

O uso da teoria estatística dos recordes e de uma distribuição de probabilidades limitada superiormente na análise de frequência de vazões máximas

Wilson Fernandes¹ & Mauro Naghettini²

RESUMO – O aprofundamento do conhecimento sobre os processos hidrológicos conduz à constatação que algumas variáveis hidrológicas, tais como os máximos anuais de vazão e de precipitação, possuem limites superiores finitos. No entanto, quase todas as distribuições de probabilidades utilizadas para modelar o comportamento probabilístico dessas variáveis são ilimitadas, o que apresenta uma contradição do ponto de vista hidrológico, justificável talvez pela maior facilidade de tratamento matemático ou talvez pela imposição de decisões tomadas nos projetos de engenharia. Este artigo explora o uso da distribuição de probabilidades EV4, limitada superiormente, na análise de frequência de vazões máximas. São analisados dois métodos para a estimação do limite superior: (i) um método determinístico onde o limite é fixado pela PMF (*Probable Maximum Flood*) e (ii) um método estatístico onde o limite é estimado pela denominada equação genérica. Além disso, descrevem-se os princípios da teoria estatística dos recordes, de uso crescente em outros países, porém ainda pouco difundida no Brasil. Tal teoria é utilizada na busca de um modelo probabilístico que permita melhor descrever as maiores enchentes ocorridas em uma região. O estudo foi conduzido com base nos dados da bacia do rio American, próximo ao reservatório de Folsom, no estado americano da Califórnia.

ABSTRACT – Our increased understanding of hydrological processes leads to acknowledge the concept that some hydrological variables, such as streamflow and rainfall maxima, are upper-bounded variables. However, almost all probability distributions that are currently employed in hydrologic practice are unbounded. This hydrologic contradiction may be justified either by the relatively simpler mathematical treatment required by unbounded distributions or by the urgency compelled by the engineering decision-taking process. This paper describes an exploratory study on the use of the upper-bounded EV4 probability distribution in the context of flood frequency analysis. Two methods for estimating the upper-bound are analyzed: (i) a deterministic procedure, according to which the upper-bound is equal to the Probable Maximum Flood (PMF) and (ii) a statistical procedure, in which the upper-bound is estimated by the so-called generic equation. In addition, this paper describes the principles of the statistical theory of records, as formulated on the basis of the EV4 distribution, with the prospect of searching for a model that better synthesizes our current experience of observing the largest floods in a given region. The exploratory study, described herein, has been applied to flood data observed in the American river near the Folsom reservoir, located in the American state of California.

Palavras-chave: Distribuições de probabilidades limitadas, Distribuição EV4, Teoria estatística dos recordes

1) Aluno de doutorado do Departamento de Engenharia Hidráulica e Recursos Hídricos da Universidade Federal de Minas Gerais, Av. Contorno, 842/8 – CEP 30110060 Belo Horizonte MG, 31 3238 1010, wilson@ehr.ufmg.br.

2) Professor associado do Departamento de Engenharia Hidráulica e Recursos Hídricos da Universidade Federal de Minas Gerais, Av. Contorno, 842/8 – CEP 30110060 Belo Horizonte MG, 31 3238 1002, naghet@netuno.lcc.ufmg.br

1 – INTRODUÇÃO

É um fato que algumas variáveis físicas possuem limites superiores claramente definidos, como por exemplo, a concentração de oxigênio dissolvido em um corpo d'água (limite superior entre 9 e 10 mg/l, a depender da temperatura da água). Outras variáveis, tais como a vazão máxima anual ou a altura de chuva diária máxima anual, possuem um limite superior embora não seja conhecido *a priori*. O desconhecimento desse limite se deve ao fato da insuficiente compreensão e quantificação de todos os processos físicos causais envolvidos. A esse respeito, é bastante conhecida a controvérsia quanto à existência da Precipitação Máxima Provável (PMP), originalmente formulada como um limite superior de produção de precipitação por uma coluna de ar atmosférico. Se de fato existe um limite inequívoco para a PMP, sua determinação fica comprometida pela insuficiente quantificação da variabilidade espaço-temporal das variáveis que lhe dão origem.

As distribuições de probabilidades utilizadas em hidrologia estatística, na sua maioria, são predominantemente ilimitadas superiormente. A razão dessa predominância deve-se ao fato que as distribuições limitadas superiormente pressupõem a estimação do limite superior a partir de uma amostra de vazões de enchentes (ou de precipitações máximas), de tamanho relativamente pequeno. Além das dificuldades inerentes à estimação, há também o receio de se prescrever um modelo distributivo que pressupõe um valor da variável aleatória cuja probabilidade de excedência deve ser nula.

Por outro lado, indo além da análise convencional de frequência de cheias vista como um mero expediente de engenharia, os geomorfólogos argumentam que não há razão científica para afirmar que as cheias de um ribeirão ou um rio qualquer não tenham um limite físico e que, portanto, possam tender ao infinito. Essa é uma longa controvérsia que remonta ao início do século XX (Horton, 1936) e que persiste até os dias de hoje.

De acordo com Takara e Tosa (1999), variáveis hidrológicas tais como a vazão ou a precipitação, podem assumir somente valores não-negativos, com um valor máximo fisicamente possível, prescrito por um limite superior finito, e devem ser analisadas sob esses aspectos. Nessa mesma linha de raciocínio, outros pesquisadores, como Boughton (1980) e Laursen (1983), recomendam que somente distribuições limitadas superiormente devam ser usadas para modelar as vazões de cheias ou as precipitações máximas.

Hosking & Wallis (1997) consideram errônea a recomendação de se usar somente distribuições limitadas e sustentam que, se o objetivo da análise de frequência é o de estimar o quantil de tempo de retorno de 100 anos, por exemplo, é irrelevante considerar como “fisicamente impossível” a ocorrência do quantil de 100.000 anos. Acrescentam que impor um limite superior ao modelo probabilístico pode comprometer a obtenção de estimativas confiáveis de quantis para os

tempos de retorno que de fato interessam. Contraponto a essa afirmação, Takara e Tosa (1999) analisaram o efeito da incorporação de limite superior aos dados de chuva e vazão por meio do uso de distribuições limitadas e apresentaram evidências de que, além do fato desses modelos serem fisicamente mais plausíveis, as estimativas dos quantis são mais acuradas, sobretudo para amostras pequenas.

De acordo com Takara et al. (1996), uma vantagem adicional dos modelos limitados decorre de sua aplicação na estimativa de cheias sob condições ou cenários de mudanças climáticas globais. Sob o cenário de aquecimento global, os limites superior e inferior das cheias são mais facilmente estimados sob os pontos de vista climatológico, meteorológico e hidrológico do que sob o ponto de vista puramente matemático ou estatístico. Distribuições limitadas podem incorporar estimativas de limites provenientes de diferentes cenários de mudanças climáticas globais permitindo prever as eventuais alterações na frequência de eventos raros.

