**A DISTRIBUIÇÃO GENERALIZADA DE VALORES
EXTREMOS NO ESTUDO DA VELOCIDADE MÁXIMA DO
VENTO EM PIRACICABA, SP**

EZEQUIEL ABRAHAM LÓPEZ BAUTISTA

Dissertação apresentada à Escola Superior de
Agricultura "Luiz de Queiroz", Universidade de São
Paulo, para obtenção do título de Mestre em
Agronomia, Área de Concentração: Estatística e
Experimentação Agronômica.

P I R A C I C A B A

Estado de São Paulo - Brasil

Abril - 2002

**A DISTRIBUIÇÃO GENERALIZADA DE VALORES
EXTREMOS NO ESTUDO DA VELOCIDADE MÁXIMA DO
VENTO EM PIRACICABA, SP**

EZEQUIEL ABRAHAM LÓPEZ BAUTISTA

Engenheiro Agrônomo

Orientador: Prof. Dr. **SILVIO SANDOVAL ZOCCHI**

Dissertação apresentada à Escola Superior de
Agricultura "Luiz de Queiroz", Universidade de São
Paulo, para obtenção do título de Mestre em
Agronomia, Área de Concentração: Estatística e
Experimentação Agronômica.

P I R A C I C A B A

Estado de São Paulo - Brasil

Abril – 2002

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

López Bautista, Ezequiel Abraham

A distribuição generalizada de valores extremos no estudo da velocidade máxima do vento em Piracicaba, SP / Ezequiel Abraham López Bautista. - - Piracicaba, 2002.

47 p.

Dissertação (mestrado) - - Escola Superior de Agricultura Luiz de Queiroz, 2002.

Bibliografia.

1. Distribuição de extremos 2. Piracicaba 3. Velocidade 4. Vento I. Título

CDD 551.55

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

Aos meus pais:

Justiniano Rafael López Velásquez (in memorian)
Basilía Bautista,

ofereço.

Aos meus irmãos:

Carlos Enrique, Cristina, Celeste e Clara Luz,

com carinho.

À minha avó:

Concepción Vidalina Echeverría De León,

dedico com amor e gratidão.

AGRADECIMENTOS

Ao Prof. Dr. Silvio Sandoval Zocchi, pela orientação e confiança no desenvolvimento deste trabalho, contribuindo no meu amadurecimento profissional.

Ao Programa Estudante Convênio de Pós-Graduação (PEC-PG) da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro conferido para a realização deste curso.

Aos docentes do Curso de Estatística e Experimentação Agronômica da Escola Superior de Agricultura “Luiz de Queiroz”, pela transmissão dos conhecimentos.

Ao Prof. Dr. Luiz Roberto Angelocci pelo fornecimento dos dados.

Aos colegas de curso, Andréia, Cristiano, Denise, Geneville, Glaucy, Janaína, Mariana e Ramiro, pela amizade e o convívio.

Aos amigos, Adriano, Alexandre, Alexandre Jordão, André, Axel, Denis, Franzini, Juliana, Maurício, Paulo, Renato, Rose, Silvano e Telde, pelo companheirismo e por todos os bons momentos de convívio.

Ao Prof. Ary Dos Santos, pelos ensinamentos da língua portuguesa.

Aos docentes da Sub-Área de Métodos de Cuantificación e Investigación da Faculdade de Agronomia da Universidade de San Carlos de Guatemala, em especial ao M.C. Víctor Manuel Alvarez Cajas, M. Sc. Marino Barrientos e ao Eng. Agr. Byron González Ramírez, pelo apoio e incentivo dados para a realização dos meus estudos de Pós-Graduação.

SUMÁRIO

	Página
LISTA DE FIGURAS	vi
LISTA DE TABELAS	viii
RESUMO.....	x
SUMMARY.....	xii
1 INTRODUÇÃO.....	1
2 REVISÃO DE LITERATURA.....	3
3 MATERIAL E MÉTODOS.....	10
3.1 Material.....	10
3.2 Métodos.....	10
3.2.1 Análise exploratória.....	10
3.2.2 Teste de aleatoriedade.....	10
3.2.3 Estimação dos parâmetros da distribuição GVE	13
3.2.4 Seleção da distribuição de valores extremos	16
3.2.5 Obtenção da matriz de variâncias e covariâncias de $\hat{\epsilon}$	17
3.2.6 Diagnóstico do ajuste	17
3.2.7 Obtenção das probabilidades de obtenção de valores extremos de velocidade de vento acima de valores pré-estabelecidos	19
3.2.8 Estimação do período de retorno	19
3.2.9 Estimativas dos níveis de retorno da distribuição GVE	19
3.2.10 Obtenção dos intervalos de confiança para os níveis de retorno	20
4 RESULTADOS E DISCUSSÃO	25
5 CONCLUSÕES	40
REFERÊNCIAS BIBLIOGRÁFICAS	41

