WATER RESOURCES RESEARCH, VOL. 34, NO. 9, PAGES 2271–2280, SEPTEMBER 1998

Modeling river flows with heavy tails

Paul L. Anderson
Department of Mathematics, Albion College, Albion, Michigan

Mark M. Meerschaert
Department of Mathematics, University of Nevada, Reno

Abstract. Recent advances in time series analysis provide alternative models for river flows in which the innovations have heavy tails, so that some of the moments do not exist. The probability of large fluctuations is much larger than for standard models. We survey some recent theoretical developments for heavy tail time series models and illustrate their practical application to river flow data from the Salt River near Roosevelt, Arizona. We also include some simple diagnostics that the practitioner can use to identify when the methods of this paper may be useful.

1. Introduction

In this paper we will discuss the application of heavy tail models to hydrology. Since many river flow time series exhibit occasional sharp spikes, a model that captures this heavy tail characteristic is important in adequately describing the series. Typically, a time series with heavy tails is appropriately transformed so that normal asymptotics apply. We propose a new model that allows a more faithful representation of the river flow without preliminary transformations. As an application, we consider the average monthly flow of the Salt River near Roosevelt, Arizona. The Salt River flow series is periodically stationary; that is, its mean and covariance functions are periodic with respect to time. We fit a periodic autoregressive moving average (ARMA) model to the data without moment assumptions [*Anderson and Meerschaert*, 1997]. We compare this model, which has stable asymptotics, to the classical model presented by *Anderson and Vecchia* [1993], which has normal asymptotics after log transforming the data, so that the innovations have finite fourth moment. Regarding the extreme value behavior of the models, we contrast the classical approach applied to the logarithms of the flow data to the alternative heavy tail approach and demonstrate how the classical approach seriously understates the probability of large fluctuations. In the concluding remarks of the paper we mention some simple diagnostics that the practitioner can use to identify when the methods of this paper may be useful.

We say that a probability distribution has heavy tails if the tails of the distribution diminish at an algebraic rate (like some power of x) rather than at an exponential rate. In this case some of the moments of this probability distribution will fail to exist. The kth moment of a probability distribution function $F(x)$ with density $f(x)$ is defined by

$$\mu_k = \int x^k \, dF(x) = \int x^k f(x) \, dx. \tag{1}$$

The mean μ and variance σ^2 are related to the first two moments by the familiar equations $\mu = \mu_1$ and $\sigma^2 = \mu_2 - \mu_1^2$. Perhaps the most familiar example of a probability distribution with heavy tails is the Cauchy distribution. If X is standard Cauchy, then the density of X is given by

$$f(x) = \frac{1}{\pi(1 + x^2)}, \tag{2}$$

and the distribution function is $F(x) = 1/2 + \pi^{-1} \arctan(x)$. Although the bell-shaped graph of the density of the Cauchy distribution appears to be similar to that of the normal law, the tails are heavier. While the density of the normal law diminishes at an exponential rate, for the Cauchy we have $f(x) \sim \pi^{-1}x^{-2}$ as $|x| \to \infty$. This causes the integral (equation (1)) defining the kth moment to diverge when $k \geq 1$, and hence the mean and standard deviation of the Cauchy are undefined.

The Cauchy and the normal laws are two examples of stable probability distributions. The basic properties of stable distributions are given by *Feller* [1971]. These distributions are the only probability distributions with the property that the sample mean of independent and identically distributed (i.i.d.) observations has the same probability distribution as one of the observations, after a linear rescaling. If $X_1, X_2, X_3, \cdots, X_n$ are independent random variables with the same stable distribution, then the sum $S_n = X_1 + \cdots + X_n$ as well as the sample mean $\bar{X} = S_n/n$ also have a stable distribution. If X_i are standard normal, then

$$\frac{X_1 + \cdots + X_n}{n^{1/\alpha}} \tag{3}$$

has the same distribution as one of the summands when $\alpha = 2$. If X_i are standard Cauchy, then (3) has the same distribution as one of the X_i when $\alpha = 1$. There are stable distributions for every value of $\alpha \in (0, 2]$. For most stable distributions, there exists no closed form expression for the density function. However, it is known that when $\alpha < 2$, we always have $f(x) \sim Cq\alpha x^{-\alpha-1}$ as $x \to -\infty$ and $f(x) \sim Cp\alpha x^{-\alpha-1}$ as $x \to +\infty$, where $C > 0$ and $0 \leq p, q \leq 1$ with $p + q = 1$. When $\alpha = 2$, the distribution is normal, so all moments exist. The moments μ_k of a stable distribution with $\alpha < 2$ exist when $k < \alpha$ and fail to exist when $k \geq \alpha$.

The normal distribution is widely applicable because of the central limit theorem, which guarantees that the distribution of the sum or sample mean of a large number of observations will

Paper number 98WR01449.
0043-1397/98/98WR-01449$09.00

be at least approximately normal. This result depends on the fact that the individual observations do not have heavy tails, so that the theoretical mean and standard deviation exist. The extended central limit theorem states that a similar result holds for heavy tail observations, except that the limiting distribution is stable. *Mandelbrot* [1963] and *Fama* [1965] argue that fluctuations in stock market prices should be modeled as stable random variables. Subsequent research by Mandelbrot and others determined that the sample paths of a stable random walk are actually random fractals with dimension α. An excellent modern reference on stable laws and processes is by *Samorodnitsky and Taqqu* [1994]. *Janicki and Weron* [1994] discuss practical methods for simulating stable stochastic processes. *Mittnik and Rachev* [1995] provide details on stable models in finance, including recent developments in the theory of option pricing. *Nikias and Shao* [1995] focus on applications of stable models in signal processing.

2. Time Series Models with Heavy Tails

Suppose that $\{\varepsilon_t\}$ are independent random variables with common distribution function $F(x) = P[\varepsilon_t \le x]$. Regular variation is a technical condition that is required for the extended central limit theorem to apply. A nonnegative Borel measurable function $R(x)$ varies regularly with index ρ provided that

$$\lim_{x\to\infty} \frac{R(\lambda x)}{R(x)} = \lambda^{\rho} \tag{4}$$

for all $\lambda > 0$. Then for any small $\delta > 0$ we always have $x^{\rho-\delta} < R(x) < x^{\rho+\delta}$ for all x sufficiently large, so that $R(x)$ behaves much like x^{ρ} for large x. For example, the functions $x^{-\alpha}$ and $x^{-\alpha}\log x$ are both regularly varying with index $-\alpha$. If $\rho = 0$, we say that $R(x)$ is slowly varying.

We say that the distribution $F(x)$ belongs to the domain of attraction of some nondegenerate random variable Y with distribution $G(x)$ if there exist real constants $a_n > 0$ and b_n such that

$$\frac{\varepsilon_1 + \cdots + \varepsilon_n - nb_n}{a_n} \Rightarrow Y, \tag{5}$$

where $\Rightarrow$ indicates convergence in distribution. The extended central limit theorem states that (5) holds with Y normal if and only if the truncated second moment function

$$\mu(x) = \int_{|u|\le x} u^2\, dF(u) \tag{6}$$

is slowly varying; (5) holds with Y nonnormal if and only if the tail function $T(x) = P[|\varepsilon_t| > x] = F(-x) + 1 - F(x)$ varies regularly with index $-\alpha$ for some $0 < \alpha < 2$ and the tails satisfy the balancing condition

$$\lim_{x\to\infty} \frac{1 - F(x)}{T(x)} = p \quad \text{and} \quad \lim_{x\to\infty} \frac{F(-x)}{T(x)} = q \tag{7}$$

for some $0 \le p, q \le 1$ with $p + q = 1$; see, for example, *Feller* [1971, p. 577]. The norming constants can be chosen to satisfy $nP[|\varepsilon_t| > a_n] \to C$ and are always of the form $a_n = n^{1/\alpha}\ell_n$, where ℓ_n is slowly varying. The domain of normal attraction, in which we assume that $a_n = n^{1/\alpha}$, is a strictly smaller class of distributions. If $\alpha > 1$, then the mean $E\varepsilon_t$ exists, and we can take $b_n = E\varepsilon_t$. If $\alpha < 1$, then the mean fails to exist, the norming constant $a_n \to \infty$ faster than $n \to \infty$, and we can let $b_n = 0$ for all n. The Pareto distribution $F(x) = 1 - Cx^{-\alpha}$ $(x \ge C^{1/\alpha})$ as well as the type II extreme value distribution $F(x) = e^{-Cx^{-\alpha}}$ $(x \ge 0)$ belong to the stable domain of attraction when $0 < \alpha < 2$. Sums of i.i.d. random variables with these distributions are approximately stable.

