

## QUANTIS DE CHEIA GEV REGIONAIS - UMA APLICAÇÃO PRÁTICA

**Fabricio Müller**

COPEL - Superintendência de Planejamento  
Coordenadoria de Planejamento da Geração - Rua Treze de Maio, 616  
80510-030 Curitiba, PR  
splcnp@mail.copel.br

**Heinz Dieter Fill**

Centro de Hidráulica e Hidrologia Prof. Parigot de Souza - CEHPAR - Caixa Postal 1309  
80001-970 Curitiba, PR  
fill@cch.copel.br

### RESUMO

Apresenta-se uma aplicação prática de estimadores de quantis de cheia, os quais utilizam a distribuição de probabilidades generalizada de extremos (GEV), e o método dos momentos com pesos probabilísticos, em estações fluviométricas do rio Tibagi combinado com análise regional. Um novo estimador regional de quantis de cheia denominado NQR(GLS) é utilizado pela primeira vez. Este estimador é ainda combinado com o estimador 2P/AR (Krüger, 1996), resultando num estimador bayesiano linear que é uma extensão daqueles propostos por Fill (1994) e Krüger (1996).

### INTRODUÇÃO

Uma importante linha de pesquisa atual, para estimativa de quantis de cheia, é a utilização da distribuição GEV, com parâmetros estimados por PWM (Probability Weighted Moments - Momentos de Pesos Probabilísticos) combinado com análise regional. Dentro desta linha de pesquisa Fill (1994) propôs o estimador NQR (Normalized Quantile Regression) e um estimador bayesiano empírico (EB) que consiste em uma combinação linear do estimador NQR e o chamado estimador 2P proposto por Lu (1991). Krüger (1996) aperfeiçoou o estimador 2P de Lu, utilizando uma regressão regional e obteve um novo: o estimador EB, robusto e mais preciso.

Este trabalho propõe um aperfeiçoamento adicional, relevante no caso

das estações utilizadas na regressão dos parâmetros apresentarem significativa correlação espacial, e que consiste em utilizar o método GLS (Generalized Least Squares - Mínimos Quadrados Generalizado) na estimativa dos coeficientes da regressão regional. Adicionalmente ilustra o método aplicando-o no estudo de vazões máximas no rio Tibagi, estado do Paraná.

### MOMENTOS-L E A DISTRIBUIÇÃO GEV

Os momentos lineares ou momentos-L são descritores das distribuições de probabilidade semelhantes aos momentos tradicionais (média, desvio-padrão, etc.) cujos estimadores são combinações lineares dos dados observados. Entre as vantagens dos estimadores dos momentos-L sobre os estimadores dos momentos tradicionais pode-se citar:

- pelo fato de serem funções lineares dos dados sofrem menos os efeitos de observações atípicas (Lu, 1991);
- desde que exista a média, sempre existirão todos os demais momentos.

Os momentos-L podem ser expressos como funções lineares dos chamados momentos de pesos probabilísticos - PWM (Greenwood et al., 1979; Stedinger et al., 1993) definidos como:

$$B_r = E[\{F(x)^r\}] \quad (1)$$

onde  $r$  é a ordem do PWM, para uma variável aleatória  $X$  com função de distribuição  $F(x)$ .

Estimadores dos PWMs empregam as observações ordenadas  $X_{(n)} \leq \dots \leq X_{(1)}$  (estatísticas de ordem), em uma amostra  $(X_i | i = 1, \dots, n)$ . Estimadores PWM não tendenciosos podem ser obtidos por (Landwehr et al., 1979):

$$\begin{aligned} b_0 &= \frac{1}{n} \sum_{i=1}^n X_{(i)} \\ b_1 &= \sum_{i=1}^{n-1} \frac{(n-i)}{n(n-i)} X_{(i)} \\ b_2 &= \sum_{i=1}^{n-2} \frac{(n-i)(n-i-1)}{n(n-i)(n-2)} X_{(i)} \end{aligned} \quad (2)$$

A partir dos PWM os estimadores dos momentos-L podem ser calculados através das fórmulas (Hosking, 1986):

$$\begin{aligned} \hat{\lambda}_1 &= b_0 \\ \hat{\lambda}_2 &= 2b_1 - b_0 \\ \hat{\lambda}_3 &= 6b_2 - 6b_1 + b_0 \end{aligned} \quad (3)$$

Para os estimadores dos momentos-L de ordem superior a três existem expressões similares.

Assim como no caso dos momentos tradicionais, os momentos-L de ordem igual ou superior a dois são expressos através de seus coeficientes adimensionais de variação ( $\tau_2$ ), assimetria ( $\tau_3$ ) e curtose ( $\tau_4$ ):

$$\begin{aligned} \tau_2 &= \lambda_2 / \lambda_1 \\ \tau_3 &= \lambda_3 / \lambda_2 \\ \tau_4 &= \lambda_4 / \lambda_2 \end{aligned} \quad (4)$$

Estimadores para  $\tau_r$  são obtidos substituindo-se  $\lambda_r$  por seu estimador  $\hat{\lambda}_r$ . Deve-se notar que embora os  $\hat{\lambda}_r$  obtidos por (2) e (3) são não tendenciosos os  $\hat{\tau}_r$  serão tendenciosos. O quantil 100p% da distribuição GEV pode ser expresso apenas como função dos dois primeiros momentos-L e pelo parâmetro de forma  $\kappa$  (Fill, 1994):

$$q_p = \lambda_1 + \frac{\lambda_2}{1-2^{-\kappa}} \left[ 1 - \frac{(-\ln p)}{\Gamma(1+\kappa)} \right], \text{ se } \kappa \neq 0, \quad (5)$$

$$q_p = \lambda_1 + \frac{\lambda_2}{\ln 2} [-0,5772 - \ln(-\ln p)], \text{ se } \kappa = 0,$$

onde

$$\kappa = 7,8950c + 2,9554c^2 \quad (6)$$

e

$$c = \frac{2}{\tau_3 + 3} - 0,6309$$

Dividindo (5) por  $\lambda_1$ , obtém-se o quantil das vazões normalizadas  $x_p = q_p / \mu_q$ , onde  $\mu_q = \lambda_1$  é a média das vazões máximas anuais:

$$x_p = 1 + K(\kappa, p) \tau_2 \quad (7)$$

onde o fator de frequência  $K(\kappa, p)$  é dado por:

$$K(\kappa, p) = \frac{1}{1-2^{-\kappa}} \left[ 1 - \frac{(-\ln p)}{\Gamma(1+\kappa)} \right], \text{ se } \kappa \neq 0, \quad (8)$$

$$K(\kappa, p) = \frac{1}{\ln 2} [-0,5772 - \ln(-\ln p)], \text{ se } \kappa = 0.$$

Estimadores dos quantis de cheia normalizados  $\hat{x}_p$  são obtidos substituindo em (7) e (8)  $\tau_2$  e  $\kappa$  por seus estimadores regionais ou locais.

