

## Curvas Envoltórias Probabilísticas para a Estimação de Cheias de Projeto

José Alexandre Pinto Coelho Filho, Mauro Naghettini

Universidade Federal de Minas Gerais

alexandre\_ufmg@yahoo.com.br; naghet@netuno.lcc.ufmg.br

Éber José de Andrade Pinto

Companhia de Pesquisa de Recursos Minerais

eber@bh.cprm.gov.br

Recebido: 20/04/10 - revisado: 01/11/10 - aceito: 21/02/11

---

### RESUMO

As curvas envoltórias regionais têm sido utilizadas desde o início da década de 1920, como meio simples de sintetizar graficamente o regime de vazões de cheias observadas em estações de monitoramento localizadas em uma ou mais regiões geográficas. De maneira geral, as curvas envoltórias de cheias representam o conhecimento disponível no que se refere à observação de vazões máximas por área de drenagem, sintetizando graficamente o regime de vazões máximas em uma determinada região. Entretanto, as curvas envoltórias são dependentes das amostras disponíveis, o que implica em limites superiores empíricos passíveis de serem excedidos. A tarefa de atribuição de uma probabilidade de superação às curvas envoltórias é complexa e tem sido objeto de recentes investigações [Castellarin et al. (2005) e Castellarin (2007)], à procura de um meio simples para a obtenção de estimativas preliminares de vazões de projeto de estruturas hidráulicas. Este artigo apresenta o emprego e a extensão das metodologias descritas nas citadas referências, na construção das curvas envoltórias probabilísticas para a bacia do alto rio São Francisco, no estado de Minas Gerais. Os resultados aqui obtidos foram considerados satisfatórios, assim demonstrando a viabilidade de se estimar quantis de cheias em pontos diversos de uma região hidrologicamente homogênea, a partir de um quantitativo mínimo de informações fluviométricas.

**Palavras-chave:** Curvas envoltórias de cheias, estimação de quantis de vazões máximas, vazões de projeto.

---

### INTRODUÇÃO

Atualmente, tem-se discutido sobre a necessidade e as diferentes possibilidades de inferir informações hidrológicas em bacias hidrográficas com escassez ou ausência de dados fluviométricos. Neste sentido, iniciativas recentes tais como as desenvolvidas pela Associação Internacional de Ciências Hidrológicas (IAHS - International Association of Hydrological Sciences), como aquela denominada PUB (Prediction in Ungauged Basins), descrita por Sivapalan et al. (2003), têm estimulado o desenvolvimento de trabalhos de concepção de métodos para a estimação de cheias em bacias hidrográficas para as quais se dispõe apenas de uma rede de monitoramento deficitária, ou até mesmo de nenhum monitoramento.

Nesse contexto, a interpretação probabilística das curvas envoltórias e a associação de uma probabilidade de igualdade ou superação aos valores raros e extremos vêm de encontro à afirmação apre-

sentada no parágrafo anterior, podendo constituir-se em uma ferramenta para o dimensionamento de estruturas hidráulicas, em locais desprovidos de redes de monitoramento.

Em geral, as curvas envoltórias de cheias constituem uma representação gráfica, em escala bi-logarítmica, das observações das máximas vazões ocorridas em diferentes estações fluviométricas versus suas respectivas áreas de drenagem, podendo sintetizar as vazões máximas em uma dada região geográfica.

As curvas envoltórias de cheias foram pioneiramente utilizadas por Jarvis (1925), que apresentou um trabalho baseado nas cheias máximas registradas em 888 estações de monitoramento fluviométrico nos Estados Unidos. No mesmo país, Crippen e Bue (1977) e Crippen (1982), aprimoraram o estudo inicial feito por Jarvis (1925), utilizando um total de 883 estações de monitoramento fluviométrico para a elaboração de dezessete curvas envoltórias, sendo uma para cada região com comportamento homogêneo do ponto de vista hidroló-

gico. Vogel et al. (2001), elaboraram um estudo abordando o trabalho de Crippen e Bue (1977) e Crippen (1982), concluindo que as curvas envoltórias obtidas por esses autores ainda eram válidas para 740 das 883 estações de monitoramento utilizadas na análise, entre os períodos de 1977 a 1994.

Recentemente, Castellarin et al. (2005) e Castellarin (2007) propuseram uma metodologia, apresentando resultados promissores, no sentido da interpretação probabilística dos métodos empíricos de construção de curvas envoltórias de cheias.

Com efeito, a associação de uma probabilidade de superação do valor esperado da curva envoltória é de grande utilidade para a estimação de vazões máximas em diversas bacias hidrográficas, em geral, e as do estado de Minas Gerais, em particular, via de regra providas de considerável potencial hídrico e desprovidas de uma eficiente rede de monitoramento hidrométrico. Assim, o presente artigo objetiva a aplicação da metodologia de interpretação probabilística associada aos valores esperados das curvas envoltórias de cheias construídas para a bacia hidrográfica do alto rio São Francisco, considerando a delimitação de regiões homogêneas do ponto de vista hidrológico.

Como decorrência, em uma segunda etapa, buscou-se aqui estender a metodologia proposta por Castellarin (2007), de forma a elaborar a curva de frequência completa de vazões máximas para bacias hidrográficas compreendidas nos limites das curvas envoltórias traçadas para a região em estudo.

Em última etapa, de posse das estimativas dos quantis de vazões máximas utilizando a metodologia das curvas envoltórias de cheias, efetuou-se a comparação dos resultados com os quantis de vazões máximas estimados pela aplicação da análise de frequência regional descrita por Hosking e Wallis (1997), e pela análise de frequência local, utilizando o método dos momentos (MOM) e o método dos momentos-L (MML). As comparações permitem considerar que os resultados foram satisfatórios, assim demonstrando a viabilidade da estimação de quantis de enchentes raras, a partir de uma coleção mínima de informações fluviométricas.

## BASES TEÓRICAS PARA A APLICAÇÃO DA METODOLOGIA

De modo genérico, uma curva envoltória de cheias pode ser posta sob a forma:

$$\ln\left(\frac{Q_{\max}}{A}\right) = a + b \ln(A) \quad (1)$$

na qual  $Q_{\max}$  denota a máxima enchente observada em uma estação fluviométrica de área de drenagem  $A$ ,  $a$  e  $b$  representam coeficientes regionais de regressão estabelecidos com base em um grande número de diferentes locais de monitoramento ao longo de uma região considerada homogênea, sob a ótica dos mecanismos físicos e climáticos que governam as ocorrências locais de enchentes extremas. Em sua interpretação probabilística das curvas envoltórias, Castellarin et al. (2005) adotaram as seguintes premissas: (i) a região que contém as estações fluviométricas em estudo é considerada homogênea, no sentido das hipóteses inerentes ao método de regionalização da cheia-índice, ou index-flood, tal como proposto por Dalrymple (1960); (ii) a relação entre a cheia-índice (ou a cheia-média)

local, representada por  $\mu_x$ , e a respectiva área de drenagem  $A$  obedece a uma lei de escala dada por:

$$\mu_x = CA^{b+1} \quad (2)$$

na qual  $b$  e  $C$  denotam constantes, sendo  $b$  coincidente com o coeficiente angular da equação (1). Desta forma, esse coeficiente pode ser expresso por meio da regressão das cheias-índice, em função de suas respectivas áreas de drenagem. Por outro lado, a estimativa do intercepto  $a$  pode ser obtida por meio da seguinte equação:

$$a = \max_{j=1, \dots, M} \left\{ \ln\left(\frac{Q_j}{A_j}\right) - b \ln(A_j) \right\} \quad (3)$$

onde  $Q_j$  representa a máxima vazão observada no local  $j$ , entre as  $M$  estações fluviométricas da região homogênea, e  $A_j$  representa a respectiva área de drenagem.

