Cross correlations among estimators of shape

Eduardo S. Martins

Fundação Cearense de Meteorologia e Recursos Hídricos (FUNCEME), Governo do Estado do Ceara, Fortaleza, Brazil

Jery R. Stedinger

School of Civil and Environmental Engineering, Cornell University, Ithaca, New York, USA

Received 16 July 2002; revised 27 July 2002; accepted 27 July 2002; published 26 November 2002.

[1] The regional variability of shape parameters (such as κ for the GEV distribution) may be described by generalized least squares (GLS) regression models that allow shape parameters to be estimated from basin characteristics recognizing the sampling uncertainty in available shape estimators. Implementation of such GLS models requires estimates of the cross-site correlation of the shape parameter estimators for every pair of sites. Monte Carlo experiments provided the information needed to identify simple power approximations of the relationships between the cross correlation of estimators of skewness γ from [log] Pearson type 3 (P3) data and of the shape parameter κ of both generalized Pareto (GP) and generalized extreme value (GEV) distributions, as functions of the intersite correlation of concurrent flows. *INDEX TERMS:* 1821 Hydrology: Floods; *KEYWORDS:* skewness coefficient, log Pearson type 3, generalized extreme value distribution, generalized least squares regression

Citation: Martins, E. S., and J. R. Stedinger, Cross correlations among estimators of shape, *Water Resour. Res.*, *38*(11), 1252, doi:10.1029/2002WR001589, 2002.

1. Introduction

[2] An important problem in flood frequency analysis is the estimation of flood quantiles and other statistics of annual flood peaks at ungaged locations, and at gaged sites with short records. Regional flood and rainfall frequency analysis is an efficient approach for improving quantile estimators at sites with short records, as well as providing estimates at ungaged locations.

[3] *Interagency Advisory Committee on Water Data* (*IACWD*) [1982], *Fiorentino et al.* [1987], *Gabrielle and Arnell* [1991], *Stedinger and Lu* [1995], *Madsen and Rosbjerg* [1997a, 1997b], and *Hosking and Wallis* [1997] advocate the use of hierarchical approaches in flood frequency analysis. For example, *IACWD* [1982] recommends use of a weighted combination of regional and at-site skewness coefficient estimators, while the mean and variance are estimated using at-site data. *Martins and Stedinger* [2000] demonstrate the value of using regional information that reduces the uncertainty in the estimated shape parameter κ of a GEV distribution. Regional regression procedures can estimate such shape parameters by relating them to physiographic characteristics, such as drainage area, main channel slope, and mean elevation.

[4] The regional variability of shape parameters can be described by use of generalized least squares (GLS) regression models that include basin characteristics as explanatory variables. *Stedinger and Tasker* [1985, 1986], *Tasker and Stedinger* [1987], *Tasker et al.* [1996], *Moss and Tasker* [1991], and *Kroll and Stedinger* [1998] report Monte Carlo experiments documenting the value of generalized least squares (GLS) procedures for estimating empirical relationships between streamflow statistics and physiographic basin characteristics. *Tasker and Driver* [1988], *Tasker and Stedinger* [1986, 1989], and *Madsen and Rosbjerg* [1997b] illustrate the use of such methods.

[5] The GLS algorithm accounts for the heterogeneity of variances and cross correlations of the residuals of a regression model. It can provide more accurate estimators in terms of mean square error (MSE) than ordinary and weighted least squares (OLS and WLS) algorithms [*Stedinger and Tasker*, 1985], and a nearly unbiased estimator of the model error variance [*Stedinger and Tasker*, 1986]. In the context of the index flood method, *Hosking and Wallis* [1988, 1997] and *Madsen and Rosbjerg* [1997a, 1997b] demonstrate that cross correlation among concurrent flows does affect the precision of quantile estimators.

[6] Application of GLS requires estimation of a sampling error covariance matrix. In general, the at-site variance of the shape parameter estimators can be obtained analytically [*Madsen and Rosbjerg*, 1997a; *Tasker and Stedinger*, 1986]. However, first-order asymptotic expressions for cross-site correlations among shape estimators are complicated and may be inaccurate [*Stedinger and Lu*, 1995]. For multivariate normal data, the effect of intersite dependence on the pth central moment estimator can be expressed in terms of the cross-site correlation of the observations

$$\rho\{\hat{\theta}_i;\hat{\theta}_j\} = \rho_{i,j}^{P} \qquad (1)$$

where $\rho\{\hat{\theta}_i;\hat{\theta}_j\}$ is the cross correlation between the pth central moment estimators at sites i and j; $\hat{\theta}_i$ is the pth central moment estimator at site i, and $\rho_{i,j}^{P}$ is the pth power of the cross-site correlation between concurrent flows of stations i and j [*Stedinger*, 1983; *Hosking and Wallis*, 1988; *Gabrielle and Arnell*, 1991; *Madsen*, 1996]. Adopting

0043-1397/02/2002WR001589

equation (1), *Madsen and Rosbjerg* [1997b] assumed that $\rho\{\hat{\kappa}_i, \hat{\kappa}_j\} \cong \rho_{1,j}^2$, where $\hat{\kappa}_i$ is the L moment estimator of the shape parameter κ of the Generalized Pareto (GP) distribution for station i, and $\rho_{i,j}$ is the cross-site correlation between concurrent flows at stations i and j. This is consistent with the fact that the shape parameter of the GP distribution determines the second moment of the GP distribution (as well as higher moments), and the estimator of the GP shape parameter primarily depends the estimator of τ_2, the L moment coefficient of variation (see equation A3).