## ESTIMADORES REGIONALIZADOS

Os parâmetros da distribuição GEV utilizada (no nosso caso  $\lambda_1$ ,  $\tau_2$  e  $\kappa$ ) podem ser estimados exclusivamente a partir de dados observados no próprio local (estimador local) ou agregando as observações em vários locais hidrológicamente similares através de um modelo regional e usando este modelo na estimativa do parâmetro (estimador regional). Utilizando apenas um parâmetro local ( $\lambda_1$ ), a

12 JUN 1998

regionalização é 1P, e quando se utiliza dois parâmetros locais ( $\lambda_1$  e  $\tau_2$ ), ela é do tipo 2P. A análise local tradicional empregando estimadores locais para os três parâmetros da distribuição é chamada 3P. Além destes 3 estimadores, ainda é possível considerar um estimador do tipo bayesiano resultante de combinação linear entre estimadores dos tipos 1P e 2P. Finalmente os estimadores regionais podem diferir entre si em função do modelo regional adotado e do método de estimar o coeficiente deste modelo.

Foram considerados seis modelos neste trabalho:

**1. modelo GEV - 3P** - três parâmetros locais estimados pelo método PWM;

**2. modelo GEV / NQR (WLS) (Fill, 1994)** - estimador do tipo 1P, onde dois parâmetros regionais são obtidos por regressão WLS (*Weighted Least Squares* - Mínimos Quadrados Ponderados) dos quantis normalizados em função da área e da declividade. O modelo utilizado neste método tem a seguinte forma:

$$\ln x_p^i = \beta_0 + \beta_1 \ln a^i + \beta_2 \ln s^i + \varepsilon_i \quad (9)$$

para  $i=1, \dots, m$  locais, onde  $a^i$  é a área de drenagem,  $s^i$  a declividade para um local  $i$  e  $x_p^i$  é o quantil normalizado para uma probabilidade  $p$  e um local  $i$  (Fórmula 7), e  $\beta_0$ ,  $\beta_1$  e  $\beta_2$  são parâmetros. Se houver interesse o modelo original pode utilizar outras características fisiográficas locais, além da área e da declividade. O sistema de equações para a obtenção dos parâmetros  $\beta_0$ ,  $\beta_1$  e  $\beta_2$ , utilizando notação matricial, é o seguinte:

$$[A^T \Lambda^{-1} A] \hat{\beta} = [A^T \Lambda^{-1} \hat{\theta}_p] \quad (10)$$

$$[\hat{\theta}_p - A\hat{\beta}]^T \Lambda^{-1} [\hat{\theta}_p - A\hat{\beta}] = m - k$$

onde  $\hat{\theta}_p = \ln \hat{x}_p$  = vetor dos logaritmos dos quantis normalizados em cada local (dimensão  $m$ );  $\hat{\beta}$  = vetor de parâmetros (dimensão  $k$ );  $k$  = número de características fisiográficas mais um (no nosso caso  $k=3$ );  $A$  = matriz de dimensões  $m \times k$ , com os elementos da primeira coluna iguais a um e as colunas contendo os logaritmos das áreas e declividades locais;  $\Lambda = \sigma_e^2 I + \Sigma$ , onde  $\sigma_e^2$  = a variância de erro do modelo (também incógnita),  $I$  = matriz identidade, e  $\Sigma$  = matriz de covariância amostral de  $\ln \hat{x}_p$ . Para o modelo NQR (WLS), os elementos fora da diagonal principal são iguais a zero, correspondendo à hipótese de independência dos ruídos da regressão.

O sistema (10) é resolvido iterativamente para a solução simultânea do vetor  $\hat{\beta}$  e da variância do modelo  $\sigma_e^2$  (no caso  $k=3$  resultando 4 equações a 4 incógnitas). Stedinger e Tasker (1985), Fill (1994) e Krüger (1996) apresentam um método para a solução do sistema (10).

**3. modelo GEV / NQR (GLS)** - utilizado pela primeira vez neste trabalho; estimador do tipo 1P, onde dois parâmetros regionais são obtidos por regressão GLS dos quantis normalizados em função da área e da declividade; a diferença entre os modelos (2) e (3) está na matriz  $\Sigma$  de covariância dos resíduos da equação de regressão. Para o modelo (2) esta matriz é diagonal, enquanto que para o modelo (3) esta matriz é cheia com seus valores obtidos por simulação Monte Carlo.

**4. modelo GEV / 2P/LS (Lu e Stedinger, 1992)** - utiliza dois parâmetros locais, e o parâmetro de forma regional é um único para todas as estações, obtido por média ponderada com o tamanho dos registros de vazões máximas locais. O valor médio do coeficiente de assimetria utilizado para o cálculo de  $k$  regional é:

12 JUL 1998

$$\hat{\tau}_3^R = \frac{\sum_{i=1}^m n_i \hat{\tau}_3^i}{\sum_{i=1}^m n_i}, \quad (11)$$

com  $\hat{\tau}_3^i$  sendo o estimador local do coeficiente de assimetria-L para o local  $i$  ( $i=1, \dots, m$ ) com  $n_i$  anos de observações.