### Estimação da probabilidade de superação da curva envoltória de cheias

Castellarin et al. (2005) propuseram um estimador empírico da probabilidade de excedência  $p$ , que considera a correlação cruzada entre as diferentes estações fluviométricas. O estimador avalia um número equivalente de seqüências estatisticamente independentes MEC para as  $M$  estações fluviométricas da região homogênea, compostas por  $n$

valores de período comum de dados fluviométricos concomitantes. Esse estimador empírico é dado conforme a seguinte equação:

$$M_{EC} = \frac{M}{1 + \rho^{\beta}(M-1)} \quad (4)$$

em que

$$\beta = \frac{1,4(nM)^{0,176}}{(1-\rho)^{0,376}} \quad (5)$$

onde  $\bar{\rho^{\beta}}$  e  $\overline{(1-\rho)^{0,376}}$  são os valores médios das respectivas funções de coeficiente de correlação  $\rho$  entre as estações  $i$  e  $j$ , com  $1 \leq i < j \leq M$ . Conforme dedução de Castellarin et al. (2005), a estimação da probabilidade de superação da curva envoltória pode ser posta sob a forma:

$$p = 1 - \frac{\hat{n}_{ef} - \eta}{\hat{n}_{ef} + 1 - 2\eta} \quad (6)$$

onde  $p$  representa uma probabilidade empírica, calculada por meio de um parâmetro de posição de plotagem, o qual é dependente da distribuição selecionada, e  $\hat{n}_{ef}$  denota o número efetivo de observações amostrais, resultante da soma dos tamanhos efetivos das amostras de todos os subconjuntos regionais.

Posteriormente, Castellarin (2007) propõe a utilização da distribuição Generalizada de Valores Extremos (GEV), que de acordo com Vogel e Douglas (2005), é considerado um modelo distributivo muito usual para descrever o comportamento de eventos extremos.

### Modelagem da correlação cruzada entre as estações

De acordo com Castellarin (2007), uma possível aproximação da modelagem da correlação cruzada entre as estações de monitoramento utilizadas no estudo consiste no cálculo dos coeficientes de correlação amostrais utilizando estimadores já propostos por Stedinger (1981). Desse modo, busca-se a correlação cruzada entre as estações de monitoramento através de fórmulas de correlação empíricas em função das distâncias entre as seções transversais dessas estações ou entre os centróides das bacias hidrográficas por elas drenadas. Por simplici-

dade adota-se, doravante no presente artigo, que a distância refere-se àquela que separa as seções transversais de duas estações fluviométricas.

Nesse contexto, Castellarin (2007) cita o trabalho de Tasker e Stedinger (1989), no qual descrevem um modelo de correlação entre as estações em função da distância, comumente utilizado em estudos de regionalização. Este modelo é dado pela seguinte equação:

$$\rho_{i,j} = \exp\left(-\frac{\lambda_1 d_{i,j}}{1 + \lambda_2 d_{i,j}}\right) \quad (7)$$

na qual,  $\rho_{i,j}$  indica a correlação entre as estações  $i$  e  $j$ , separadas pela distância  $d_{i,j}$ , e  $\lambda_1$  e  $\lambda_2$  denotam coeficientes a serem estimados pelo método dos mínimos quadrados. Para Castellarin (2007), a equação (7) pode ser aplicada para a análise da tendência do decaimento gradual do grau de correlação cruzada em função da distância entre as estações utilizadas no estudo. Ainda, de acordo com o mesmo autor, uma função de correlação mais simples e passível de utilização pode ser posta da seguinte maneira:

$$\rho_{i,j} = \exp(-\lambda d_{i,j}) \quad (8)$$

onde,  $\rho_{i,j}$  indica a correlação entre as estações  $i$  e  $j$ , separadas pela distância  $d_{i,j}$ , e  $\lambda$  denota o coeficiente de decaimento exponencial da correlação cruzada para a região em estudo.

### Estimação do número efetivo de observações amostrais

De acordo com Castellarin (2007), a condição de períodos em comum entre todos os dados de vazões máximas anuais entre as séries hidrológicas, registradas em estações fluviométricas de uma região considerada homogênea, é raramente verificada na prática da hidrologia, inviabilizando a aplicação da equação (4).

Dessa maneira, o mesmo autor propõe a estimação de um número efetivo  $n_{ef}$  de observações amostrais, segundo o seguinte raciocínio: suponha que uma determinada região homogênea contenha  $M$  estações de monitoramento, onde cada estação possua  $n$  registros de vazões máximas anuais. Primeiro, torna-se necessário identificar os registros

únicos de vazão máxima anual, ou seja, o número de vezes,  $n_1$ , em que o registro de vazão máxima é presente em apenas um determinado ano ao longo da série histórica, não ocorrendo nenhum outro valor de vazão em comum no mesmo ano em nenhuma das  $M$  estações. Desta forma, por definição, essas observações são efetivas.

Assim, os registros de vazões máximas remanescentes, ou seja,  $(n-n_1)$  pertencentes às  $M$  estações inseridas dentro de uma região homogênea, são subdivididos em subconjuntos regionais  $N_{sub}$ . Desta maneira, quando ocorrerem dois ou mais registros de vazões máximas em comum para um determinado ano ou sequência de anos da série hidrológica, sendo esses registros únicos para esse mesmo período, eles irão constituir um subconjunto regional  $N_{sub}$ . Essa operação é ilustrada na Figura 1, a qual mostra a definição de subconjuntos regionais hipotéticos.

Desta forma, após definidos os registros em comum de vazões que constituem um subconjunto regional composto por  $L_s$  postos de monitoramento, dentre as  $M$  estações constituintes de uma região considerada homogênea, pode-se definir o conceito de número efetivo de observações utilizado na análise. Considerando os subconjuntos regionais formados a partir dos registros fluviométricos disponíveis para a região considerada homogênea, com registros de vazões máximas anuais em comum para um ou mais anos ao longo da série hidrológica, composto por  $L_s$  estações fluviométricas, o número efetivo de observações corresponde àqueles a serem efetivamente utilizados no estudo, dentre os registros das estações, de maneira a reduzir a correlação cruzada entre essas estações de monitoramento.

Nesse contexto, a equação para o cálculo do número efetivo de observações amostrais é definida da seguinte forma:

$$n_{ef} = n_1 + \sum_{s=1}^{N_{sub}} \frac{L_s \cdot I_s}{1 + \left[ \frac{L_s \cdot I_s}{\rho^\beta} \right] (L_s - 1)} \quad (9)$$

onde

$$\beta = \frac{1,4(L_s \cdot I_s)^{p,176}}{(1-\rho)^{0,376}} \quad (10)$$

em que os termos  $\rho^\beta$  e  $(1-\rho)^{0,376}$  possuem o mesmo significado daqueles presentes nas equações

(4) e (5),  $I_s$  denota o número de vezes em que uma combinação de dados de vazão em comum ocorreu única e exclusivamente para aquele ano, ou anos, ao longo da série hidrológica, que caracterizam a formação de um subconjunto regional. Para o exemplo apresentado na Figura 1, tem-se três subconjuntos regionais, onde os valores de  $L_s$  e  $I_s$  são, respectivamente: subconjunto 1 (2 e 3), subconjunto 2 (4 e 8), subconjunto 3 (2 e 3).