[7] This paper uses Monte Carlo simulation to derive an approximation of the relationship between the cross-site correlation among concurrent flows and the correlation of estimators of the skewness coefficient γ based upon [log] Pearson type 3 (P3) data, and of the shape parameter κ of both GEV and GP models. P3 data can correspond to the logarithms of log-Pearson type III data [*IACWD*, 1982]. The regional variability of such shape parameters can then be described by a GLS regression model that allows the shape parameters to be estimated from basin characteristics, while accounting for the heterogeneity among their variances and the cross correlations among shape parameter estimators.

2. Monte Carlo Experiments

[8] Monte Carlo experiments provided estimates of $\rho(\hat{\kappa}_x, \hat{\kappa}_y)$ for both GEV and GP distributions, and $\rho(\hat{\gamma}_x, \hat{\gamma}_y)$ for P3 data, given the observed cross correlation between the concurrent observations at any two sites $\hat{\rho}_{x,y}$. These experiments were performed with simulated pairs of annual flood or precipitation series (x, y) with the sample sizes in Table 1, and intersite correlation coefficients for concurrent values in the range $-0.75 \leq \rho_{x,y} \leq 0.95$; this range surely exceeds the range of practical interest in hydrology given that negative cross correlations are unlikely in practice, though the results with negative correlations are of intellectual curiosity and may have applications in other fields. The x series has a total record length of $n_{xy} + n_x$, and y series a record length of $n_{xy} + n_y$. Here n_{xy} is the record length for the common period.

[9] Three experiments were performed employing, respectively, P3, GEV and GP distributions. The shape parameters for these experiments were in the ranges: $-0.3 \leq \kappa \leq +0.1$ for both GEV and GP experiments, and $0 \leq \gamma \leq +1$ for the P3 experiment. (If $X \sim$ P3 with skewness coefficient γ, then $-X \sim$ P3 has skew $-\gamma$; thus the correlation of skew estimators for $\gamma_x = \gamma_y < 0$ is the same positive value as obtained for $\gamma_x = \gamma_y > 0$.)

[10] The experiments had four fundamental steps.

1. The first step was to generate bivariate normal deviates z_x and z_y with cross correlation $\rho_{x,y}$, and sample sizes $n_{xy} + n_x$ and $n_{xy} + n_y$, respectively.

2. These normal deviates were transformed to have the desired marginal distribution, as in the work of *Hosking and Wallis* [1988].

3. The sample intersite correlation coefficient $\hat{\rho}_{x,y}$ between the x and y series was computed.

4. For the GEV and GP experiments, L moment estimators of κ_x and κ_y (see Appendix A, equation A1 and A3) were computed using the entire x and y series, respectively: $(\hat{\kappa}_x, \hat{\kappa}_y)$. For the P3 experiment, estimators of γ_x and γ_y (see Appendix A, equation A4) were computed using the entire x and y series, respectively: $(\hat{\gamma}_x, \hat{\gamma}_y)$.

Table 1. Sample Sizes of the Series in the Monte Carlo Experiments Used to Verify the Effect of Sample Sizes Differences on the Relationships Between $\hat{\rho}_{x,y}$ and $\hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)$ for Both GEV and GP Data and Between $\hat{\rho}_{x,y}$ and $\hat{\rho}(\hat{\gamma}_x, \hat{\gamma}_y)$ for P3 Data

Series	n_x	n_{xy}	n_y
1	0	40	0
2	20	40	0
3	40	40	0
4	80	40	0
5	20	40	20
6	40	40	40
7	80	40	80

Steps (i)–(iv) were repeated s=10,000 times to compute the mean $\bar{\rho}_{x,y}$ of $\hat{\rho}_{x,y}$ and the needed intersite correlation coefficients for the shape parameters using

$$\hat{\rho}_{\hat{\theta}_x,\hat{\theta}_y} = \frac{\sum_{i=1}^{s}\left(\left(\hat{\theta}_i^x - \bar{\theta}^x\right)\left(\hat{\theta}_i^y - \bar{\theta}^y\right)\right)}{\left(\sum_{i=1}^{s}\left(\hat{\theta}_i^x - \bar{\theta}^x\right)^2\right)\left(\sum_{i=1}^{s}\left(\hat{\theta}_i^y - \bar{\theta}^y\right)^2\right)} \tag{2}$$

where $\bar{\theta}^x = \sum_{i=1}^{s} \frac{\hat{\theta}_i^x}{s}$ and $\bar{\theta}^y = \sum_{i=1}^{s} \frac{\hat{\theta}_i^y}{s}$. Here θ represents the shape parameter of interest.