**5. modelo GEV / 2P/AR (Krüger, 1996; Krüger e Fill, 1996)** - utiliza dois parâmetros locais, e o parâmetro de forma  $\kappa$  regional é obtido por regressão em função da área e da declividade a partir do modelo:

$$\ln \kappa = \alpha_0 + \alpha_1 \ln a^i + \alpha_2 \ln s^i + \varepsilon_i \quad (12)$$

**6. modelo GEV - Bayesiano NQR (GLS) + 2P/AR** - pondera um valor entre os modelos (3) e (5). Utilizado também pela primeira vez neste trabalho. O quantil bayesiano é dado por

$$\hat{x}_p^{BAY} = b \hat{x}_p^{NQR} + (1-b) \hat{x}_p^{2P} \quad (13)$$

onde

$$b = \frac{\text{var}(\hat{x}_p^{2P})}{\text{var}(\hat{x}_p^{NQR}) + \text{var}(\hat{x}_p^{2P})} \quad (14)$$

Fill (1994) e Krüger (1996) apresentam os resultados para  $\text{var}(\hat{x}_p^{2P})$  e  $\text{var}(\hat{x}_p^{NQR})$  para outros modelos bayesianos: o de Fill pondera um valor entre os modelos (2) e (4); e o de Krüger, entre os modelos (2) e (5).

### OBTENÇÃO DA MATRIZ DE COVARIÂNCIA AMOSTRAL GLS

O modelo de regressão GLS considera não-nula a covariância entre as vazões máximas em estações distintas. A matriz de covariância deste modelo foi estimada por simulação Monte Carlo seguindo os passos:

1. seleção da probabilidade de não-excedência  $p$  dos quantis de cheia;
2. obtenção dos valores de  $\bar{\kappa}$  ( $\kappa$  médio) e  $\bar{\tau}_2$  ( $\tau_2$  médio) dados por

$$\bar{\tau}_2 = \frac{\sum_{i=1}^m n_i \tau_2^i}{\sum_{i=1}^m n_i} \quad (15)$$

$$\bar{\kappa} = \frac{\sum_{i=1}^m n_i \kappa^i}{\sum_{i=1}^m n_i} \quad (16)$$

onde  $m$  é o número de estações incluídas na regionalização,  $n_i$  é o número de anos de observação da estação  $i$  e  $\tau_2^i$  e  $\kappa^i$  são os valores locais do coeficiente de variação-L ( $L-Cv$ ) e do parâmetro de forma.

3. obtenção do valor de  $\bar{n}$  ( $n$  médio). O valor de  $\bar{n}$  é dado pelo valor inteiro arredondado da média dos valores de  $n_i$ ,  $i=1, \dots, m$ ;
4. geração de duas séries de variáveis aleatórias normais padronizadas de tamanho  $\bar{n}$  com coeficiente de correlação  $\rho(z)$  (escolhidos neste caso iguais a 0,2 ; 0,4 ; 0,6 ; 0,8), tais que

$$z_{2,i} = \rho(z) z_{1,i} + \varepsilon_i \quad (17)$$

onde  $\varepsilon_i$  é uma variável aleatória normal independente com média zero e variância unitária, e  $z_{1,i}$  e  $z_{2,i}$  são os elementos correspondentes das duas séries.

5. transformação das séries normais  $z_{j,i}$  em séries GEV  $x_{j,i}$  ( $j=1, 2$ ,  $i=1, \dots, \bar{n}$ ), com parâmetros  $\bar{\kappa}$  e  $\bar{\tau}_2$  através da transformação:

$$x_{j,i} = 192 \left[ \bar{\kappa} \cdot \Phi(z_{j,i}) \right]^{1/192} \quad (18)$$

onde  $\Phi$  é função de distribuição acumulada normal padrão e  $K[\bar{\kappa}, \Phi(z_{ji})]$  é obtido pela Fórmula (8);

6. cálculo dos estimadores  $\hat{\kappa}_j$  e  $\hat{\tau}_{2j}$ , para cada uma das duas séries  $x_{ji}$ ,  $j=1,2$ , e cálculo do coeficiente de correlação amostral  $\hat{\rho}(x)$  entre as séries GEV  $x_{1j}$  e  $x_{2j}$ ;
7. obtenção dos estimadores dos quantis  $\hat{x}_{pj}$ ,  $j=1,2$  para a probabilidade de não-excedência  $p$  selecionada no passo 1) através da Fórmula (7);
8. repetição dos passos 4) a 7) um número suficientemente grande de vezes. No presente estudo utilizaram-se 50000 repetições;
9. cálculo da variância e da covariância das séries de  $\hat{x}_{pj}$ ,  $j=1,2$ , obtidos na repetição do passo 8);
10. estimativa do  $\rho(x)$  correspondente ao  $\rho(z)$  preestabelecido no passo 3), através da média dos valores de  $\hat{\rho}(x)$  obtidos nas repetições do passo 7);
11. repetição dos passos 3) a 9) para diferentes valores de  $\rho(z)$  preestabelecidos, obtendo-se  $\text{var}(\hat{x}_p)$  e  $\text{cov}(\hat{x}_{p1}, \hat{x}_{p2})$  em função de  $\rho(z)$ . O valor da  $\text{var}(\hat{x}_p)$  é obtido como média entre  $\text{var}(\hat{x}_{p1})$  e  $\text{var}(\hat{x}_{p2})$ ;
12. interpolando-se entre os valores de  $\rho(x)$  (passo 6), obtém-se o valor de  $\text{cov}(\hat{x}_{p1}, \hat{x}_{p2})$  e  $\text{var}(\hat{x}_p)$  para os valores do coeficiente de correlação  $\hat{\rho}(x)$  estimado entre as estações usadas na regionalização;
13. multiplicam-se os valores de  $\text{var}(\hat{x}_p)$  e  $\text{cov}(\hat{x}_{p1}, \hat{x}_{p2})$  respectivamente por  $\bar{n}/n_i$  e  $\bar{n}/n_j$  ( $n_{ij}$  = período comum de observações nas estações  $i$  e  $j$ );
14. repete-se todo o procedimento para outras probabilidades de não

excedência dos quantis, obtendo-se uma matriz de covariâncias para cada valor de  $p$  (neste estudo escolheu-se  $p = 0,1 ; 0,01$  e  $0,001$ );

