

XI SIMPÓSIO DE RECURSOS HIDRÍCOS DO NORDESTE

DISTRIBUIÇÃO GENERALIZADA DE VALORES EXTREMOS (GEV): METODOLOGIAS PARA ESTIMATIVA DE PARÂMETROS

Benedito Martins Gomes¹; Wagner Alessandro Pansera² & Manoel Moises Ferreira de Queiroz³

RESUMO – O conhecimento da distribuição das precipitações intensas no tempo e no espaço tem grande importância no planejamento das atividades agrícolas. A distribuição generalizada de valores extremos é caracterizada por três parâmetros, englobando as três formas assintóticas de valores extremos. Existem diversas metodologias para estimativa dos parâmetros da GEV, entre elas: momentos, momentos-L, momentos-LH e verossimilhança. Na prática, o que interessa são os quantis associados aos períodos de retorno, e não os parâmetros da GEV propriamente. Portanto, este trabalho teve por objetivo comparar os quantis estimados pelas diferentes metodologias estudadas. Observou-se que os momentos, momentos-L e momentos-LH produzem estimativas semelhantes, já a verossimilhança foi a mais discrepante.

ABSTRACT– Distribution of intense rainfall in time and space has great importance for agricultural activities planning. Distribution of generalized extreme value is characterized by three parameters, including the three asymptotic forms of extreme values. There are several methodologies to GEV parameters estimating, among them: moments, L-moments, LH-moments and likelihood. In practice, what matters is the quantile associated with return periods, and not GEV parameters. Therefore, this study aimed to compare the quantiles estimated by different methods studied. It was observed that moments, L-moments and LH-moments produce similar estimates, since the likelihood was the most discrepant.

Palavras-Chave – momentos, chuvas intensas, GEV.

INTRODUÇÃO

O conhecimento da distribuição das precipitações pluviiais no tempo e no espaço tem grande importância no planejamento das atividades agrícolas. A modelagem da precipitação é de grande utilidade para diversos campos do conhecimento, destacando-se sua aplicação na área da Engenharia Agrícola, em que os modelos que consideram a chuva como um processo aleatório, e que são utilizados para simular sequências de chuvas mantendo as características estatísticas da série histórica, são conhecidos como modelos estocásticos. Um modelo estocástico é aquele que possui parâmetros obtidos a partir dos dados observados e que, apesar de não ser possível prever

1) Afiliação: Universidade Estadual do Oeste do Paraná, Rua Universitária - Jardim Universitário, Cascavel/PR, (45) 32203175, (45) 33247922, bmgomes@unioeste.br

2) Afiliação: Universidade Estadual do Oeste do Paraná, Rua Universitária - Jardim Universitário, Cascavel/PR, (45) 32203175, (45) 33247922, wapansera@yahoo.com.br

3) Afiliação: Universidade Federal de Campina Grande, Rua Jairo Viera Feitosa – Bairro dos Pereiros, Pombal/PB, (83) 3431-4016, (83) 34314016, ppgsa@ccta.ufcg.edu.br

precisamente os valores que assumirá a precipitação no futuro, existe a possibilidade de fazer avaliações das probabilidades associadas aos valores futuros (Calgaro et al., 2009).

As chuvas constituem-se na principal entrada de água em uma bacia hidrográfica, e a sua quantificação, bem como o conhecimento da forma como se distribui temporal e espacialmente são fundamentais em estudos relacionados à necessidade de irrigação, disponibilidade de água para abastecimento doméstico e industrial, erosão do solo, controle de inundações, entre outros (Damé et al., 2008).

