Switching the pooling similarity distances: Mahalanobis for Euclidean

Juraj M. Cunderlik[1] and Donald H. Burn[2]

Received 12 May 2005; revised 6 October 2005; accepted 11 November 2005; published 14 March 2006.

[1] In recent years, catchment similarity measures based on flood seasonality have become popular alternatives for identifying hydrologically homogeneous pooling groups used in regional flood frequency analysis. Generally, flood seasonality pooling measures are less prone to errors and are more robust than measures based on flood magnitude data. However, they are also subject to estimation uncertainty resulting from sampling variability. Because of sampling variability, catchment similarity in flood seasonality can significantly deviate from the true similarity. Therefore sampling variability should be directly incorporated in the pooling algorithm to decrease the level of pooling uncertainty. This paper develops a new pooling approach that takes into consideration the sampling variability of flood seasonality measures used as pooling variables. A nonparametric resampling technique is used to estimate the sampling variability for the target site, as well as for every site that is a potential member of the pooling group for the target site. The variability is quantified by Mahalanobis distance ellipses. The similarity between the target site and the potential site is then assessed by finding the minimum confidence interval at which their Mahalanobis ellipses intersect. The confidence intervals can be related to regional homogeneity, which allows the target degree of regional homogeneity to be set in advance. The approach is applied to a large set of catchments from Great Britain, and its performance is compared with the performance of a previously used pooling technique based on the Euclidean distance. The results demonstrate that the proposed approach outperforms the previously used approach in terms of the overall homogeneity of delineated pooling groups in the study area.

Citation: Cunderlik, J. M., and D. H. Burn (2006), Switching the pooling similarity distances: Mahalanobis for Euclidean, *Water Resour. Res.*, *42*, W03409, doi:10.1029/2005WR004245.

1. Introduction

[2] Regional flood frequency analysis (RFFA) is a common method used for the estimation of design floods at ungauged locations and at gauged locations with data insufficient for a reliable estimation of flood quantiles for the required return period. RFFA usually combines a geographic technique that transfers hydrologic information from similar gauged locations to an ungauged site, with a statistical technique that assesses the probability of flood events at the ungauged site.

[3] Regional regression models that relate flood quantiles to catchment physiographic and climatic characteristics were the first RFFA models. In 1960, T. Dalrymple developed the so-called index flood method, which has become one of the most frequently used RFFA techniques. Its principles were reflected in most RFFA approaches that followed, such as the American Bulletin 17B [*U.S. Water Resources Council*, 1981], the iterative search technique [*Wiltshire*, 1985], the fractional membership method [*Acreman and Sinclair*, 1986], the region-of-influence method [*Burn*, 1990], and the method of hierarchical regions [*Gabriele and Arnell*, 1991].

[4] A geographic regionalization, leading to spatially contiguous regions, was initially used for the selection of catchments for RFFA. *Acreman and Wiltshire* [1989] were among the first who suggested using geographically discontinuous catchments. This idea was further developed by *Burn* [1990] into the region-of-influence (ROI) approach. Geographically discontinuous catchments are called a pooling group, and the process of forming the groups is known as pooling [*Reed et al.*, 1999]. The concept of ROI laid out the foundations for site-focused pooling. In site-focused pooling, a pooling group is specifically tailored to a target site of interest and a given return period [*Cunderlik and Burn*, 2002b].

[5] Regions and pooling groups are usually delineated according to similarity of catchments in physiographic and climatic characteristics [*Mosley*, 1981; *Wiltshire*, 1985; *Acreman and Sinclair*, 1986; *Ando*, 1990; *Pearson*, 1991; *Institute of Hydrology*, 1999; *Ouarda et al.*, 2001]. In recent years, similarity measures based on flood seasonality have become popular alternatives for identifying hydrologically homogeneous pooling groups [*Ouarda et al.*, 1993; *Black and Werritty*, 1997; *Burn*, 1997; *Lecce*, 2000; *Castellarin et al.*, 2001; *Cunderlik and Burn*, 2002a, 2002b; *Ouarda et al.*, 2001, *Cunderlik et al.*, 2004a]. The main advantage of this approach is that flood date data used for describing flood

[1]Conestoga-Rovers and Associates, Waterloo, Ontario, Canada.
[2]Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada.

0043-1397/06/2005WR004245

seasonality are practically error-free and are more robust than flood magnitude data. Also, in flood seasonality based RFFA the flood magnitude data are used only for testing regional homogeneity and not directly in the pooling process. Finally, the above mentioned studies demonstrated that pooling based on flood seasonality tends to include catchments with similar physiographic properties and leads to homogeneous groups and effective quantile estimations.

[6] Several methods for describing flood seasonality have emerged during the last decade. *Bayliss and Jones* [1993], *Magilligan and Graber* [1996], *Burn* [1997], and others defined homogeneous pooling groups by means of the mean day of flood (MDF) (directional mean) and the flood occurrence variability measure $\bar{r}$ (complement of the directional variance). *Ouarda et al.* [1993] and *Ashkar et al.* [1993] derived a graphical technique based on plotting the mean annual number of exceedances against time, for each station and for increasing base levels, for determining flood seasonality from peaks-over-threshold data. *Krasovskaia* [1997] introduced a method for the hierarchical aggregation of monthly streamflow series into seasonality types by means of minimization of an entropy-based objective function. *Black and Werritty* [1997], *Lecce* [2000], and *Cunderlik and Burn* [2002b] described flood seasonality by means of the relative frequencies of flood occurrence. *Castellarin et al.* [2001] combined flood and rainfall seasonality into a compound measure that describes the mean delay between the two phenomena. *Cunderlik and Burn* [2002a] explored the sensitivity of a detailed descriptor of flood seasonality to the record length and to the length of the overlapping period.

[7] All of the above studies identified homogeneous pooling groups based on catchment similarity in flood seasonality but did not consider estimation uncertainty due to sampling variability. *Cunderlik et al.* [2004a, 2004b] demonstrated the importance of addressing sampling variability in the estimation of flood seasonality. The authors showed that the effect of sampling variability can produce flood seasonality significantly different from the "true" (population) flood seasonality. The objective of this paper is therefore to develop a new site-focused pooling method that takes into consideration the sampling variability in the measures describing flood seasonality. A nonparametric resampling technique is used to estimate the sampling variability for the target site, as well as for every site that is a potential member of the pooling group for the target site. The variability is quantified by the Mahalanobis distance [*Mahalanobis*, 1930, 1936]. In contrast to previously used pooling similarities based on the Euclidean distance, the Mahalanobis distance takes into account not only the average value, but also the variance and covariance of the pooling variables. The new pooling method not only ranks catchments according to similarity with the focused (target) site, but it also quantifies how much two catchments differ in their probability distribution of flood seasonality. This adds a new dimension to site-focused pooling because the target degree of regional homogeneity in flood seasonality can be set in advance.