The density of most stable distributions cannot be expressed in closed form, and it is most convenient to specify the class of stable distributions in terms of their characteristic functions. The characteristic function or Fourier transform of a random variable Y with distribution function $G(x)$ is $\phi(u) = Ee^{iuY} = \int e^{iux}\, dG(x)$. If Y is stable, then the characteristic function is

$$\phi(u) = \exp\{iu\zeta - \sigma^{\alpha}|u|^{\alpha}[1 - i\theta\, \mathrm{sgn}(u)\tan(\pi\alpha/2)]\} \tag{8a}$$

when $\alpha \ne 1$ or

$$\phi(u) = \exp\{iu\zeta - \sigma|u|[1 + i\theta(2/\pi)\, \mathrm{sgn}(u)\ln|u|]\} \tag{8b}$$

when $\alpha = 1$; see, for example, *Janicki and Weron* [1994] or *Samorodnitsky and Taqqu* [1994]. We say that Y is stable with index α, location parameter ζ, scale parameter σ, and skewness θ. Note that $\alpha \in (0, 2]$, $\gamma \in \mathrm{IR}$, $\sigma \ge 0$, and $\theta \in [-1, 1]$. If we choose the norming constants a_n in (5) to satisfy $nP[|\varepsilon_t| > a_n] \to C$, then the stable limit Y will have scale factor satisfying $\sigma^{\alpha} = C\Gamma(1 - \alpha)\cos(\pi\alpha/2)$ when $\alpha \ne 1$ and $\sigma = C\pi/2$ when $\alpha = 1$. The parameter C is called the dispersion and is sometimes used as an alternative scale parameter for stable laws. The parameters in the balance equations (equation (7)) are related to the skewness of the limit Y by $p = (1 + \theta)/2$ and $q = (1 - \theta)/2$. If $\alpha > 1$ and $b_n = E\varepsilon_t$ or $\alpha < 1$ and $b_n = 0$, then the location parameter of the limit Y in (5) is $\zeta = 0$. If $\alpha \le 1$, then EY does not exist, but if $\alpha > 1$, then $EY = \zeta$, and, in particular, we obtain $EY = 0$ when we center to zero expectation. If Y_n are i.i.d. stable with the same distribution as Y, then $\Sigma\, k_n Y_n$ is also stable with the same stable index and skewness parameters. The location parameter of $\Sigma\, k_n Y_n$ is $\Sigma\, k_n\zeta$. The scale parameter of $\Sigma\, k_n Y_n$ is $\sigma(\Sigma\, |k_n|^{\alpha})^{1/\alpha}$, and so, in particular, the scale parameter of kY is $k\sigma$. The dispersion of $\Sigma\, k_n Y_n$ is $C\, \Sigma\, |k_n|^{\alpha}$.

Example: Use of the dispersion C rather than the scale factor σ is preferable for domains of attraction. Suppose for example that $0 < \alpha < 2$ and that the random variables ε_t are i.i.d. Pareto with distribution $F(x) = 1 - Cx^{-\alpha}$. Then $P[\varepsilon_t > x] = Cx^{-\alpha}$, and if we let $a_n = n^{1/\alpha}$, then $nP[\varepsilon_t > a_n] = nC(a_n)^{-\alpha} = C$, so that

$$\frac{\varepsilon_1 + \cdots + \varepsilon_n - nb_n}{n^{1/\alpha}} \Rightarrow Y, \tag{9}$$

where Y is stable with index α and dispersion C. Since the random variables are positive, we have $p = 1$ and $q = 0$ in the balance equations, so that Y has skewness 1. If $\alpha > 1$ and we center to zero expectation, then Y has mean zero. If we suppose instead that the random variables ε_t are i.i.d. with a type II extreme value distribution $F(x) = e^{-Cx^{-\alpha}}$, or more generally that $1 - F(x) = Cx^{-\alpha}(1 + o(x))$, then we still have $P[\varepsilon_t > x] \sim Cx^{-\alpha}$, and the same asymptotics apply.

Suppose that the i.i.d. sequence $\{\varepsilon_t\}$ represents the innovations process of a time series. If ε_t has a finite fourth moment, then normal asymptotics apply, but if ε_t has an infinite fourth moment, the asymptotics are governed by stable laws. Infinite fourth moments occur when the tail function $T(x) = P[|\varepsilon_t| > x]$ varies regularly with index $-\alpha$ for some $0 < \alpha < 4$. Then

$T(x) \to 0$ about as fast as $x^{-\alpha} \to 0$ as $x \to \infty$. A summary of the basic asymptotic theory for moving averages of random variables with heavy tails is given by *Brockwell and Davis* [1991, pp. 535–545]. *Brockwell and Davis* [1991, p. 535] advise that "any time series which exhibits sharp spikes or occasional bursts of outlying observations suggests the possible use" of these methods. *Kokoszka and Taqqu* [1994] and *Mikosch et al.* [1995] consider ARMA models with infinite variance innovations, while *Kokoszka* [1996] and *Kokoszka and Taqqu* [1996] discuss prediction and parameter estimation for infinite variance fractional autoregressive-integrated moving average (ARIMA) models. *Bhansali* [1993] gives a general method for parameter estimation for linear infinite variance processes. *Anderson and Meerschaert* [1997] develop the basic asymptotic theory for periodic moving averages in the case where the innovations have infinite fourth moment.

Anderson and Vecchia [1993] employ a periodic ARMA model for time series in which both the mean and the covariance structure of the process vary with the season. We will say that $\tilde{X}_t$ follows a PARMA$_\nu(p, q)$ model (a periodic ARMA(p, q) model with period ν) if there exists an i.i.d. sequence $\{\varepsilon_t\}$ such that

$$X_t - \sum_{j=1}^{p} \phi_t(j) X_{t-j} = \sigma_t \varepsilon_t - \sum_{j=1}^{q} \theta_t(j) \sigma_{t-j} \varepsilon_{t-j} \qquad (10)$$

holds almost surely for all t, where $X_t = \tilde{X}_t - \mu_t$. The model parameters μ_t, $\phi_t(j)$, $\theta_t(j)$, and σ_t are all assumed to be periodic with the same period ν. Anderson and Vecchia obtain asymptotic results for the sample autocovariances and sample autocorrelations of periodic ARMA processes in the case where the sequence $\{\varepsilon_t\}$ has finite fourth moment. Some of these results can also be obtained from the work of *Tjostheim and Paulsen* [1982] by rewriting the process as a stationary vector time series of dimension ν. *Adams and Goodwin* [1995] discuss parameter estimation for the periodic ARMA model with finite fourth moments. Although the periodic ARMA can be rewritten as a stationary vector ARMA process, the prediction problem is not the same. For example, monthly data can be used to form a vector of yearly observations, but the one-step prediction should be based on the data from previous months of the same year as well as the data from previous years. Forecasting for the periodic model including the multivariate case is considered by *Ula* [1993]. *Gardner and Spooner* [1994] include an extensive review of results on periodic time series models with finite fourth moments and their applications in signal processing. *Tiao and Grupe* [1980] demonstrate the pitfalls of ignoring seasonal behavior in time series modeling. Seasonal variations in the mean of time series data can easily be removed by a variety of methods. However, when the variance (or dispersion, in the infinite variance case) as well as the mean varies with the season, then the use of periodic time series models is indicated.