A caracterização das chuvas intensas é imprescindível para solucionar problemas de interesse da engenharia, de modo especial o controle do escoamento superficial, em áreas urbanas e rurais. O uso de técnicas e modelos hidrológicos permite a caracterização e a quantificação do escoamento superficial originado de chuvas intensas. A meta do hidrólogo não é simplesmente quantificar um evento hidrológico que ocorreu, mas ter a capacidade de prever a ocorrência de eventos extremos e suas possíveis consequências. Atualmente se utiliza cada vez mais modelos matemáticos que têm sido associados a Sistemas de Informações Geográficas (SIG) que permitem uma descrição espacial das características físicas da área de estudo. Apesar dessas considerações, os modelos apresentam deficiências em virtude das simplificações e adaptações necessárias para a sua utilização e também em razão de dificuldades na determinação dos parâmetros a serem utilizados (Santos et al., 2010).

A teoria de valores extremos tem como objetivo o estudo estatístico de fenômenos de risco elevado com grande impacto. A distribuição generalizada de valores extremos (GEV) é caracterizada por três parâmetros (escala, forma e posição) englobando as três formas assintóticas de valores extremos. Eventos extremos podem ser definidos como o máximo de uma variável aleatória em um determinado período.

Existem diversas metodologias para estimativa dos parâmetros da GEV, entre elas: Momentos (M), Momentos-L (ML), Momentos-LH (MLH), Máxima Verossimilhança (MVS).

O método MVS exige um maior esforço computacional, pelo fato de envolver soluções numéricas de sistemas de equações não lineares e implícitas. O método dos M é método de estimação mais simples. Entretanto, os estimadores desse método são, em geral, de qualidade inferior e menos eficientes do que os estimadores de MVS, particularmente para distribuições de três ou mais parâmetros. O método dos ML produz estimadores de parâmetros comparáveis, em qualidade, àqueles produzidas pelo método da MVS, com a vantagem de exigirem um menor esforço computacional para a solução de sistemas de equações menos complexas (Naghetini e Pinto, 2007). Os MLH são uma generalização dos ML para caracterizar a parte superior das distribuições e os valores extremos dos dados. No cálculo dos MLH é intencionalmente ignorada a

parte inferior dos dados, por dar mais ênfase na parte superior dos dados para indicação antecipada do valor extremo futuro (Wang, 1997).

Na prática o que interessa são os quantis associados aos níveis de probabilidade de não-excedência ou período de retorno, e não os parâmetros da GEV propriamente.

Uma vez calculados ou estimados valores pontuais, mapas podem ser obtidos por meio de métodos de interpolação que usam os valores amostrados para estimativas de valores em locais com ausência ou informações restritas de precipitação. O processo de interpolação é comumente realizado, utilizando métodos que consideram as informações uniformemente distribuídas na região de interesse ou que consideram apenas os dados pontuais mais próximos, ignorando outros efeitos que podem interferir na distribuição das chuvas. Várias técnicas são utilizadas para a interpolação de dados de chuvas, tais como o inverso do quadrado da distância e o método das isoietas. A geoestatística, que é baseada na teoria das variáveis regionalizadas, vem sendo grandemente usada porque permite obter correlação entre amostras vizinhas, o que possibilita a estimativa em locais não-amostrados, sem tendência e com variância mínima (Montebeller et al., 2007).

Dessa forma para que o processo de interpolação gere informações confiáveis é necessário que as estimativas dos quantis também sejam. Como os quantis são diretamente afetados pelos parâmetros da distribuição de probabilidade GEV e existem diversas formas de estimá-los, há necessidade de estudar seus efeitos nas estimativas dos quantis.

Sabendo-se da importância do conhecimento das precipitações extremas para conservação do solo e água, este trabalho teve por objetivo avaliar o efeito das metodologias de obtenção dos parâmetros da GEV na estimativa de quantis.

MATERIAL E MÉTODOS

Os dados para confecção deste trabalho foram obtidos via sistema de informações hidrológicas HIDROWEB da Agência Nacional de Águas (ANA). Foram usadas 227 estações pluviométricas localizadas no estado do Paraná (Figura 1), sendo utilizados registros de máximas diárias anuais (1976/2006).