2. Description of Flood Seasonality

[8] The seasonality of floods can be described by means of directional statistics [*Fisher*, 1993]. This method is based on defining individual dates of flood occurrence as a directional variable and then calculating the directional mean and variance. The Julian day of flood occurrence (Day_i) can be converted to an angular value (θ_i) using

$$\theta_i = Day_i \frac{2\pi}{N_D}; \qquad 0 \le \theta_i \le 2\pi, \tag{1}$$

where N_D is the number of days in a year. In this sense, a date of flood occurrence represents a vector with unit magnitude and a direction given by θ_i. The directional mean $\bar{\theta}$, or the mean day of flood MDF, is calculated as the addition of unit vectors

$$\bar{\theta} = \tan^{-1}\left(\frac{\bar{y}}{\bar{x}}\right); \quad \bar{x} \ne 0; \quad \text{MDF} = \bar{\theta}\frac{N_D}{2\pi}, \tag{2}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} \cos(\theta_i); \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} \sin(\theta_i), \tag{3}$$

where n is the number of samples for a given site. The mean direction ($\bar{\theta}$) represents a directional location measure of a sample consisting of dates of flood occurrence.

[9] A convenient measure of dispersion (variability) of the individual dates of flood occurrence around the mean value can be defined as [*Bayliss and Jones*, 1993]

$$\bar{r} = \sqrt{\bar{x}^2 + \bar{y}^2}; \quad 0 \le \bar{r} \le 1. \tag{4}$$

The variable $\bar{r}$ represents a dimensionless dispersion measure. A value close to 1 indicates that all events in the sample are tightly grouped about the mean direction, while a value closer to zero indicates that there is greater variability in the occurrence of flood events. The complement $(1 - \bar{r})$ is the directional variance.

[10] The directional mean and the variance measure defined in polar coordinates by $\bar{\theta}$ and $\bar{r}$ can be represented in Cartesian coordinates by the average coordinates $\bar{x}$ and $\bar{y}$. The coordinates describe flood seasonality at a given site and can be used to measure similarity (distance) in flood seasonality between two or more sites. The similarity has been previously measured by the Euclidean distance

$$Dis_{i,j} = \sqrt{\left(\bar{x}_i - \bar{x}_j\right)^2 + \left(\bar{y}_i - \bar{y}_j\right)^2}, \tag{5}$$

where $\bar{x}_i$ and $\bar{y}_i$ ($\bar{x}_j$ and $\bar{y}_j$) are the coordinates defined in equation (3) for site i and j, respectively. As follows from equation (5), the Euclidean distance only accounts for the average values and cannot be used to describe the patterns of site-specific sampling variability of correlated variables.

3. Sampling Variability in Flood Seasonality

[11] The effect of sampling variability can be effectively modeled by various parametric and nonparametric resampling methods. In the nonparametric bootstrap resampling, synthetic records are generated by randomly drawing observations from the observed record. Let us consider a record of length n observed at a site i. The pair of average

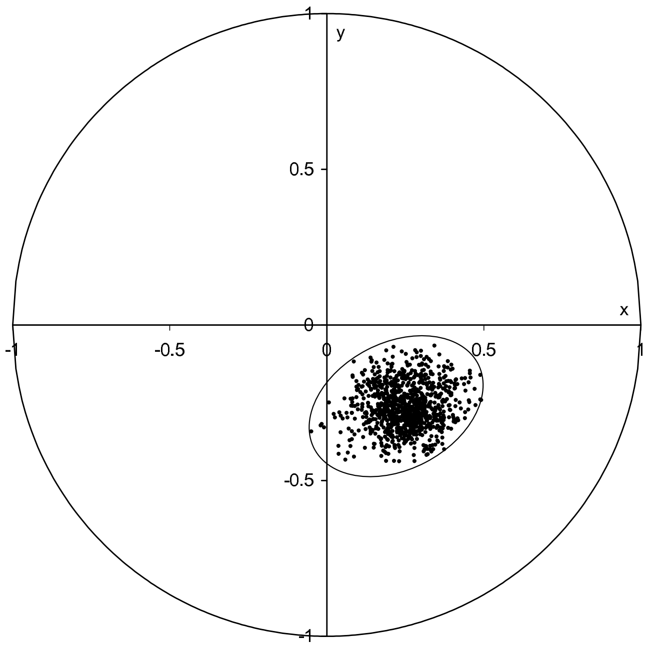

Figure 1. Sampling variability of flood seasonality measures $\bar{x}$ and $\bar{y}$ determined from 60 peaks-over-threshold events from 1966 to 1985 at Findhorn at Forres (ID 7002).

coordinates $\bar{x}_i$ and $\bar{y}_i$ calculated from the observed record according to equation (3) describes flood seasonality at the site in terms of both the mean ($\bar{\theta}$) and variance ($\bar{r}$) of dates of flood occurrence. The sampling variability (uncertainty) of flood seasonality described by $\bar{x}_i$ and $\bar{y}_i$ can be assessed by bootstrapping the observed record N_B-times, generating N_B n-sample long synthetic records, each with flood seasonality characterized by $\bar{x}^B_{i,k}$ and $\bar{y}^B_{i,k}$, $k = 1, \ldots N_B$. The capability of characterizing the sampling variability using this approach is affected by the length of the available sample record, as subsequently discussed.

[12] Figure 1 shows an example of a cloud of $N_B = 1000$ pairs of $\bar{x}^B_{i,k}$ and $\bar{y}^B_{i,k}$ values generated from a 20-year-long record from Findhorn at Forres (ID 7002) in Great Britain. The shape of the cloud changes only minimally when $N_B \geq 1000$. By plotting the bootstrapped values $\bar{x}^B_{i,k}$ and $\bar{y}^B_{i,k}$ around the original flood seasonality measures $\bar{x}_i$ and $\bar{y}_i$, it can be seen that the sampling variability in all directions is not equal. Thus the Euclidean distance is not an appropriate tool for measuring distances between points defined by correlated variables such as $\bar{x}^B_{i,k}$ and $\bar{y}^B_{i,k}$ values in Figure 1.

