AJUSTE DE VAZÕES VIA DISTRIBUIÇÃO GEV E MOMENTOS LH EM MATLAB

Manoel Moisés F. de Queiroz[1] & Fazal Hussain Chaudhry[2]

Resumo - A distribuição de probabilidade generalizada de valores extremos (GEV), introduzida por Jenkinson (1955), tem encontrado muitas aplicações em hidrologia. Ela tem sido utilizada para modelar uma extensa variedade de extremos naturais, incluindo cheias, chuvas, velocidade do vento, temperaturas e outros máximos. Estudos têm mostrado que estimadores de máxima verossimilhança dos parâmetros da GEV em pequenas amostras são instáveis, podendo fornecer valores absurdos do parâmetro de forma, onde são recomendados estimadores de momentos LH, uma generalização de momentos L, que são baseados na combinação linear de estatísticas de altas ordens, introduzidas para caracterizar a parte mais alta da distribuição e os valores extremos dos dados. Contudo, não existem programas computacionais em ambiente Matlab para PC, que modele eventos extremos via momentos LH. O principal propósito deste trabalho é apresentar uma rotina em Matlab que provê a modelação de eventos extremos através da GEV, utilizando momentos LH e o teste de qualidade de ajuste proposto por Wang (1998). Como resultado são apresentados os 3 parâmetros estimados da GEV, os valores das taxas de momentos LH: coeficiente de variação, assimetria, curtose e os valores do teste de qualidade de ajuste.

Abstract - The generalized extreme-value (GEV) distribution, introduced by Jenkinson (1955), has found many applications in hydrology. It has been used to model a wide variety of natural extremes, including floods, rainfall, wind speeds, temperature, and other maxims. Previous studies show that small-sample maximum-likelihood estimator parameters are unstable and demonstrates that absurd values of the GEV shape parameter can be generated. It is recommended that LH moment estimators, a generalization of L moments, that are based on linear combinations of higher-order statistics, should be introduced for characterizing the upper part of distributions and larger events in data. However, there are no computer packages available in Matlab for PC that model extreme events by LH moments. The main purpose of this paper is to present a routine in Matlab to model extreme events by GEV using LH moments and goodness-of-fit test proposed by Wang (1998). The

[1] Prof. Adjunto da Área de Recursos Hídricos e Saneamento Ambiental, CCET- UNIOESTE, Rua Universitária, 2069, Jardim Universitário, Fone (45)220-3262, Fax (45)324-4566, CEP 85819-110, Cascavel-PR; mfqueiroz@unioeste.br
[2] Prof. Titular do Departamento de Hidráulica e Saneamento - EESC/USP, CxP 359, Fone (16)274-9266, Fax (16)274-9212; fazal@sc.usp.br.

results are GEV-parameters, coefficient of variation, skewness, kurtosis and goodness-of-fit test values.

Palavras-chaves - análise de freqüência, momentos LH, distribuição GEV, eventos extremos

INTRODUÇÃO

A distribuição de probabilidade generalizada de valores extremos (GEV), introduzida por Jenkinson (1955), que combina os três possíveis tipos de distribuições de valores extremos em uma única forma, tem sido utilizada para representar a distribuição de valores extremos em diferentes campos, principalmente em análise de freqüência de cheias, com considerável aceitação para descrever fluxo máximo de cheias anuais. Na prática, a distribuição GEV tem sido usada para modelar uma extensa variedade de extremos naturais, incluindo, além de cheias, chuvas, velocidade do vento, temperaturas, picos de ondas do mar e outros máximos (Martins & Stedinger, 2000; Queiroz, 2002).

Procedimentos computacionais para estimação dos parâmetros da GEV, utilizando estimadores de máxima verossimilhança foram propostos por Otten & Van Montfort (1980), Prescott & Walden (1980) e Hosking (1985). Contudo, esses estimadores são muito instáveis quando aplicados a pequenas amostras (Hosking et al., 1985) e podem gerar valores absurdos do parâmetro de forma (k) da GEV (Martins & Stedinger, 2000). São recomendados estimadores de momentos LH, uma generalização de momentos L (Hosking, 1990), que são baseados na combinação linear de estatísticas de altas ordens, introduzidas para caracterizar, também, a parte mais alta da distribuição e os valores extremos dos dados, através das ordens LH mais altas, com vistas à estimação de cheias com grandes períodos de retornos (Wang, 1997). A ordem LH igual a zero é similar a momentos L. Assim, a utilização de momentos LH possibilita o ajuste da distribuição, GEV, aos dados amostrais, desde a sua forma descritiva feita por momentos L até a caracterização dos valores mais altos para uma análise preditiva.

Desta forma, considerando a performance e versatilidade do Matlab, desenvolveu-se um algoritmo que realiza o ajuste da distribuição GEV a dados amostrais de valores extremos através de 5 níveis de combinação lineares das estatísticas de ordens LH (0, 1, 2, 3 e 4).