Admitindo que as vazões em um determinado curso d'água possuem um limite superior, surge a questão crucial de estimação desse limite. A resposta a essa questão tem sido objeto de estudo desde a década de 1930, quando Horton desenvolveu um método estatístico para estimação de vazões máximas, o qual pressupunha a existência de um limite superior para a vazão (Horton, 1936). No decorrer do século XX, grandes esforços foram realizados na direção da determinação desse limite, sobretudo aqueles referentes ao cálculo da precipitação máxima provável (PMP), e da correspondente cheia máxima provável (PMF, do acrônimo em inglês *Probable Maximum Flood*). Na origem, a PMP foi definida como o limite máximo teórico, fisicamente possível, da precipitação, para uma determinada duração, em uma determinada área geográfica, em uma determinada época do ano (WMO, 1986). A PMF, por sua vez, decorre da transformação da PMP por meio de modelos de transformação chuva-vazão.

Recentemente, Vogel et al. (2001), Castellarin et al. (2007), Castellarin et al. (2005), England (2005) têm se dedicado ao desenvolvimento de metodologias que possibilitam a análise das denominadas curvas envoltórias. As curvas envoltórias refletem o conjunto da experiência de observações dos recordes de vazões registradas em um certo número de bacias hidrográficas. Vazão recorde é definida, neste artigo, como a maior cheia registrada em uma bacia. As curvas envoltórias, as quais são obtidas a partir da relação gráfica entre a cheia recorde e a área de drenagem de vários postos fluviométricos em uma região, permitem avaliar o limite superior das vazões em uma determinada bacia a partir dos recordes observados na região da qual a bacia faz parte. Se por um lado, as curvas envoltórias não permitem a definição exata do limite, uma vez que estão associadas a uma probabilidade de não excedência diferente de zero (Castellarin et al., 2005), por outro, elas permitem definir um valor mínimo para o limite.

Além da aplicação natural do conceito da curva envoltória à estimação do limite superior para as vazões, England (2005) descreve pelo menos 3 outras aplicações: (1) em estudos de cheias máximas em bacias não-monitoradas; (2) em estudos onde se deseja comparar estimativas probabilísticas de vazões de projeto, para vertedores de novas barragens e para aqueles já existentes, e (3) como uma forma de avaliar a pertinência dos valores encontrados para a PMF.

O comportamento da curva envoltória é dado pelos recordes que a compõem. A análise estatística do comportamento dos recordes é denominada teoria dos recordes e foi primeiramente abordada por Chandler (1952). Há uma grande quantidade de métodos para a análise do comportamento das vazões máximas anuais. No entanto, de acordo com Douglas e Vogel (2006), a literatura sobre o comportamento dos recordes ainda é bastante escassa. Ainda de acordo com esses mesmos autores, isso se deve principalmente a dois fatores: o primeiro refere-se ao fato de que a teoria probabilística usada para descrever eventos recordes é muito mais complexa que a teoria usada para descrever os máximos anuais; o segundo está relacionado ao reduzido tamanho das amostras de recordes observados. De fato, enquanto o número de cheias anuais em um determinado posto é igual ao número de anos de dados observados, há, por definição, somente um recorde para cada estação fluviométrica.

Uma revisão da literatura disponível sobre a aplicação da teoria dos recordes em hidrologia mostrou que somente distribuições ilimitadas têm sido utilizadas. No entanto, conforme menção anterior, o uso de distribuições limitadas tem vantagens sob o ponto de vista estatístico e, sobretudo, é mais adequado sob o ponto de vista físico. Nesse contexto, o objetivo deste artigo é o de explorar uma aplicação da teoria dos recordes, com o uso de uma distribuição limitada, na análise de frequência de vazões máximas anuais.

Embora haja um crescente desenvolvimento de técnicas relacionadas ao uso de curvas envoltórias e, por consequência, do uso da teoria dos recordes ao redor do mundo (vide, por exemplo, Mimikou (1984) na Grécia; Kadoya (1992) no Japão), no Brasil, não há registros de aplicações dessa teoria. Nesse sentido, esse artigo tem como objetivo inicial difundir essas teorias e incentivar futuros estudos de curvas envoltórias e da aplicação de distribuições limitadas na análise de frequência de cheias em bacias brasileiras.

2 – UM MODELO DISTRIBUTIVO COM LIMITE SUPERIOR: A DISTRIBUIÇÃO EV4

2.1 – Distribuições limitadas

Botero (2006) analisou o comportamento das vazões máximas anuais com base em três distribuições de probabilidades limitadas superiormente: a distribuição EV4, proposta por Kanda (1981) para analisar o comportamento probabilístico de abalos sísmicos e de ventos extremos no Japão; a distribuição de valores extremos transformada (TDF), proposta por Eliasson (1994), que a

utilizou na análise de frequência de precipitações extremas na Islândia e nos Estados Unidos; e a distribuição Lognormal de 4 parâmetros (LN4), proposta por Takara e Loebis (1996), com base na variável transformada de Slade (Slade, 1936 *apud*. Botero, 2006), empregada na análise de frequência de precipitações extremas no Japão e Indonésia. Botero (2006) desenvolveu uma extensa análise do comportamento dessas distribuições a partir de dados fluviométricos (dados sistemáticos) e paleohidrológicos (dados não sistemáticos) de bacias espanholas e concluiu que a distribuição com melhor capacidade descritiva era a EV4.

Neste artigo, faz-se uma análise da distribuição dos recordes de uma variável aleatória distribuída conforme o modelo EV4. Embora futuros estudos possam mostrar que outras distribuições limitadas produzam melhores resultados do que os publicados por Botero (2006), não é objetivo aqui escolher o melhor ajuste e sim mostrar a capacidade descritiva de um modelo limitado superiormente na análise de recordes de vazão.

2.2 – Formulação da EV4

A função de probabilidades acumulada, a função densidade de probabilidade e a função de quantis de uma variável distribuída de acordo com o modelo EV4 são dadas, respectivamente, pelas seguintes equações:

$$F(x) = \exp \left\{ - \left[\frac{\alpha - x}{\beta(x - \varepsilon)} \right]^\lambda \right\} \quad (1)$$

com $\varepsilon < x \leq \alpha$,

$$f(x) = \frac{\lambda(\alpha - x)^{\lambda-1}(\alpha - \varepsilon)}{\beta^\lambda(x - \varepsilon)^{\lambda+1}} \exp \left\{ - \left[\frac{\alpha - x}{\beta(x - \varepsilon)} \right]^\lambda \right\} \quad (2)$$

$$x = \frac{\alpha + \beta\varepsilon(-\ln F)^{1/\lambda}}{\beta(-\ln F)^{1/\lambda} + 1} \quad (3)$$

onde α e ε são respectivamente os limites superior e inferior de X , β é o parâmetro de escala e λ é o parâmetro de forma.