A distribuição GEV é descrita pela equação 1:

$$F(x) = P(X \leq x) = \exp \left[- \left(1 - k \frac{x-u}{\alpha} \right)^{\frac{1}{k}} \right] \quad (1)$$

Em que: u é um parâmetro de posicionamento, α é um parâmetro de escala (com $\alpha > 0$) e k é um parâmetro de forma.

Os parâmetros da GEV foram estimados por: Máxima Verossimilhança (Petersen-Øverleir e Reitan, 2009), Momentos e Momentos-L (Naghattini e Pinto, 2007), Momentos-LH (Queiroz e

Chaudhry, 2006). O ajuste da distribuição teórica aos dados de máximos anuais foi averiguado pelo teste de Kolmogorov-Smirnov (Eq.02), sendo que as estações reprovadas foram retiradas da análise.

$$D_N = \sup_{-\infty < x < \infty} |F_N(x) - F_X(x)| \quad (2)$$

Em que: F_N frequência empírica e F_X frequência teórica estimada pela distribuição GEV.

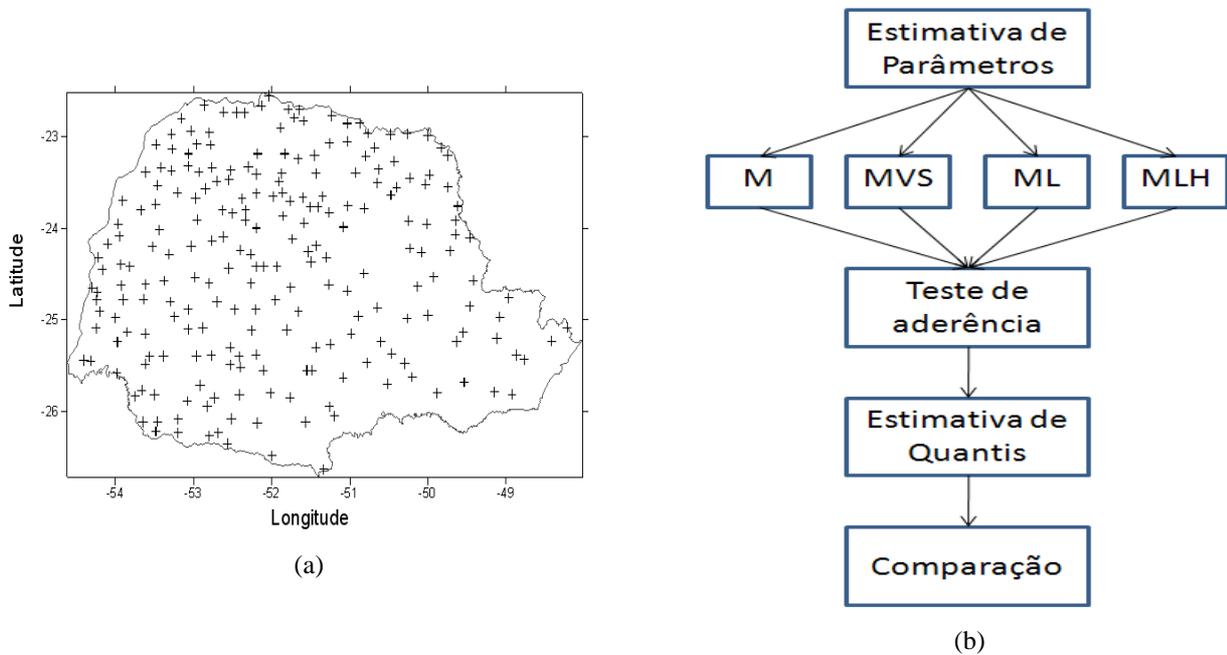


Figura 1 – (a) Localização das estações (b) Fluxograma de pesquisa

Para o tipo de evento hidrológico considerado foi escolhida a metodologia de estimativa de parâmetros que melhor se ajustou a distribuição de frequência da amostra. Um critério de escolha, consiste em escolher a metodologia que minimiza o valor da equação 3.

