

# ANÁLISE BAYESIANA NO ESTUDO DO TEMPO DE RETORNO DAS PRECIPITAÇÕES PLUVIAIS MÁXIMAS EM JABOTICABAL (SP)<sup>1</sup>

Bayesian analysis for estimating the return period of maximum precipitation at Jaboticabal São Paulo state, Brazil

Luiz Alberto Beijo<sup>2</sup>, Mário Javier Ferrua Vivanco<sup>3</sup>, Joel Augusto Muniz<sup>4</sup>

## RESUMO

Dados históricos de precipitação máxima são utilizados para realizar previsões de chuvas extremas, cujo conhecimento é de grande importância na elaboração de projetos agrícolas e de engenharia hidráulica. A distribuição generalizada de valores extremos (GEV) tem sido aplicada com frequência nesses tipos de estudos, porém, algumas dificuldades na obtenção de estimativas confiáveis sobre alguma medida dos dados têm ocorrido devido ao fato de que, na maioria das situações, tem-se uma quantidade escassa de dados. Uma alternativa para obter melhorias na qualidade das estimativas seria utilizar informações dos especialistas de determinada área em estudo. Sendo assim, objetiva-se neste trabalho analisar a aplicação da Inferência Bayesiana com uma distribuição *a priori* baseada em quantis extremos, que facilite a incorporação dos conhecimentos fornecidos por especialistas, para obter as estimativas de precipitação máxima para os tempos de retorno de 10 e 20 anos e seus respectivos limites superiores de 95%, para o período anual e para os meses da estação chuvosa em Jaboticabal (SP). A técnica Monte Carlo, via Cadeias de Markov (MCMC), foi empregada para inferência *a posteriori* de cada parâmetro. A metodologia Bayesiana apresentou resultados mais acurados e precisos, tanto na estimação dos parâmetros da distribuição GEV, como na obtenção dos valores de precipitação máxima provável para a região de Jaboticabal, apresentando-se como uma boa alternativa na incorporação de conhecimentos *a priori* no estudo de dados extremos.

**Termos para indexação:** Inferência bayesiana, distribuição *a priori*, conhecimento *a priori*, quantis extremos, técnica Monte Carlo via Cadeias de Markov, precipitação pluvial máxima.

## ABSTRACT

Historical maximum rainfall data are used to forecast extreme rainfall, which is important to elaborate agricultural and hydraulic engineering projects. Generalized Extreme Value Distribution (GEV) has been applied in such type of studies. Since those values are extracted from the upper (or lower) tail of the original distribution, a scarce amount of data is obtained in most cases, which may be a problem acquiring reliable estimates about some measure of interest. An alternative to overcome this potential problem would be the use of information available from experts in the area. Therefore, this paper intended to analyze the application of the Bayesian Inference using *a priori* distribution based on extreme quantiles, which facilitates the incorporation of the information supplied by the experts in order to determine the punctual and the 95% upper limit estimates of the probable maximum precipitation for return periods of 10 and 20 years, yearly and monthly in Jaboticabal, São Paulo State, Brazil. Markov Chain Monte Carlo (MCMC) methods were used to a posterior inference of each parameter. Bayesian inference yielded more suitable and accurate results in the estimation of the parameters of the GEV distribution as well as in the determination of the values of the probable maximum precipitation estimates for Jaboticabal. It turned out as an interesting way of incorporating prior knowledge to the study of extreme data.

**Index terms:** Bayesian Inference, prior distribution, prior knowledge, extreme quantiles Markov Chain Monte Carlo method, Maximum rainfall.

(Recebido em 27 de abril de 2006 e aprovado em 26 de março de 2008)

## INTRODUÇÃO

A chuva é um fenômeno natural de extrema importância para a sobrevivência humana. Na agricultura, por exemplo, está diretamente associada desde a germinação das sementes até o momento de colheita de uma determinada cultura. Entretanto, quando ocorrem

chuvas intensas, que são aquelas que registram um grande volume de água precipitado, em um curto espaço de tempo (TUCCI, 2001), seus efeitos, geralmente, passam a ser danosos. Na agricultura, podem causar erosões dos solos e alagamentos, podendo provocar a perda de plantações inteiras, enquanto que, em outros casos, podem causar enchentes, deslizamentos, rompimentos de diques e

<sup>1</sup>Parte da tese do primeiro autor apresentada à Universidade Federal de Lavras – Lavras – MG.

<sup>2</sup>Matemático, Doutor em Estatística e Experimentação Agropecuária – Departamento de Ciências Exatas/DCE – Universidade Federal de Alfenas/UNIFAL – Rua Gabriel Monteiro da Silva, 714 – Centro – 37130-000 – Alfenas, MG – luizbeijo@unifal-mg.edu.br

<sup>3</sup>Estatístico, Doutor em Estatística e Experimentação Agropecuária, Professor Titular – Departamento de Ciências Exatas/DEX – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – 37200-000 – Lavras, MG – ferrua@ufla.br

<sup>4</sup>Engenheiro Agrônomo, Doutor em Estatística e Experimentação Agropecuária, Professor Adjunto – Departamento de Ciências Exatas/DEX – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – 37200-000 – Lavras, MG – joamuniz@ufla.br

represas, soterramentos entre outros, levando, em alguns casos, à perdas de vidas humanas.