Distribuição GEV

A função de distribuição generalizada de valores extremos – GEV, que engloba as três formas assintóticas de distribuição de valores extremos conhecidas como valor extremo do tipo I (VEI), valor extremo do tipo II (VEII) e valor extremo do tipo III (VEIII) (Fisher Tippett, 1928; Gumbel, 1958), é definida, segundo Jenkinson (1955), como segue:

$$F(x) = P(X \le x) = \exp\left[-\left(1 - k\frac{x-u}{\alpha}\right)^{\frac{1}{k}}\right] \quad (1)$$

Quando o parâmetro k se aproxima de zero a equação (1) converge para a distribuição Gumbel

$$F(x) = P(X \le x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right] \quad (2)$$

sendo

$-\infty < x < +\infty$, $k=0$ - distribuição VEI (distribuição Gumbel)

$\varepsilon \le x < +\infty$, $k<0$ - distribuição VEII

$-\infty < x \le \omega$, $k>0$ - distribuição VEIII

onde *u* é um parâmetro de posicionamento com $-\infty < u < +\infty$, α é um parâmetro de escala com $0 < \alpha < +\infty$ e *k* é um parâmetro de forma com $-\infty < k < +\infty$. Assim, quando k>0, o limite superior da distribuição assintótica VEIII torna-se $\omega = u + \alpha/k$ e quando k<0 o limite inferior da distribuição assintótica VEII torna-se $\varepsilon = u + \alpha/k$.

O p-ésimo quantil da distribuição GEV é dado pelas seguintes relações decorrentes das equações (1) e (2):

$$x_p = u + \frac{\alpha}{k}\left[1 - (-\ln(p))^k\right], \qquad k \neq 0, \quad 0<p<1 \quad (3a)$$

$$x_p = u - \alpha \ln[-\ln(p)], \qquad k=0, \quad 0<p<1 \quad (3b)$$

Combinando a equação (2) com a variável reduzida de Gumbel (z): $z = (x - u)/\alpha$, obtém-se $F(x) = exp[-exp(-z)]$, que resulta em:

$$z = \ln[-\ln(F(x))] \quad (4)$$

A variável reduzida de Gumbel também se relaciona com o período de retorno (T), T=1/F(x). Logo, a equação (4) pode ser usada para definir z com respeito às distribuições VEI, VEII e VEIII. Assim, em um gráfico x *versus* z define-se o comportamento das três formas de distribuições de valores extremos, com relação à posição de plotagem de x, como mostra a Figura 1.

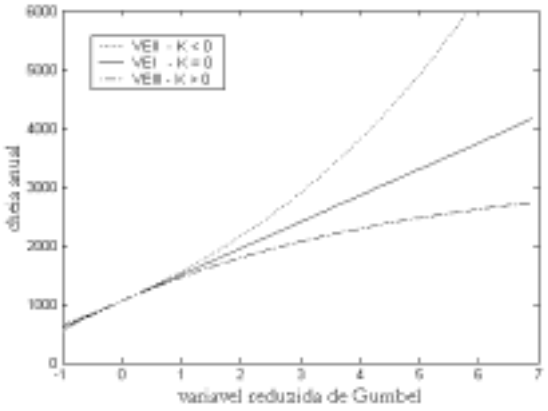

Figura 1: Distribuição das três formas de valores extremos representados pela GEV.

Momentos LH

Seja X_1, X_2,....,X_n uma amostra aleatória de uma população com função densidade de probabilidade $f(x)$ e função distribuição $F(x)$, e seja $x_{1:n} \le x_{2:n} \le \cdots \le x_{n:n}$ as estatísticas de ordem obtidos da amostra acima. O valor esperado do i-ésimo menor valor da variável é dado através da seguinte expressão (Hosking, 1990).

$$E\left[X_{i:n}\right] = \frac{n!}{(n-i)!(i-1)!}\int_0^1 x\left[F(x)\right]^{i-1}\left[1-F(x)\right]^{n-i}\,dF(x) \qquad (5)$$

Dada uma amostra de tamanho n retirada de uma distribuição $F(x)=P(X\le x)$, baseado na combinação linear das mais elevadas estatísticas de ordem e na equação (5), os momentos LH são definidos como:

$$\lambda_1^{\eta} = E\left[X_{(\eta+1):(\eta+1)}\right] \qquad (6a)$$

$$\lambda_2^{\eta} = \tfrac{1}{2}E\left[X_{(\eta+2):(\eta+2)} - X_{(\eta+1):(\eta+2)}\right] \qquad (6b)$$

$$\lambda_3^{\eta} = \tfrac{1}{3}E\left[X_{(\eta+3):(\eta+3)} - 2X_{(\eta+2):(\eta+3)} + X_{(\eta+1):(\eta+3)}\right] \qquad (6c)$$

$$\lambda_4^{\eta} = \tfrac{1}{4}E\left[X_{(\eta+4):(\eta+4)} - 3X_{(\eta+3):(\eta+4)} + 3X_{(\eta+2):(\eta+4)} - X_{(\eta+1):(\eta+4)}\right] \qquad (6d)$$

onde λ_1^{η}, maior valor esperado na amostra de tamanho $\eta+1$, corresponde a uma medida de posicionamento da distribuição; λ_2^{η}, metade da diferença entre o maior e segundo maior valor esperado na amostra de tamanho $\eta+2$, caracteriza a expansão da parte superior da distribuição; λ_3^{η}, reflete como está a assimetria da parte superior da distribuição, através dos três maiores valores esperados na amostra de tamanho $\eta+3$; e λ_4^{η} prover uma medida da pontiagudeza da parte superior da distribuição através dos quatros maiores valores esperados na amostra de tamanho $\eta+4$.