## APLICAÇÃO NA BACIA DO RIO TIBAGI

Para o cálculo das vazões máximas esperadas da bacia do rio Tibagi, foram inicialmente utilizadas 30 estações dentro ou muito próximas da mesma. A Tabela 1 apresenta estas 30 estações juntamente com o número de anos de observação, localização, área, declividade e coeficientes  $\kappa$  e  $\tau_2$  obtidos localmente. Para testar a adequação dos dados ao modelo utilizou-se o ajuste de regressão de  $\tau_2$  (Fill, 1994; Krüger, 1996; Krüger e Fill, 1996):

$$\ln \tau_2^i = \alpha_0^i + \alpha_1^i \ln a^i + \alpha_2^i \ln s^i + \varepsilon_i \quad (19)$$

com parâmetros estimados por WLS e OLS (*Ordinary Least Squares* - Mínimos Quadrados Simples). A Figura 1a apresenta a adequação dos dados ao modelo WLS, através de todas as 30 estações da Tabela 1. A partir deste teste inicial, fizeram-se várias tentativas para reduzir a dispersão dos dados, chegando-se a um número final de 17 estações selecionadas, hachuradas na Tabela 1. A Figura 1b apresenta o ajuste final de aderência ao modelo WLS (Fórmula 16), na mesma escala da Figura 1a. É fácil observar que as 17 estações selecionadas se ajustam melhor ao modelo que as 30 estudadas anteriormente, pela comparação entre as Figuras 1a e 1b. A Tabela 2 apresenta o erro do modelo e os coeficientes  $\alpha_j$  para a Fórmula (16), tanto para as estações escolhidas inicialmente como para aquelas da seleção final, para regressões OLS e WLS. Como as melhores abordagens são aquelas em que se obtém menores erros do modelo e coeficientes mais significativos, percebe-se, como era esperado, que a abordagem WLS produziu resultados melhores que a abordagem OLS, e que a seleção das 17 estações da escolha final resultou melhor que as 30 iniciais.

Tabela 1. Estações selecionadas para a regionalização de vazões máximas na bacia do rio Tibagi (estações hachuradas: 17 escolhidas na seleção final).

Estação	Rio	Número de observações	Declividade (m/km)	Área (Km <sup>2</sup> )	$\tau_2$	K
Tamanduá	Jaguariaíva	19	6,16	1780	0,422	-0,316
Salto do Itararé	Itararé	39	3,07	5218	0,274	-0,060
Tomazina	das Cinzas	65	5,56	2015	0,322	-0,159
Andirá	das Cinzas	63	1,37	5622	0,300	0,104
Figueira	Laranjinha	22	4,43	1060	0,282	-0,228
Santa Cruz	Tibagi	30	1,85	1340	0,264	0,040
Lajeado	Imbituva	16	1,50	1319	0,120	0,537
Uvaia	Tibagi	22	1,33	4440	0,270	-0,056
Engenheiro Rosaldo Leitão	Tibagi	20	0,49	5731	0,249	-0,056
Balsa do Pitangui	Tibagi	8	0,60	6940	0,224	0,498
Barra do Pitangui	Pitangui	17	1,10	1015	0,286	0,059
Tibagi	Tibagi	61	0,83	8550	0,270	-0,196
Fazenda Manzanilha	Iapó	19	2,71	2180	0,223	0,066
Fazenda Manzanilha Sr2	Iapó	8	2,71	2180	0,224	0,073
Telêmaco Borba	Tibagi	16	1,06	13400	0,403	-0,377
Barra do Ribeirão das Antas	Tibagi	28	0,81	15600	0,271	-0,133
São Jerônimo	Tibagi	8	0,53	17800	0,234	0,370
Porto Londrina	Tibagi	18	0,51	18768	0,239	0,325
Cebolão	Tibagi	8	0,97	20000	0,260	0,478
Jataizinho	Tibagi	64	1,05	21400	0,320	-0,230
Ponte Preta	Congonhas	19	2,78	370	0,286	0,440
Santa Fé	Bandeirante do Norte	15	4,68	1220	0,301	-0,262
Vila Silva Jardim	Pirapó	29	2,61	4627	0,245	-0,066
Rio dos Patos	dos Patos	65	3,18	1090	0,256	-0,088
Tereza Cristina	Ivaí	39	3,47	3570	0,244	0,025
Porto Espanhol	Ivaí	31	1,38	8600	0,235	-0,019
Porto Monteiro	Alonzo	22	1,70	2620	0,207	0,038
Ubá do Sul	Ivaí	29	0,41	12500	0,209	0,237
Vila Rica	Ivaí	11	0,99	19300	0,238	1,017
Salto Natal	Mourão	22	5,00	852	0,297	-0,355
Porto Bananeiras	Ivaí	22	1,54	24200	0,158	0,111
Castro	Iapó	26	3,57	1193	0,391	-0,339
Barra do Ribeirão Raposo Tavares	Corumbataí	30	0,90	3240	0,221	0,086
Santa Terezinha	Laranjinha	64	2,16	3240	0,285	0,072