O fato dos projetos hidráulicos em geral serem concebidos considerando o custo mínimo, associado a um risco admissível de falha, requer a previsão de grandezas hidrológicas de grande magnitude, tais como máximas vazões ou precipitações, que podem vir a ocorrer em certa localidade. Assim, as séries de máximos valores são empregadas para ajuste, segundo a lei probabilística que melhor descreva o processo, possibilitando extrapolações (VIEIRA et al., 1991).

O conhecimento da precipitação máxima provável, que é definida como a quantidade de precipitação considerada como o limite superior estatístico (físico) numa dada bacia, para uma determinada duração, é de extrema importância para o dimensionamento adequado de trabalhos de conservação do solo, estradas, barragens e drenagem, entre outros. Muitas organizações vinculadas à segurança de barragens recomendam, explicitamente, o estudo de precipitações máximas para o caso de grandes obras, visto que, elas são projetadas para suportar o limite máximo conhecido dos eventos meteorológicos da região (TUCCI, 2001).

Jenkinson (1955) propôs que os três tipos de distribuições de valores extremos (Gumbel, de Fréchet e de Weibull) fossem representados numa forma paramétrica única, designada por *distribuição generalizada de valores extremos* (GEV). A distribuição GEV tem sido utilizada com grande frequência em estudos de fenômenos ambientais, principalmente para solucionar problemas relacionados à áreas de Engenharia, entre os quais, o estudo da precipitação pluvial máxima. Desde 1975, a *Natural Environmental Research Council* (NERC) recomenda a distribuição GEV, para a análise de frequências de enchentes no Reino (MARTINS & STEDINGER, 2000).

A partir dos valores extremos de amostras, vários métodos podem ser utilizados para obtenção dos estimadores dos parâmetros dessa distribuição, entre os quais, o mais aplicado tem sido o método da máxima verossimilhança. Segundo Smith (1985), os métodos baseados em verossimilhança são preferidos devido à teoria dos estimadores de máxima verossimilhança ser bem compreendida e as inferências serem facilmente modificadas ao incorporar modelos com estruturas mais complexas. Entretanto, as condições de regularidade, para estimação pelo método da máxima verossimilhança não são necessariamente satisfeitas; quando  $-1 \leq \xi \leq -0,5$ , os estimadores de máxima verossimilhança existem, mas não satisfazem às condições de regularidades e quando  $\xi < -1$ , os estimadores de máxima verossimilhança não existem.

Segundo Coles & Dixon (1999), um argumento que tem sido usado contra os estimadores de máxima verossimilhança se refere às suas propriedades, em pequenas amostras. E, como em análises de precipitações máximas, na maioria dos casos, se dispõe de pequenas séries de dados, faz-se necessária a aplicação de uma metodologia que possibilite a otimização do uso dos dados disponíveis.

De acordo com Paulino et al. (2003), uma das vantagens da Inferência Bayesiana está no fato de permitir a incorporação de informações passadas (*a priori*), melhorando, dessa forma, o processo de inferência. Porém, a aplicação da Inferência Bayesiana em modelagem de valores extremos é muito recente. Conforme Coles & Powell (1996), até o ano de 1995, havia poucos trabalhos ligando os dois temas. Apenas na última década, alguns autores têm sugerido o procedimento Bayesiano para incorporar o conhecimento na estrutura dos modelos na inferência de valores extremos e para reduzir as incertezas da estimação dos parâmetros e de quantis da distribuição GEV. Coles & Tawn (1996) utilizaram conhecimentos de profissionais que trabalham na área para construir suas informações *a priori*, as quais incorporam a dependência entre os parâmetros associados ao modelo extremo, podendo incorporar o conhecimento baseado em quantis com uma distribuição Gama. Stephenson (2002) e Stephenson & Tawn (2004), sugerem um procedimento Bayesiano para incorporar o conhecimento prévio na estrutura dos teoremas de modelos na inferência de valores extremos, reduzindo a incerteza da estimação dos parâmetros.

Sendo assim, objetiva-se, neste trabalho, analisar a aplicação da Inferência Bayesiana, com uma distribuição *a priori* baseada em quantis extremos que facilite a incorporação dos conhecimentos fornecidos por especialistas, para obter as estimativas de precipitação máxima para os tempos de retorno de 10 e 20 anos e seus respectivos limites superiores de 95%, para o período anual e para os meses da estação chuvosa de Jaboticabal (SP). E também, comparar em relação à precisão e acurácia, os resultados dessa metodologia com os fornecidos pelo método da máxima verossimilhança.

## MATERIALE MÉTODOS

Os dados utilizados foram obtidos a partir dos registros pluviográficos fornecidos pelo Departamento de Ciências Exatas da Faculdade de Ciências Agrárias e Veterinária da UNESP, campus Jaboticabal, estado de São

Paulo. A estação Agrometeorológica do Campus encontra-se nas seguintes coordenadas geográficas; latitude de 21°15'22''S, longitude 48°18'58''W e altitude de 595 m. O clima da região é Cwa, de acordo com a classificação climática de Köppen, com chuvas no verão e inverno relativamente seco (VASCONCELLOS, 1998).