Quando $\eta=0$, momentos LH tornam-se iguais aos momentos L. Quando η aumenta, os momentos LH refletem mais e mais as características da parte superior da distribuição e dos valores extremos máximos dos dados. Momentos LH são chamados momentos L_1, momentos L_2, ... para $\eta = 1, 2, ...$ respectivamente. Normalizando os momentos LH, obtém-se o coeficiente de variação LH (τ_2^{η}), assimetria (τ_3^{η}) e curtose (τ_4^{η}), respectivamente, como:

$$\tau_2^{\eta} = \frac{\lambda_2^{\eta}}{\lambda_1^{\eta}}; \qquad \tau_3^{\eta} = \frac{\lambda_3^{\eta}}{\lambda_2^{\eta}}; \qquad \tau_4^{\eta} = \frac{\lambda_4^{\eta}}{\lambda_2^{\eta}} \qquad (7)$$

ordenando a amostra em $x_{(1)} \le x_{(2)} \le ..... \le x_{(n)}$, a estimativa dos momentos LH é feita como segue:

$$\hat{\lambda}_1^{\eta} = \frac{1}{{}^{n}C_{\eta+1}} \sum_{i=1}^{n} {}^{i-1}C_{\eta} x_{(i)} \tag{8a}$$

$$\hat{\lambda}_2^{\eta} = \frac{1}{2} \frac{1}{{}^{n}C_{\eta+2}} \sum_{i=1}^{n} \left({}^{i-1}C_{\eta+1} - {}^{i-1}C_{\eta}\, {}^{n-1}C_1 \right) x_{(i)} \tag{8b}$$

$$\hat{\lambda}_3^{\eta} = \frac{1}{3} \frac{1}{{}^{n}C_{\eta+3}} \sum_{i=1}^{n} \left({}^{i-1}C_{\eta+2} - 2\,{}^{i-1}C_{\eta+1}\, {}^{n-i}C_1 + {}^{i-1}C_{\eta}\, {}^{n-i}C_2 \right) x_{(i)} \tag{8c}$$

$$\hat{\lambda}_4^{\eta} = \frac{1}{4} \frac{1}{{}^{n}C_{\eta+4}} \sum_{i=1}^{n} \left({}^{i-1}C_{\eta+3} - 3\,{}^{i-1}C_{\eta+2}\, {}^{n-i}C_1 + 3\,{}^{i-1}C_{\eta+1}\, {}^{n-i}C_2 - {}^{i-1}C_{\eta}\, {}^{n-i}C_3 \right) x_{(i)} \tag{8d}$$

onde

$${}^{m}C_j = \binom{m}{j} = \frac{m!}{j!(m-j)!} \qquad \text{e} \qquad {}^{m}C_j = 0 \ \text{ quando } \ j > m.$$

Estimativas dos parâmetros da distribuição GEV:

Dada uma amostra, os três parâmetros *u*, α e *u* da distribuição GEV podem ser estimados, considerando a estimativa dos momentos LH amostrais através das equações (8a), (8b), (8c) e (8d), para um valor selecionado de η e $\kappa \neq 0$, como segue (Wang, 1977):

$$\lambda_1^{\eta} = u + \frac{\alpha}{\kappa}\left[1 - \Gamma(1+k)(\eta+1)^{-k}\right] \tag{9a}$$

$$\lambda_2^{\eta} = \frac{(\eta+2)\alpha\Gamma(1+\kappa)}{2!\kappa}\left[-(\eta+2)^{-\kappa} + (\eta+1)^{-\kappa}\right] \tag{9b}$$

$$\lambda_3^{\eta} = \frac{(\eta+3)\alpha\Gamma(1+\kappa)}{3!\kappa}\left[-(\eta+4)(\eta+3)^{-\kappa} + 2(\eta+3)(\eta+2)^{-\kappa}\right) \left(-(\eta+2)(\eta+1)^{-\kappa}\right] \tag{9c}$$

$$\lambda_4^{\eta} = \frac{(\eta+4)\alpha\Gamma(1+\kappa)}{4!\kappa}\left[-(\eta+6)(\eta+5)(\eta+4)^{-\kappa} + 3(\eta+5)(\eta+4)(\eta+3)^{-\kappa}\right) \left(-3(\eta+4)(\eta+3)(\eta+2)^{-\kappa} + (\eta+3)(\eta+2)(\eta+1)^{-\kappa}\right] \tag{9d}$$

Assim, os parâmetros *u*, α e *k* da distribuição GEV podem estão ser estimados, substituindo os três primeiros momentos LH nas equações (9a), (9b) e (c) pelo os seus respectivos estimadores amostrais nas equações (8a), (8b) e (8c), para cada valor de η selecionado.