$$\Delta^2 = \sum_{i=1}^n [F_X - F_N]^2 \quad (3)$$

Estimados os parâmetros da GEV é possível calcular os quantis associados aos períodos de retorno, conforme equação 4.

$$x_T = u + \frac{\alpha}{k} \left[1 - \left(-\ln \left(1 - \frac{1}{T} \right) \right)^k \right] \quad (4)$$

Em que: T é o período de retorno em anos.

Para comparar os quantis estimados pela metodologia que obteve melhor ajuste com os demais, foram utilizados o coeficiente de correlação e o índice de concordância. A dispersão dos valores em torno da média é dada pelo coeficiente de correlação. Este índice indica o grau de dispersão dos dados obtidos, ou seja, o erro aleatório. A concordância se refere à aproximação das

metodologias; pode ser avaliada graficamente pelo afastamento dos pontos cotados no gráfico de regressão em relação à reta de valores iguais, 1:1. Tais índices são estimados pelas equações 5 e 6.

$$r = \frac{\sum_{i=1}^n (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 \sum_{i=1}^n (P_i - \bar{P})^2}} \quad (5)$$

$$d = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (|P_i - \bar{O}| - |O_i - \bar{O}|)^2} \quad (6)$$

Em que: n é o número de estações consideradas, O representa a metodologia de melhor ajuste e P é a metodologia que está sendo comparada.

Foram utilizados também o erro médio absoluto (EMA) e a raiz do erro quadrático médio (RSME), conforme equações 7 e 8.

$$EMA = \frac{1}{n} \sum_{i=1}^n \left| \frac{P_i - O_i}{O_i} \right| \quad (7)$$

$$RSME = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad (8)$$

RESULTADOS E DISCUSSÃO

A tabela 1 mostra um resumo dos parâmetros da GEV, referentes às 227 estações pluviométricas estudadas.

Tabela 1 – Resumo dos parâmetros da GEV para as metodologias de estimativa estudadas

	M			MLH			MVS			ML		
	u	α	k									
Min	55,67	12,74	-0,21	55,23	8,87	-0,45	54,15	10,38	-0,59	55,25	10,40	-0,43
Q1	75,60	18,00	-0,04	74,94	15,40	-0,16	75,19	16,82	-0,09	75,12	16,99	-0,12
Med	81,01	20,64	0,03	80,63	18,93	-0,02	80,81	19,17	0,02	80,48	19,72	-0,02
Q3	87,78	23,65	0,11	87,06	23,82	0,13	87,59	22,26	0,12	87,33	22,15	0,08
Max	119,67	39,48	0,58	114,82	82,11	1,16	119,34	33,39	0,61	118,15	36,92	0,53

Min – Mínimo, Q1 – Quartil 1, Med – Mediana, Q3 – Quartil 3, Max - Máximo

O parâmetro de forma (k) variou entre -0,6 e 0,6 para as metodologias M, ML e MVS. Já para a metodologia MLH variou entre -0,45 e 1,16. O parâmetro de posição (u) variou de 55 a 120. O parâmetro de escala variou entre 10 e 40 para as metodologias M, ML e MVS; para a metodologia MLH variou entre 9 e 82. Observa-se que aproximadamente metade das estações tem parâmetro de forma positivo e a outra metade negativa, como demonstrada pela mediana.

A figura 2 mostra o teste de Kolmogorov-Smirnov para as metodologias de estimativa de parâmetros da GEV, lembrando que para $n=31$ e 5% de probabilidade o valor crítico é 0,238. Não se observam estações reprovadas no teste de Kolmogorov-Smirnov para as metodologias dos momentos, momentos-L e verossimilhança. Porém a metodologia dos momentos-LH observa-se seis estações reprovadas no teste de aderência. As estações reprovadas foram retiradas de todas as metodologias para as análises subsequentes.

