

**CONSTRUÇÃO DE UMA PRIORI PARA OS  
PARÂMETROS DO MODELO DE VALORES  
EXTREMOS GENERALIZADO BASEADA EM  
QUANTIS COM DISTRIBUIÇÃO GUMBEL**

**LUIZ ALBERTO BEIJO**

**2006**

**LUIZ ALBERTO BEIJO**

**CONSTRUÇÃO DE UMA PRIORI PARA OS PARÂMETROS DO  
MODELO DE VALORES EXTREMOS GENERALIZADO BASEADA  
EM QUANTIS COM DISTRIBUIÇÃO GUMBEL**

Tese apresentada à Universidade Federal de Lavras,  
como parte das exigências do Programa de  
Pós-Graduação em Agronomia, área de concentração  
em Estatística e Experimentação Agropecuária, para  
a obtenção do título de “Doutor”.

Orientador:

Prof. Dr. Joel Augusto Muniz

Co-orientador:

Prof. Dr. Mário Javier Ferrua Vivanco

**LAVRAS  
MINAS GERAIS - BRASIL**

**2006**

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da  
Biblioteca Central da UFLA**

Beijo, Luiz Alberto

Construção de uma priori para os parâmetros do modelo de valores extremos generalizado baseada em quantis com distribuição Gumbel / Luiz Alberto Beijo.

-- Lavras : UFLA, 2006.

98p. : il.

Orientador: Joel Augusto Muniz

Tese (Doutorado) – UFLA.

Bibliografia.

1. Metodologia Bayesiana. 2. Distribuição *a priori*. 3. Conhecimento a priori.  
4. Quantis extremos. 5. Intervalo HPD. 6. Técnica MCMC. 7. Precipitação pluvial máxima. I. Universidade Federal de Lavras. II. Título.

CDD-519.544  
-634.88891

**LUIZ ALBERTO BEIJO**

**CONSTRUÇÃO DE UMA PRIORI PARA OS PARÂMETROS DO  
MODELO DE VALORES EXTREMOS GENERALIZADO BASEADA  
EM QUANTIS COM DISTRIBUIÇÃO GUMBEL**

Tese apresentada à Universidade Federal de Lavras,  
como parte das exigências do Programa de  
Pós-Graduação em Agronomia, área de concentração  
em Estatística e Experimentação Agropecuária, para  
a obtenção do título de “Doutor”.

APROVADA em, 22 de março de 2006.

Prof. Dr. Vicente Garibay Cancho	ICMC - USP
Prof. Dr. Mário Javier Ferrua Vivanco	UFLA
Prof <sup>a</sup> Dr <sup>a</sup> Thelma Sáfadi	UFLA
Prof. Dr. Luiz Gonsaga de Carvalho	UFLA

Prof. Dr. Joel Augusto Muniz  
UFLA  
(Orientador)

**LAVRAS  
MINAS GERAIS - BRASIL**

*A sabedoria que vem do alto é, antes de tudo, pura, pacífica, humilde, compreensiva, cheia de misericórdia e bons frutos, sem discriminações e sem hipocrisia. Na verdade, um fruto justiça é semeado na PAZ para aqueles que trabalham pela PAZ.*  
*Tg 3, 17-18.*

*Se alguém de vocês tem falta de sabedoria, que peça a DEUS, e Ele a dará, porque é generoso e dá sem condições. Todavia é preciso pedir com fé, sem duvidar...*  
*Tg 1, 5-6.*

À glória de meu Pai Celestial,

Aos meus pais, Sebastião Vicente Beijo e Geraldina Lourenço Beijo,

Aos meus irmãos, Nino e Nina,

Às minhas sobrinhas Bibi e Vivi,

À Tia Tonica (*in memoriam*),

DEDICO

À minha eterna professora de Matemática, Dona Dorinha,

À verdadeira Orientadora, Dona Maria das Graças,

OFEREÇO

## AGRADECIMENTOS

Agradeço

A DEUS, por tudo.

À Universidade Federal de Lavras (UFLA), ao Programa de Pós-Graduação em Estatística e Experimentação Agropecuária e ao Departamento de Ciências Exatas (DEX), pela oportunidade de realização do curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pela bolsa de estudo concedida.

Ao professor Joel Augusto Muniz e ao professor Mário Javier Ferrua Vivanco, pela atenção, orientação, amizade e incentivo.

Aos professores membros da banca examinadora, Vicente Garibay Cancho, Thelma Sáfadi e Luiz Gonsaga de Carvalho, pelas sugestões e colaboração.

Ao professor Fabrício “O Matemático”, pelo auxílio e sugestões.

Aos professores do departamento de Ciências Exatas, pelos conhecimentos transferidos.

A todos os funcionários do departamento de Ciências Exatas, pela competência e prestabilidade.

Aos meus colegas de curso, pelo companheirismo e carinho de verdadeiros irmãos, em especial ao Washington, Cirillo e Eric.

Aos colegas da nova turma, pela fraternidade que vocês proporcionaram.

À tia Mariinha e à minha prima Glória, pelas orações e força nos momentos difíceis.

Aos professores, funcionários e colegas da Escola Estadual Secretário Tristão da Cunha, pelos ensinamentos que foram o início da caminhada.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

## SUMÁRIO

	Página
LISTA DE TABELAS.....	i
RESUMO.....	iii
ABSTRACT .....	iv
1 - INTRODUÇÃO.....	1
2 - REFERENCIAL TEÓRICO.....	3
2.1 - Teoria de Valores Extremos .....	3
2.1.1 - Distribuição Generalizada de Valores Extremos.....	5
2.1.2 - Aplicações da Distribuição GEV.....	7
2.1.3 - Inferência dos Parâmetros .....	8
2.2 - Inferência Bayesiana .....	10
2.2.1 - Inferência Bayesiana na Distribuição GEV.....	13
2.3 - Simulação Estocástica via Cadeias de Markov .....	16
2.3.1 - Algoritmo Metropolis-Hastings .....	16
2.3.2 - Verificação de Convergência.....	18
2.3.2.1 - Diagnóstico de Geweke.....	19
2.3.2.2 - Critério de Raftery e Lewis .....	20
2.4 - Estudos de Precipitações Máximas.....	21
2.5 - Teste para Desvio de Aleatoriedade .....	23
3 - MATERIAL E MÉTODOS.....	25
3.1 - Dados Simulados .....	25
3.2 - Dados de Precipitação Máxima de Jaboticabal .....	26
3.3 - Função de Densidade de Probabilidade da Distribuição GEV .....	27
3.4 - Função de Verossimilhança.....	27
3.5 - Probabilidades de Ocorrência de Precipitações Máximas .....	29
3.5.1 - Estimação do Período de Retorno .....	29

3.5.2 - Estimaco do Nvel de Retorno da Distribuico GEV .....	30
4 - CONSTRUO DA DISTRIBUIO A <i>PRIORI</i> FUNDAMENTADA EM QUANTIS EXTREMOS .....	31
4.1 - Incorporaco das Informaces .....	37
4.2 - Distribuico <i>a Posteriori</i> .....	39
5 - RESULTADOS E DISCUSSO .....	40
5.1 - Anlise dos Dados Simulados .....	40
5.2 - Estudo das Precipitaes Mximas de Jaboticabal .....	46
6 - CONCLUSES .....	54
7 - REFERNCIAS BIBLIOGRFICAS .....	56
8 - ANEXOS .....	64



## LISTA DE TABELAS

	<b>Página</b>
<b>TABELA 1</b> Informações <i>a priori</i> e valores calculados das medianas e dos quantis 90% para $q_{pi}$ com os respectivos parâmetros da Gumbel, associados à distribuição <i>a priori</i> para análise do período anual.....	38
<b>TABELA 2</b> Média e desvio padrão a posteriori dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de dependência de Raftery e Lewis (FD) e p-valor do critério de Geweke para as amostras de n=30 .....	41
<b>TABELA 3</b> Média e desvio padrão a posteriori dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de dependência de Raftery e Lewis (FD) e p-valor do critério de Geweke para as amostras de n=50. ....	43
<b>TABELA 4</b> Média e desvio padrão a posteriori dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de dependência de Raftery e Lewis (FD) e p-valor do critério de Geweke para as amostras de n=100 .....	45
<b>TABELA 5</b> Resultado do teste para desvio de aleatoriedade ( <i>run test</i> ) das séries de dados estudados.....	48
<b>TABELA 6</b> Média e desvio padrão a posteriori dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de convergência de Raftery e Lewis (FD) e p-valor Critério de Geweke .....	49
<b>TABELA 7</b> Maior altura de precipitação (mm) ocorrida nos meses mais chuvosos e no ano entre 1992 e 2001, em Jaboticabal (SP) .....	50

<b>TABELA 8</b>	Estimativas das precipitações pluviais máximas (mm) nos meses mais chuvosos e no ano na região de Jaboticabal-SP, para os tempos de retorno de 10 e 20 anos e seus respectivos intervalos HPD de 95 % e limite superior (LS) de 95%.....	51
-----------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

## RESUMO

BEIJO, Luiz Alberto. **Construção de uma priori para os parâmetros do modelo de valores extremos generalizado baseada em quantis com distribuição Gumbel**. Lavras: UFLA, 2006. 98p. Tese (Doutorado em Estatística e Experimentação Agropecuária)\*

A distribuição generalizada de valores extremos (GEV) desempenha um papel fundamental em estudos relacionados a medições físicas, em que é aplicada com a finalidade de descrever o comportamento de eventos raros. Como esses valores são extraídos da cauda superior (ou inferior) da distribuição original, isto faz com que, na maioria das situações, se tenha uma quantidade escassa de dados, o que pode dificultar a obtenção de estimativas confiáveis sobre alguma medida de interesse dos mesmos. Uma alternativa para se obter uma melhoria na qualidade das estimativas seria utilizar informações que os especialistas de determinada área em estudo possuam. Sendo assim, propõe-se neste trabalho uma distribuição *a priori* alternativa para a abordagem Bayesiana da distribuição GEV, que facilite a incorporação dos conhecimentos fornecidos por especialistas. A forma proposta para introduzir o conhecimento consiste em utilizar quantis da distribuição GEV, assumindo que estes têm uma distribuição de Gumbel. Isto facilita a incorporação da informação, uma vez que quantis extremos e os parâmetros da distribuição de Gumbel são medidas familiares a especialistas da área. Para avaliar o desempenho desta nova abordagem, foram simuladas 36 séries de dados, de diferentes tamanhos de amostras e diferentes valores de parâmetros. Objetivou-se também 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). A técnica Monte Carlo via Cadeias de Markov (MCMC) foi empregada para inferência *a posteriori* de cada parâmetro. A metodologia apresentou bons resultados na estimação dos parâmetros da distribuição GEV, para os diferentes tamanhos de amostras, e 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.

---

\* Comitê Orientador: Prof. Dr. Joel Augusto Muniz - UFLA (Orientador) e Prof. Dr. Mário Javier Ferrua Vivanco - UFLA (Co-orientador)

## ABSTRACT

BEIJO, Luiz Alberto. **Construction of one priori for the parameters of the generalized extreme values model based in quantis with Gumbel distribution.** Lavras: UFLA, 2006. 98p. Thesis (Doctorate in Statistics and Agricultural Experimentation). \*

The generalized extreme value distribution (GEV) plays a key role in studies related to physical measurements, where it is applied to describe the behavior of the extreme values, or rare event. As these values are extracted from the upper (or lower) tail of the original distribution, a scarce amount of data is obtained in most cases. This can be usually a problem in the acquirement of reliable estimates on some measure of interest. An alternative to overcome this potential problem, and as a consequence, to obtain an improvement in the quality of the estimates would be to use available information that certain specialists possess. One theme of this paper is to propose an alternative prior distribution for the Bayesian approach of the GEV distribution that makes easier the incorporation of the information supplied by specialists. The form proposed to introduce the prior knowledge of experts consists of using quantis of the GEV distribution, assuming that they have a Gumbel distribution. This approach makes easier the incorporation of the information, once extreme quantis and Gumbel distribution parameters are standard quantities to specialists in extreme values analyses. To evaluate the performance of this new approach, 36 series of data, for samples of different sizes and different values of parameters have been simulated. It was also aimed to determine the punctual and the upper limit 95% estimates of the probable maximum rainfall for return periods of 10 and 20 years, for the annual and month periods of the rainy station at region of Jaboticabal, Sao Paulo States, Brazil. Posterior inference is obtained through Markov Chain Monte Carlo (MCMC) methods. The proposed prior presented good results in the estimation of parameters of GEV distribution, for samples of different sizes, and in the determination of values of probable maximum rainfall estimates for the region of Jaboticabal. It presented itself as a good alternative in the incorporation of prior knowledge in the study of extreme data.

---

\* Guidance Committee: Prof. DSc. Joel Augusto Muniz - UFLA (Adviser) and Prof. DSc. Mário Javier Ferrua Vivanco - UFLA

## 1 – INTRODUÇÃO

A teoria de valores extremos é uma metodologia utilizada para descrever o comportamento dos pontos mais afastados da média de uma distribuição (observações extremas), e tem sido de vital importância para o planejamento de atividades sujeitas aos efeitos de eventos extremos. Desde os fundamentos da teoria de valores extremos por Fisher & Tippett (1928) e a criação da distribuição Generalizada de Valores Extremos, que habitualmente é denotada por distribuição GEV do inglês “*Generalized Extreme Value*” por Jenkinson (1955), várias aplicações desta teoria têm sido implementadas com sucesso em muitas áreas da ciência como: Hidrologia, Climatologia, Ciências Atuariais, Engenharia e em Economia e Finanças, nas quais o interesse maior está na obtenção de quantis extremos desta distribuição.

Sendo a distribuição Generalizada de Valores Extremos uma função de densidade, é natural que seus parâmetros sejam estimados a partir dos valores extremos de amostras. Vários métodos podem ser utilizados para obtenção dos estimadores dos parâmetros desta distribuição e várias sugestões foram propostas, entre elas, técnicas gráficas, estimadores baseados no método dos momentos, estatística de ordem, método dos momentos de probabilidade ponderada, método de máxima verossimilhança; método de máxima verossimilhança penalizada e, recentemente, métodos Bayesianos.

Apesar do enorme potencial de aplicação da abordagem Bayesiana, o emprego destes métodos em estudos, que envolvem valores extremos ainda é tímido. 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, entre estes, Coles & Tawn (1996), Coles & Powell (1996), Stephenson (2002) e Stephenson &

Tawn (2004). Porém, algumas dificuldades ainda existem para a aplicação desta metodologia, sendo que uma delas é decidir qual é a melhor maneira de se introduzir o conhecimento *a priori*, obtido de especialistas das áreas em estudo, para melhorar a modelagem dos valores extremos.

Sendo assim, propõe-se neste trabalho desenvolver uma distribuição *a priori* para a abordagem Bayesiana da distribuição GEV, que facilite a incorporação de conhecimentos fornecidos por especialistas, de maneira que se possa atingir uma maior acurácia nos resultados. Esta distribuição baseia-se na introdução do conhecimento *a priori* por meio de quantis extremos, seguindo a sugestão de Coles & Tawn (1996), porém, com algumas alterações no método desenvolvido pelos autores. Uma delas é de trabalhar diretamente com quantis extremos, e não, com a diferença entre eles. Uma vantagem desta estratégia é que é possível assumir que os quantis extremos têm uma distribuição de Gumbel, sendo que esta distribuição possui parâmetros que são familiares aos especialistas da área, possibilitando assim uma forma prática de se introduzir um conhecimento *a priori*, com o objetivo de obter maior acurácia nas estimativas dos parâmetros e dos quantis da distribuição GEV, otimizando o uso dos dados disponíveis.

Para avaliar o desempenho da distribuição *a priori* que está sendo proposta no presente trabalho, foram geradas 36 séries de dados para diferentes tamanhos de amostras e diferentes valores de parâmetros, séries estas que são abordadas na seção 3.1.

Na seção 3.2 serão apresentados detalhes sobre as séries de dados de precipitação máximas de Jaboticabal. Objetivou-se, no estudo destas séries de dados, avaliar o desempenho da distribuição *a priori* na determinação do nível de retorno de precipitações máximas, em relação a acurácia.

## 2 - REFERENCIAL TEÓRICO

### 2.1 – Teoria de Valores Extremos

Os fundamentos da teoria de valores extremos foram inicialmente expostos por Fisher & Tippett (1928), que definiram os três tipos possíveis (I, II e III) de distribuições assintóticas de valores extremos, conhecidas como de Gumbel, de Fréchet e de Weibull, respectivamente. No entanto, o primeiro a estudar e formalizar a aplicação estatística destas distribuições foi Gumbel (1958), cuja metodologia tem sido freqüentemente aplicada à máxima anual de séries de dados referentes a precipitações e vazões de rios. (Bautista, 2002).

De acordo com Souza & Silva (1999), uma das aplicações da teoria dos valores extremos é dada em função da estimabilidade das probabilidades e quantis com razoável precisão nos limites dos valores encontrados dentro da amostra. Porém, a modelagem destes valores corresponde a um ajuste apenas da cauda da distribuição dos dados originais. Uma outra vantagem na teoria dos valores extremos, segundo Leadbetter et al. (1983), refere-se à extensão dos resultados para outras estatísticas; neste aspecto, suponha-se que a teoria dos valores extremos seja para o valor máximo, assim, os resultados poderão ser estendidos para o valor mínimo devido à relação:  $-\text{Max}(-X_1, \dots, -X_n) = \text{Min}(X_1, \dots, X_n)$ . Entretanto, é importante reconhecer que devido às particularidades dos valores extremos (máximo/mínimo), inerentes a cada problema, tem-se que o comportamento de cada um dos valores extremos deve ser estudado separadamente para cada amostra. Devido a este fato, em muitas situações, os valores mínimos poderão apresentar comportamento diferente dos máximos.

A teoria dos valores extremos ocupa-se, em primeiro lugar, com a distribuição exata, tanto de mínimos como de máximos, e, em segundo lugar, com suas distribuições assintóticas. A condição essencial, para ambos os casos, é que a distribuição inicial, a partir da qual os extremos foram extraídos, e seus parâmetros, permaneçam constantes de uma amostra a outra, ou as mudanças que têm acontecido, ou acontecerão, possam ser determinadas ou eliminadas. Outra condição é que os extremos observados devem ser extremos de amostras de dados independentes. Gumbel (1958) defende algumas razões para que estas condições não sejam consideradas como críticas. A distribuição dos valores extremos pode ser caracterizada pelos parâmetros clássicos: média, mediana e moda, entre as quais, a moda exerce um papel muito importante (Vivanco, 1994).

Segundo Leadbetter et al. (1983), ao se estudar o comportamento extremo de uma seqüência, um aspecto de interesse é a função de distribuição,  $F_n(x)$ , de  $M_n = \max(X_1, \dots, X_n)$  para  $n$  grande, isto é, seja  $\{X_i\}$ ,  $i \in N$  uma sucessão de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.), com função de distribuição comum  $F(x)$ , então a distribuição de  $M_n$  é dada por:

$$F_n(x) = \Pr(M_n \leq x) = \{F(x)\}^n .$$

Dadas as constantes de normalização  $\{a_n, b_n (> 0)\}$ , o teorema de Fisher-Tippet estabelece que

$$\lim_{n \rightarrow \infty} \Pr\{a_n^{-1}(M_n - b_n) \leq x\} = \lim_{n \rightarrow \infty} F^n(a_n + b_n x) \xrightarrow{d} L(x),$$

em que,  $L(x)$  é uma função de distribuição não degenerada.

A teoria assintótica de valores extremos (Fisher & Tippett, 1928 e Gnedenko, 1943) estabelece que, considerando os parâmetros de posição e de escala,  $L(x)$  tem de ser de uma das três formas:



$$\Lambda(x) = \exp(-e^{-x}); \quad (i)$$

$$\Phi_{\alpha}(x) = \exp(-x^{-\alpha}); \quad x > 0; \quad \alpha > 0; \quad (ii)$$

$$\Psi_{\alpha}(x) = \exp(-(-x)^{\alpha}); \quad x < 0; \quad \alpha > 0 \quad (iii)$$

Estes três tipos de distribuições podem ser representados numa forma paramétrica única, designada por *distribuição generalizada de valores extremos* “GEV”, conforme proposto por Jenkinson (1955). Maiores detalhes sobre teoremas da teoria de valores extremos podem ser encontrados em Gumbel (1958) e Leadbetter et al. (1983).

### 2.1.1 - Distribuição Generalizada de Valores Extremos

A distribuição generalizada de valores extremos (GEV), proposta por Jenkinson (1955), tem função de distribuição acumulada de probabilidade dada por:

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

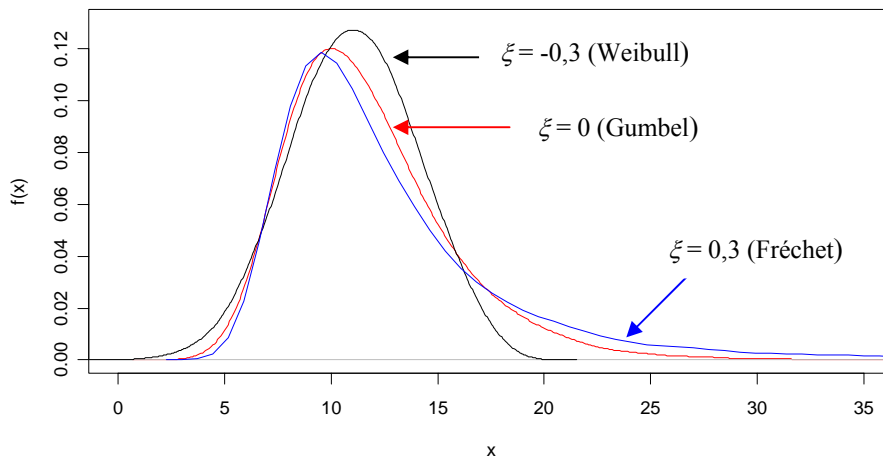
definida em,  $-\infty < x < \mu - \frac{\sigma}{\xi}$ , para  $\xi < 0$  e  $\mu - \frac{\sigma}{\xi} < x < \infty$ , para  $\xi > 0$ , sendo  $\mu, \sigma$  e  $\xi$  os parâmetros de posição, escala e de forma respectivamente, com  $\sigma > 0$ .

Conforme Raynal (1997), as distribuições de valores extremos de Fréchet e de Weibull correspondem aos casos particulares de (1) em que  $\xi > 0$  e  $\xi < 0$ , respectivamente, e com  $\lim_{\xi \rightarrow 0} F(x)$ , tem-se que (1) corresponderá à função de distribuição acumulada de Gumbel que é dada por:

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

definida em  $-\infty < x < \infty$ .

A Figura 1 apresenta os gráficos da função densidade de probabilidade da GEV para  $\xi = -0,3$  (Weibull),  $\xi \rightarrow 0$  (Gumbel) e  $\xi = 0,3$  (Fréchet), com  $\mu = 20$  e  $\sigma = 3$ , a partir dos quais pode se observar que o parâmetro  $\xi$  determina a natureza das caudas da distribuição, ou seja, a velocidade com que as caudas se aproximam de zero. Nota-se que, quanto maior o valor do parâmetro  $\xi$ , mais lenta é essa velocidade.



**FIGURA 1:** Função densidade de probabilidade da distribuição generalizada de valores extremos (GEV) para  $\xi = -0,3$ ,  $\xi = 0$  e  $\xi = 0,3$  com  $\mu = 20$  e  $\sigma = 3$ .

### 2.1.2 - Aplicações da Distribuição GEV

Desde sua origem com Fisher & Tippett (1928), e posteriores desenvolvimentos com Coles & Tawn (1991), Coles & Powell (1996), Ferro & Segers (2003), Gnedenko (1943), Gumbel (1958), Hann (1977), Jenkinson (1955), Leadbetter et al. (1983), Pickands (1975), Smith & Naylor (1997), Stephenson (2002), Stephenson & Tawn (2004), várias aplicações da teoria de valores extremos têm sido implementadas, com sucesso, em muitas áreas da ciência. Dentre sua vasta aplicação, a distribuição GEV tem sido utilizada com frequência em estudos de fenômenos ambientais principalmente para solucionar problemas relacionados a diversas áreas de Engenharia, e em aplicações que envolvem modelos financeiros e de seguros, além de outras.

Segundo Martins & Stedinger (2000), a distribuição GEV foi recomendada para a análise de frequências de enchentes no Reino Unido pelo *Natural Environmental Research Council* (NERC), em 1975. Essa distribuição também tem sido utilizada para a análise de vários tipos de dados relacionados à medidas físicas como:

- Frequências regionais de vazões, em estudos realizados por Chowdhury et al. (1991) e Hosking et al. (1985a);
- Estudos sobre velocidades máximas de vento, como em Bautista (2002), Ross (1987) e Whalen et al. (2004);
- Estimar precipitações máximas prováveis como, em trabalhos de Banzatto (1973) e Coles et al. (2003), entre outros.

A aplicação da distribuição GEV, em modelos financeiros baseia-se principalmente na modelagem dos chamados valores de Risco (VaR's). Com o objetivo de investigar a possibilidade de se obter melhoria significativa do ponto de vista do controle de risco, Souza (1999) e Souza & Silva (1999) utilizaram a teoria de valores extremos aplicada às séries financeiras de retornos de C - Bond

e retornos de Telebrás e concluíram que esta teoria é mais indicada para modelagem das caudas. Outros pesquisadores têm estudado nesta área, Kearns & Pagan (1997), McNeil & Frey (2000), entre outros.

A distribuição GEV também tem sido aplicada em outras áreas como o estudo de níveis de poluentes tanto na água, como na atmosfera. Segundo Sharma et al. (1999), a teoria de valores extremos oferece uma alternativa prática para o estudo de concentrações extremas de poluentes em centros urbanos.

Muitas organizações vinculadas à segurança de barragens recomendam explicitamente o estudo de precipitações máximas para o caso de grandes obras, nas quais o galgamento envolve grandes riscos (Tucci, 2001).

### **2.1.3 – Inferência dos Parâmetros**

A distribuição generalizada de valores extremos apresenta, na sua forma geral, três parâmetros desconhecidos: de posição, de escala e de forma. Destes, é de grande interesse a estimação do parâmetro de forma  $\xi$ , o qual determina o comportamento da cauda. Para fazer inferências sobre esses parâmetros, segundo Coles & Dixon (1999), inúmeras sugestões foram propostas, entre elas, técnicas gráficas, estimadores baseados no método dos momentos; utilização de estatísticas de ordem; método dos momentos de probabilidade ponderada; método da máxima verossimilhança; método da máxima verossimilhança penalizada e métodos Bayesianos.

