (2000), Statistics toolbox for use with MATLAB, Natick, Mass.

Whitfield, P. H., and A. J. Cannon (2000), Polar plotting of seasonal hydrologic and climatic data, *Northwest Sci.*, *74*(1), 76–80.

B. Bobée, J. M. Cunderlik, and T. B. M. J. Ouarda, Institut National de la Recherche Scientifique, Eau, Terre, and Environnement, University of Québec, 2800 Einstein, C.P. 7500, Sainte-Foy, Québec, Canada G1V 4C7. (juraj_cunderlik@inrs-ete.uquebec.ca)