[11] Caution should be exercised before applying these results to estimators of shape from a GP/Poisson partial duration series (PDS) analysis [*Stedinger et al.*, 1993; *Martins and Stedinger*, 2001] wherein peaks-over-threshold series are developed for two sites. In that framework, particularly if the cross correlation is small, or thresholds with very different exceedance probabilities, one will not always record concurrent peaks greater than the threshold at both sites. This could result in a different cross correlation between κ estimators. In that situation, the cross correlation between the estimators may be described by the adjustment employing the total number of observations used to compute each estimator with a concurrent length corresponding to the number of events that resulted in floods peaks at both sites. The Monte Carlo simulations reported here use GP pairs (x_i, y_i), which for the concurrent period of record always contain both x and y values so there is no issue whether or not both values are observed, as there could be with two PDS.

[12] The next step was to derive a relationships between $\hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)$ for both GEV and GP models, or $\hat{\rho}(\hat{\gamma}_x, \hat{\gamma}_y)$ for P3 data, and the observed intersite correlation of concurrent values $\hat{\rho}_{x,y}$. For $\rho_{xy} = 0$ (independent series), the correlation between the shape parameters must be zero. Similarly, when $\rho_{xy} = 1$, one should see a perfect correlation between the shape parameters. When the shape parameter is controlled by the third moment, negative values of ρ_{xy} will result in negative values for the cross-site correlations: $\hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)$ for the GEV model and $\hat{\rho}(\hat{\gamma}_x, \hat{\gamma}_y)$ for the P3 distribution. These arguments and equation (1) motivate use of the simple power function

$$\hat{\rho}_{\hat{\theta}_x,\hat{\theta}_y} = \mathrm{Sign}\left(\hat{\rho}_{x,y}\right) cf_{xy} \left|\hat{\rho}_{x,y}\right|^{\delta} \quad \text{with} \quad cf_{xy} = \frac{n_{xy}}{\sqrt{\left(n_{xy} + n_x\right)\left(n_{xy} + n_y\right)}} \tag{3}$$

though the same exponent δ need not provide the best fit to both positive and negative correlations with the asymmetric

Table 2. Estimated δ Values for the Monte Carlo Experiments (Series 1) P3 Distribution

γ for P3	δ	R^2
$\gamma_x = \gamma_y = 0.0$	2.8	0.998
$\gamma_x = \gamma_y = 0.2$	2.9	0.998
$\gamma_x = \gamma_y = 0.4$	3.0	0.998
$\gamma_x = \gamma_y = 0.6$	3.0	0.995
$\gamma_x = \gamma_y = 0.8$	3.2	0.991
$\gamma_x = \gamma_y = 1.0$	3.3	0.988
$\gamma_x = 0.0, \gamma_y = 1.0$	3.3	0.999

GEV distribution, as it should for the P3 distribution for which negative skews $-\gamma$ are obtained by using the negative of a P3 variate with positive skew γ.

[13] This power function was fit to the $\{\hat{\rho}_{x,y}, \hat{\rho}(\hat{\theta}_x, \hat{\theta}_y)\}$ pairs. Sign $(\hat{\rho}_{x,y})$ is plus or minus one depending upon the sign of $\hat{\rho}_{x,y}$. The only parameter to be identified, δ, was obtained by minimizing the mean square error between observed and fitted values. The cf_{xy} term in equation (3) accounts for the sample size differences between the two series, following *Stedinger and Tasker* [1985, equation 4]. The validity of this function will be demonstrated.

[14] When the shape parameter is controlled by the second moment, as it is with the GP model, negative correlation among flows ($\rho_{xy} < 0$) generated positive cross-site correlation $\hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)$ among the κ estimators. With normal flows such behavior is also observed between variance estimators, where $\hat{\rho}\{\hat{\sigma}_x^2; \hat{\sigma}_y^2\} = \rho_{xy}^2$, as follows from equation (1). For the shape parameter of the GP distribution, the model in equation (3) without the term Sign $(\hat{\rho}_{x,y})$ was fit to the $\{\hat{\rho}_{x,y}, \hat{\rho}(\hat{\theta}_x, \hat{\theta}_y)\}$ pairs when $\hat{\rho}_{x,y}$ is negative and also to the pairs when $\hat{\rho}_{x,y}$ is positive, to identify two exponents δ^- and δ^+, respectively.

3. Results

[15] The model in equation (3) was fit to $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\gamma}_x, \hat{\gamma}_y)\}$ pairs generated in the P3 Monte Carlo experiments, where $\bar{\rho}_{x,y}$ is the average of the 10,000 sample correlation. Table 2 reports the R^2 and δ values identified. Generally, $2.8 \leq \delta \leq 3.3$, depending upon the skewness coefficient ($0 \leq \gamma \leq +1$). Figure 1 displays the data and the fitted power function for $\gamma_x = \gamma_y = 0.2$, and $\gamma_x = 0$ with $\gamma_y = +1$. For the normal model (obtained with $\gamma = 0$), the cross correlation among third moments should depend upon ρ_{xy}^3, whereas the variance should depend upon ρ_{xy}^2 (see equation 1). Because the skewness coefficient depends upon sample estimators of both of these statistics, one would expect δ to fall between 2 and 3 when $\gamma = 0$, as it did. For nonnormal distributions with $\gamma > 0.6$, $\delta > 3$ indicating that the cross correlations among the skewness estimators are smaller than for normal data (because ρ^δ decreases with increasing δ for $0 < \rho < 1$). The $\hat{\rho}(\hat{\gamma}_x, \hat{\gamma}_y)$ values for series 2–7, corresponding to different combinations of nonconcurrent and concurrent records as indicated in Table 1, were divided by cf_{xy} to check how close the corrected values are to the series 1 results. As shown in Figures 1a and 1b, the coefficient cf_{xy} works reasonably well for skewness coefficients in the range $-1 \leq \gamma \leq +1$ when the skewness coefficients are equal, and with different skewness coefficients for the x and y series, as is the case when $\gamma_x = 0$ with $\gamma_y = +1$.