[13] The Mahalanobis measure is a convenient alternative to adequately account for the correlation between pooling variables. The Mahalanobis distance is based on both the mean and variance of the predictor variables, as well as the covariance matrix of all the variables. The region of constant Mahalanobis distance around the mean forms an ellipse in two-dimensional (2-D) space, or an ellipsoid or hyperellipsoid when more variables are used. If we have a vector $\vec{u}$ containing the values $\bar{x}_i$ and $\bar{y}_i$ for site i and a vector $\vec{v}$ containing the values $\bar{x}_j$ and $\bar{y}_j$ for site j, then the Mahalanobis distance is calculated as

$$D = \sqrt{(\vec{u} - \vec{v})^T W^{-1} (\vec{u} - \vec{v})}, \tag{6}$$

where W is the covariance matrix of vectors $\vec{u}$ and $\vec{v}$. If the variables are uncorrelated and of equal variance, the Mahalanobis distance is identical to the Euclidean distance. Figure 1 shows the Mahalanobis distance ellipse fit to the resampled data from Findhorn at Forres (ID 7002).

4. New Flood Seasonality Pooling Method

[14] Denote flood seasonality measures $\bar{x}_i$ and $\bar{y}_i$ at the target site i, and flood seasonality measures $\bar{x}_j$ and $\bar{y}_j$ at a site j that is a potential member of the pooling group for the target site. Resampling the data record for the target site can be used to estimate the covariance matrix given in equation (6), which is then used to define the Mahalanobis distance measure. However, in addition to there being uncertainty in the seasonality measures for the target site, there is also uncertainty in the seasonality measures for each site that is a potential member of the target pooling group. Therefore the uncertainty in all seasonality measure estimates should be incorporated in the pooling method.

[15] Incorporating the uncertainty in flood seasonality measures for all sites can be accomplished by resampling not only the record from the target site, but also records from all sites that are potential members of the target pooling group. In this way, each site will have a unique "coordinate system" defined by the Mahalanobis distance. Since the distance for each site to the target site is found using a different site-specific similarity measure, it is not straightforward to define a similarity measure that can be used to rank the closeness of each site to the target site. The solution proposed is to define the similarity based on the minimum confidence interval at which the Mahalanobis ellipse for a site and the ellipse for the target site intersect. A lower confidence interval implies greater similarity. The confidence intervals can be calculated by the following procedure:

[16] 1. Resample N_B pairs of the seasonality measures $\bar{x}^B_{i,k}$ and $\bar{y}^B_{i,k}$ from the data record.

[17] 2. Calculate the covariance matrix of vectors $\vec{u}$ and $\vec{v}$ for the N_B values of $\bar{x}^B_{i,k}$ and $\bar{y}^B_{i,k}$ and use this to define a site-specific coordinate system for the Mahalanobis distance.

[18] 3. Determine the Mahalanobis distance D between each resampled pair $\bar{x}^B_{i,k}$ and $\bar{y}^B_{i,k}$ and the original values $\bar{x}_i$ and $\bar{y}_i$ and rank these values.

[19] 4. Define the (1-a) $\times$ 100% confidence interval as the ellipse that passes through the point that has the xth largest Mahalanobis distance (for example, if a = 0.05 and $N_B = 1000$, then $x = 50$).

[20] An example of a theoretical confidence ellipse is given in Figure 2. The ellipse is described by its center Z_0, semimajor axis a and semiminor axis b, and the orientation of the ellipse α. The parameters of a confidence ellipse with the Mahalanobis distance smaller or equal to the distance determined by the confidence interval can be estimated via the method of least squares. First, the center of the ellipse is defined by the Cartesian coordinates of the seasonality measure estimated from the observed record ($Z_0[\bar{x}, \bar{y}]$), and the orientation α is estimated as the average slope of linear regressions $\bar{x}_i = f(\bar{y}_i)$ and $\bar{y}_i = f(\bar{x}_i)$, where $f(\ )$ denotes a functional relationship. In the next step, the ellipse is

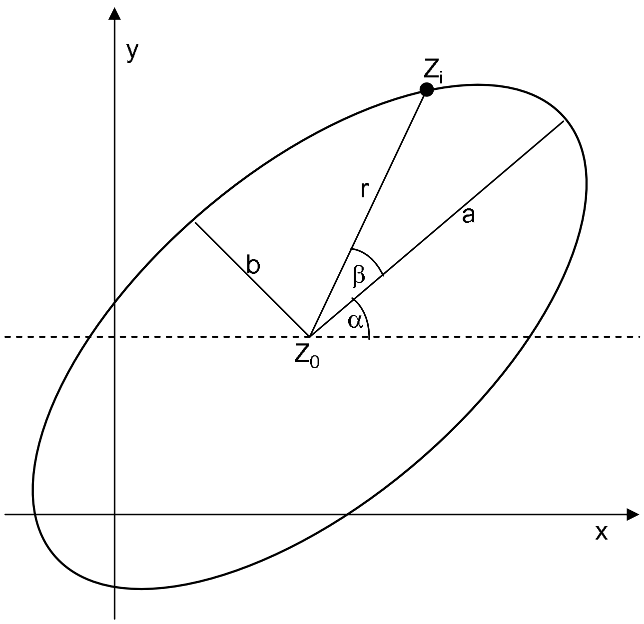

Figure 2. Definition of the Mahalanobis confidence ellipse. The ellipse is defined by center Z_0, semimajor axis a, semiminor axis b, and its orientation α.

transformed so that it is centered at the origin with axis coincident with the Cartesian axis ($Z'_0[0, 0]$ and $\alpha' = 0°$). The transformation is accomplished by utilizing the equation of ellipse in polar coordinates

$$\begin{aligned} x' &= r\cos\beta \\ y' &= r\sin\beta, \end{aligned} \tag{7}$$

where r is the distance of a point Z_i from the ellipse center Z_0 (see Figure 2), defined as

$$r^2 = \frac{b^2a^2}{b^2\cos^2\beta + a^2\sin^2\beta} \tag{8}$$

and β is the angle between the line $Z_0 - Z_i$ and the semimajor axis. The transformed ellipse can then be described in Cartesian coordinates by a simple form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{9}$$

where a and b are the semimajor and semiminor ellipse axis in the x and y directions, respectively. Such an ellipse can be fitted to a set of points $\{\bar{x}_i, \bar{y}_i\}$, $i = 1, 2, \ldots, N_B$ by the method of least squares. If we define

$$f_{a,b}(x,y) = 1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}, \tag{10}$$

then the method of least squares finds the values of a and b that minimize the function

$$\chi^2 = \sum\{f_{a,b}(x_i, y_i)\}^2 \tag{11}$$

by setting the partial derivates of χ^2 with respect to a and b to zero. *Sussman* [1999] showed that the a and b that minimize χ^2 solve the system of two equations

$$\begin{aligned} \left(\sum x_i^2\right)a^2b^2 - \left(\sum x_i^2y_i^2\right)a^2 - \left(\sum x_i^4\right)b^2 &= 0 \\ \left(\sum y_i^2\right)a^2b^2 - \left(\sum x_i^2y_i^2\right)b^2 - \left(\sum y_i^4\right)a^2 &= 0. \end{aligned} \tag{12}$$