As observações referem-se às precipitações pluviiais diárias máximas, expressas em altura de lâmina d'água (mm), referentes a 46 anos de observação do período compreendido entre janeiro de 1956 a dezembro de 2001. Os dados foram agrupados em períodos de um ano e mensal; utilizaram-se apenas os meses da estação chuvosa (de setembro a março), extraindo-se a máxima precipitação pluviial diária observada de cada período,

formaram-se os vetores de máximos. Foram utilizadas as observações de 1956 a 1990 para calcular a precipitação pluviial diária máxima provável, para os tempos de retorno de 10 e 20 anos e extraiu-se a maior observação de cada série do período restante (1991 a 2001), para verificar a acurácia das estimativas pontuais e intervalares.

Com o intuito de verificar se as observações satisfaziam à pressuposição de independência, exigida na função de verossimilhança, realizou-se o teste para desvio de aleatoriedade, sendo que, para a sua realização, foi utilizado o pacote T SERIES do Sistema Computacional Estatístico R.

A função de densidade de probabilidade da distribuição GEV, é dada por,

$$f(x) = \frac{1}{\sigma} \left\{ \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{-\left( \frac{1 + \xi}{\xi} \right)} \exp \left\{ - \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{-\left( \frac{1}{\xi} \right)} \right\} \right\} \quad (1)$$

definida em,  $-\infty < x < \mu - \frac{\sigma}{\xi}$ , para  $\xi < 0$  e  $\mu - \frac{\sigma}{\xi} < x < \infty$ , para  $\xi > 0$ , e quando  $\xi \rightarrow 0$ , tem-se que essa

corresponde à função de densidade de probabilidade Gumbel, dada por:

$$f(x) = \frac{1}{\sigma} \exp \left\{ - \left( \frac{x - \mu}{\sigma} \right) - \exp \left[ - \left( \frac{x - \mu}{\sigma} \right) \right] \right\} \quad (2)$$

que é definida em,  $-\infty < x < \infty$ .

Supondo-se que há independência entre as observações, a função de verossimilhança será,

$$L(\mathbf{x}|\theta) = \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\left( \frac{1 + \xi}{\xi} \right)} \right\} \exp \left\{ \sum_{i=1}^n \left[ - \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\left( \frac{1}{\xi} \right)} \right] \right\} \quad (3)$$

para,  $\mu - \frac{\sigma}{\xi} > x_{(n)}$  e  $\xi < 0$ ,  $\mu - \frac{\sigma}{\xi} < x_{(1)}$  e  $\xi > 0$ . Caso contrário,  $L(\mathbf{x}|\mu, \sigma, \xi)$  não existe.

Os estimadores de máxima verossimilhança de  $\mu, \sigma$  e  $\xi$  são obtidos pela solução do sistema de equações não lineares formado pelas derivadas de

primeira ordem do logaritmo da equação (3), em relação a cada parâmetro, igualadas a zero, isto é, pela solução de:

$$\frac{\partial l(\mu, \sigma, \xi)}{\partial \mu} = 0 \Rightarrow \frac{1}{\sigma} \sum_{i=1}^n \left( \left( 1 + \hat{\xi} - w_i \frac{1}{\hat{\xi}} \right) \hat{\sigma}^{-1} \right) = 0$$

$$\frac{\partial l(\mu, \sigma, \xi)}{\partial \sigma} = 0 \Rightarrow -\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n \left\{ (x_i - \hat{\mu}) \left[ (1 + \xi) - w_i^{\frac{1}{\xi}} \right] w_i^{-1} \right\} = 0 \quad (4)$$

$$\frac{\partial l(\mu, \sigma, \xi)}{\partial \xi} = 0 \Rightarrow \sum_{i=1}^n \left\{ \left( 1 - w_i^{-\frac{1}{\xi}} \right) \left[ \frac{1}{\xi^2} \ln(w_i) - \frac{(x_i - \hat{\mu})}{\xi \hat{\sigma} w_i} \right] - \frac{(x_i - \hat{\mu})}{\hat{\sigma} w_i} \right\} = 0$$

em que,  $w_i = 1 + \xi \left( \frac{x_i - \hat{\mu}}{\hat{\sigma}} \right)$

Visto que o sistema de equações (4) não possui solução analítica, deve-se utilizar um método iterativo para obtenção de uma solução numérica.

A precipitação máxima provável para um determinado tempo de retorno  $T$  pode ser determinada pela seguinte expressão, que corresponde ao quantil da distribuição GEV.

$$y_t = q_p = \mu + \frac{\sigma}{\xi} \left[ (x_p)^{-\xi} - 1 \right] \quad (5)$$

em que,  $x_p = -\ln(1 - p)$

Dados de valores extremos são normalmente escassos, principalmente em dados meteorológicos; logo, informações de profissionais que atuam na área podem ser essenciais para a complementação das informações contidas nos dados. Portanto, é razoável esperar que um “profissional da área” forneça informações relevantes *a priori* sobre o comportamento extremo, uma vez que ele tem conhecimentos específicos sobre as características dos dados em estudo. Outro fato importante está na possibilidade do uso de conhecimento específico das características de eventos

extremos em regiões vizinhas as quais se pretende estudar, que podem se caracterizar como importantes informações *a priori*.