LISTA DE FIGURAS

Figura		Página
1	Função densidade de probabilidade da distribuição generalizada de valores extremos (GVE) para $\xi = -0,3$ (Weibull), $\xi \rightarrow 0$ (Gumbel) e $\xi = 0,3$ (Fréchet) com $m=10$ e $s=2,6$	5
2	Região de 95% de confiança para os parâmetros da distribuição de Gumbel	23
3	Gráfico de radar para representar as medidas de tendência central média e mediana, da variável velocidade máxima de vento (km h^{-1}) em cada um dos meses do ano	26
4	Gráficos de caixa (<i>box plot</i>) para a variável velocidade máxima de vento para cada um dos meses do ano	27
5	Gráficos quantil-quantil para diagnóstico do ajuste da distribuição de Gumbel aos dados de velocidade máxima de vento para os doze meses do ano	32
6	Gráficos de radar para representar a estimativa pontual do nível de retorno (km h^{-1}) para períodos de retorno de 5 e 10 anos, em cada um dos meses do ano e seus respectivos intervalos de confiança obtidos através dos métodos delta (a) e da estatística de razão de verossimilhança (b)	38

7	Gráficos de radar para representar a estimativa pontual do nível de retorno (km h ⁻¹) para períodos de retorno de 50 e 100 anos, em cada um dos meses do ano e seus respectivos intervalos de confiança obtidos através dos métodos delta (a) e da estatística de razão de verossimilhança (b)	39
---	---	----

LISTA DE TABELAS

Tabela	Página
1 Velocidades máximas de vento mensais (km h^{-1}) registradas na estação agrometeorológica da ESALQ, Piracicaba, nos períodos 1956 a 1971 e de 1974 a 2000	11
2 Estatísticas descritivas da variável aleatória velocidade máxima mensal (km h^{-1}) de vento, nos períodos de 1956 a 1971 e de 1974 a 2000, em Piracicaba, SP	25
3 Números totais de valores menores (n_1) e maiores (n_2) do que a mediana, número total de seqüências (v) e valores críticos ($v_{0,05;n1,n2}$) segundo cada mês do ano, para uso no teste de chorrilho	28
4 Estimativas dos parâmetros da distribuição generalizada de valores extremos e respectivas variâncias e covariâncias estimadas, para cada um dos meses do ano	29
5 Intervalos de 95 % de confiança do parâmetro de forma (ξ) e valores da estatística de razão de verossimilhança modificada (T_{LR}^*) para cada um dos meses do ano	30
6 Estimativas dos parâmetros me s da distribuição de Gumbel e correspondentes variâncias e covariâncias estimadas	31

7	Resultados do teste de Kolmogorov-Smirnov para verificação da qualidade do ajuste da distribuição de Gumbel aos dados de velocidade máxima de ventos (km h^{-1}).....	33
8	Probabilidades de ocorrência de rajadas máximas mensais de vento com velocidade acima de 40, 50, 60, 70, 80, 90 e 100 km h^{-1} , a 10 m acima do nível do solo, para os 12 meses do ano, em Piracicaba, SP	34
9	Períodos de retorno estimados para os maiores valores de velocidade máxima de vento (km h^{-1}) registrados em cada um dos meses do ano, nos períodos de 1956 a 1971 e de 1974 a 2000	35
10	Níveis de retorno (km h^{-1}) estimados e limites inferior (LI) e superior (LS) de seus respectivos intervalos de 95 % de confiança, para os períodos de retorno 5, 10, 50 e 100 anos, obtidos através do método delta	36
11	Níveis de retorno (km h^{-1}) estimados e limites inferior (LI) e superior (LS) de seus respectivos intervalos de 95 % de confiança, para os períodos de retorno 5, 10, 50 e 100 anos obtidos através da estatística de razão de verossimilhança....	37