A denominação EV4, dada por Kanda (1981), decorre do fato de essa tender às distribuições de valores extremos quando a variável X tende aos limites α e ε . Quando X tende ao limite inferior ε , a EV4 tende para uma EV2 (Distribuição de Fréchet) com parâmetro de escala igual a $(\alpha - \varepsilon)/\beta$. Por outro lado, quando X tende ao limite superior α , a EV4 tende a uma EV3 (Distribuição de Weibull) com parâmetro de escala igual a $\beta(\alpha - \varepsilon)$.

Botero (2006), verificou que não é possível obter uma equação de momentos para a EV4 escrita de acordo com a equação 1. No entanto, definindo $\varepsilon = 0$, conforme será visto a seguir, e

fazendo uma mudança de variáveis da forma $Z = 1/X$, pode-se mostrar que a equação de momentos para a EV4 toma a seguinte forma:

$$\mu_z^r = \frac{1}{\alpha^r} \sum_{i=0}^r \binom{r}{i} \cdot \beta^{r-i} \cdot \Gamma\left(\frac{r-i}{\lambda} + 1\right) \quad (4)$$

onde μ_z^r é o r -ésimo momento central da variável Z e $\Gamma(\cdot)$ representa a função gama.

2.3 – Estimação dos parâmetros

A vazão máxima anual pode assumir somente valores não negativos. Desta forma, torna-se intuitivo definir $\varepsilon = 0$ para aplicações da EV4 com dados fluviométricos. A rigor, ε deve ser superior a zero, uma vez que, na prática, a menos que o rio fique seco por mais de um ano, sempre haverá um valor mínimo de vazão maior do que zero. Por outro lado, a prescrição de um limite inferior para as vazões máximas anuais de um rio é controvertida e não há um consenso quanto a uma metodologia a ser empregada para tal. Takara e Tosa (1999) verificaram que os melhores ajustes são obtidos quando $\varepsilon = 0$ e que à medida que o limite inferior se aproxima do mínimo amostral, a EV4 perde a capacidade de descrever o comportamento das variáveis máximas anuais. Por essa razão e por maior conveniência do ponto de vista matemático, nas aplicações que se seguem, será fixado o limite inferior em zero.

Botero (2006) lista três possibilidades para a estimação dos parâmetros da EV4:

1. Estimação pelo método de máxima verossimilhança

Neste caso, trata-se de encontrar, por algum método numérico, os parâmetros que maximizam a função de verossimilhança dada a seguir.

$$L(\lambda, \beta, \varepsilon, \alpha) = n \ln(\lambda) + n \ln(\alpha - \varepsilon) - n\lambda \ln(\beta) + \\ + (\lambda - 1) \sum_{i=1}^n \ln(\alpha - x_i) - (\lambda + 1) \sum_{i=1}^n \ln(x_i - \varepsilon) - \sum_{i=1}^n \left[\frac{\alpha - x_i}{\beta(x_i - \varepsilon)} \right]^\lambda \quad (5)$$

Botero (2006) verificou que o valor de α que maximiza a função de verossimilhança é dado pelo máximo amostral. Portanto, em se empregando o método de estimação por máxima verossimilhança, não se espera que o valor encontrado para α permita estudar o limite superior da variável.

2. Atribuição de um valor extremo ao parâmetro α

Neste caso, pode-se atribuir ao limite superior o valor da PMF que é, por suposição, o maior valor possível para as vazões em um determinado local. No entanto, o método de estimação da PMF está longe de ser um consenso entre os pesquisadores e, assim, a estimativa de α por meio desse valor, certamente não irá refletir um valor inequívoco para o limite superior. Apesar disso, Botero (2006) verificou que, na falta de informações paleohidrológicas, uma estimativa plausível para o limite superior pode ser a PMF.

Definido o valor de α , pode-se encontrar os demais parâmetros por meio do método da máxima verossimilhança.

3. Equação genérica

Kijko (2004) desenvolveu um método para a estimação da maior magnitude possível de um terremoto, o equivalente à PMF para as vazões de um rio, baseado na distribuição do recorde (maior valor) de uma variável. Tal método, denominado de equação genérica, é desenvolvido da forma descrita a seguir.

Seja uma variável X pertencente ao intervalo $[x_{\min}, x_{\max}]$ cuja distribuição acumulada de probabilidade seja dada por $F_X(x)$. Dada uma amostra com n valores, classificada do menor para o maior valor, sendo X_n o valor correspondente ao máximo observado, a distribuição de X_n , conforme será visto no próximo item, é dada por:

$$F_{X_n}(x) = \begin{cases} 0 & \text{para } x < x_{\min} \\ [F_X(x)]^n & \text{para } x_{\min} \leq x \leq x_{\max} \\ 1 & \text{para } x > x_{\max} \end{cases} \quad (6)$$

O valor esperado de X_n é

$$E[X_n] = \int_{x_{\min}}^{x_{\max}} x \cdot d[F_X(x)]^n \quad (7)$$

Integrando por partes, tem-se

$$E[X_n] = x_{\max} - \int_{x_{\min}}^{x_{\max}} [F_X(x)]^n dx \quad (8)$$

Partindo do pressuposto que o melhor estimador para $E[X_n]$ é o máximo valor observado, Kijko (2004) demonstrou que

$$x_{\max} = x_{\max, obs} + \int_{x_{\min}}^{x_{\max}} [F_X(x)]^n dx \quad (9)$$

A estimação dos parâmetros utilizando a equação genérica é feita iterativamente, conforme as seguintes etapas:

- estima-se todos os parâmetros pelo método de máxima verossimilhança, com $g_0 = \text{máximo observado}$. Com isso, tem-se $\{k_i, v_i, a_i, g_i = \text{máx.obs.}\}$
- com os parâmetros $\{k_i, v_i, a_i, g_i\}$ estima-se o valor de g_{i+1} pela equação genérica;
- com o valor de g_{i+1} estima-se os novos parâmetros $\{k_{i+1}, v_{i+1}, a_{i+1}\}$ pelo método de máxima verossimilhança; e
- repete-se os passos b e c até que se tenha $g_{i+1} - g_i \leq \delta$, onde δ é um número pequeno (10^{-4} , por exemplo).

Além dos métodos propostos por Botero (2006), pode-se estimar os parâmetros pelo método dos momentos. O método dos momentos é bastante conhecido e consiste em igualar os k primeiros momentos populacionais aos seus respectivos estimadores amostrais. No caso da EV4, partindo-se da equação 4, tem-se:

$$\frac{\beta}{\alpha} g_1 + \frac{1}{\alpha} = \hat{\mu}_z \quad (10)$$

$$\frac{\beta}{\alpha} \cdot \sqrt{g_2 - g_1^2} = \hat{\sigma}_z \quad (11)$$

$$\frac{g_3 - 3g_2g_1 + 2g_1^3}{\sqrt{(g_2 - g_1^2)^3}} = \hat{\gamma}_z \quad (12)$$

onde $\hat{\mu}_z$, $\hat{\sigma}_z$ e $\hat{\gamma}_z$ são, respectivamente, a média, o desvio padrão e o coeficiente de assimetria amostral da variável $Z = 1/X$ e $g_r = \Gamma\left(\frac{r}{\lambda} + 1\right)$.