Vários estudos têm sido feitos quanto a se determinar os melhores estimadores. Hosking et al. (1985b) estudaram as propriedades assintóticas do estimador de  $\xi$  (para os valores entre  $-0,5$  e  $0,5$ ), assim como propriedades para pequenas amostras, tendo chegado à conclusão que o estimador por eles proposto (estimador do método dos momentos de probabilidade ponderada)

apresenta um pequeno viés, que decresce rapidamente quando o tamanho da amostra aumenta.

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. O mesmo autor citou que, devido aos limites da distribuição GEV dependerem de seus parâmetros, as condições de regularidade para a estimação pelo método da máxima verossimilhança não são necessariamente satisfeitas. Assim, por meio de um estudo cuidadoso obteve os seguintes resultados:

- i) quando  $\xi > -0,5$ , os estimadores de máxima verossimilhança satisfazem completamente as condições de regularidades;
- ii) quando  $-1 \leq \xi \leq -0,5$ , os estimadores de máxima verossimilhança existem, mas não satisfazem as condições de regularidades;
- iii) quando  $\xi < -1$ , os estimadores de máxima verossimilhança não existem.

Coles & Dixon (1999) comentam que um argumento que tem sido usado contra os estimadores de máxima verossimilhança se refere as suas propriedades em pequenas amostras e propuseram o uso do estimador de máxima verossimilhança penalizada e o compararam, aplicando dados anuais máximos de níveis do mar de 58 estações da costa do Reino Unido, com o estimador do método dos momentos de probabilidade ponderada e o estimador de máxima verossimilhança, concluindo que o método proposto por eles apresentou-se melhor que os outros dois na estimação de quantis. Hosking et al. (1985b) argumentaram que, em pequenas amostras, o estimador do método dos momentos de probabilidade ponderada é superior ao estimador de máxima verossimilhança, em termos de viés e erro quadrático médio.

Conforme Neves & Martins (1996), existe uma vasta literatura sobre a estimação do parâmetro  $\xi$ , cujos estimadores se baseiam num número  $m(n)$  de estatísticas ordinais de topo de uma amostra de dimensão  $n$ , por exemplo, Fraga Alves (1992), Hill (1975), Pickands (1975) e Themido Pereira (1994), numa perspectiva semi-paramétrica. Segundo esta abordagem, não se requer uma forma para a função de distribuição das  $n$  variáveis aleatórias, mas apenas o seu comportamento na cauda onde se pretende fazer inferência. As autoras apresentaram um estimador simples para  $\xi$ , baseado em quantis empíricos que foram selecionados de forma a minimizarem a variância assintótica do estimador considerado. Restringindo o estudo ao domínio de valores  $\xi \in ]-1/2; 1[$ , afirmando que este parâmetro se encontra, para a maioria das aplicações práticas, neste domínio de valores, e recorrendo a um procedimento de simulação, realizaram, para várias dimensões de amostras e vários valores do parâmetro de forma, 5000 repetições de cada situação e calcularam as estimativas médias e o erro quadrático médio empírico associado, comparando-o com o estimador obtido pelo “método dos momentos de probabilidade ponderada”, apresentado por Hosking et al.(1985b) e concluíram que o estimador proposto, no caso de grandes amostras, apresentou boas propriedades, visto que o viés e erro quadrático médio diminuíram à medida que se aumentou o tamanho da amostra.

## **2.2 – Inferência Bayesiana**

A abordagem Bayesiana surgiu no meio científico no século XX, mais precisamente na década de 30, com base em alguns estudos teóricos, como o de Jeffreys (1939); porém, estes estudos exigiam resoluções de integrais complexas, o que fez com que os métodos Bayesianos ficassem mais algumas décadas em desuso.

Na década de 90, o problema dos cálculos integrais foi solucionado de maneira alternativa com o trabalho de Gelfand et al. (1990), que exploraram um recurso de simulação dinâmica, o algoritmo Amostrador de Gibbs, elaborado por Geman & Geman (1984), que utiliza a teoria das Cadeias de Markov. Os resultados obtidos foram excelentes e este algoritmo se consolidou como uma das ferramentas mais utilizadas em análise Bayesiana.

A metodologia Bayesiana consiste de uma informação *a priori* dos dados amostrais e do cálculo da densidade *a posteriori* dos parâmetros. A utilização de informação *a priori* em inferência Bayesiana requer a especificação de uma distribuição a priori para a quantidade de interesse  $\theta$  (parâmetro). Esta distribuição deve representar probabilisticamente o conhecimento que se tem sobre  $\theta$  antes da realização do experimento. A distribuição *a posteriori* de um parâmetro  $\theta$  contém toda a informação probabilística a respeito deste parâmetro e sua função de densidade *a posteriori* é a melhor descrição do processo de inferência.

Utilizando-se a distribuição a priori  $P(\theta)$ , e a função de verossimilhança,  $L(S_n|\theta)$ , torna-se possível, por meio do Teorema de Bayes, obter a distribuição *a posteriori* de  $\theta$ ,  $P(\theta|S_n)$ , de forma que qualquer conclusão seja feita a partir desta distribuição (Box & Tiao, 1992). A expressão matemática do teorema é dada a seguir:

$$P(\theta|S_n) = \frac{L(S_n|\theta)P(\theta)}{\int L(S_n|\theta)P(\theta)d\theta},$$

sendo  $S_n = \{y_1, y_2, \dots, y_n\}$  a amostra. Como o denominador não depende de  $\theta$  tem-se:

$$P(\theta|Y) \propto L(Y|\theta)P(\theta), \quad (3)$$

ou seja, a expressão acima pode ser entendida como:

$$\text{Posteriori} \propto \text{Verossimilhança} \times \text{Priori},$$

em que,  $\propto$  representa proporcionalidade.

Segundo Rosa (1998), para se inferir com relação a qualquer elemento de  $\theta$ , a distribuição *a posteriori* conjunta dos parâmetros,  $P(\theta | y)$ , deve ser integrada em relação a todos os outros elementos que a constituem. Assim, se o interesse do pesquisador se concentra sobre determinado conjunto de  $\theta$ , por exemplo,  $\theta_1$ , tem-se a necessidade da obtenção da distribuição  $P(\theta_1 | y)$ , dada por:

$$P(\theta_1 | y) = \int_{\theta \neq \theta_1} P(\theta | y) d\theta_{\theta \neq \theta_1} .$$

A integração da distribuição conjunta *a posteriori* para a obtenção das marginais geralmente não é analítica e, portanto, necessita de algoritmos iterativos especializados como Metropolis-Hastings e Amostrador de Gibbs. Estes algoritmos utilizam a simulação de Monte Carlo para gerar valores a partir de distribuições de probabilidades conhecidas e a teoria das Cadeias de Markov para representar a dependência entre os parâmetros; portanto, fazem uso do método conhecido como MCMC, Monte Carlo – Markov Chain, (Sorensen, 1996).

Mediante a necessidade do uso de recursos computacionais na metodologia Bayesiana, houve um aumento significativo de sua aplicação na última década, devido ao grande desenvolvimento da tecnologia computacional. Assim, esta metodologia tem sido utilizada em vários segmentos da pesquisa científica. Em Inferência Bayesiana, as incertezas sobre uma estimativa pontual são também descritas por intervalo, o qual, nesta teoria, é chamado de intervalo de credibilidade, que é construído, baseado na distribuição *a posteriori*. Segundo Paulino et al. (2003), a região de credibilidade é definida numericamente e admite uma interpretação probabilística direta e inequívoca. Afirmam ainda que existe uma infinidade de regiões de credibilidade com o mesmo grau de credibilidade e que normalmente interessa selecionar aquela que englobe todos



os valores de  $\theta$  mais credíveis *a posteriori*. Esta região é dita a região de alta densidade *a posteriori*, abreviadamente região HPD (Highest Probability Density).

### 2.2.1 - Inferência Bayesiana na Distribuição GEV

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 e os que existiam eram, em sua maioria, aplicados à distribuição Weibull, em estudos de análise de confiabilidade e com prioris não informativas, entre estes podem ser citados: Basu (1964), Bhattacharya (1967), Holla (1966) e Sinha & Sloan (1988). Posteriormente foram implementados trabalhos nesta mesma linha, porém com prioris informativas, com destaque para Berger & Sun (1993), Dey & Lee (1992), Smith & Naylor (1987), Singpurwalla & Song (1988).

Chechile (2001) apresenta uma função de densidade de probabilidade a priori conjugada para a distribuição de Gumbel e utiliza o algoritmo exato de Monte Carlo, para avaliar as propriedades da distribuição *a posteriori* dos parâmetros posição e escala. O autor destaca algumas vantagens da abordagem proposta por ele, em relação à abordagem clássica, entre estas, a aplicabilidade em pequenas amostras e a obtenção de intervalos HPD.

A aplicação da abordagem Bayesiana junto à distribuição GEV é algo mais recente ainda e, em um dos primeiros casos estudados, Coles & Tawn (1996) consideraram os conhecimentos de especialistas da área para formular informações *a priori* como base para uma análise Bayesiana de precipitações extremas de uma estação situada no sudoeste da Inglaterra. Os autores argumentam que a formulação da priori em termo dos parâmetros da GEV não é

uma abordagem prática e sugerem, como alternativa, a dedução da informação *a priori* por meio de quantis extremos, justificando que esta é uma medida (um valor) na qual é mais provável que um especialista possa quantificar mais precisamente suas crenças *a priori* sobre um comportamento extremo. Os autores ainda destacaram que é totalmente irreal não acreditar que um especialista possa fornecer informações importantíssimas *a priori*. No caso específico de precipitações máximas, um especialista possui um conhecimento extensivo em meteorologia e detalhes topográficos, que podem conduzir a uma melhor modelagem, quando se deseja antecipar chuvas extremas em uma região específica.

A particular formulação adotada por Coles & Tawn (1996) teve como princípio extrair (elicitare) a informação *a priori* em termos dos quantis da distribuição de precipitação máxima anual. Tendo que, se  $F(q_p) = 1 - p$ , em que  $F(\square)$  é a função de distribuição GEV, que tem o quantil para uma determinada probabilidade dado por:

$$q_p = \mu + \frac{\sigma(x_p^{-\xi} - 1)}{\xi}, \text{ em que } , x_p = -\log(1 - p).$$

Os autores sugerem 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,  $q_{p_1} < q_{p_2} < q_{p_3}$ . Posteriormente trabalhar com as diferenças  $(\tilde{q}_{p_1}, \tilde{q}_{p_2}, \tilde{q}_{p_3})$ , tal que,  $\tilde{q}_{p_i} = q_{p_i} - q_{p_{i-1}}$  para  $i = 1, 2, 3$ , com  $\tilde{q}_{p_0} = e_1$ , e  $e_1$  é o menor valor físico da variável em estudo. Distribuições independentes *a priori* para as diferenças de quantis são assumidas como:  $\tilde{q}_{p_i} \square \text{ gama } (\alpha_i, \beta_i)$ , com  $\alpha_i, \beta_i > 0$ . Segundo os autores as informações podem ser obtidas por meio de especialistas tendo conhecimento da mediana ou outro quantil de  $\tilde{q}_{p_i}$  para valores específicos de  $p_i$ .

Obtendo então *a priori* para  $\theta = (\mu, \sigma, \xi)$  dada por:

$$P(\theta) \propto J(\theta) \prod_{i=1}^3 \tilde{q}_i^{\alpha-1} \exp\{-\tilde{q}_i/\beta_i\},$$

em que  $J(\theta)$  é o Jacobiano da transformação de  $(\tilde{q}_{p1}, \tilde{q}_{p2}, \tilde{q}_{p3})$  para  $\theta = (\mu, \sigma, \xi)$ .

Coles & Powell (1996) construíram uma priori usando a distribuição normal multivariada em  $(\mu, \log \sigma, \xi)$ , baseando as estimativas destes hiperparâmetros em outras localizações, onde o processo de interesse foi observado. Utilizando a abordagem Bayesiana, como forma de incorporar informações *a priori* de outras regiões, Walshaw (2000) aplicou a distribuição *a priori* proposta por Coles & Powell (1996), para estudar via mistura de modelos a velocidade máxima de ventos em Boston e Key West (USA).

Uma recomendação de Coles & Pericchi (2003) é que, embora testes padrões possam indicar que a distribuição GEV possa ser reduzida para a distribuição Gumbel, deve-se tomar cuidado em se tomar essa decisão, pois isto é uma estratégia de risco, devido ao fato da distribuição GEV estimar melhor quantis extremos. Outra afirmação dos autores é que existe uma superioridade do ponto de vista prático em se adotar uma abordagem inferencial Bayesiana em estudos de extremos, visto que esta é uma técnica que permite expressar as incertezas de uma forma clara e realizar previsões.

Behrens et al. (2004) adotaram a priori proposta por Coles & Tawn (1996), modificando-a para ser usada na distribuição generalizada de Pareto (GPD) com estimação de threshold, sugerindo uma maneira alternativa na análise de eventos extremos, na qual a inferência é baseada na mistura de modelos, sendo o modelo Gama, para as observações anteriores ao threshold e a GPD, para as posteriores. Os autores aplicaram o modelo sugerido em dados de índices financeiros da Nasdaq 100 e a séries de dados simulados e obtiveram boas estimativas para os parâmetros, no estudo via simulação.

## 2.3 – Simulação Estocástica via Cadeias de Markov

Da distribuição *a posteriori* conjunta pode-se obter a distribuição marginal de cada parâmetro, integrando-se esta em relação aos outros parâmetros do modelo, porém, nem sempre é possível a obtenção das distribuições marginais correspondentes analiticamente.

Os métodos Monte Carlo via Cadeias de Markov (MCMC) são uma alternativa aos métodos não iterativos em problemas mais complexos. A idéia é obter uma amostra da distribuição de interesse e calcular estimativas amostrais de características desta distribuição. Dentre os métodos MCMC, os mais utilizados são o Amostrador de Gibbs e o Metropolis-Hastings. Será abordado apenas o último, que será utilizado no presente trabalho.

### 2.3.1 - Algoritmo Metropolis-Hastings

O algoritmo de Metropolis foi apresentado inicialmente por Metropolis et al. (1953) e generalizado por Hastings (1970), resultando no algoritmo Metropolis-Hastings. Este método é usado, geralmente, quando é difícil gerar amostras da condicional completa a posteriori. Neste caso, gera-se o valor do parâmetro, a partir de uma distribuição proposta e esse é aceito, ou não, com certa probabilidade de aceitação. Este mecanismo de correção garante a convergência da cadeia para a distribuição de equilíbrio.

Para descrever o algoritmo, suponha que a distribuição de interesse é a distribuição *a posteriori*  $P(\theta|Y)$  e que a distribuição proposta é  $q(\cdot|\theta)$ , o algoritmo pode ser descrito nos seguintes passos:

**Passo 1:** Atribua valor inicial ao parâmetro  $\theta = \theta^0$  e inicie o contador de iterações da cadeia em  $j = 0$ .

**Passo 2:** Obtenha um novo valor  $\theta^*$ , gerado a partir da função de transição  $q(\cdot|\theta)$ .

**Passo 3:** Calcule a probabilidade de aceitação do novo valor  $\alpha(\theta, \theta^*)$  dado por:

$$\alpha(\theta, \theta^*) = \min \left\{ 1, \frac{P(\theta^*|Y)q(\theta|\theta^*)}{P(\theta|Y)q(\theta^*|\theta)} \right\}.$$

**Passo 4:** Gere  $u$  de uma distribuição Uniforme[0,1] e faça  $\theta^j = \theta^*$  se  $u < \alpha$ , caso contrário rejeite e faça  $\theta^j = \theta^{j-1}$ .

**Passo 5:** Incremente o contador  $j$  e volte ao passo 2.

Assim para  $j$  suficientemente grande,  $\theta^1, \dots, \theta^j$  é uma amostra da distribuição *a posteriori*  $P(\theta|Y)$ . Para o caso vetorial  $\theta = (\theta_1, \dots, \theta_k)$  a densidade de transição será dada por  $q(\cdot|\theta)$  e uma probabilidade de aceitação dada por  $\alpha(\theta, \theta^*)$ , e deve-se proceder da mesma forma.

Segundo Hastings (1970), as principais características deste método para amostragem de uma distribuição com densidade  $p(x)$  são:

- (a) A computação depende apenas de  $p(x)$ , somente por meio das razões na forma  $p(x')/p(x)$ , em que  $x'$  e  $x$  são pontos amostrados. Logo a constante normalizadora não precisa ser conhecida e não se faz necessária a fatorização de  $p(x)$ .
- (b) Uma seqüência de amostras é obtida pela simulação da cadeia de Markov. As amostras obtidas são, por conseqüência, correlacionadas, e os cálculos de erros, relacionados às estimativas, podem requerer mais cuidado do que apenas com amostras independentes.

Maiores detalhes sobre este algoritmo podem ser encontrados em Gamerman (1996), Ghib & Greenberg (1995) e Hastings (1970).

### 2.3.2 – Verificação de Convergência

Os métodos de MCMC são uma ótima ferramenta para resolução de muitas dificuldades práticas na análise Bayesiana, porém, alguns problemas quanto ao seu uso podem surgir. De acordo com Cowles & Carlin (1996), pode-se destacar: a incerteza sobre qual o número necessário de iterações para que se garanta que a cadeia convergiu para o estado de equilíbrio, a possibilidade das iterações iniciais da amostra serem influenciadas pelos valores iniciais dos parâmetros e, ainda, o fato das seqüências de valores apresentarem correlações entre si.

Não existe uma técnica geral para resolver estas questões, mas pode-se sempre avaliar a convergência das cadeias, detectando problemas fora do período de aquecimento. Para eliminar uma possível autocorrelação das cadeias, deve-se selecionar a partir do burn-in (que é o descarte da parte inicial das cadeias) um valor a cada  $k$  iterações. Uma análise de convergência em métodos de simulação pode ser feita preliminarmente, analisando os gráficos ou medidas descritivas dos valores. Os gráficos mais freqüentes são o gráfico de  $\theta$  ao longo das iterações (traço) e um gráfico da estimativa da distribuição a posteriori de  $\theta$  (um histograma ou uma densidade kernel). As estatísticas usuais são a média, o desvio padrão e os quantis.

Uma outra avaliação de convergência em métodos de MCMC pode ser feita, usando algumas técnicas de diagnóstico de convergência. As técnicas mais conhecidas são descritas por Geweke (1992) e Heidelberger & Welch (1983), que usam resultados baseados em análise espectral, Raftery & Lewis (1992) que permitem calcular quantas iterações são necessárias para uma cadeia atingir a distribuição estacionária, por meio da estimação de quantis a posteriori com uma precisão previamente fixada e Gelman & Rubin (1992), que utilizaram resultados baseados na análise de variância clássica para duas ou mais

cadeias simuladas com valores iniciais diferentes. Estes métodos e outros foram comparados no trabalho de Cowles & Carlin (1996), que concluíram que não se pode afirmar qual deles é o mais eficiente. Vale destacar que as técnicas de Geweke, Heidelberger-Welch, Raftery-Lewis, Gelman-Rubin e outras estão implementadas no pacote BOA, executável no freeware R.

Além da análise gráfica e das medidas descritivas, os diagnósticos de Geweke (1992) e Raftery & Lewis (1992) serão usados neste trabalho para avaliar a convergência das cadeias geradas.

### 2.3.2.1 - Diagnóstico de Geweke

O diagnóstico de Geweke (1992) foi desenvolvido para indicar a convergência da média *a posteriori* da quantidade amostrada,  $g(\theta)$ , em uma única cadeia, isto é, os valores  $g(\theta^{(i)})$  são calculados a cada iteração da simulação, formando uma série temporal. A partir da cadeia, a variância assintótica  $S_g(0)$  da medida de  $g(\theta)$  pode ser estimada. De uma forma resumida, o diagnóstico consiste nos seguintes passos:

**Passo 1:** Divide-se a cadeia gerada de tamanho  $N$  em duas seqüências, em que a primeira possui as  $N_a$  primeiras iterações e a segunda, as  $N_b$  últimas iterações. Geweke (1992) sugere que as médias devem ser construídas após algumas iterações iniciais terem sido descartadas e que sejam usados os valores de  $N_a = 0,1N$  e  $N_b = 0,5N$ .

**Passo 2:** Calculam-se as médias  $\hat{g}_a$  e  $\hat{g}_b$  e as variâncias assintóticas  $\hat{S}_a(0)$  e  $\hat{S}_b(0)$ , respectivamente. As variâncias são determinadas pela estimação da densidade espectral.

Pode-se mostrar que, se as razões  $\frac{N_a}{N}$  e  $\frac{N_b}{N}$  são fixas e  $N \rightarrow \infty$ ,

$$\frac{\hat{g}_a - \hat{g}_b}{\sqrt{\frac{\hat{S}_a(0)}{N_a} + \frac{\hat{S}_b(0)}{N_b}}} \square N(0,1).$$

Logo, um teste pode ser construído e, se a diferença padronizada entre as médias for muito grande, existe indicação de ausência de convergência.

### 2.3.2.2 - Critério de Raftery e Lewis

O critério sugerido por Raftery & Lewis (1992) é um método que estima quantas iterações são necessárias para que a cadeia apresente convergência, propondo quantidades necessárias para isto. Deve-se ter uma seqüência da cadeia com um  $N$  mínimo, que é o número mínimo de iterações requeridas para se obter a precisão necessária de estimação. O método fornece as estimativas do “burn-in”, o número de iterações que devem ser computadas  $N$  (total), a distância mínima de uma iteração a outra, para se obter a subamostra aproximadamente independente “thin”. Esses valores são calculados mediante especificações para garantir que um quantil  $u$  de uma determinada função de densidade  $f(\theta)$  seja estimada com precisão pré definida. Outra saída importante é o fator de dependência, que é responsável pelo acréscimo multiplicativo ao número de iterações necessárias para se alcançar a convergência, devido à autocorrelação dentro da seqüência. Segundo Raftery & Lewis (1992), se este fator for maior que 5, pode-se dizer que a convergência ainda não foi obtida, necessitando reparametrizar o modelo em estudo.

Uma revisão mais detalhada sobre diagnósticos de convergência pode ser vista em Cowles & Carlin, (1996) e Nogueira (2004).



## 2.4 - Estudos de Precipitações Máximas

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, seus efeitos passam, geralmente, 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 represas, soterramentos, entre outros, levando, em alguns casos, a perdas de vidas humanas.

A precipitação é entendida em Hidrologia como toda água na forma líquida ou sólida, proveniente do meio atmosférico, que atinge a superfície terrestre. Esta é a maior classe de hidrometeoros, devendo ser distinguida de outras classes tais como a de nuvens, a de nevoeiros, a de orvalho, a de geada, entre outras. Vale ressaltar que os dois últimos tipos de hidrometeoros, orvalho e geada, formam-se junto ao solo, não devendo, portanto, ser confundidos com precipitação (Pinto, 1995). A precipitação pluvial é medida em milímetros (mm) de altura de lâmina d'água, para uma interpretação prática, considerando que tenha ocorrido uma precipitação pluvial de 100 mm, isto equivale ao acúmulo de 100 litros de água em 1m<sup>2</sup>.

Segundo Tucci (2001), a precipitação máxima é entendida como a ocorrência extrema, com duração, distribuição temporal e espacial crítica para uma área ou bacia hidrográfica. A disponibilidade de longas séries de precipitações é, em geral, muito mais freqüente que a de vazão e o estudo das precipitações máximas é um dos caminhos para conhecer-se a vazão de enchente de uma bacia.

O conhecimento da precipitação máxima provável que, de acordo com Tucci (2001), 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. 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).

Muitas organizações vinculadas à segurança de barragens recomendam explicitamente o estudo de precipitações máximas para o caso de grandes obras, nas quais o galgamento envolve grandes riscos (Tucci 2001).

Dentre sua vasta aplicação, a distribuição GEV tem sido utilizada com grande frequência em estudos de fenômenos ambientais, principalmente para solucionar problemas relacionados a áreas de Engenharia, entre os quais, o estudo da precipitação pluvial máxima.

Analisando a série de precipitação máxima diária compreendida no período de 1951 a 1999 da Venezuela, Coles & Pericchi (2003) tentaram prever uma precipitação de 410mm, que causou mais de 30 mil mortes em 1999. Os autores verificaram que a redução do modelo GEV para o modelo de Gumbel era aceitável pelo teste da razão de verossimilhança, ou seja, a distribuição de Gumbel tinha sido considerada perfeitamente adequada, como um modelo, para determinar os parâmetros relacionados aos eventos extremos. Porém, uma implicação foi encontrada com referência ao quantil extremo, os autores verificaram que o modelo de Gumbel foi totalmente ineficiente na predição, enquanto a distribuição GEV teve um resultado mais aceitável.

Na grande maioria dos estudos de precipitações máximas, nem sempre os pesquisadores têm se preocupado em determinar uma estimativa intervalar para as precipitações máximas prováveis, obtendo apenas estimativas pontuais para certos períodos de retorno, as quais podem não ser os melhores valores a serem adotados em cálculos de engenharia. (Entende-se por período de retorno o intervalo médio de tempo em longo prazo, ou número de anos que separam um evento de dimensão conhecida de outro evento com dimensão igual ou superior).

Vargas et al. (1999) analisaram o erro associado a alguns métodos de estimação de quantis máximos mais utilizados em estudos de hidrologia aplicada no Chile, sugerindo trabalhar com limite superior a certo nível de confiança, afirmando que este é mais seguro, e é mais improvável de que seja superado, de acordo com o risco de falha aceitável para a obra.

## 2.5 - Teste para Desvios na Aleatoriedade

Uma das pressuposições, para que se possa utilizar a metodologia para a estimação dos parâmetros da distribuição GEV, é que as observações de cada período, ano ou mês, sejam, ao longo do tempo, independentes. Logo, para testar a hipótese de que a série de  $n$  observações amostrais  $(x_1, x_2, \dots, x_n)$ , para um certo período, é aleatória, foi utilizado o teste conhecido como “*run test*”, apresentado por Lehmann (1975), o qual é utilizado na detecção de desvios na aleatoriedade de uma seqüência de medições quantitativas no tempo, ocasionados por tendência ou periodicidade.

Como passo inicial deste teste, obtém-se o valor da mediana ( $M_d$ ), em seguida, obtém-se, para os dados sem ordenar, os valores da variável indicadora  $I_i (i = 1, 2, \dots, n)$ , definida por:

$$I_i = \begin{cases} 1 & \text{se } x_i \geq M_d \\ -1 & \text{se } x_i < M_d \end{cases}.$$

Desta forma, é gerado um vetor com seqüências  $(I_1, I_2, \dots, I_n)$  de números 1 (um) positivos e negativos, para as quais se testa a aleatoriedade.