A DISTRIBUIÇÃO GENERALIZADA DE VALORES EXTREMOS NO ESTUDO DA VELOCIDADE MÁXIMA DO VENTO EM PIRACICABA, SP

Autor: EZEQUIEL ABRAHAM LÓPEZ BAUTISTA

Orientador: Prof. Dr. SILVIO SANDOVAL ZOCCHI

RESUMO

A teoria dos valores extremos desempenha um papel fundamental na modelagem de eventos associados a probabilidades muito pequenas ou eventos raros. Os modelos probabilísticos baseados nesta teoria visam prever, a partir de um conjunto de valores máximos de um processo ambiental registrado num período relativamente curto (30 anos, por exemplo), os valores máximos esperados num período maior de tempo (50, 100 ou mais anos), que para o caso específico dos ventos, são de grande utilidade no planejamento de estruturas civis. Este trabalho consistiu no ajuste da distribuição generalizada de valores extremos (GVE), que inclui como casos particulares, as distribuições Gumbel, Fréchet e de Weibull (definidas na teoria dos valores extremos), aos dados de velocidade máxima mensal de ventos registrados durante um período de 43 anos (1956 a 1971 e 1974 a 2000) na estação agrometeorológica da Escola Superior de Agricultura “Luiz de Queiroz” em Piracicaba, Estado de São Paulo. Para a estimação dos parâmetros dessa distribuição, foi utilizado o método da máxima verossimilhança. O ajuste aos dados foi avaliado através dos gráficos quantil-quantil e do teste de Kolmogorov-Smirnov. A partir do ajuste inicial da distribuição GVE, a distribuição de Gumbel demonstrou

ser a mais adequada para modelar os dados de velocidade máxima de vento em todos os meses do ano. Observou-se também, que os meses de setembro a dezembro apresentaram as maiores velocidades máximas de vento. Ventos com velocidades acima de 60 km h^{-1} , considerados como muito fortes, também apresentaram-se neste período do ano. Por último, foram obtidos os níveis de retorno para os períodos de retorno 5, 10, 50 e 100 anos, e construídos seus respectivos intervalos de 95% de confiança, através dos métodos delta e da estatística de razão de verossimilhança.

THE GENERALIZED EXTREME VALUE DISTRIBUTION TO STUDY MAXIMUM WIND SPEED IN PIRACICABA, SP

Author: EZEQUIEL ABRAHAM LÓPEZ BAUTISTA

Adviser: Prof. Dr. SILVIO SANDOVAL ZOCCHI

SUMMARY

The extreme value theory plays a fundamental role in modeling of events associated to very small probabilities or rare events. The aim of the probabilistic models based on this theory is to predict, from a set of maximum values of an environmental process recorded on a relatively short period (e.g. 30 years), the expected maximum values in a greater period (50, 100 or more years). For the specific case of wind, these values are very useful for the planing and development of civil structures. This work is concerned with the fitting of the generalized extreme value (GEV) distribution, which includes as specific cases, the Gumbel, Fréchet and Weibull distributions (defined in the extreme value theory), to the maximum wind speeds recorded monthly during a 43 years period (1956 to 1971 and 1974 to 2000) in the agrometeorological station of the College of Agriculture “Luiz de Queiroz”, in Piracicaba, SP (Brazil). For the estimation of parameters of the GVE distribution the method of the maximum likelihood was used. The fitting to the data was evaluated through the quantil-quantil graph and the Kolmogorov-Smirnov test. From the initial fitting of the GEV distribution, we concluded that the Gumbel distribution was the most suitable to model the maximum wind speed for all months. It was observed that the September to December period showed the highest values of

maximum wind speed. This period also showed winds with speeds above 60 km h^{-1} , considered as very strong. Finally, we obtained the return levels for the return periods of 5, 10, 50 and 100 years, and we constructed their respective 95% confidence intervals, through the delta and the likelihood ratio statistics methods.

1 INTRODUÇÃO

O vento tem importância muito grande na atividade humana. Na agricultura, por exemplo, está diretamente associado ao desenvolvimento das plantas, ao facilitar as trocas de calor, de dióxido de carbono e de vapor d'água entre a atmosfera e a vegetação, além de ajudar no processo de polinização das flores e também poder ser utilizado como fonte de energia (energia eólica). Entretanto, quando se registram ventos de velocidades elevadas, normalmente de curta duração, os seus efeitos passam, geralmente, a ser danosos, provocando o estímulo excessivo à evapotranspiração, o acamamento das plantas, a queda de flores e frutos, a quebra de galhos e arrancamento de plantas, causando a erosão dos solos e a deformação da paisagem.