Anderson and Meerschaert [1997] compute the asymptotic distribution of the sample autocovariance and sample autocorrelation function for periodic moving averages of i.i.d. random variables with heavy tails. Any periodic ARMA process can be expressed in terms of a periodic moving average

$$X_t = \sum_{j=-\infty}^{\infty} \psi_t(j) \varepsilon_{t-j}, \qquad (11)$$

where the moving average parameters $\psi_t(j)$ are all assumed to be periodic with the same period ν. Suppose that $\{\varepsilon_t\}$ are i.i.d. and that their common distribution $F(x)$ has regularly varying tails with index $-\alpha$ for some $\alpha > 2$. Then $\sigma^2 = E\varepsilon_t^2 < \infty$, and the autocovariance and autocorrelation at season i and lag ℓ are

$$\gamma_i(\ell) = \sigma^2 \sum_j \psi_i(j) \psi_{i+\ell}(j + \ell) \qquad (12)$$

$$\rho_i(\ell) = \frac{\gamma_i(\ell)}{\sqrt{\gamma_i(0)\gamma_{i+\ell}(0)}} = \frac{\sum_j \psi_i(j)\psi_{i+\ell}(j+\ell)}{\sqrt{\sum_j \psi_i(j)^2 \sum_j \psi_{i+\ell}(j)^2}}.$$

The sample mean at season i and the sample autocovariance and sample autocorrelation at season i and lag ℓ are

$$\hat{\mu}_i = N^{-1} \sum_{n=0}^{N-1} X_{n\nu+i}$$

$$\hat{\gamma}_i(\ell) = N^{-1} \sum_{n=0}^{N-1} (X_{n\nu+i} - \hat{\mu}_i)(X_{n\nu+i+\ell} - \hat{\mu}_{i+\ell}) \qquad (13)$$

$$\hat{\rho}_i(\ell) = \frac{\hat{\gamma}_i(\ell)}{\sqrt{\hat{\gamma}_i(0)\hat{\gamma}_{i+\ell}(0)}}.$$

When $\nu = 1$, all of these formulas reduce to the more familiar case of a stationary moving average in which the ψ weights do not vary with the season. The periodic moving average model (equation (11)) is nonstationary, but it is mathematically equivalent to a stationary vector moving average. If we let $Z_t = (\varepsilon_{t\nu}, \cdots, \varepsilon_{(t+1)\nu-1})'$, then the random vectors Z_t are i.i.d. with mean zero and covariance matrix $\sigma^2 I$, where I is the $\nu \times \nu$ identity matrix and $\sigma^2 = E\varepsilon_t^2 < \infty$ for $\alpha > 2$. If we let $Y_t = (X_{t\nu}, \cdots, X_{(t+1)\nu-1})'$, then we can rewrite (11) in the form

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j}, \qquad (14)$$

where Ψ_t is the $\nu \times \nu$ matrix with ij entry $\psi_i(t\nu + i - j)$, and we number the rows and columns $0, 1, \cdots, \nu - 1$ for ease of notation. We define the sample autocovariance by $\hat{\Gamma}(h) = N^{-1} \sum_{t=0}^{N-1} (Y_t - \hat{\mu})(Y_{t+h} - \hat{\mu})'$, where $\hat{\mu} = N^{-1} \sum_{t=0}^{N} Y_t$ is the sample mean, and the autocovariance matrix by $\Gamma(h) = E(Y_t - \mu)(Y_{t+h} - \mu)'$, where $\mu = EY_t$. Note that the ij entry of $\Gamma(h)$ is $\gamma_i(h\nu + j - i)$ and likewise for $\hat{\Gamma}(h)$. The autocorrelation matrix $R(h)$ has ij entry equal to $\rho_i(h\nu + j - i)$ and likewise for the sample autocorrelation matrix $\hat{R}(h)$. The ij term of $R(h)$ is also called the cross correlation of the i and j components of the vector process at lag h. In this application it represents the correlation between season i at year t and season j at year $t + h$.

We will say that the i.i.d. sequence $\{\varepsilon_t\}$ is $RV(\alpha)$ if $P[|\varepsilon_t| > x]$ varies regularly with index $-\alpha$ and $P[\varepsilon_t > x]/P[|\varepsilon_t| > x] \to p$ for some $p \in [0, 1]$. If ε_t is $RV(\alpha)$, then Z_t has i.i.d. components with regularly varying tails, and we will also say that Z_t is $RV(\alpha)$. If $\alpha > 2$, then Z_t belongs to the domain of attraction of a multivariate normal law whose components are i.i.d. univariate normal. *Loretan and Phillips* [1994] find that the price fluctuations of currency exchange rates and stock market prices often follow an $RV(\alpha)$ model with $2 < \alpha < 4$. In this case, $\sigma^2 = E\varepsilon_t^2 < \infty$, but $E\varepsilon_t^4 = \infty$. Since the innovations

ε_t have a finite variance, the sample autocorrelations for the stationary moving average model are asymptotically normal; see, for example, *Brockwell and Davis* [1991, proposition 7.3.8]. The following result shows that when $2 < \alpha < 4$, the sample autocorrelations of the periodic moving average model are asymptotically stable.

Theorem: Suppose that X_t is a periodic moving average of the $RV(\alpha)$ sequence ε_t with $2 < \alpha < 4$ and that $P[\varepsilon_t > x]/P[|\varepsilon_t| > x] \rightarrow p$ for some $p \in [0, 1]$. Then $\sigma^2 = E\varepsilon_t^2 < \infty$, $\hat{\mu}_i$, $\hat{\gamma}_i(\ell)$, $\hat{\rho}_i(\ell)$ are consistent estimators of μ_i, $\gamma_i(\ell)$, $\rho_i(\ell)$, respectively, and

$$N^{-1/2}(\hat{\mu} - \mu) \Rightarrow W, \tag{15}$$

where W is a Gaussian random vector with mean zero and covariance $\sigma^2(\Sigma_j \Psi_j)(\Sigma_j \Psi_j)'$; for some $a_N \rightarrow \infty$ we have

$$Na_N^{-2}[\hat{\gamma}_i(\ell) - \gamma_i(\ell)] \Rightarrow C_{i\ell} = \sum_{r=0}^{\nu-1} C_r(i, \ell)S_r$$
$$Na_N^{-2}[\hat{\rho}_i(\ell) - \rho_i(\ell)] \Rightarrow D_{i\ell} = \sum_{r=0}^{\nu-1} D_r(i, \ell)S_r \tag{16}$$

jointly in $i = 0, \cdots, \nu - 1$ and $\ell = 0, \cdots, h$, where

$$D_r(i, \ell) = \frac{\rho_i(\ell)}{\gamma_i(\ell)} C_r(i, \ell) - \frac{\rho_i(\ell)}{2\gamma_i(0)} C_r(i, 0) - \frac{\rho_i(\ell)}{2\gamma_{i+\ell}(0)} C_r(i + \ell, 0) \tag{17}$$

and $C_r(i, \ell) = \Sigma_j \psi_i(j\nu + i - r)\psi_{i+\ell}(j\nu + i + \ell - r)$. We can always choose a_N so that $NP[|\varepsilon_t| > a_N] \rightarrow C$ for some $C > 0$. Then $S_0, S_1, \cdots, S_{\nu-1}$ are i.i.d. $\alpha/2$ stable with mean zero, skewness 1, and dispersion C_i $C_{i\ell}$ is $\alpha/2$ stable with mean zero, skewness 1, and dispersion $C\ \Sigma_r\ |C_r(i, \ell)|^{\alpha/2}$; and $D_{i\ell}$ is $\alpha/2$ stable with mean zero, skewness 1, and dispersion $C\ \Sigma_r\ |D_r(i, \ell)|^{\alpha/2}$.

Proof. Since $\alpha > 2$, we have $\sigma^2 = E\varepsilon_t^2 < \infty$; see, for example, *Feller* [1971, XVII.5]. Then we can apply *Brockwell and Davis*'s [1991] theorem 11.2.2 to the vector process (equation (14)) to obtain (15), where the form of the covariance follows from the fact that Z_t has covariance matrix $\sigma^2 I$. Since $N^{1/2} \rightarrow \infty$, it follows that $\hat{\mu} \rightarrow \mu$ in probability, and hence $\hat{\mu}_i \rightarrow \mu_i$ in probability for each $i = 0, 1, \cdots, \nu - 1$. Then $\hat{\mu}_i$ is a consistent estimator of μ_i. Theorem 2.2 and corollaries 2.3 and 2.4, along with the remark following the proof of theorem 2.2, from *Anderson and Meerschaert* [1997] yield the first convergence in (16), and then theorem 3.1 and corollaries 3.2 and 3.3 in the same paper yield the second convergence. *Anderson and Meerschaert* [1997] show that in (16) we can always choose $NP[|\varepsilon_t| > a_N] \rightarrow 1$, and then the random variables S_r in the limit are i.i.d. stable with mean zero, index $\alpha/2$, skewness 1, and dispersion 1. Choose $C > 0$ and define $\tilde{a}_N = C^{-1/\alpha}a_N$. Then