Seja  $K$  a variável aleatória número total de seqüências de uns negativos e positivos na série  $(I_1, I_2, \dots, I_n)$ . Sob a hipótese de que a seqüência é aleatória a variável  $k$  é uma variável aleatória com:

Média  $= (n/2) + 1$ , e variância  $= (n(n-2))/(4(n-1))$ .

Se o valor  $k$  for significativamente diferente da média, ou seja,  $P(K < k_c) < \alpha$ , rejeita-se a hipótese de aleatoriedade.

## 3 - MATERIAL E MÉTODOS

### 3.1 - Dados Simulados

A fim de que diversas situações pudessem ser avaliadas usando-se a distribuição *a priori* proposta, foram geradas via simulação 36 séries de dados para diferentes tamanhos de amostras, e diferentes valores dos parâmetros da distribuição GEV, adotando-se os seguintes valores:

- Tamanho de amostras:  $n = (30; 50; 100)$ .
- Parâmetro posição  $\mu = (40; 70)$ .
- Parâmetro escala  $\sigma = (10; 20)$ .
- Parâmetro forma  $\xi = (-0,30; 0,10; 0,20)$ .

Estes valores foram escolhidos com base em valores médios que ocorrem em dados meteorológicos extremos (precipitações máximas e ventos máximos), encontrados na literatura. Procurou-se, também, com o tamanho das amostras, representar situações reais. O tamanho  $n=30$  representa a faixa mínima recomendada pela Organização Mundial de Meteorologia, que é de 30 anos. O tamanho máximo, igual a 100, é devido ao fato de que não se tem conhecimento de séries históricas de dados meteorológicos com duração superior a este valor. Outro fato a ser destacado são as propriedades assintóticas da distribuição GEV, indicando que, se os resultados forem satisfatórios com  $n=100$ , amostras maiores tenderão a apresentar resultados iguais ou melhores.

As simulações foram realizadas, utilizando-se o pacote “EVD” do Sistema Computacional Estatístico R, conforme R Development Core Team (2004), que possui funções específicas para esta atividade.

### **3.2 – Dados de Precipitação Máxima de Jaboticabal**

Os dados utilizados foram obtidos a partir dos registros pluviográficos fornecidos pelo Departamento de Ciências Exatas da Faculdade de Ciências Agrárias e Veterinárias da UNESP, campus Jaboticabal, estado de São Paulo. A estação Agrometeorológica do Campus encontra-se nas seguintes coordenadas geográficas; latitude de 21<sup>o</sup>15'22''S, longitude 48<sup>o</sup>18'58''W e altitude de 595m. O clima da região é Cwa, de acordo com a classificação climática de Köppen, com chuvas de verão e relativamente seco no inverno (Vasconcellos, 1998).

As observações referem-se às precipitações pluviárias 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, estando de acordo com a faixa mínima recomendada pela Organização Mundial de Meteorologia que é, segundo Pinto (1999), de 30 anos.

Os dados foram agrupados em períodos de um ano e mensal; utilizaram-se apenas os meses da estação chuvosa (de setembro a março), extraíndo-se a máxima precipitação pluvial diária, observada de cada período, formaram-se os vetores de máximos. Foram utilizadas as observações de 1956 a 1990 para calcular a precipitação pluvial 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, realizou-se o teste para desvio de aleatoriedade, conforme descrito na seção (2.5) e para a sua realização foi utilizado o pacote TSERIES do Sistema Computacional Estatístico R.

### 3.3 – Função de Densidade de Probabilidades da Distribuição GEV

Derivando-se as funções de distribuição acumulada de probabilidade dada em (1) e (2) em relação a  $x$ , obtém-se, respectivamente, a função de densidade de probabilidade da distribuição GEV, dada por:

$$f(x) = \frac{1}{\sigma} \left\{ \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{\frac{1+\xi}{\xi}} \exp \left\{ - \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\} \right\}, \quad (4)$$

definida em,  $-\infty < x < \mu - \frac{\sigma}{\xi}$ , para  $\xi < 0$  e  $\mu - \frac{\sigma}{\xi} < x < \infty$ , para  $\xi > 0$ , e a função de densidade de probabilidade Gumbel, dada por:

$$f(x) = \frac{1}{\sigma} \exp \left\{ - \left( \frac{x - \mu}{\sigma} \right) - \exp \left[ - \left( \frac{x - \mu}{\sigma} \right) \right] \right\}, \quad (5)$$

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

### 3.4 - Função de Verossimilhança

Seja  $X_1, X_2, \dots, X_n$  uma série de  $n$  variáveis aleatórias, independentes e identicamente distribuídas, com distribuição GEV e  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  uma série de observações. Supondo que há independência entre as observações, a função de verossimilhança é dada por,

$$L(\mathbf{x}|\theta) = \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{\frac{1+\xi}{\xi}} \right\} \exp \left\{ \sum_{i=1}^n \left\{ - \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\} \right\} \quad (6)$$

em que,  $L(\mathbf{g})$  assume valores diferentes de zero, para  $\xi < 0$ , se todos os valores de  $x_i$  ( $i=1,2,\dots,n$ ) forem menores do que  $\mu - \frac{\sigma}{\xi}$ , ou seja, se  $\mu - \frac{\sigma}{\xi} > x_{(n)}$ , sendo  $x_{(n)}$  o maior valor da série de observações, e para  $\xi > 0$ , se todos os valores  $x_i$  ( $i=1,2,\dots,n$ ) forem maiores do que  $\mu - \frac{\sigma}{\xi}$ , ou seja,  $\mu - \frac{\sigma}{\xi} < x_{(1)}$ , sendo  $x_{(1)}$  o menor valor da série de observações. Caso contrário,  $L(\mathbf{x}|\mu, \sigma, \xi) = 0$ .

Logo, o logaritmo da função de verossimilhança é dado por:

$$\begin{aligned}
l(\mathbf{x}|\mu, \sigma, \xi) &= \ln[L(\mathbf{x}|\mu, \sigma, \xi)] \\
&= -n \ln \sigma - \left(\frac{1+\xi}{\xi}\right) \sum_{i=1}^n \ln \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^n \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \\
&= \sum_{i=1}^n \left\{ -\ln \sigma - \left(\frac{1+\xi}{\xi}\right) \ln \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right] - \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \right\}, \quad (7)
\end{aligned}$$

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.

No caso particular da distribuição GEV com  $\xi \rightarrow 0$ , ou seja, a distribuição de Gumbel, tem-se que o logaritmo da função de verossimilhança é dados por:

$$l(\mathbf{x}|\mu, \sigma) = \sum_{i=1}^n \left\{ -\ln \sigma - \left(\frac{x_i - \mu}{\sigma}\right) - \exp\left(\frac{x_i - \mu}{\sigma}\right) \right\}. \quad (8)$$



### 3.5 – Probabilidades de Ocorrência de Precipitações Máximas

A probabilidade  $P$  de que ocorra uma precipitação pluvial máxima maior que um certo valor  $y$  é estimada, utilizando-se a função de distribuição acumulada da seguinte maneira:

$$P(Y > y) = 1 - F(y) = 1 - \exp \left\{ - \left[ 1 + \xi \left( \frac{y - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\}. \quad (9)$$

Cujo limite para  $\hat{\xi}$  tendendo a zero é dada por:

$$P(Y > y) = 1 - \exp \left\{ - \exp \left[ - \left( \frac{y - \mu}{\sigma} \right) \right] \right\}, \quad (10)$$

em que,  $y$  é um valor precipitação de máxima do período e  $0 < y < \infty$ .

#### 3.5.1 – Estimação do Período de Retorno

Seja  $A$  um evento e  $T$  o tempo aleatório entre ocorrências consecutivas de eventos  $A$ . O valor médio  $\tau$  da variável  $T$  é denominado período de retorno do evento  $A$ . No caso em estudo,  $A$  é o evento: “precipitação máxima que excede um determinado valor  $x$ ”, cuja probabilidade  $P(A)$  é dada por  $1 - F(y)$ . Logo, o período de retorno para esse evento é dado por:

$$\tau = \frac{1}{P(A)} = \frac{1}{1 - F(y)},$$

sendo  $\tau$  geralmente expresso em anos.

### 3.5.2 – Estimação do Nível de Retorno da Distribuição GEV

O nível de retorno ( $y_t$ ), associado ao período  $\tau$ , é obtido a partir da solução da equação:

$$\int_{-\infty}^{y_t} f(y; \theta) dx = 1 - p,$$

em que  $p = 1/\tau$ , ou seja,

$$F(q_p) = 1 - p. \quad (11)$$

Ao inverter a equação acima e relacionando-a com (9), chega-se a  $y_t$ , que equivale ao quantil da distribuição GEV, ou seja, aplicando-se logaritmo neperiano, tem-se,

$$\ln \left\{ \exp \left[ - \left[ 1 + \xi \left( \frac{y - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} \right] \right\} = \ln(1 - p) \Rightarrow \left[ 1 + \xi \left( \frac{y - \mu}{\sigma} \right) \right]^{\frac{1}{\xi}} = -\ln(1 - p),$$

elevando os dois lados da igualdade por  $(-\xi)$  e isolando-se  $y$  obtém-se,

$$y_t = q_p = \mu + \frac{\sigma}{\xi} \left[ (x_p)^{-\xi} - 1 \right], \quad (12)$$

para  $\xi \neq 0$ , e cujo limite para  $\xi$ , tendendo a zero, é dado por:

$$q_p = \mu - \sigma \left[ \ln(x_p) \right], \quad (13)$$

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

A estimativa  $\hat{y}_t$  do nível de retorno  $y_t$  para períodos de retorno  $p = 1/\tau$ , é obtida por substituição das estimativas de  $\mu, \sigma$  e  $\xi$ .

#### 4 – CONSTRUÇÃO DA DISTRIBUIÇÃO A *PRIORI* FUNDAMENTADA EM QUANTIS EXTREMOS

Dados de valores extremos são normalmente escassos, principalmente em dados meteorológicos; logo, informações de “especialistas” podem ser essenciais para a complementação das informações contidas nos dados. Portanto, é razoável esperar que um “especialista” 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*.

Sendo assim, propõe-se neste trabalho uma *a priori* baseada na abordagem utilizada por Coles & Tawn (1996), que trata da extração de informações com uma parametrização com qual os especialistas estão mais familiarizados, que é utilização de quantis. Além da vantagem da familiaridade com os quantis extremos, os mesmos autores ainda ressaltam o fato de que é improvável que o conhecimento *a priori* do comportamento extremo possa ser extraído diretamente em termos dos parâmetros da distribuição GEV. Mesmo se a forma para cada parâmetro estiver disponível, isto pode não significar qual é claramente *a priori* conjunta apropriada, destacando-se a dificuldade de se identificar certas dependências entre os parâmetros escala  $\sigma$  e forma  $\xi$ .

Conseqüentemente, devem-se evitar situações em que se adotam *prioris* independentes para os parâmetros de distribuição GEV, como sugerido por Smith & Naylor (1987). Para se extrair informações de um especialista com uma parametrização, que corresponda a uma escala com qual ele tenha familiaridade e com a qual a dependência natural entre as especificações *a priori* seja construída, a utilização de quantis parece ser mais coerente. A abordagem dos

quantis foi primeiramente utilizada por Crowder (1992), que sugeriu a construção da priori no espaço de probabilidades para quantis fixos.

Adotando os quantis como forma de incorporar o conhecimento a priori tem-se que  $F(q_p) = 1 - p$ , em que  $F(\square)$  é a função de distribuição GEV dada em (1) que tem o quantil para uma determinada probabilidade dado por:

$$q_p = \mu + \frac{\sigma}{\xi} (x_p^{-\xi} - 1), \text{ em que, } x_p = -\ln(1 - p), \quad (14)$$

Sugere-se 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 desses quantis são independentes e admitindo como priori marginal a distribuição de Gumbel com os parâmetros  $a_i$  (posição) e  $b_i$  (escala), tem-se:

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

Existe uma justificativa sobre esta construção. Como se está trabalhando 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:

$$P(\theta) \propto J(\theta) \prod_{i=1}^3 \frac{1}{b_i} \exp \left\{ - \left( \frac{q_{p_i} - a_i}{b_i} \right) - \exp \left[ - \left( \frac{q_{p_i} - a_i}{b_i} \right) \right] \right\}, \quad (15)$$

em que,  $J(\theta)$  é o Jacobiano da transformação de  $(q_{p_1}, q_{p_2}, q_{p_3})$  para  $\theta = (\mu, \sigma, \xi)$ , dado por:

$$J(\theta) = \frac{\sigma}{\xi^2} \left| \sum_{\substack{i,j \in \{1,2,3\} \\ i < j}} (-1)^{i+j} (x_i x_j)^{-\xi} \ln(x_j/x_i) \right|, \quad (16)$$

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

A demonstração da expressão (16) é dada a seguir:

### Prova

Seguindo a definição de jacobiano, dada por Mood et al. (1974), tem-se que:

$$J(\theta) = \begin{vmatrix} -\frac{\sigma}{\xi^2} [\xi x_1^{-\xi} \ln(x_1) + x_1^{-\xi} - 1] & \frac{x_1^{-\xi}}{\xi} - \frac{1}{\xi} & 1 \\ -\frac{\sigma}{\xi^2} [\xi x_2^{-\xi} \ln(x_2) + x_2^{-\xi} - 1] & \frac{x_2^{-\xi}}{\xi} - \frac{1}{\xi} & 1 \\ -\frac{\sigma}{\xi^2} [\xi x_3^{-\xi} \ln(x_3) + x_3^{-\xi} - 1] & \frac{x_3^{-\xi}}{\xi} - \frac{1}{\xi} & 1 \end{vmatrix}$$

Calculando o determinante, tem-se:

$$\begin{aligned} J(\theta) = & \left\{ -\frac{\sigma}{\xi^2} [\xi x_1^{-\xi} \ln(x_1) + x_1^{-\xi} - 1] \left[ \frac{x_2^{-\xi}}{\xi} - \frac{1}{\xi} \right] \right\} + \\ & + \left\{ -\frac{\sigma}{\xi^2} [\xi x_2^{-\xi} \ln(x_2) + x_2^{-\xi} - 1] \left[ \frac{x_3^{-\xi}}{\xi} - \frac{1}{\xi} \right] \right\} + \\ & + \left\{ -\frac{\sigma}{\xi^2} [\xi x_3^{-\xi} \ln(x_3) + x_3^{-\xi} - 1] \left[ \frac{x_1^{-\xi}}{\xi} - \frac{1}{\xi} \right] \right\} + \\ & - \left\{ -\frac{\sigma}{\xi^2} [\xi x_3^{-\xi} \ln(x_3) + x_3^{-\xi} - 1] \left[ \frac{x_2^{-\xi}}{\xi} - \frac{1}{\xi} \right] \right\} + \\ & - \left\{ -\frac{\sigma}{\xi^2} [\xi x_1^{-\xi} \ln(x_1) + x_1^{-\xi} - 1] \left[ \frac{x_3^{-\xi}}{\xi} - \frac{1}{\xi} \right] \right\} + \\ & - \left\{ -\frac{\sigma}{\xi^2} [\xi x_2^{-\xi} \ln(x_2) + x_2^{-\xi} - 1] \left[ \frac{x_1^{-\xi}}{\xi} - \frac{1}{\xi} \right] \right\}. \end{aligned}$$

Efetuada-se o produto, tem-se que:

$$\begin{aligned}
J(\theta) = & \left| \frac{\sigma}{\xi^2} \left\{ \left[ \frac{\xi(x_1x_2)^{-\xi}}{\xi} \ln(x_1) + \frac{(x_1x_2)^{-\xi}}{\xi} - \frac{x_2^{-\xi}}{\xi} - \frac{\xi x_1^{-\xi}}{\xi} \ln(x_1) - \frac{x_1^{-\xi}}{\xi} + \frac{1}{\xi} \right] + \right. \\
& + \left[ \frac{\xi(x_2x_3)^{-\xi}}{\xi} \ln(x_2) + \frac{(x_2x_3)^{-\xi}}{\xi} - \frac{x_3^{-\xi}}{\xi} - \frac{\xi x_2^{-\xi}}{\xi} \ln(x_2) - \frac{x_2^{-\xi}}{\xi} + \frac{1}{\xi} \right] + \\
& + \left[ \frac{\xi(x_1x_3)^{-\xi}}{\xi} \ln(x_3) + \frac{(x_1x_3)^{-\xi}}{\xi} - \frac{x_1^{-\xi}}{\xi} - \frac{\xi x_3^{-\xi}}{\xi} \ln(x_3) - \frac{x_3^{-\xi}}{\xi} + \frac{1}{\xi} \right] + \\
& - \left[ \frac{\xi(x_2x_3)^{-\xi}}{\xi} \ln(x_3) + \frac{(x_2x_3)^{-\xi}}{\xi} - \frac{x_2^{-\xi}}{\xi} - \frac{\xi x_3^{-\xi}}{\xi} \ln(x_3) - \frac{x_3^{-\xi}}{\xi} + \frac{1}{\xi} \right] + \\
& - \left[ \frac{\xi(x_1x_3)^{-\xi}}{\xi} \ln(x_1) + \frac{(x_1x_3)^{-\xi}}{\xi} - \frac{x_3^{-\xi}}{\xi} - \frac{\xi x_1^{-\xi}}{\xi} \ln(x_1) - \frac{x_1^{-\xi}}{\xi} + \frac{1}{\xi} \right] + \\
& \left. - \left[ \frac{\xi(x_1x_2)^{-\xi}}{\xi} \ln(x_2) + \frac{(x_1x_2)^{-\xi}}{\xi} - \frac{x_1^{-\xi}}{\xi} - \frac{\xi x_2^{-\xi}}{\xi} \ln(x_2) - \frac{x_2^{-\xi}}{\xi} + \frac{1}{\xi} \right] \right\} \right| \quad (17)
\end{aligned}$$

Simplificando, obtém-se

$$\begin{aligned}
J(\theta) = & \frac{\sigma}{\xi^2} \left| \left\{ -(x_1x_2)^{-\xi} [\ln(x_2) - \ln(x_1)] + (x_1x_3)^{-\xi} [\ln(x_3) - \ln(x_1)] + \right. \right. \\
& \left. \left. -(x_2x_3)^{-\xi} [\ln(x_3) - \ln(x_2)] \right\} \right|.
\end{aligned}$$

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

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

conforme se queria demonstrar.

Agora, quando  $\xi=0$ , a distribuição a priori é definida por continuidade, usando-se:

$$\lim_{\xi \rightarrow 0} q_{pi} = \mu - \sigma \ln(x_i), \quad i = 1, 2, 3.$$

e

$$\lim_{\xi \rightarrow 0} J(\theta) = \frac{\sigma}{2} \left| \sum_{\substack{i,j \in \{1,2,3\} \\ i < j}} (-1)^{i+j} \ln(x_i) \ln(x_j) \ln(x_j/x_i) \right|. \quad (18)$$

Isto é demonstrado a seguir.

**Prova**

Outra forma de simplificar a expressão (17) é da seguinte maneira;

$$J(\theta) = \left[ \overbrace{-\frac{\sigma}{\xi^2} \left\{ \ln(x_1) \left[ (x_1 x_2)^{-\xi} - x_1^{-\xi} \right] - \ln(x_2) \left[ (x_1 x_2)^{-\xi} - x_2^{-\xi} \right] \right\}}^{K1} + \right. \\ \left. \overbrace{-\frac{\sigma}{\xi^2} \left\{ \ln(x_3) \left[ (x_1 x_3)^{-\xi} - x_3^{-\xi} \right] - \ln(x_1) \left[ (x_1 x_3)^{-\xi} - x_1^{-\xi} \right] \right\}}^{K2} + \right. \\ \left. \overbrace{-\frac{\sigma}{\xi^2} \left\{ \ln(x_2) \left[ (x_2 x_3)^{-\xi} - x_2^{-\xi} \right] - \ln(x_3) \left[ (x_2 x_3)^{-\xi} - x_3^{-\xi} \right] \right\}}^{K3} \right].$$

$$\text{Logo, } \lim_{\xi \rightarrow 0} J(\theta) = \lim_{\xi \rightarrow 0} K1 + \lim_{\xi \rightarrow 0} K2 + \lim_{\xi \rightarrow 0} K3$$

Resolvendo  $\lim_{\xi \rightarrow 0} K1$ , tem-se,

$$\lim_{\xi \rightarrow 0} K1 = \lim_{\xi \rightarrow 0} \left[ -\frac{\sigma}{\xi^2} \left\{ \ln(x_1) \left[ (x_1 x_2)^{-\xi} - x_1^{-\xi} \right] - \ln(x_2) \left[ (x_1 x_2)^{-\xi} - x_2^{-\xi} \right] \right\} \right].$$

Aplicando-se a regra de l'Hôpital, de acordo com Larson et al.(1998), tem-se:

$$\lim_{\xi \rightarrow 0} K1 = \sigma \lim_{\xi \rightarrow 0} \left[ -\ln(x_1) \left[ \frac{(x_1 x_2)^{-\xi} \ln(x_1 x_2)^{-1} - x_1^{-\xi} \ln(x_1)^{-1}}{2\xi} \right] + \right. \\ \left. + \ln(x_2) \left[ \frac{(x_1 x_2)^{-\xi} \ln(x_1 x_2)^{-1} - x_2^{-\xi} \ln(x_2)^{-1}}{2\xi} \right] \right],$$

Novamente, aplicando-se a regra de l'Hôpital, obtém-se:

$$\lim_{\xi \rightarrow 0} K1 = \frac{\sigma}{2} \lim_{\xi \rightarrow 0} \left| -\ln(x_1) \left[ (x_1 x_2)^{-\xi} \ln(x_1 x_2) \ln(x_1 x_2) - x_1^{-\xi} \ln(x_1) \ln(x_1) \right] + \right. \\ \left. + \ln(x_2) \left[ (x_1 x_2)^{-\xi} \ln(x_1 x_2) \ln(x_1 x_2) - x_2^{-\xi} \ln(x_2) \ln(x_2) \right] \right|,$$

resolvendo-se o limite, chega-se ao seguinte resultado:

$$\lim_{\xi \rightarrow 0} K1 = \frac{\sigma}{2} \left| -\ln(x_1) \left\{ [\ln(x_1 x_2)]^2 - [\ln(x_1)]^2 \right\} + \right. \\ \left. + \ln(x_2) \left\{ [\ln(x_1 x_2)]^2 - [\ln(x_2)]^2 \right\} \right|.$$

Empregando-se a propriedade de logaritmo e simplificando, obtém-se;

$$\lim_{\xi \rightarrow 0} K1 = \frac{\sigma}{2} \left| -\ln(x_1) \left\{ [\ln(x_1)]^2 + [\ln(x_2)]^2 + 2 \ln(x_1) \ln(x_2) - [\ln(x_1)]^2 \right\} + \right. \\ \left. + \ln(x_2) \left\{ [\ln(x_2)]^2 + [\ln(x_1)]^2 + 2 \ln(x_1) \ln(x_2) - [\ln(x_2)]^2 \right\} \right|,$$

logo,

$$\lim_{\xi \rightarrow 0} K1 = \frac{\sigma}{2} \left| \ln(x_1) [\ln(x_2)]^2 - \ln(x_2) [\ln(x_1)]^2 \right| \\ \lim_{\xi \rightarrow 0} K1 = \frac{\sigma}{2} \left| -\ln(x_1) \ln(x_2) [\ln(x_2) - \ln(x_1)] \right|.$$

Por analogia, pode-se calcular  $\lim_{\xi \rightarrow 0} K2$  e  $\lim_{\xi \rightarrow 0} K3$  que se obtém o seguinte resultado:

$$\lim_{\xi \rightarrow 0} K2 = \frac{\sigma}{2} \left| \ln(x_1) \ln(x_3) [\ln(x_3) - \ln(x_1)] \right| \text{ e}$$

$$\lim_{\xi \rightarrow 0} K3 = \frac{\sigma}{2} \left| -\ln(x_2) \ln(x_3) [\ln(x_3) - \ln(x_2)] \right|.$$

Portanto,  $\lim_{\xi \rightarrow 0} J(\theta)$  pode ser expresso por:

$$\lim_{\xi \rightarrow 0} J(\theta) = \frac{\sigma}{2} \left| -\ln(x_1) \ln(x_2) [\ln(x_2) - \ln(x_1)] + \right. \\ \left. + \ln(x_1) \ln(x_3) [\ln(x_3) - \ln(x_1)] + \right. \\ \left. - \ln(x_2) \ln(x_3) [\ln(x_3) - \ln(x_2)] \right|,$$



ou,

$$\lim_{\xi \rightarrow 0} J(\theta) = \frac{\sigma}{2} \left| \sum_{\substack{i,j \in \{1,2,3\} \\ i < j}} (-1)^{i+j} \ln(x_i) \ln(x_j) \ln(x_j/x_i) \right|.$$

#### 4.1 - Incorporação das Informações

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 especialistas 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 estes valores.

Para exemplificar esta aplicação, considerando o estudo dos dados de precipitação máxima de Jaboticabal, especificamente para o período anual, tomaram-se, como informações *a priori*, certas características da análise de precipitações máximas de Lavras-MG, no 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. Baseando-se nas características dos dados, esperou-se que para os períodos de retorno de 10, 100 e 200 anos, ou seja, para níveis de retorno com  $p = (0,1 ; 0,01 \text{ e } 0,005)$  as precipitações máximas apresentem os seguintes valores de medianas e quantis de 90% (vide Tabela 1).

**TABELA 1:** Informações *a priori* e valores calculados das medianas e dos quantis 90% para  $q_{pi}$  com os respectivos parâmetros da Gumbel, associados à distribuição *a priori* para análise do período anual.

Tempo de retorno (anos)	Valores a priori (mm)		Parâmetros da Gumbel		Valores calculados (mm)	
	Mediana	Quantil 90%	$a_i$	$b_i$	Mediana	Quantil 90%
10	60,0	70,0	60,0	3,0	60,7	64,5
100	95,0	110,0	90,0	7,0	92,6	105,8
200	120,0	140,0	115,0	10,0	118,6	137,5

Vale destacar que não existe uma expressão para obter os valores de medianas e quantis de 90%, os valores aqui utilizados foram obtidos junto aos autores citados acima, baseando-se em seus conhecimentos.

Mas uma idéia de como estes valores foram obtidos pode ser expressa em termos das estimativas dos parâmetros de posição e escala da distribuição de Gumbel, obtidas por Beijo et al. (2005), que foram respectivamente 67,6 e 20,8. Aplicando-se estes valores na expressão (13), e utilizando  $p = (0,1 ; 0,01 \text{ e } 0,005)$ , podem-se calcular as estimativas de precipitações máximas para o tempo de retorno de 10, 100 e 200 anos, que seriam 114 mm, 163 mm e 178 mm, respectivamente. Como estes são os valores máximos estimados, é coerente confiar que medianas e quantis de 90% tenham os valores apresentados na Tabela 1.