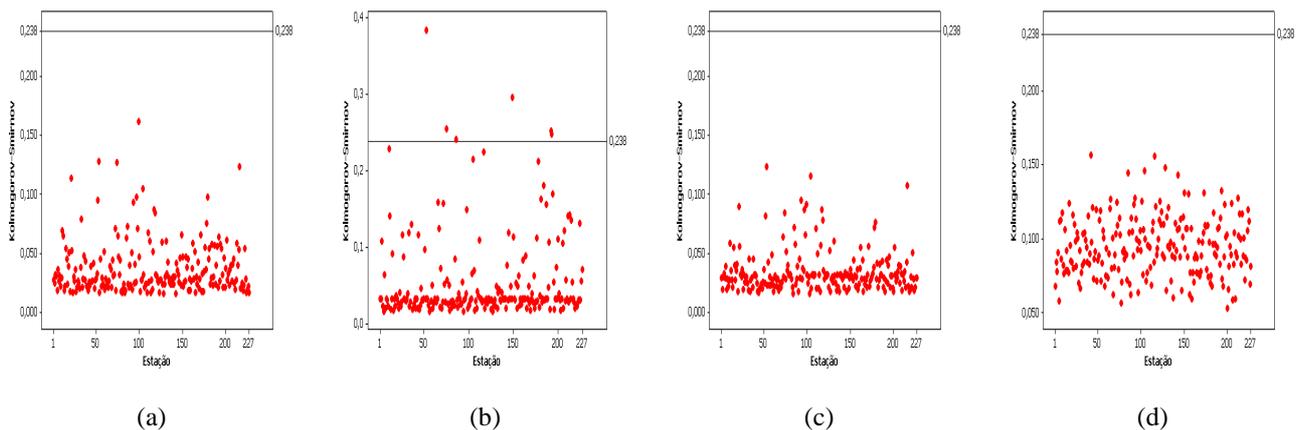


Figura 2 – Teste de aderência para as metodologias: (a) M, (b) MLH, (c) MVS, (d) ML.

Usando o critério da equação 3 foram escolhidas as metodologias de melhor ajuste para cada estação, conforme figura 3.

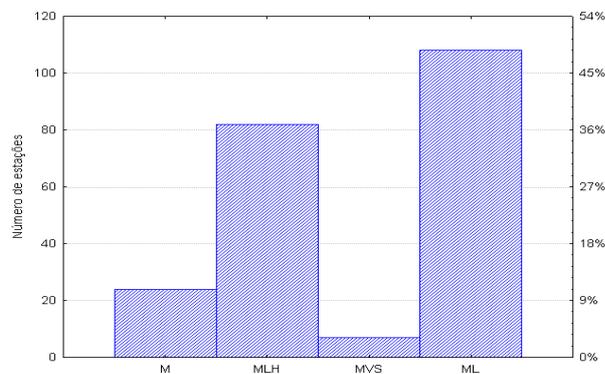


Figura 3 – Histograma de frequência da metodologia de estimativa de parâmetros de melhor ajuste

Observa-se na figura 3 que a metodologia MLH obteve melhor ajuste em 37% das estações e os ML em 50% das estações. Nota-se que as duas metodologias representam 87% das estações estudadas.

Observa-se na tabela 2 a metodologia dos MLH obteve o maior coeficiente de variação, além disso, seus valores variaram entre 0,0118 e 0,2888 (maior variação observada). A metodologia dos

ML obteve o menor coeficiente de variação e seus valores variaram entre 0,0097 e 0,0911 (menor variação observada). Portanto, foi escolhida a metodologia dos ML para comparar com as demais.