$$NP[|\varepsilon_t| > \tilde{a}_N] = \frac{P[|\varepsilon_t| > C^{-1/\alpha}a_N]}{P[|\varepsilon_t| > a_N]} \cdot NP[|\varepsilon_t| > a_N] \rightarrow (C^{-1/\alpha})^{-\alpha} \cdot 1 = C \tag{18}$$

using the fact that $T(x) = P[|\varepsilon_t| > x]$ varies regularly with index $-\alpha$. Recall from section 2 that if Y_n are i.i.d. stable with mean zero, index $\alpha/2$, skewness 1, and dispersion C, then $\Sigma\ k_nY_n$ is also stable with mean zero, index $\alpha/2$, skewness 1, and dispersion $C\ \Sigma\ |k_n|^{\alpha/2}$, and, in particular, kY_n is stable with mean zero, index $\alpha/2$, skewness 1, and dispersion $C|k|^{\alpha/2}$. If we replace a_N with $\tilde{a}_N$ in (16), then we multiply the random variables S_r in the limit by $(C^{-1/\alpha})^{-2} = C^{2/\alpha}$, and so the resulting random variables have dispersion $(C^{2/\alpha})^{\alpha/2} = C$. Then the dispersion of the first limit in (16) is $C\ \Sigma_r|C_r(i, \ell)|^{\alpha/2}$, and the dispersion of the second limit is $C\ \Sigma_r|D_r(i, \ell)|^{\alpha/2}$. *Feller* [1971, XVII.5] shows that a_N varies regularly with index $1/\alpha$, and then *Feller* [1971, VIII.8] shows that for any $\delta > 0$ we have $N^{1/\alpha-\delta} < a_N < N^{1/\alpha+\delta}$ for all large N. Then $N^{1-2/\alpha-\delta} < Na_N^{-2} < N^{1-2/\alpha+\delta}$ for all large N. Since $\alpha > 2$, we have $1 - 2/\alpha - \delta > 0$ for all $\delta > 0$ sufficiently small, and so $Na_N^{-2} \rightarrow \infty$. Then it follows from (16) that $\hat{\gamma}_i(\ell) \rightarrow \gamma_i(\ell)$ and $\hat{\rho}_i(\ell) \rightarrow \rho_i(\ell)$ in probability. This concludes the proof.

3. Detecting Heavy Tails in Time Series Data

In this section we consider the general problem of detecting heavy tails in time series data and estimating the tail parameter α. We illustrate the general problem with a data set representing monthly river flows. The simplest probability model with heavy tails is the Pareto random variable X, whose distribution function is defined by $F(x) = 1 - Cx^{-\alpha}$, where $\alpha > 0$, $C > 0$, and $x \geq C^{1/\alpha}$. If $\alpha \in (0, 2)$, then X is in the domain of attraction of a stable law with index α; otherwise, X is attracted to a normal law. Since a stable law X with index α and dispersion C satisfies $P(X > x) \sim Cpx^{-\alpha}$ and $P(X < -x) \sim Cqx^{-\alpha}$ as $x \rightarrow \infty$, it is reasonable to model a stable law as having Pareto tails. Unfortunately, the MLE for the Pareto distribution parameters α, C is undefined. *Hill* [1975] solves this problem by computing the maximal likelihood estimator (MLE) of the conditional distribution of the r largest-order statistics of X given that they all exceed some fixed number D. Hill calculates that

$$\hat{H}_r = r^{-1} \sum_{i=1}^{r} \ln X_{(i)} - \ln X_{(r+1)} \tag{19}$$

is the conditional maximum likelihood estimator of $1/\alpha$ conditional on $X_{(r+1)} \geq D$, where $X_{(1)} \geq X_{(2)} \geq \cdots$ are the order statistics of a random sample $X_1 \cdots X_n$. We can approximate α by $\hat{\alpha}_r = 1/\hat{H}_r$. *Resnick and Stărică* [1995] show that this procedure yields a consistent estimator of the tail index α for stationary moving average models where the innovations have regularly varying probability tails with index $-\alpha$. Their result can be understood by noting that since the largest observations in a heavy tail time series model tend to be widely spaced in time, they resemble i.i.d. observations.

Figure 1 shows a time series plot for 72 years of river flow data from October 1912 to September 1983. The data measures the average monthly flow rate of the Salt River near Roosevelt, Arizona, in cubic feet per second (historical streamflow data are available on the Web at http://water.usgs.gov). The graph shows occasional sharp spikes characteristic of heavy tail data. We applied *Hill*'s [1975] estimator to the largest $r = 10$ order statistics of the Salt River data and obtained $\alpha = 3.182$, and for $r = 20$ we get $\alpha = 3.023$. This indicates that the Salt River data have heavy tails with infinite fourth moment but finite variance. *Hall* [1982] shows that Hill's estimator is consistent and asymptotically normal with variance approximately α^2/r. For $r = 20$ and $\alpha = 3.023$ we obtain $\sigma = 0.676$, and a z test of $H_0 : \alpha = 4$ versus $H_a : \alpha < 4$ or $H_0 : \alpha = 2$

versus $H_a : \alpha > 2$ has p value 0.07. Increasing r yields a smaller standard deviation, but Hall's theorem assumes that $r/n \to 0$ as $n \to \infty$, so that one should base Hill's estimator on a vanishingly small percentage of the data. We do not expect that the data are exactly Pareto, just that the tails are approximately Pareto, so using only the largest few order statistics is appropriate. Hill's estimator for C is given by

$$\hat{C}_r = \frac{r+1}{n} X_{(r+1)}^{\hat{\alpha}_r}. \tag{20}$$

For the raw Salt River data $r = 10$ we obtain $C = 1.3546 \times 10^{10}$, and for $r = 20$ we obtain $C = 3.34 \times 10^9$.

We will now consider three alternative estimators for the tail index α, including the robust estimator of *Meerschaert and Scheffler* [1998]. These alternate estimation methods also suggest that the true value of α is around 3.0 for the Salt River data. *Hosking and Wallis* [1987] discuss the generalized Pareto (GP) model and its applications to hydrology. If U is standard exponential, then $X = a(1 - e^{-kU})/k$ is GP. Here $k \in \mathbb{R}$ and $a > 0$, and it is not hard to compute that $P(X > x) = (1 - kx/a)^{1/k}$ for $x \geq 0$. If $k < 0$, then $X - a/k$ is Pareto with $\alpha = -1/k$ and $C = (-k/a)^{1/k}$. If $k = 0$, we let $P(X > x) = e^{-x/a}$, which makes the distribution function continuous at $k = 0$. If $k > 0$, then $X \in [0, a/k]$. For the Salt River raw data the MLE for the GP parameters is $a = 618.951$ and $k = -0.315618$, so that the data fit a Pareto model with $\alpha = 3.168$ and $C = 2.8 \times 10^{10}$. *Smith* [1984] shows that this MLE estimate of k is consistent and asymptotically normal with variance approximately equal to $(1 - k)^2/n$. For the raw Salt River data, $(1 - k)^2/n = 0.0015227$, so the standard deviation is 0.0390. A z test based on this statistic comparing $H_0 : k = -0.25$ versus $H_a : k < -0.25$ (that is, $H_0 : \alpha = 4$ versus $H_a : \alpha < 4$) rejects H_0 at the 95% level. Similarly, we can reject $\alpha = 2$ in favor of $\alpha > 2$ with $z = 4.72$ and $p = 0.000$. Note that these standard deviations are much smaller than for *Hill*'s [1975] estimator, and the p values are much smaller, because we are using all of the data.

Mandelbrot [1963] uses a kind of probability plot to demonstrate graphically that the fluctuations in certain cotton exchange prices have heavy tails with $\alpha \approx 1.7$. Suppose that $X_1, \cdots, X_n$ are i.i.d. Pareto with distribution function $F(x)$. Then $F(x) = P[X > x] = Cx^{-\alpha}$, and so $\log F(x) = \log C - \alpha \log x$. Ordering the data as before so that $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(n)}$, we should have approximately that $x = X_{(r)}$ when $F(x) = r/n$. Then a plot of $\log X_{(r)}$ versus $\log(r/n)$ should be approximately linear with slope $-\alpha$. Figure 2 uses this method to illustrate the heavy tail distribution of the Salt River data. It shows the Salt River data along with the line that fits a Pareto model with parameters $\alpha = 3.023$ and $C = 3.34 \times 10^9$ (from *Hill*'s [1975] estimator with $r = 20$). The graph indicates that the true α lies between 2.9 and 3.1.