As soluções do sistema formado pelas equações 10, 11 e 12 fornecem as estimativas dos parâmetros da EV4. Vale salientar, entretanto, que a solução do sistema não oferece uma boa alternativa para a estimação do parâmetro α uma vez que o limite superior tende a ser o máximo amostral.

3 – TEORIA DOS RECORDES

A literatura sobre teoria dos recordes e sua aplicação em hidrologia estatística vem aumentando a partir da década de 90. Conforme apontado por Douglas e Vogel (2006), já há uma vasta gama de estudos sobre o comportamento, métodos de estimação e inferência estatística dos recordes provenientes das mais variadas distribuições de probabilidades. No entanto, não há aplicação da teoria dos recordes para distribuições limitadas. Os princípios da teoria dos recordes podem ser sumariados do modo como se descreve a seguir.

Seja uma variável aleatória X com função acumulada de probabilidade dada por $F_X(x)$. O máximo de uma série independente $X_i, i = 1, \dots, n$ é

$$Y_{\max} = \max(X_1, X_2, \dots, X_n) \quad (13)$$

A função acumulada de probabilidades de $Y_{\max}$ é dada por

$$\begin{aligned} F_Y(y) &= P(Y_{\max} \leq y) \\ &= P(X_1 \leq y, X_2 \leq y, \dots, X_n \leq y) \\ &= P(X_1 \leq y)P(X_2 \leq y) \dots P(X_n \leq y) \\ &= F_1(y)F_2(y) \dots F_n(y) \\ &= [F_X(y)]^n \end{aligned} \quad (14)$$

Da equação 14 pode-se verificar que a distribuição do valor máximo de uma amostra de n elementos independentes, depende da distribuição-mãe de X e do tamanho da amostra. Se X se distribui de acordo com uma EV4, tem-se:

$$\begin{aligned} F_Y(y) &= \left[\exp \left\{ - \left[\frac{\alpha - y}{\beta(y - \varepsilon)} \right]^\lambda \right\} \right]^n \\ &= \exp \left\{ - n \left[\frac{\alpha - y}{\beta(y - \varepsilon)} \right]^\lambda \right\} \end{aligned} \quad (15)$$

Desta forma, se X segue uma distribuição EV4, o recorde também seguirá uma EV4 com parâmetro de escala dado por $\beta n^{-1/\lambda}$. Pode-se mostrar que a equação de momentos para a variável $1/Y$ é dada por:

$$\mu_{1/Y}^r = \frac{1}{\alpha^r} \sum_{i=0}^r \binom{r}{i} \cdot \frac{\beta^{r-i}}{n^{\frac{r-i}{\lambda}}} \cdot \Gamma \left(\frac{r-i}{\lambda} + 1 \right) \quad (16)$$

4 – APLICAÇÃO

A combinação da teoria de recordes e de distribuições limitadas superiormente foi aplicada aos dados fluviométricos do rio American próximo ao reservatório de Folsom, localizado no estado americano da Califórnia. Essa bacia foi escolhida pela disponibilidade de dados sistemáticos e não sistemáticos, das estimativas de PMF e de diversos estudos abordando tanto as vazões máximas anuais quanto as vazões recordes. Além disso, este estudo está inserido em um projeto de pesquisa no Departamento de Engenharia Hidráulica e Recursos Hídricos da UFMG, o qual explora a análise de frequência com base em métodos bayesianos que, por sua vez, fazem uso de dados não sistemáticos. No Brasil, a disponibilidade de dados não sistemáticos é praticamente nula, o que impossibilitaria a aplicação da pesquisa em desenvolvimento.

De acordo com England (2002), a bacia do rio American, no eixo da barragem de Folsom, drena uma área de cerca de 4820 km². O reservatório de Folsom está localizado imediatamente a montante da cidade de Sacramento e foi construído no início da década de 50. A capacidade de armazenamento é de 1245 hm³ e seu vertedor foi dimensionado para uma cheia máxima de 16990 m³/s. A figura 1 mostra a localização da bacia.

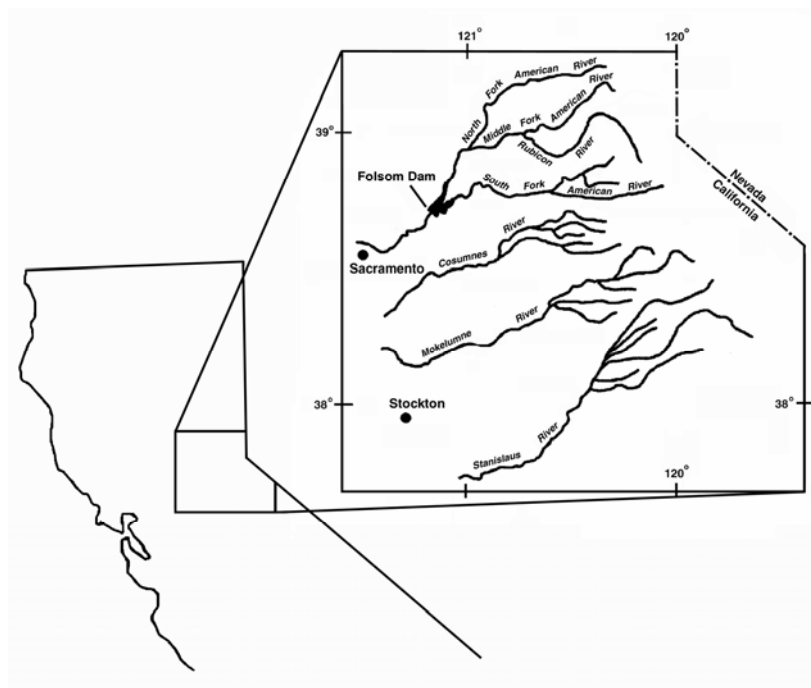


Figura 1 – Localização da bacia do rio American no estado da Califórnia (fonte: England, 2002)

O monitoramento de vazões é feito pelo USGS (United State Geological Survey) através da estação 11446500 (American River at Fair Oaks), localizada a jusante do reservatório de Folsom com área de drenagem de 4890 km². A estação 11446500 possui dados de 1904 até a data atual. No entanto, uma vez que a barragem de Folsom foi construída em 1955, foram utilizados somente os dados até a sua construção, de forma a não inserir dados afetados pela regularização de vazão proporcionada pelo reservatório. Assim, foram utilizados 51 dados de vazão máxima instantânea anual na análise. A PMF estimada para a bacia é de 17.414 m³/s. A figura 2 ilustra a variabilidade das vazões naturais máximas anuais da estação referenciada.