Conjunto de estações - original												
Ano	1	2	...	...	...	...	...	...	...	...	...	n
Estação 1												
Estação 2												
Estação (M-1)												
Estação M												
Subconjunto	$n_1$	1	2	3	$n_1$	1	2	3	1	$n_1$		

Conjunto de estações - rearranjado												
Ano												
Estação 1												
Estação 2												
Estação (M-1)												
Estação M												
Subconjunto	$n_1$	1	2	3								

Figura 1 - Exemplo de definição dos subconjuntos regionais

### Estimação da posição de plotagem

De acordo com Castellarin (2007), a estimação da posição de plotagem por meio da distribuição Generalizada de Valores Extremos (GEV) é a que fornece os melhores resultados para a aplicação da metodologia para a interpretação probabilística das curvas envoltórias. Assim, conforme apresentado por Arnell et al. (1986), reportado no trabalho de Castellarin (2007), uma estimativa da posição de plotagem para o cálculo de uma probabilidade empírica  $\eta$  segundo uma distribuição de frequência GEV pode ser aproximada pela seguinte relação linear:

$$\eta(\kappa) \cong \frac{\exp(\gamma) - 1}{\exp(\gamma)} - \frac{\pi^2}{12\exp(\gamma)} \kappa \cong 0,439 - 0,462\kappa \quad (11)$$

onde  $\gamma = 0,5772$  representa a constante de Euler e  $\kappa$  é o parâmetro de forma da distribuição GEV. Para Castellarin (2007), a aplicação da equação (11)

para a estimação da posição de plotagem segundo uma distribuição GEV traz vantagens em relação às outras constantes na literatura, uma vez que a referida equação é dependente apenas do parâmetro de forma  $\kappa$ .

#### Estimação da probabilidade de superação da curva envoltória

Após o cálculo da probabilidade empírica e do número efetivo de observações amostrais a serem utilizados no estudo, pode-se estimar por meio da aplicação da equação (6) uma probabilidade de superação  $p$ , cujo inverso, denota o tempo de retorno associado ao valor esperado da curva envoltória traçada para uma determinada região homogênea.

#### Extensão da metodologia proposta por Castellarin (2007)

Sabendo-se a probabilidade de superação, e conseqüentemente o tempo de retorno, do valor esperado da curva envoltória, pode-se estimar a vazão para esse mesmo período de retorno em qualquer bacia pertencente àquela região homogênea, na qual a distribuição Generalizada de Valores Extremos seja considerada o modelo probabilístico regional. Assim, conhecido o quantil de vazão estimado através da envoltória, denotado por  $X_T$ , e admitindo que a região em estudo tenha a distribuição GEV como o modelo distributivo para descrever o comportamento de vazões máximas anuais, e com a média das máximas vazões obtida pela lei de escala proposta por Castellarin et al. (2005), pode-se estender a metodologia e construir a curva de frequência completa para aquela bacia em particular, ou para qualquer outra localizada na área de estudo.

De fato, uma vez estimado o parâmetro de forma regional  $\kappa$ , por meio do método dos momentos-L, tal como descrito por Hosking e Wallis (1997), as estimativas dos outros parâmetros, a saber, o de escala  $\alpha$  e o de posição  $\beta$  da distribuição GEV podem ser calculadas por meio da solução do sistema formado pelas equações (12) e (13), que representam a média e a função de quantis da distribuição GEV, respectivamente:

$$\mu_X = \beta + \frac{\alpha}{\kappa} [1 - \Gamma(1 + \kappa)] \quad (12)$$

$$x(T) = \beta + \frac{\alpha}{\kappa} \left\{ 1 - \left[ -\ln \left( 1 - \frac{1}{T} \right) \right]^\kappa \right\} \quad (13)$$

Conhecidas as estimativas dos parâmetros de forma, escala e posição da distribuição de frequências GEV, pode-se construir a curva de frequência completa para qualquer bacia de área de drenagem compreendida nos limites da curva envoltória traçada para a região em estudo.

#### Avaliação de desempenho da metodologia estendida

Uma vez efetuada a estimação dos quantis de vazões máximas de interesse, utilizando a metodologia que associa um tempo de retorno às das curvas envoltórias de cheias, tal como descrito por Castellarin (2007), estendida para o cálculo da curva total de frequência, a comparação é efetuada com os quantis de vazões máximas estimados pela regionalização hidrológica utilizando os momentos-L, descrita por Hosking e Wallis (1997), e também pela análise de frequência local, utilizando as estimativas dos parâmetros da distribuição GEV pelo método dos momentos (MOM) e pelo método dos momentos-L (MML).

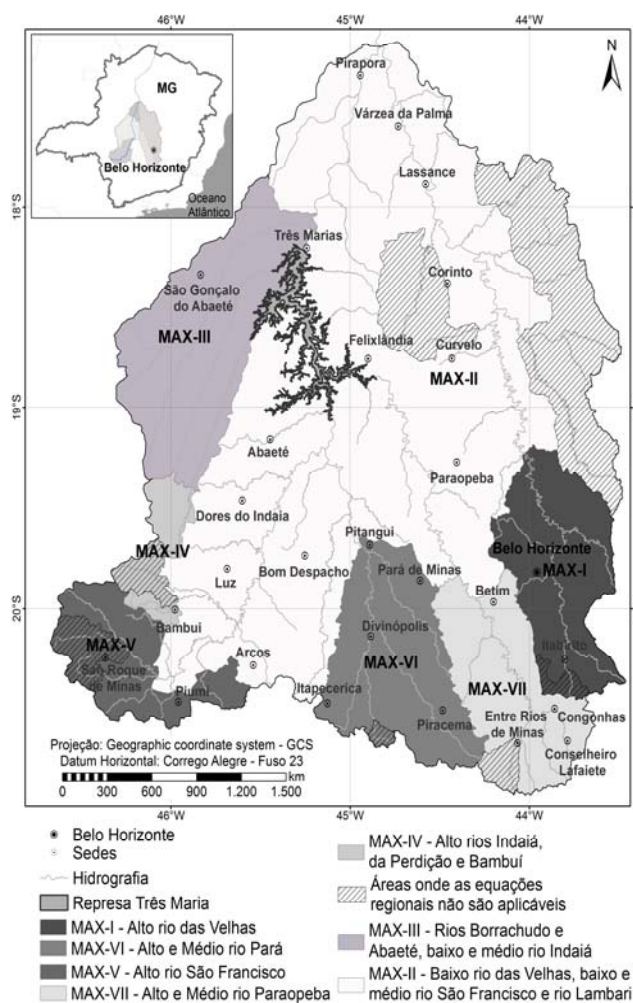
Por fim, pode-se calcular o desvio médio percentual para cada região homogênea, onde tem-se a análise dos desvios percentuais entre as estimativas dos valores de quantis de vazões máximas obtidos tanto pela metodologia das curvas envoltórias de cheias, quanto pela regionalização utilizando os momentos-L.

#### APLICAÇÃO DA METODOLOGIA PROPOSTA

Para a aplicação da metodologia proposta, a Companhia de Pesquisas de Recursos Minerais (CPRM, 2001) forneceu um estudo de regionalização de vazões máximas anuais para a porção mineira da bacia do rio São Francisco, abrangendo as sub-bacias 40 e 41, conforme a divisão da Agência Nacional de Energia Elétrica (ANEEL). A escolha desta área como região para a aplicação da metodologia deve-se à qualidade, quantidade e disponibilidade de informações hidrológicas fornecidas através da rede hidrométrica operada na bacia.