[16] The power function in equation (3) was also fit to $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)\}$. Table 3 reports the R^2 and δ values identified. Figures 2a and 2b shows the $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)\}$ pairs obtained with the GEV experiments for $\kappa_x = \kappa_y = -0.3$, and $\kappa_x = -0.3$, $\kappa_y = +0.1$, respectively. The results from the κ/GEV experiments show that the exponent for intersite dependence can be larger or smaller than the value of 3 for normal variates. (See equation 1.) Generally, $2.4 \leq \delta \leq 3.1$, depending upon the shape parameter κ ($-0.30 \leq \kappa \leq 0.10$). The values of $\hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)$ for series 2–7 were divided by cf_{xy} to test this correction. As one can see in Figures 2a and 2b, the corrected series 2–7 values are very close to those for series 1. This was observed for shape parameters in the range $-0.3 \leq \kappa \leq +0.1$, and also for cases with different shape parameters, including $\kappa_x = -0.3$, $\kappa_y = +0.1$.

[17] The power model without the term Sign $(\hat{\rho}_{x,y})$ was fit to $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)\}$ pairs generated in the GP Monte Carlo experiments. Table 3 reports R^2, δ^- and δ^+ values. Figures 3a and 3b shows the $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)\}$ pairs obtained with the GP experiments for $\kappa_x = \kappa_y = -0.3$, and $\kappa_x = -0.3$ with $\kappa_y = +0.1$, respectively. For positive values of the average cross-site correlation $\bar{\rho}_{x,y}$, the effect of intersite dependence is close to, but larger ($\delta < 2$) than those predicted by the value of 2 employed by *Madsen and Rosbjerg* [1997a]. Generally, $1.7 \leq \delta^+ \leq 1.9$, depending upon the shape parameter κ ($-0.30 \leq \kappa \leq 0.10$). For negative values of the average cross-site correlation $\bar{\rho}_{x,y}$ the effect of intersite dependence can be much less than that predicted by the value of 2 employed by *Madsen and Rosbjerg* [1997a]: generally, $2.2 \leq \delta^- \leq 5.8$, depending upon the shape parameter κ. One can see in Figures 3a and 3b that the corrected 2–7 series correlations yield values very close to those for series 1. This was observed for shape parameter in the range $-0.3 \leq \kappa \leq +0.1$, and also for cases with different shape parameters, including $\kappa_x = -0.3$ with $\kappa_y = +0.1$.

4. Conclusions

[18] Monte Carlo experiments provided the information needed to develop simple power approximations of the relationships between the cross correlation of estimators of the skewness coefficient γ from P3 data, and the shape parameter κ of both GEV and GP distributions, as a function of the cross correlation of concurrent values. These relationships can be used with the regional average value of the skewness coefficient γ, or of κ, to calculate the cross correlation among shape estimators for different sites. A factor (cf_{xy}) accounts for the lengths of concurrent and nonconcurrent records for any two sites: it works reasonably well for the shape parameters in the range considered ($-0.30 \leq \kappa \leq 0.10$; $-1 \leq \gamma \leq 1$).

[19] For the P3 data, the exponent δ is in the range $2.8 \leq \delta \leq 3.3$. For positive cross-site correlation between concurrent flows, the effect of positive cross correlation among concurrent flows on estimators of the shape parameter of GEV distribution ($2.4 \leq \delta \leq 3.1$) is about the same as observed for the skewness coefficient for the P3 distribution, and less than when estimating the shape parameter of the GP distribution ($1.7 \leq \delta^+ \leq 1.9$). Caution should be