Nas construções, por sua vez, ventos fortes causam o destelhamento de prédios, a queda de redes de transmissão de energia elétrica e a vibração de estruturas e equipamentos. A importância dos efeitos deste fenômeno meteorológico está, portanto, intimamente ligada ao desenvolvimento da tecnologia dos materiais e da engenharia estrutural. Segundo Blessmann (1986), o vento passou a ser um grande problema à medida que as construções tornaram-se mais altas e foi utilizado material em quantidade cada vez menor. De forma geral, a previsão probabilística da ocorrência de ventos extremos é de vital importância para o planejamento das atividades sujeitas a seus efeitos adversos, e uma forma de modelar esses eventos, é utilizar a teoria dos valores extremos proposta por Fisher & Tippett (1928). Segundo esta teoria, existem três tipos de distribuições assintóticas de valores extremos, a tipo I ou de Gumbel, a tipo II ou de Fréchet e a tipo III ou de Weibull.

Essas distribuições são freqüentemente utilizadas para estimar probabilidades de ocorrência de ventos acima de valores pré-estabelecidos ou para se prever o valor máximo de velocidade de vento em determinados períodos de tempo. Segundo Raynal (1997), um problema que surge na prática, é o da escolha da distribuição de valores extremos mais adequada para uma amostra de dados, em particular. Como alternativa este autor sugere a utilização da distribuição generalizada de valores extremos, desenvolvida por Jenkinson (1955), que pode ser considerada

uma família de distribuições, que inclui os três tipos de distribuições assintóticas de valores extremos como casos particulares.

O presente trabalho foi desenvolvido com o objetivo principal de apresentar e implementar a metodologia para ajustar a distribuição generalizada de valores extremos aos dados de velocidade máxima mensal de ventos em Piracicaba, e obter a probabilidade de ocorrência mensal de valores extremos de velocidades de vento acima de 40, 50, 60, 70, 80, 90 e 100 km h⁻¹, estimar o período de retorno para o maior valor de velocidade máxima de vento registrado em cada um dos meses do ano, e determinar as velocidades máximas para períodos de retorno de 5, 10, 50 e 100 anos (e seus respectivos intervalos de confiança).

2 REVISÃO DE LITERATURA

Os fundamentos da teoria dos valores extremos foram inicialmente expostos por Fisher & Tippett (1928), que definiram os três tipos possíveis (I, II e III) de distribuições assintóticas dos 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 Gumbel¹, citado por Jenkinson (1955), cuja metodologia tem sido frequentemente aplicada à máxima anual de séries de dados referentes a vazões de rios. Posteriormente, Barricelli² e Brooks & Carruthers³, citados por Jenkinson (1955), perceberam que ao utilizar essa metodologia, os valores máximos de temperatura e precipitação pluvial previstos para períodos longos, eram superestimados e subestimados, respectivamente, e propuseram algumas modificações para corrigir este problema.

Outras contribuições importantes para o estudo de valores extremos foram dadas por Gnedenko (1943), que mostrou as condições necessárias e suficientes para a existência das distribuições assintóticas dos valores extremos e determinou que as caudas dessas distribuições, ou seja, a parte que trata dos valores máximos ou mínimos menos frequentes, podem ser modeladas por alguns tipos de distribuições contínuas. Por exemplo, as caudas da distribuição de Gumbel correspondem às distribuições exponencial, gama, normal ou log-normal, as da distribuição de Fréchet seguem uma distribuição de Cauchy, Pareto ou t de Student e as da distribuição de Weibull seguem uma distribuição uniforme.

Um problema que surgiu desde que foram propostas as distribuições de valores extremos foi identificar o tipo de distribuição mais adequada para uma determinada amostra de dados. Para solucionar este problema, diversos procedimentos foram propostos, como por exemplo, os

¹ GUMBEL, E.J. Les valeurs extremes des distributions statistiques. *Annales de l'institute Henri Poincaré*, v.5, p.115-158, 1935.