Um destaque deve ser dado à situação de se estar utilizando a distribuição de Gumbel como distribuição dos quantis, pelo fato de seus parâmetros serem mais familiares aos pesquisadores que têm experiência em estudos relacionados a valores extremos.

## 4.2 - Distribuição *a posteriori*

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

$$P(\theta|\mathbf{x}) \propto J(\theta) \prod_{j=1}^3 \frac{1}{b_j} \exp \left\{ - \left( \frac{q_j - a_j}{b_j} \right) - \exp \left[ - \left( \frac{q_j - a_j}{b_j} \right) \right] \right\} \times \\ \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\left( \frac{1+\xi}{\xi} \right)} \right\} \exp \left\{ \sum_{i=1}^n \left[ - \left[ 1 + \xi \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right] \right\}, \quad (19)$$

Calcular a integral de (19) analiticamente apresenta problemas, pois esta não tem uma forma fechada e a técnica Monte Carlo via Cadeias de Markov (MCMC) foi empregada para superar esta dificuldade. O algoritmo utilizado foi de Metropolis, apresentado por Hastings (1970), que está implementado no pacote MCMCpack do Sistema Computacional Estatístico R, sendo que é utilizado como distribuição candidata para os parâmetros uma distribuição Normal Multivariada.

Realizou-se um processo com 420000 iterações, sendo descartada as 20000 primeiras (“burn-in”) 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).

## 5 - RESULTADOS E DISCUSSÃO

### 5.1 – Análise dos Dados Simulados

A convergência da cadeia de cada parâmetro foi monitorada por meio da visualização gráfica do traço e do histograma, apresentados no Anexo B, 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 nas Tabelas 2, 3 e 4 . De acordo com Geweke (1992), existirão evidências contra a convergência, se o p-valor analisado for menor que um nível de significância fixado pelo pesquisador. Considerando um nível de significância de 5%, pode-se verificar que, para todos os parâmetros nas diversas séries analisadas, não houve nenhuma evidência contra a convergência dos mesmos.

Com base nas Tabelas 2, 3 e 4 , vê-se que o valor do fator de dependência de Raftery e Lewis (FD) atingiu valores desejados que, de acordo com Raftery & Lewis (1992), caracteriza-se uma não convergência da cadeia se o fator apresentar valor maior que “cinco”, não havendo evidências contra a convergência de cada um dos parâmetros da distribuição GEV para todas as amostras estudadas. Pode-se notar ainda que os erros Monte Carlo apresentaram valores baixos, confirmando a boa qualidade das estimativas dos parâmetros.

Um resultado importante em relação ao algoritmo de Metropolis é a taxa de aceitação, que ficou entre 0,25 e 0,42, permanecendo dentro da faixa recomendada em literatura que, segundo Gelman et al. (2003), deve ser entorno de 0,23 e 0,44, dependendo do número de parâmetros.

**TABELA 2** - Média e desvio padrão *a posteriori* dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de dependência de Raftery e Lewis (FD) e p-valor do critério de Geweke para as amostras de  $n=30$ .

Par.	Valor Real	Média Posteriori	HPD 95%		Desvio Padrão	Erro MC	FD	Geweke p-valor
			LI	LS				
$\mu$	40,0	36,864	32,892	40,775	2,042	0,0329	0,9907	0,875
$\sigma$	10,0	11,108	8,357	14,163	1,507	0,0311	1,0320	0,789
$\xi$	-0,30	-0,089	-0,453	0,034	0,157	0,0031	1,0977	0,558
$\mu$	40,0	39,437	35,117	43,757	2,206	0,0396	1,0320	0,325
$\sigma$	10,0	10,163	7,111	13,361	1,630	0,0251	0,9907	0,249
$\xi$	0,10	0,110	0,007	0,331	0,116	0,0029	1,0534	0,799
$\mu$	40,0	38,600	33,727	43,345	2,466	0,0418	1,0109	0,233
$\sigma$	10,0	10,634	6,110	15,868	2,637	0,0513	0,9704	0,857
$\xi$	0,20	0,222	-0,018	0,414	0,129	0,0032	1,0755	0,495
$\mu$	40,0	40,501	35,120	46,543	3,005	0,0616	1,0320	0,713
$\sigma$	20,0	20,902	15,329	26,694	3,020	0,0717	1,0755	0,494
$\xi$	-0,30	-0,122	-0,417	0,030	0,156	0,0042	1,1439	0,296
$\mu$	40,0	43,217	33,285	52,607	5,128	0,0981	1,0320	0,439
$\sigma$	20,0	25,527	18,153	33,594	4,042	0,0661	1,0755	0,323
$\xi$	0,10	0,123	-0,029	0,349	0,123	0,0028	1,0109	0,652
$\mu$	40,0	41,572	36,198	47,499	2,942	0,0418	1,0534	0,983
$\sigma$	20,0	24,927	18,069	33,324	4,066	0,0638	1,0320	0,816
$\xi$	0,20	0,275	0,110	0,438	0,083	0,0015	2,2419	0,294
$\mu$	70,0	66,247	61,629	70,880	2,320	0,0582	1,0526	0,098
$\sigma$	10,0	11,028	8,261	13,840	1,495	0,0329	1,0789	0,103
$\xi$	-0,30	-0,057	-0,313	0,032	0,105	0,0029	1,7368	0,600
$\mu$	70,0	72,226	67,386	76,713	2,388	0,0379	1,0320	0,159
$\sigma$	10,0	11,647	8,001	15,650	1,962	0,0308	1,0870	0,199
$\xi$	0,10	0,090	-0,035	0,334	0,114	0,0024	1,0534	0,783
$\mu$	70,0	69,725	65,498	74,115	2,193	0,0361	1,0109	0,532
$\sigma$	10,0	10,524	7,168	14,008	1,835	0,0340	0,9907	0,107
$\xi$	0,20	0,257	0,062	0,433	0,090	0,0019	2,0993	0,799
$\mu$	70,0	66,284	61,611	70,818	2,349	0,0445	1,0109	0,729
$\sigma$	20,0	11,008	8,324	13,964	1,492	0,0248	1,0109	0,161
$\xi$	-0,30	-0,057	-0,314	0,030	0,104	0,0028	1,1919	0,367
$\mu$	70,0	75,437	65,814	83,668	4,592	0,0954	1,1439	0,506
$\sigma$	20,0	22,799	15,202	30,140	3,872	0,0687	1,0320	0,503
$\xi$	0,10	0,073	-0,039	0,306	0,105	0,0035	1,0320	0,387
$\mu$	70,0	68,923	60,927	77,011	4,116	0,0820	1,0534	0,568
$\sigma$	20,0	17,448	11,190	25,622	3,824	0,0805	1,0534	0,248
$\xi$	0,20	0,237	0,007	0,423	0,116	0,0031	1,0755	0,863

Analisando-se a Tabela 2, pode-se verificar, para o caso do tamanho de amostra 30, que a média *a posteriori* apresentou valores de estimativas mais próximas dos valores reais para os parâmetros  $\mu$  e  $\sigma$ , enquanto que para o parâmetro  $\xi$  as estimativas apresentaram valores mais distantes dos valores reais, porém, vale ressaltar que nestes casos o intervalo HPD de 95% abrange todos os valores reais. Uma menor acurácia obtida nas estimativas parâmetro  $\xi$ , ocorreu para o valor negativo deste parâmetro. Uma justificativa para este fato é que o parâmetro  $\xi$  negativo indica uma tendência da distribuição GEV para o modelo Weibull, que geralmente é utilizado na análise de valores mínimos. Como a distribuição *a priori*, que foi desenvolvida no presente trabalho, foi construída baseada em quantis extremos máximos, seria razoável esperar que não apresentasse uma boa acurácia na obtenção das estimativas de  $\xi$  negativo. Isto também justifica o fato dos valores de desvio padrão serem mais elevados em relação à média para este caso.

A situação de baixa acurácia obtida para as estimativas do parâmetro  $\xi$ , principalmente para o caso negativo, também pode ser verificada analisando-se as Figuras 1B a 12B do Anexo B, nas quais se nota que os histogramas para este parâmetro apresentaram uma forma mais achatada e os verdadeiros valores não se situaram na região mais densa. Resultados semelhantes aos obtidos por Behrens et al.(2004) que, embora utilizando uma mistura de modelos para analisar valores extremos, obtiveram baixa acurácia e baixa precisão na estimação do parâmetro forma no caso de pequenas amostras.

**TABELA 3** - Média e desvio padrão *a posteriori* dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de dependência de Raftery e Lewis (FD) e p-valor do critério de Geweke para as amostras de  $n=50$ .

Par.	Valor Real	Média Posteriori	HPD 95%		Desvio Padrão	Erro MC	FD	Geweke p-valor
			LI	LS				
$\mu$	40,0	39,486	37,208	41,818	1,173	0,0165	1,0534	0,409
$\sigma$	10,0	10,931	9,400	12,527	0,818	0,0139	1,0534	0,610
$\xi$	-0,30	-0,321	-0,450	-0,196	0,068	0,0012	1,0109	0,205
$\mu$	40,0	39,959	37,262	42,947	1,504	0,0431	1,0263	0,872
$\sigma$	10,0	10,414	7,862	12,647	1,222	0,0304	1,0789	0,281
$\xi$	0,10	0,092	-0,026	0,298	0,102	0,0029	1,0526	0,176
$\mu$	40,0	40,210	36,668	43,890	1,848	0,0278	0,974	0,425
$\sigma$	10,0	11,623	8,867	14,652	1,496	0,0226	0,970	0,456
$\xi$	0,20	0,236	0,088	0,376	0,073	0,0013	1,120	0,980
$\mu$	40,0	41,368	36,768	45,892	2,319	0,0397	1,1207	0,967
$\sigma$	20,0	18,484	14,790	22,476	2,008	0,0407	1,0109	0,377
$\xi$	-0,30	-0,279	-0,453	0,008	0,129	0,0033	0,9907	0,783
$\mu$	40,0	41,474	34,599	48,618	3,686	0,0542	0,9704	0,381
$\sigma$	20,0	22,034	16,140	28,620	3,232	0,0543	0,9506	0,559
$\xi$	0,10	0,171	-0,017	0,351	0,111	0,0015	1,0320	0,898
$\mu$	40,0	40,777	36,626	44,872	2,135	0,0295	1,0534	0,625
$\sigma$	20,0	22,130	17,838	26,626	2,299	0,0366	1,0109	0,585
$\xi$	0,20	0,291	0,154	0,412	0,066	0,0012	2,1516	0,856
$\mu$	70,0	68,143	64,609	72,021	1,959	0,0404	1,1053	0,537
$\sigma$	10,0	10,871	8,565	13,255	1,237	0,0207	1,0263	0,299
$\xi$	-0,30	-0,204	-0,504	0,022	0,187	0,0059	1,0789	0,665
$\mu$	70,0	71,496	68,095	75,209	1,842	0,0269	1,0109	0,702
$\sigma$	10,0	11,193	8,547	14,221	1,491	0,0229	0,9704	0,678
$\xi$	0,10	0,155	-0,018	0,335	0,107	0,0019	1,0534	0,813
$\mu$	70,0	70,274	66,777	73,747	1,775	0,0294	0,9907	0,977
$\sigma$	10,0	10,477	7,820	13,470	1,475	0,0285	1,0109	0,216
$\xi$	0,20	0,222	0,007	0,382	0,100	0,0021	1,0755	0,803
$\mu$	70,0	66,308	60,627	72,214	3,012	0,0478	1,0977	0,116
$\sigma$	20,0	18,844	15,281	22,793	1,973	0,0334	1,0534	0,156
$\xi$	-0,30	-0,081	-0,316	0,023	0,111	0,0019	0,9907	0,825
$\mu$	70,0	67,716	60,404	74,827	3,775	0,0591	1,0320	0,821
$\sigma$	20,0	21,049	14,841	27,794	3,384	0,0583	0,9907	0,737
$\xi$	0,10	0,177	-0,013	0,348	0,113	0,0022	1,1439	0,676
$\mu$	70,0	67,697	61,732	73,253	2,894	0,0467	1,0109	0,434
$\sigma$	20,0	18,783	14,248	23,231	2,429	0,0428	0,9907	0,111
$\xi$	0,20	0,292	0,176	0,415	0,060	0,0010	1,0534	0,882

Para o caso do tamanho de amostra 50 (Tabelas 3), nota-se que houve uma melhora na obtenção das estimativas dos valores reais dos parâmetros, principalmente para parâmetro  $\xi$ , em que os valores estimados convergiram melhor para os verdadeiros valores dos parâmetros. Pode-se verificar, analisando-se as Tabelas 1 e 2, que houve uma redução nos valores do desvio padrão quando se aumentou o tamanho da amostra de 30 para 50, considerando por exemplo o caso de  $\mu = 40$ ,  $\sigma = 10$  e  $\xi = -0,30$ , tem-se respectivamente que o desvio padrão reduziu de 2,042; 1,507 e 0,157 para 1,173; 0,818 e 0,068. Isto indica uma melhora significativa na estimação dos parâmetros, com o aumento do tamanho amostral. Esta redução do valor do desvio padrão também pode ser verificada analisando-se as Figuras 1B a 24B do anexo B, nas quais nota-se que os histogramas dos parâmetros, no caso das amostras de tamanho 50, apresentaram uma forma menos “achatada” do que no caso das amostras de tamanho 30 e os verdadeiros valores se situaram na região mais densa.

Uma observação a ser feita nas Tabelas 2, 3 e 4, refere-se à redução do valor do desvio padrão e do erro Monte Carlo, à medida que se aumenta o tamanho da amostra. Embora utilizando uma modelagem diferente para estudar valores extremos, Behrens et al.(2004) obtiveram resultados semelhantes, em que, conseguiram melhores estimativas dos parâmetros reais das séries simuladas quando houve o aumento do tamanho da amostra.

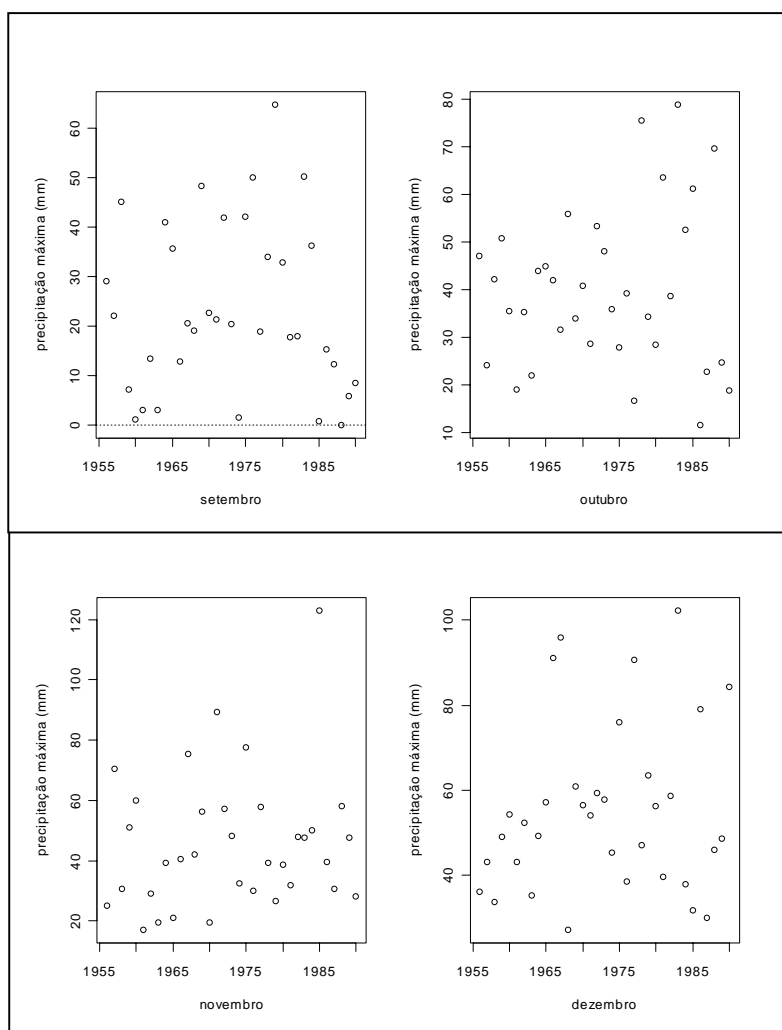


**TABELA 4** - Média e desvio padrão *a posteriori* dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de dependência de Raftery e Lewis (FD) e p-valor do critério de Geweke para as amostras de  $n=100$ .

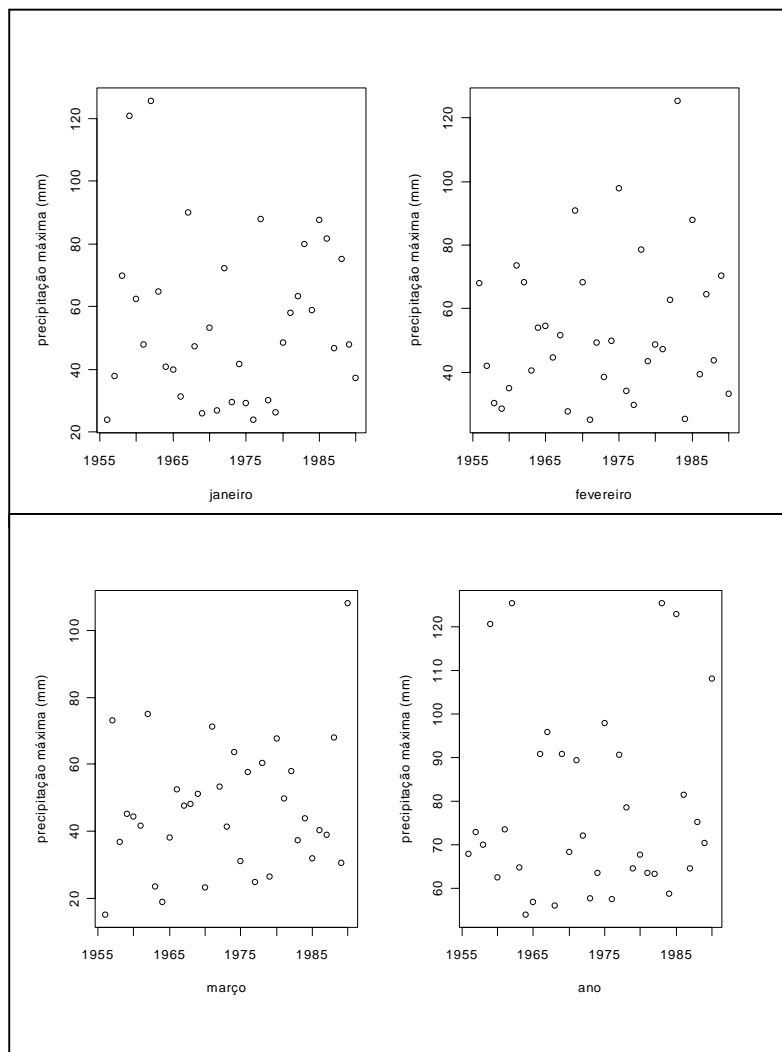
Par.	Valor Real	Média Posteriori	HPD 95%		Desvio Padrão	Erro MC	FD	Geweke p-valor
			LI	LS				
$\mu$	40,0	38,912	36,530	41,253	1,199	0,0176	1,0320	0,791
$\sigma$	10,0	11,176	9,530	12,798	0,838	0,0130	0,9704	0,911
$\xi$	-0,30	-0,324	-0,459	-0,187	0,071	0,0012	0,9909	0,342
$\mu$	40,0	40,263	38,088	42,383	1,126	0,0159	1,0109	0,375
$\sigma$	10,0	9,454	7,873	11,071	0,806	0,0129	0,9506	0,285
$\xi$	0,10	0,118	-0,020	0,309	0,104	0,0021	0,9704	0,831
$\mu$	40,0	40,054	37,819	42,318	1,146	0,0271	1,0533	0,949
$\sigma$	10,0	9,946	8,219	11,867	0,952	0,0231	1,0109	0,359
$\xi$	0,20	0,224	0,087	0,359	0,069	0,0030	2,0448	0,750
$\mu$	40,0	40,895	37,345	44,430	1,811	0,0312	1,0110	0,977
$\sigma$	20,0	22,255	19,413	25,406	1,532	0,0218	0,9906	0,862
$\xi$	-0,30	-0,345	-0,471	-0,220	0,065	0,0011	1,0321	0,375
$\mu$	40,0	40,794	36,220	45,668	2,373	0,0778	1,0789	0,325
$\sigma$	20,0	20,115	16,329	24,459	2,115	0,0798	1,0000	0,273
$\xi$	0,10	0,231	0,009	0,375	0,090	0,0047	2,6842	0,292
$\mu$	40,0	41,094	38,174	44,107	1,542	0,0226	1,0109	0,400
$\sigma$	20,0	18,545	16,046	21,527	1,424	0,0215	0,9706	0,865
$\xi$	0,20	0,267	0,162	0,386	0,057	0,0011	2,3097	0,580
$\mu$	70,0	69,009	66,930	71,334	1,127	0,0158	0,9907	0,550
$\sigma$	10,0	10,155	8,606	11,649	0,781	0,0132	1,0109	0,922
$\xi$	-0,30	-0,349	-0,474	-0,204	0,071	0,0012	1,0534	0,506
$\mu$	70,0	69,485	67,089	72,056	1,278	0,0236	1,0320	0,888
$\sigma$	10,0	11,043	9,395	12,767	0,882	0,0162	0,9704	0,212
$\xi$	0,10	0,162	0,007	0,287	0,077	0,0021	2,4111	0,626
$\mu$	70,0	70,467	68,395	72,902	1,167	0,0190	0,9907	0,941
$\sigma$	10,0	10,222	8,395	12,068	0,949	0,0157	0,9907	0,159
$\xi$	0,20	0,251	0,115	0,385	0,069	0,0011	2,0758	0,474
$\mu$	70,0	67,010	62,085	71,538	2,414	0,0426	1,0109	0,512
$\sigma$	20,0	21,974	18,890	25,439	1,723	0,0268	0,9704	0,922
$\xi$	-0,30	-0,308	-0,445	-0,175	0,072	0,0012	1,1439	0,903
$\mu$	70,0	68,831	63,591	73,879	2,620	0,0386	1,0109	0,512
$\sigma$	20,0	22,238	18,583	26,200	2,020	0,0335	0,9704	0,922
$\xi$	0,10	0,154	0,007	0,296	0,080	0,0018	1,1439	0,903
$\mu$	70,0	70,632	66,719	74,703	2,089	0,0297	0,9907	0,629
$\sigma$	20,0	18,142	15,120	21,576	1,663	0,0246	0,9818	0,748
$\xi$	0,20	0,284	0,179	0,392	0,055	0,0013	2,2077	0,273

## 5.2 – Estudo das Precipitações Máximas de Jaboticabal

Pode-se verificar nas Figuras 2 e 3, que poucas das precipitações máximas que ocorreram entre 1956 e 1990 tiveram valores superiores a 100 mm.



**FIGURA 2** - Representação gráfica da série de precipitações diárias máximas (em mm) no município de Jaboticabal (SP) dos meses de setembro, outubro, novembro e dezembro entre os anos de 1956 e 1990.



**FIGURA 3** - Representação gráfica da série de precipitações diárias máximas (em mm) no município de Jaboticabal (SP) dos meses de janeiro, fevereiro, março e 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 5 que não existem evidências estatísticas para rejeitar a hipótese de aleatoriedade.

**TABELA 5** - 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

A convergência da cadeia de cada parâmetro foi monitorada por meio da visualização gráfica do traço e do histograma, apresentados no Anexo A 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 6. 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.

Com base na Tabela 6, vê-se que o valor do fator de convergência de Raftery e Lewis (FD) atingiu valores desejados ( $FD < 5$ ), não havendo, no entanto, evidências contra a convergência de cada um dos parâmetros da distribuição GEV, para todos os períodos estudados. Pode-se notar ainda que os erros Monte Carlo apresentaram valores baixos, confirmando a boa qualidade das estimativas dos parâmetros.

**TABELA 6** - Média e desvio padrão *a posteriori* dos parâmetros da distribuição GEV, com seus respectivos intervalos HPD 95%, erro Monte Carlo, fator de convergência de Raftery e Lewis (FD) e p-valor do Critério de Geweke .

Período	Par.	Média Posteriori	HPD 95%		Desvio Padrão	Erro MC	FD	Geweke p-valor
			LI	LS				
Setembro	$\mu$	14,242	9,904	19,389	2,545	0,0436	0,9907	0,534
	$\sigma$	14,152	10,298	18,602	2,193	0,0364	1,0320	0,438
	$\xi$	0,218	-0,016	0,390	0,115	0,0025	1,1439	0,982
Outubro	$\mu$	32,644	28,556	37,372	2,303	0,0358	0,9907	0,582
	$\sigma$	15,947	12,538	20,218	2,005	0,0332	1,0320	0,841
	$\xi$	0,232	0,009	0,379	0,097	0,0015	1,0109	0,833
Novembro	$\mu$	40,777	36,870	44,616	2,098	0,0648	1,0263	0,255
	$\sigma$	19,574	15,467	24,232	2,341	0,0693	1,1368	0,198
	$\xi$	0,207	0,009	0,401	0,112	0,0041	1,0789	0,069
Dezembro	$\mu$	45,753	41,045	50,257	2,377	0,0359	1,0320	0,385
	$\sigma$	15,439	11,286	19,922	2,261	0,0373	0,9704	0,857
	$\xi$	0,096	-0,031	0,322	0,109	0,0023	0,9907	0,778
Janeiro	$\mu$	43,931	39,391	48,602	2,341	0,0381	1,0109	0,927
	$\sigma$	20,287	14,990	26,502	3,023	0,0505	1,0109	0,712
	$\xi$	0,078	-0,035	0,308	0,102	0,0020	1,0320	0,576
Fevereiro	$\mu$	40,750	37,370	44,636	1,897	0,0282	0,9907	0,613
	$\sigma$	16,014	12,042	20,414	2,200	0,0326	0,9616	0,373
	$\xi$	0,387	0,283	0,488	0,053	0,0008	1,0109	0,873
Março	$\mu$	42,614	38,213	46,483	2,184	0,0576	1,0547	0,695
	$\sigma$	19,432	15,324	24,051	2,294	0,0344	1,0109	0,141
	$\xi$	0,108	-0,033	0,320	0,110	0,0042	1,1207	0,914
Ano	$\mu$	64,379	60,305	68,682	2,115	0,0310	0,9704	0,556
	$\sigma$	10,810	7,230	14,840	2,018	0,0332	1,0534	0,540
	$\xi$	0,417	0,301	0,529	0,058	0,0009	0,9907	0,662

Analisando-se a Tabela 6, pode-se verificar que 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. Uma possível causa deste acontecimento é o fato de se ter utilizado uma amostra de tamanho 36, e, como verificado na seção anterior, estimativas dos parâmetros com menor desvio padrão foram obtidas quando se utilizou a amostra de tamanho 50 ou superior. Embora utilizando 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=0,03 e desvio padrão=0,07). Pode-se observar ainda, na Tabela 6, que os erros Monte Carlo apresentaram valores bem reduzidos em relação à média, o que indica que a influência da técnica na obtenção das estimativas dos parâmetros foi insignificante.