The system in equation (12) has the solution

$$\begin{aligned} a^2 &= \frac{\left(\sum x_i^2y_i^2\right)^2 - \left(\sum x_i^4\right)\left(\sum y_i^4\right)}{\left(\sum x_i^2y_i^2\right)\left(\sum y_i^2\right) - \left(\sum x_i^2\right)\left(\sum y_i^4\right)} \\ b^2 &= \frac{\left(\sum x_i^2y_i^2\right)^2 - \left(\sum x_i^4\right)\left(\sum y_i^4\right)}{\left(\sum x_i^2y_i^2\right)\left(\sum x_i^2\right) - \left(\sum x_i^4\right)\left(\sum y_i^2\right)}. \end{aligned} \tag{13}$$

Other methods for ellipse fitting are given by *Cui et al.* [1996], *Lei and Wong* [1999], and *Fitzgibbon et al.* [1999].

[21] By means of the outlined methodology, confidence ellipses are fitted to the target site and to all sites that are potential members of the pooling group of the target site. The similarity between the target site i and the potential site j is then assessed by finding the minimum confidence interval at which their ellipses intersect. If the two ellipses intersect at the confidence interval of 0%, then the target site is, in terms of estimated flood seasonality, identical with the site j. If the confidence interval is 99%, it means that there is a 1% probability that the sample flood seasonality of both sites originated from the same population. If two ellipses do not intersect, they have different flood seasonality distributions and hence should not be grouped in the same pooling group.

[22] The minimum confidence intervals can also be understood in terms of regional homogeneity. This adds a new dimension to site-focused pooling because the target degree of regional homogeneity can be set in advance. The lower is the confidence interval, the higher is the similarity of the two sites in terms of flood seasonality, and the higher is the overall homogeneity of the pooling group, in terms of flood seasonality.

[23] The confidence interval at which two ellipses intersect can be found iteratively. The confidence interval is increased by a step chosen according to the required accuracy. A step of 0.1% should be adequate for most practical applications as it provides high accuracy and is still computationally effective. A tolerance e is also required to account for the size of the iteration step and the number of bootstrap simulations. The intersecting interval is found if

$$D_{C_i} \le e, \tag{14}$$

where D_{C_i} is the minimum distance between the ellipses for the confidence interval C_i. The tolerance should be larger for larger iteration steps and smaller for larger number of simulations. The tolerance term e assures that

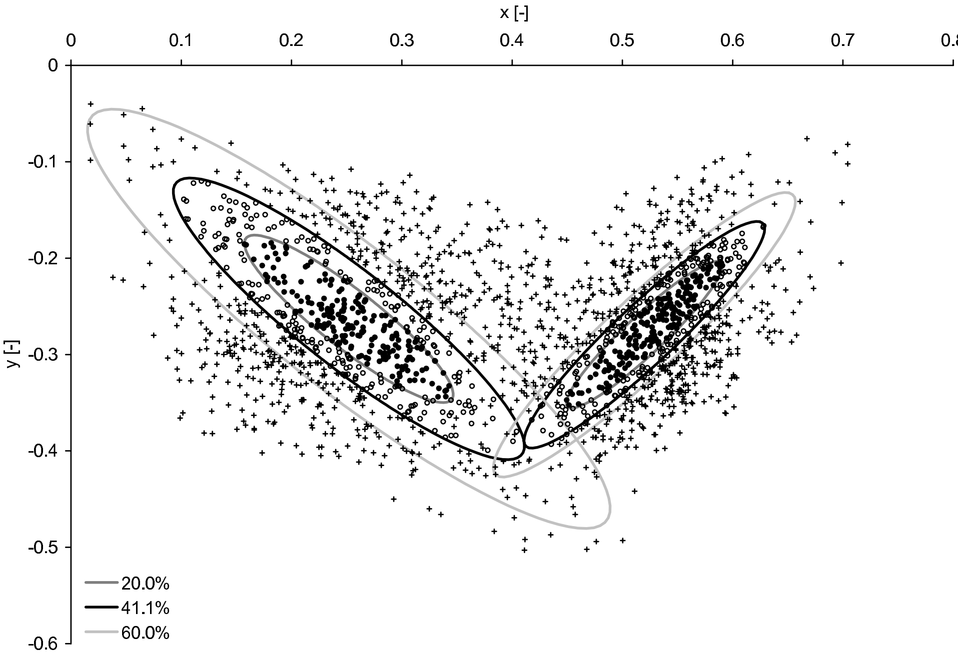

Figure 3. Demonstration of the new flood seasonality pooling method. Different symbols represent points with the Mahalanobis distance enclosed by 20.0% and 41.1% confidence intervals and outside of 41.1% confidence interval.

the intersection of two ellipses is always found for a chosen step. Figure 3 describes the iterative algorithm using data from Great Britain. The two ellipses intersect at the 41.1% confidence interval. Figure 4 illustrates the minimum distance between the two Mahalanobis ellipses plotted in Figure 3, expressed in the Euclidean metric, as a function of the confidence interval. This relation reflects the bivariate pattern of sampling variability in the flood seasonality of the two sites shown in Figure 3.

[24] In the site-focused pooling, the minimum confidence intervals at which the confidence ellipse of the target site intersects with the ellipses of the potential sites are ranked and the sites with the smallest minimum confidence intervals are included in the pooling group for the target site. Different criteria, such as the 5T rule [*Institute of Hydrology*, 1999] described later in the text, can be applied to set the upper limit to the size of the pooling group. In this method the size can also be defined by the confidence interval (e.g., including only sites with the intersecting confidence interval ≤50%).