Para facilitar o procedimento computacional, Wang, (1977) propôs uma equação aproximada para o cálculo de k, tomando como base as equações (9b) e (9c) e a equação (7) que define τ_3^{η}, a qual corresponde a:

$$k = \alpha_0 + \alpha_1\left[\tau_3^{\eta}\right] + \alpha_2\left[\tau_3^{\eta}\right]^2 + \alpha_3\left[\tau_3^{\eta}\right]^3 \tag{10}$$

onde os coeficientes $\alpha_0, \alpha_2 \; e \; \alpha_3$ variam em função de η (Tabela 1). Uma vez que *k* é obtido de (10), as equações (9b) e (9a) fornecem respectivamente α e *u*.

Tabela 1 - Valores dos coeficientes da equação (10)*.

η	α_0	α_1	α_2	α_3
0	0,2849	-1,8213	0,8140	-0,2835
1	0,4823	-2,1494	0,7269	-0,2103
2	0,5914	-2,2351	0,6442	-0,1616
3	0,6618	-2,4548	0,5733	-0,1273
4	0,7113	-2,5383	0,5142	-0,1027

*WANG (1997)

Análises de dados observados e dados obtidos via simulação Monte Carlo, mostraram que momentos LH reduzem as influências indesejáveis que os menores eventos amostrais podem exercer na estimação de eventos com grandes períodos de retorno, comparado ao uso de momentos L (Wang, 1997b, Queiroz, 2002).

Razões de momentos e coeficiente de variação LH

As razões de momentos e coeficiente de variação LH da distribuição GEV são calculadas de acordo com as relações em (7) e equações (9a), (9b), (9c) e (9d), resultando nas seguintes expressões:

$$\tau_2^{\eta} = \frac{1}{2}\frac{(\eta+2)\alpha\Gamma(1+k)\left[-(\eta+2)^{-k}+(\eta+1)^{-k}\right]}{ku+\alpha\left[1-\Gamma(1+k)(\eta+1)^{-k}\right]} \qquad (11a)$$

$$\tau_3^{\eta} = \frac{1}{3}\frac{(\eta+3)}{(\eta+2)}\left[\frac{-(\eta+4)(\eta+3)^{-k}+2(\eta+3)(\eta+2)^{-k}-(\eta+2)(\eta+1)^{-k}}{-(\eta+2)^{-k}+(\eta+1)^{-k}}\right] \qquad (11b)$$

$$\tau_4^{\eta} = \frac{1}{12}\frac{(\eta+4)}{(\eta+2)}\left[\frac{-(\eta+6)(\eta+5)(\eta+4)^{-k}+3(\eta+5)(\eta+4)(\eta+3)^{-k}}{-(\eta+2)^{-k}+(\eta+1)^{-k}}\right)$$

$$+\left(\frac{-3(\eta+4)(\eta+3)(\eta+2)^{-k}+(\eta+3)(\eta+2)(\eta+1)^{-k}}{-(\eta+2)^{-k}+(\eta+1)^{-k}}\right] \qquad (11c)$$

onde $\tau_2^{\eta}, \tau_3^{\eta}, \tau_4^{\eta}$ são, respectivamente, o coeficiente de variação LH, assimetria LH e curtose LH e $\Gamma(\bullet)$ a função gama. As razões de momentos LH e coeficiente de variação LH amostrais são obtidas, relacionando os momentos LH amostrais, indicados através das equações (8a), (8b), (8c) e (8d) seguindo as fórmulas em (7), onde os momentos LH da distribuição GEV são substituídos pelo os seus respectivos estimadores.

Teste de qualidade de ajuste da distribuição GEV via momentos LH

A distribuição GEV pode ser ajustada para uma série de dados, igualando os seus três primeiros momentos LH aos respectivos momentos LH amostrais, como mostrado anteriormente (Wang, 1997). A curtose LH da população $\left(\tau_4^\eta\right)$ é uma função da assimetria LH populacional $\left(\tau_3^\eta\right)$, onde ambos dependem apenas do parâmetro de forma k. Como o valor estimado da curtose LH amostral $\left(\hat{\tau}_4^\eta\right)$ não é usado no ajuste da distribuição, para determinar qual é a chance, estatisticamente pequena, de $\hat{\tau}_4^\eta$ ocorrer se a verdadeira distribuição subjacente é uma distribuição GEV. Dado um estimador amostral particular $\hat{\tau}_3^\eta$, precisa-se conhecer $p\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta\right)$. Porém, não é tão simples encontrar $p\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta\right)$ quando a mesma depende de τ_3^η da população. Contudo, dado $\hat{\tau}_3^\eta$, é possível inferir τ_3^η populacional, usando o teorema de Bayes, mostrado a seguir:

$$p\left(\tau_3^\eta \mid \hat{\tau}_3^\eta\right) \propto p\left(\hat{\tau}_3^\eta \mid \tau_3^\eta\right) p\left(\tau_3^\eta\right) \tag{12}$$

onde $p\left(\tau_3^\eta\right)$ é uma distribuição a priori que pode ser informativa ou não informativa. Assim, dado $p\left(\tau_3^\eta \mid \hat{\tau}_3^\eta\right)$ encontra-se:

$$p\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta\right) = \int P\left(\hat{\tau}_4^\eta \mid \tau_3^\eta, \hat{\tau}_3^\eta\right) P\left(\tau_3^\eta \mid \hat{\tau}_3^\eta\right) d\tau_3^\eta \tag{13}$$

onde $p\left(\hat{\tau}_3^\eta \mid \tau_3^\eta\right)$ em (12) e $p\left(\hat{\tau}_4^\eta \mid \tau_3^\eta, \hat{\tau}_3^\eta\right)$ em (13) podem ser derivada de $p\left(\hat{\tau}_4^\eta, \hat{\tau}_3^\eta \mid \tau_3^\eta\right)$, principalmente, utilizando simulação Monte Carlo. Portanto, teoricamente, é possível comparar $\hat{\tau}_4^\eta$ com $p\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta\right)$ para inferir se a distribuição subjacente é significativamente diferente da distribuição GEV. Não obstante, proceder por tal inferência requer muito esforço computacional que, em geral, não é prático (Wang, 1998).