Analisando as Tabelas 7 e 8, pode-se verificar que a estimativa pontual de precipitação máxima ( $\hat{y}_p$ ) e seu respectivo limite superior de 95%, para um tempo de retorno de 10 anos, foram menores que a maior precipitação ocorrida entre os anos de 1991 a 2001, tanto para o mês de fevereiro como para o período anual. Pode-se concluir que houve uma subestimação da quantidade de precipitação máxima provável. Porém, pode-se verificar, ainda nas Tabelas 7 e 8, 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 7.

**TABELA 7** - 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

Como em análises de precipitações máximas sempre se está interessado no evento extremo máximo, sugere-se neste caso utilizar o limite superior como valor de referência. Estes 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.

**TABELA 8** - Estimativas das precipitações pluviométricas 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 HPD de 95 % e limite superior (LS) de 95%.

Período	Tempo de Retorno					
	10 anos			20 anos		
	$\hat{y}_p$	HPD	LS (95%)	$\hat{y}_p$	HPD	LS (95%)
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

Na Tabela 8, escolhendo 20 anos como o tempo de retorno adequado de ocorrer uma determinada precipitação pluvial máxima diária para o período anual, uma interpretação das informações pode ser feita do seguinte modo:

espera-se 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 HPD de 95% (107,7 mm; 148,7 mm) e limite superior de 95% igual a 146,9 mm, respectivamente.

Utilizando a metodologia proposta por Coles & Tawn (1996), Beijo et al. (2005b) também objetivaram obter as estimativas de precipitação máxima para Jaboticabal. Os resultados estão apresentados na Tabela 1E (vide Anexo E), e, comparando-se os resultados desta Tabela com as da Tabela 8, pode-se verificar que os valores das estimativas de precipitação máxima obtidos utilizando-se a distribuição *a priori* proposta no presente trabalho, apresentaram maior acurácia do que os valores fornecidos pela priori proposta por Coles & Tawn (1996). Destaca-se a estimativa do mês de fevereiro, em que a precipitação pluvial de 137,2 mm só seria prevista tendo como base o limite superior do intervalo HPD de 99% para um tempo de retorno de 50 anos (vide Tabela 1E do Anexo E). Pela Tabela 8, pode-se verificar que esta precipitação seria prevista, tendo como base o limite superior de 95% para um tempo de retorno de 20 anos, ou seja, houve uma redução no tempo de previsão de 30anos, mostrando que o resultado obtido via distribuição *a priori*, proposta no presente trabalho, foi mais acurado que o fornecido pela modelagem de Coles & Tawn (1996).

Do ponto de vista prático, esta maior acurácia nos resultados é muito importante em termos de segurança e economia. Considerando, por exemplo, que um engenheiro desejasse projetar uma obra hidráulica (um dique, uma curva de nível, entre outras) que tem um tempo de vida (permanência) de 20 anos, ele adotaria em seus cálculos a estimativa de 113,3 mm (vide Tabela 1E do Anexo E) de precipitação máxima para este tempo. Porém, em um período de tempo menor, ocorreu uma precipitação pluvial de 137,2 mm (vide Tabela 7), o que poderia acarretar sérios danos. No caso da construção de uma acurva de



nível, esta 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, esta subestimação da precipitação máxima poderia causar rompimentos destas obras, o que seria um alto risco à segurança.

## 6 – CONCLUSÕES

Diante dos resultados obtidos pode-se concluir que:

- a) A distribuição *a priori* desenvolvida, baseando-se em quantis extremos com distribuição Gumbel, apresentou resultados satisfatórios na estimação dos parâmetros da distribuição GEV, para os diferentes tamanhos de amostras e diferentes valores paramétricos.
- b) A distribuição *a priori*, que foi proposta no presente trabalho, mostrou-se adequada para a incorporação de conhecimentos de especialistas no estudo de dados extremos, para a obtenção das estimativas de precipitações máximas, para certos níveis de retorno na região Jaboticabal (SP), e apresentou uma maior acurácia do que a distribuição *a priori* proposta por Coles & Tawn (1996).
- c) 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 (SP).
- d) No caso da análise precipitações máximas da região de Jaboticabal (SP), em que  $n=36$ , espera-se que a acurácia para as estimativas de precipitações máximas para níveis de retorno menores melhore com o aumento do tamanho da amostra.
- e) As séries geradas, aplicando-se a distribuição *a priori* proposta, não apresentaram nenhum problema quanto à convergência.

Em estudos futuros, pretende-se avaliar a aplicabilidade da distribuição *a priori* desenvolvida no presente trabalho junto ao modelo de valores extremos generalizado, com utilização de threshold. Outro fato a ser verificado nestes estudos seria analisar a utilização de outros métodos numéricos, para os cálculos em relação à distribuição *a posteriori*, entre eles, aproximação de Laplace, aproximação via quadratura Gaussiana e amostragem por importância. Dando mais um passo à frente, pode-se incorporar, no modelo bayesiano aqui proposto, uma componente espacial e sazonal para estudar principalmente séries históricas de dados extremos.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

BANZATTO, D. A. **Estimativa das precipitações máximas prováveis no Estado de São Paulo**. 1973. 119 p. Tese (Doutorado em Ciências) - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal.

BASU, A. P. Estimates of reliability for some distributions useful in life testing. **Technometrics**, Alexandria, v. 6, n. 2, p. 215-219, 1964.

BAUTISTA, E. A. L. **A distribuição generalizada de valores extremos no estudo da velocidade máxima do vento em Piracicaba, SP**. 2002. 47 p. Dissertação (Mestrado em Agronomia) – Universidade de São Paulo, Piracicaba, SP.

BEHRENS, C. N.; LOPES, H. F.; GAMERMAN, D. Bayesian Analysis of extreme events with threshold estimation. **Statistical Modeling**, Londres, v. 4, n. 3, p. 227-244, Oct. 2004.

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. 2005a.

BEIJO, L. A. ; MUNIZ, J. A.; VIVANCO, M. J. F. Intervalo de máxima densidade a posteriori (HPD) do tempo de retorno no estudo de precipitações máximas via distribuição GEV. In: 11º SEAGRO - Simpósio de Estatística Aplicada à Experimentação Agrônômica & 50ª Reunião Anual da RBRAS - Região Brasileira da Sociedade Internacional de Biometria, 2003, Lavras. **Programa e Resumos do 11º SEAGRO & 50ª Reunião Anual da RBRAS**, Londrina:UEL, 2005b. CD-Room.

BERGER, J. O.; SUN, D. Bayesian analysis for the poly-Weibull distribution. **Journal of the American Statistical Association**, Alexandria, v. 88, n. 424, p. 1412-1418, Dec. 1993.

BHATTACHARYA, S. K. Bayesian approach to life testing and reliability estimation. **Journal of the American Statistical Association**, Alexandria, v. 62, n. 317, p. 48-62, Mar. 1967.

BOX, G. E. P.; TIAO, G. **Bayesian inference in statistical analysis**. New York: Wiley, 1992

CHECHILE, R. A. Bayesian Analysis of Gumbel Distributed Data. **Communications in Statistics: theory and methods**, New York, v. 30, n. 3, p. 485-496, 2001.

CHOWDHURY, J. U.; STENDINGER, J. R.; LU, L. H. Goodness of - Fit Test for Regional Generalized Extreme Value Flood Distributions. **Water Resources Research**, Washington, v. 27, n. 7, p. 1765-1776, July 1991.

COLES, S. G.; DIXON, J. Likelihood-based inference for extreme value models. **Extremes**, Tyresoe, v. 2, n. 1, p. 5-23, 1999.

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.; PERICCHI, L. R. Anticipating catastrophes through extreme value modelling. **Applied Statistics**, London, v. 52, n. 3, p. 405-416, 2003.

COLES, S. G.; PERICCHI, L. R.; SISSON S. A full probabilistic approach to extreme rainfall modelling. **Journal of Hidrology**, Amsterdam, v. 273, n. 1/4, p. 35-50, Mar. 2003.

COLES, S. G ; TAWN, J. A. Modelling extreme multivariate events. **Journal of the Royal Statistical Society**, London, v. 53, n. 2, p. 377-392, 1991.

COLES, S. G.; TAWN, J. A. A Bayesian analysis of extreme rainfall data. **Applied Statistics**, Oxford, v. 45, n. 4, p. 463-478, 1996.

COWLES, M. K.; CARLIN, B. P. Markov chain Monte Carlo convergence diagnostics: a comparative review. **Journal of the American Statistical Association**, Alexandria, v. 91, n. 434, p. 883-904, June 1996.

CROWDER, M. Bayesian priors based on a parameter transformation using the distribution function. **Annals of the Institute of Statistical Mathematics**, Dordrecht, v. 44, n. 3, p. 405-416, Sept. 1992.

DEY, D. K.; LEE, T. M. Bayes computation for life testing and reliability. **IEEE Transactions in Reliability**, New York, v. 41, n. 4, p. 621-626, Dec. 1992.

FERRO, C. A. T.; SEGERS, J. Inference for clusters of extreme values. **Journal of the Royal Statistical Society B**, London, v. 65, n. 2, p. 545-556, 2003.

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

FRAGA ALVES, M. I. **Inferência Estatística em Modelos Extremais**. 1992. Tese (Doutoramento em Estatística) - Faculdade de Ciências de Lisboa.

GAMERMAN, D. **Simulação Estocástica Via Cadeias de Markov**. São Paulo: Associação Brasileira de Estatística, 1996. 196 p.

GELFAND, A. E.; HILLS, S. E.; RACINE-POON, A.; SMITH, A. E. M. Illustration of Bayesian inference in normal data models using Gibbs sampling. **Journal of the American Statistical Association**, Alexandria, v. 85, n. 410, p. 972-985, June 1990.

GELMAN, A.; CARLIN, J.B.; STERN, H.S.; RUBIN, D.B. **Bayesian data analysis**. London: Chapman Hall, 2003. 668 p.

GELMAN, A.; RUBIN, D. B. A single series from the Gibbs sampler provides a false sense of security. In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Ed.). **Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting**,. Oxford: Clarendon Press, 1992. p. 625-631.

GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Washington, v. 6, n. 6, p. 721-741, 1984

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. (Ed.). **Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting**, Oxford: Clarendon Press, 1992. p. 625-631.

CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hastings algorithm. **The American Statistician**, Alexandria, v. 49, n. 4, p. 327-335, Nov. 1995.

GNEDENKO, B. V. Sur la distribution limite du terme maximum d'une série aléatoire. **Annales des Mathématiques**, Clermont-Ferrand, v. 44, p. 423-453, 1943.

GUMBEL, E. J. **Statistics of Extremes**. New York: Columbia University Press, 1958. 375 p.

HAAN, C. T. **Statistical methods in Hidrology**. Ames, Iowa: The Iowa State University Prees, 1977. 377 p.

HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, London, v. 57, n. 1, p. 97-109, Apr. 1970.

HEIDELBERGER, P.; WELCH, P. Simulation run lengh control in the presence of an initial transient. **Operations Research**, Landing, v. 31, n. 6, p. 1109-1144, 1983.

HILL, B. M. A simple general approach to inference about the tail of a distribution. **Annals of Statistics**, Hayward, v. 3, n. 5, p. 1163-1174, 1975

HOLLA, M. S. Bayesian estimates of reliability function. **Australian Journal Statistics**, Sydney, v. 8, n. 1, p. 32-35, 1966.

HOSKING, J. R. M.; WALLIS, J. R.; WOOD, E. F. An appraisal of the regional flood frequency procedure in the UK flood studies report. **Hydrological Science Journal**, Wallingford, v. 30, n. 1, p. 85-109, 1985a.

HOSKING, J. R. M.; WALLIS, J. R.; WOOD, E. F. Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. **Technometrics**, Alexandria, v. 27, n. 3, p. 251-261, Aug. 1985b.

JEFFREYS, H. **The times of P, S, and SKS, and the velocities Of P and S**; Monthly Notices of the Royal Astronomical Society. Monthly Notices R. A. S. Geophys, p. 498-533, 1939.

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.

KEARNS, P.; PAGAN, A. R. Estimating the Density Tail Index for Financial Time Series. **Review of Economics and Statistics**, Cambridge, v. 79, n. 2, p. 171-175, May 1997.

- LARSON, R. E.; HOSTETLER, R. P.; EDWARDS, B. H. **Cálculo com aplicações**. Tradução Alfredo Alves de Faria ; revisão técnica Eliana Farias e Soares . 4. ed. Rio de Janeiro: LTC, 1998. 711 p. Título original: Brief calculus with applications.
- LEADBETTER, M. R.; LINDGREN, G.; ROOTZEN, H. **Extremes and related properties of random sequences and processes**. New York: Springer-Verlag, 1983. 336 p.
- LEHMANN, E. L. **Nonparametric Statistical Methods Based on Ranks**. New York : McGraw-Hill, 1975. 457 p.
- 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.
- MCNEIL, A.; FREY, R. Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach. **Journal of Empirical Finance**, Amsterdam, v. 7, p. 271-300, 2000.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equations of state calculations by fast computing machines. **Journal of Chemical Physics**, New York, v. 21, p. 1087-1092, 1953.
- MOOD, A. M.; GAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**. 3. ed. New York: J. Wiley & Sons, 1974. 564 p
- NEVES, M.; MARTINS, M. J. Estimador simples do parâmetro de forma da distribuição generalizada de valores extremos. In: CONGRESSO ANUAL DA SOCIEDADE PORTUGUESA DE ESTATÍSTICA FUNCHAL, 4., 1996.
- NOGUEIRA, D. A. **Proposta e avaliação de critérios de convergência para o método de Monte Carlo via cadeias de Markov: casos uni e multivariados**. 2004. 121 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Curso de Pós-graduação em Agronomia/Experimentação Agropecuária, Universidade Federal de Lavras, Lavras, MG.
- PAULINO, D. C.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: FCG, 2003. 446 p.



PICKANDS, J. Statistical Inference using extreme order statistics. **Annals of Statistics**, Hayward, v. 3, n. 1, p. 119-131, 1975.

PINTO, F. R. L. **Equações de intensidades-duração-freqüência da precipitação para os estados do Rio de Janeiro e Espírito Santo: estimativa e espacialização**. 1999. 70 p. Dissertação (Mestrado em Engenharia Agrícola) – Curso de Pós-graduação em Engenharia Agrícola, Universidade Federal de Viçosa, Viçosa, MG.

RAFTERY, A L.; LEWIS, S. One long run with diagnostics: implementation strategies for Markov chain Monte Carlo [Comment]. **Statistical Science**, Hayward, v. 7, n. 4, p. 493-497, Nov. 1992.

R DEVELOPMENT CORE TEAM. **The R Manuals :Current Version: 1. 9. 1**. 2004. Disponível em: <<http://www.r-project.org>>. Acesso em: 2006.

RAYNAL, J. A. Sobre el uso del dominio de atracción para la identificación de valores extremos para máximos. **Ingeniería Hidráulica en México**, Mexico, v. 12, n. 1, p. 57-62, 1997.

ROSA, G. J. M. **Análise Bayesiana de modelos lineares mistos robustos via Amostrador de Gibbs**. 1998. 57 p. Tese (Doutorado em Estatística e Experimentação Agronômica) – Universidade de São Paulo, Piracicaba, SP.

ROSS, W. H. A peaks over threshold analysis of extreme wind speeds. **Canadian Journal of Statistic**, Ontario, v. 15, n. 4, p. 328-335, Dec. 1987.

SHARMA, P.; KHARE, M.; CHAKARVARTI, S. P. Application of Extreme Value Theory for Predicting Violations of Air Quality Standards for an Urban Road Intersection. **Transportation Research Part D: Transport and Environment**, Oxford, v. 4, n. 3, p. 201-216, May 1999.

SINGPURWALLA, N. D.; SONG, M. S. Reliability analysis using Weibull lifetime data and expert opinion. **IEEE Transactions in Reliability**, New York, v. 37, n. 3, p. 340-347, Aug. 1988.

SINHA, S. K.; SLOAN, J. B. Bayes estimation of the parameters and reliability function of the 3-parameter Weibull distribution. **IEEE Transactions in Reliability**, New York, v. 37, n. 4, p. 364-368, Oct. 1988.

SMITH, R. L. Maximum likelihood estimation in a class of nonregular cases. **Biometrika**, London, v. 72, n. 1, p. 67-90, 1985.

SMITH, R. L.; NAYLOR J. C. A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. **Applied Statistics**, v. 36, n. 3, p. 358-369, 1987.

SMITH, R. L.; TAWN, J. Á.; COLES, S. G. Markov chain models for threshold exceedances. **Biometrika**, London, v. 84, n. 2, p. 249-268, June 1997.

SORENSEN, D. **Gibbs Sampling in quantitative genetic**. Copenhagen: Foulun, 1996. 186 p.

SOUZA, L. A. E. **Valor em risco em Épocas de crise**. 1999. 121 p. Dissertação (mestrado em Economia) - Universidade de São Paulo. Faculdade de Economia Administração e Contabilidade, São Paulo, SP.

SOUZA, L. A. E. de; SILVA, M. E. da. **Teoria de valores extremos para cálculo de VaR**. São Paulo, julho de 1999. Disponível em: <[www.risktech.com.br/PDFs/var-evt.pdf](http://www.risktech.com.br/PDFs/var-evt.pdf)> Acesso em : 29 ago. 2003.

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: 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.

THEMIDO PEREIRA, M. T. **Contribuições à teoria de valores extremos**. 1994. Tese (Doutorado em Estatística) - Universidade de Lisboa - Faculdade de Ciências, Lisboa.

TUCCI, C. M. **Hidrologia: ciência e aplicação**. Porto Alegre: UFRGS, 2001. 943 p.

VARGAS, M. X.; HARAMBOUR, F. P.; CHIBBARO, R. S. Uso de intervalos de confianza em la detrmínación de caudales de diseño em zonas sin información registrada. **Anales del Instituto de Ingenieros de Chile**, v. 11, n. 2, Ago. 1999.

VASCONCELLOS, S. L. B. **Análise da quantidade e distribuição pluviométrica na região de Jaboticabal - SP.** 1998. 63 p. Dissertação (Mestrado em Agronomia) – Curso de Pós-graduação em Agronomia, Escola Superior de Agricultura Luiz de Queirós, Universidade de São Paulo.

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.

VIVANCO, M. J. F. **Análise de valores extremos no tratamento estatístico da corrosão de equipamentos.** Campinas, 1994. 107 p. Dissertação (Mestrado em Estatística) – Curso de Pós-graduação em Estatística, Universidade Estadual de Campinas, 1994.

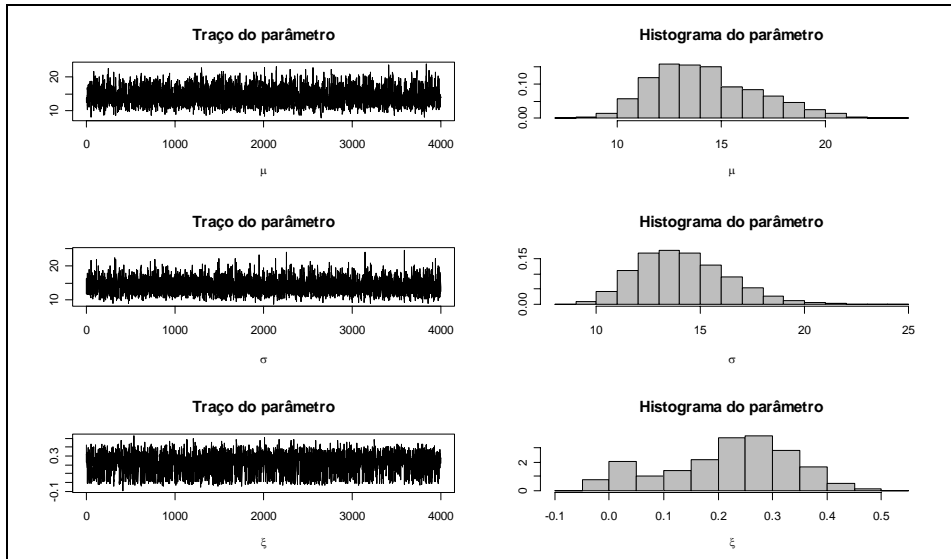
WALSHAW, D. Modelling extreme wind speeds in regions prone to hurricanes. **Journal of the Royal Statistical Society Series C - Applied Statistics**, London, v. 49, n. 1, p. 51-62, 2000.

WHALEN, T. M.; SAVAGE, G. T.; JEONG, G. D. An evaluation of the self-determined probability-weighted moment method for estimating extreme wind speeds. **Journal of Wind Engineering and Industrial Aerodynamics**, Amsterdam, v. 92, n. 3/4, p. 219-239, Mar. 2004.

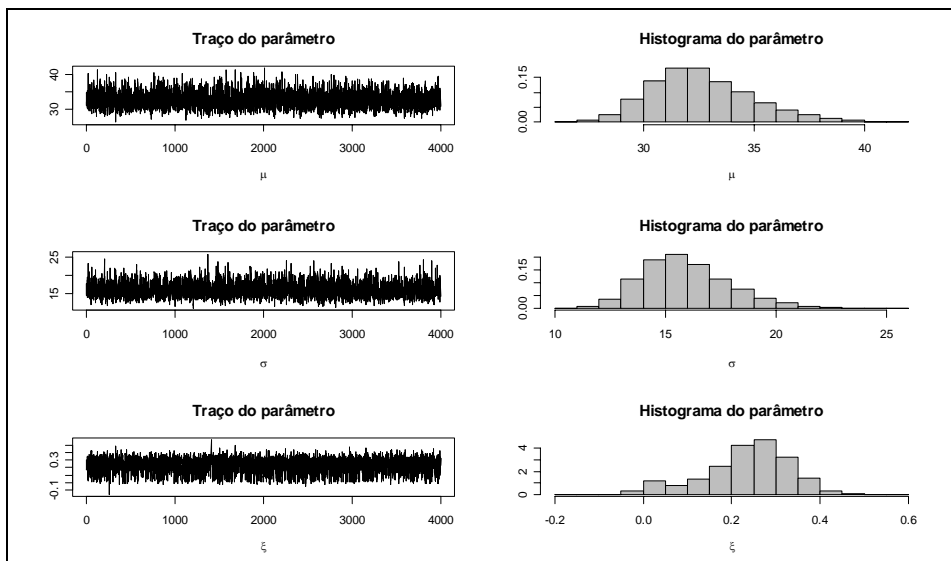
## 8 - ANEXOS

### ANEXO A

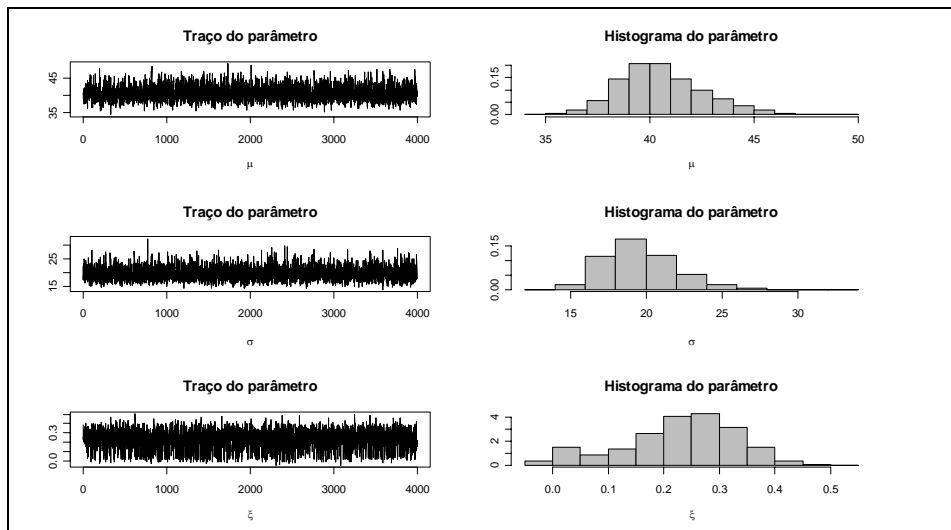
FIGURA		Página
1 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para setembro.....	65
2 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para outubro.....	65
3 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para novembro.....	66
4 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para dezembro.....	66
5 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para janeiro.....	67
6 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para fevereiro.....	67
7 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para março.....	68
8 A	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para o período anual.....	68



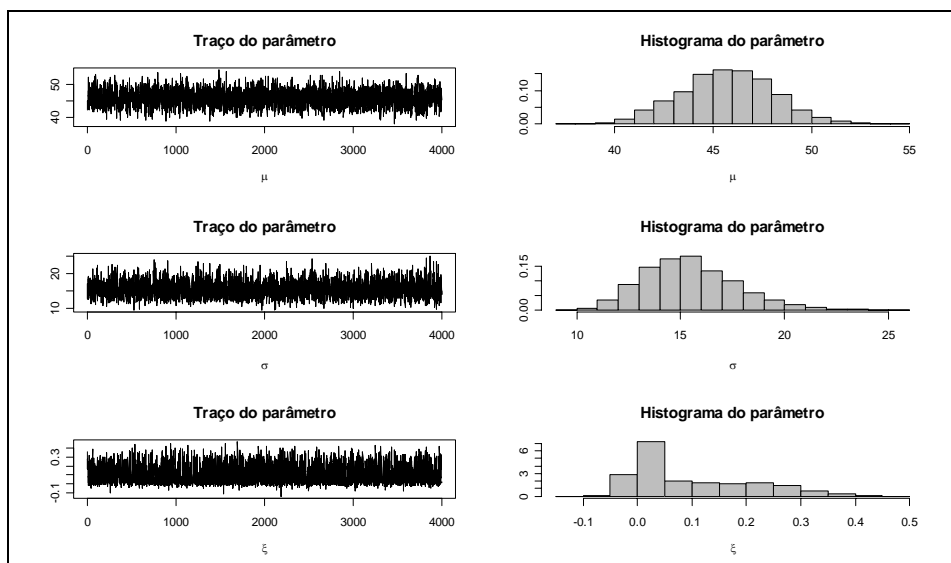
**FIGURA 1 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para setembro.