Outro aspecto importante a ser mencionado é o fato das regiões homogêneas já estarem definidas conforme estudo produzido por CPRM (2001). São elas: Região MAX I – Alto rio das Velhas, Região MAX II – Baixo rio das Velhas, baixo e médio rio São Francisco e rio Lambari, Região MAX III – Rios Borrachudo e Abaeté, baixo e médio rio Indaia,

Região MAX IV – Alto rios Indaiá, da Perdição e Bambuí, Região MAX V – Alto rio São Francisco, Região MAX VI – Alto e Médio rio Pará, Região MAX VII – Alto e Médio rio Paraopeba. A Figura 2 ilustra os limites geográficos das 7 regiões homogêneas mencionadas. Assim, foram usadas, a priori, as regiões delimitadas no referido estudo, sujeitas a posterior delimitação de novas regiões homogêneas. Conforme menção anterior, o trabalho de Castellarin (2007) pressupõe que a distribuição de frequências a ser utilizada seja a Generalizada de Valores Extremos (GEV).



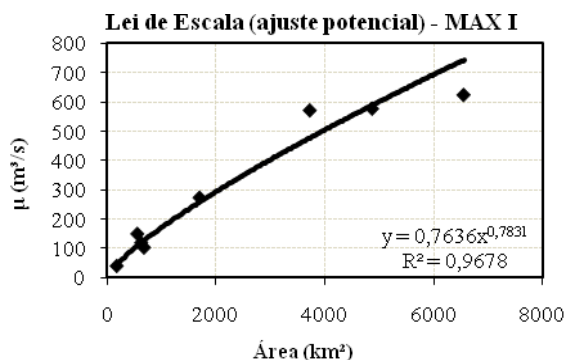
**Figura 2 - Limites geográficos das regiões homogêneas definidas por CPRM (2001)**

Dessa maneira, algumas regiões homogêneas, previamente definidas por CPRM (2001), foram verificadas sob o critério da aderência. Para tal, utilizou-se o conjunto de rotinas computacionais em linguagem Fortran-77 propostas por Hosking e Wal-

lis (1997), analisando os dados sob o ponto de vista da discordância, da aderência e da homogeneidade. Entretanto, as regiões MAX III e MAX IV, definidas como homogêneas por CPRM (2001), não apresentaram bons resultados para a medida de aderência para uma distribuição de frequências GEV. Na busca dos ajustes desejados, optou-se por trabalhar com o rearranjo de alguns postos fluviométricos pertencentes às regiões MAX II, MAX III e MAX IV, uma vez que estas são fronteiriças, permitindo que as estatísticas locais não prevalecessem sobre as características físicas locais nestas mesmas regiões. Nesse contexto, optou-se pela exclusão da estação fluviométrica Fazenda São Félix (40975000), pertencente à região homogênea MAX III, uma vez que sua permanência implicaria na inviabilidade de aceitação conjunta dos critérios de discordância, homogeneidade e aderência propostos por Hosking e Wallis (1997). Além disso, a estação Barra do Funchal (40930000), pertencente à região homogênea MAX IV, foi transladada para a região homogênea MAX II.

Tais procedimentos, tanto de exclusão quanto de deslocamento de estações de monitoramento de uma localidade a outra, são recomendados por Hosking e Wallis (1997) desde que exista um sentido físico em seu emprego, ou seja, que as estatísticas locais não prevaleçam sobre as características físicas locais dessas regiões.

#### Elaboração das curvas envoltórias empíricas



**Figura 3 - Ajuste potencial sobre a lei de escala para a região MAX I**

O traçado das curvas envoltórias faz-se por meio da equação (2), pressupondo uma lei de escala entre a cheia-índice local e a respectiva área de drenagem. Nesse sentido, para cada estação fluviométrica de uma região homogênea, foram tomadas as médias de suas vazões máximas anuais. Em seguida, foram elaborados gráficos das áreas de drenagem

das estações fluviométricas versus suas respectivas cheias médias anuais. Após o ajuste potencial, têm-se os gráficos tais como exemplificados pelas Figuras 3 e 4.

Com o valor do coeficiente angular, a ser obtido pela subtração da unidade ao valor do expoente do ajuste potencial, pode-se calcular o coeficiente linear da curva envoltória por meio da aplicação direta da equação (3). Este procedimento é repetido para cada posto de uma região considerada homogênea, e o máximo valor do coeficiente linear, dentre os postos pertencentes àquela região, é utilizado para o cálculo da envoltória.

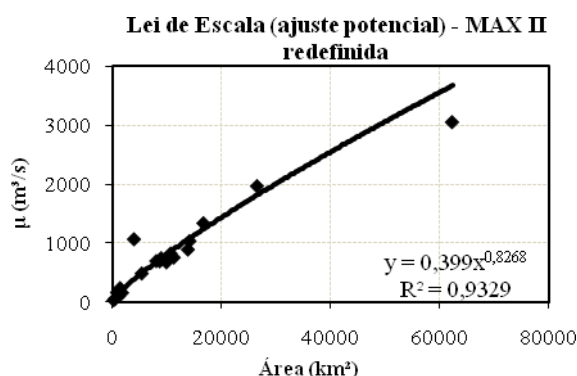


Figura 4 - Ajuste potencial sobre a lei de escala para a região MAX II redefinida

O cálculo da curva envoltória é decorrente da manipulação da equação (1). Na aplicação, foram tomados os logaritmos naturais do quociente dos valores das vazões máximas anuais das séries registradas pelas estações fluviométricas, pertencentes a uma mesma região homogênea, pelas respectivas áreas de drenagem, e compilados no eixo das ordenadas do gráfico da curva envoltória.

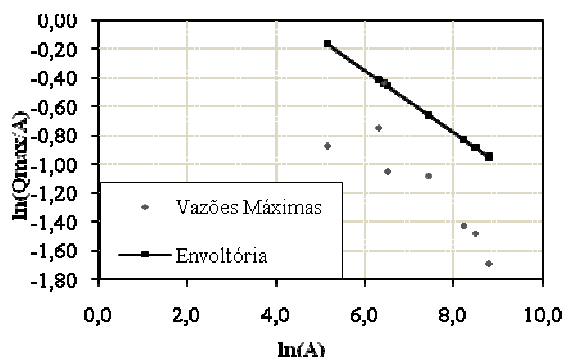


Figura 5 - Curva envoltória traçada para a região MAX I

O eixo das abscissas denota os logaritmos naturais dos valores da área de drenagem das respectivas estações componentes da bacia. As Figuras 5 e 6 apresentam as curvas envoltórias traçadas para as regiões MAX I e MAX II, embora esse procedimento tenha sido realizado para as 7 regiões homogêneas.

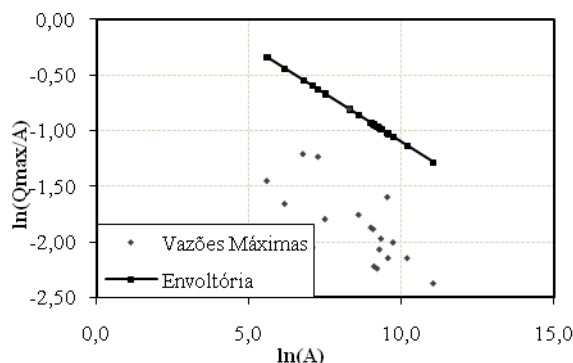


Figura 6 - Curva envoltória traçada para a região MAX II redefinida

#### Redução da variabilidade espacial contida nos dados amostrais

Castellarin (2007) propõe os modelos apresentados pelas equações (7) e (8), no sentido de se reduzir a variabilidade espacial contida nos dados amostrais das estações fluviométricas. O mesmo autor relata que a aplicação tanto da equação (7) quanto da equação (8) implica em resultados semelhantes. Nesse sentido fez-se a opção pela utilização da equação (8), pela sua maior facilidade de ajuste. Ainda, de acordo com Castellarin (2007), qualquer um dos dois modelos pode ser calibrado por meio de dois procedimentos distintos, a saber: (i) para toda a área de estudo, e (ii) utilizando as estações agrupadas por região homogênea.