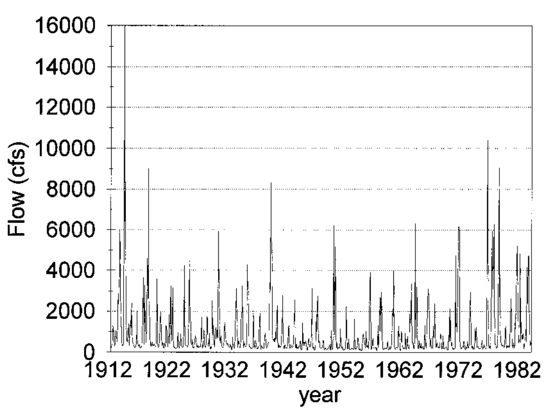

Figure 1. Monthly streamflow for the Salt River near Roosevelt, Arizona, from October 1912 to September 1983 (1 cfs = 2.8317×10^{-2} m^{3}s).

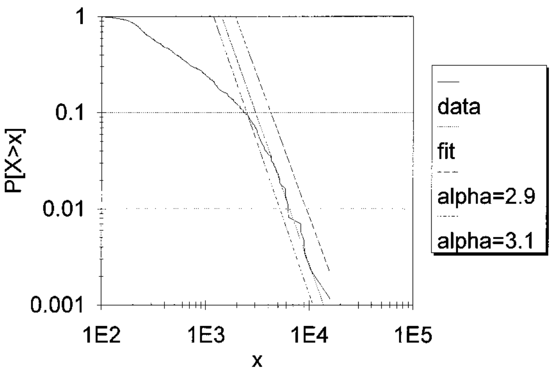

Figure 2. Probability plot for the Salt River showing Pareto tail with α near 3.0.

Meerschaert and Scheffler [1998] propose a robust estimator of $1/\alpha$ given by

$$\hat{\gamma}_n = \frac{\ln \sum_{i=1}^{n} (X_i - \bar{X})^2 - 2 \ln 2}{2 \ln n + \gamma + \ln \pi - 2 \ln 2}, \tag{21}$$

where $\gamma \doteq 0.5772$ is Euler's constant. Here $X_1, \cdots, X_n$ is the (unsorted) data, and $\bar{X}_n$ is the sample mean. When $0 < \alpha < 2$, this formula yields a consistent estimator for $1/\alpha$ for a broad range of time series models, including periodic moving averages, whose innovations have regularly varying tails with index α. If $\alpha > 2$, then the estimator $\hat{\gamma}_n \to \frac{1}{2}$ in probability as $n \to \infty$. Since we believe that $2 < \alpha < 4$, we apply the estimator to the squared data, which has tail index $\alpha/2$. The *Meerschaert and Scheffler* [1998] estimator is unbiased when the data are Pareto-like with $C = 1$. For a Pareto this can be accomplished by dividing by $C^{1/\alpha}$, and so we divided the raw data by $C^{1/\alpha} = 1413$, obtained from *Hill*'s [1975] estimator with $r = 20$. Then we applied the above estimator of α and doubled the result (since the index of the squared data is $\alpha/2$) to get 3.104 as the Meerschaert-Scheffler estimate of α for the raw Salt River data. Since the results of Hill's estimator are consistent with the estimates obtained by a variety of alternative methods, we are fairly confident that the Salt River flow has heavy tails with α near 3.

4. A Periodic ARMA Model for the Salt River

In this section we illustrate the application of heavy tail time series methods by fitting a periodic ARMA model to the Salt River flow data. Let $\bar{X}_t$ denote the average flow rate in cubic feet per second t months after October 1912 of the Salt River near Roosevelt, Arizona. The tail parameter α dictates our modeling approach. If $\alpha > 4$, then the time series has finite fourth moments, and the classical approach based on normal asymptotics is appropriate. If $\alpha < 2$, then both the autocovariance and autocorrelation of the time series are undefined. Preliminary estimates of the tail index α in section 3 indicate

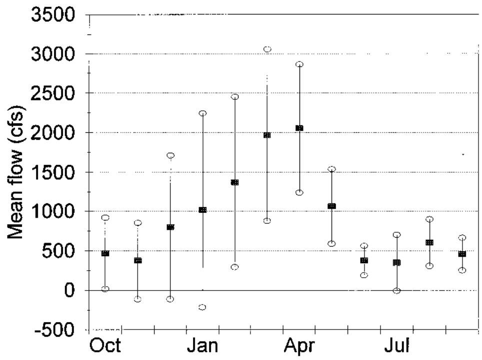

Figure 3. Monthly sample means for the Salt River, including 95% confidence bands.

that $2 < \alpha < 4$, so the probability distribution of $\tilde{X}_t$ has heavy tails, with a finite variance but infinite fourth moment. Since both the mean and the correlation structure of the process vary significantly by month, we will fit a periodic ARMA model. We find that a PARMA$_{12}$(1, 0) model is sufficient to capture the most important features of the data. The model is

$$X_t - \phi_t X_{t-1} = \sigma_t \varepsilon_t, \tag{22}$$

where $X_t = \tilde{X}_t - \mu_t$ is the mean-standardized process, ε_t is the standardized heavy-tailed innovations process, and the model parameters μ_t, ϕ_t, and σ_t are all periodic with the same period $\nu = 12$, so that for example we have $\mu_{t\nu+i} = \mu_i$ for all t. Since $\alpha > 2$, the theorem given in section 2 implies that the sample mean $\hat{\mu}_i$, sample autocovariance function $\hat{\gamma}_i(\ell)$, and the sample autocorrelation function $\hat{\rho}_i(\ell)$ are consistent estimators of the true mean, autocovariance, and autocorrelation. We first estimate the mean flow μ_i for month $i = 0, 1, \cdots, 11$ by the sample mean $\hat{\mu}_i$ on the basis of $N = 72$ years of flow data. We adapt the procedure of *Brockwell and Davis* [1991, p. 407] to obtain simultaneous 95% confidence intervals for $\mu_0, \cdots, \mu_{11}$. By the theorem given in section 2 we have $N^{-1/2}(\hat{\mu}_i - \mu_i) \Rightarrow W_i$, where W_i is normal with mean zero and variance

$$v_i = 2\pi f_i(0) = \sum_{j=-\infty}^{\infty} \gamma_i(j), \tag{23}$$

where $f_i(\omega)$ is the spectral density for the ith seasonal component. The lag window estimator

$$2\pi\hat{f}(0) = \sum_{|h|\le r} \left(1 - \frac{|h|}{r}\right) \hat{\gamma}_i(h) \tag{24}$$

is consistent when $r = r_n \to \infty$ in such a way that $r_n/n \to 0$. Then, with probability $1 - \delta$,

$$|\hat{\mu}_i - \mu_i| \le \Phi_{1-\delta/(2\nu)} \sqrt{2\pi\hat{f}_i(0)/N} \qquad i = 0, 1, \cdots, 11 \tag{25}$$

is approximately true for large r and N. Here Φ_p is the p percentile of a standard normal distribution. We apply this formula with $r = 10$ and $\delta = 0.05$ ($\nu = 12$ and $N = 72$) to obtain the confidence bands in Figure 3. Even by this extremely conservative method we are able to reject the hypothesis that the seasonal means are equal, since, for example, the confidence intervals for months 6 and 8 do not overlap. Since the seasonal means differ significantly, we mean correct the data by season. Next we compute the sample autocorrelation func-