[25] The proposed pooling approach can be applied to the regional flood frequency estimation at ungauged sites when the flood seasonality descriptors for the ungauged sites are derived indirectly from catchment climatic char-

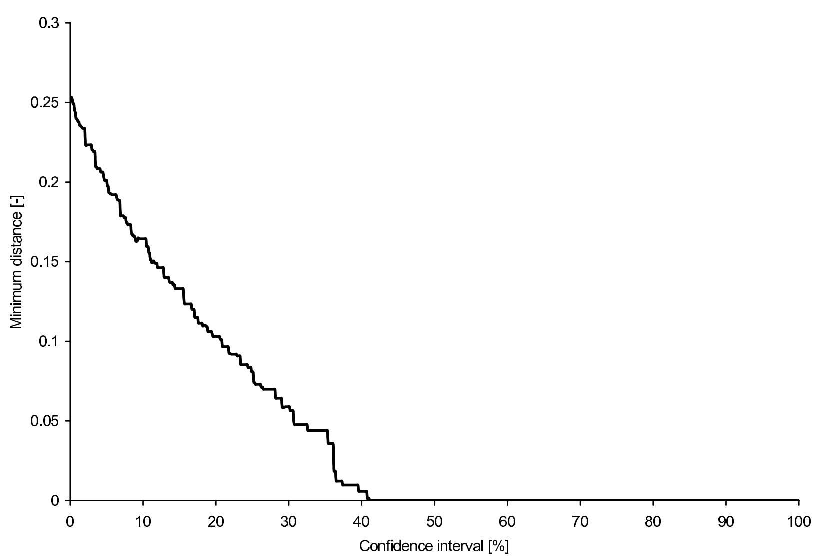

Figure 4. Minimum distances between two Mahalanobis ellipses as a function of the confidence interval.

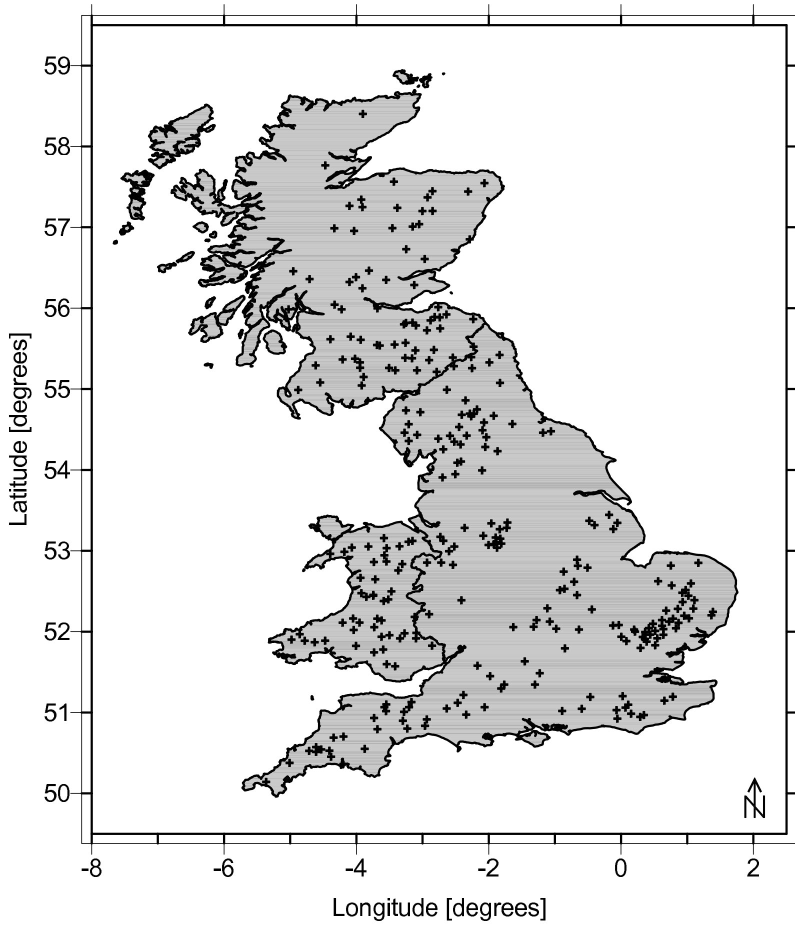

Figure 5. The 294 hydrometric sites selected for the study.

acteristics or by means of geostatistical mapping, as presented by *Cunderlik and Burn* [2002b].

5. Performance Evaluation

[26] The performance of the proposed pooling approach depends on several factors including the sample record lengths and the degree of overlap, the parent distributions of flood seasonality, and the degree of regional heterogeneity of the pooling group. The nonparametric resampling of very short sample records may produce misleading clouds of synthetic mean flood occurrences. Similar clouds and Mahalanobis ellipses can be then generated even for rather different parent flood seasonality distributions. It should be remembered that the bootstrapping provides only an approximated representation of the real parent distributions. If the observed records do not overlap, they may reflect different low-frequency climatic periods that affect the variability in flood seasonality. Parameters of complex flood seasonality distributions (e.g., multimodal) are much harder to capture than the parameters of simple (unimodal) distributions. Finally, the approach may perform differently in pooling groups highly heterogeneous in terms of flood seasonality than in groups with similar distribution of flood occurrences. Future research should explore the impact of record length, and other factors, on the calculation of the similarity measures used in this research.

[27] *Cunderlik et al.* [2004b] performed a detailed performance study focused on the determination of flood seasonality from hydrologic records. The study evaluated different sampling models depending on the estimation method, flood seasonality type, and sample record length. The authors recommended a minimum of 30 annual maximum flood observations to obtain reliable estimates of flood seasonality. Most of these findings apply also to the approach proposed in this paper, and therefore the performance analysis is not repeated here.