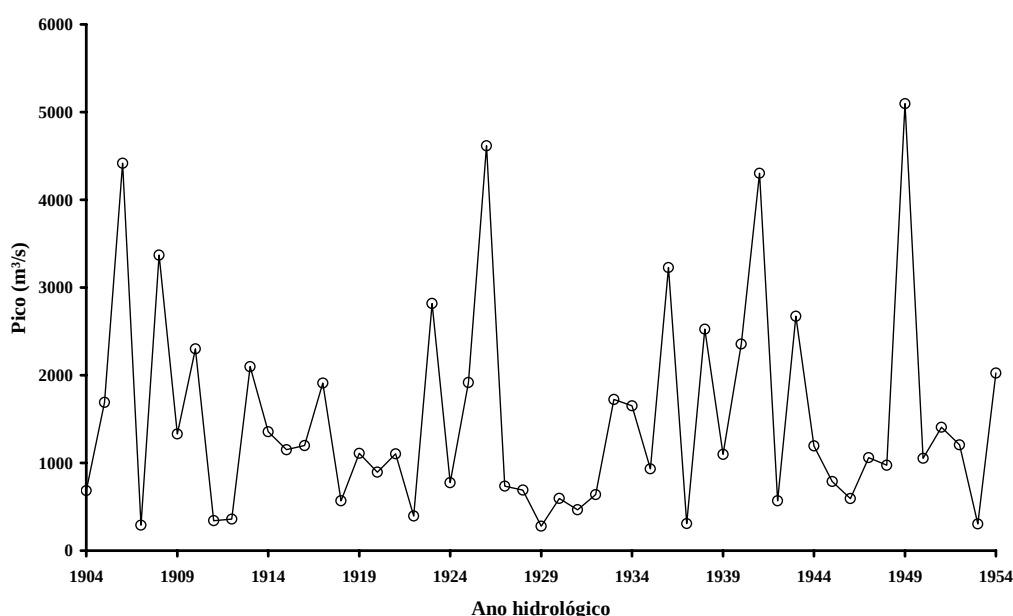


Figura 2 – Vazões máximas anuais da estação 11446500 (American River at Fair Oaks)

Os parâmetros da EV4 foram calculados por dois métodos:

- Método 1 – fixou-se o limite superior pela PMF e os demais parâmetros foram estimados pelo método da máxima verossimilhança;
- Método 2 – os parâmetros foram estimados pelo método da equação genérica.

A tabela a seguir mostra os resultados encontrados.

Tabela 1 – Parâmetros da EV4 para os dados da estação 11446500

Parâmetro	Método 1	Método 2
α	17414	5829
β	21,1818	6,0393
λ	1,3031	1,0911

A figura 3 mostra os ajustes de acordo com os dois métodos de estimação. A grande diferença dos quantis obtidos pelos dois métodos decorre do fato de a PMF estimada para o local ser bastante superior ao máximo observado, alterando substancialmente a cauda superior em direção ao limite. Embora o ajuste pela equação genérica tenha se mostrado mais adequado, o ajuste pelo método da máxima verossimilhança, com o limite superior fixado na PMF, parece ser fisicamente mais plausível pelas razões a serem apontadas a seguir.

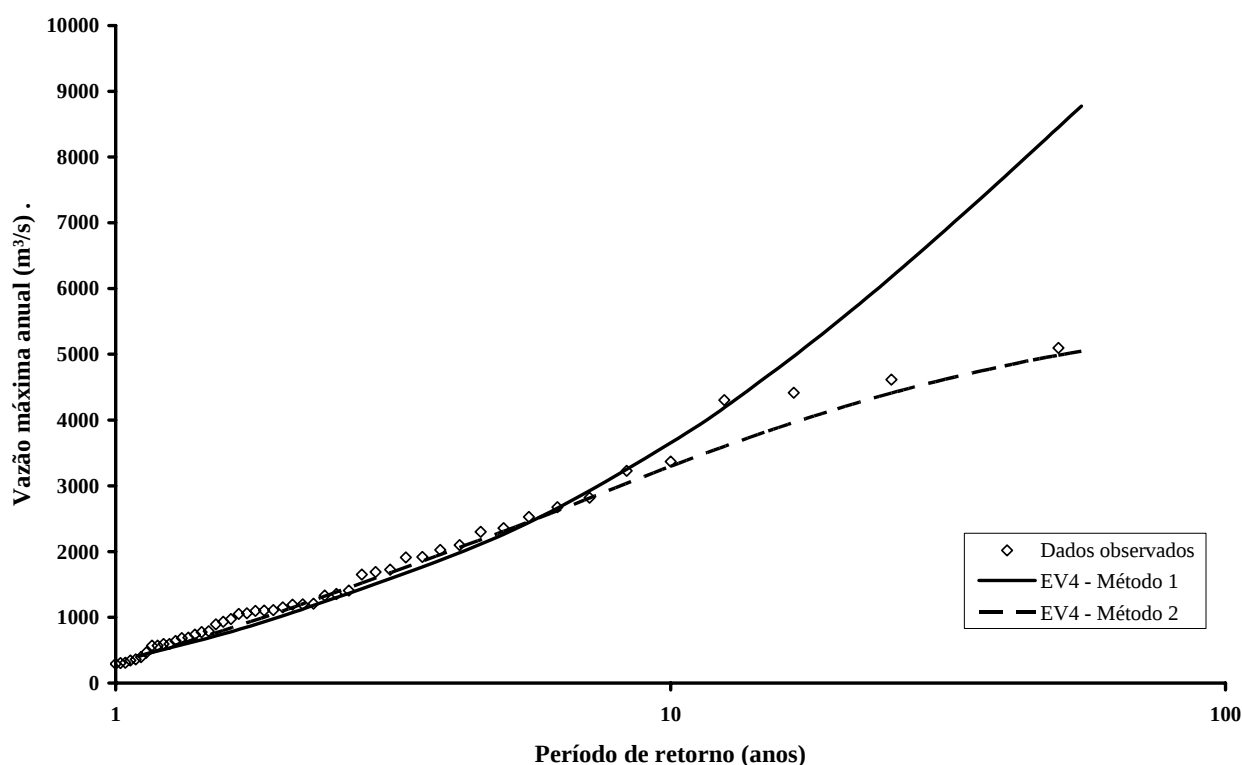


Figura 3 – Ajuste da EV4 pela equação genérica e pelo método da máxima verossimilhança

A análise do recorde amostral pode ser feita diretamente pela aplicação da equação 15. O recorde amostral é 5097 m³/s. A probabilidade de excedência desse recorde pode ser calculada pela seguinte equação:

$$p_n = 1 - F_Y(y) \quad (17)$$

Pelos dois métodos utilizados tem-se:

$$p_n^{\text{método 1}} = 0,95$$

$$p_n^{\text{método 2}} = 0,58$$

O valor p_n dado pelo método 1 está condizente com o senso comum de que a ocorrência de uma cheia superior às cheias já ocorridas é bastante provável. Por outro lado, o valor calculado pelo método 2 está mais condizente com os valores encontrados por Douglas e Vogel (2006).