Wang (1998) desenvolveu um teste de qualidade de ajuste da GEV com base em $p\left(\hat{\tau}_4^\eta \mid \tau_3^\eta = \hat{\tau}_3^\eta\right)$ como uma aproximação de $p\left(\hat{\tau}_4^\eta \mid \tau_3^\eta, \hat{\tau}_3^\eta\right)$, além de assumir que, as distribuições de $\hat{\tau}_3^\eta$ e de $\hat{\tau}_4^\eta$ da GEV, seguem uma distribuição conjunta normal. Para descrição completa da distribuição conjunta normal, precisa-se conhecer a média, desvio padrão e coeficiente de correlação dos estimadores amostrais $\hat{\tau}_3^\eta$ e $\hat{\tau}_4^\eta$. Suas médias são assumidas para serem os valores populacionais de τ_3^η e τ_4^η, respectivamente, onde é negligenciado algum erro de estimação. Os desvios padrão e coeficiente de correlação, denotados como $\sigma\left(\hat{\tau}_3^\eta\right), \sigma\left(\hat{\tau}_4^\eta\right) e\, \rho\left(\hat{\tau}_3^\eta, \hat{\tau}_4^\eta\right)$, respectivamente, ambos são funções de τ_3^η e do tamanho da amostra e podem ser encontrados através de simulação Monte Carlo.

A distribuição condicional amostral de $\hat{\tau}_4^\eta$ quando $\tau_3^\eta = \hat{\tau}_3^\eta$ é normalmente distribuída com média τ_4^η e desvio padrão dado como segue (Wang, 1998):

$$\sigma\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta = \tau_3^\eta\right) = \sigma\left(\hat{\tau}_4^\eta\right)\left[1 - \rho^2\left(\hat{\tau}_3^\eta, \hat{\tau}_4^\eta\right)\right]^{\frac{1}{2}} \tag{14}$$

Um teste de hipótese de que uma série de dados vem da distribuição ajustada, pode ser conduzido na base da estimativa amostral $\hat{\tau}_4^\eta$ através da comparação da seguinte estatística (WANG, 1998):

$$Z_W = \frac{\hat{\tau}_4^\eta - \tau_4^\eta}{\sigma\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta = \tau_3^\eta\right)} \tag{15}$$

com valores críticos de uma distribuição normal padrão.

O desvio padrão em (15), $\sigma\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta = \tau_3^\eta\right)$, é função de τ_3^η e do tamanho da amostra e pode ser calculado com $\sigma\left(\hat{\tau}_4^\eta\right) e\ \rho\left(\hat{\tau}_3^\eta, \hat{\tau}_4^\eta\right)$ através da equação (14), usando simulação Monte Carlo. Para evitar o enorme esforço computacional envolvido nas várias fases do teste, Wang (1998) propôs a seguinte aproximação:

$$\sigma^2\left(\hat{\tau}_4^\eta \mid \hat{\tau}_3^\eta = \tau_3^\eta\right) = \frac{b}{n} + \frac{c}{n^2} \tag{16}$$

onde

$$b = b_0 + b_1\left[\tau_3^\eta\right] + b_2\left[\tau_3^\eta\right]^2 + b_3\left[\tau_3^\eta\right]^3 + b_4\left[\tau_3^\eta\right]^4 \tag{17}$$

$$c = c_0 + c_1\left[\tau_3^\eta\right] + c_2\left[\tau_3^\eta\right]^2 + c_3\left[\tau_3^\eta\right]^3 + c_4\left[\tau_3^\eta\right]^4 \tag{18}$$

os coeficientes em (17) e (18) variam com os valores de η e estão apresentados na Tabela 2.

Tabela 2 - Valores dos coeficientes das equações (17) e (18)*

η	b_0	b_1	b_2	b_3	b_4	c_o	c_1	c_2	c_3	c_4
0	0.0745	0.0555	0.0067	-0.3090	0.2240	1.0100	-0.0282	-2.9336	4.0801	-1.0874
1	0.0579	-0.0328	0.1524	-0.4102	0.2672	1.3403	-0.8291	-3.8777	9.5371	-5.7866
2	0.0488	-0.0527	0.1620	-0.3856	0.2566	1.8800	-2.2233	-2.5825	10.4350	-7.3887
3	0.0380	-0.0309	0.0354	-0.1233	0.0878	2.6784	-4.8418	3.5255	2.3736	-3.2076
4	0.0241	0.0024	-0.0813	0.0733	-0.0210	3.7793	-8.3485	11.5170	-7.9095	1.9459

*WANG (1998)

A equação (16) tem a mesma forma daquela utilizada por Chowdhury et al.,(1991) para coeficiente de variação L e assimetria L. A mesma assegura que a variância é assintótica e inversamente proporcional ao tamanho da amostra. O segundo termo assegura tal efeito em amostras pequenas.