² BARRICELLI, N.A. Les plus grands et les plus petits maxima ou minima annuels d'une variable climatique. *Archive for Mathematik og Naturvidenskab* v.46, n.6, 1943.

³ BROOKS, C.E.P.; CARRUTHERS, N. *Handbook of statistical methods in meteorology*. London: HMSO, 1953. p.131-134.

de Van Monfort (1978), Tiago de Oliveira (1981) e Hosking et al. (1985b), e o da curvatura, desenvolvido por Castillo (1988). Para a implementação destes métodos, embora não seja necessário estimar os parâmetros dos três tipos de distribuições de valores extremos, proporcionam uma solução apenas aproximada ao problema da identificação (Raynal, 1997). Este autor sugere que a forma mais correta de fazer essa identificação seja através da estimação dos parâmetros da distribuição generalizada de valores extremos (GVE) desenvolvida por Jenkinson (1955). Essa distribuição apresenta como casos particulares, os três tipos de distribuições de valores extremos, e tem função de distribuição acumulada de probabilidade dada por:

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

definida em, $-\infty < x < m - s/x$ para $x < 0$, $-\infty < x < +\infty$ para x tendendo a zero, $m - s/x < x < +\infty$ para $x > 0$, sendo m , s e x os parâmetros de locação, escala e de forma respectivamente, com $s > 0$.

As distribuições de valores extremos de Fréchet e de Weibull correspondem aos casos particulares de (1) em que $x > 0$ e $x < 0$, respectivamente. Como limite de $F(x)$ com x tendendo a zero tem-se que:

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

que é a função de distribuição acumulada de Gumbel com parâmetros de locação e de escala m e s , respectivamente, com $s > 0$.

Derivando-se (1) em relação a x , obtém-se a função densidade de probabilidade da distribuição GVE, dada por:

$$f(x) = \frac{1}{s} \left[1 + x \left(\frac{x - m}{s} \right) \right]^{-\left(\frac{1+x}{x}\right)} \exp \left\{ - \left[1 + x \left(\frac{x - m}{s} \right) \right]^{-\frac{1}{x}} \right\}, \quad (3)$$

definida em, $-\infty < x < m - s/x$ para $x < 0$ e $m - s/x < x < +\infty$ para $x > 0$, cujo limite para x tendendo a zero, é:

$$f(x) = \frac{1}{s} \left\{ \exp \left(- \frac{x - m}{s} \right) \exp \left[- \exp \left(- \frac{x - m}{s} \right) \right] \right\}, \quad (4)$$

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

A Figura 1 apresenta os gráficos da função densidade de probabilidade para $x = -0,30$ (Weibull), x tendendo a zero (Gumbel) e $x = 0,30$ (Fréchet), com $m = 10$ e $s = 2,6$, a partir de onde pode se observar que o parâmetro x 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 parâmetro ξ , mais lenta é essa velocidade (Coronado, 2000).

Figura 1 –Função densidade de probabilidade da distribuição generalizada de valores extremos (GVE) para $x = -0,3$ (Weibull), $x \rightarrow 0$ (Gumbel) e $x = 0,3$ (Fréchet), com $m=10$ e $s=2,6$.

Através da Figura 1 e da função (3) observa-se, ainda, que a distribuição de Fréchet corresponde a um modelo com cauda inferior finita ($x > 1,33$) e cauda superior infinita. Por outro lado, no caso da distribuição de Weibull, a cauda superior é finita ($x < 18,67$), motivo pelo qual, segundo Holmes & Moriarty (1999), esta distribuição é mais apropriada para estudar variáveis que têm limitações em magnitude por razões geofísicas como o caso de variáveis associadas a certos fenômenos naturais, como por exemplo, terremotos e ventos.

Para fazer inferências sobre os parâmetros da distribuição GVE, inúmeras sugestões foram propostas, entre elas, técnicas gráficas, estimadores baseados no método dos momentos e no método da máxima verossimilhança. Aspectos computacionais da estimação dos parâmetros dessa distribuição através do método da máxima verossimilhança foram desenvolvidos por

Jenkinson⁴, citado por Otten & Montfort (1980), por Prescott & Walden (1980, 1983), Hosking (1985) e por Martins & Stedinger (2000).