Douglas e Vogel (2006) analisaram o comportamento das cheias recordes de 1490 bacias americanas tendo como referência a distribuição Generalizada de Valores Extremos (GEV). De acordo com os autores, se as cheias recordes são independentes espacialmente, então os valores teóricos de p_n devem ter mediana 0,50 e quartis inferiores e superiores em 0,25 e 0,75, respectivamente. A mediana dos valores p_n dos recordes das bacias analisadas foi de 0,45. A diferença em relação ao valor esperado foi atribuída à possível correlação espacial entre os recordes. Tal fato é plausível, uma vez que cheias extremas tendem a ser causadas por eventos chuvosos frontais, que afetam várias bacias ao mesmo tempo. O histograma da figura 4 ilustra os resultados encontrados por Douglas e Vogel (2006).

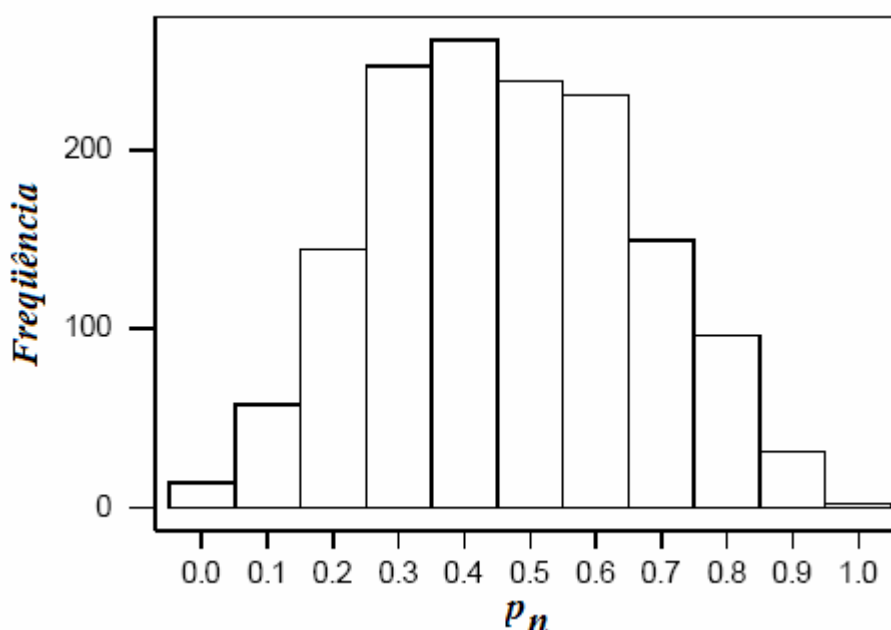


Figura 4 – Histograma de valores p_n para bacias americanas (*adap.* Douglas e Vogel, 2006)

Como pode ser verificado o valor encontrado pelo método 1 é pouco freqüente para as bacias americanas. Já o valor encontrado pelo método 2 apresenta uma maior freqüência, embora esteja longe do valor médio de 0,45 obtido com o modelo distributivo GEV. Douglas e Vogel (2006) analisaram, ainda, o comportamento dos recordes quando a vazão máxima anual segue as

distribuições Wakeby, Log-Pearson tipo III e Lognormal de três parâmetros e encontraram resultados semelhantes aos obtidos para o modelo GEV.

A capacidade do modelo dado pela equação 15 de descrever os recordes provenientes de uma distribuição EV4 foi verificada por meio do método de Monte Carlo. Para tal, foram geradas 1000 séries aleatórias de vazões máximas anuais de tamanho 51 a partir dos parâmetros da tabela 1 e da equação 3. Para cada série gerada, foi selecionado o recorde, o que deu origem a duas amostras de recordes simulados de tamanho 1000. Na sequência foi atribuída uma probabilidade empírica para cada recorde simulado, com base na posição de plotagem de Gringorten e foi estimado o recorde correspondente a essa probabilidade pela inversa da equação 15. As figuras 5 e 6 mostram os resultados encontrados para cada conjunto de parâmetros.

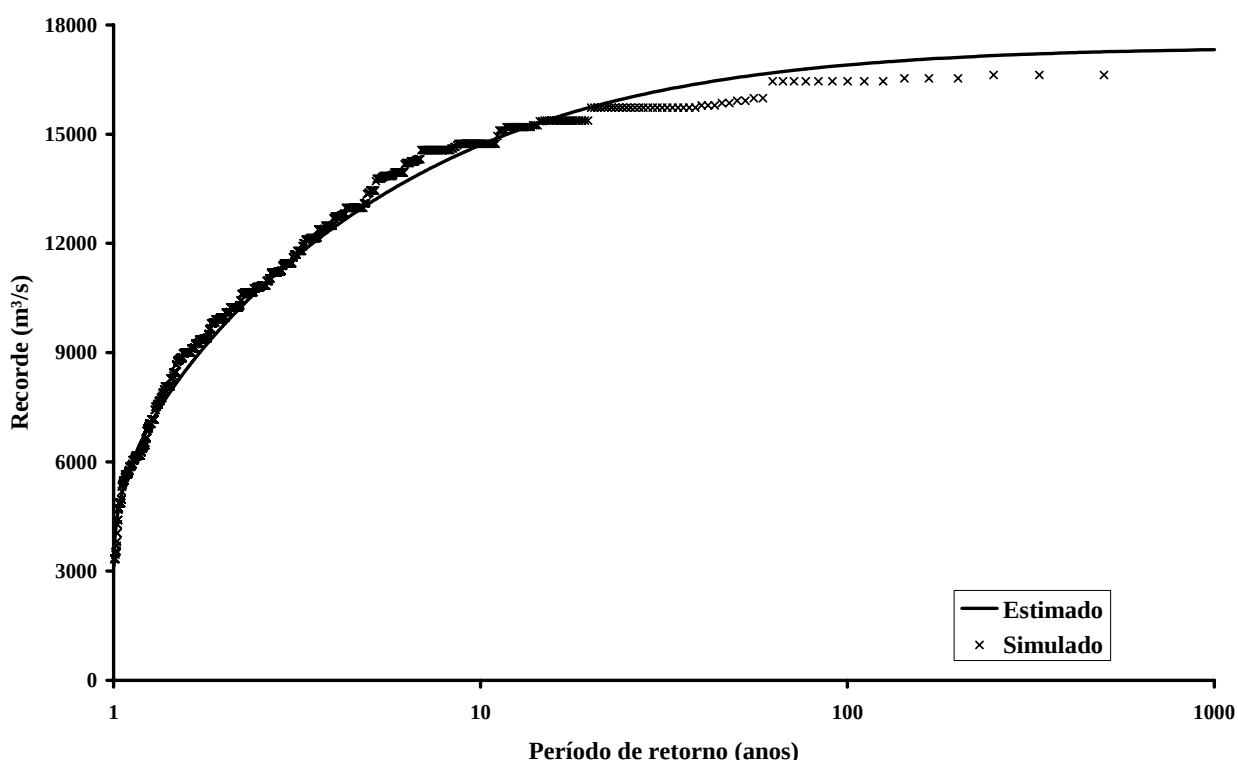


Figura 5 – Recordes simulados e estimados com os parâmetros do método 1.