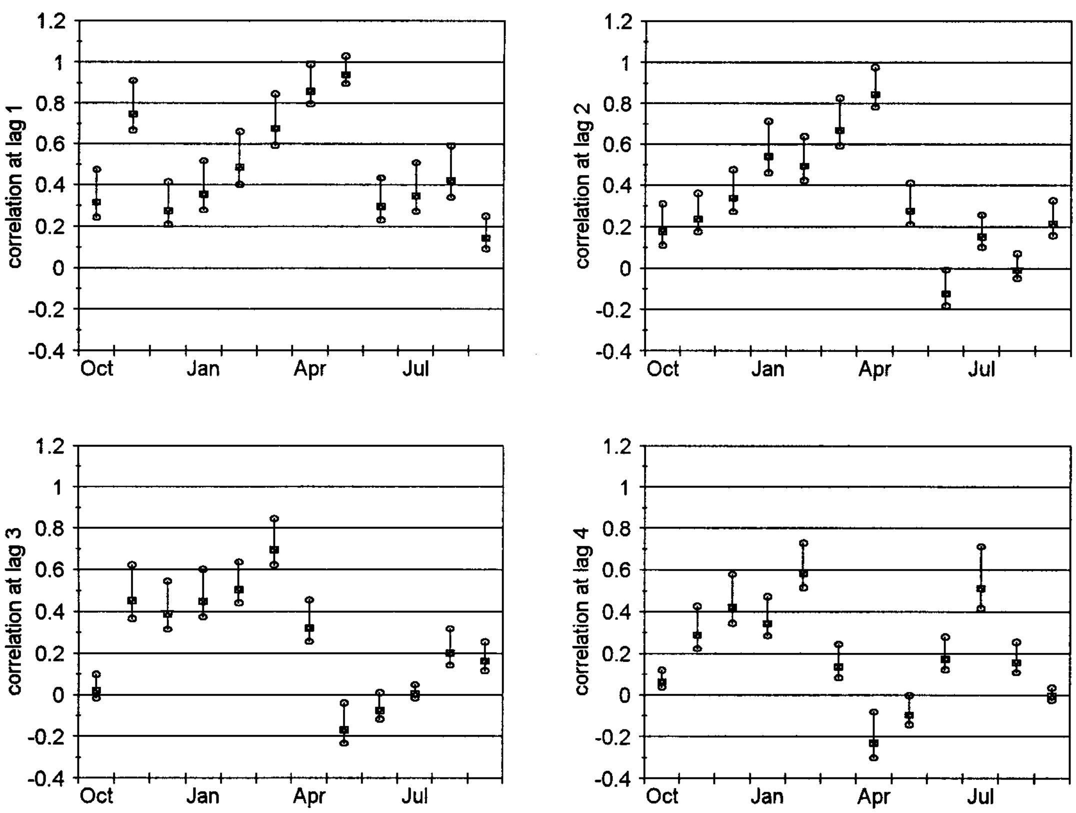

Figure 4. Monthly autocorrelations for the Salt River, including 95% confidence bands.

tion $\hat{\rho}_i(\ell)$, which is displayed in Figure 4 along with the appropriate 95% confidence bands, for the first few lags. By the theorem given in section 2 we have $Na_N^{-2}[\hat{\rho}_i(\ell) - \rho_i(\ell)] \Rightarrow D_{i\ell}$, where $D_{i\ell}$ is stable with index $\alpha/2$, mean zero, skewness 1, and dispersion $d_{i\ell} = \Sigma_r |D_r(i, \ell)|^{\alpha/2}$. Since the dispersion depends on the model parameters, it is necessary to fit a model before constructing confidence intervals. Using consistent estimators (discussed below) of the model parameters, we compute the dispersion $d_{i\ell}$. The scale factor for $D_{i\ell}$ is given by $\sigma_{i\ell}^{\alpha/2} = d_{i\ell}\Gamma(2 - \alpha/2) \cos(\pi\alpha/4)/(1 - \alpha/2)$. Then, with probability $1 - \delta$,

$$\hat{\rho}_i(\ell) + s_{\delta/(2\nu)}\sigma_{i\ell}a_N^2/N \le \rho_i(\ell) \le \hat{\rho}_i(\ell) + s_{1-\delta/(2\nu)}\sigma_{i\ell}a_N^2/N \quad (26)$$

$$i = 0, 1, \cdots, 11$$

is approximately true for large N and fixed ℓ. Here s_p is the p percentile of a stable distribution with mean zero, skewness 1, and scale factor 1, which can be obtained from the accurate tables of *McCulloch and Panton* [1996]. Notice that the stable confidence intervals are asymmetric because of the skewness of the limit $D_{i\ell}$. We apply this formula with ($\nu = 12$ and $N = 72$) $\delta = 0.05$ and $a_N = (CN)^{1/\alpha}$, where α and C are obtained from *Hill*'s [1995] estimator with $r = 20$, to obtain the confidence bands in Figure 4. Since both the mean and the autocorrelation function vary significantly by month, it is appropriate to employ a periodic ARMA model. We have not been able to find any simple transformation of the data, such as differencing, which would allow the use of a stationary time series model. Note also that although the mean- and variance-standardized data appear to fit an AR(1) model quite well on the basis of the partial autocorrelation function (graph not shown), this is misleading because the autocorrelation function of this series is the same as for the original data and clearly shows a seasonal pattern to the covariance structure. Failure to consider the possibility of seasonal variations in the covariance structure can easily lead to a misspecified model, which is quite serious in this case since a stationary AR(1) model has normal asymptotics for the autocovariance function (ACF), while a periodic AR(1) model has stable asymptotics for the ACF.

We estimate the model parameters using the method of moments. A simple calculation shows that for a $\text{PARMA}_{12}(1, 0)$ model,

$$\phi_t = \rho_{t-1}(1)\sqrt{\gamma_t(0)/\gamma_{t-1}(0)}$$
$$\sigma_t = \sqrt{\gamma_t(0)(1 - \rho_{t-1}(1)^2)}, \quad (27)$$

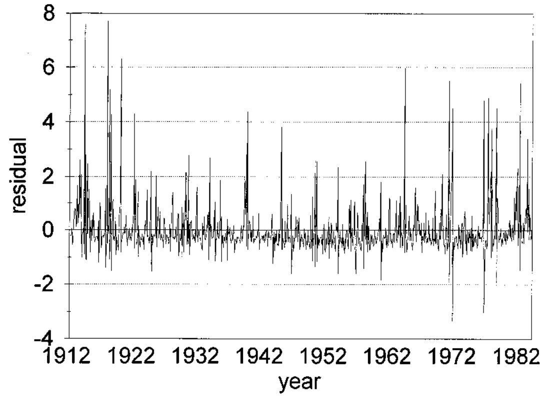

Figure 5. Residual plot for $\text{PARMA}_{12}(1, 0)$ model for the Salt River.

and since $\alpha > 2$, we can substitute $\hat{\gamma}_i(0)$ and $\hat{\rho}_i(1)$ to obtain consistent estimators of the model parameters. Table 1 shows the resulting model parameter estimates. It follows from (22) that $\sigma_t\varepsilon_t = X_t - \phi_t X_{t-1}$, where $X_t = \tilde{X}_t - \mu_t$ is the mean standardized process. Then the estimated residuals can be obtained using

$$\hat{\varepsilon}_t = \frac{\tilde{X}_t - \hat{\phi}_t\tilde{X}_{t-1} - \hat{\mu}_t + \hat{\phi}_t\hat{\mu}_{t-1}}{\hat{\sigma}_t}. \quad (28)$$

Figure 5 shows a time series plot of the standardized residuals. Figure 6 plots the autocorrelation function for the estimated residuals. Since $\alpha > 2$, the usual normal asymptotics apply here, and so the 95% confidence band at $1.96(N\nu)^{-1/2}$ is appropriate. Although a few of the autocorrelations lie slightly outside of this band, there is no periodic pattern in the autocorrelations, and we consider the model fit to be adequate. The stable index α which appears in the asymptotics of the autocovariance and autocorrelation of the raw time series data is the same as for the innovations ε_t. If we apply *Hill*'s [1975] estimator to the absolute residuals, we should get approximately the same value of α as for the raw data. In fact, an application of Hill's estimator with $r = 14$ yields $\hat{\alpha} = 3.067$ (and $\hat{C} = 0.0397$), which is similar to the estimates obtained above using the raw data and well within the asymptotic confidence intervals for those estimates. Naturally, the appropriate value of r is somewhat smaller for the residuals, since one large innovation in a linear time series model can result in