[28] The first performance analysis of the proposed pooling approach considered in this study involves comparing the homogeneity of delineated pooling groups with the homogeneity of groups identified by the previous pooling approach based on the Euclidean distance similarity. Both pooling approaches are applied to a large set of catchments. Catchments are added to a pooling group until the number of station-years exceeds $5T$, where T is the focused return period (i.e., for $T = 100$ years the pooling process is terminated when the number of station-years is greater than 500). The confidence interval is not applied as a criterion for terminating the pooling process so that the performance evaluation will reflect only the definition of catchment similarity. The homogeneity of pooling groups is then evaluated by means of the Hosking and Wallis regional homogeneity test [*Hosking and Wallis*, 1997]. The test is based on a comparison of the regional variance of L-moment ratios in a pooling group with the

Table 1. Homogeneity of Pooling Groups Delineated by the ED and MD Methods for Target Return Periods T = 50, 100, and 200 Years

	ED Pooling			MD Pooling		
T, years	H1	H2	H3	H1	H2	H3
			T = *50*			
Average	4.268	1.611	0.921	4.029	1.585	0.870
Standard Deviation	1.461	1.270	1.312	1.250	1.040	1.267
Maximum	7.444	4.481	3.546	7.451	3.712	3.403
Minimum	1.422	−0.372	−1.781	0.752	−0.408	−1.762
			T = *100*			
Average	5.977	2.432	1.475	5.765	2.161	1.212
Standard Deviation	1.525	1.332	1.458	1.487	1.222	1.309
Maximum	9.482	5.949	4.985	9.630	4.906	4.231
Minimum	3.004	−0.792	−1.415	3.020	−0.977	−0.996
			T = *200*			
Average	8.326	3.523	2.240	8.194	2.988	1.752
Standard Deviation	1.601	1.250	1.264	1.566	0.971	1.116
Maximum	12.054	5.806	4.503	11.889	5.444	4.390
Minimum	5.822	1.575	−0.004	4.703	0.026	−1.524

average variance of the L-moment ratios given from N simulations of groups influenced only by sampling variability. Depending on which L-moment ratios are used, three test statistics are defined: H1 if L-CV is used, H2 if L-CV and L-CS are used, and H3 if L-CS and L-CK are used. The region should be regarded as "acceptably homogeneous" if the test statistic H < 1, "possibly heterogeneous" if 1 ≤ H ≤ 2, and "definitely heterogeneous" if H > 2 [*Hosking and Wallis*, 1997]. All three test statistics, H1 (based on L-CV), H2 (based on L-CV and L-CS), and H3 (based on L-CS and L-CK), are evaluated here, and the pooling group is regarded as sufficiently homogeneous when H ≤ 2. In addition to the number of homogeneous groups identified by each pooling approach, the average values, standard deviation, and maximum and minimum of H are also recorded for different scenarios of target return period.

[29] The second performance analysis considered in this study involves comparing the two pooling approaches in terms of regional flood quantile estimation using a jackknife procedure [see, e.g., *Efron and Tibshirani*, 1993]. In this procedure, one site at a time is regarded as ungaged. The remaining sites are treated as gaged and are used for building a pooling group for the ungaged site. The data from the pooling group are then used to estimate the T-year flood (T = 50, 100, and 200 years) at the ungaged site according to the index-flood procedure [*Dalrymple*, 1960]. The regional estimate of the T-year flood is compared to the "true" T-year flood at the ungaged site, calculated from at-site data available at the site. This procedure is carried out for the proposed Mahalanobis and the Euclidean pooling approaches separately. The performance of the pooling approaches is then assessed by the BIAS and RMSE measures defined as

$$\mathrm{BIAS} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{Q}_T^i - Q_T^i}{Q_T^i}\right) \times 100\ [\%] \qquad (15)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{Q}_T^i - Q_T^i}{Q_T^i}\right)^2} \times 100\ [\%], \qquad (16)$$

where $\hat{Q}_T^i$ is the regional "estimated" T-year flood at the site i, Q_T^i is the at-site "true" T-year flood at the site i, and N is the total number of sites used in the performance evaluation. The regional and at-site flood quantiles are dimensionless (the index flood is set to unity) to avoid the problem of index-flood estimation and its effect on performance evaluation.

6. Application

[30] The proposed pooling approach was tested on a large set of catchments from Great Britain (GB). GB was chosen because of its dense network of long-record hydrometric

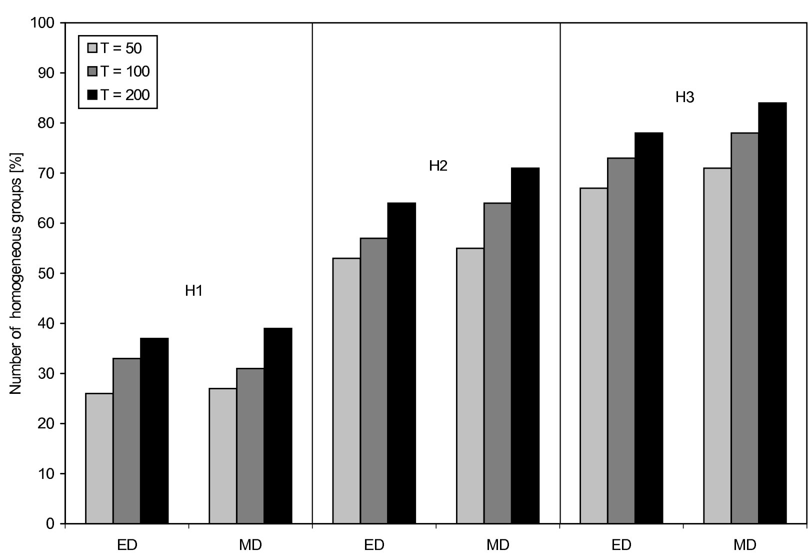

Figure 6. Relative number of homogeneous groups (H ≤ 2) delineated by the ED and MD methods for target return periods T = 50, 100, and 200 years.

Table 2. Performance Measures BIAS and RMSE Calculated for the ED and MD Pooling Methods for Target Return Periods $T = 50$, 100, and 200 Years

	ED Pooling		MD Pooling	
T, years	BIAS, %	RMSE, %	BIAS, %	RMSE, %
50	1.984	4.151	1.266	3.868
100	2.234	6.424	1.960	5.127
200	2.694	8.221	2.232	6.874

stations. The territory of GB is part of the climate region with prevailing baroclinic conditions all year. A detailed description of the climate of GB is given by, e.g., *Manley* [1970]. Flooding arises in GB mostly from frontal cyclonic precipitation, which is most frequent in winter [*Hayden*, 1988]. The description of flood regimes in GB is given by *Bayliss and Jones* [1993], *Black and Werritty* [1997], *Institute of Hydrology* [1999], and *Cunderlik and Burn* [2002b]. The data used in the analysis were obtained from the Flood Estimation Handbook (FEH) flood peak data CD-ROM [*Institute of Hydrology*, 1999]. Peaks-over-threshold (POT) data were used in the analysis since they provide more information on flood seasonality than the annual maximum (AM) data [*Cunderlik et al.*, 2004b]. A 20-year-long period from 1966 to 1985 was selected as a common observation period to avoid possible effects of low-frequency variability in flood seasonality estimated from different observation periods. For every site that overlapped this high-data density period, 60 maximum POT events were identified and included in the database. Thus all sites had the same record length. The catchments included in the database had to be essentially rural with minimal flood attenuation by reservoirs and lakes. A total of 294 sites were selected for the study. Figure 5 shows the selected sites on a map of GB.