Tabela 2 – Estatística descritiva do critério de escolha

	Média	CV	Mínimo	Q1	Mediana	Q3	Máximo
M	0,0504	57,97	0,0095	0,0312	0,0442	0,0601	0,2073
MLH	0,0578	87,92	0,0118	0,0282	0,0413	0,0670	0,2888
MVS	0,0693	62,41	0,0142	0,0414	0,0542	0,0894	0,2239
ML	0,0381	45,77	0,0097	0,0248	0,0359	0,0467	0,0911

A metodologia dos ML foi comparada com as demais pelo coeficiente de correlação, índice de concordância, EMA e RSME (usando todas as estações que foram aprovadas pelo teste de aderência), conforme figura 4.

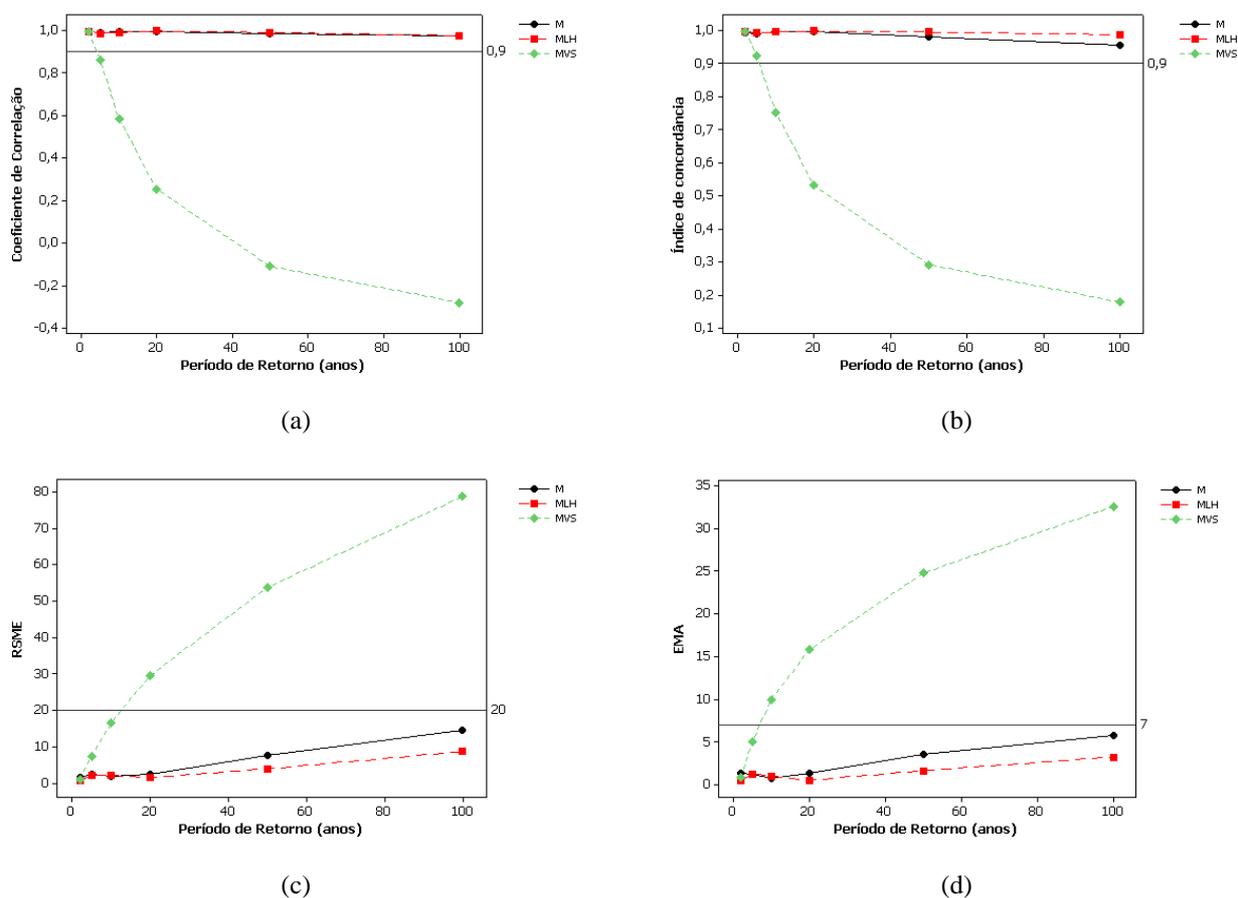


Figura 4 – Comparação entre ML e demais metodologias. (a) Correlação, (b) concordância, (c) RSME, (d) EMA.

Observa-se pela figura 4 que as metodologias dos MLH e M corroboram com os quantis obtidos pelos ML, porque os coeficientes de correlação e os índices de concordância estão acima de 0,9. O RSME dos MLH pode chegar até 9mm para período de retorno de 100 anos, já para os M

podem chegar até 15mm. O EMA dos MLH e M podem chegar até 6% para os períodos de retorno usuais.

Conforme a figura 4 mostra a metodologia MVS é a mais discrepante em relação aos ML. Observa-se que o aumento no período do retorno as diferenças ficam cada vez maiores. O RSME mostra que a diferença pode chegar até 80mm para períodos de retorno de 100 anos, o EMA mostra uma discrepância média de aproximadamente 35%.

CONCLUSÃO

Conclui-se pelos resultados obtidos neste trabalho que as metodologias dos momentos, momentos-L e momentos-LH produzem resultados semelhantes na estimativa de quantis. Os resultados mais discrepantes foram obtidos pela metodologia da verossimilhança, em relação aos momentos-L.