Por outro lado, para obter as estimativas de máxima verossimilhança dos parâmetros da distribuição GVE, Prescott & Walden (1983) e Coles (1999) recomendaram o uso do método de Newton-Raphson. Hosking (1985), por sua vez, modificou esse método para melhorar a velocidade e a taxa de convergência, e quatro anos mais tarde, MacLeod (1989) alterou-o novamente a fim de prevenir alguns erros na hora de executá-lo, e otimizar a inversão da matriz Hessiana.

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 GVE dependerem de seus parâmetros, as condições de regularidade para a estimação pelo método da máxima verossimilhança não são necessariamente satisfeitas, tal como acontece nas distribuições lognormal, Weibull e gama de três parâmetros. Assim, através de um estudo cuidadoso, obteve os seguintes resultados:

- i) quando $x > -0,5$, os estimadores de máxima verossimilhança são completamente regulares,
- ii) quando $-1 < x < -0,5$, os estimadores de máxima verossimilhança existem, mas são não regulares,
- iii) quando $x < -1$, os estimadores de máxima verossimilhança não existem.

Estudos realizados por Hosking et al. (1985b) sobre estimação dos parâmetros da distribuição GVE através do método da máxima verossimilhança, utilizando simulação computacional de amostras de dados, revelaram que podem surgir problemas de não convergência no processo iterativo de Newton-Raphson devido ao não cumprimento das condições de regularidade. Situações similares foram reportadas por Martins & Stedinger (2000).

Apesar dos problemas que podem ocorrer quando $x < -0,5$, esta situação, segundo Smith (1985), é extremamente rara para dados ambientais, e correspondem a distribuições com cauda superior muito curta e finita. De uma forma geral, ao se trabalhar com dados reais, os

⁴ JENKINSON, A.F. Statistics of extremes, in estimation of maximum floods. World Meteorology Office, nota técnica n.98, cap. 5, p.183-227. Geneva, Suíça. 1969.

valores do parâmetro x geralmente se encontram no intervalo $(-0,5;0,5)$ (Hosking et al.,1985b). Brabson & Palutikof (2000), por sua vez concluíram, a partir de um estudo com simulações, que $x \in (-0,25; 0,25)$, situação também bastante freqüente, garante a eficiência do processo de obtenção de estimativas de máxima verossimilhança dos parâmetros. Assim, na prática, a estimação por máxima verossimilhança geralmente é válida e regular como um procedimento para a inferência.

Com relação a outros métodos de estimação, Hosking et al. (1985b) mostraram que os estimadores dos parâmetros da distribuição GVE, obtidos através do método dos momentos de probabilidade ponderada ou do método de momentos L, eram preferíveis em comparação com os obtidos através do método da máxima verossimilhança, em termos de viés e variância para amostras cujos tamanhos variaram entre 15 e 100. Em contrapartida, Smith (2001) comentou que nenhum destes últimos métodos de estimação desenvolvidos tem o poder e generalidade que possui o método da máxima verossimilhança.

Quanto às aplicações, a distribuição GVE tem sido utilizada em muitas áreas diferentes, sendo, uma das principais, a engenharia estrutural, onde o objetivo principal é projetar estruturas que resistam aos níveis mais extremos de certos processos ambientais. Segundo Martins & Stedinger (2000), essa distribuição foi recomendada para a análise de freqüências de enchentes no Reino Unido pelo *Natural Environment Research Council* (NERC), em 1975, e para freqüências de chuvas nos Estados Unidos por Willeke et al.⁵, em 1995.

Essa distribuição também foi utilizada para a análise de freqüências regionais de vazões, nos estudos realizados por Hosking et al. (1985a), Wallis & Wood (1985), Lettenmaier et al. (1987), Hosking & Wallis (1988), Chowdhury et al. (1991), Madsen et al. (1997) e Beirlant & Matthys (2001).

Para estudos sobre velocidades máximas de vento, tradicionalmente tem sido utilizada a distribuição de Gumbel (também conhecida como de Fisher-Tippett), caso particular da GVE, como pode ser observado nos trabalhos de Simiu & Filliben (1976), Ross (1987), Grigoriu (1984), Gusella (1991), Abild et al. (1992) e Walshaw (1994). No entanto, Simiu & Heckert (1996) e Holmes & Moriarty (1999) criticaram o uso dessa distribuição, argumentando que para períodos de retorno altos, há uma tendência a predizer valores ilimitados e irreais de velocidades

⁵ WILLEKE, G.E.; HOSKING, J.R.M.; WALLIS, J.R.; GUTTMAN, N.B. **The national drought atlas**. Fort Belvoir : Institute for Water Resources and U.S. Army Corps of Engenniers, 1995.

máximas de vento, e concluíram que a distribuição de Weibull é a mais apropriada para modelar velocidades máximas de vento de origem extratropical (excluindo áreas de tornados). Galambos & Macri (1999), por outro lado, fizeram algumas críticas ao trabalho de Simiu & Heckert (1996), concluindo que para alguns locais a distribuição de Gumbel é o modelo mais adequado e realista para estudar o comportamento das velocidades extremas de vento.