Adotando os quantis como forma de incorporar o conhecimento *a priori*, pode-se, a partir da expressão (5), expressar a informação *a priori* em termos de  $(q_{p_1}, q_{p_2}, q_{p_3})$ , com  $p_1 > p_2 > p_3$ , conseqüentemente,  $q_{p_1} < q_{p_2} < q_{p_3}$ . Assumindo que, *a priori*, esses quantis são independentes e admitindo-se como *priori* marginal a distribuição de Gumbel com os parâmetros  $a_i$  (posição) e  $b_i$  (escala), tem-se:

$$q_{pi} \sim \text{Gumbel}(a_i, b_i), \text{ em que, } -\infty < a_i < \infty \text{ e } b_i > 0$$

Ao se trabalhar com valores extremos,  $q_{p_1}, q_{p_2}$  e  $q_{p_3}$  devem ser altos quantis e, também, podem ser considerados valores extremos. Portanto, sua distribuição deve corresponder

a uma distribuição de valores extremos, como é o caso da Gumbel.

Assim, a distribuição *a priori* para  $\theta = (\mu, \sigma, \xi)$  será dada por:

$$P(\theta) \propto J(\theta) \prod_{i=1}^3 \frac{1}{b_i} \exp \left\{ - \left( \frac{q_{pi} - a_i}{b_i} \right) - \exp \left[ - \left( \frac{q_{pi} - a_i}{b_i} \right) \right] \right\} \quad (6)$$

em que,  $J(\theta)$  é o Jacobiano da transformação de  $(q_{p1}, q_{p2}, q_{p3})$  para  $\theta = (\mu, \sigma, \xi)$ , dado por:

$$J(\theta) = \frac{\sigma}{\xi^2} \left| \sum_{\substack{i,j \in \{1,2,3\} \\ i < j}} (-1)^{i+j} (x_i x_j)^{-\xi} \ln(x_j/x_i) \right| \quad (7)$$

e,  $x_i = -\ln(1 - p_i)$ , para  $i = 1, 2, 3$ .

Os parâmetros  $a_i$  e  $b_i$  da distribuição de Gumbel têm suas estimativas determinadas por meio das medidas de posição e escala na confiança *a priori*. A idéia é de obter-se informações de profissionais que atuam na área e utilizar o conhecimento a respeito da mediana e do quantil 90% para cada  $q_{pi}$  e, depois, obter-se estimativas para os parâmetros da distribuição de Gumbel, que correspondam a esses valores.

Para a análise dos dados de precipitação máxima de Jaboticabal, tomaram-se, como informações *a priori*,

certas características da análise de precipitações máximas de Lavras-MG, em estudo realizado por Beijo et al. (2005), que, analisando as estatísticas descritivas dos dados, afirmaram que as precipitações máximas ocorridas em Lavras-MG e Jaboticabal-SP apresentam semelhanças.

Dada a densidade *a priori* e a função de verossimilhança, pode-se então determinar a densidade *a posteriori* dada por:

$$P(\theta|\mathbf{x}) \propto J(\theta) \prod_{j=1}^3 \frac{1}{b_j} \exp \left\{ - \left( \frac{q_j - a_j}{b_j} \right) - \exp \left[ - \left( \frac{q_j - a_j}{b_j} \right) \right] \right\} \times \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\left( \frac{1+\xi}{\xi} \right)} \right\} \exp \left\{ \sum_{i=1}^n \left[ - \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right] \right\} \quad (8)$$

Calcular analiticamente a integral da equação (8) apresenta problemas, pois esta não tem uma forma fechada. Sendo empregada a técnica Monte Carlo via Cadeias de Markov (MCMC) foi para superar esta dificuldade. O algoritmo utilizado foi de Metropolis-Hastings, apresentado por Hastings (1970), que está implementado no pacote MCMCpack do Sistema Computacional Estatístico R. Utilizando-se como distribuição candidata para os parâmetros uma distribuição Normal Multivariada. Realizou-se um processo com 420000 iterações, sendo que foram descartadas as 20000 primeiras (“burn-in”), com o objetivo de eliminar interferências dos “chutes” iniciais, e para assegurar a independência da amostra, considerou-se um

espaçamento entre os pontos amostrados de tamanho 100 (“thinning”), ou seja, obteve-se uma amostra de tamanho 4000 para cada parâmetro. Foram utilizados os pacotes EVDBAYES, MCMCpack e BOA (Bayesian Output Analysis) do Sistema Computacional Estatístico R, conforme R Development Core Team (2004).

A monitoração da convergência das cadeias será realizada por meio da análise gráfica, das medidas descritivas e dos diagnósticos de Geweke (1992) e Raftery & Lewis (1992).