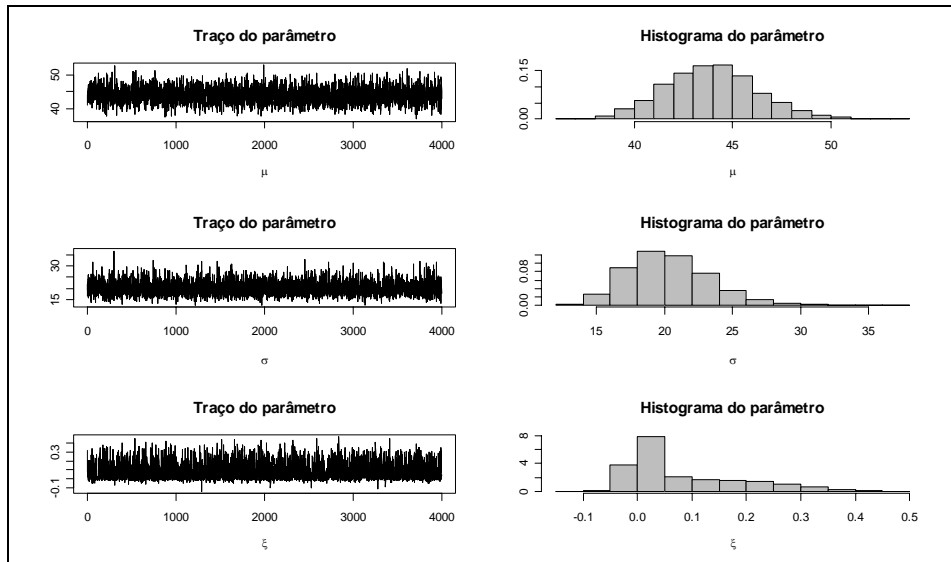
**FIGURA 2 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para outubro.



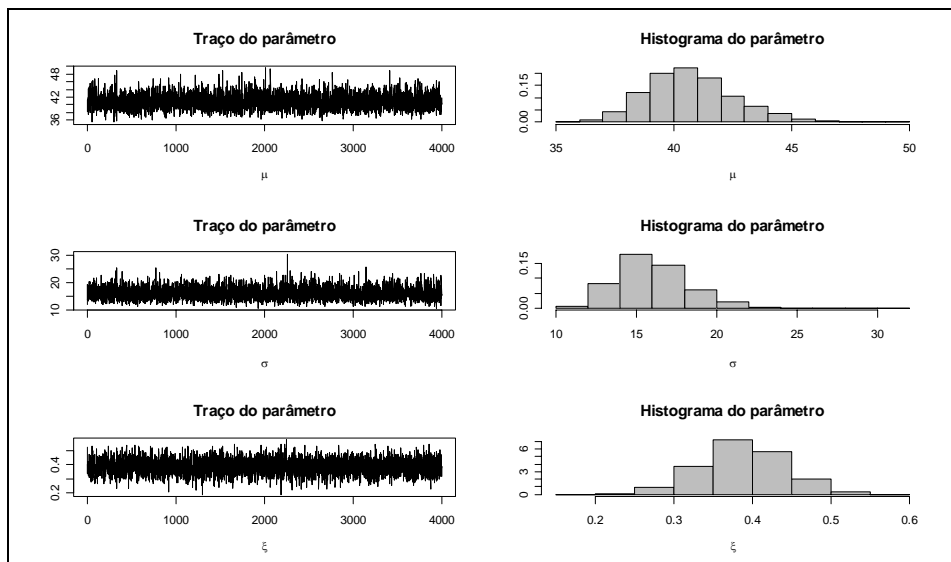
**FIGURA 3 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para novembro.



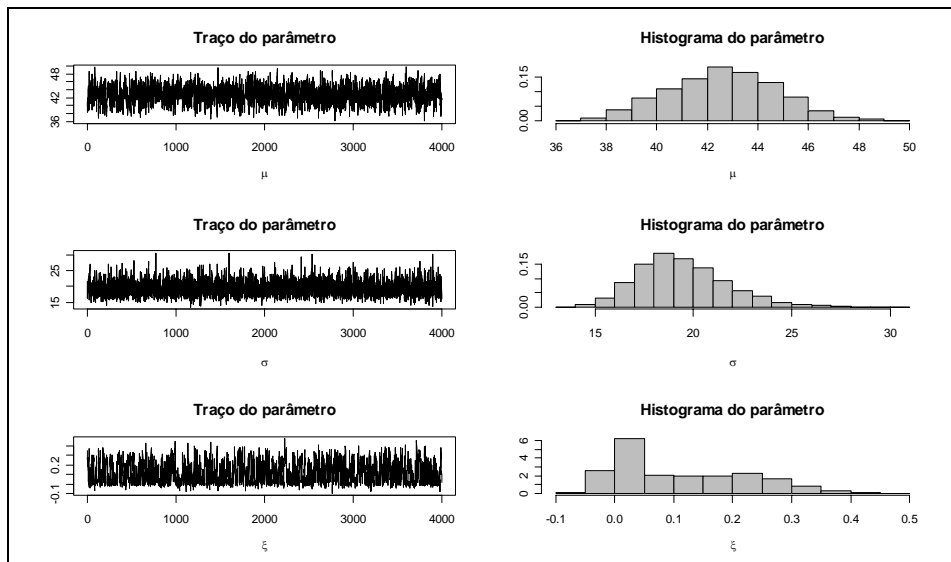
**FIGURA 4 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para dezembro.



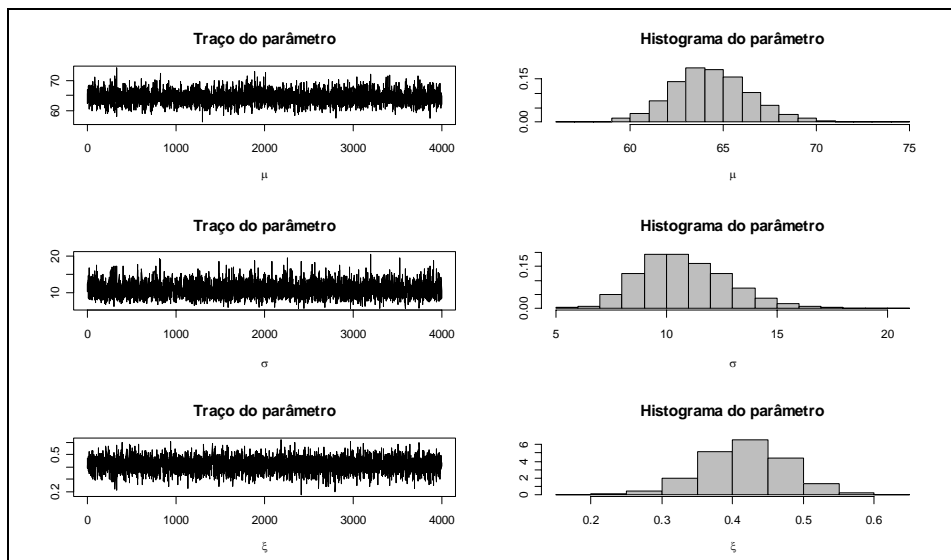
**FIGURA 5 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para janeiro.



**FIGURA 6 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para fevereiro.



**FIGURA 7 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para março.



**FIGURA 8 A:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para o período anual.



## ANEXOS

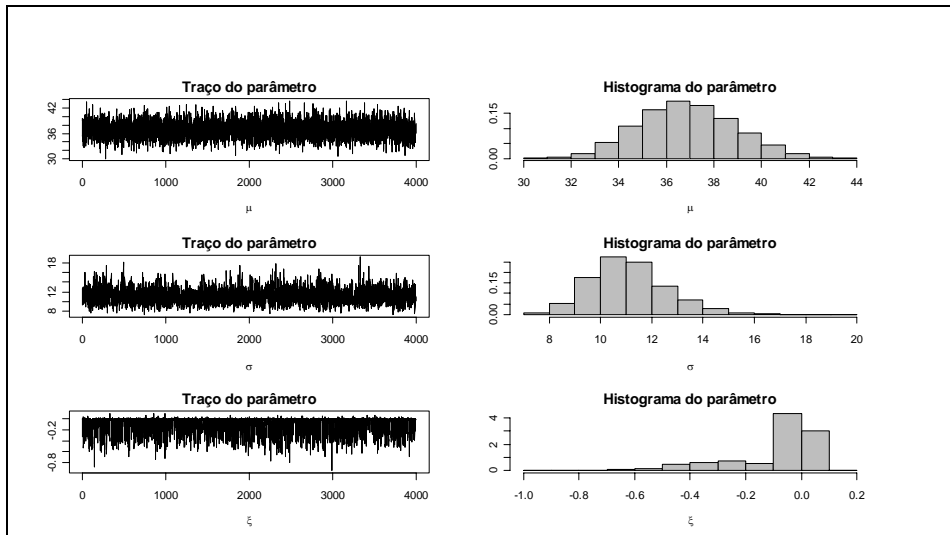
### ANEXO B

FIGURA		Página
1 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 40, \sigma = 10$ e $\xi = -0,30$ .....	73
2 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 40, \sigma = 10$ e $\xi = 0,10$ .....	73
3 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 40, \sigma = 10$ e $\xi = 0,20$ .....	74
4 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 40, \sigma = 20$ e $\xi = -0,30$ .....	74
5 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 40, \sigma = 20$ e $\xi = 0,10$ .....	75
6 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 40, \sigma = 20$ e $\xi = 0,20$ .....	75
7 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 70, \sigma = 10$ e $\xi = -0,30$ .....	76
8 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 70, \sigma = 10$ e $\xi = 0,10$ .....	76

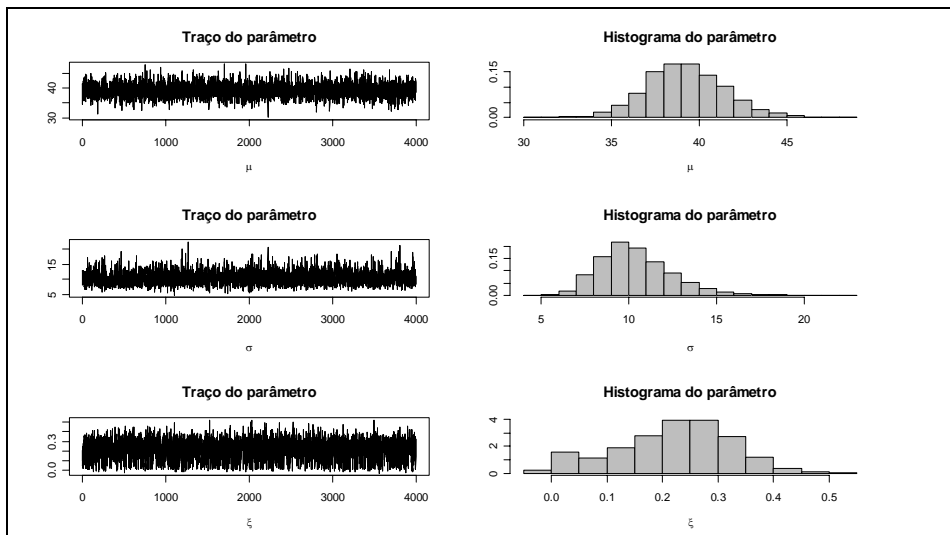
<b>FIGURA</b>		<b>Página</b>
9 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 70, \sigma = 10$ e $\xi = 0,20$ .....	77
10 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 70, \sigma = 20$ e $\xi = -0,30$ .....	77
11 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 70, \sigma = 20$ e $\xi = 0,10$ .....	78
12 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 30, \mu = 70, \sigma = 20$ e $\xi = 0,20$ .....	78
13 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 40, \sigma = 10$ e $\xi = -0,30$ .....	79
14 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 40, \sigma = 10$ e $\xi = 0,10$ .....	79
15 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 40, \sigma = 10$ e $\xi = 0,20$ .....	80
16 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 40, \sigma = 20$ e $\xi = -0,30$ .....	80
17 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 40, \sigma = 20$ e $\xi = 0,10$ .....	81

<b>FIGURA</b>		<b>Página</b>
18 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 40, \sigma = 20$ e $\xi = 0,20$ .....	81
19 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 70, \sigma = 10$ e $\xi = -0,30$ .....	82
20 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 70, \sigma = 10$ e $\xi = 0,10$ .....	82
21 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 70, \sigma = 10$ e $\xi = 0,20$ .....	83
22 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 70, \sigma = 20$ e $\xi = -0,30$ .....	83
23 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 70, \sigma = 20$ e $\xi = 0,10$ .....	84
24 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 50, \mu = 70, \sigma = 20$ e $\xi = 0,20$ .....	84
25 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 40, \sigma = 10$ e $\xi = -0,30$ .....	85
26 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 40, \sigma = 10$ e $\xi = 0,10$ .....	85

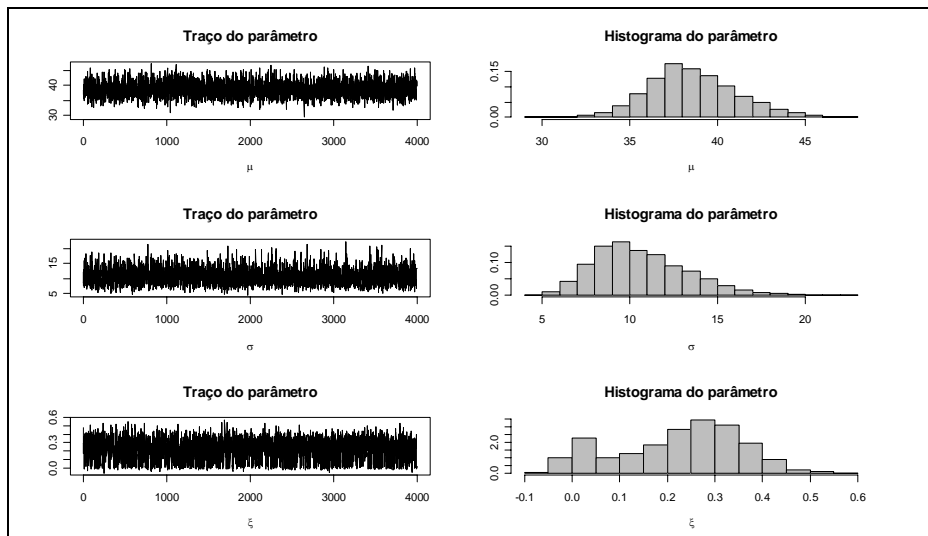
<b>FIGURA</b>		<b>Página</b>
27 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 40, \sigma = 10$ e $\xi = 0,20$ .....	86
28 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 40, \sigma = 20$ e $\xi = -0,30$ .....	86
29 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 40, \sigma = 20$ e $\xi = 0,10$ .....	87
30 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 40, \sigma = 20$ e $\xi = 0,20$ .....	87
31 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 70, \sigma = 10$ e $\xi = -0,30$ .....	88
32 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 70, \sigma = 10$ e $\xi = 0,10$ .....	88
33 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 70, \sigma = 10$ e $\xi = 0,20$ .....	89
34 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 70, \sigma = 20$ e $\xi = -0,30$ .....	89
35 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 70, \sigma = 20$ e $\xi = 0,10$ .....	90
36 B	Traços e histogramas das marginais dos parâmetros $\mu, \sigma$ e $\xi$ da distribuição GEV para a série de dados simulados com $n = 100, \mu = 70, \sigma = 20$ e $\xi = 0,20$ .....	90



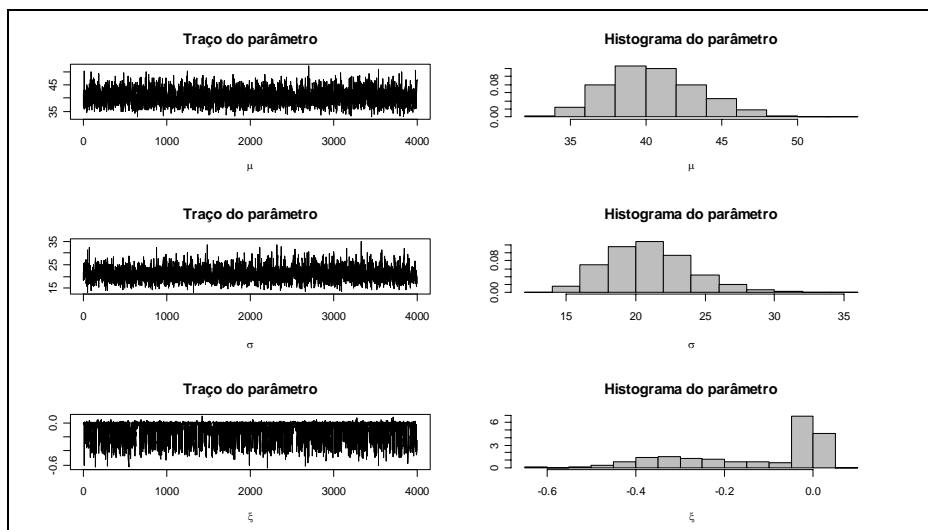
**FIGURA 1 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 40, \sigma = 10$  e  $\xi = -0,30$ .



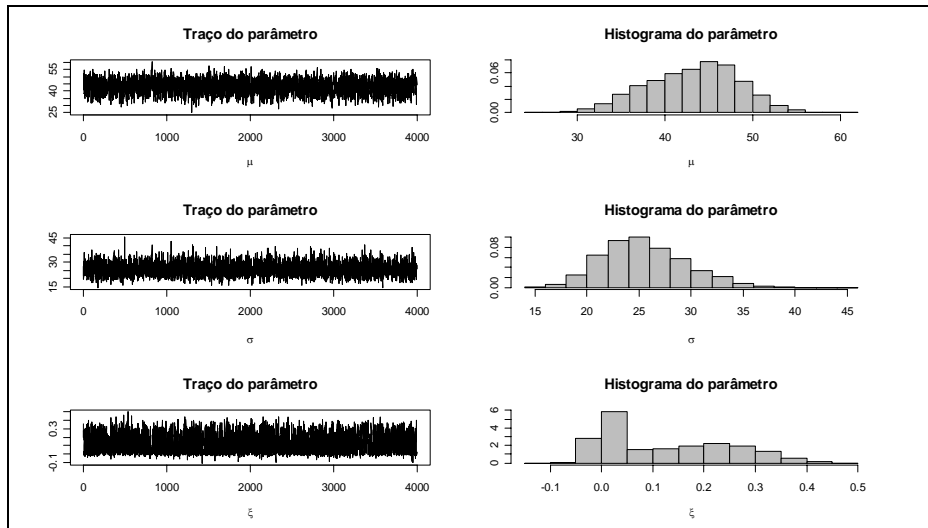
**FIGURA 2 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 40, \sigma = 10$  e  $\xi = 0,10$ .



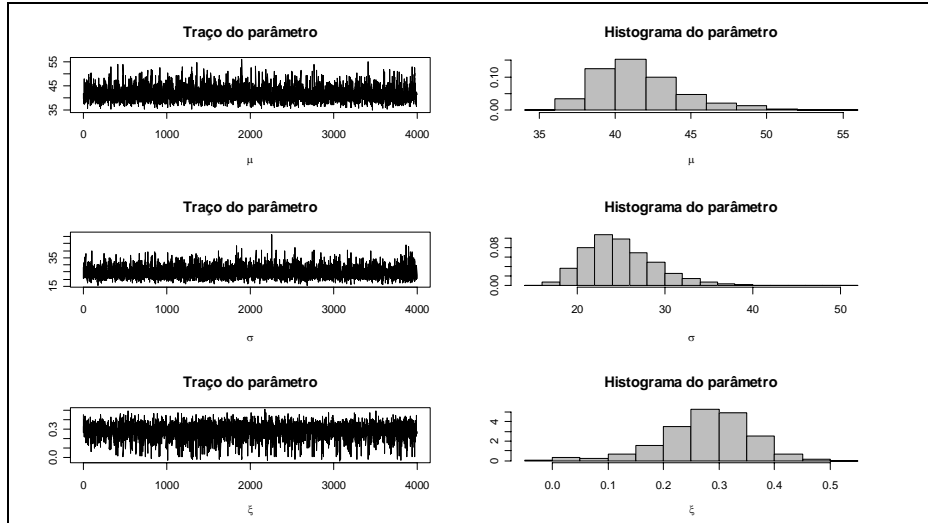
**FIGURA 3 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30$ ,  $\mu = 40$ ,  $\sigma = 10$  e  $\xi = 0,20$ .



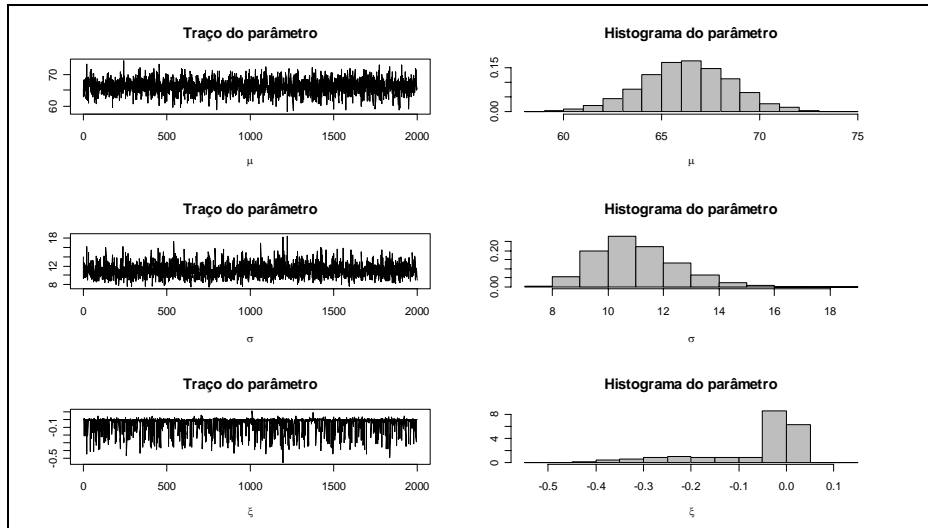
**FIGURA 4 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30$ ,  $\mu = 40$ ,  $\sigma = 20$  e  $\xi = -0,30$ .



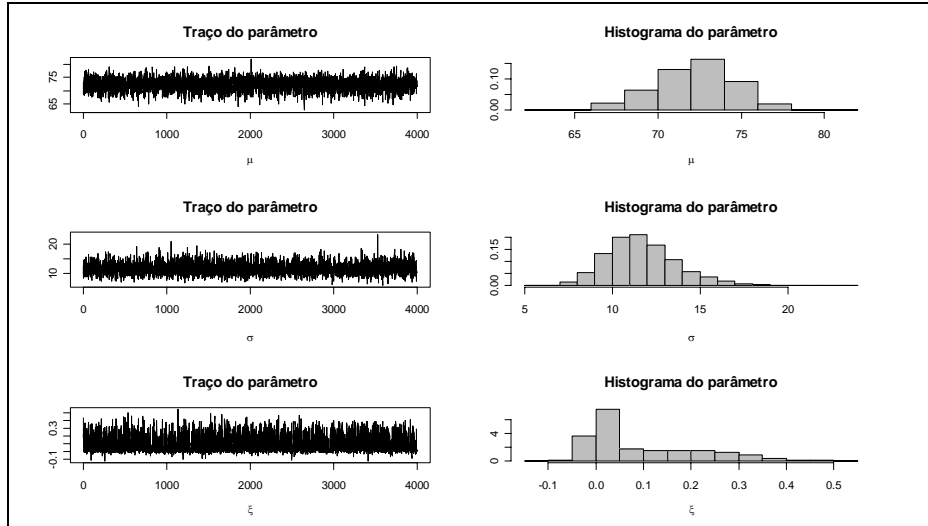
**FIGURA 5 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 40, \sigma = 20$  e  $\xi = 0,10$ .



