WATER RESOURCES RESEARCH, VOL. 37, NO. 6, PAGES 1723–1731, JUNE 2001

Frequency of record-breaking floods in the United States

Richard M. Vogel
Department of Civil and Environmental Engineering, Tufts University, Medford, Massachusetts

Antigoni Zafirakou-Koulouris
Devry Institute of Technology, Long Island City, New York

Nicholas C. Matalas
Vienna, Virginia

Abstract. The theory of record-breaking processes offers a framework for understanding extreme events which is nearly independent of the theory of extremes. The mathematical theory of record-breaking processes is applied to the problem of identifying nonstationarity in hydrological records. A record flood event is simply an event which exceeds all previous events. The probability distribution and first four moments of the number of record events in an n-year period, R, are derived for a serially independent process. The variance of estimates of the mean, standard deviation, and coefficient of variation R is also derived. In addition, approximate confidence intervals are derived for the mean number of record-breaking events in a region with spatially correlated flood series. Using these results, in combination with 1571 flood records in the United States, we document that the average number of record breaking flood events over n-year periods ranging from [10, 80] behaved as if the annual flood series were serially independent for all regions of the United States. However, when spatial correlation of the flood records is ignored, as is the case in many previous studies, it appears as if flood records are not serially independent in the western and Midwestern regions of the United States. These results emphasize the importance of accounting for the spatial correlation structure of hydrologic records when performing regional hypothesis tests.

1. Introduction

In a stationary hydrologic world, persistent or not, all flood and drought records will eventually be broken. This is even true if the probability density function (pdf) of floods or droughts are bounded above or below, respectively. A record event is defined as an event whose magnitude exceeds or is exceeded by all previously recorded events. There is a rich literature on the mathematics of record events [see *Glick*, 1978; *Nevzorov*, 1987; *Nagaraja*, 1988; *Nevzorov and Balakrishnan*, 1998]. The mathematics of record events for the case of univariate sequences is so well developed, that a textbook exists [*Arnold et al.*, 1998]. *Nevzorov* [1987] describes applications of the theory of record events to sports records, traffic jams, rainfall, and other problems. The mathematics literature on record events is quite advanced, and general results exist for most univariate problems relating to record events for independent and identically distributed (iid) random variables. There is a growing literature on the mathematics of record events in multivariate sequences. Since the iid assumption is the only assumption required for most theoretical results pertaining to record events, the theory of records has been suggested for testing the iid assumption [*Foster and Stuart*, 1954]. One goal of this study is to introduce a mathematical framework for evaluating the frequency of record events in hydrology. Another goal is to examine the record breaking properties of historical flood observations in the United States to determine whether or not they behave like serially independent observations. Our final goal is to evaluate the influence of spatial correlation of flood observations on our ability to detect whether or not annual maximum flood series exhibit serial dependence.

Paper number 2001WR900019.
0043-1397/01/2001WR900019$09.00

Most previous literature relating to floods and droughts is an application of extreme value theory and the primary focus is usually on the pdf of either observed annual maximum (or minimum) or partial duration series events. There is no general direct relationship between the theory of extreme value distributions and the distribution of record breaking events because, the theory of extremes ignores the time order of the observations, whereas time plays an essential role in the record breaking process. Instead of focusing on the original random variable such as the series of annual maximum floods, the theory of records focuses on the time series of record events. Hence the theory of records offers a framework for understanding extreme events which is nearly independent of the theory of extremes. There are a few special cases in which the pdf of the original random variable is related to the pdf of the series of record events. Examples of such cases include the Weibull, Pareto, Gumbel, and Power Function pdf's [see *Arnold et al.*, 1998, chapter 2].

There are now a very large number of theoretical and applied studies which have examined the properties of various models of flood and drought frequency when fit to observations (for a review, see *Bobee and Rasmussen* [1995]). Conclusions derived from such studies are limited because one never

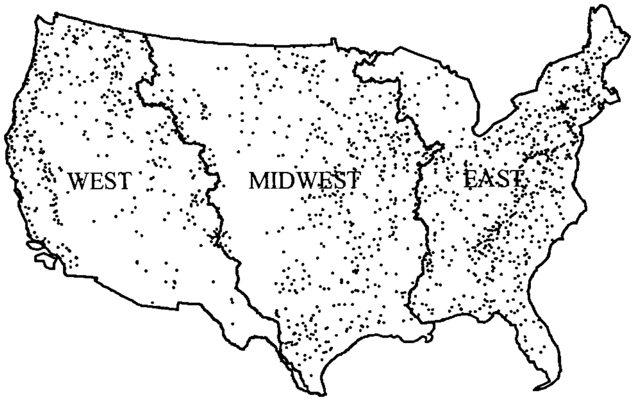

Figure 1. Location of streamflow gaging stations in the conterminous United States and definition of eastern, midwestern, and western regions.

knows which pdf best describes observed flood series. For that reason, many studies have constructed Monte Carlo experiments, generating floods from plausible yet artificial distribution functions. Monte Carlo experiments can lead to definitive conclusions regarding parameter estimation and hypothesis testing, however only with respect to the flood models chosen. Furthermore, such experiments are usually based on the assumption that flood (or drought) events are independent and identically distributed (iid) events over time. The theory of record events enables us to construct experiments with observed flood data without having to resort to a pdf assumption. Record-breaking statistics are nonparametric.

We begin our evaluations with a discussion of U.S. flood observations and envelope curves which provide an upper bound on record-breaking flood experiences to date. Next we summarize and derive some new properties of the record-breaking process for serially independent but spatially correlated events. Finally, we apply that theory to compare the record-breaking frequency of actual floods with our theoretical relations in order to detect any serial dependence.

2. U.S. Flood Observations

We employ the data set compiled by *Slack et al.* [1993], which comprises average streamflow values recorded on a daily, monthly and annual basis in the conterminous United States as well as in Alaska, Hawaii, and the Caribbean. This Hydro-Climatic Data Network (HCDN) is a subset of the much larger U.S. Geological Survey streamflow gaging network and is intended for use in "climate-sensitive" studies. The HCDN contains records spanning the time period 1874–1988 and is available on a CD-ROM and on the World Wide Web. For the purpose of this study, only data pertaining to the 48 conterminous states are considered. The United States is partitioned into three regions, east, midwest, and west as shown in Figure 1, which also illustrates the stream gage locations.