[31] For every selected site a pooling group was identified by the previously used Euclidean distance site-focused pooling (ED) and by the proposed Mahalanobis distance pooling (MD). The pooling groups were targeted at return periods T = 50, 100, and 200 years. The homogeneity of the pooling groups was assessed by the Hosking and Wallis homogeneity test. Table 1 summarizes the results. The average values of the test statistics obtained by the MD method were 2–6% (H1), 2–15% (H2), and 6–22% (H3) lower than the average values obtained by the ED method. The average H1 values were high for both methods, but this is the strictest homogeneity measure. The overall variability of the H measures obtained by the MD method dropped by 2–14% (H1), 8–22% (H2), and 3–12% (H3). The H2 measure had the lowest variability among the Hosking and Wallis test measures.

[32] In terms of the maximum test values, the ED pooling led to slightly better results in H1 test for T = 50 and 100 years than the MD pooling. However, the maximum H2 and H3 values were 6–18% and 3–15% lower for the MD method than for the ED maximum values. The minimum values rose between 1 and 30% for the H3 measure for groups targeted at T = 50 years, and for H1 and H3 values for T = 100 years pooling groups. In the remaining scenarios, minimum H values dropped by 1–98%. The average size of a pooling group delineated by the MD method (15.1 sites for T = 50 years, 29.8 sites for T = 100 years, and 59.3 sites for T = 200 years) was very similar to the size of the ED groups (15.3 sites for T = 50 years, 30.0 sites for T = 100 years, and 59.5 sites for T = 200 years).

[33] Figure 6 compares the relative number of homogeneous groups (with H $\leq$ 2) delineated by the ED and MD methods for return periods T = 50, 100, and 200 years. The difference between the evaluated methods in the

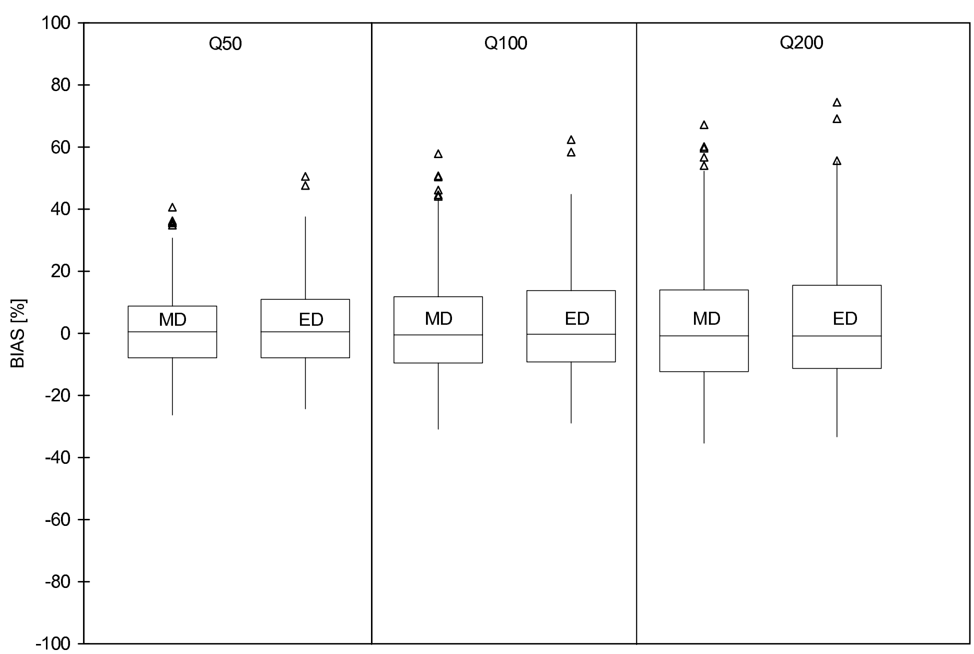

Figure 7. BIAS for the ED and MD pooling methods for target return periods T = 50, 100, and 200 years.

number of homogeneous groups increased as the return period increased. The number of homogeneous groups increased with the return period (size of the group), which is a common drawback of most homogeneity tests. The difference in the number of homogeneous groups between the two approaches was minimal for H1 (4–6%), larger for H3 (5–7%), and most significant for H2 (4–12%).

[34] The second performance analysis carried out in this study compared the performance of the MD and ED pooling approaches in terms of regional quantile estimation. Both the "regional" and "true" quantiles were estimated assuming a generalized-logistic distribution according to recommendations in the FEH [*Institute of Hydrology*, 1999]. Table 2 compares the performance measures BIAS and RMSE for the two pooling approaches. The performance measures were consistently better for the MD approach for all evaluated return periods. In terms of BIAS, the values for the ED approach were 15–50% higher than the values for the MD approach. The differences between the two approaches in terms of RMSE were not as high. The RMSE values for the ED approach were 7–25% higher than the RMSE values for the MD approach. Figure 7 illustrates the BIAS values for both approaches in box plots. It can be seen that the spread of the differences between the estimated and true flood quantiles is higher for the ED approach.

7. Conclusions

[35] Sampling variability adds considerable uncertainty to flood seasonality, especially when estimated from short records. The effect of sampling variability has not yet been addressed in regional flood frequency studies that use flood seasonality measures as pooling variables.

[36] This paper outlined a new flood seasonality approach that simulates sampling variability in flood seasonality by a nonparametric resampling technique. Resampled data were used to estimate the covariance matrix for the target site, as well as for the sites that are potential members of the pooling group for the target site. The covariance matrix was incorporated in the Mahalanobis distance ellipse that reflects the sampling uncertainty and was established for every site. The similarity between the target site and the potential site was then assessed by finding the minimum confidence interval at which the two ellipses intersect. Since the confidence interval reflects the sampling distribution of the observed data, the new pooling measure also quantifies how much two catchments differ in their probability distribution of flood seasonality. This adds a new dimension to site-focused pooling, because the target degree of regional homogeneity can be set in advance. The homogeneity testing is thus integrated into the pooling algorithm.