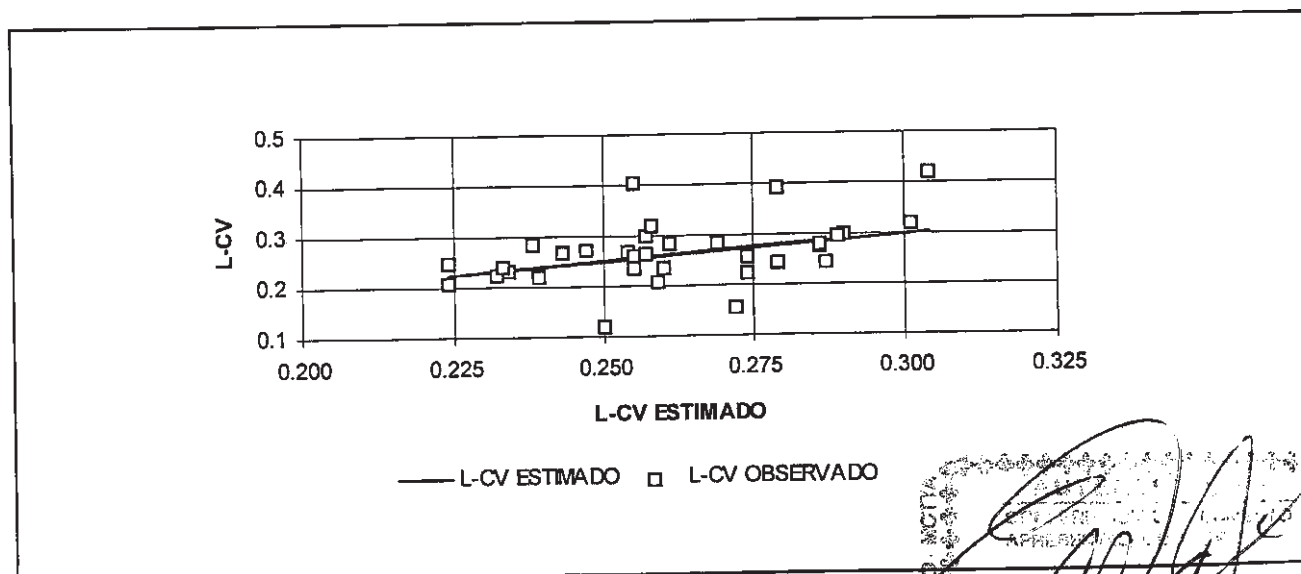


Figura 1a. Adequação do modelo de regressão WLS de L-Cv para as 30 estações da Tabela 1.

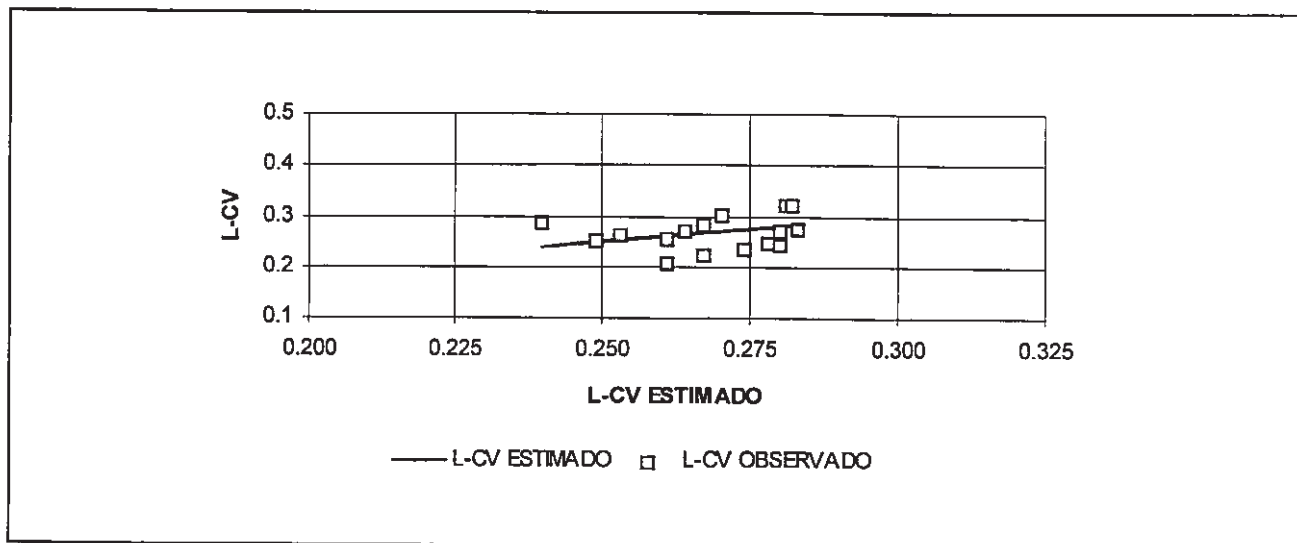


Figura 1b. Adequação do modelo de regressão WLS de L-Cv para as 17 estações hachuradas da Tabela 1.

Tabela 2. Resultados da regressão de L-Cv para as estações selecionadas para a bacia do rio Tibagi.

Método	Seleção	Erro do modelo $\sigma_e^2$	Intercepto $\alpha_0$	Coefficientes $\ln(a)$ $\alpha_1$	$\ln(s)$ $\alpha_2$
OLS	30 estações	0,0334	-1,670	0,0300	0,137
WLS	30 estações	0,0257	-1,648	0,0287	0,133
OLS	17 estações	-0,0025	-1,520	0,0196	0,0431
WLS	17 estações	$< 10^{-6}$	-1,810	0,0544	0,0734

As correlações amostrais entre as máximas anuais observadas intra-posto são significativas na maioria dos casos, como mostrado na Tabela 3, que apresenta os coeficientes de correlação amostrais para postos com período comum de dados igual ou superior a 15 anos. Isto indica a conveniência de se utilizar uma regressão GLS ao invés do modelo WLS utilizado por Fill (1994), Krüger (1996) e Krüger e Fill (1996). Devido ao curto período de dados em comum entre a maioria das estações, o que implica em erros amostrais elevados (Tasker e Stedinger, 1989), as correlações amostrais passaram por um processo de "alisamento". Este "alisamento" foi feito como segue: seja  $\rho_{ij}$  o coeficiente de correlação entre as vazões máximas da estação  $i$  e da estação  $j$ , e  $d_{ij}$  a distância (em km) entre estas duas estações. Tasker e Stedinger (1989) sugerem uma regressão não-linear na forma:

$$\rho_{ij} = \exp \left\{ \left[ \frac{d_{ij}}{\alpha d_{ij} + 1} \right] \ln \theta \right\} \quad (20)$$

onde  $\alpha$  e  $\theta$  são parâmetros. A expressão (20) pode ser transformada em

$$\frac{\ln \rho_{ij}}{d_{ij}} = \alpha (-\ln \rho_{ij}) + \ln \theta \quad (21)$$

o que permite a estimativa de  $\alpha$  e  $\theta$  por regressão OLS. A Figura 2 apresenta, nas ordenadas, as correlações observadas e estimadas pela Fórmula (20), para estações com 25 ou mais anos de observação em comum, e nas abscissas, a distância em km. Preferiu-se utilizar 25 anos de período em comum, e não 15, como na Tabela 3, para aumentar a representatividade dos parâmetros da curva de regressão.