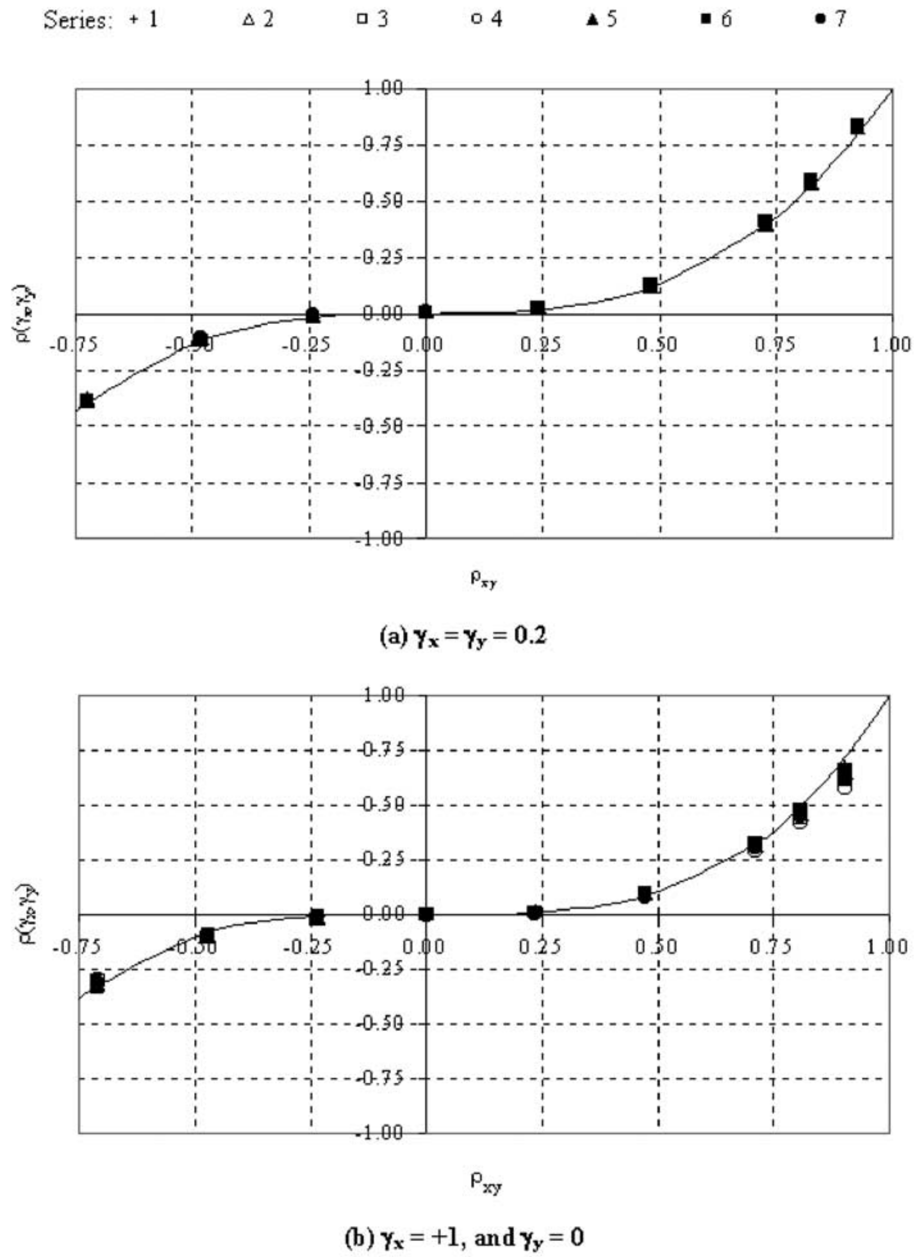

Figure 1. Monte Carlo results for the P3 experiments: $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\gamma}_x, \hat{\gamma}_y)\}$ pairs when (a) $\gamma_x = \gamma_y = 0.2$; (b) $\gamma_x = +1$ and $\gamma_y = 0$. Model fitted to series 1. Series 1–7 refers to different combinations of sample sizes for series x and y (see Table 1).

exercised before applying these results to models of partial duration series because any two series may not correspond to sets of concurrent events: some events may show up in one series but not the other. However, the cf_{xy} adjustment can be employed to represent the fraction of the events that were observed at both sites.

[20] Use of the relationships derived here allows construction of generalized least squares (GLS) regression models that correctly incorporate the sampling error and cross correlation among GEV or LP3 shape parameter estimators. Such models should provide better estimators of the shape parameter of a GEV or LP3 distribution for a

Table 3. Estimated δ Values for the Monte Carlo Experiments (Series 1) GEV and GP Distributions

	GEV		GP	
κ	δ	R^2	δ^-/δ^+	R^2
$\kappa_x = \kappa_y = -0.3$	2.4	0.997	2.2/1.8	0.999
$\kappa_x = \kappa_y = -0.2$	2.6	0.998	2.5/1.8	0.999
$\kappa_x = \kappa_y = -0.1$	2.8	0.998	3.0/1.8	0.999
$\kappa_x = \kappa_y = 0.0$	2.9	0.998	4.0/1.8	0.999
$\kappa_x = \kappa_y = 0.1$	3.1	0.998	5.8/1.7	0.999
$\kappa_x = -0.3$, $\kappa_y = +0.1$	2.6	0.996	2.7/1.9	0.999