Rotina em Matlab para momentos LH e teste de Wang (1998)

Os procedimentos de ajuste da distribuição GEV através de momentos LH (Wang, 1997) e do teste de qualidade de ajuste para a mesma, proposto por Wang (1998), estão sistematizados através da seguinte rotina em Matlab que procede ao ajuste para qualquer tamanho de amostra.

```
function [k,α,u,cv,ca,cc,z_w]=fmomlh(x)
%--------------------------------------------------------------------------
%Calcula os 4 primeiros momentos-LH
% determina os parâmetros (k,α,u) da GEV para η=0, 1, 2, 3 e 4
% determina as razões de momentos LH (cv,ca,cc ) para η=0, 1, 2, 3 e 4
% determina os valores do Teste de WanG - z_w para η=0, 1, 2, 3 e 4
%--------------------------------------------------------------------------
x=sort(x);
N=length(x);
for j=1:5
    n=j-1;
if n= =0
   k1=N;
   k2=k1*(N-1)/2;
   k3=k2*(N-2)/3;
   k4=k3*(N-3)/4;
elseif n= =1
   k1=N*(N-1)/2;
   k2=k1*(N-2)/3;
   k3=k2*(N-3)/4;
   k4=k3*(N-4)/5;
elseif n= =2
   k1=N*(N-1)*(N-2)/6;
   k2=k1*(N-3)/4;
   k3=k2*(N-4)/5;
   k4=k3*(N-5)/6;
elseif n= =3
   k1=N*(N-1)*(N-2)*(N-3)/24;
   k2=k1*(N-4)/5;
   k3=k2*(N-5)/6;
   k4=k3*(N-6)/7;
else
   k1=N*(N-1)*(N-2)*(N-3)*(N-4)/120;
   k2=k1*(N-5)/6;
   k3=k2*(N-6)/7;
   k4=k3*(N-7)/8;
end
sigma1=0; sigma2=0; sigma3=0; sigma4=0;
for i=1:N
```

```
if n= =0
   F1=i-1;
   F2=1;
   F4=F1*(i-2)/2;
   F6=F4*(i-3)/3;
elseif n= =1
   F2=i-1;
   F1=F2*(i-2)/2;
   F4=F1*(i-3)/3;
   F6=F4*(i-4)/4;
elseif n= =2
   F2=(i-1)*(i-2)/2;
   F1=F2*(i-3)/3;
   F4=F1*(i-4)/4;
   F6=F4*(i-5)/5;
elseif n= =3
   F2=(i-1)*(i-2)*(i-3)/6;
   F1=F2*(i-4)/4;
   F4=F1*(i-5)/5;
   F6=F4*(i-6)/6;
else
   F2=(i-1)*(i-2)*(i-3)*(i-4)/24;
   F1=F2*(i-5)/5;
   F4=F1*(i-6)/6;
   F6=F4*(i-7)/7;
end
    F3=N-i;
    F5=F3*(N-i-1)/2;
    F7=F5*(N-i-2)/3;
coef1=F2;
coef2=F1-F2.*F3;
coef3=F4-2*F1*F3+F2*F5;
coef4=F6-3*F4*F3+3*F1*F5-F2*F7;
sigma1=sigma1+coef1*x(i);
sigma2=sigma2+coef2*x(i);
sigma3=sigma3+coef3*x(i);
sigma4=sigma4+coef4*x(i);
end
LH1(j)=sigma1./k1;
LH2(j)=sigma2./(2*k2);
LH3(j)=sigma3./(3*k3);
LH4(j)=sigma4./(4*k4);
cv(j)=LH2(j)./LH1(j); % coeficiente de variação
ca(j)=LH3(j)./LH2(j); % Coeficiente de assimetria
cc(j)=LH4(j)./LH2(j); % Coeficiente de curtose
end

%Cálculo dos parâmetros da GEV
%coeficientes para o cálculo de k (Wang,1997)
Ao=[0.2849 0.4823 0.5914 0.6618 0.7113];
A1=[-1.8213 -2.1494 -2.3351 -2.4548 -2.5383];
A2=[0.8140 0.7269 0.6442 0.5733 0.5142];
```

```
A3=[-0.2835 -0.2103 -0.1616 -0.1273 -0.1027];
for i=1:5
ni=i-1;
k(i)=Ao(i)+A1(i)*CS(i)+A2(i)*CS(i)^2+A3(i)*CS(i)^3;
α(i)=(LH2(i)*2*k(i))/((ni+2)*(-(ni+2)^(-k(i))+(ni+1)^(-k(i)))*gamma(1+k(i)));
u(i)=LH1(i)-α(i)*(1-gamma(1+k(i))*(ni+1)^(-k(i)))/k(i);
end

%---------- teste de hipótese proposto por Wang (1998)----------

b=[0.0745  0.0555 0.0067 -0.3090 0.2240
   0.0579 -0.0328 0.1524 -0.4102 0.2672
   0.0488 -0.0527 0.1620 -0.3856 0.2566
   0.0380 -0.0309 0.0354 -0.1233 0.0878
   0.0241 0.0024 -0.0813 0.0733 -0.0210];
c=[1.0100 -0.0282 -2.9336 4.0801 -1.0874
   1.3403 -0.8291 -3.8777 9.5371 -5.7866
   1.8800 -2.2233 -2.5825 10.435 -7.3887
   2.6784 -4.8418 3.5255 2.3736 -3.2076
   3.7793 -8.3485 11.517 -7.9095 1.9459];
for j=1:5
  bo(j)=b(j,1)+b(j,2)*ca(j)+b(j,3)*ca(j)^2+b(j,4)*ca(j)^3+b(j,5)*ca(j)^4;
  co(j)=c(j,1)+c(j,2)*ca(j)+c(j,3)*ca(j)^2+c(j,4)*ca(j)^3+c(j,5)*ca(j)^4;
  sig(j)=bo(j)/n+co(j)/n^2;
  t2(j)=(j+1)*α(j)*gamma(1+k(j))*(-(j+1)^(-k(j))+(j)^(-k(j)))/(2*k(j));
  t4(j)=(j+3)*α(j)*gamma(1+k(j))*(-(j+5)*(j+4)*(j+3)^(-k(j))+3*(j+4)*(j+3)*(j+2)^(-k(j))-
3*(j+3)*(j+2)*(j+1)^(-k(j))+(j+2)*(j+1)*(j)^(-k(j)))/(4*3*2*k(j));
  cct(j)=t4(j)/t2(j);
z(j)=abs(cc(j)-cct(j))/(sig(j))^0.5;
end
%----------
```