## RESULTADOS E DISCUSSÃO

Pode-se verificar na Figura 1, que poucas das precipitações máximas que ocorreram entre 1956 e 1990 tiveram valores superiores a 100 mm.

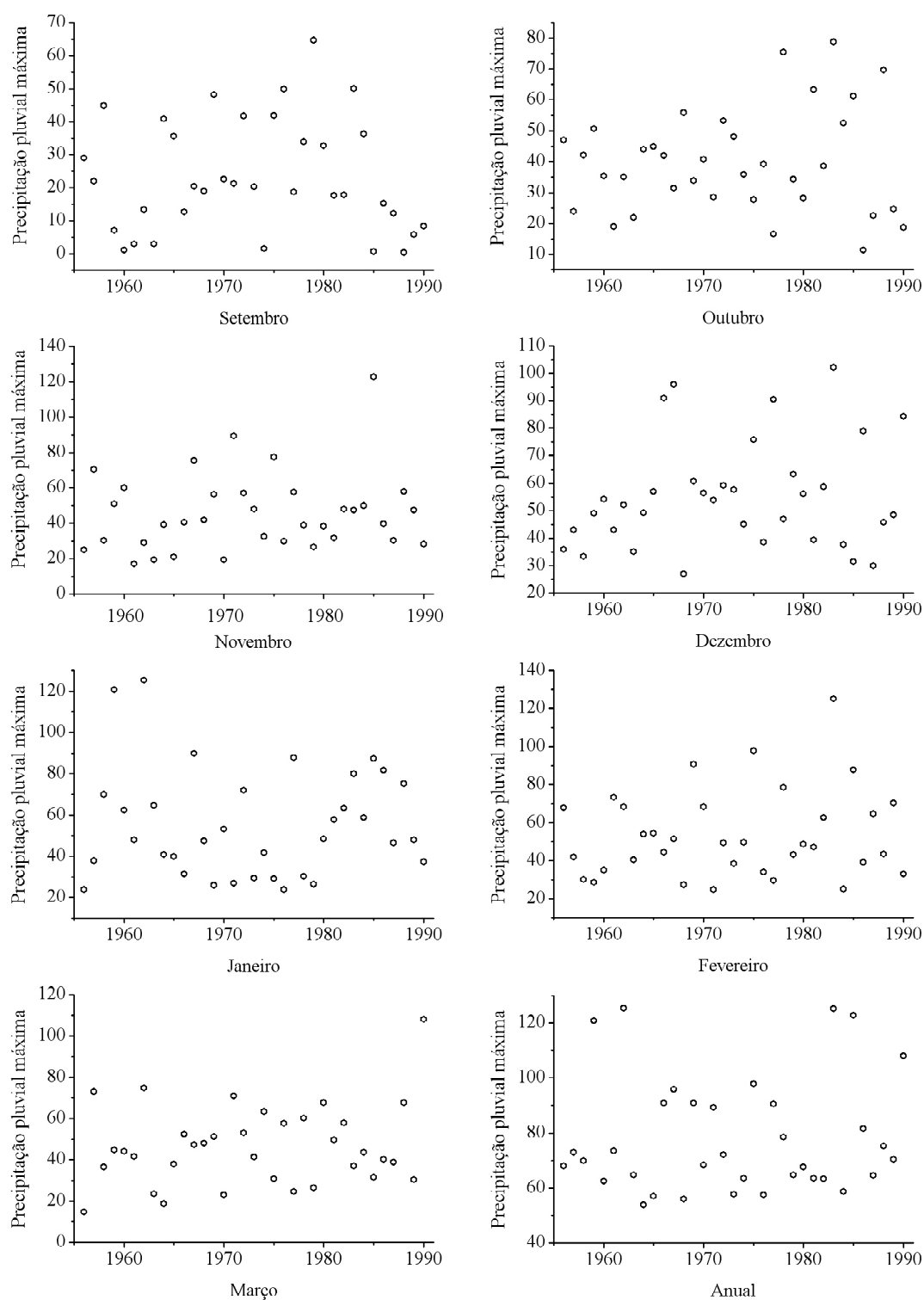


Figura 1 – Representação gráfica da série de precipitações pluviométricas diárias máximas (mm) no município de Jaboticabal (SP), dos meses de setembro a março, e para o período anual entre os anos de 1956 e 1990.

Adotando-se um nível de significância de 5%, para o teste para desvio de aleatoriedade das séries de dados, nota-se pelos p-valores apresentados na Tabela 1 que não existem evidências estatísticas, para rejeitar a hipótese de aleatoriedade.

A convergência da cadeia de cada parâmetro foi monitorada por meio da visualização gráfica do traço, do histograma e dos critérios disponíveis no pacote BOA do programa R, não existindo evidências contra a convergência. Um dos critérios de convergência analisado foi o Critério de Geweke, cujos resultados são apresentados na Tabela 2. De acordo com Geweke (1992), existirão

evidências contra a convergência se o p-valor analisado for menor que um nível de significância desejado. Considerando um nível de significância de 5%, pode-se verificar que, para todos os parâmetros nos diversos períodos estudados, não houve nenhuma evidência contra a convergência dos mesmos. Sendo confirmado pelo valor do fator de convergência de Raftery e Lewis (1992), que atingiu valores desejados ( $FD < 5$ ). Pode-se notar ainda que os erros Monte Carlo apresentaram valores baixos, confirmando a boa qualidade das estimativas dos parâmetros.