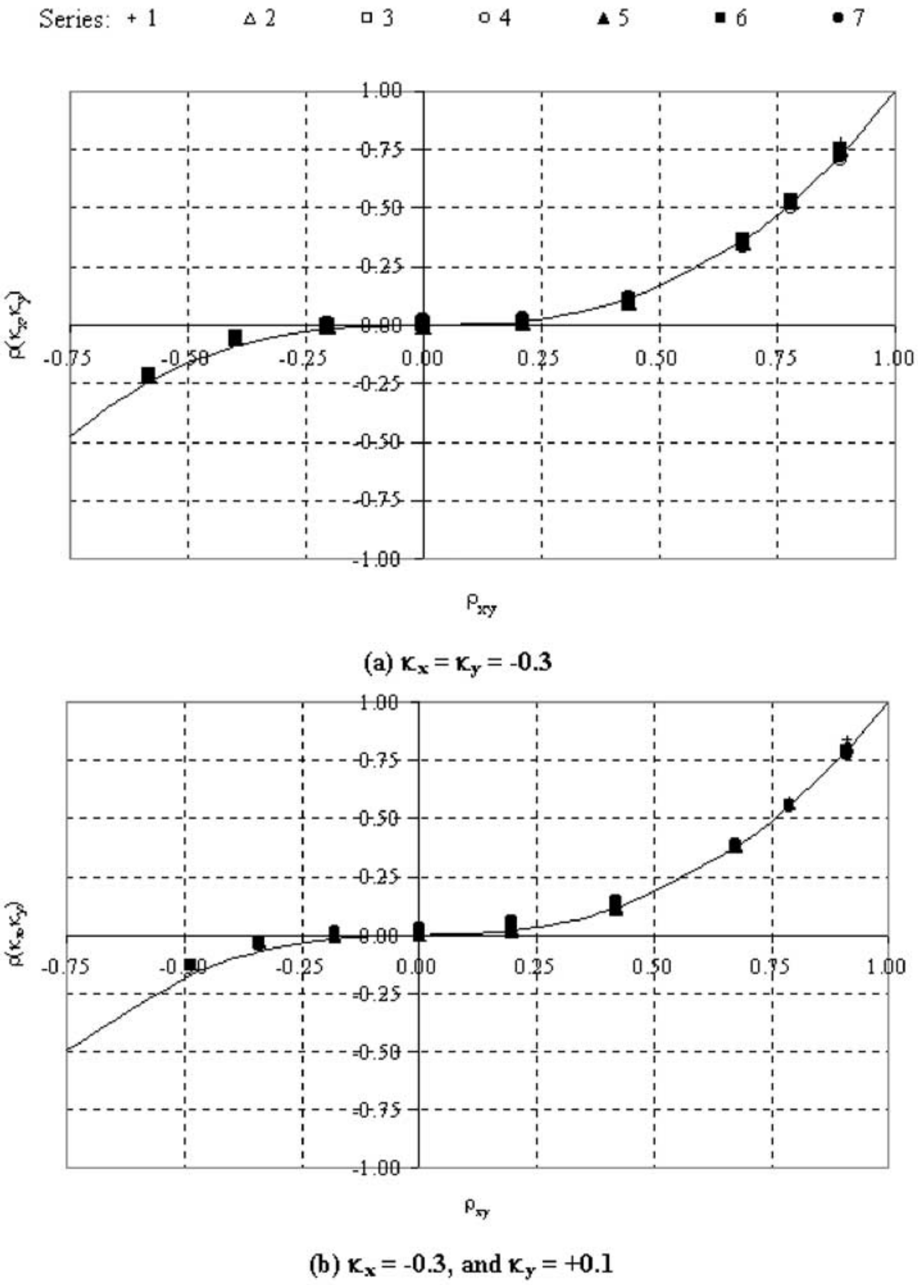

Figure 2. Monte Carlo results for the GEV experiments: $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)\}$ pairs when (a) $\kappa_x = \kappa_y = -0.3$; (b) $\kappa_x = -0.3$ and $\kappa_y = +0.1$. Model fitted to series 1. Series 1–7 refers to different combinations of sample sizes for series x and y (see Table 1).

gauged or ungauged site, as well as better and unbiased estimators of the precision of such regional shape estimators.

Appendix A: Shape Parameters Estimators

A1. GEV and GP Shape Parameters

A1.1. $\hat{\kappa}$/GEV [L Moment]

[21] The L moment estimator for the shape parameter of the GEV distribution [*Hosking et al.*, 1985] was obtained using

$$\hat{\kappa} = 7.8590\ c + 2.9554\ c^2, \tag{A1}$$

with $c = 2/(3 + \hat{\tau}_3) - \log(2)/\log(3)$. Here the final $\hat{\kappa}$ function is a very good approximation for $\hat{\kappa}$ in the range (−0.5, 0.5) [*Hosking et al.*, 1985]. The L moment estimators $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\tau}_2 = \hat{\lambda}_2/\hat{\lambda}_1$ (L-CV) and $\hat{\tau}_3 = \hat{\lambda}_3/\hat{\lambda}_2$ (L skewness) were obtained by using the unbiased estimator of the first three PWMs β_r [*Landwehr et al.*, 1979; *Hosking and Wallis*, 1995]:

$$b_r = \sum_{i=1}^{n}\left[\frac{(i-1)(i-2)(i-3)\ldots(i-r)}{n(n-1)(n-2)\ldots(n-r)} x_{(i)}\right] \qquad r = 0, 1, 2, \ldots \tag{A2}$$