No Brasil, Ortolani (1986) verificou que as velocidades instantâneas máximas de vento têm uma variação significativa com a época do ano, a latitude e a altitude no Estado de São Paulo. Por sua vez, Wagner et al. (1987) ajustaram a distribuição de Gumbel às velocidades máximas diárias de vento em Londrina e Ponta Grossa, no Paraná, e determinaram as velocidades máximas para períodos de retorno de 2, 10 e 20 anos. Na região de Botucatu, Estado de São Paulo, Martins (1993) estudou o regime de ventos em cinco locais e concluiu que para essa região, os meses de maio e outubro apresentaram a menor e a maior intensidades de velocidade de vento, respectivamente. Em outro estudo, Camargo et al. (1994) ajustaram as distribuições normal e gama aos dados de rajadas máximas diárias de Campinas, Estado de São Paulo, para cada um dos meses do ano e obtiveram os valores das probabilidades mensais de ocorrência de rajadas acima de 10, 20, 30, 40 e 50 km h⁻¹.

Em Piracicaba, Estado de São Paulo, Angelocci et al (1995) ajustaram as distribuições de probabilidades normal, gama, de Weibull e de Gumbel para ventos máximos em cada mês do ano. Baseando-se em uma série de registros de 20 anos (de 1975 a 1994) e no teste de Kolmogorov-Smirnov (Siegel, 1956), recomendaram as distribuições normal, gama e de Weibull para a estimativa da probabilidade de ocorrência de rajadas de ventos na localidade em qualquer mês, comentando que o ajuste da distribuição de Gumbel não foi satisfatório para os meses de fevereiro, abril e novembro. Utilizando a distribuição normal, estimaram as probabilidades de ocorrência de vento com velocidades iguais ou maiores do que 40, 50, 60, 70, 80, 90 e 100 km h⁻¹.

Nas localidades de Ribeirão Preto e Jaú, ambas no Estado de São Paulo, Pires et al. (1999) utilizaram registros disponíveis das velocidades máximas de rajadas de vento em 25 anos, e ajustaram as distribuições normal, gama e de Weibull em cada mês do ano. Concluíram que em todos os meses ocorreu um bom ajuste de todas essas distribuições.

As distribuições de valores extremos também têm sido utilizadas na estimativa das precipitações máximas prováveis, como pode ser observado em estudos realizados por Occhipinti & Santos (1965) na cidade de São Paulo, por Banzatto (1973), no Estado de São

Paulo, e por Krishnam & Kushwaha (1978), na Índia. Para estudos sobre duração das estiagens mais longas do ano no Triângulo Mineiro, Silva & Assis (1979) utilizaram a distribuição de Gumbel. Carvalho (1982) também ajustou essa distribuição de valores extremos aos dados de estiagem nas localidades de Avaré, Itapeva e Presidente Prudente, no Estado de São Paulo, estimou a duração máxima provável (em número de dias secos consecutivos) da estiagem mais longa do ano e a probabilidade dessa estimativa não ser ultrapassada.

Exemplos de aplicação da distribuição GVE em estudos sobre meio ambiente foram apresentados por Thas et al. (1997), que fizeram uma revisão dos principais conceitos de estatísticas de valores extremos, e ressaltaram a importância da aplicação da distribuição GVE para modelagem estatística de concentrações de poluentes em rios, que serviram de base para o estabelecimento de planos de manejo da qualidade d'água.