Tabela 1 – Resultado do teste para desvio de aleatoriedade (*run test*) das séries de dados estudados.

Setembro	Outubro	Novembro	Dezembro	Janeiro	Fevereiro	Março	Ano
0,123	0,603	0,609	0,326	0,234	0,101	0,202	0,228

Tabela 2 – Estimativas dos parâmetros da distribuição GEV obtidas via método da máxima verossimilhança e Inferência Bayesiana, erro Monte Carlo, fator de convergência de Raftery e Lewis (FD) e p-valor do Critério de Geweke.

Período	Parâmetro	Máxima Verossimilhança		Bayesiana		Convergência		
		Estimativa	Desvio Padrão	Média Posteriori	Desvio Padrão	Erro MC	FD	Geweke p-valor
Setembro	$\mu$	15,613	2,877	14,242	2,545	0,0436	0,9907	0,534
	$\sigma$	13,979	2,207	14,152	2,193	0,0364	1,0320	0,438
	$\xi$	-0,041	0,192	0,218	0,115	0,0025	1,1439	0,982
Outubro	$\mu$	32,840	2,750	32,644	2,303	0,0358	0,9907	0,582
	$\sigma$	14,180	1,997	15,947	2,005	0,0332	1,0320	0,841
	$\xi$	-0,090	0,143	0,232	0,097	0,0015	1,0109	0,833
Novembro	$\mu$	35,157	2,892	40,777	2,098	0,0648	1,0263	0,255
	$\sigma$	14,769	2,236	19,574	2,341	0,0693	1,1368	0,198
	$\xi$	0,124	0,152	0,207	0,112	0,0041	1,0789	0,069
Dezembro	$\mu$	45,706	2,802	45,753	2,377	0,0359	1,0320	0,385
	$\sigma$	12,241	2,132	15,439	2,261	0,0373	0,9704	0,857
	$\xi$	0,077	0,156	0,096	0,109	0,0023	0,9907	0,778
Janeiro	$\mu$	41,777	3,749	43,931	2,341	0,0381	1,0109	0,927
	$\sigma$	17,612	3,061	20,287	3,023	0,0505	1,0109	0,712
	$\xi$	0,1834	0,217	0,078	0,102	0,0020	1,0320	0,576
Fevereiro	$\mu$	41,929	3,124	40,750	1,897	0,0282	0,9907	0,613
	$\sigma$	15,372	2,495	16,014	2,200	0,0326	0,9616	0,373
	$\xi$	0,164	0,184	0,387	0,053	0,0008	1,0109	0,873
Março	$\mu$	38,437	2,986	42,614	2,184	0,0576	1,0547	0,695
	$\sigma$	15,782	2,134	19,432	2,294	0,0344	1,0109	0,141
	$\xi$	-0,054	0,116	0,108	0,110	0,0042	1,1207	0,914
Ano	$\mu$	66,775	2,377	64,379	2,115	0,0310	0,9704	0,556
	$\sigma$	11,712	2,093	10,810	2,018	0,0332	1,0534	0,540
	$\xi$	0,361	0,1945	0,417	0,058	0,0009	0,9907	0,662

Analisando-se a Tabela 2, pôde-se verificar que as estimativas dos parâmetros  $\mu$  e  $\sigma$  não divergiram muito entre as duas metodologias de inferência, exceto para os meses de novembro e março, em que as estimativas fornecidas pela metodologia Bayesiana apresentaram valores mais elevados. Entretanto, para o parâmetro  $\xi$  houve uma maior divergência entre as estimativas, sendo que, em todos os períodos, com exceção do mês de janeiro, a metodologia Bayesiana forneceu valores mais elevados. Um destaque deve ser dado aos meses de setembro, outubro e março, em que, as estimativas de máxima verossimilhança apresentaram valores negativos, enquanto que as bayesianas apresentaram valores positivos, isto pode influenciar na obtenção de estimativas de precipitações máximas. Os valores do desvio padrão para os parâmetros  $\mu$  e  $\sigma$  foram baixos em relação à média, enquanto que, para o parâmetro  $\xi$  os valores do desvio padrão foram mais elevados, porém, no caso desse parâmetro, o método Bayesiano forneceu estimativas mais precisas, uma vez que apresentou valores do desvio padrão menores, chegando em alguns casos (fevereiro e ano) a ser mais de três vezes inferior. Embora utilizando-se uma série de dados de ventos máximos de 48 anos, Coles & Powell (1996) também obtiveram um valor elevado do desvio padrão em relação à média *a posteriori* do parâmetro forma da distribuição GEV (média igual a 0,03 e desvio padrão igual a 0,07).