Most flood studies employ annual maximum instantaneous flow records; hence such records might be preferred to the annual maximum daily flows included in the HCDN. On the other hand, annual maximum daily flows are more reliably measured and are a measure of extremes that can reveal trends. According to J. R. Slack (personal communication, 1998), annual maximum daily and annual maximum instantaneous streamflows occur on the same day 73% of the time and, on average, the average daily flow is 80% of the instantaneous maximum flow; the standard deviation of the ratio of the average daily over the instantaneous flows is 0.2.

3. Bounds on Flood Experience

One traditional approach to investigating the record breaking behavior of floods in a region is to create an envelope curve. An envelope curve is a plot of the observed "flood of record" at many sites in a region versus their drainage areas, with an envelope drawn in such a way as to enclose all flood experience to date. Envelope curves sidestep the issue of flood frequency yet they are quite useful in practice, because they provide a bound on flood experience. Envelope curves provide a motivation for the remainder of this study which introduces a statistically rigorous approach for investigating the record breaking properties of floods.

Jarvis [1925] first introduced an envelope curve for flood experience in the United States as of 1924. Later *Crippen and Bue* [1977] updated the Jarvis bound on flood experience in the United States using the flood of record experienced in the United States as of 1974 on 883 basins with drainage areas less than 16,000 km^2 (10,000 miles^2). In the intervening years from 1924 to 1974 the envelope curve developed by *Jarvis* [1925] was exceeded many times in spite of a claim of paleoflood evidence of a natural upperbound to flood magnitudes [*Enzel et al.*, 1993]. *Matalas* [1997] summarizes the Crippen and Bue and Jarvis envelope curves along with the flood data compiled by *Crippen and Bue* [1977] and a world-wide data set compiled through *UNESCO* [1971a, 1971b] for rivers whose drainage areas are one or two orders of magnitude larger than the rivers compiled by Crippen and Bue. Kirby [see *Matalas*, 1997] compiled the experience as of 1994 at 740 of the gaging stations considered by Crippen and Bue. *Matalas* [1997] documents that during the subsequent period 1974–1994 record breaking floods occurred at 88 of the 740 sites considered by Crippen and Bue but still the Crippen-Bue results bounded even this more recent flood experience. In this study we tested to see if flood experience in the United States summarized by the larger group of 1571 stations in the HCDN [*Slack et al.*, 1993] exceeds the Crippen and Bue bound. Again, the Crippen-Bue bound still remains a bound to U.S. flood experience. Unless we have already experienced a natural upper bound on floods in the United States the theory of record-breaking events assures that in time, even the Crippen-Bue bound will be exceeded in spite of arguments to the contrary by *Enzel et al.* [1993]. Yet interestingly, the expected waiting time until the next record flood (or drought) is infinite [*Chandler*, 1952].

4. Number of Record-Breaking Events in a Sequence of Length *n*

Let $X_1, X_2, \ldots, X_n$, represent a sequence of annual maximum flood observations, where n is the total number of time periods for which records are available. The observation X_i is a record high if X_i exceeds all previous records in the sequence, or if and only if $X_i = \max(X_1, X_2, \ldots, X_i)$. The trials at which record highs occur in the original sequence may be expressed as the series of binary variates

$$Y_i = \begin{cases} 1 & X_i = \max(X_1, X_2, \ldots, X_i) \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Let R denote the number of record-breaking events in an n-year period where

$$R = \sum_{i=1}^{n} Y_i. \tag{2}$$

In this initial study we focus our attention on the theoretical properties of the number of record high events R in a series of length n. Other properties such as the waiting time between record events and the time of occurrence of each record event provide substance for future investigations. Similarly, one could focus attention on droughts instead of floods. If the max function in (1) is replaced by min, one obtains the lower record events. Alternatively, one can switch from upper to lower record events by replacing the original sequence by $-X_1$, $-X_2, \ldots, -X_n$. Some initial theoretical results are taken from the mathematics literature and others are introduced here for the first time.

4.1. Probability Distribution of the Number of Record Events

David and Barton [1962] first introduced an expression for the exact probability mass function (pmf) for the number of upper and lower record events in an n-year period. The exact pmf of R can be expressed compactly using Stirling numbers of the first kind S_n^r, where

$$P[R = r] = \frac{|S_n^r|}{(n)!} \tag{3}$$

with

$$S_n^r = \sum_{k=0}^{n-r} (-1)^k \binom{n-1+k}{n-r+k} \binom{2n-r}{n-r-k} \cdot \left[\frac{1}{k!} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^{n-r+k} \right],$$

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

The result in (3) is given in numerous other papers however, the result reported by *Glick* [1978] has a typographic error. A much simpler expression which yields exactly the same pmf as in (3) is $P_n[R = r]$ defined by the recursion

$$P_j[R = r] = \left(1 - \frac{1}{j}\right) P_{j-1}[R = r] + \left(\frac{1}{j}\right) P_{j-1}[R = r - 1] \tag{4}$$

for $r \geq 1$ and $j \geq 2$ with the initial values $P_1[R = 0] = 0$ and $P_1[R = 1] = 1$. *Glick* [1978] also reports the asymptotic result for large sample sizes

$$P[R = r] = \frac{[\ln(n)]^{r-1}}{n \cdot (r-1)!}. \tag{5}$$

Figure 2 compares the exact and asymptotic pmf and cumulative mass function (cmf) of the number of record floods in an n-year period. The cmf is defined by $P[R \leq r] = \sum_{k=1}^{r} P[R = k]$. For values of $n > 20$ the exact expressions for the pmf of R are extremely large numerical problems, making the asymptotic formula quite useful in those situations.