Table 1. Model Parameters

i	$\hat{\gamma}_i(0)^{1/2}$	$\hat{\rho}_i(1)$	$\hat{\phi}_i$	$\hat{\sigma}_i$
0	796.92	0.31630	0.33978	788.82
1	361.33	0.74372	0.14341	342.88
2	1290.12	0.27360	2.75546	862.43
3	2055.65	0.35421	0.43595	1977.22
4	1792.86	0.48316	0.30891	1676.53
5	1950.48	0.67288	0.52567	1707.71
6	1566.30	0.85675	0.54035	1158.67
7	1060.65	0.93832	0.58017	547.00
8	303.17	0.29441	0.26821	104.82
9	397.33	0.34682	0.38585	379.72
10	483.83	0.42057	0.42231	453.80
11	333.63	0.14225	0.29001	302.79

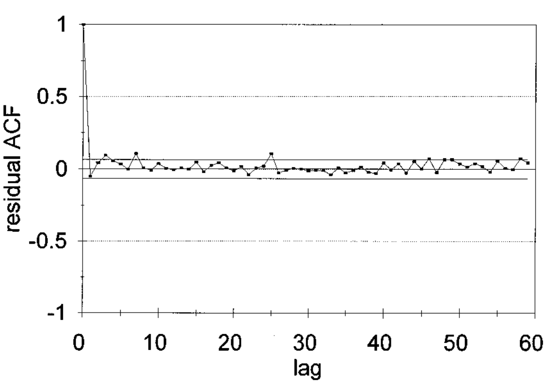

Figure 6. Residual autocorrelation function for $\text{PARMA}_{12}(1, 0)$ model for the Salt River.

several large observations. Using the same value of r for the residuals yields a somewhat smaller estimate of α, but this is misleading. Applying the quadratic estimator of *Meerschaert and Scheffler* [1998] to the residuals yields $\hat{\alpha} = 3.20$, which is also consistent with our results obtained from the raw data. Since the residuals are standardized, there is no need to rescale. The MLE for the generalized Pareto model applied to the positive residuals yields $\hat{\alpha} = 2.83$, which is also within the confidence bands.

5. Extreme Values

Suppose that $\{\varepsilon_t\}$ are independent random variables with common distribution function $F(x) = P[\varepsilon_t \le x]$. The same technical condition of regular variation required for the extended central limit theorem is also relevant for extreme values. We say that the distribution $F(x)$ belongs to the extremal domain of attraction of some nondegenerate random variable Z with distribution $H(x)$ if there exist real constants $a_n > 0$ and b_n such that

$$\frac{\max\{\varepsilon_1, \cdots, \varepsilon_n\} - nb_n}{a_n} \Rightarrow Z, \tag{29}$$

where $\Rightarrow$ indicates convergence in distribution. For probability distributions with heavy tails the possible limit distributions are called the type II max-stable distributions. They are of the form $H(x) = \exp(-Cx^{-\alpha})$ for some $C > 0$ and $\alpha > 0$. The probability distribution $F(x)$ belongs to the extremal domain of attraction of this type II max-stable law if and only if the tail $1 - F(x)$ varies regularly with index $-\alpha$. In this case we can take a_n to satisfy $nP[\varepsilon_t > a_n] \to C$ and $b_n = 0$. The norming constants always satisfy $a_n = n^{1/\alpha}\ell_n$, where ℓ_n is slowly varying. The limit Z is called max-stable because if Z_n are i.i.d. with the same distribution as Z, then $\max\{Z_1, \cdots, Z_n\}$ has the same distribution as Z after a linear rescaling. Since the largest observations in the Salt River data are widely spaced in time, they should resemble i.i.d. observations. In this section we will use the model (equation (29)) to predict the extreme value behavior of the Salt River flow.

In section 3 we use *Hill*'s [1975] estimator with $r = 20$ to obtain estimates of $\alpha = 3.023$ and $C = 3.34 \times 10^9$ for the Salt River data. Then for large n the maximum flow M_n over n consecutive months has approximately the same distribution as $n^{1/\alpha} Z$, where Z has a type II max-stable distribution with parameters α and C. The p percentile of this distribution is given by $P[n^{1/\alpha}Z \le p] = \eta(p)$, where

$$\eta(p) = \left(\frac{nC}{-\ln p}\right)^{1/\alpha}, \tag{30}$$

and, in particular, the median of the maximum flow over n months is $1595n^{1/\alpha}$. The predicted maximum flows for 10, 50, 100, and 500 years are 7800, 13,000, 17,000, 28,000 cfs (1 cfs = 2.8317×10^{-2} m^3/s). Note that the predicted maximum increases algebraically with n because of the heavy tails. For the Salt River data the largest observation over this 72 year period was 15,990 cfs, which is between the 50 year flood level and the 100 year flood level. There was one observation above the 50 year flood level and six observations above the 10 year flood level in 72 years. The probability distribution of M_n also has heavy tails, so that a value far above the median would not be surprising. For example, the 75th percentile of the 100 year maximum flow is over 22,000 cfs, and the 90th percentile is over 30,000 cfs.

A lognormal model gives a much poorer fit to the data. *Leadbetter et al.* [1980] show that if $X_1, X_2, X_3, \cdots$ are independent normal random variables with mean μ and variance σ^2 and $M_n = \max\{X_1, \cdots, X_n\}$, then

$$M_n - \sigma\sqrt{\log n} \to \mu \tag{31}$$

in probability, so $M_n \approx \mu + \sigma\sqrt{\log n}$ for large n. Taking X_t to be the natural logarithm of the Salt River flow at month t, the maximum flow over n consecutive months is $\exp(M_n) \approx \exp(\mu + \sigma\sqrt{\log n})$. We compute $\mu \approx \bar{X} = 6.2139$ and $\sigma \approx s = 0.9954$, and then the predicted maximum flows for 10, 50, 100, and 500 years are 4400, 6200, 7000, and 9400 cfs. In 72 years, there were two observations exceeding 9400 and six observations exceeding 7000, indicating that the lognormal model seriously understates the probability of major floods.

Also of interest is the number of exceedances over a given level $L > 0$ over a given period. For large m the number of observations in m months that exceeds L is approximately Poisson with mean $mCL^{-\alpha}$; see, for example, *Leadbetter et al.* [1980, theorem 3.1.1]. If we take L to be the p percentile of the maximum flow over a period of n months, then the mean number of exceedances in m months is

$$\mu = mC\left[\left(\frac{nC}{-\ln p}\right)^{1/\alpha}\right]^{-\alpha} = (m/n)(-\ln p), \tag{32}$$

which indicates, for example, that the probability of a river flow exceeding its N year flood level (defined as the median of the probability distribution of the maximum flow over N years) more than once in any given N year period is only 0.15, and the probability of exceeding its N year flood level more than twice is only 0.03. The expected number of exceedances of the 10 year flood level over a 72 year period is 7.2 ln 2, which is embarrassingly close to the actual number of six exceedances observed in 72 years of data.

6. Concluding Remarks

Hydrologic streamflows can exhibit both heavy tails and nonstationarity. Data analysis should include an examination of the seasonal mean and standard deviation. If the seasonal standard deviation varies significantly, removing the seasonal mean is insufficient to produce a stationary time series, and a periodic ARMA model may be appropriate. We advise screening for heavy tails whenever a streamflow exhibits annual order-of-magnitude fluctuations. Occasional sharp spikes in a time series plot, numerous outliers, or a histogram can also indicate heavy tails. When heavy tails arise, the tail parameter α should be estimated using one or more of the methods mentioned in this paper. This parameter governs the extreme behavior (i.e., flood levels), and if $\alpha < 4$, it also determines the rate of convergence of the sample autocorrelations. Taking logarithms removes the heavy tail but may distort flood level predictions.

The difficulty in estimating the tail parameter α seems to lie at the heart of an ongoing controversy in finance. *Mandelbrot* [1963] and *Fama* [1965] argue that variations in stock market prices and currency exchange rates follow a stable distribution with $0 < \alpha < 2$, while *Loretan and Phillips* [1994] and *Jansen and de Vries* [1991] use *Hill*'s [1975] estimator to compute that these price fluctuations have heavy tails with $2 < \alpha < 4$.