No presente trabalho efetuou-se a calibração da equação (8) utilizando os dois procedimentos mencionados. A aplicação da equação (8) inicia-se com o cálculo da correlação entre as estações  $i$  e  $j$ , cujas seções transversais são separadas pela distância  $d_{i,j}$ , abrangendo todas as estações presentes no estudo, ou agrupadas por região homogênea. Assim, pode-se construir a representação gráfica das distâncias entre as estações versus os valores dos coeficientes de correlação. A Figura 7 ilustra os resultados obtidos para a calibração da equação (8) para toda a área de estudo.

Tais como em outros estudos dessa natureza (Castellarin, 2007), verifica-se aqui a enorme disper-

Tabela 1 - Síntese completa dos resultados relativos à aplicação da metodologia

Cálculos Efetuados	Calibração do modelo de correlação para toda a região						
	MAX I	MAX II redefinida	MAX III redefinida	MAX IV redefinida	MAX V	MAX VI	MAX VII
Número de estações	8	19	4	3	5	8	14
Número de observações	245	545	74	98	146	193	328
Número de observações únicas ( $n_1$ )	9	1	8	22	9	7	3
Número de observações efetivas ( $n_{ef}$ )	104,11	216,37	47,58	63,23	71,07	106,06	116,06
Parâmetro de forma da GEV ( $k$ )	0,0297	0,0655	-0,0839	0,0544	-0,0896	-0,0658	-0,0037
Coefficiente angular ( $b$ )	-0,2169	-0,1732	0,1815	-0,8248	0,1518	-0,2851	-0,0999
Coefficiente linear ( $a$ )	0,9563	0,6277	-1,9300	3,4350	-1,1430	1,2444	-0,2896
Parâmetro ( $C$ )	0,7636	0,3990	0,0657	17,538	0,0856	0,9096	0,2319
Parâmetro do modelo de correlação ( $\lambda$ )	-0,006	-0,006	-0,006	-0,006	-0,006	-0,006	-0,006
Probabilidade empírica ( $\eta$ )	0,4253	0,4087	0,4778	0,4139	0,4804	0,4694	0,4407
Subconjuntos amostrais ( $N_{sub}$ )	21	53	10	4	9	20	42
Probabilidade de excedência ( $p$ )	0,0055	0,0027	0,0110	0,0092	0,0073	0,0050	0,0048
Tempo de retorno (anos)	181	366	91	108	136	200	208
Cálculos Efetuados	Calibração do modelo de correlação por região homogênea						
	MAX I	MAX II redefinida	MAX III redefinida	MAX IV redefinida	MAX V	MAX VI	MAX VII
Número de estações	8	19	4	3	5	8	14
Número de observações	245	545	74	98	146	193	328
Número de observações únicas ( $n_1$ )	9	1	8	22	9	7	3
Número de observações efetivas ( $n_{ef}$ )	104,11	196,33	43,66	85,21	124,37	128,88	169,22
Parâmetro de forma da GEV ( $k$ )	0,0297	0,0655	-0,0839	0,0544	-0,0896	-0,0658	-0,0037
Coefficiente angular ( $b$ )	-0,2169	-0,1732	0,1815	-0,8248	0,1518	-0,2851	-0,0999
Coefficiente linear ( $a$ )	0,9563	0,6277	-1,9300	3,4350	-1,1430	1,2444	-0,2896
Parâmetro ( $C$ )	0,7636	0,3990	0,0657	17,538	0,0856	0,9096	0,2319
Parâmetro do modelo de correlação ( $\lambda$ )	-0,006	-0,005	-0,004	-0,082	-0,038	-0,014	-0,019
Probabilidade empírica ( $\eta$ )	0,4253	0,4087	0,4778	0,4139	0,4804	0,4694	0,4407
Subconjuntos amostrais ( $N_{sub}$ )	21	53	10	4	9	20	42
Probabilidade de excedência ( $p$ )	0,0055	0,0030	0,0119	0,0069	0,0042	0,0041	0,0033
Tempo de retorno (anos)	181	332	84	146	239	243	303

são entre os pontos, a qual, provavelmente, reflete a incapacidade da distância entre as estações, apenas, em explicar a correlação entre as correspondentes vazões máximas anuais. Certamente, a inclusão de outras variáveis explicativas, mais representativas das distâncias relativas entre as posições prevalentes dos mecanismos geradores das enchentes nas bacias em análise, poderia trazer algum benefício neste contexto, embora com a desvantagem de uma implementação muito mais complexa.

No quadro de sua inserção na metodologia de interpretação probabilística de curvas envoltórias de cheias, a representação da correlação cruzada por meio de um método simples, tal como o modelo de decaimento exponencial, conforme sugestão de Castellarin (2007), apresenta-se como uma alternativa parcimoniosa, viável e de execução prática.

Esse procedimento, segundo Castellarin (2007), busca a consideração da correlação cruzada entre as estações de monitoramento utilizadas no estudo e o conseqüente uso apenas das obser-

vações independentes entre os postos. Dessa forma, para cada uma das 7 regiões homogêneas, são definidos os subconjuntos regionais e o número de observações únicas.

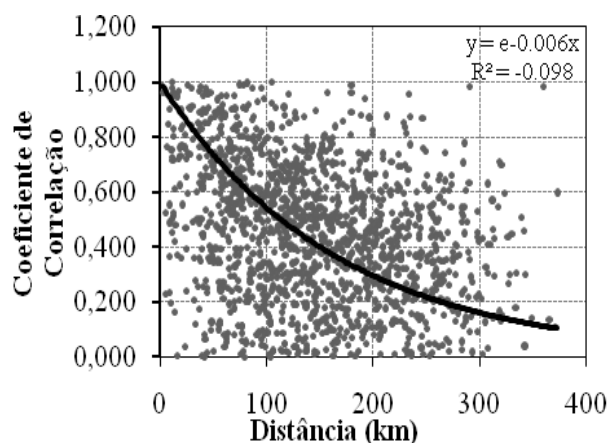


Figura 7 - Calibração da equação (8) para toda a área de estudo



### Cálculo do número efetivo de observações amostrais

O cálculo do número efetivo de observações amostrais é executado aplicando-se as equações (9) e (10). A próxima etapa consiste na obtenção do grau de correlação entre os subconjuntos regionais, por meio da utilização do modelo de correlação cruzada proposto conforme a equação (8), seja esta calibrada por região homogênea, ou para toda a área de estudo.

Com efeito, pode-se entrar com a distância entre as estações pertencentes aos subconjuntos regionais formados para uma região homogênea em um gráfico tal como apresentado pela Figura 7, que denota o modelo exponencial calibrado para toda a área de estudo, e assim obtêm-se o coeficiente de correlação entre os dados das estações que constituem os subconjuntos regionais. Análise similar pode ser realizada utilizando a equação (8) calibrada por região homogênea. Repetindo esses procedimentos para todos os subconjuntos regionais formados para cada uma das 7 regiões homogêneas, pode-se calcular os termos  $(1 - \rho)^{0,376}$  e  $\rho^{\beta}$  das equações (10) e (9), respectivamente, efetuando o cálculo do número de observações efetivas.