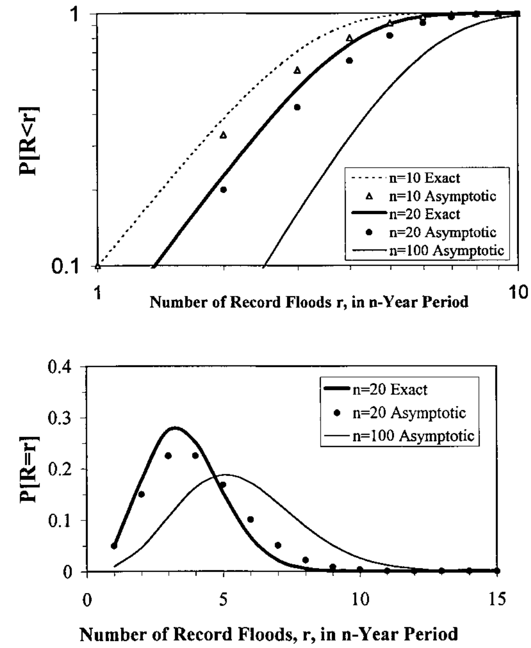

Figure 2. The cumulative mass function $P[R \leq r]$ and the probability mass function of the number of record floods in an n-year period.

4.2. Moments of the Number of Record Breaking Events

The first observation is defined to be a record event. The second observation has an equal chance of being smaller or larger than the first; that is, its probability of surpassing the initial record is exactly 1/2. With probability 1/3 the third event will be a new maximum, since the third observation is equally likely to be smallest, equal, or largest. For an independent sequence all ranks are equally likely hence the maximum rank has probability $1/i$. The following results for the mean and variance of Y and R are due to *Glick* [1978]. The expectation of Y_i, is

$$E[Y_i] = Y_i P\{Y_i = 1\} = 1/i. \tag{6}$$

Similarly,

$$\text{Var}[Y_i] = E[Y_i^2] - (E[Y_i])^2 = 1/i - 1/i^2. \tag{7}$$

In general,

$$E[Y_i^k] = 1/i. \tag{8}$$

The mean and variance of R are then

$$\mu_R = \sum_{i=1}^{n} 1/i, \tag{9}$$

$$\sigma_R^2 = \sum_{i=1}^{n} 1/i - \sum_{i=1}^{n} 1/i^2. \tag{10}$$

A focus of this study is on the sampling properties of estimates of μ_R and σ_R. For this purpose, the moment ratios, skewness and kurtosis are required. Neither of these statistics are reported in the literature so they are derived here. *Zafirakou-Koulouris* [2000] derived the skewness of R, denoted $\gamma_R = E[(R - \mu_R)^3]/\sigma_R^3$ as

$$\gamma_R = \frac{\sum_{i=1}^{n} 1/i - 3\sum_{i=1}^{n} 1/i^2 + 2\sum_{i=1}^{n} 1/i^3}{\left(\sum_{i=1}^{n} 1/i - \sum_{i=1}^{n} 1/i^2\right)^{3/2}}. \tag{11}$$

An alternative to deriving the moments of R is to use the pmf in either (3) or (4) to obtain the noncentral moments of R using

$$E[R^k] = \sum_{i=1}^{n} P_n[R = i]i^k. \tag{12}$$

Then the central moments of R are simply

$$E[(R - \mu_R)^k] = \sum_{j=1}^{k} \binom{n}{j} E[R^{n-j}](-\mu_R)^j, \tag{13}$$

which leads to expressions for the moment ratios, coefficient of variation $C_v[R]$, skewness γ_R, and kurtosis κ_R, respectively:

$$C_v[R] = \frac{E[(R - \mu_R)^2]^{1/2}}{\mu_R} \tag{14}$$

$$\gamma_R = \frac{E[(R - \mu_R)^3]}{E[(R - \mu_R)^2]^{3/2}} \tag{15}$$

$$\kappa_R = \frac{E[(R - \mu_R)^4]}{E[(R - \mu_R)^2]^2} \tag{16}$$

$$\kappa_R \cong 3.19 - \frac{1.42}{n} - \frac{5.43}{n^2} - 0.00419\sqrt{n}. \tag{17}$$

The approximation in (17) is accurate to at least three decimal places for $4 \leq n \leq 100$.

The mean, coefficient of variation, skewness, and kurtosis of R are displayed in Figure 3 as a function of n. For large samples the skewness approaches 0.4, and the kurtosis converges to 3.13. From the central limit theorem, one might expect R to follow a normal distribution for large samples, since it is the sum of the random variable Y. Figure 4 compares a moment ratio diagram for R with a moment ratio diagram for the lognormal, Gamma, and Weibull pdf's. See *Bobee et al.* [1993] for a discussion of moment ratio diagrams. The relation between γ_R and $C_v[R]$ in Figure 4 is based on a variation of the record length from $n = 3$ to $n = 20{,}000$ with points denoted for the cases $n = 10$ and $n = 10{,}000$. Figures 3 and 4 illustrate that in spite of the central limit theorem, the tail behavior of the distribution of R differs significantly from other common distributions even for large sample sizes.

4.3. Sampling Properties of Estimates of Moments R for Independent Samples

Our later experiments with actual flood data require knowledge of the sampling properties of moment estimators of R. Such sampling properties are required to derive confidence intervals for regional average estimates of R, which enable us to evaluate whether actual flood records are in accord with the theory of record events for independent processes. It is always useful to report the sampling properties of statistics along with any sample statistic, and it is common practice to summarize the sampling variability of a statistic by reporting the statistics' standard error. This section derives the standard error of sample estimates of the mean μ_R, standard deviation σ_R, and coefficient of variation $C_v(R)$ of R for serially independent samples.