RESULTADOS E DISCUSSÕES

São mostrados a seguir os resultados do processo de ajuste da GEV a uma série de dados de cheias anuais e de vazões mínimas de sete dias, feitos através do uso do programa desenvolvido em Matlab. Os valores dos parâmetros da distribuição GEV podem ser apresentados, juntamente com as razões de momentos LH e valores do Teste de Wang, para os diferentes valores de η, em uma tabela desenhada numa janela de Figura como apresentado na Tabela 3. Os ajustes da distribuição GEV referentes aos cinco valores de η (LH0, LH1, LH2, LH3 e LH4) são plotados em um gráfico juntamente com os dados observados de vazões (Figura 2). O melhor ajuste determinado a partir do menor valor do teste de Wang é exibido com os valores de cheias em outro gráfico, como mostrado na Figura 2. O mesmo procedimento pode ser aplicado para ajustar vazões mínimas. Na citada figura é apresentado um gráfico de vazões mínimas de sete dias obtidos da mesma série de vazões diária observadas na estação fluviométrica sob o código 64685000 localizada no Paraná.

O programa realiza e exibe graficamente, ao mesmo tempo, os 5 ajustes da GEV aos dados de eventos hidrológicos máximos, referentes aos 5 níveis de combinações lineares (η). Dependendo do comportamento dos dados, o ajuste pode resultar em valores de $k>0$ ou $k<0$ para os diferentes valores de η, definindo o tipo de valor extremo (VEI, VEII ou VEIII). Na prática, quando $-0.04<k<0.04$ o ajuste se aproxima da distribuição Gumbel. Para $\eta=0$, o ajuste de momentos LH corresponde aos momentos L que atribui o mesmo peso para os dados durante o processo de ajuste. À medida que o valor de η aumenta, os valores amostrais mais altos recebem maiores pesos.

Tabela 3: Valores dos parâmetros da GEV, das razões de momentos LH e do Teste de Wang.

Tabela de Resultados

LH	Parâmetros da GEV			Taxas de momentos LH			Teste de Wang
	k	alfa	u	cv	ca	ck	
0	0.2164	997	3106	0.167	0.038	0.078	0.7295
1	0.2753	1065	3101	0.112	0.100	0.060	0.6187
2	0.3142	1126	3082	0.089	0.123	0.041	0.8683
3	0.3619	1217	3040	0.077	0.126	0.042	0.6580
4	0.3954	1291	2998	0.068	0.128	0.062	0.0820

Tamanho da amostra n = 44

k - parâmetro de forma
alfa- parâm. de escala
u - posicionamento
cv - coef. de variação
ca - coef. de assimetria
ck - coef. de curtose

"aperte uma tecla para ver os gráficos"