Independentemente do método de estimação de parâmetros utilizado, verifica-se o bom ajuste do modelo dado pela equação 15.

Com a finalidade de se avaliar a influência do tamanho da amostra procedeu-se a uma nova aplicação do método de Monte Carlo. A partir dos parâmetros obtidos pelo método 1, foram geradas 1000 séries de tamanhos 10, 50, 100, 250, 500 e 1000. Da mesma forma como realizada para a série de tamanho 51, definiu-se a probabilidade empírica para cada recorde e estimou-se o recorde por meio da equação 15. A figura 7 ilustra os resultados encontrados. Por questão de clareza e espaço, não são mostrados os recordes simulados.

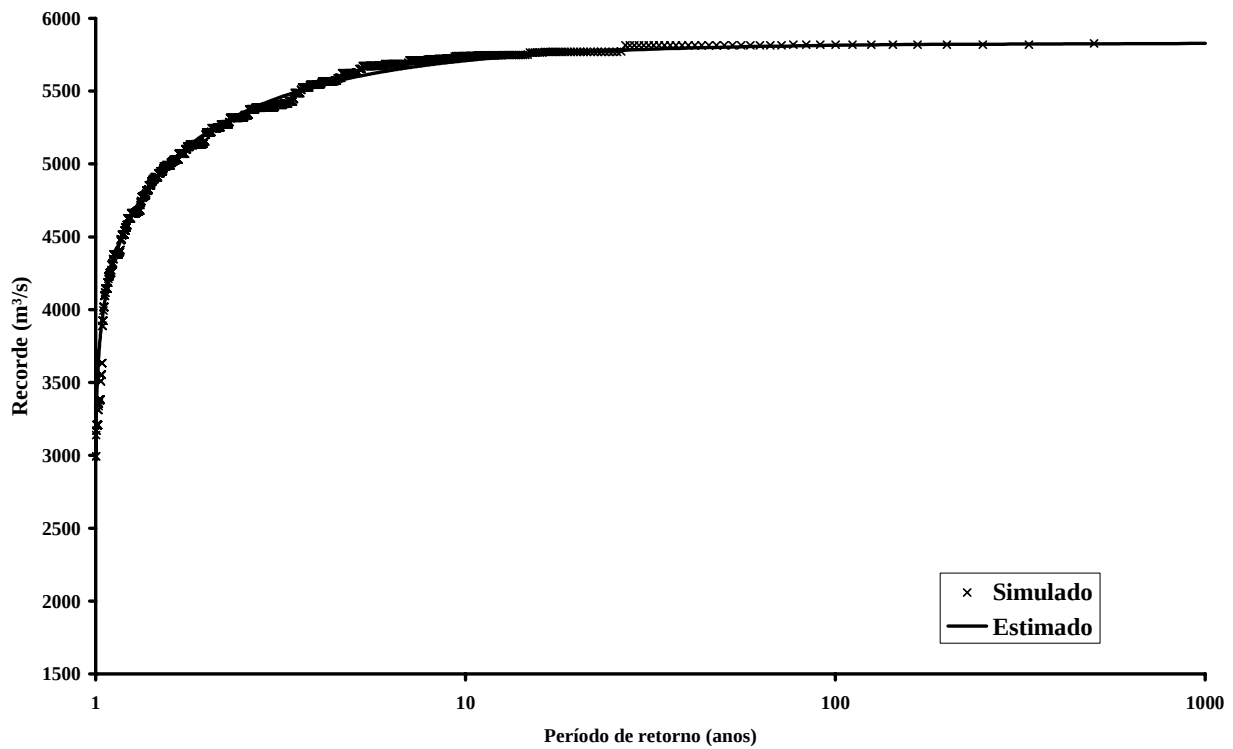


Figura 6 – Registros simulados e estimados com os parâmetros do método 2.

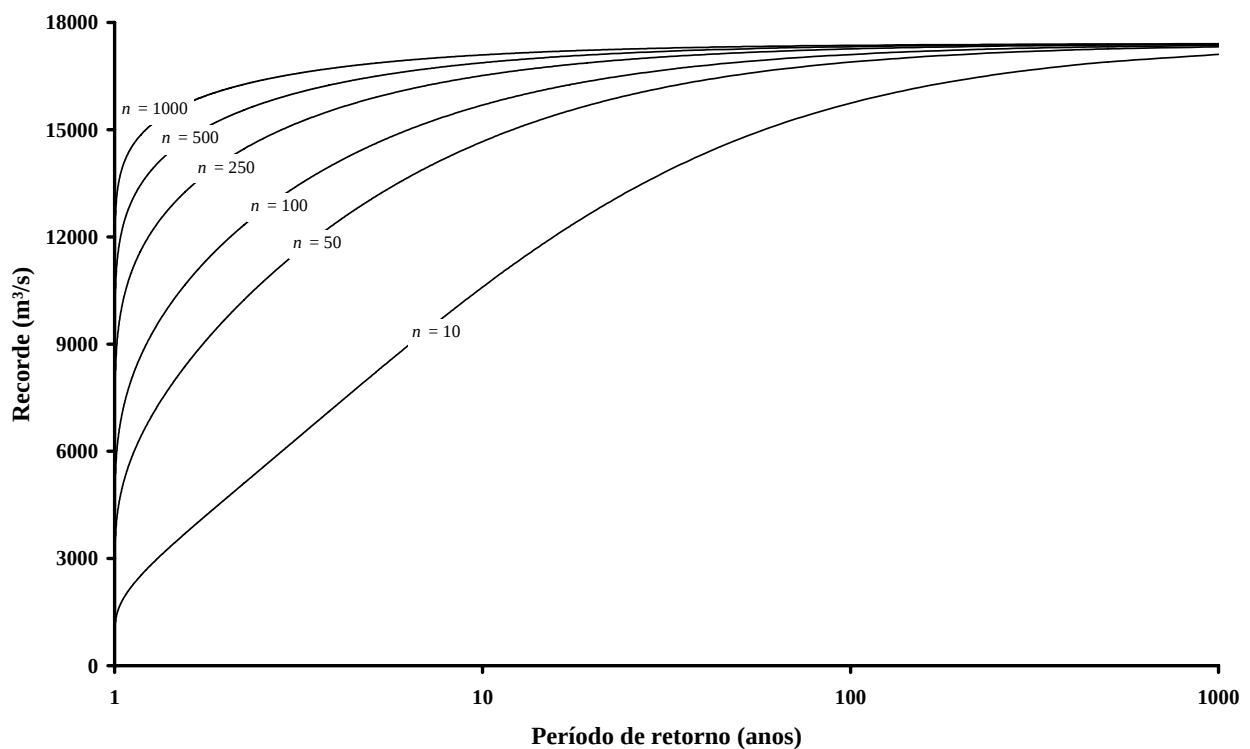


Figura 7 – Registros simulados e estimados com os parâmetros do método 1, para amostras de tamanhos variados.