Tabela 3. Correlações espaciais entre estações com mais de 15 anos de observação em comum na bacia do rio Tibagi.

nº Estação	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 Salto do Itararé	1,00																
2 Tomazina	0,90	1,00															
3 Santa Cruz	0,31	0,24	1,00														
4 Tibagi	0,55	0,57	0,70	1,00													
5 Jataizinho	0,63	0,79	0,46	0,85	1,00												
6 Santa Fé	0,35	0,31	(*)	0,05	0,20	1,00											
7 Rio dos Patos	0,31	0,39	0,55	0,69	0,59	(*)	1,00										
8 Tereza Cristina	0,33	0,37	(*)	0,68	0,55	(*)	0,73	1,00									
9 Figueira	(*)	0,95	(*)	0,54	0,87	(*)	0,39	0,40	1,00								
10 Uvaia	(*)	0,52	(*)	0,89	0,66	(*)	0,85	0,82	0,51	1,00							
11 Eng. Rosaldo Leitão	(*)	0,50	(*)	0,91	0,67	(*)	0,85	0,83	0,52	0,99	1,00						
12 Barra do Pitangui	(*)	0,30	(*)	0,52	0,28	(*)	0,40	0,52	0,19	0,65	(*)	1,00					
13 Fazenda Manzanilha	(*)	0,57	(*)	0,75	0,77	(*)	0,34	0,35	0,56	0,54	0,63	0,21	1,00				
14 Barra do Rib. das Antas	(*)	0,63	(*)	0,94	0,79	(*)	0,65	0,71	0,59	0,82	0,84	0,46	0,76	1,00			
15 Vila Silva Jardim	(*)	0,26	(*)	0,55	0,47	(*)	0,42	0,19	0,15	0,39	0,42	0,09	0,42	0,54	1,00		
16 Porto Espanhol	(*)	0,44	(*)	0,76	0,66	(*)	0,84	0,94	0,51	0,87	0,90	0,51	0,48	0,79	0,29	1,00	
Porto Monteiro SR1	(*)	0,39	(*)	0,66	0,61	(*)	0,60	0,52	0,45	0,60	0,58	0,05	0,47	0,73	0,60	0,56	1,00

(\*) período de observação comum menor que 15 anos

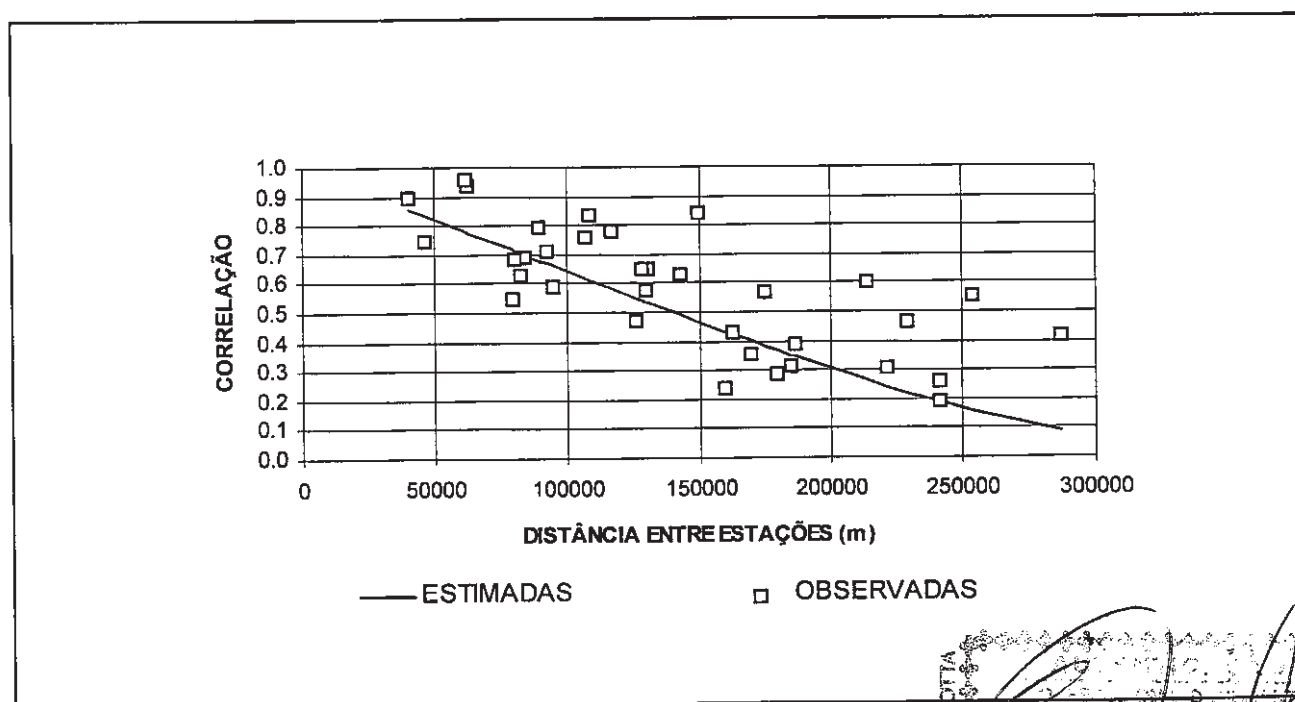


Figura 2. Correlações entre as máximas anuais observada e estimada pela Fórmula (20), contra distância em km, para as estações na bacia do rio Tibagi com 25 ou mais anos de observação em comum.



No caso do modelo (3) obteve-se, para a série final com 17 estações,  $\bar{n}$ ,  $\bar{\tau}_2$  e  $\bar{\kappa}$ , respectivamente iguais a 35, 0,273 e -0,0975. A partir destes dados, seguindo o método descrito na seção anterior, obtiveram-se  $cov(x_{p1}, x_{p2})$  e  $var(x_p)$  para cada um dos quatro valores escolhidos de  $\rho(z)$  e que estão mostrados na Tabela 4. O valor final de  $\rho(x)$  na tabela foi calculado pela média dos vários  $\hat{\rho}(x)$  obtidos. Para o modelo NQR(GLS), foram calculadas três matrizes  $\Sigma$ , uma para cada valor de  $\rho$  selecionado. A diagonal principal foi calculada com valores iguais a  $var(\hat{x}_p)$  ajustados pelo período de observação de cada estação; os outros valores da matriz são as  $cov(\hat{x}_{p1}, \hat{x}_{p2})$  obtidas por interpolação linear entre as correlações observadas alisadas (Fórmula 20) e os valores de  $\rho(x)$  obtidos na simulação Monte Carlo. Consideram-se os limites inferior e superior para a covariância, que são: para  $\rho(x) = 0$ ,  $cov(\hat{x}_{p1}, \hat{x}_{p2}) = 0$ ; para  $\rho(x) = 1$ ,  $cov(\hat{x}_{p1}, \hat{x}_{p2}) = var(\hat{x}_p)$ . Para o modelo (2), NQR(WLS), as matrizes  $\Sigma$  são diagonais, compostas pelos mesmos elementos da diagonal principal da matriz  $\Sigma$  do modelo NQR(GLS).