**FIGURA 6 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 40, \sigma = 20$  e  $\xi = 0,20$ .

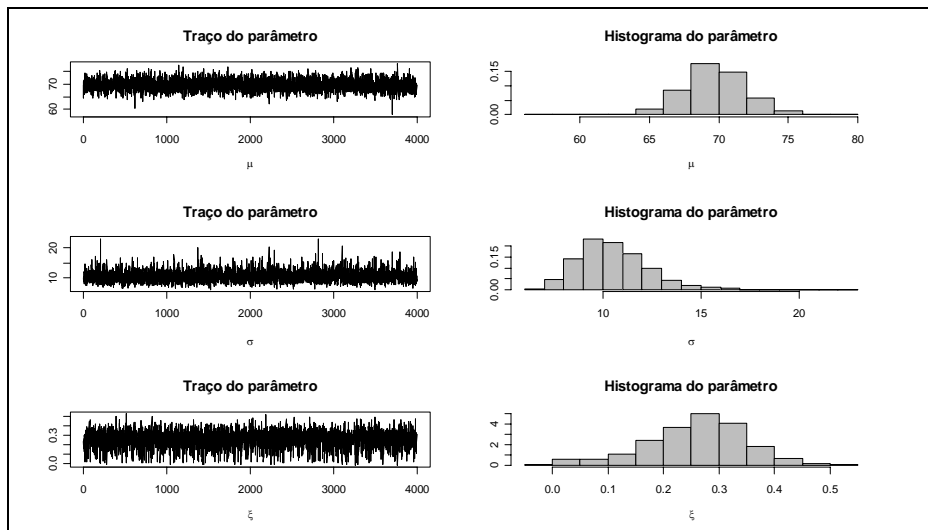


**FIGURA 7 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 70, \sigma = 10$  e  $\xi = -0,30$ .

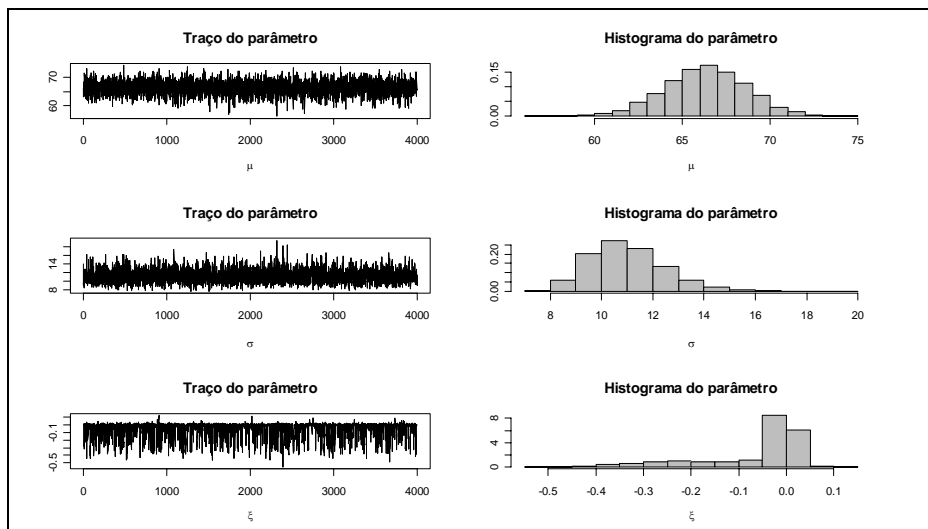


**FIGURA 8 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 70, \sigma = 10$  e  $\xi = 0,10$ .

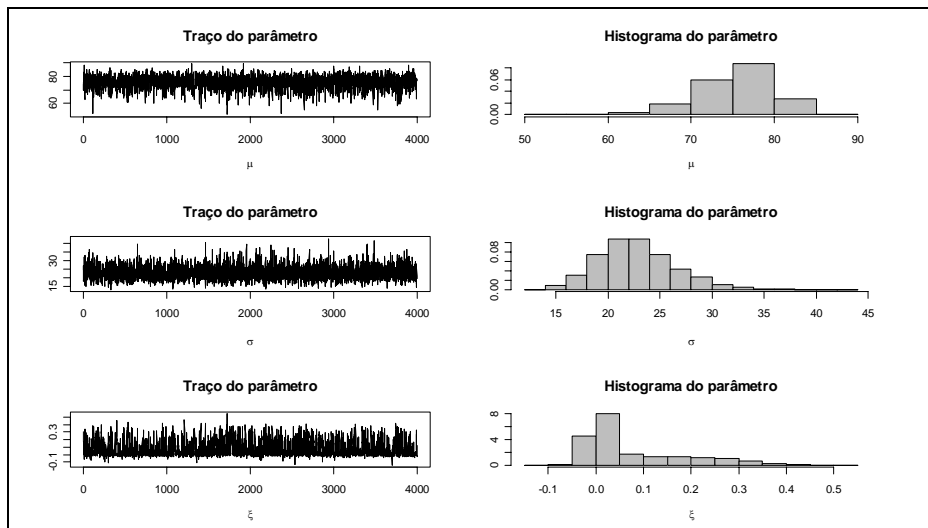




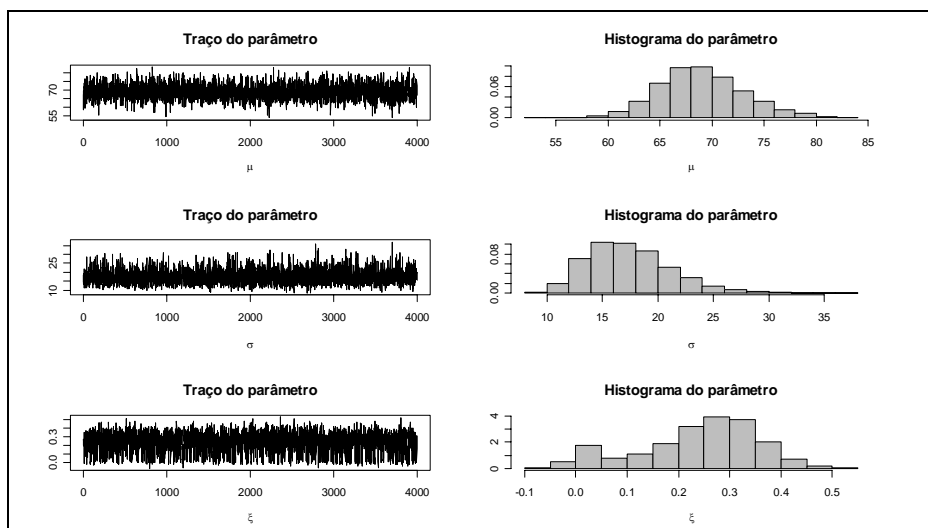
**FIGURA 9 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 70, \sigma = 10$  e  $\xi = 0,20$ .



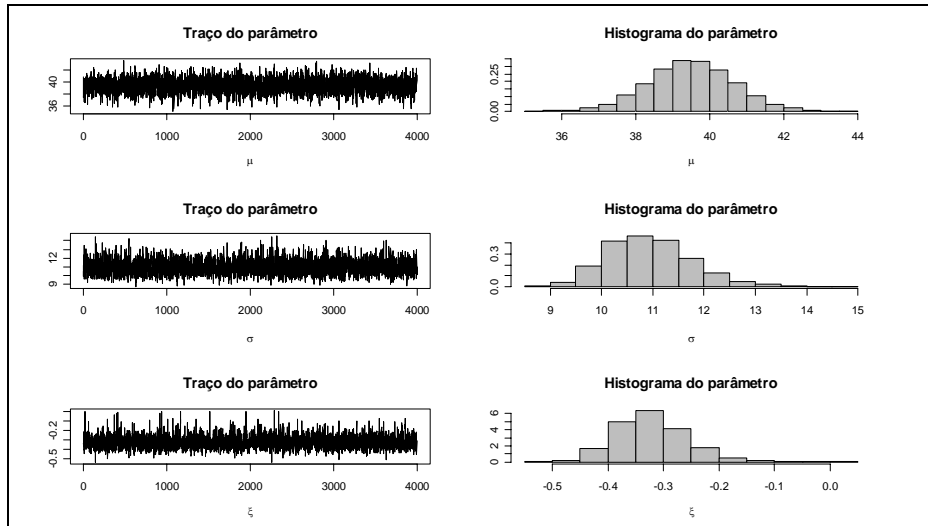
**FIGURA 10 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 70, \sigma = 20$  e  $\xi = -0,30$ .



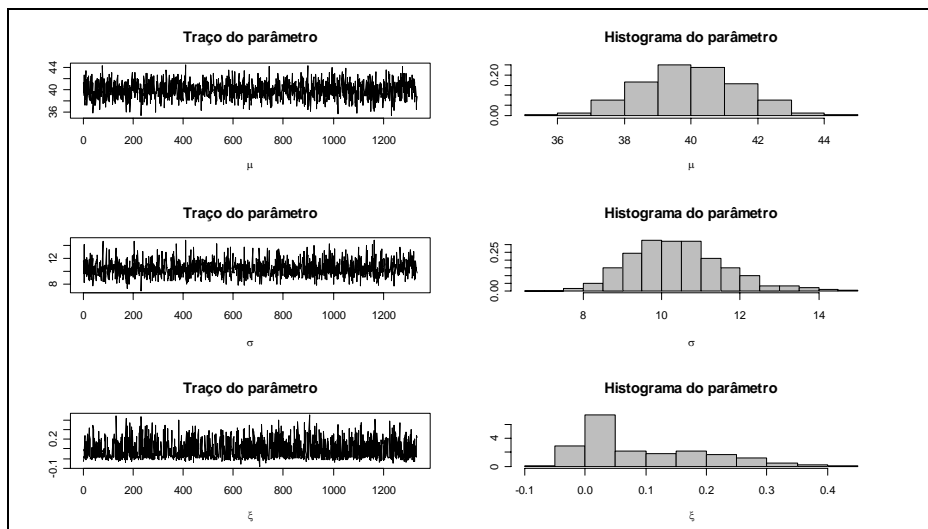
**FIGURA 11 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 70, \sigma = 20$  e  $\xi = 0,10$ .



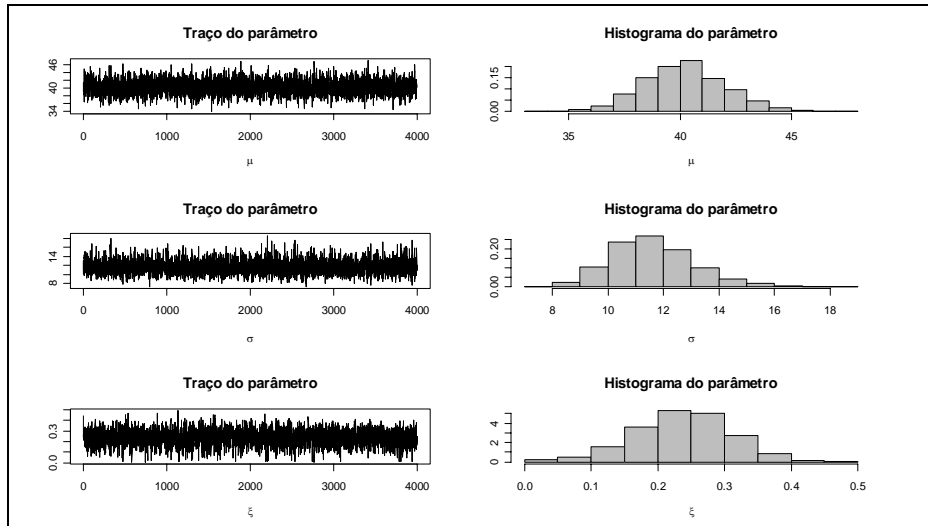
**FIGURA 12 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 30, \mu = 70, \sigma = 20$  e  $\xi = 0,20$ .



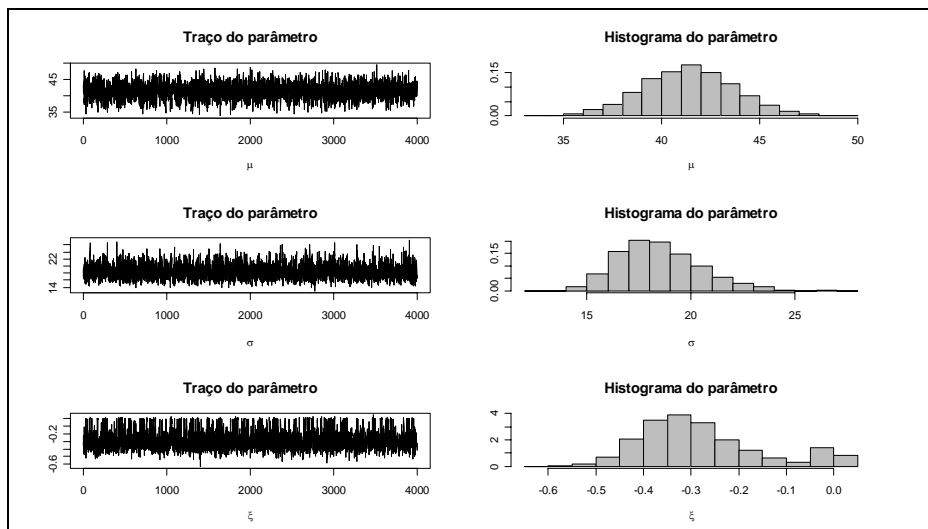
**FIGURA 13 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 40, \sigma = 10$  e  $\xi = -0,30$ .



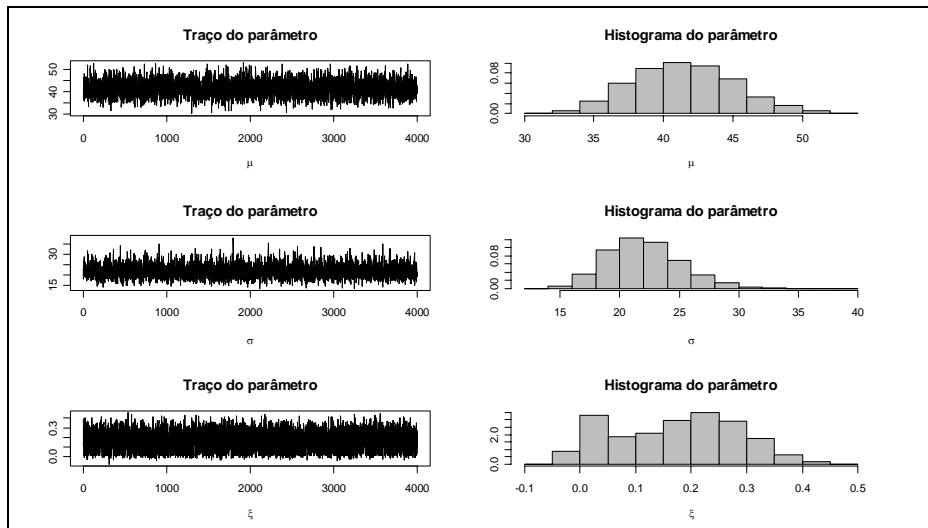
**FIGURA 14 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 40, \sigma = 10$  e  $\xi = 0,10$ .



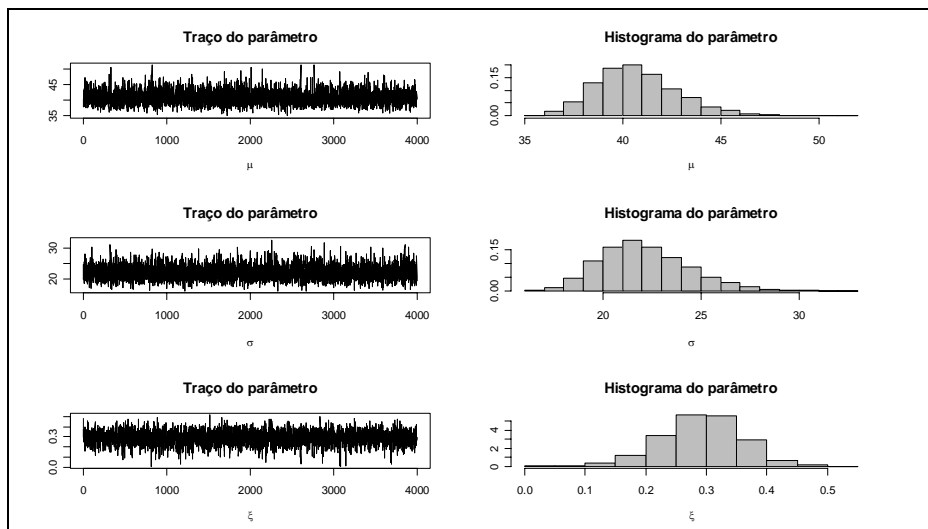
**FIGURA 15 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 40, \sigma = 10$  e  $\xi = 0,10$ .



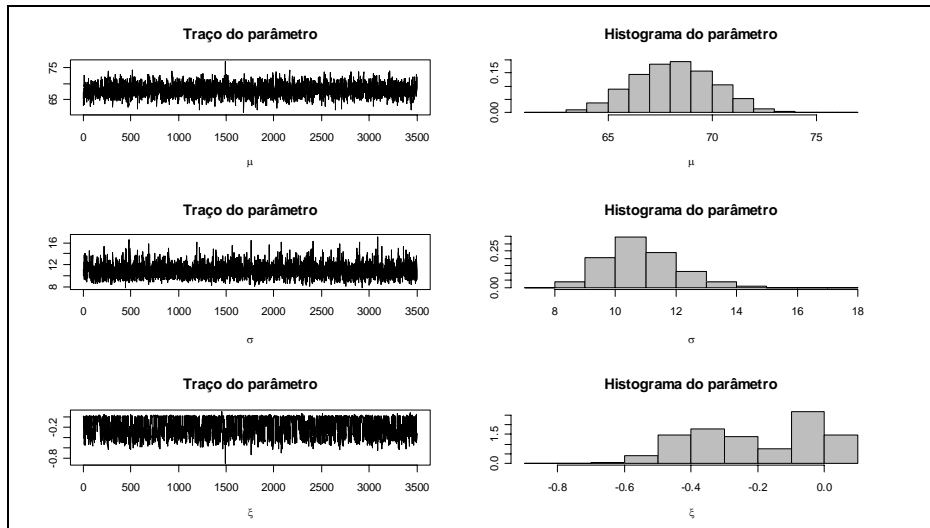
**FIGURA 16 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 40, \sigma = 20$  e  $\xi = -0,30$ .



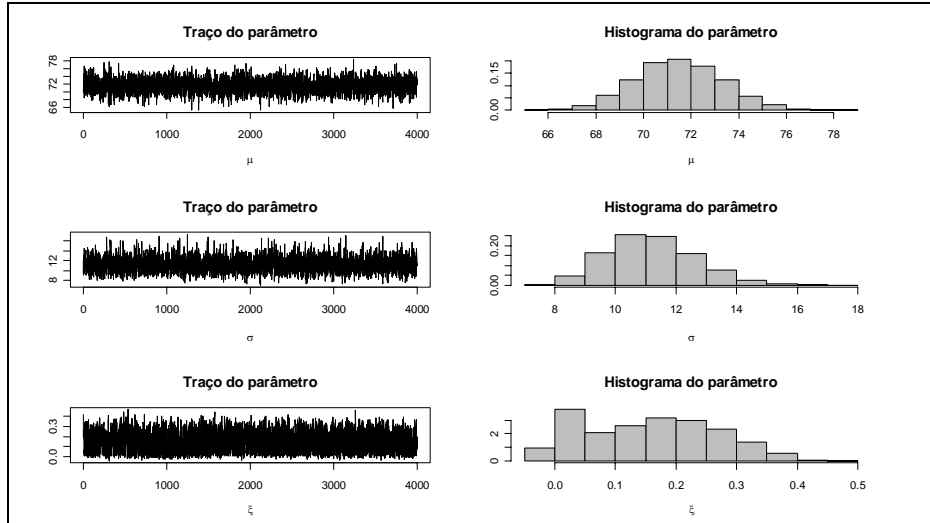
**FIGURA 17 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 40, \sigma = 20$  e  $\xi = 0,10$ .



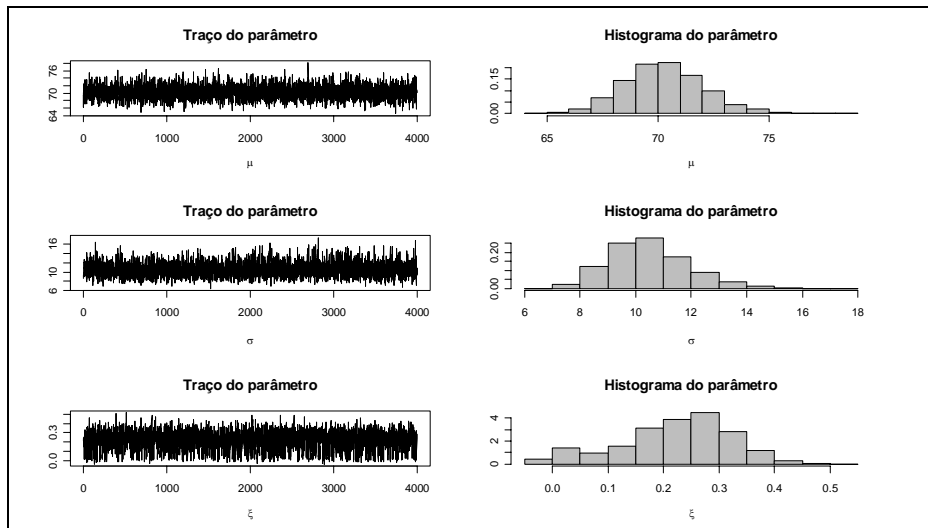
**FIGURA 18 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 40, \sigma = 20$  e  $\xi = 0,20$ .



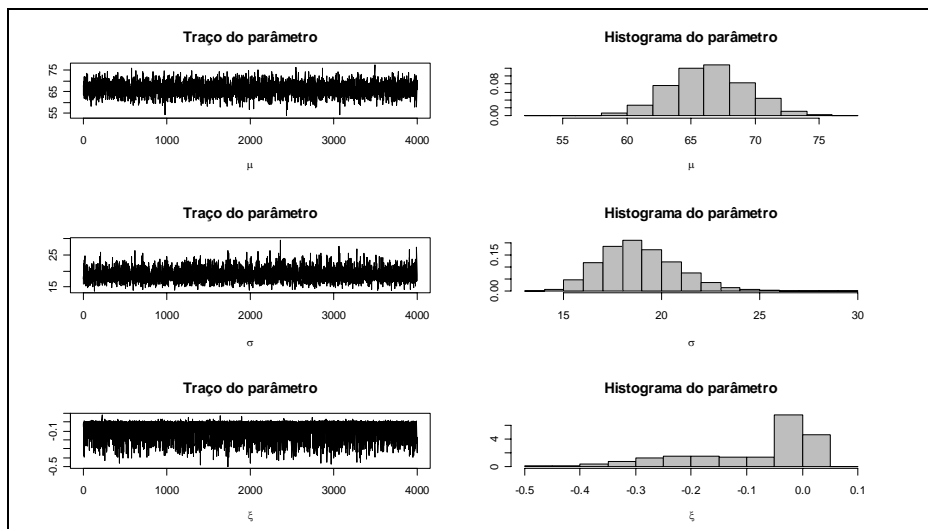
**FIGURA 19 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 70, \sigma = 10$  e  $\xi = -0,30$ .



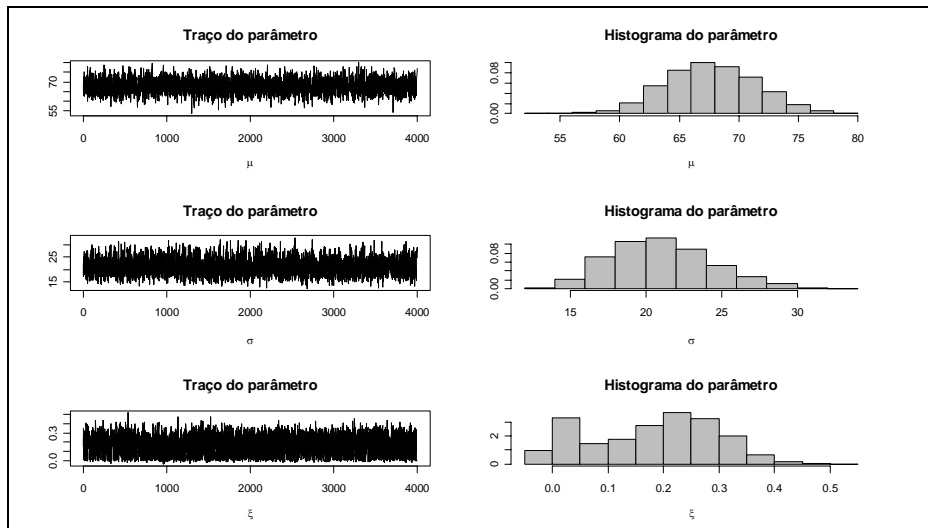
**FIGURA 20 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 70, \sigma = 10$  e  $\xi = 0,10$ .



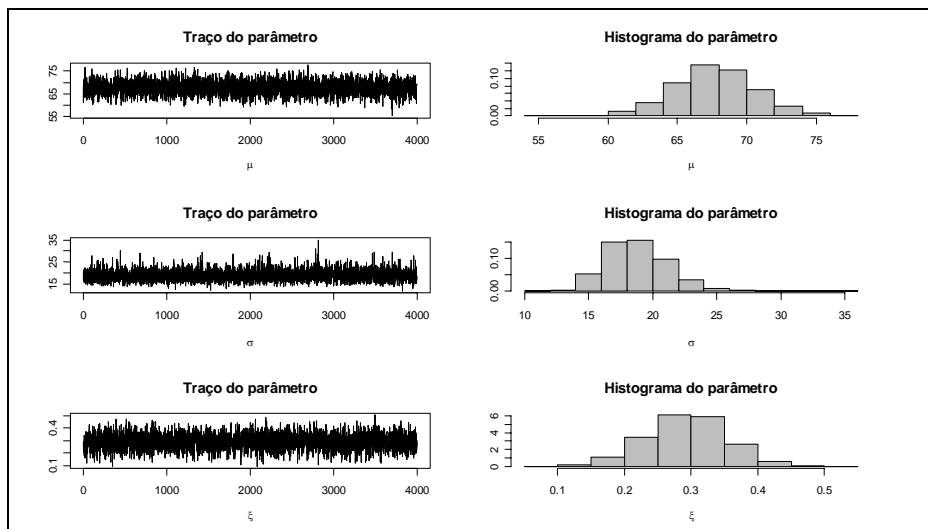
**FIGURA 21 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50$ ,  $\mu = 70$ ,  $\sigma = 10$  e  $\xi = 0,20$ .