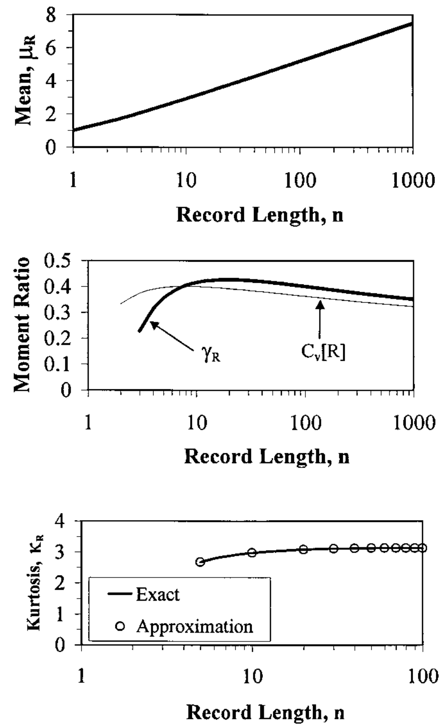

Figure 3. The mean, coefficient of variation, skewness, and kurtosis of R as a function of record length n.

The sampling variance of an estimate of the mean number of record floods $\bar{R}$ in an n-year period for a region of s (spatially) independent (samples) sites is

$$\text{Var}(\bar{R}) = \frac{\sigma_R^2}{s}, \tag{18}$$

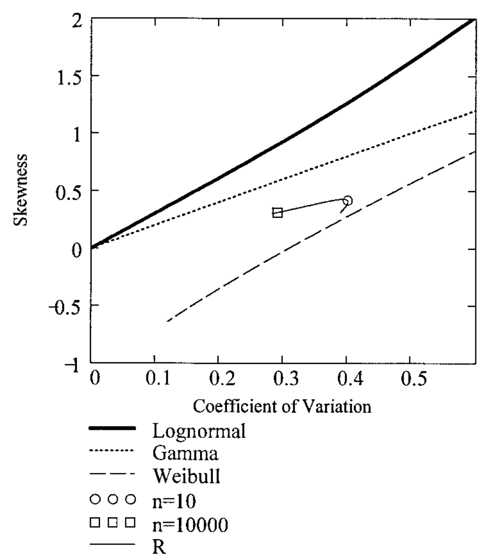

Figure 4. Moment ratio diagram illustrating relationship between skewness and coefficient of variation for the lognormal, gamma, Weibull distributions, and the distribution of R.

where

$$\bar{R} = \frac{1}{s} \sum_{i=1}^{s} R_i \tag{19}$$

and σ_R^2 is given in (10). Here the s sites are assumed to have an equal record length n. *Kendall and Stuart* [1977] document that (18) is an exact result for any independent random variable. An exact result for the sampling variance of the standard deviation, does not exist. Instead, we employ the first-order approximation derived by *Kendall and Stuart* [1977] and others:

$$\text{Var}\ (s_R) \approx \frac{(\kappa_R - 1)\sigma_R^2}{4s}, \tag{20}$$

where σ_R^2 is given in (10), κ_R is given in (16) and (17), and s_R denotes a sample estimate of the standard deviation of R given by

$$s_R = \sqrt{\frac{1}{(s-1)} \sum_{i=1}^{s} (R_i - \bar{R})^2}. \tag{21}$$

Kendall and Stuart [1977] also provide a first-order estimate of the sampling variance of an estimate of the coefficient of variation

$$\text{Var}\ (\hat{C}_v) \approx \frac{(C_v)^2}{s} \left\{ \frac{\mu_4 - \mu_2^2}{4\mu_2^2} + \frac{\mu_2}{\mu_1^2} - \frac{\mu_3}{\mu_2\mu_1} \right\}, \tag{22}$$

where $\mu_k = E[(R - \mu_R)^k]$ denotes the kth central moment defined in (13). Substitution of the first four central moments of R into (22) leads to

$$\text{Var}\ (\hat{C}_v[R]) \approx \frac{(C_v[R])^2}{s} \left\{ \frac{\kappa_R\sigma_R^4 - \sigma_R^4}{4\sigma_R^4} + \frac{\sigma_R^2}{\mu_R^2} - \frac{\gamma_R\sigma_R}{\mu_R} \right\}, \tag{23}$$

where the sample coefficient of variation is defined by $\hat{C}_v[R] = s_R/\bar{R}$ and the theoretical coefficient of variation is defined by $C_v[R] = \sigma_R/\mu_R$.

4.4. Influence of Spatial Correlation on Sampling Properties of Moments of *R*

Observations of floods in a region are correlated in space and this influences the sampling properties of the moments of the number of record events. Assume in a given region that there are s sites (samples), each with a flood flow sequence of length n, where all sequences are concurrent. Let ρ denote a fixed spatial correlation between the observations at all sites in the region. Let R_i denote the number of record events at site i. Then the average number of record breaking events in the region $\bar{R}$, is still given by (19). If the sequences are spatially independent of each other, then $\rho = 0$, and the mean and variance of the average number of records in the region are given by $E[\bar{R}] = \mu_R$ and $\text{Var}[\bar{R}] = \sigma_R^2/s$. On the other extreme, if the s sequences in a region are perfectly correlated with one another, $\rho = 1$ in which case $R_i = R\ \forall\ i$ so that the regional average number of record events $\bar{R}$ is equal to all the at-site values of R. In this case the mean and variance of $\bar{R}$ are $E[\bar{R}] = \mu_R$ and $\text{Var}[\bar{R}] = \sigma_R^2$. The mean number of record events in a region does not depend upon the degree of spatial dependence among the flow sequences, whereas the same cannot be said about the variance of $\bar{R}$. In general, when flow sequences are spatially correlated, Var $[\bar{R}]$ will depend on the number of sites (samples) in the region, s, and the spatial correlation of the number of record events across the sites which we term ρ_R. *Matalas and Langbein* [1962], *Stedinger* [1983], and others document the variance of the regional mean of a variable subject to cross correlation among the stations. In this case, one obtains

$$\text{Var}\ [\bar{R}] = [1 + (s-1)\rho_R]\sigma_R^2/s, \tag{24}$$

which reduces to Var $[\bar{R}] = \sigma_R^2/s$ when $\rho_R = 0$ and to Var $[\bar{R}] = \sigma_R^2$ when $\rho_R = 1$. The current literature on record-breaking processes does not provide a theoretical relationship between the cross correlation of flow sequences ρ and the cross correlation of the number of record events ρ_R; hence we resort to a Monte Carlo experiment. Multivariate normal flow sequences were generated for $s = 2$ and $s = 3$ sites with cross correlation ρ among the flow sequences ranging from 0 to 1 and sample sizes ranging from 10 to 100. The variance of the average number of record events was estimated from 100,000 Monte Carlo experiments and the results are reported in Figure 5 using solid circles. Also shown in Figure 5, using solid lines, is the empirical approximation

$$\text{Var}\ [\bar{R}] = \left[1 + (1-s)\left[\frac{0.326\rho}{\rho - 1.337}\right]\right] \text{Var}\ [R]/s, \tag{25}$$

which is shown to provide a good fit to the Monte Carlo results. Equation (25) is useful for approximating the inflation in the variance of the mean number of record events in a region