Para o modelo (4), o valor médio de  $\kappa$  obtido para as 17 estações definitivas foi de -0,0975. Já a regressão para o modelo (5) apresentou o seguinte resultado:

$$\kappa = 0,3885 + \ln(s) * (-0,0552) + \ln(s) * (-0,04285) + \varepsilon \quad (22)$$

com erro do modelo próximo a zero.

Os quantis (em  $m^3/s$ ) obtidos para o modelo GEV local mais os cinco modelos de regionalização analisados são apresentados na Tabela 5.

## CONCLUSÕES

De um ponto de vista puramente conceitual, desde que se dispõe de um

período de observação razoável nas estações ( $n > 10$  anos) parece indicado recomendar o modelo (6) para estimar os quantis de cheia nos vários locais devido aos fatos:

- i. as simulações Monte Carlo realizadas por Krüger (1996) e Krüger e Fill (1996) mostraram que os estimadores 2P baseados no método (5) são mais precisos que os baseados no método (4);
- ii. os modelos 1P são melhores em situações onde há períodos curtos de dados,  $\tau_2$  elevado e erros de modelo pequenos e os modelos 2P têm melhor desempenho na situação oposta. Como, em um caso real, não se sabe a priori o verdadeiro  $\tau_2$  nem o erro de modelo, os modelos bayesianos, que ponderam modelos 1P e 2P são mais robustos fornecendo resultados mais precisos para uma variada gama de valores de  $\tau_2$  e  $\sigma_e^2$  (ver Fill, 1994; Krüger, 1996; Krüger e Fill, 1996).
- iii. a regressão por GLS é vantajosa, pois ao contrário da WLS leva em conta as correlações espaciais entre estações. Não levando em conta estas correlações, diminui a precisão dos estimadores dos quantis de cheia (Stedinger e Tasker, 1985; Hosking e Wallis, 1988; Fill, 1994). Existem outros modelos bayesianos, ponderando valores entre os modelos (2) e (4) (Fill, 1994) e (2) e (5) (Krüger, 1996; Krüger e Fill, 1996), mas não foram levados em conta neste trabalho.

Entretanto, em casos onde não se dispõe de dados medidos localmente, ou apenas um período muito curto está disponível ( $n < 10$  anos), um modelo 1P pode ser o mais indicado, até porque a estimativa do  $\tau_2$  local ou não é possível ou está sujeita a um erro amostral considerável. No caso da ausência total de dados até a estimativa da vazão máxima média só pode ser feita por métodos de regionalização.

O uso de estimadores 3P (puramente locais) não é recomendável e menos que exista um período de observações superior a

Tabela 4. Resultados da simulação Monte Carlo para  $\text{var}(\hat{x}_p)$  e  $\text{cov}(\hat{x}_{p1}, \hat{x}_{p2})$ .

$p$	$\rho(z)$	$\rho(x)$	$\text{var}(\hat{x}_{p1})$	$\text{var}(\hat{x}_{p2})$	$\text{cov}(\hat{x}_{p1}, \hat{x}_{p2})$
0,1	0,2000	0,1893	0,0226	0,0228	0,0178
0,01	0,2000	0,1884	0,2444	0,2465	0,0920
0,001	0,2000	0,1881	1,3507	1,3778	0,1336
0,1	0,4000	0,3829	0,0225	0,0226	0,0181
0,01	0,4000	0,3798	0,2456	0,2481	0,1029
0,001	0,4000	0,3825	1,3723	1,4300	0,2112
0,1	0,6000	0,5800	0,0226	0,0228	0,0192
0,01	0,6000	0,5800	0,2462	0,2489	0,1267
0,001	0,6000	0,5780	1,3837	1,4328	0,3896
0,1	0,8000	0,7844	0,0226	0,0228	0,0207
0,01	0,8000	0,7848	0,2464	0,2512	0,1731
0,001	0,8000	0,7837	1,4001	1,4481	0,7694

100 anos (Fill, 1994). Na Tabela 5 verifica-se também que este estimador apresenta um valor bastante diferenciado em relação aos vários modelos regionais.

### AGRADECIMENTOS

Os autores agradecem ao engenheiro Cláudio Marchand Krüger pelas sugestões, e à engenheira Martha von Borstel Sugai pelo auxílio no decorrer da execução deste trabalho.

### REFERÊNCIAS

FILL, H. D. (1994). *Improving flood quantile estimates using regional information*. Ithaca: Cornell Univ. 264p. Tese (Doutorado) - Cornell Univ.

GREENWOOD, J. A, J. M. LANDWEHR, N. C. MATALAS e J. R. WALLIS (1979). Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form. *Water Resources Res.*, 15(6), p. 1049-1054, 1979.

HOSKING, J.R.M (1986) *The theory of probability weighted moments*. Yorkton Heights: IBM Research Division 160p. Research Report RC 12210.

HOSKING, J.R.M.. e WALLIS, J.R. (1988). The effect of intersite dependence on regional flood

frequency analysis. *Water Resources Res.*, 24(4). p. 588-600.

KRÜGER, C. (1996). *Estimador de quantis de cheias baseado em assimetria regional*. Curitiba: UFPR. 134p. Dissertação (Mestrado).

KRÜGER, C. E H.D. FILL (1996) Um estimador de cheias baseado em assimetria regional. *Revista Brasileira de Recursos Hídricos*. v.1, n.2. p.107-122.

LANDWEHR, J., N.C. MATALAS, e J.R. WALLIS (1979). Probability Weighted Moments Compared with some Traditional Techniques in Estimating Gumbel Parameters and Quantiles. *Water Resources Res.* 15(5), p. 1055-1064.