A Tabela 1 apresenta a síntese completa dos resultados obtidos pela aplicação da metodologia aos dados da bacia do alto rio São Francisco.

### Estimação dos quantis de vazões máximas utilizando o tempo de retorno associado às curvas envoltórias de cheias

O quantil de vazão máxima anual  $X_T$  pode ser obtido utilizando-se uma distribuição de frequência GEV, ajustada para o período de retorno associado às curvas envoltórias traçadas para as 7 regiões homogêneas presentes no estudo. A média das vazões máximas anuais pode ser estimada pelo ajuste potencial sobre a lei de escala proposta por Castellarin et al. (2005). Dessa forma, pode-se aplicar o sistema formado pelas equações (12) e (13), que representam a média e a função de quantis da distribuição GEV, respectivamente.

Para a aplicação dessas equações, resta a estimação dos parâmetros de forma  $\kappa$ , escala  $\alpha$  e posição  $\beta$ , da distribuição GEV. O parâmetro de forma regional foi estimado aplicando-se o método dos momentos-L. Estimado o parâmetro de forma regional para cada uma das 7 regiões homogêneas,

resta o cálculo dos parâmetros de posição e escala, os quais podem ser obtidos por meio das equações (12) e (14):

$$\beta = \mu_X - \frac{\alpha}{\kappa} [1 - \Gamma(1 + \kappa)] \quad (14)$$

Com efeito, a manipulação matemática do sistema formado pelas equações (12) e (14) gera a expressão (15), a qual é utilizada para a estimação do parâmetro de escala da distribuição GEV:

$$\alpha = \frac{(X_T - \mu_X) \kappa}{[-1 + \Gamma(1 + \kappa)] + \{1 - [-\ln(1 - 1/T)]^{\kappa}\}} \quad (15)$$

Estimado o parâmetro de escala, o mesmo pode ser levado à equação (14) para a obtenção do parâmetro de posição da distribuição GEV. De posse da estimativa dos três parâmetros que descrevem o comportamento da distribuição GEV, pode-se traçar a curva de frequência completa para a bacia com área de drenagem compreendida nos limites da curva envoltória. A próxima etapa consiste na comparação dos resultados obtidos àqueles referentes à análise de frequência local, usando o método dos momentos (MOM) e o método dos momentos-L (MML) para a estimação dos parâmetros da distribuição de frequência GEV. Em seguida, faz-se também a comparação daqueles quantis estimados pela aplicação da metodologia das curvas envoltórias com aqueles obtidos através da análise de frequência regional utilizando os momentos-L. Uma vez que neste estudo foram analisadas 61 estações de monitoramento fluviométrico, optou-se por apresentar os gráficos com as respectivas curvas de frequências, apenas para as estações com o maior número de registros fluviométricos para as regiões homogêneas MAX I e MAX II. As Figuras 8 e 9 ilustram essas curvas, onde os termos sobrescritos (1) e (2) indicam, pela ordem, a estimação do tempo de retorno associado à curva envoltória utilizando a equação de correlação calibrada para toda a região e por região homogênea.

Finalmente foi calculado o desvio médio percentual para as estações pertencentes a cada região homogênea, onde tem-se a análise dos desvios percentuais entre as estimativas dos valores de quantis de vazões máximas obtidos tanto pela metodologia das curvas envoltórias de cheias, quanto pela regionalização utilizando os momentos-L, conforme apresentado na Tabela 2. Dessa maneira, pode-se notar que, para um dado período de retor-

Tabela 2 - Desvios médios por estação para as regiões homogêneas

Região MAX I			Região MAX II redefinida		
T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>	T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>
2	-2,1	-2,1	2	-5,3	-5,5
5	13,4	13,4	5	47,0	48,0
10	19,0	19,0	10	68,7	70,2
50	26,3	26,3	50	99,2	101,4
100	28,3	28,3	100	108,0	110,4
500	31,6	31,6	500	122,9	125,6
1000	32,7	32,7	1000	127,8	130,6
10000	35,2	35,2	10000	139,8	142,9
Região MAX III redefinida			Região MAX IV redefinida		
T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>	T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>
2	2,9	2,7	2	0,8	1,0
5	-6,0	-5,5	5	-0,7	-1,6
10	-9,5	-8,8	10	-1,3	-2,8
50	-14,4	-13,3	50	-2,3	-4,6
100	-15,8	-14,5	100	-2,6	-5,1
500	-18,1	-16,7	500	-3,1	-6,0
1000	-18,9	-17,4	1000	-3,2	-6,3
10000	-20,7	-19,0	10000	-3,7	-7,1
Região MAX V			Região MAX VI		
T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>	T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>
2	-1,7	0,7	2	-4,7	-4,0
5	18,8	13,5	5	12,6	10,9
10	27,3	18,9	10	19,6	17,0
50	39,4	26,5	50	29,3	25,4
100	42,9	28,6	100	32,1	27,8
500	48,8	32,3	500	36,8	31,8
1000	50,7	33,5	1000	38,3	33,1
10000	55,4	36,5	10000	42,0	36,3
Região MAX VII			Os termos sobrescritos (1) e (2) acima dos "Desvios Médios" indicam a estimação do tempo de retorno associado à curva envoltória utilizando a equação de correlação calibrada para toda a região e por região homogênea, respectivamente.		
T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>			
2	-3,0	-2,2			
5	12,3	9,8			
10	18,6	14,7			
50	27,3	21,6			
100	29,9	23,6			
500	34,2	27,0			
1000	35,6	28,1			
10000	39,1	30,8			

Tabela 3 - Síntese dos cálculos efetuados para as regiões MAX II-a e MAX II-b

Cálculos Efetuados	MAX II-a <sup>1</sup>	MAX II-b <sup>1</sup>	MAX II-a <sup>2</sup>	MAX II-b <sup>2</sup>
Número de estações	6	5	6	5
Número de observações	158	146	158	146
Número de observações únicas ( $n_I$ )	8	16	8	16
Número de observações efetivas ( $n_{ef}$ )	119,54	96,67	101,89	106,83
Parâmetro de forma da GEV ( $k$ )	-0,0948	0,0061	-0,0948	0,0061
Coefficiente angular ( $b$ )	-0,2405	-0,1217	-0,2405	-0,1217
Coefficiente linear ( $a$ )	0,6980	-0,7707	0,6980	-0,7707
Parâmetro ( $C$ )	0,6847	0,2169	0,6847	0,2169
Parâmetro do modelo de correlação ( $\lambda$ )	-0,006	-0,006	-0,003	-0,009
Probabilidade empírica ( $\eta$ )	0,4828	0,4362	0,4828	0,4362
Subconjuntos amostrais ( $N_{sub}$ )	13	14	13	14
Probabilidade de excedência ( $p$ )	0,0043	0,0058	0,0051	0,0053
Tempo de retorno (anos)	231	172	197	190

Os termos sobrescritos (1) e (2) indicam a estimativa do tempo de retorno associado à curva envoltória utilizando a equação de correlação calibrada para toda a região e por região homogênea, respectivamente, nesta tabela, na Tabela 4 e nas Figuras 10 e 11.