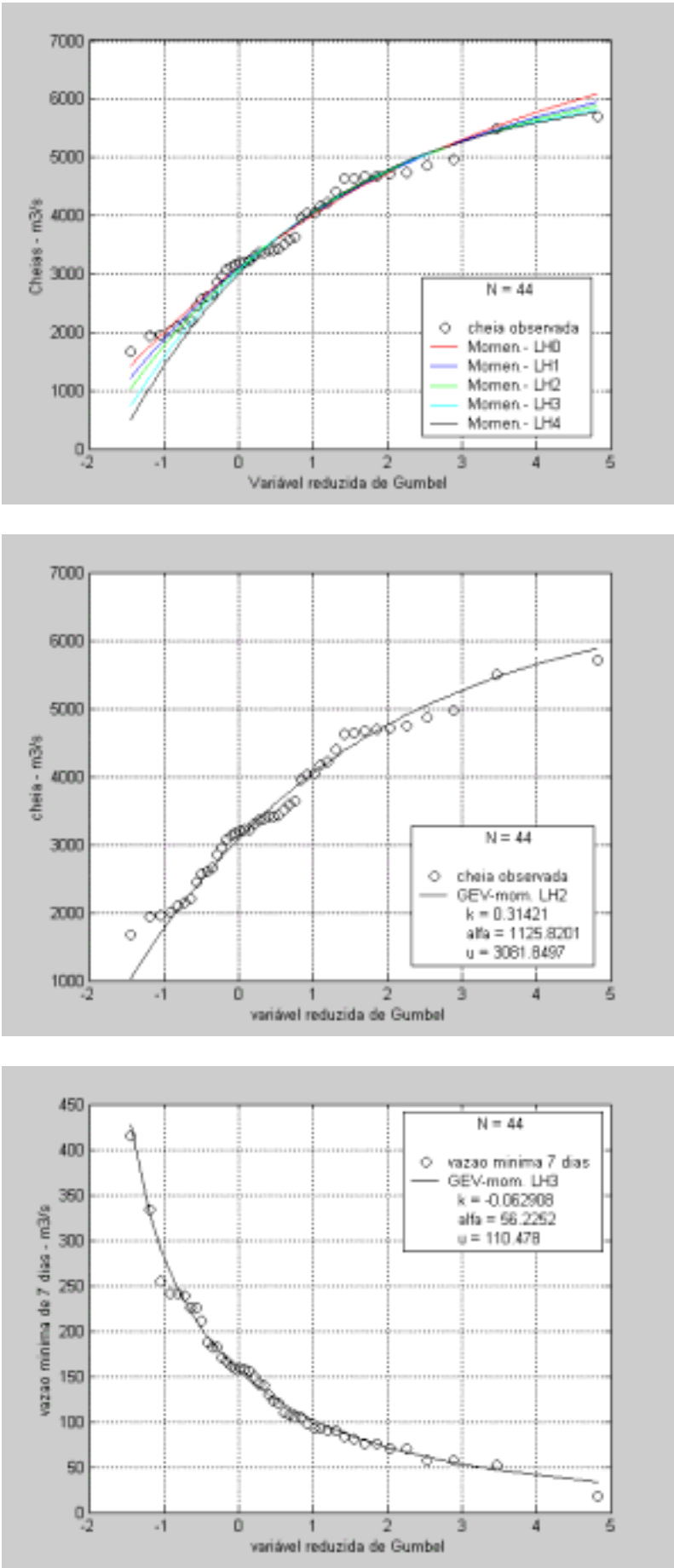

Figura 2: Ajustes da distribuição GEV as séries de cheias anuais e vazões mínimas de sete dias.

CONCLUSÃO

O programa apresenta ótima performance, independente do tamanho da amostra, onde o uso de momentos LH reduz as influências indesejáveis de pequenas amostras na estimação de eventos de grande período de retorno. O programa fornece como resultados, os 3 parâmetros da distribuição GEV, as razões de momentos LH que permite traçar o diagrama de momentos LH.

BIBLIOGRAFIA

CHOWDHURY, J. U.; STEDINGER, J. R.; LU, L. (1991). Goodness-of-fit test for regional generalized extreme value flood distributions. *Water Resour. Res*., v.27, n.7, p.1765-1776.

GUMBEL, E. J. (1958). *Statistics of Extreme*. Columbia University Press, New York, 396p.

FISHER, R. A.; TIPPETT, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, v.24, p.180-190.

HOSKING, J.R.M.; WALLIS, J.R.; WOOD, E.F. (1985). Estimation of the generalized extreme-value distribution by the method of probability weighted moments. *Technometrics*, v.27, n.3, p.251-261.

HOSKING, J.R.M., 1990, L-moments: analysis and estimation of distribution using linear combinations of order statistics, *J.R. Statit. Soc. Ser*. B, 52(2), 105-124, 1990.

JENKINSON, A..F., 1955, The frequency distribution of the annual maximum (or minimum) of meteorological elements, *Quarterly Journal of the Royal Meteorological Society*, 81, 158-171.

MARTINS, E. S. e STEDINGER, J.R., 2000, Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources. Research*, 36(3)*:*737-744.

OTTEN, A.; VAN MONFORT, M. A. J. (1980). Maximum-likelihood estimation of the general extreme-value distribution parameters. *J. of Hydrology*, v.47, p.187-192.

PRESCOTT, P.; WALDEN, A. T. (1980). Maximum - likelihood estimation of the parameters of the generalized extreme - value distribution. *Biometrika*, v.67, p.723-724.

QUEIROZ, M. M. F. de (2002). Análise de cheias anuais segundo distribuição generalizada. Tese (Doutorado) – Escola de Engenharia de São Carlos – Universidade de São Paulo.

WANG, Q.J., 1997, LH moments for statistical analysis of extreme events, *Water Resources Research*, 33(12): 2841-2848.

WANG, Q. J. (1998). Approximate goodness-of-fit test of fitted generalized extreme value distribution using LH moments. *Water Resou. Res*, v.34, n.12, p.3497-3502.