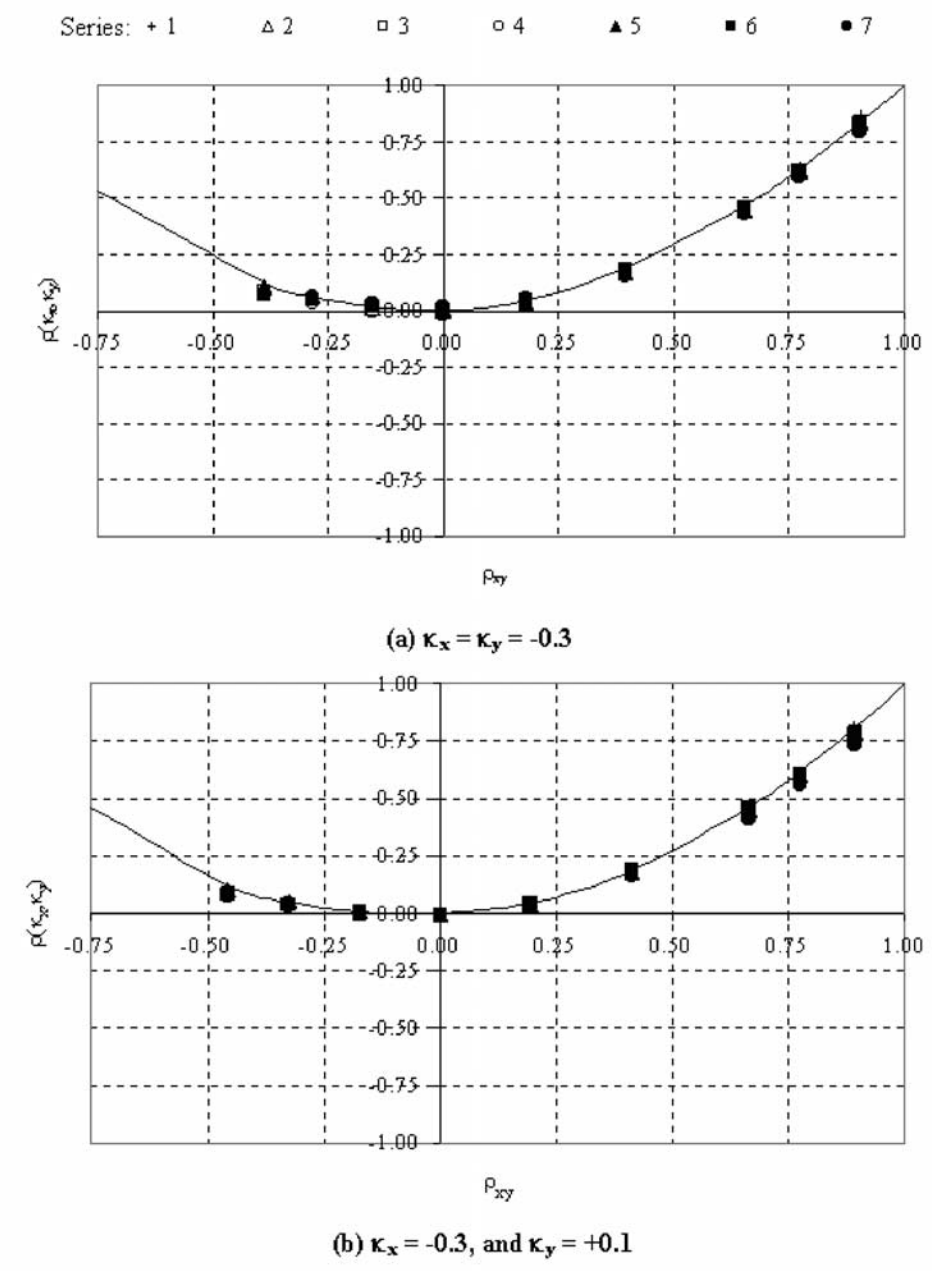

Figure 3. Monte Carlo results for the GP experiments: $\{\bar{\rho}_{x,y}, \hat{\rho}(\hat{\kappa}_x, \hat{\kappa}_y)\}$ pairs when (a) $\kappa_x = \kappa_y = -0.3$; (b) $\kappa_x = -0.3$ and $\kappa_y = +0.1$. Model fitted to series 1. Series 1–7 refers to different combinations of sample sizes for series x and y (see Table 1).

where the $x_{(i)}$ are the ordered observations from a sample of size n $\{x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}\}$, and where $\lambda_1 = \beta_0$, $\lambda_2 = 2\beta_1 - \beta_0$, and $\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$ [*Hosking*,1990; see also *Wang*, 1996].

A1.2. $\hat{\kappa}$/GP [L Moment]

[22] The L moment estimator for the GP distribution [*Hosking and Wallis*, 1987; *Madsen and Rosbjerg*, 1997a] for the shape parameter is

$$\hat{\kappa} = \frac{1}{\hat{\tau}_2} - 2 \quad \text{(A3)}$$

where the L moment estimators $\hat{\lambda}_1, \hat{\lambda}_2$, and $\hat{\tau}_2 = \hat{\lambda}_2 / \hat{\lambda}_1$ (L-CV) are obtained by using the unbiased estimator for the first two PWMs given in equation (A2).

A2. Skewness

[23] A nearly unbiased estimate of the skewness used here is

$$\hat{\gamma} = \frac{n}{(n-1).(n-2)s^3} \sum_{i=1}^{n} (x_i - x)^3 \quad \text{(A4)}$$

where $\bar{x}$ is the mean and s is the standard deviation of the series $\{x_i\}$, which can represent either the flows or their logarithms depending upon the domain of interest. Use of a different function of n (as done by *Tasker and Stedinger* [1986]), would not affect the correlation between skewness estimators.

[24] **Acknowledgments.** We thank the Brazilian Government agency CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) and the Ceara State Government (Brazil) who provided support for much of this research. We also thank Charles Kroll, Gary Tasker, and an anonymous referee who provided reviews and useful comments.