**FIGURA 22 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50$ ,  $\mu = 70$ ,  $\sigma = 20$  e  $\xi = -0,30$ .

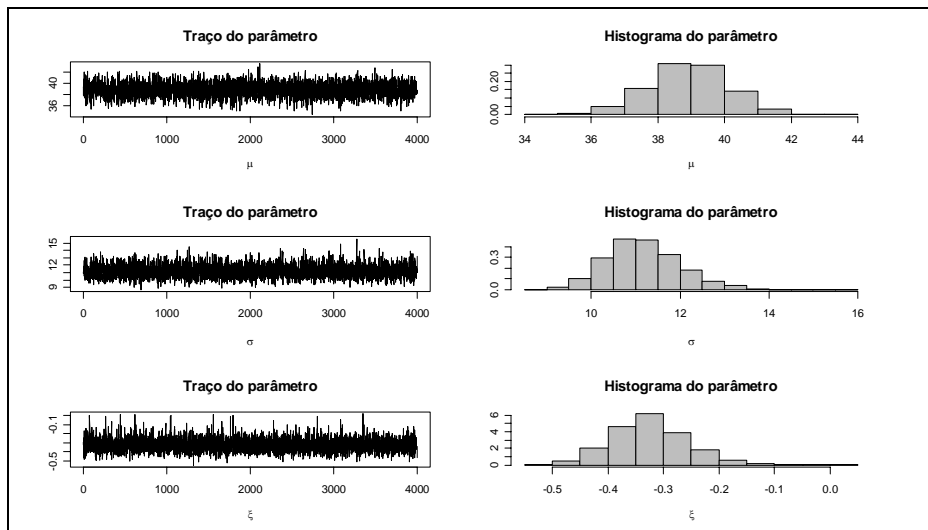


**FIGURA 23 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 70, \sigma = 20$  e  $\xi = 0,10$ .

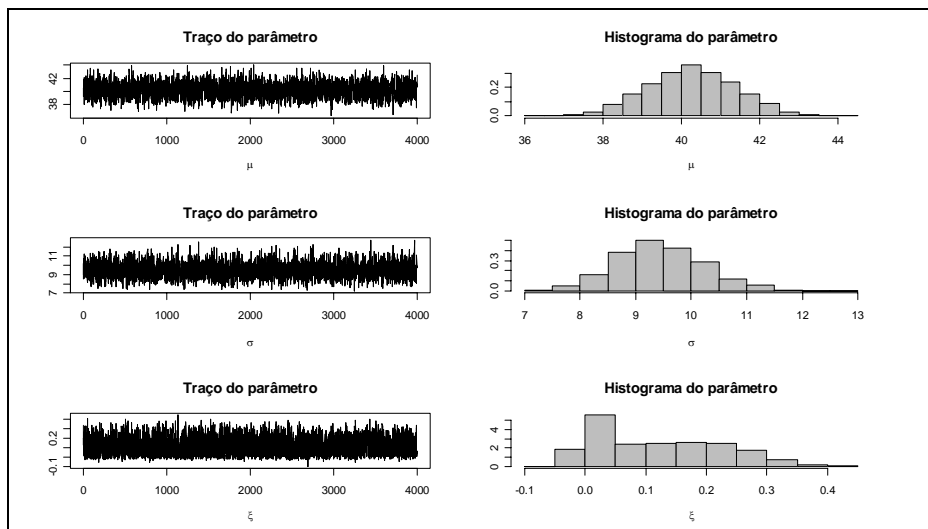


**FIGURA 24 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 50, \mu = 70, \sigma = 20$  e  $\xi = 0,20$ .

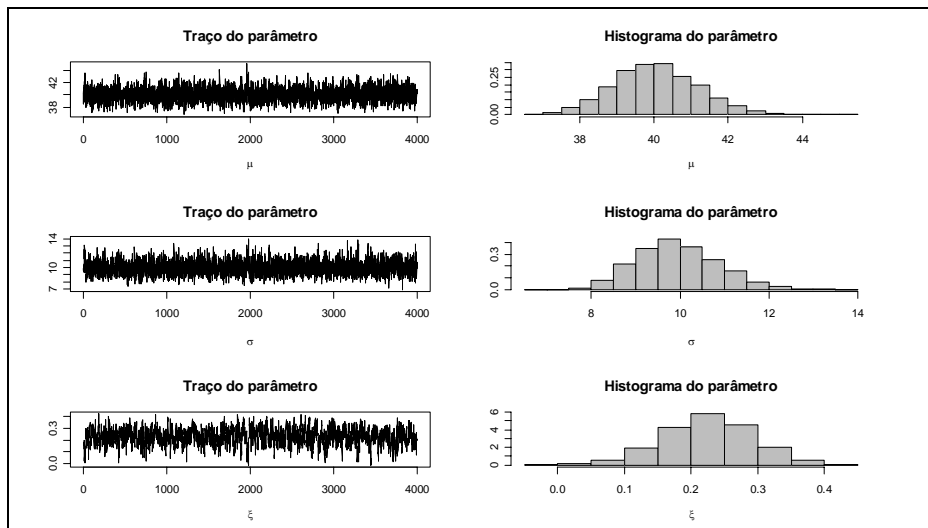




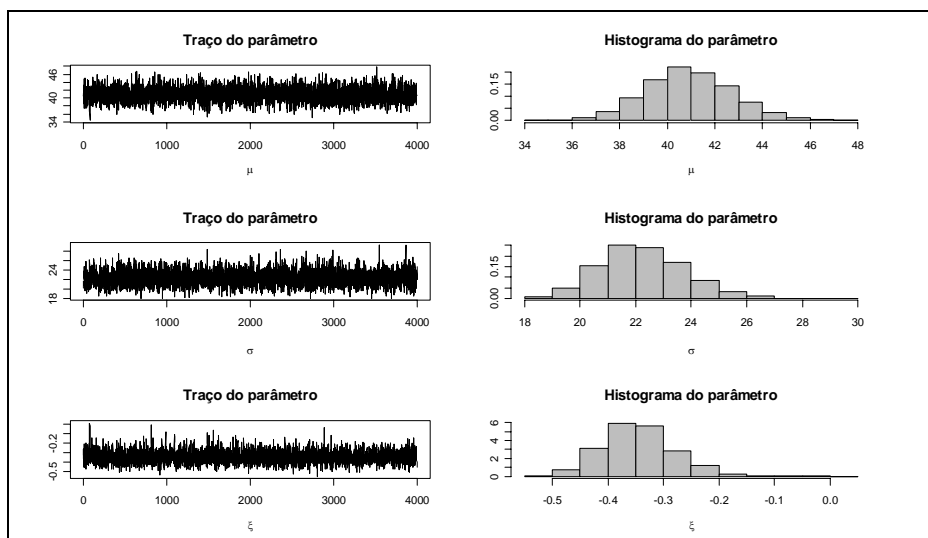
**FIGURA 25 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 40, \sigma = 10$  e  $\xi = -0,30$ .



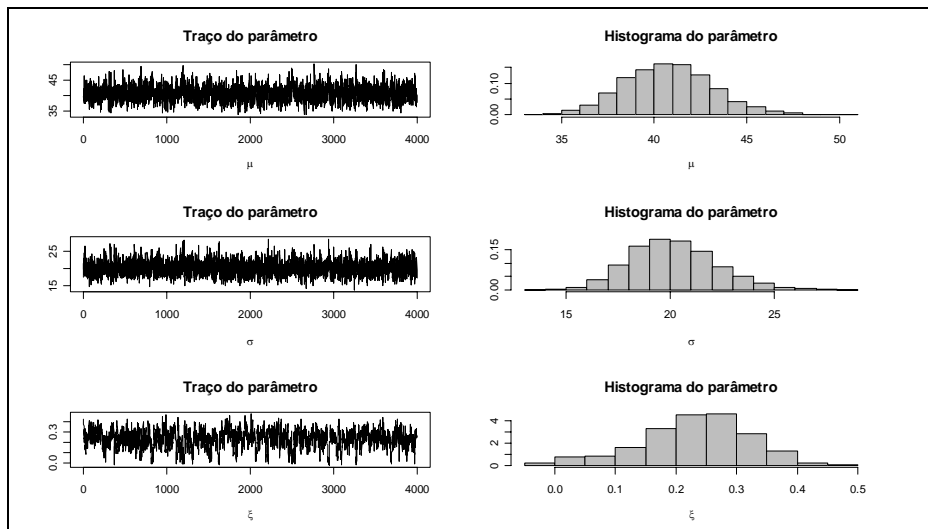
**FIGURA 26 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 40, \sigma = 10$  e  $\xi = 0,10$ .



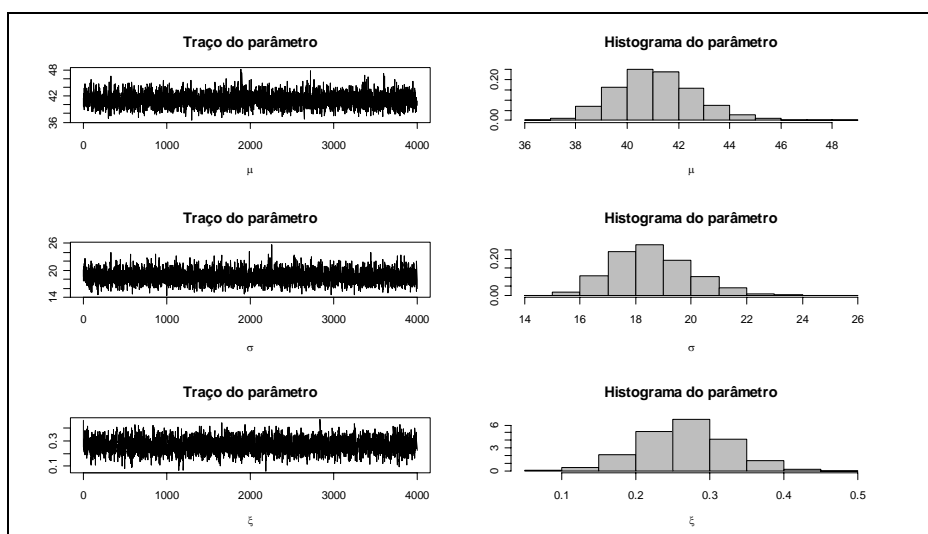
**FIGURA 27 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100$ ,  $\mu = 40$ ,  $\sigma = 10$  e  $\xi = 0,20$ .



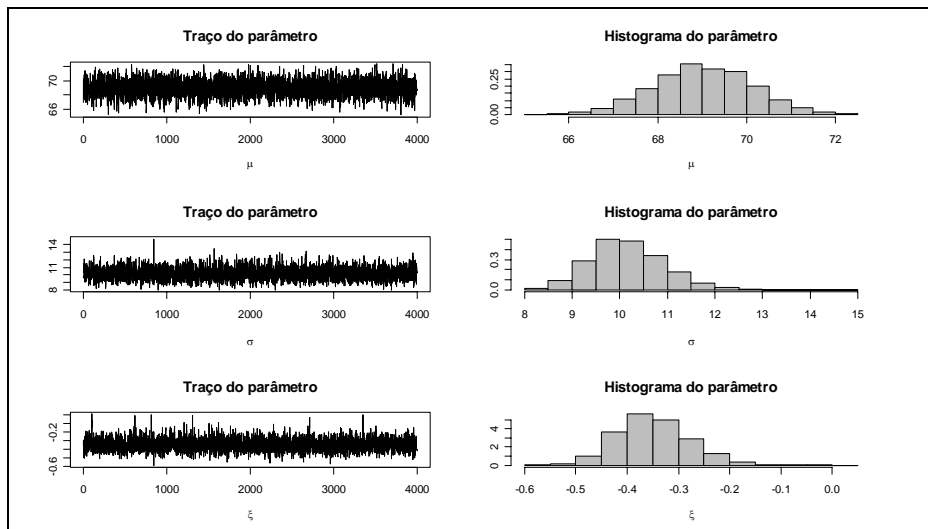
**FIGURA 28 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100$ ,  $\mu = 40$ ,  $\sigma = 20$  e  $\xi = -0,30$ .



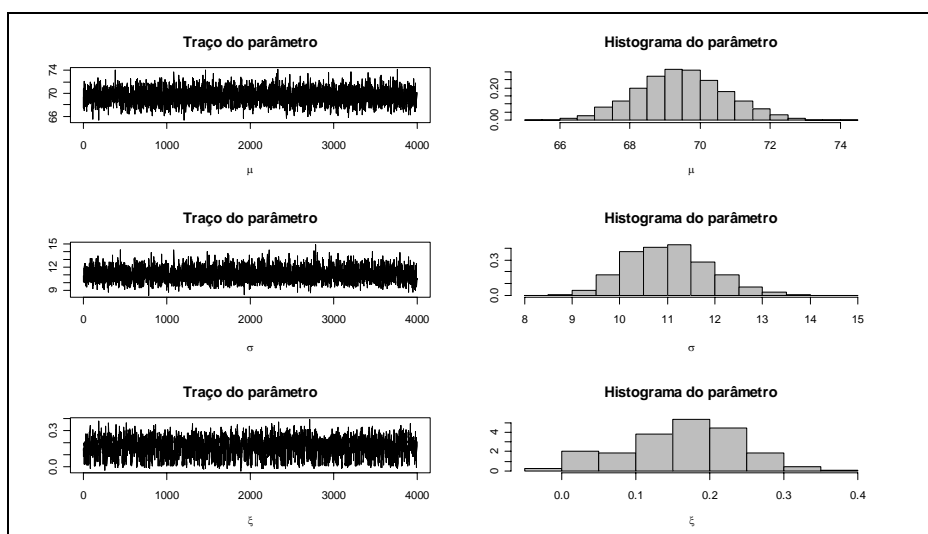
**FIGURA 29 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 40, \sigma = 20$  e  $\xi = 0,10$ .



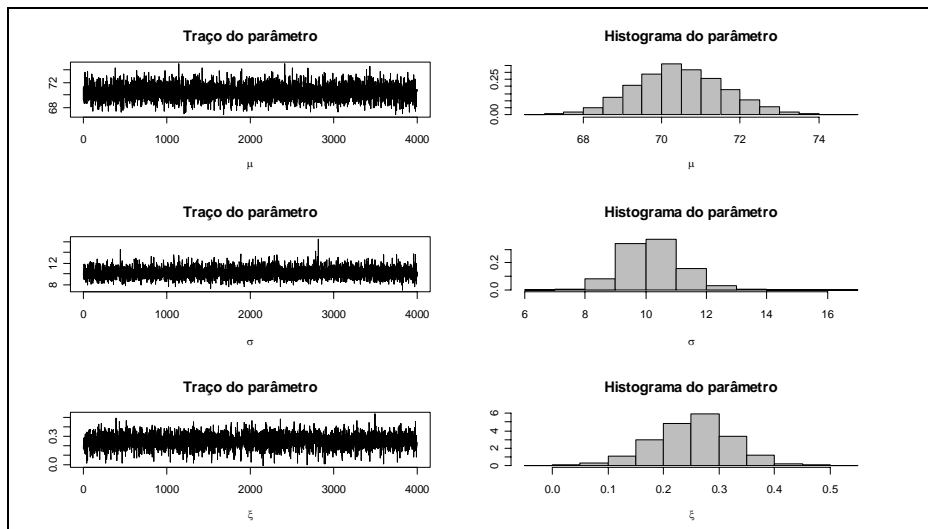
**FIGURA 30 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 40, \sigma = 20$  e  $\xi = 0,20$ .



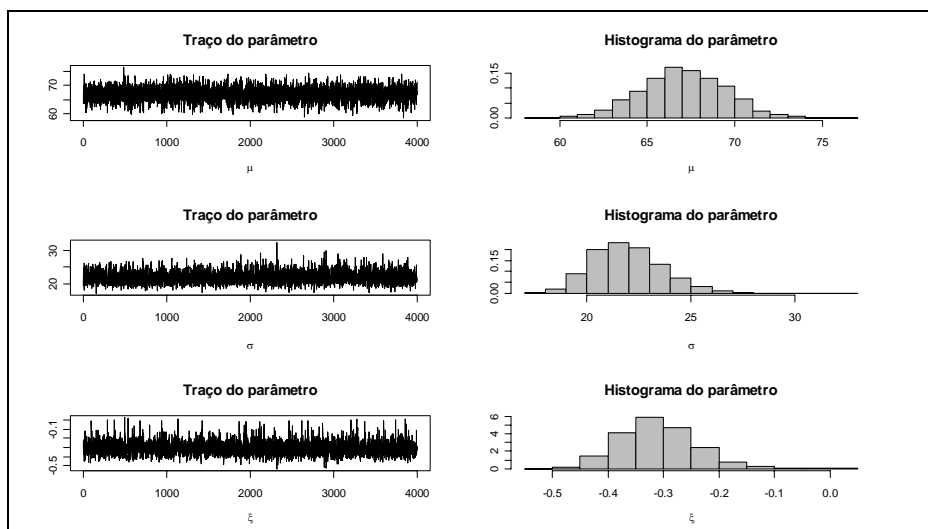
**FIGURA 31 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 70, \sigma = 10$  e  $\xi = -0,30$ .



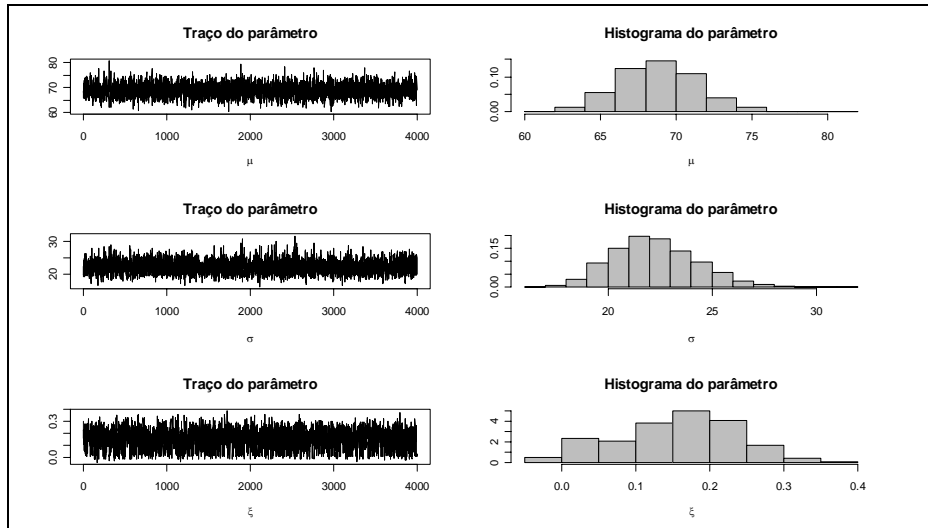
**FIGURA 32 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 70, \sigma = 10$  e  $\xi = 0,10$ .



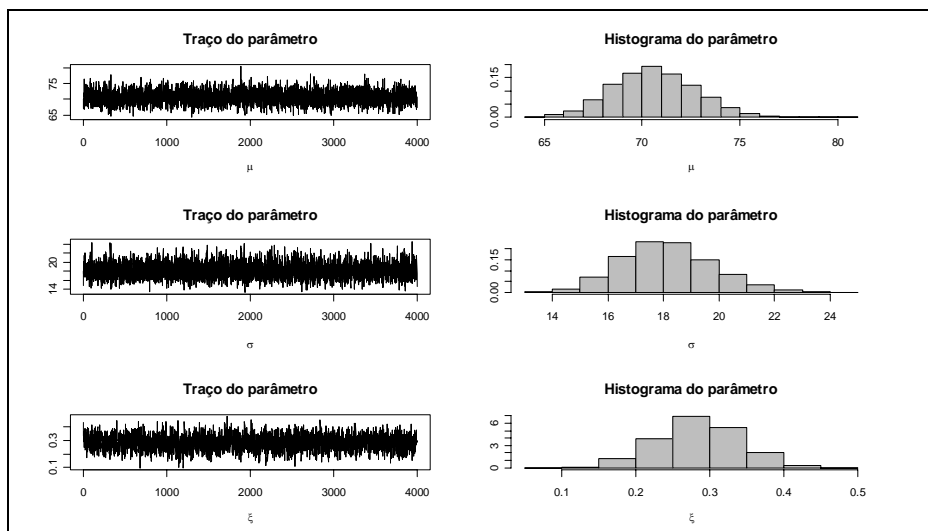
**FIGURA 33 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100$ ,  $\mu = 70$ ,  $\sigma = 10$  e  $\xi = 0,20$ .



**FIGURA 34 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100$ ,  $\mu = 70$ ,  $\sigma = 20$  e  $\xi = -0,30$ .



**FIGURA 35 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 70, \sigma = 20$  e  $\xi = 0,10$ .



**FIGURA 36 B:** Traços e histogramas das marginais dos parâmetros  $\mu, \sigma$  e  $\xi$  da distribuição GEV para a série de dados simulados com  $n = 100, \mu = 70, \sigma = 20$  e  $\xi = 0,20$ .

## ANEXOS C

	Página
<b>PROGRAMA 1C</b> Rotina R para gerar as cadeias via MCMC.....	92
<b>PROGRAMA 2C</b> Rotina R para cálculo dos quantis da Distribuição Generalizada de valores Extremos.....	94

### Programa 1C: Rotina R para gerar as cadeias via MCMC

**#Criando a Função posteriori dada pela expressão (19)#**

```
postm<-function(pa) {  
  
  quant<-numeric(3)  
  p<-numeric(3)  
  k<-numeric(3)  
  y<-numeric(3)  
  for (i in 1:3){  
    if (abs(pa[3])<=0.001){  
      p[i]=log(-log(1-pr[i]))  
      quant[i]=pa[1]-(log(pa[2])*p[i])  
    }  
    else{  
      p[i]=-log(1-pr[i])  
      quant[i]=(pa[1]+((log(pa[2])/pa[3])*  
(p[i]^(-pa[3]))-1)))  
    }  
  
    if (abs(pa[3])<=0.001){  
      lj=((p[1]*p[2])*(p[2]-p[1]))-((p[1]*p[3])*  
(p[3]-p[1]))+((p[2]*p[3])*(p[3]-p[2]))  
  
      lj=log(abs(lj))+log(log(pa[2]))-log(2)  
    }  
  
    else {  
      lj=((p[1]*p[2])^(-pa[3]))*(log(p[2]/p[1]))  
      +((p[2]*p[3])^(-pa[3]))*(log(p[3]/p[2]))  
      -((p[1]*p[3])^(-pa[3]))*(log(p[3]/p[1]))  
  
      lj=log(abs(lj))+log(log(pa[2]))-log(pa[3]^2)  
    }  
    for (i in 1:3){  
      k[i]=(quant[i]-a[i])/b[i]  
      y[i]=((1/b[i])*(exp(-k[i]))*(exp(-exp(-k[i]))))  
    }  
  
    ld=log(y[1])+log(y[2])+log(y[3])  
  
    prime=ld+lj ## Indica a distribuição a priori ##  
  }  
}
```



```

#Função de verossimilhança implementada no pacote EVDBAYES#

vero=gevlik2(par=pa,x)

postm=prime+vero ## Criou a posteriori ##
}

## Gerando um vetor de dados de tamanho "n" a partir da
distribuição GEV, com os parâmetros posição=loc,
escala=scale e forma=shape.##

x<-rgev(n=100, loc=70, scale=20, shape=0.2)

## Declarando as probabilidades e os hiperparâmetros##

pr<-c(0.1,0.01,0.005)
a<-c(60,90,110)
b <-c(2,6,10)

## MCMCmetrop1R: Função implementada no pacote MCMCPack do
R que realiza o algoritmo de Metropolis com passeio
aleatório ###

ptmc <- MCMCmetrop1R(postm, theta.init=c(65.5,23,0.12),
thin=100, mcmc=420000, burnin=20000, tune=c(1.2,0.90,0.25),
seed=1,logfun=TRUE, ...)

raftery.diag(ptmc)
plot(ptmc)
summary(ptmc)

```

**Programa 2C:** Rotina R para cálculo dos quantis da Distribuição Generalizada de Valores Extremos

```
k<-length(ptmc[,1])
loc <- numeric(k)
escala <- numeric(k)
forma <- numeric(k)

prob=c(0.1,0.05)
np<-length(prob)
q1 <- numeric(k)
q2 <- numeric(k)

for (i in 1:k){
    loc[i] <- ptmc[i,1]
    escala[i] <- ptmc[i,2]
    forma[i] <- ptmc[i,3]

    q1[i]<-(loc[i]+((escala[i]/forma[i])*(((log(1-prob[1]))^
    (-forma[i]))-1)))
    q2[i]<-(loc[i]+((escala[i]/forma[i])*(((log(1-prob[2]))^
    (-forma[i]))-1)))
    }

nit<-length(q1)
qo<-matrix(0,nit,2)
qo[,1]<-q1
qo[,2]<-q2
```

## ANEXOS D

	Página
<b>TABELA 1D</b> Precipitações pluviiais máximas (mm) ocorridas nos meses mais chuvosos na região de Jaboticabal-SP entre 1956 e 1990.....	96

**TABELA 1D** - Precipitações pluviiais máximas (mm) ocorridas nos meses mais chuvosos da região de Jaboticabal-SP entre 1956 e 1990

	Setembro	Outubro	Novembro	Dezembro	Janeiro	Fevereiro	Março	Ano
1956	29	47	25	36	24	68	15	68
1957	22	24	70,6	43	38	42	73	73
1958	45	42,2	30,5	33,5	70	30,2	36,6	70
1959	7,2	50,8	51	49	120,8	28,6	45	120,8
1960	1,2	35,4	60	54,2	62,5	35	44,2	62,5
1961	3	19	17	43	48	73,5	41,7	73,5
1962	13,4	35,2	29	52,2	125,5	68,4	75	125,5
1963	3	22	19,6	35,2	64,8	40,6	23,5	64,8
1964	41	44	39,2	49,2	41	54	18,7	54
1965	35,6	45	21	57	40	54,5	38	57
1966	12,8	42	40,5	91	31,5	44,5	52,5	91
1967	20,5	31,5	75,5	96	90	51,5	47,5	96
1968	19	56	42	27	47,5	27,5	48	56
1969	48,2	33,9	56,3	60,7	26	90,9	51,2	90,9
1970	22,6	40,8	19,4	56,5	53,3	68,4	23,1	68,4
1971	21,3	28,5	89,4	53,9	27	25	71,1	89,4
1972	41,8	53,4	57,3	59,3	72,1	49,4	53,2	72,1
1973	20,3	48,1	48,2	57,7	29,5	38,5	41,4	57,7
1974	1,6	35,9	32,6	45,2	41,8	49,8	63,5	63,5
1975	42	27,8	77,5	76	29,2	98	31	98
1976	50	39,3	30	38,5	24	34,1	57,6	57,6
1977	18,8	16,7	57,7	90,6	87,8	29,6	24,7	90,6
1978	33,9	75,5	39,1	47	30,3	78,6	60,3	78,6
1979	64,7	34,3	26,6	63,4	26,4	43,3	26,5	64,7
1980	32,8	28,3	38,5	56,1	48,5	48,7	67,7	67,7
1981	17,7	63,5	31,9	39,5	57,9	47,2	49,8	63,5
1982	17,9	38,7	48	58,7	63,4	62,7	57,9	63,4
1983	50,1	78,9	47,6	102,3	80	125,4	37,2	125,4
1984	36,3	52,5	50	37,7	58,8	25,2	43,7	58,8
1985	0,7	61,3	123	31,6	87,5	87,8	31,7	123
1986	15,3	11,5	39,7	79	81,6	39,4	40,3	81,6
1987	12,3	22,6	30,5	29,9	46,7	64,6	38,9	64,6
1988	0	69,7	58	45,8	75,2	43,6	67,8	75,2
1989	5,9	24,7	47,6	48,6	48	70,4	30,4	70,4
1990	8,4	18,8	28,2	84,4	37,4	33,2	108,1	108,1

## ANEXOS E

	Página
<b>TABELA 1E</b> Estimativas das precipitações pluviais máximas (mm) nos meses mais chuvosos na região de Jaboticabal-SP, para diferentes tempos de retorno e limites inferior (LI) e superior (LS) de seus respectivos intervalos HPD de 99 % obtidas via priori proposta por Coles & Tawn (1996) .....	98

**TABELA 1E** - Estimativas das precipitações pluviiais máximas (mm) nos meses mais chuvosos na região de Jaboticabal-SP, para diferentes tempos de retorno e limites inferior (LI) e superior (LS) de seus respectivos intervalos HPD de 99 % obtidas via priori proposta por Coles & Tawn (1996).

Período	Período de Retorno											
	10 anos			20 anos			25 anos			50 anos		
	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS
Setembro	38,0	49,8	62,9	48,6	62,5	78,0	51,7	66,7	83,0	62,2	80,2	99,9
Outubro	56,8	68,5	82,2	71,1	86,4	102,7	77,0	92,8	110,8	94,9	114,9	136,7
Novemb.	58,4	71,1	85,9	72,7	89,1	106,5	78,7	95,5	114,5	97,3	117,8	140,1
Dezembro	81,5	91,7	102,8	90,2	102,7	115,3	93,4	106,5	120,1	104,9	119,2	137,0
Janeiro	64,0	78,3	94,6	79,0	97,0	116,7	84,5	103,6	124,6	102,6	126,6	149,9
Fevereiro	62,5	76,7	91,1	78,2	94,6	113,3	83,0	101,0	120,3	103,1	123,2	148,3
Março	61,2	70,1	79,6	71,4	82,4	93,1	74,9	86,4	97,6	86,5	99,3	113,6

Fonte: Beijo et al. (2005b)