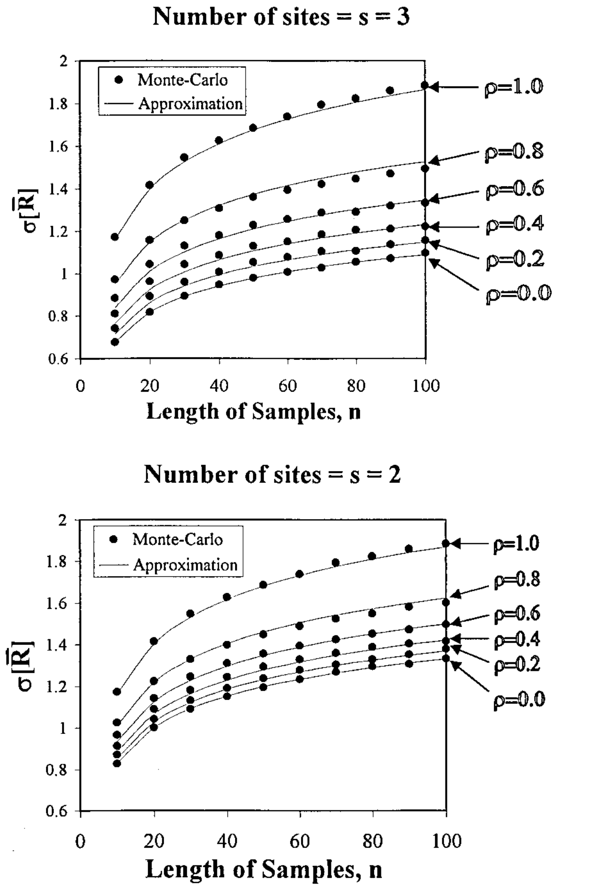

Figure 5. Plot of relationship between standard deviation of $\bar{R}$ based on Monte Carlo experiments and approximate relation given in (25).

which results from spatial correlation of the flow records equal to $\rho > 0$. Note that for the two extreme cases: (1) $\rho = 0$, (25) yields $\mathrm{Var}\,[\bar{R}] = \sigma_R^2/s$ and (2) $\rho = 1$, (25) yields $\mathrm{Var}\,[\bar{R}] \cong \sigma_R^2$ as expected. Since (25) was only developed for small s, its extrapolation later on for large s is speculative; however, it does enable us to approximate the impact of spatial correlation on our evaluations of the frequency of record-breaking properties of floods. A more definitive approach to dealing with spatial correlation of the flood records would have been to employ the bootstrap as suggested by *Walker* [1999] and *Douglas et al.* [2000], or the approach used by *Lettenmaier et al.* [1994].

4.5. Approximate Confidence Intervals for Moments of Number of Record Events

Ideally, exact confidence intervals for each of the moment estimators would be available to enable us to determine, with a specified level of confidence, whether observed flood observations follow the theory of record events for serially independent processes. Owing to the complexity of the distribution of R, the sampling distributions of moment estimators of R, such as $\bar{R}$, s_R, and $\hat{C}_v[R]$ are even more complex than the sampling distribution of R; hence we resort to approximate confidence intervals due to Chebyshev [see *Ross*, 1994]. We realize that Chebyshev confidence intervals are only a crude approximation; however, they are expedient here because they can be easily parameterized to document the influence of spatial correlation on the width of the derived intervals. Chebyshev's inequality for any random variable X with mean μ and variance σ^2 is

$$P[|X - \mu| \geq c] \leq \frac{\sigma^2}{c^2}, \tag{26}$$

where c is a constant equal to one half the width of the confidence interval. In this study we set $c = 3\sigma$, which implies $P[|X - \mu| \geq 3\sigma] \leq 0.11\bar{1}$. For example, application of this result to the statistic $\bar{R}$ leads to an approximate 89% confidence interval equal to $[\bar{R} - 3\sqrt{\mathrm{Var}\,[\bar{R}]}, \bar{R} + 3\sqrt{\mathrm{Var}\,[\bar{R}]}]$. Analogous confidence intervals are constructed for the statistics s_R and $\hat{C}_v[R]$.

5. Record-Breaking Behavior of U.S. Flood Observations

In the following experiments, we report estimates of the regional average, standard deviation and coefficient of variation of R as a function of n for the three geographic regions illustrated in Figure 1. Sample estimates of skewness and kurtosis are known to be significantly biased, so they are not calculated here [*Wallis et al*., 1974; *Vogel and Fennessey*, 1993]. In each of Figures 6, 7, and 8 we compare theoretical and sample estimates of the mean, standard deviation, and coefficient of variation of the number of record floods in an n-year period for the eastern, midwestern, and western regions of the United States, respectively. Also shown are the approximate 89% Chebyshev confidence intervals for each statistic. The heavy confidence intervals denote intervals based on the assumption of spatial independence ($\rho = 0$) of the flood observations. The light weight confidence intervals (shown only for μ_R) are based on the assumption that the cross correlation of the flood observations is equal to the average cross correlation of flow records for all sites in the region.