LU, L. H. (1991) *Statistical methods for regional flood frequency investigations*. Ithaca: Cornell University. Tese (Doutorado) - Cornell University.

LU, L.H e STEDINGER, J.R. (1992). Variance of two and three-parameter GEV/PWM quantiles estimators: formulae, confidence intervals, and a comparison. *J. of Hydrol.*, 138. p. 247-267.

STEDINGER, J.R. e TASKER, G.D. (1985). Regional hydrologic analysis 1. Ordinary, weighted and generalized least squares compared. *Water Resources Res.*, 21(9). p.1421-1432.

STEDINGER, J.R., VOGEL, R.M. e GEORGIUO, E.F. (1993) Frequency analysis of extreme events. In: *Handbook of Hydrology*, cap.18. editor-chefe D.J. Maidment. New York: Mc Graw Hill.

TASKER, G.D. e STEDINGER, J.R. (1989). An operational GLS model for hydrologic regression. *J. of Hydr.* 11(1) p.361-367.

Handwritten signature and stamp: "12 JUN 1998" and "ESTADÍSTICO".

Tabela 5. Quantis ( $m^3/s$ ) para os postos utilizados, por um modelo local e 5 modelos de regionalização.

Posto	Recorrência (anos)	Modelos GEV					Bayesiano
		3P	NQR (WLS)	NQR (GLS)	2P/LS	2P/AR	
Salto Do Itararé	10	602,65	607,44	616,35	601,64	600,17	604,42
	100	993,86	1093,96	1102,76	1029,50	1063,95	1082,32
	1000	1435,76	1753,90	1700,64	1556,70	1683,00	1694,28
Tomazina	10	457,79	434,23	441,62	460,31	460,09	456,37
	100	872,40	759,13	767,90	819,89	825,96	804,81
	1000	1461,60	1174,49	1139,03	1262,96	1284,54	1209,82
Figueira	10	262,90	261,24	267,21	266,80	267,27	267,25
	100	522,16	430,66	457,55	459,80	443,52	450,87
	1000	954,85	626,50	675,72	697,60	642,38	663,11
Santa Cruz	10	192,30	188,92	194,07	191,77	192,20	192,67
	100	287,31	300,01	336,16	325,07	306,88	318,92
	1000	372,32	421,56	514,09	489,32	429,10	469,62
Uvaia	10	592,79	585,36	597,90	591,72	592,05	594,05
	100	970,08	990,11	1070,77	1008,79	998,97	1038,88
	1000	1391,89	1491,87	1694,91	1522,68	1488,40	1626,67
Eng. Rosaldo Leitão	10	700,29	697,08	715,56	699,10	700,25	705,15
	100	1125,07	1128,69	1298,06	1167,90	1127,12	1218,47
	1000	1600,74	1633,03	2137,87	1745,54	1607,38	1933,09
Barra Do Pitangui	10	141,32	131,93	136,24	140,99	141,50	139,42
	100	211,06	199,02	235,05	243,86	222,12	229,47
	1000	270,67	265,54	364,09	370,62	300,83	338,26
Tibagi	10	1179,60	1176,28	1200,07	1190,84	1190,44	1191,98
	100	2232,34	2022,96	2192,71	2030,86	2041,65	2092,94
	1000	3863,66	3112,22	3572,86	3065,90	3104,21	3322,76
Fazenda Manzanilha	10	324,95	346,71	354,12	324,37	324,65	333,72
	100	459,90	583,10	619,78	527,05	518,30	566,84
	1000	573,39	869,31	944,05	776,78	746,52	863,01
Barra do Rib. das Antas	10	2212,81	2235,56	2266,33	2218,32	2208,92	2227,82
	100	3914,56	4098,98	4197,10	3785,83	3987,30	4115,54
	1000	6189,64	6739,55	6831,76	5717,27	6473,13	6746,51
Jataizinho	10	3728,99	3576,24	3620,96	3790,33	3767,38	3736,17
	100	7713,08	6651,70	6766,65	6740,80	7208,00	6983,81
	1000	14400,00	11114,76	11129,59	10376,27	12154,38	11446,96
Santa Fé	10	145,38	142,19	145,26	148,61	148,82	147,14
	100	306,55	237,52	249,56	260,37	253,41	250,89
	1000	597,40	350,32	368,84	398,09	374,24	370,20
Vila Silva Jardim	10	537,03	561,04	570,34	536,29	535,60	545,27
	100	867,31	992,02	1018,26	892,58	909,78	961,36
	1000	1245,32	1561,02	1575,60	1331,60	1393,70	1507,52
Rio Dos Patos	10	347,64	350,26	359,00	347,47	348,16	349,59
	100	579,51	566,86	616,42	584,48	558,04	571,68
	1000	859,13	810,14	921,41	876,53	787,98	827,64
Tereza Cristina	10	1715,99	1799,10	1828,59	1710,93	1708,99	1735,49
	100	2555,31	3179,45	3236,90	2843,74	2892,81	3027,20
	1000	3333,42	4991,30	4934,59	4239,54	4416,31	4697,14
Porto Espanhol	10	2909,83	3066,78	3118,12	2903,25	2898,25	2953,88
	100	4459,82	5444,05	5680,20	4779,55	4900,69	5264,68
	1000	6049,68	8637,56	9093,95	7091,46	7532,29	8486,42
Porto Monteiro SR1	10	1095,58	1187,31	1214,98	1092,91	1094,21	1123,86
	100	1561,03	1965,32	2142,85	1742,47	1760,30	1879,97
	1000	1979,53	2888,90	3329,20	2542,81	2398,43	2866,36

## **Regional GEV Flood Quantiles - An Application**

### **ABSTRACT**

*This paper presents an application of flood quantile estimators using GEV (Generalized Extreme Value) distribution and PWM (Probability Weighted Moments) estimation combined with regional analysis at Tibagi River in Brazil. A new flood quantile estimator called NQR (GLS) is used for the first time. This estimator may be combined with a 2P/AR estimator (Krüger, 1996; Krüger and Fill, 1996) resulting in a Bayesian Linear Estimator which is an extension of those proposed by Fill (1994), Krüger (1996) and Krüger and Fill (1996).*