O ajuste em cada caso foi bastante semelhante àquele encontrado para amostras de tamanho 51. Em geral, a qualidade do ajuste aumenta com o tamanho da amostra. O comportamento exibido

na figura 7 já era esperado, uma vez que com o aumento do tamanho da amostra, cresce a probabilidade de que recordes maiores sejam obtidos em cada série. Isso explica o fato de a distribuição para $n = 1000$ tender mais rapidamente para o limite de que a distribuição para $n = 10$. A mesma simulação foi realizada para os parâmetros do método 2 (não mostrada aqui) e os resultados confirmaram aqueles obtidos para o método 1.

5 – CONCLUSÕES

Esse artigo apresentou uma aplicação da teoria dos recordes aplicada a uma população limitada superiormente. A teoria dos recordes vem sendo utilizada com bastante frequência em várias partes do mundo. No entanto, no Brasil esse método ainda não foi aplicado. Nesse sentido, esse artigo teve como objetivo difundir a metodologia e incentivar estudos futuros, sobretudo aqueles ligados às curvas envoltórias as quais fazem uso dos princípios da referida teoria.

Por outro lado, uma revisão bibliográfica mostrou que a teoria dos recordes tem sido aplicada somente com distribuições de probabilidades ilimitadas. Do ponto de vista hidrológico, é bastante plausível o argumento pelo uso de distribuições limitadas. Variáveis hidrológicas tais como a precipitação e a vazão máxima anual certamente possuem um limite superior e esse fato deve ser levado em consideração na análise estatística. Do ponto de vista prático, verifica-se que as distribuições limitadas são mais complicadas que as distribuições ilimitadas e apresentam mais problemas numéricos quando da estimação de seus parâmetros. Esse fato explica, parcialmente, o uso preferencial das distribuições ilimitadas atualmente.

A despeito dos problemas das distribuições limitadas, apresentou-se a distribuição EV4, com limite inferior fixado em zero e limite superior indeterminado. Foram apresentados, também, quatro métodos de estimação de seus parâmetros com destaque para o método da máxima verossimilhança com o limite superior pré-fixado e para o método da equação genérica. Neste estudo exploratório, a EV4 demonstrou possuir capacidade descritiva na modelagem de eventos extremos de vazão e potencial de uso quando o enfoque contempla as distribuições de probabilidade com limite superior. Este estudo exploratório apresenta os resultados preliminares de uma investigação em andamento sobre o emprego conjunto de análise bayesiana, distribuições limitadas superiormente e de informações históricas e paleohidrológicas sobre cheias extremas.

BIBLIOGRAFIA

BOTERO, B. A. (2006). *Estimación de crecidas de alto período de retorno mediante funciones de distribución com limite superior e información no sistemática*, tese de doutorado apresentada ao Departamento de Ingeniería Hidráulica y Medio Ambiente da Universidad Politécnica de Valencia, Espanha.

- BOUGHTON, W. C. (1980). "A frequency distribution for annual floods". Water Resources Research, v. 16, pp. 347-354.
- CASTELLARIN, A.; VOGEL, R. M.; MATALAS, N. C. (2005). "Probabilistic behavior of a regional envelop curve". Water Resources Research, v. 41, W06018. doi: 10.1029/2004WR003042.
- CASTELLARIN, A.; VOGEL, R. M.; MATALAS, N. C. (2007). "Multivariate probabilistic regional envelopes of extreme floods". Journal of Hydrology, n. 336, pp. 376-390.
- CHANDLER, K. N. (1952). "The distribution and frequency of record values". J. R. Stat. Soc., B 14, pp. 220-228.
- DOUGLAS, E. M.; VOGEL, R. M. (2006). "Probabilistic behavior of floods of Record in the United States". Journal of Hydrologic Engineering, v. 11, pp. 482-488, doi: 10.1061/(ASCE)1084-0699(2006)11:5(482).
- ELÍASSON, J. (1994). "Statistical estimates of PMP values". Nordic Hydrology, v. 25, pp. 301-312.
- ENGLAND, J. F. (2002). *Flood hazard analysis: Folsom dam Central Valley Project California*. Dam Safety Office, Bureau of Reclamation, 128 p.
- ENGLAND, J. F. (2005). *Envelope curve probabilities for dam safety*, U.S. Department of the Interior, Bureau of Reclamation, 21 p.
- HORTON, R. E. (1936). "Hydrologic conditions as affecting the results of the application of methods of frequency analysis to flood records". U.S. Geological Survey Water-Supply Papers, n. 771, pp. 433-449.
- HOSKING, J. R. M.; WALLIS, J. R. (1997). *Regional Frequency Analysis - an approach based on L-moments*. Cambridge University Press, 1st Edition, Cambridge University Press, 222 p.
- KADOYA, M. (1992). "Study on record flood peaks in Japan", Proc. Jpn. Acad., Ser. B, n. 68, pp. 133-138.
- KANDA, J. (1981). "A new value distribution with lowers and upper limits for earthquake motions and wind speeds". Theoretical and Applied Mechanics, v. 31, University of Tokyo Press, pp. 351-360.
- KIJKO, A. (2004). "Estimation of the maximum earthquake magnitude, m_{max} ". Pure and Appl. Geophys., v. 161, pp. 1655-1681, doi: 10.1007/S00024-004-2531-4.
- LAURSEN, E. M. (1983). Comment on "Paleohydrology of southwestern Texas" by KOCHER, R. C., BAKER, V. R., PATTON, P. C. Water Resources Research, v. 19, pp. 1339.
- MIMIKOU, M. (1984). "Envelope curves for extreme flood events in northwestern and western Greece". Journal of Hydrology., n. 67, pp. 55-66.

TAKARA, K.; TOSA, K. (1999). “*Storm and flood frequency analysis using PMP/PMF estimates*” in Proceedings of International Symposium on Floods and Droughts, Nanjing, Out. 1999, China, pp. 7-17.

TAKARA, K.; LOEBIS, J. (1996). “*Frequency analysis introducing probable maximum hydrologic events: Preliminary studies in Japan and in Indonesia*” in LOEBIS, J., editor, Proceedings of international symposium on comparative research on hydrology and water resources in Southeast Asia and the Pacific, Indonesian National Committee for International Hydrological Programme, pp. 67-76.

TAKARA, K.; TAKASAO, T.; TOMOSUGI, K. (1996). “*Possibility and necessity of paradigm shift in hydrologic frequency analysis*” in Proceedings of International Conference on Water Resources and Environment Research, Kyoto, Japan, v. 1, pp. 435-442.

VOGEL, R. M.; ZAFIRAKOU-KOULOURIS, A.; MATALAS, N. C. (2001). “*Frequency of record-breaking floods in the United States*”. Water Resources Research, v. 37, n. 6, pp. 1723-1731.

WMO – World Meteorological Organization (1986). *Manual for Estimation of Probable Maximum Precipitation*, Operational Hydrologic Report No. 1, WMO No. 332, 2nd edition, Geneva, 270 p.