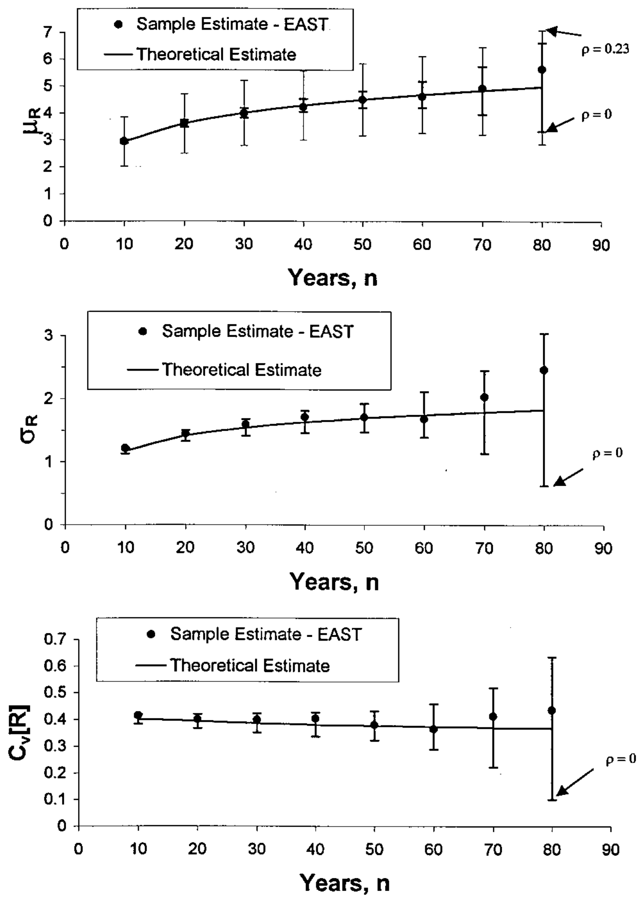

Figure 6. Comparison of sample and theoretical estimates of the mean $E[R]$, standard deviation, $s[R]$, and coefficient of variation $C_v[R]$ of R as a function of n for the eastern region.

Average cross correlations of the annual maximum flow records in the eastern, midwestern, and western regions of the United States are 0.23, 0.19, and 0.42, respectively [*Walker*, 1999]. The average value of ρ for the entire United States is 0.28. These sample estimates of the average spatial correlation of the annual maximum flood series were computed for all possible pairs of observations which had at least 10 years of record in common. Employing the regional average value of cross correlation is the simplest approach to describe the distribution of spatial correlations in a region. *Stedinger* [1983] and *Hosking and Wallis* [1988] also used regional average values of cross correlation to describe the dependence between flow series at different sites. *Douglas et al*. [2000] compared the use of regional trend tests of U.S. flood records based on (1) regional average spatial cross correlations and (2) the bootstrap approach for preserving the empirical regional distribution of the spatial dependence of flood observations. They found good agreement between these two approaches. Nevertheless, our use of a regional average spatial cross correlation is a gross simplification because the complex spatial and temporal climatic mechanisms which give rise to flood observations will lead to spatial correlation structures which depend significantly upon how the regions are defined.

In computing the moments of R all possible nonoverlapping sets of n-year periods within the HCDN database are considered. Table 1 reports the number of such nonoverlapping n-year periods available in each region. The reason that the

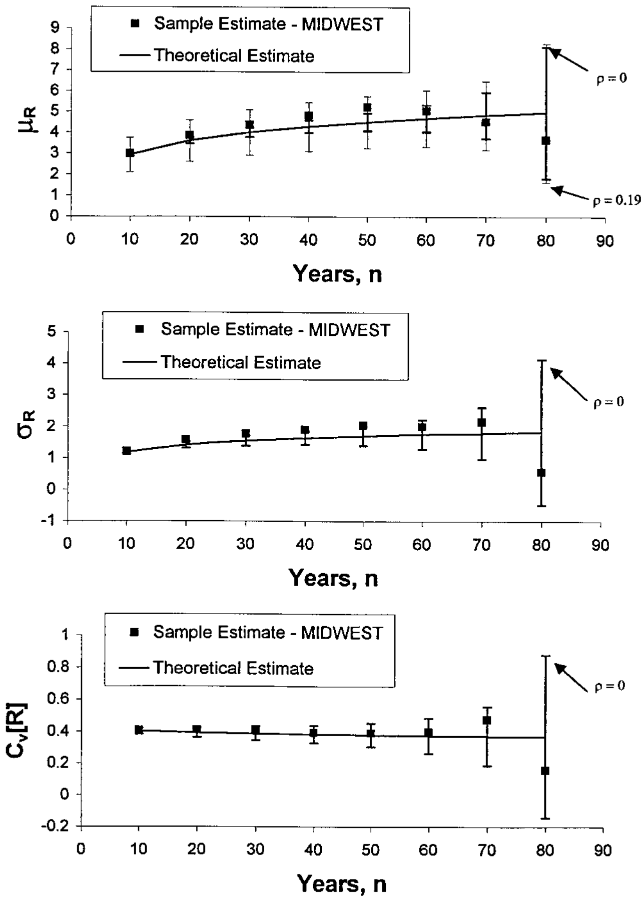

Figure 7. Comparison of sample and theoretical estimates of the mean $E[R]$, standard deviation, $\sigma[R]$, and coefficient of variation $C_v[R]$ of R as a function of n for the midwestern region.