[37] The proposed approach was applied to a large set of catchments from Great Britain, and its performance was compared with the performance of a previously used pooling based on the Euclidean distance. The results demonstrated that the proposed approach outperformed the previous approach in terms of overall homogeneity of the delineated pooling groups in the study area. The Euclidean distance pooling performed better than the proposed method in terms of maximum H1 values and minimum H3 values, both for return periods of 50 and 100 years. In all other H indicators (minimum, maximum, and mean) and all return periods, the proposed approach led to better results than the Euclidean distance pooling. The decrease in the performance of the proposed method in some H1 indicators is likely related to sample characteristics of the selected data set. Further analysis based on different data sets is needed to confirm this assumption.

[38] In terms of flood quantile estimation, the proposed pooling method outperformed the traditional Euclidean pooling method. The BIAS for the ED approach was 15–50% higher than the BIAS for the MD approach. In terms of RMSE, the ED approach led to RMSE 7–25% higher than the MD approach.

[39] Further work is needed to explore the degree of regional homogeneity quantified by the Mahalanobis confidence interval. Also, the shapes of relations between ellipse distances and the confidence intervals, describing the bivariate sampling variability of flood seasonality between two sites, deserve further research, as they can be used as auxiliary pooling criteria. Future research should also explore the impact of record length on the calculation of the Mahalanobis confidence interval.

[40] Finally, a promising topic is the application of the proposed pooling approach to the estimation of flood quantiles in ungauged catchments and its use with alternative indexes of catchment similarity.

[41] **Acknowledgment.** The authors appreciate the helpful comments from Bellie Sivakumar and two anonymous reviewers.

References

Acreman, M. C., and C. D. Sinclair (1986), Classification of drainage basins according to their physical characteristics: An application for flood frequency analysis in Scotland, *J. Hydrol.*, *84*, 365–380.

Acreman, M. C., and S. E. Wiltshire (1989), The regions are dead; long live the regions: Methods of identifying and dispensing with regions for flood frequency analysis, *IAHS Publ.*, *187*, 175–188.

Ando, Y. (1990), Regionalization of parameters using basin geology, land use, and soil type for use in a storm rainfall-runoff relationship, *IAHS Publ.*, *191*, 211–218.

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*, 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.

Castellarin, A., D. H. Burn, and A. Brath (2001), Assessing the effectiveness of hydrological similarity measures for flood frequency analysis, *J. Hydrol.*, *241*, 270–285.

Cui, Y., J. Weng, and H. Reynolds (1996), Estimation of ellipse parameters using optimal minimum variance estimator, *Pattern Recognit. Lett.*, *17*, 309–316.

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

Cunderlik, J. M., T. B. M. J. Ouarda, and B. Bobée (2004b), Determination of flood seasonality from hydrological records, *Hydrol. Sci. J.*, *49*(3), 511–526.

Dalrymple, T. (1960), Flood frequency analysis, *U.S. Geol. Surv. Water Supply Pap.*, *1543-A*, 11–51.
Efron, B., and R. J. Tibshirani (1993), *An Introduction to Bootstrap*, 436 pp., CRC Press, Boca Raton, Fla.
Fisher, N. I. (1993), *Statistical Analysis of Circular Data*, Cambridge Univ. Press, New York.
Fitzgibbon, A., M. Pilu, and R. B. Fisher (1999), Direct least square fitting of ellipses, *IEEE Trans. Pattern Anal. Mach. Intell.*, *21*(5), 476–480.
Gabriele, S., and N. Arnell (1991), A hierarchical approach to regional flood frequency analysis, *Water Resour. Res.*, *27*, 1281–1289.
Hayden, B. P. (1988), Flood climates, in *Flood Geomorphology*, edited by V. R. Baker et al., pp. 13–27, John Wiley, Hoboken, N. J.
Hosking, J. R. M., and J. R. Wallis (1997), *Regional Frequency Analysis: An Approach Based on L-moments*, 223 pp., Cambridge Univ. Press, New York.
Institute of Hydrology (1999), *Flood Estimation Handbook*, vol. 3, Wallingford, UK.
Krasovskaia, I. (1997), Entropy-based grouping of river flow regimes, *J. Hydrol.*, *202*, 173–191.
Lecce, S. A. (2000), Seasonality of flooding in North Carolina, *Southeast. Geogr.*, *41*(2), 168–175.
Lei, Y., and K. C. Wong (1999), Ellipse detection based on symmetry, *Pattern Recognit. Lett.*, *20*, 41–47.
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.
Mahalanobis, P. C. (1930), On tests and measures of groups divergence, *J. Asiat. Soc. Bengal*, *26*, 541–588.
Mahalanobis, P. C. (1936), On the generalized distance in statistics, *Proc. Natl. Inst. Sci. India A*, *2*, 49–55.
Manley, G. (1970), The climate of the British Isles, in *World Survey of Climatology*, vol. 5, *Climates of Northern and Western Europe*, edited by C. C. Wallen, pp. 81–133, Elsevier, New York.
Mosley, M. P. (1981), Delimitation of New Zealand hydrologic regions, *J. Hydrol.*, *49*, 173–192.
Ouarda, T. B. M. J., F. Ashkar, and N. El-Jabi (1993), Peaks over threshold model for seasonal flood variations, in *Engineering Hydrology: Proceedings of the Symposium*, edited by C. Y. Kuo, pp. 341–346, Am. Soc. Civ. Eng., Reston, Va.
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.
Pearson, C. P. (1991), Regional flood frequency analysis for small New Zealand basins: 2. Flood frequency groups, *N. Z. J. Hydrol.*, *30*(2), 77–90.
Reed, D. W., D. Jakob, A. J. Robson, D. S. Faulkner, and E. J. Stewart (1999), Regional frequency analysis: A new vocabulary, in *Hydrological Extremes: Understanding, Predicting, Mitigating*, *IAHS Publ.*, *255*, 237–250.
Sussman, G. (1999), Ellipse fitting. Presented at Celestial Mechanics Study Group, July 21 1999, http://members.bellatlantic.net/~vze2vrva/ellipse_fitting.html.
U. S. Water Resources Council (1981), Guidelines for determining flood flow frequency, *Bull. 17B*, Hydrol. Comm., Washington, D. C.
Wiltshire, S. E. (1985), Grouping basins for regional flood frequency analysis, *Hydrol. Sci. J.*, *30*, 151–159.

D. H. Burn, Department of Civil Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada, N2L 3G1.

J. M. Cunderlik, Conestoga-Rovers and Associates, 651 Colby Drive, Waterloo, Ontario, Canada, N2V 1C2. (jcunderlik@craworld.com)