Tabela 4 - Desvios médios por estação para as regiões MAX II-a e MAX II-b

Região MAX II-a			Região MAX II-b		
T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>	T (anos)	Desvio Médio (%) <sup>1</sup>	Desvio Médio (%) <sup>2</sup>
2	-2,3	-2,8	2	0,6	0,7
5	6,2	7,2	5	-0,9	-1,3
10	10,3	12,0	10	-1,6	-2,2
50	16,7	19,5	50	-2,5	-3,5
100	18,7	21,8	100	-2,8	-3,8
500	22,3	26,0	500	-3,3	-4,5
1000	23,5	27,4	1000	-3,4	-4,7
10000	26,6	31,0	10000	-3,8	-5,2

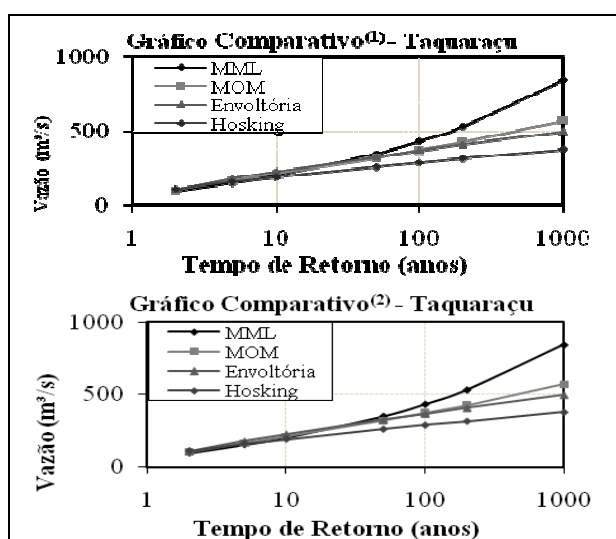


Figura 8 - Curvas de frequência (Região MAX I)

no,  $T = 100$  anos, por exemplo, o desvio médio percentual entre a aplicação da interpretação probabilística da curva envoltória de cheias e a utilização da metodologia descrita por Hosking e Wallis (1997), para a região homogênea MAX VI, oscila em torno de 25%.

Conforme pode ser observado na Tabela 2, e também na Figura 9, os resultados obtidos para a região homogênea MAX II redefinida foram discrepantes em relação às demais estações. Os desvios médios por estação para essa região variaram em média aproximadamente 70%, para um período de retorno de apenas 10 anos, comparados aos quantis de vazão estimados pela regionalização utilizando os momentos-L. Acredita-se que uma das justificativas para tais resultados seja a existência de regiões de comportamento hidrológico singular, inseridas dentro da região homogênea MAX II redefinida. Nesse sentido, optou-se por uma nova redefinição desta região, sendo ela subdividida nas regiões MAX II-a e MAX II-b. Após

essa redefinição, foram repetidos os procedimentos descritos ao longo do presente artigo para essas localidades. A Tabela 3 apresenta a síntese dos novos resultados. As Figuras 10 e 11 ilustram, pela ordem, as curvas de frequência obtidas utilizando o tempo de retorno estimado pelas curvas envoltórias das regiões MAX II-a e MAX II-b. Os termos sobrescritos (1) e (2) indicam a estimação do tempo de retorno associado à curva envoltória utilizando a equação de correlação calibrada para toda a região e por região homogênea (MAX II-a e MAX II-b), respectivamente. Para a análise dos desvios percentuais entre as estimativas dos valores de quantis de vazões máximas obtidos tanto pela metodologia das curvas envoltórias de cheias, quanto pela metodologia de regionalização utilizando os momentos-L, efetuou-se o cálculo dos desvios médios percentuais para as regiões homogêneas MAX II-a e MAX II-b, apresentados na Tabela 4.

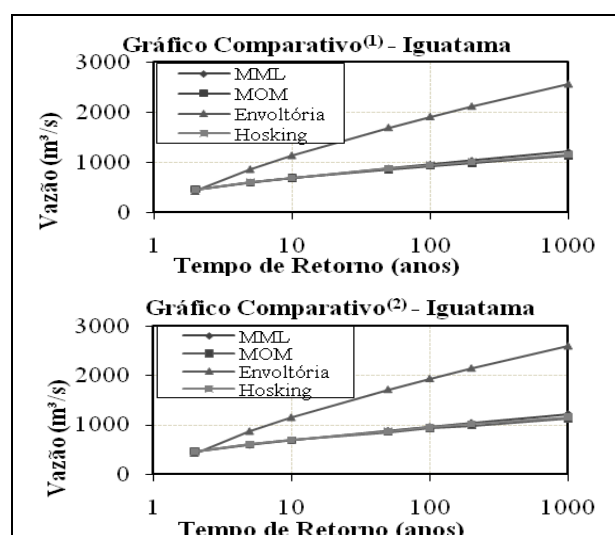


Figura 9 - Curvas de frequência (Região MAX II - redefinida)

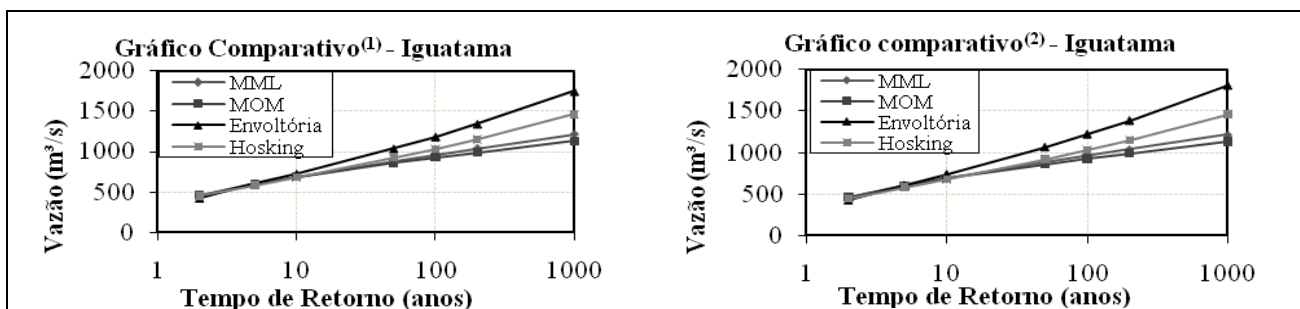


Figura 10 - Curva de frequência para a estação Iguatama

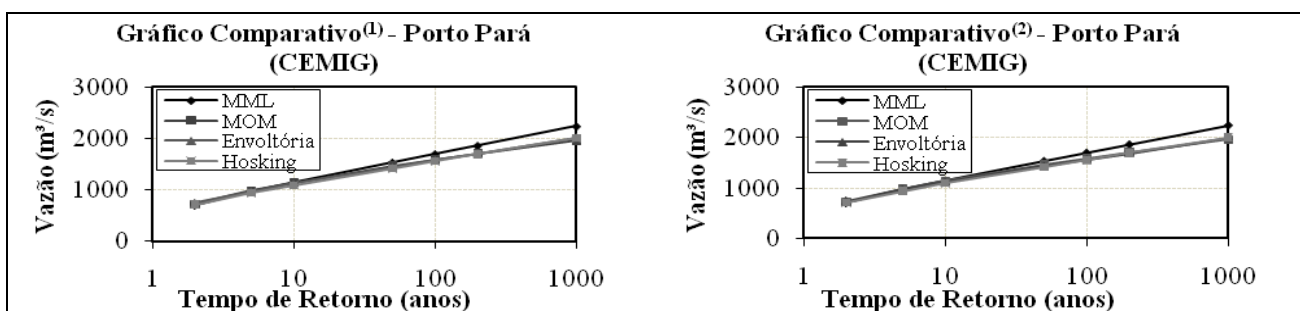


Figura 11 - Curva de frequência para a estação Porto Pará (CEMIG)