McCulloch [1995] points out that Hill's estimator often gives very poor estimates of α when the data is actually stable with $\alpha \in (1.5, 2.0)$. The authors verified McCulloch's claim by computing Hill's estimator for the largest 5% of 1000 simulated stable random variables with $\alpha = 1.8$. We used the simulation method of *Chambers et al*. [1976] as implemented by John Nolan (personal communication, 1997); see also *Weron* [1996]. In repeated simulations the resulting estimates were near 3.0 in most trials. Of course, it is possible that some economic time series have $\alpha < 2$ and others have $2 < \alpha < 4$, but the discrepancy thus far seems more a matter of modeling assumptions. Maximum likelihood estimation of the stable index will always yield $0 < \alpha \leq 2$ since the model assumes that α lies in this range. On the other hand, although the tails of a stable model have Pareto-like tails, it is not clear that this proportion of the data is large enough to allow accurate estimation of the tail index by methods such as Hill's estimator. Further research is needed to establish the asymptotic theory of the various α estimators in this case, especially for nonstationary time series models. For the hydrologic data we examined, the existing α estimators seem to be adequate.

The $PARMA_{12}(1,0)$ used in this paper is not intended to represent the best possible model for the Salt River streamflow. This model does remove most of the serial correlation (see Figure 6), and our moment estimates for the model parameters are easily obtained. Our primary goal for this model was to check the tail estimates of the α parameter for the model residuals, which can be assumed to be i.i.d. The fact that the α estimates for the residuals agree with the estimates for the raw data suggests that heavy tails can be detected prior to time series modeling. This is important in practice, since the asymptotics of the sample moments depend on the tail behavior (see the theorem and the discussion preceeding it in section 2). In a forthcoming paper (P. Anderson et al., manuscript in preparation, 1998) we establish the consistency of an innovations algorithm for periodic time series data, and we use this algorithm to obtain a better model fit for the monthly Salt River data. It would be interesting to compare the model fit and forecasting performance of that model with the lognormal model used by *Anderson and Vecchia* [1993].

The Salt River data were aggregated from daily averages. It would also be possible to fit the daily data using a $PARMA_{365}(p, q)$ model, but this involves significantly more parameters. Monthly averages allow a more parsimonious model, and it seems unlikely that day of the week or month effects are significant in this hydrologic time series. For other applications where the number of periods per cycle ν is large, it seems reasonable to employ discrete Fourier transform methods as done by *Anderson and Vecchia* [1993] to reduce the number of parameters. Their paper also includes a test for detecting hidden periodicities. Long-term variations in river flow, including El Niño effects, could be addressed by fitting a model where the period is more than 1 year.

Davis and Resnick [1985] show that if $\{\varepsilon_t\}$ belongs to the extremal domain of attraction of some type II max-stable law, then the stationary moving average process $X_t = \sum c_j \varepsilon_{t-j}$ also belongs to the same extremal domain of attraction (assuming that some mild conditions on the moving average parameters c_j are met). Their point process argument highlights the Poisson nature of the extreme order statistics. Since the largest observations in a heavy tail time series model tend to be widely spaced in time, they should resemble i.i.d. observations. Establishing the appropriate asymptotic theory for nonstationary time series models remains an open question. For the hydrologic data we examined, the model (29) seems to provide a reasonable fit.

Acknowledgment. The authors would like to thank John Nolan, Department of Mathematics and Statistics, American University, for providing the FORTRAN code to simulate stable random variables.

References

Adams, G., and C. Goodwin, Parameter estimation for periodic ARMA models, *J. Time Ser. Anal*., *16*, 127–145, 1995.

Anderson, P., and M. Meerschaert, Periodic moving averages or random variables with regularly varying tails, *Ann. Stat*., *25*, 771–785, 1997.

Anderson, P., and A. Vecchia, Asymptotic results for periodic autoregressive moving-average processes, *J. Time Ser. Anal*., *14*, 1–18, 1993.

Bhansali, R., Estimation of the impulse response coefficients of a linear process with infinite variance, *J. Multivariate Anal*., *45*, 274–290, 1993.

Brockwell, P., and R. Davis, *Time Series: Theory and Methods*, 2nd ed., Springer-Verlag, New York, 1991.

Chambers, J., C. Mallows, and B. Stuck, A method for simulating stable random variables, *JASA J. Am. Stat. Assoc*., *71*, 340–344, 1976.

Davis, R., and S. Resnick, Limit theory for moving averages of random variables with regularly varying tail probabilities, *Ann. Probab*., *13*, 179–195, 1985.

Fama, E., The behavior of stock market prices, *J. Bus*., *38*, 34–105, 1965.

Feller, W., *An Introduction to Probability Theory and Its Applications*, vol. 2, 2nd ed., John Wiley, New York, 1971.

Gardner, W., and C. Spooner, The cumulant theory of cyclostationary time-series, I, Foundation, *IEEE Trans. Signal Proc*., *42*, 3387–3408, 1994.

Hall, P., On some simple estimates of an exponent of regular variation, *J. R. Stat. Soc*., *Ser. B*, *44*, 37–42, 1982.

Hill, B., A simple general approach to inference about the tail of a distribution, *Ann. Stat*., 1163–1173, 1975.

Hosking, J. R. M., and J. R. Wallis, Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, *29*, 339–349, 1987.

Janicki, A., and A. Weron, *Simulation and Chaotic Behavior of α Stable Stochastic Processes*, Marcel Dekker, New York, 1994.

Jansen, D., and C. de Vries, On the frequency of large stock market returns: Putting booms and busts into perspective, *Rev. Econ. Stat*., *23*, 18–24, 1991.

Kokoszka, P., Prediction of infinite variance fractional ARIMA, *Probab. Math. Stat*., *16*, 65–83, 1996.

Kokoszka, P., and M. Taqqu, Infinite variance stable ARMA processes, *J. Time Series Anal*., *115*, 203–220, 1994.

Kokoszka, P., and M. Taqqu, Parameter estimation for infinite variance fractional ARIMA, *Ann. Stat*., *24*, 1880–1913, 1996.

Leadbetter, M., G. Lindgren, and H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York, 1980.

Loretan, M., and P. Phillips, Testing the covariance stationarity of heavy-tailed time series, *J. Empirical Finan*., *1*, 211–248, 1994.

Mandelbrot, B., The variation of certain speculative prices, *J. Bus*., *36*, 394–419, 1963.

McCulloch, J. H., Measuring tail thickness in order to estimate the stable index α: A critique, *J. Bus. Econ. Stat*., *15*, 74–81, 1995.

McCulloch, J. H., and D. B. Panton, Precise tabulation of the maximally-skewed stable distributions and densities, *Comput. Stat. Data Anal*., *23*, 307–320, 1996.

Meerschaert, M., and H. P. Scheffler, A simple robust estimator for the thickness of heavy tails, *J. Stat. Plann. Inference*, in press, 1998.

Mikosch, T., T. Gadrich, C. Klüppenberg, and R. Adler, Parameter estimation for ARMA models with infinite variance innovations, *Ann. Stat*., *23*, 305–326, 1995.

Mittnik, S., and S. Rachev, *Modelling Financial Assets with Alternative Stable Models*, John Wiley, New York, 1995.

Nikias, C., and M. Shao, *Signal Processing With Alpha Stable Distributions and Applications*, John Wiley, New York, 1995.

Resnick, S., and C. Stărică, Consistency of Hill's estimator for dependent data, *J. Appl. Probab.*, *32*, 139–167, 1995.

Samorodnitsky, G., and M. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models With Infinite Variance*, Chapman and Hall, New York, 1994.

Smith, R. L., Threshold methods for sample extremes: Statistical extremes and applications (Vimeiro, 1983), *NATO ASI Ser. Ser. C*, D. Reidel, Norwell, Mass., *131*, 621–638, 1984.

Tiao, G., and M. Grupe, Hidden periodic autoregressive-moving average models in time series data, *Biometrika*, *67*, 365–373, 1980.

Tjostheim, D., and J. Paulsen, Empirical identification of multiple time series, *J. Time Series Anal.*, *3*, 265–282, 1982.

Ula, T., Forecasting of multivariate periodic autoregressive moving-average processes, *J. Time Series Anal.*, *14*, 645–657, 1993.

Weron, R., On the Chambers-Mallows-Stuck method for simulating skewed stable random variables, *Stat. Probab. Lett.*, *28*, 165–171, 1996.

P. L. Anderson, Department of Mathematics, Albion College, Albion, MI 49224. (e-mail: anderson@knot.math.unr.edu)

M. M. Meerschaert, Department of Mathematics, University of Nevada, Business Building, Room 601, Reno, Nevada 89557-0045. (e-mail: mcubed@unr.edu)

(Received August 4, 1997; revised March 31, 1998;
accepted April 28, 1998.)