Analisando as Tabelas 3 e 4, pôde-se verificar que a estimativa pontual de precipitação máxima ( $\hat{y}_p$ ), para um tempo de retorno de 10 anos, fornecidas pela metodologia Bayesiana superaram os valores das maiores precipitações ocorridas entre os anos de 1991 a 2001, exceto para os meses de fevereiro e o período anual. Enquanto que, das estimativas fornecidas pelo método da máxima verossimilhança, somente no mês de janeiro houve um bom resultado. Porém, pôde-se verificar, ainda nas Tabelas 3 e 4, que para o mês de fevereiro e para o período anual, a precipitação máxima somente seria prevista se fosse adotado como referência o limite superior 95%, do período de retorno de 20 anos.

Portanto, o limite superior de 95% mostrou-se como valor de maior acurácia em relação aos valores reais apresentados na Tabela 3.

Como em análises de precipitações máximas sempre se está interessado no evento extremo máximo, sugere-se nesse caso utilizar o limite superior como valor de referência. Esses resultados concordam com Coles & Pericchi (2003) que, embora utilizando um modelo com componente sazonal, conseguiram melhor resultado para a predição de uma precipitação de 410 mm, ocorrida na Venezuela, quando utilizaram o limite superior do intervalo de 95% de confiança do nível de retorno obtido com o ajuste da distribuição GEV.

Na Tabela 4, escolhendo-se 20 anos como o tempo de retorno adequado de ocorrer uma determinada precipitação máxima diária para o período anual, pôde-se interpretar as informações do seguinte modo: espera-se, especificamente no caso Bayesiano que, em um tempo médio de 20 anos, o valor da precipitação diária máxima no ano seja superior a 127,7 mm, tendo um intervalo de alta densidade a posteriori de 95% (107,7 mm; 148,7 mm) e limite superior de 95% igual a 146,9 mm, respectivamente.

Comparando-se, de forma geral, os resultados fornecidos pela Inferência Bayesiana e pelo método da máxima verossimilhança (Tabela 4), pôde-se verificar que os valores das estimativas de precipitação máxima obtidos, utilizando-se a distribuição *a priori* via Inferência Bayesiana, apresentaram maior acurácia do que os valores fornecidos. Destaca-se a estimativa do mês de fevereiro, que no caso da máxima verossimilhança não conseguir abranger o verdadeiro valor de precipitação máxima nem mesmo quando se adota o limite superior de 95% do tempo de retorno de 20 anos. Ou seja, houve uma melhor previsão quando se utilizou a metodologia bayesiana.

Do ponto de vista prático, essa maior acurácia nos resultados é muito importante em termos de segurança e economia, pois quando, em um período de tempo menor, ocorre uma precipitação pluvial maior do que estava previsto, isso pode acarretar sérios danos. No caso da construção de uma curva de nível, ela não suportaria o volume de água e, conseqüentemente, causaria erosões do solo e soterramentos de plantações, causando sérios prejuízos ao meio ambiente e aos proprietários. Em alguns casos, como na construção de diques e represas, essa subestimação da precipitação pluvial máxima poderia causar rompimentos das obras, o que seria um alto risco à segurança. Esses resultados concordam com os obtidos

Tabela 3 – Maior altura de precipitação pluvial (mm), ocorrida nos meses mais chuvosos e no ano entre 1991 e 2001, em Jaboticabal (SP).

Setembro	Outubro	Novembro	Dezembro	Janeiro	Fevereiro	Março	Ano
53,4	80,0	97,5	77,3	83,5	137,2	78,5	137,2



Tabela 4 – Estimativas das precipitações pluviiais máximas (mm) nos meses mais chuvosos e no ano em Jaboticabal-SP, para os tempos de retorno de 10 e 20 anos e seus respectivos intervalos de 95 % e limite superior (LS) de 95%, obtidas via Inferência Bayesiana e método da máxima verossimilhança

Método	Período	Tempo de Retorno					
		10 anos			20 anos		
		$\hat{y}_p$	Intervalo 95%	LS (95%)	$\hat{y}_p$	Intervalo 95%	LS (95%)
Bayesiano	Setembro	55,4	[ 43,8; 66,8]	65,8	73,8	[57,6; 94,0]	90,9
	Outubro	80,0	[ 67,4; 93,4]	92,0	101,6	[79,8; 122,6]	121,4
	Novembro	97,7	[ 82,7; 114,3]	112,6	123,1	[94,5; 150,7]	149,2
	Dezembro	84,7	[ 74,0; 97,2]	95,6	99,3	[83,4; 118,7]	116,5
	Janeiro	93,9	[ 80,9; 109,7]	107,3	112,2	[91,9; 135,8]	133,2
	Fevereiro	98,2	[ 84,0; 114,0]	112,2	130,0	[108,2; 153,8]	151,4
	Março	92,7	[78,0; 108,7]	107,3	112,1	[89,8; 141,0]	138,1
	Ano	104,5	[91,3; 120,0]	117,8	127,7	[107,7; 148,7]	146,9
Máxima Verossimilhança	Setembro	45,6	[34,5; 56,7]	54,8	54,8	[37,7; 71,6]	69,1
	Outubro	61,7	[52,0; 71,4]	69,8	69,8	[42,0; 97,6]	93,1
	Novembro	74,8	[56,4; 93,2]	90,2	87,8	[62,1; 113,5]	109,3
	Dezembro	80,9	[66,1; 95,7]	93,3	93,1	[70,6; 115,6]	111,9
	Janeiro	90,1	[67,2; 112,1]	109,4	104,8	[74,6; 135,0]	130,1
	Fevereiro	83,9	[64,2; 103,6]	100,4	97,0	[70,7; 123,3]	119,0
	Março	71,9	[60,5; 83,3]	81,4	81,7	[66,2; 97,2]	94,7
	Ano	107,2	[87,6; 126,8]	123,6	129,2	[84,7; 173,7]	166,4