References

Fiorentino, M., S. Gabrielle, F. Rossi, and P. Versace, Hierarchical approach for regional flood frequency analysis, in *Regional Flood Frequency Analysis*, edited by V. P. Singh, pp. 35–49, D. Reidel, Norwell, Mass., 1987.

Gabrielle, S., and N. Arnell, A hierarchical approach to regional flood frequency analysis, *Water Resour. Res.*, *27*(6), 1281–1289, 1991.

Hosking, J. R. M., L-moments: Analysis and estimation of distributions using linear combinations of order statistics, *J. R. Stat. Soc., Ser. B*, *52*(2), 105–124, 1990.

Hosking, J. R. M., and J. R. Wallis, Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, *29*(3), 339–349, 1987.

Hosking, J. R. M., and J. R. Wallis, The effect of intersite dependence on regional frequency analysis, *Water Resour. Res.*, *24*(1), 588–600, 1988.

Hosking, J. R. M., and J. R. Wallis, A comparison of unbiased and plotting position estimators of L-moments, *Water Resour. Res.*, *31*(8), 2019–2025, 1995.

Hosking, J. R. M., and J. R. Wallis, *Regional Frequency Analysis: An Approach Based on L-moments*, Cambridge Univ. Press, New York, 1997.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood, Estimation of the generalized extreme-value distribution by the method of probability weighted moments, *Technometrics*, *27*(3), 251–261, 1985.

Interagency Advisory Committee on Water Data (IACWD), Guidelines for determining flood flow frequency, *Bull. 17B*, Off. of Water Data Coordination, U.S. Geol. Surv., Reston, Va., 1982.

Kroll, C. N., and J. R. Stedinger, Regional hydrologic analysis: Ordinary and generalized least squares revisited, *Water Resour. Res.*, *34*(1), 121–128, 1998.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis, Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, *Water Resour. Res.*, *15*(5), 1055–1064, 1979.

Madsen, H., At-site and regional modelling of extreme hydrologic events, Ph.D. thesis, 188 pp., Tech. Univ. of Denmark, Lyngby, Feb. 1996.

Madsen, H., and D. Rosbjerg, The partial duration series method in regional index-flood modeling, *Water Resour. Res.*, *33*(4), 737–746, 1997a.

Madsen, H., and D. Rosbjerg, Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling, *Water Resour. Res.*, *33*(4), 771–782, 1997b.

Martins, E. S., and J. R. Stedinger, Generalized maximum likelihood GEV quantile estimators for hydrologic data, *Water Resour. Res.*, *36*(3), 737–744, 2000.

Martins, E. S., and J. R. Stedinger, Generalized maximum likelihood Pareto-Poisson flood risk analysis for partial duration series, *Water Resour. Res.*, *37*(10), 2559–2567, 2001.

Moss, M. E., and G. D. Tasker, An intercomparison of hydrological network-design technologies, *J. Hydrol. Sci.*, *36*(3), 209–221, 1991.

Stedinger, J. R., Estimating a regional flood frequency distribution, *Water Res. Res.*, *19*(2), 503–510, 1983.

Stedinger, J. R., and L.-H. Lu, Appraisal of regional and index flood quantile estimators, *Stochastic Hydrol. Hydraul.*, *9*(1), 49–75, 1995.

Stedinger, J. R., and G. Tasker, Regional hydrologic analysis, 1, Ordinary, weighted and generalized least squares compared, *Water Res. Res.*, *21*(9), 1421–1432, 1985. (Correction, *Water Resour. Res.*, *22*(5), 844, 1986.)

Stedinger, J. R., and G. Tasker, Regional hydrologic analysis, 2, Model-error estimators, estimation of sigma and log-Pearson type 3 distributions, *Water Resour. Res.*, *22*(10), 1487–1499, 1986.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, Frequency analysis of extreme events, in *Handbook of Hydrology*, edited by D. Maidment, chap. 18, McGraw-Hill, New York, 1993.

Tasker, G. D., and N. E. Driver, Nationwide regression models for predicting urban runoff water quality at unmonitored sites, *Water Resour. Bull.*, *24*(5), 1091–1101, 1988.

Tasker, G., and J. R. Stedinger, Estimating generalized skew with weighted least squares regression, *J. Water Resour. Plann. Manage.*, *112*(2), 225–237, 1986.

Tasker, G., and J. R. Stedinger, Regional regression of flood characteristics employing historical information, *J. Hydrol.*, *96*, 255–264, 1987.

Tasker, G. D., and J. R. Stedinger, An operational GLS model for hydrologic regression, *J. Hydrol.*, *111*(1–4), 361–375, 1989.

Tasker, G., S. A. Hodge, and C. S. Barks, Region of influence regression for estimating the 50-year flood at ungaged sites, *Water Resour. Bull.*, *32*(1), 163–170, 1996.

Wang, Q. J., Direct sample estimators of L moments, *Water Resour. Res.*, *32*(12), 3617–3619, 1996.

E. S. Martins, Fundação Cearense de Meteorologia e Recursos Hídricos (FUNCEME), Governo do Estado do Ceara, Fortaleza CE 60115-121, Brazil.

J. R. Stedinger, School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853-3501, USA. (jrs5@cornell.edu)