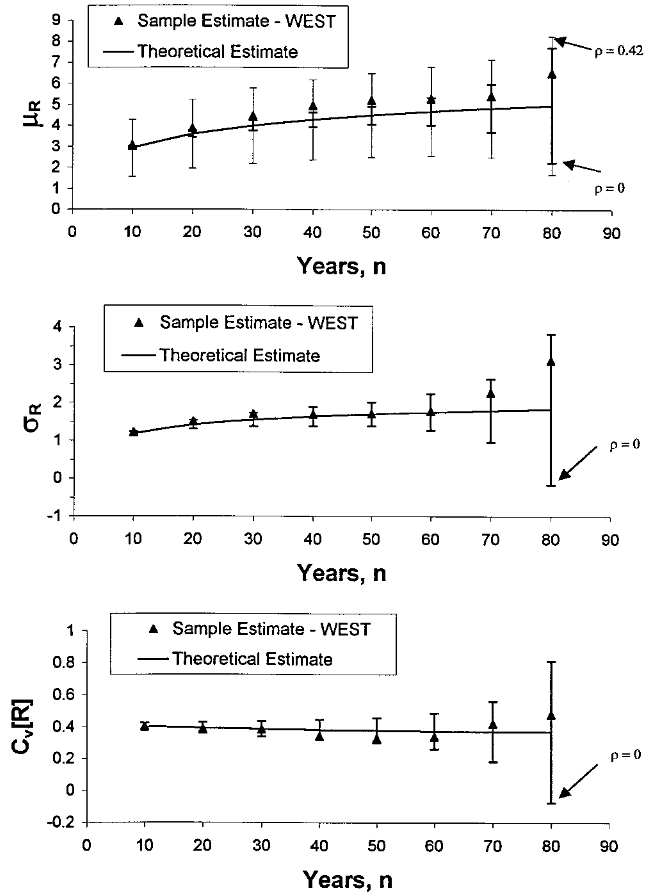

Figure 8. Comparison of sample and theoretical estimates of the mean $E[R]$, standard deviation, $\sigma[R]$, and coefficient of variation $C_v[R]$ of R as a function of n for the western region.

confidence intervals widen as n increases is due to the fact that in each region, the number of nonoverlapping sets of n-year samples decreases as n increases. The confidence intervals reflect the increasing uncertainty associated with our ability to determine properties of record-breaking events as n increases. If we had used smaller regions than defined in Figure 1, the confidence intervals would have widened. If the sample estimates of $\bar{R}$ reported in top graph of Figures 6–8 fall within the reported 89% confidence intervals for μ_R (which account for cross correlation), we conclude that the flood series in that region are serially independent since that was the only assumption required for the theoretical analysis. Note that the confidence intervals for μ_R, which account for the spatial correlation of the flood observations are much wider than the confidence intervals which assume spatial independence.

In general, Figures 6–8 illustrate that when one accounts for the spatial correlation of the flood observations, the observed regional mean $\bar{R}$ falls within the 89% confidence intervals for μ_R for all three regions of the United States. However, if the flood observations are assumed to be spatially independent (which they are not), we would conclude that flood observations in the midwestern and western regions of the United States are serially dependent. Hence our results indicate that flood observations in the eastern United States are consistent with theory of record-breaking phenomena for serially independent processes.

6. Conclusions

The mathematical theory of record-breaking processes was applied to the problem of identifying nonstationarity in hydrological records. The theory of record-breaking phenomena was introduced for serially independent processes. Since serial independence is the only prerequisite assumption to this theory, it is possible to test the hypothesis of serial independence by studying statistics of the record-breaking process. The probability distribution and the first four moments of the number of record breaking events in an n-year period R were introduced for a serially independent process. Sampling properties of es-

Table 1. Number of n-Year Samples s Available in Each Region

Record Length n	East	Midwest	West
10	2680	1919	1561
20	1164	838	665
30	650	432	350
40	418	278	200
50	270	151	136
60	112	66	63
70	36	23	22
80	11	3	4
Total	5341	3710	3001

timators of the mean, standard deviation, and the coefficient of variation of R were derived. In addition, confidence intervals were derived for the mean of R for spatially correlated observations and for the standard deviation and coefficient of variation of R for serially independent observations. Regional average values of R were estimated using flood observations at 1571 watersheds across the United States. Comparisons of the theory introduced with the regional record-breaking behavior of U.S. flood observations led to the following conclusions:

1. Observations of annual maximum flood observations do not appear to exhibit significant serial dependence anywhere in the United States. This conclusion is drawn from the fact that the observed regional mean number of record-breaking events falls within its 89% confidence intervals for the eastern, midwestern, and western regions of the United States. However, since these confidence intervals characterize spatial correlation of flood observations using an average regional value of correlation, the intervals may be slightly wider than they should be. If that is the case, then there may be slight evidence of serial dependence of flood records in the midwestern and western regions of the United States but not in the eastern regions of the United States. Other schemes for handling spatial correlation of the observations within a hypothesis testing framework include a bootstrap approach [*Wilks*, 1997; *Walker*, 1999; *Douglas et al*., 2000] or a logarithmic correlation model [*Lettenmaier et al*., 1994; *Bradley*, 1998], either of which may lead to more reliable confidence intervals than was reported in Figures 6–8.

2. Many previous studies that attempted to evaluate either nonstationarity or persistence of streamflow observations have ignored the spatial correlation among the observations. For example, *Lins and Slack* [1999] ignored spatial correlation of the streamflow records in their national study of trends in streamflow using the same database used here. *Lins and Slack* [1999] found that between 9 and 13% of their sites exhibited significant trends using a 5% level Mann-Kendall test. If they had accounted for the cross correlation of the streamflow observations, it is likely that they would not have found any significant trends at all [see *Douglas et al*., 2000]. Many other examples of trend studies exist which ignored the cross correlation of the flow records. This study has shown that ignoring the spatial correlation of the flows would lead one to conclude that there is significant serial dependence of flood observations in the western and midwestern regions of the United States. However, when one accounts for the cross correlation of the flood observations in the analysis, we conclude that flood observations are serially independent throughout the United States. This conclusion agrees with the results of *Douglas et al*. [2000], who found that when the spatial correlation structure of flood records is properly accounted for, historical flood records in the United States do not exhibit significant upward or downward trends.

3. In a nationwide assessment of streamflow trends which did account for the impact of spatial correlation of streamflows, *Lettenmaier et al*. [1994] found significant trends throughout the United States. Their study differs from this one in many respects: (1) They examined trends in monthly streamflows instead of flood flows; (2) they used a different and slightly smaller data set of streamflows than was used here; (3) they examined trends in streamflow unlike this study which examines serial dependence, in general; and (4) they divided the country into nine regions instead of three as reported in Figure 1. With all these differences it is not surprising they drew different conclusions than this study. They noted a very high degree of spatial coherence among trends, so that the significant uptrends and downtrends tended to occur in clusters within their smaller regions.