Portanto, na estimativa dos parâmetros da distribuição de probabilidade GEV deve-se optar pela metodologia dos momentos-L, pois ajustou-se aos dados de todas as estações estudadas e produz bom ajuste na grande maioria delas.

BIBLIOGRAFIA

CALGARO, M.; ROBAINA, A. D.; PEITER, M. X.; BERNARDON, T. Variação espaço-temporal dos parâmetros para a modelagem estocástica da precipitação pluvial diária no Rio Grande do Sul. Engenharia Agrícola, Jaboticabal, v.29, n.2, p.196-206, 2009.

DAMÉ, R. de C. F.; TEIXEIRA, C. F. A.; TERRA, V. S. S. Comparação de diferentes metodologias para estimativa de curvas intensidade-duração-frequência para pelotas – RS. Engenharia Agrícola, Jaboticabal, v.28, n.2, p.245-255, 2008.

MONTEBELLER, C. A.; CEDDIA, M. B.; CARVALHO, D. F. de; VIEIRA, S. R.; FRANCO, E. M. Variabilidade espacial do potencial erosivo das chuvas no estado do Rio de Janeiro. Engenharia Agrícola, Jaboticabal, v.27, n.2, p.426-435, 2007

NAGHETTINI, M.; PINTO, E. J. de A. Hidrologia Estatística. Belo Horizonte: CPRM, 2007. 552p.

PETERSEN-ØVERLEIR, A.; REITAN, T. Accounting for rating curve imprecision in flood frequency analysis using likelihood-based methods. Journal of Hydrology, Amsterdam, v.366, n., p.89-100, 2009.

QUEIROZ, M. M. F. de; CHAUDHRY, F. H. Análise de eventos hidrológicos extremos, usando-se a distribuição GEV e momentos LH. Revista Brasileira de Engenharia Agrícola e Ambiental, v.10, n.2, p.381–389, 2006.

SANTOS, G. G.; GRIEBELER, N. P.; OLIVEIRA, L. F. C. de. Chuvas intensas relacionadas à erosão hídrica. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.14, n.2, p.115–123, 2010.

WANG, Q. J. LH moments for statistical analysis of extreme events. *Water Resources Research*, Washington, v.33, n.12, p.2841- 2848, 1997.