## DISCUSSÃO DOS RESULTADOS

A análise dos resultados contidos nas Figuras 8 a 11 mostra que os quantis de vazões estimados, para diferentes períodos de retorno, são semelhantes tanto para o tempo de retorno associado à curva envoltória utilizando o modelo de correlação calibrado para toda a área de estudo, quanto por região homogênea. Para o presente estudo, portanto, o tempo de retorno estimado para a curva envoltória diferiu segundo a maneira pela qual o modelo descrito pela equação (8) foi calibrado. Entretanto, para a estimação dos quantis de vazão máxima, as diferenças entre os tempos de retorno associados às curvas envoltórias traçadas para as 7 regiões homogêneas, tanto pela calibração da equação (8) para toda a área de estudo ou por região homogênea, foram pouco expressivas. Diferenças significativas, tanto para a estimativa da média das vazões máximas, quanto das vazões máximas ocorreram para a região homogênea MAX II. Tal constatação pode ser observada na Figura 9, e também na Tabela 2, onde, para um tempo de retorno de apenas 10 anos, os desvios médios ficaram em torno de 70%. Desta forma, a região ho-

mogênea MAX II redefinida foi subdividida nas regiões MAX II-a e MAX I-b. Em seguida, foram executados os mesmos procedimentos metodológicos descritos no presente artigo para essas duas regiões. De posse dos resultados, apresentados na Tabela 4, observa-se a sensível diminuição dos desvios percentuais após a subdivisão da região MAX II redefinida, com resultados próximos àqueles obtidos para as outras regiões. A análise conjunta da Tabela 4, e também das Figuras 10 e 11, mostram resultados similares àqueles obtidos para as demais regiões homogêneas.

## CONCLUSÃO

Os resultados aqui obtidos podem ser considerados satisfatórios, quando comparados àqueles resultantes de metodologias consagradas, como a análise regional por momentos-L. Ressalta-se, entretanto, que requer um conjunto menor de informações fluviométricas disponíveis em uma região. Com efeito, a partir de um quantitativo mínimo de informações, ou seja, a curva envoltória traçada para uma região considerada homogênea

e a estimativa do parâmetro de forma da distribuição GEV, pode-se estimar quantis de vazões de cheias para qualquer localidade interna à região, com área de drenagem compreendida pelos limites extremos da curva envoltória.

A metodologia aqui descrita constitui uma pequena extensão daquela proposta por Castellarin (2007), no sentido da obtenção não apenas da probabilidade de superação do valor esperado da curva envoltória regional, como também da curva de frequência completa de vazões máximas anuais em um dado local. Tal extensão se faz com o requisito de estimação do parâmetro de forma regional da distribuição GEV, que, neste trabalho, foi implementada com o emprego dos momentos-L regionais. Em aplicações mais amplas, talvez seja possível empreender a elaboração de mapas de isolinhas de parâmetros de forma da distribuição GEV, estendendo-se por vastas áreas geográficas, de modo a facilitar o seu emprego, em conjunto com as curvas envoltórias regionais.

## AGRADECIMENTOS

Os autores agradecem à CPRM – Companhia de Pesquisa de Recursos Minerais, pelo fornecimento dos dados aqui usados. Os autores também agradecem à CAPES, pela bolsa de mestrado de José Alexandre Pinto Coelho Filho, à FAPEMIG (PPM-00158-09) e ao CNPq (301133/2009-3) pelos auxílios concedidos, e, finalmente, aos revisores anônimos pelas valiosas sugestões.

## REFERÊNCIAS

- ARNELL, N. W., M. BERAN, e J. R. M. HOSKING, Unbiased plotting positions for the general extreme value distribution, *Journal of Hydrology*, 86, 59– 69, 1986.
- CASTELLARIN, A.; VOGEL, R. M.; MATALAS, N. C. Probabilistic behavior of a regional envelope curve. *Water Resources Research*, 41, W06018, 2005.
- CASTELLARIN, A. Probabilistic envelope curves for design flood estimation at ungauged sites. *Water Resources Research*, 43, 2007.
- CPRM - COMPANHIA DE PESQUISA DE RECURSOS MINERAIS – Superintendência Regional de Belo Horizonte. Regionalização de vazões sub-bacias 40 e 41, Convênio 015/2000 ANEEL – 013/CPRM/2000, Relatório Final. Belo Horizonte, 2001, 5v.
- CRIPPEN, J. R.; BUE, C. D. Maximum flood flows in the conterminous United States. *Geological Survey Water Supply Paper 1887*, US GPO, Washington, 1977.
- CRIPPEN, J. R. Envelope curves for extreme flood events, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, 108(HY10), 1208– 1212, 1982.
- DALRYMPLE, T., Flood-frequency analysis, *Manual of Hydrology: Part.3. Flood-flow Techniques*, Geological Survey Water Supply Paper 1543-A, U.S. GPO, Washington, D.C., 80p., 1960.
- HOSKING, J. R. M.; WALLIS, J. R. *Regional Frequency Analysis: An Approach Based on L-moments*. 1a ed. New York: Cambridge University Press, 1997, 224 p.
- JARVIS, C. S. Flood flow characteristics. *Transactions ASCE*, 88, p. 985-1032, 1925.
- SIVAPALAN, M.; TAKEUSHI, K.; FRANKS, S. W.; GUPTA, V. K.; KARAMBIRI, H.; LAKSHMI, V.; MCDONNELL, J. J.; MENDIONDO, E. M.; O'CONNELL, P. E.; OKI, T.; POMEROY, J. W.; SCHERTZER, D.; UHLENBROOK, S.; ZEHE, E. IAHS Decade on predictions in ungauged basins (PUB), 2003-2012: shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, 48(6), p. 857-880, 2003.
- STEDINGER, J. R. Estimating correlations in multivariate streamflow models, *Water Resour. Res.*, 17(1), 200– 208, 1981.
- TASKER, G. D., e J. R. STEDINGER, An operational GLS model for 971 hydrologic regression, *Journal of Hydrology*, 361– 375, 1989.
- VOGEL, R. M., A. Zafirakou-Koulouris, and N. C. Matalas. Frequency of record breaking floods in the United States, *Water Resour. Res.*, 37(6), 1723– 1731, 2001.
- VOGEL, R. M., Regional Calibration of Watershed Models, Chapter 3 in *Watershed Models*, V.P. Singh and D.F. Frevert editors, CRC Press, pp.47-71, 2005.

## *Flood Frequency Estimates at Ungauged Sites as Derived From Probabilistic Regional Envelope Curves*

## ABSTRACT

*Regional envelope curves have been used since the beginning of the 20th century as convenient summaries of the largest flood discharges observed at all gauging stations located in one or more geographical regions.*

*Usually the envelope curve for a specific region relates the upper bound on all observed floods of record to the corresponding drainage areas. Regional envelope curves have been used by practitioners worldwide as a simple and expeditious procedure to provide preliminary estimates of design floods for hydraulic structures, in spite of the fact that they have been usually regarded as deterministic graphical tools and, as such, they were not associated with probabilistic statements. In this paper, we applied and extended the methods introduced by Castellarin et al. (2005) and Castellarin (2007) to construct probabilistic flood envelope curves to the upper São Francisco river basin, located in the the Brazilian state of Minas Gerais.*

**Key-words:** *Envelope curves for floods, estimates of maximum discharges, design floods.*