4. The theory of record breaking events provides a comprehensive mathematical framework for evaluating the frequency and magnitude of extreme events. Most applications of the theory of records have been to sports such as in determinations of the longest winning streak in professional basketball, or the fastest mile. Many applications to water resources are possible, and it is hoped that some will be inspired by this study.

Acknowledgments. Although the research described in this article has been funded in part by the U.S. Environmental Protection Agency through grant number R 824992-01-0 to Tufts University, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no endorsement should be inferred. The authors are indebted to Jery Stedinger, Tim Cohn, and Upmanu Lall for helpful suggestions which led to considerable improvements in the manuscript.

References

Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja, *Records*, *Ser. Probability Stat*., John Wiley, New York, 1998.

Bobee, B., and P. F. Rasmussen, Recent advances in flood frequency analysis, *U.S. Natl. Rep. Int. Union Geol. Geophys*., *1991–1994*, *Rev. Geophys*., *33*, 1111–1116, 1995.

Bobee, B., L. Perreault, and F. Ashkar, Two kinds of moment ratio diagrams and their applications in hydrology, *Stochastic Hydrol. Hydraul*., *7*, 41–65, 1993.

Bradley, A. A., Regional frequency analysis methods for evaluating changes in hydrologic extremes, *Water Resour. Res*., *34*(4), 741–750, 1998.

Chandler, K. N., The distribution and frequency of record values, *J. R. Stat. Soc*., *Ser. B*, *14*, 220–228, 1952.

Crippen, J. R., and C. D. Bue, Maximum flood flows in the conterminous United States, *U.S. Geol. Surv. Water Supply Pap*., *1887*, 52 pp., 1977.

David, F. N., and D. E. Barton, *Combinatorial Chance*, pp. 178–183, Hafner, New York, 1962.

Douglas, E. M., R. M. Vogel, and C. N. Kroll, Trends in floods and low flows in the United States: Impact of spatial correlation, *J. Hydrol*., *240*(1–2), 90–105, 2000.

Enzel, Y., L. L. Ely, P. K. House, V. R. Baker, and R. H. Webb, Paleoflood evidence for a natural upper bound to flood magnitudes in the Colorado river basin, *Water Resour. Res*., *29*(7), 2287–2297, 1993.

Foster, F. G., and A. Stuart, Distribution-free tests in time-series based on the breaking of records, with discussion, *J. R. Stat. Soc*., *Ser. B*., *16*, 1–22, 1954.

Glick, N., Breaking records and breaking boards, *Am. Math. Mon*., *85*(1), 2–26, 1978.

Hosking, J. R. M., and J. R. Wallis, The effect of intersite dependence on regional flood frequency analysis, *Water Resour. Res*., *24*(4), 588–600, 1988.

Jarvis, C. S., Flood flow characteristics, *Trans. Am. Soc. Civ. Eng*., *88*, 985–1032, 1925.

Lettenmaier, D. P., E. F. Wood, and J. R. Wallis, Hydro-climatological trends in the continental United States, 1948–1988, *J. Clim*., *7*, 586–607, 1994.

Lins, H. F., and J. R. Slack, Streamflow trends in the United States, *Geophys. Res. Lett*., *26*(2), 227–230, 1999.

Kendall, M. G., and A. Stuart, *The Advanced Theory of Statistics*, vol. 1, Chap. 10, 4th ed., Macmillan, Old Tappan, N. J., New York, 1977.

Matalas, N. C., Stochastic hydrology in the context of climate change, *Clim. Change*, *37*, 89–101, 1997.

Matalas, N. C., and W. B. Langbein, Information content of the mean, *J. Geophys. Res*., *67*(9), 3441–3448, 1962.

Nagaraja, H. N., Record values and related statistics—A review, *Commun. Stat. Theory Meth*., *17*, 2223–2238, 1988.

Nevzorov, V. B., Records, *Theory Prob. Appl.*, *32*(2), 201–228, 1987.
Nevzorov, V. B., and N. Balakrishnan, A record of records, in *Handbook of Statistics*, vol. 16, *Order Statistics: Theory and Methods*, edited by N. Balakrishnan and C. R. Rao, pp. 515–570, North-Holland, New York, 1998.
Ross, S., *A First Course in Probability*, Prentice-Hall, Englewood Cliffs, N. J., 1994.
Slack, J. R., A. M. Lumb, and J. M. Landwehr, Hydro-Climatic Data Network (HCDN), *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *93-4076*, 1993.
Stedinger, J. R., Estimating a regional flood frequency distribution, *Water Resour. Res.*, *19*(2), 503–510, 1983.
UNESCO, *Discharge of Selected Rivers of the World,* vol. III, *Monthly and Annual Discharges Recorded at Various Selected Stations (From Start of Observations up to 1964),* Paris, 1971a.
UNESCO, *Discharge of Selected Rivers of the World,* vol. III, *Mean Monthly and Extreme Discharges (1965–1969)* Paris, 1971b.
Vogel, R. M., and N. Fennessey, *L* moment diagrams should replace product moment diagrams, *Water Resour. Res.*, *29*(6), 1745–1752, 1993.
Walker, F. R., Jr., Long term variability in flood arrival rates: The presence of annual flood event clustering, Ph.D. dissertation, Cornell Univ., Ithaca, N. Y., 1999.
Wallis, J. R., N. C. Matalas, and J. R. Slack, Just a moment!, *Water Resour. Res.*, *10*(2), 281–289, 1974.
Wilks, D. S., Resampling hypothesis tests for autocorrelated fields, *J. Clim.*, *10*, 65–82, 1997.
Zafirakou-Koulouris, A., Statistical methods for floods and droughts, Ph.D. dissertation, Tufts Univ., Medford, Mass., 2000.

N. C. Matalas, 709 Glyndon Street, Southeast, Vienna, VA 22180.
R. M. Vogel, Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155. (rvogel@tufts.edu)
A. Zafirakou-Koulouris, Devry Institute of Technology, Long Island City, NY 11101.

(Received September 2, 1999; revised December 22, 2000; accepted January 12, 2001.)