por Coles & Powell (1996), que obtiveram melhores estimativas dos parâmetros da distribuição GEV e de precipitações máximas para certos períodos de retorno, quando aplicaram a metodologia Bayesiana, destacando-se que ela oferece resultados mais confiáveis que os da máxima verossimilhança.

### CONCLUSÕES

A metodologia Bayesiana, junto com a distribuição a priori, mostrou-se adequada para a incorporação de conhecimentos de especialistas no estudo de dados extremos, para a obtenção das estimativas dos parâmetros da distribuição GEV e das precipitações máximas, para certos níveis de retorno na região Jaboticabal (SP), e apresentou uma maior precisão e acurácia do que o método da máxima verossimilhança.

As estimativas intervalares, em especial o limite superior de 95%, fornecem valores de maior acurácia de precipitações máximas prováveis e podem ser usadas para auxiliar no planejamento de obras de engenharia hidráulica e agrícola na região de Jaboticabal.

As séries geradas, aplicando-se a metodologia bayesiana com distribuição *a priori* proposta, não apresentaram nenhum problema quanto à convergência.

### REFERÊNCIAS BIBLIOGRÁFICAS

- BEIJO, L. A.; MUNIZ, J. A.; CASTRO NETO, P. Estudo do tempo de retorno das precipitações máximas em Lavras (MG) pela distribuição de valores extremos do tipo I. **Ciência e Agrotecnologia**, Lavras, v. 29, n. 3, p. 657-667, maio/jun. 2005.
- COLES, S. G.; DIXON, J. Likelihood-based inference for extreme value models. **Extremes**, Tyresoe, v. 2, n. 1, p. 5-23, Mar. 1999.
- COLES, S. G.; PERICCHI, L. R. Anticipating catastrophes through extreme value modelling. **Applied Statistics**, London, v. 52, n. 3, p. 405-416, Oct. 2003.
- COLES, S. G.; POWELL, E. A. Bayesian methods in extreme value modelling: a review and new developments. **International Statistics Review**, Voorburg, v. 64, n. 1, p. 119-136, Apr. 1996.

- COLES, S. G.; TAWN, J. A. A Bayesian analysis of extreme rainfall data. **Applied Statistics**, Oxford, v. 45, n. 4, p. 463-478, 1996.
- GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Eds.). **Bayesian statistics 4**. Oxford: Clarendon, 1992. p. 625-631.
- HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, London, v. 57, n. 1, p. 97-109, Apr. 1970.
- JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, Reading, v. 81, n. 348, p. 158-171, Apr. 1955.
- MARTINS, E. S.; STEDINGER, J. R. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. **Water Resources Research**, Washington, v. 36, n. 3, p. 737-744, Mar. 2000.
- PAULINO, D. C.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: FCG, 2003. 446 p.
- RAFTERY, A. L.; LEWIS, S. One long run with diagnostics: implementation strategies for Markov chain Monte Carlo. **Statistical Science**, Hayward, v. 7, n. 4, p. 493-497, 1992.
- R DEVELOPMENT CORE TEAM. **The R manuals**: current. Version: 1.9.1. 2004. Disponível em: <<http://www.r-project.org>>. Acesso em: 10 out. 2006.
- SMITH, R. L. Maximum likelihood estimation in a class of nonregular cases. **Biometrika**, London, v. 72, n. 1, p. 67-90, 1985.
- STEPHENSON, A. G. **A user's guide to the evdbayes package**. Version 1.0. 2002. Disponível em: <<http://www.maths.lancs.ac.uk/~stephena/>>. Acesso em: 12 out. 2006.
- STEPHENSON, A. G.; TAWN, J. A. Bayesian inference for extremes: accounting for the three extremal types. **Extremes**, Clermont Ferrand, v. 7, n. 4, p. 291-307, Dec. 2004.
- TUCCI, C. M. **Hidrologia**: ciência e aplicação. Porto Alegre: UFRGS, 2001. 943 p.
- VASCONCELLOS, S. L. B. **Análise da quantidade e distribuição pluviométrica na região de Jaboticabal - SP**. 1998. 63 f. Dissertação (Mestrado em Agronomia) – Escola Superior de Agricultura Luiz de Queirós, Piracicaba, 1998.
- VIEIRA, S. R.; LOMBARDI NETO, F.; BURROWS, I. T. Mapeamento de chuva diária máxima provável para o estado de São Paulo. **Revista Brasileira de Ciência do Solo**, Campinas, v. 15, n. 1, p. 93-98, jan./abr. 1991.