Piegorsch et al (1998), por sua vez, aplicaram a teoria dos valores extremos ao estudo dos níveis extremos de ozônio na troposfera, que é a camada da atmosfera mais próxima da superfície terrestre, onde ocorrem os principais fenômenos meteorológicos, como por exemplo, processos de condensação de vapor d'água, de precipitação pluvial, movimentação de massas de ar etc. Níveis padrão de ozônio são baseados tipicamente no número de valores que excedem alguma medida de severidade de ozônio (por exemplo, valor máximo a uma determinada hora, ou a média de cada 8 horas) em relação a algum patamar específico, por exemplo, 120 ou 80 partes por bilhão (ppb). O principal interesse, neste caso, é o de monitorar se as taxas deste nível estão crescendo ou decrescendo com o tempo.

Segundo Sharma et al. (1999) a teoria dos valores extremos oferece uma alternativa prática para o estudo de concentrações extremas de poluentes em centros urbanos. Em um estudo realizado por Medici et al. (2000), foram analisados valores extremos das médias diárias de concentração de monóxido de carbono (CO), na cidade de São Paulo, entre janeiro de 1994 e dezembro de 1997, usando as distribuições generalizada de Pareto e generalizada de valores extremos. Estes autores concluíram que em um período de retorno de seis meses pode-se esperar uma concentração de CO cujo valor é aproximadamente o dobro do padrão nacional de qualidade de ar imposto para este poluente. Em um estudo similar, Küchenhoff & Thamerus (1996) analisaram valores extremos de médias diárias de concentrações de ozônio e dióxido de carbono registrados em dois centros de monitoramento na cidade de Munich (Alemanha), durante janeiro de 1980 e outubro de 1992.

Outras áreas de aplicação envolvem modelos financeiros e de seguros (Longin, 2000), modelos de telecomunicações, estudo da resistência de materiais, resistência à corrosão etc.

3 MATERIAL E MÉTODOS

3.1 Material

Os dados de velocidades máximas mensais de vento (expressos em km h^{-1}) a 10 m acima do nível do solo foram obtidos nos períodos de 1956 a 1971 e de 1974 a 2000, a partir de registros de anemógrafo do tipo universal, marca Fuess, localizado na estação agrometeorológica da Escola Superior de Agricultura “Luiz de Queiroz”, em Piracicaba (latitude $22^{\circ}42'30''\text{S}$, longitude $47^{\circ}30'00''\text{W}$, e altitude 545 m). De cada mês foi selecionado o valor de velocidade máxima, para formar a série de valores máximos, apresentada na Tabela 1. Foram utilizados registros de 42 anos para os meses de janeiro e setembro, e 43 anos para os restantes meses do ano, devido à indisponibilidade de algumas observações.

3.2 Métodos

Os métodos adotados para o estudo da série de valores máximos de velocidade de ventos, são apresentados a seguir.

3.2.1 Análise exploratória

Consistiu na estimação de medidas de tendência central (média e mediana), de dispersão (variância, desvio padrão, coeficiente de variação e amplitude interquartílica), de assimetria e construção de um gráfico de caixas (*box plot*), para a variável aleatória velocidade máxima de vento, em cada mês do ano.

3.2.2 Teste de aleatoriedade

Uma das pressuposições para que se possa utilizar a metodologia para a estimação dos parâmetros da distribuição GVE (a ser apresentada no item 3.2.3), é que as observações para cada mês, a longo do tempo sejam independentes. Assim, para testar a hipótese de que a série de

Tabela 1. Velocidades máximas de vento mensais (km h^{-1}) registradas na estação agrometeorológica da ESALQ, Piracicaba, nos períodos de 1956 a 1971 e de 1974 a 2000.

Ano	Janeiro	Fevereiro	Março	Abril	Mai	Junho	Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
1956	63,00	54,00	46,80	57,60	64,44	39,60	64,80	72,00	97,20	60,12	50,40	53,64
1957	61,20	65,16	87,12	46,80	45,00	82,80	54,00	55,08	102,96	57,60	83,16	75,60
1958	57,96	62,28	53,28	43,20	50,04	42,48	50,04	42,48	90,00	58,32	105,12	61,92
1959	63,72	61,20	55,80	80,28	54,00	68,76	51,12	72,72	46,80	49,68	104,40	65,16
1960	65,16	72,72	64,80	41,76	75,96	108,00	50,40	49,32	47,16	81,00	70,92	58,68
1961	79,92	65,16	61,56	66,60	48,60	82,08	36,00	57,24	54,00	86,40	69,12	68,40
1962	49,32	68,40	66,60	60,84	43,20	45,00	39,60	40,32	72,00	67,68	75,60	64,80
1963	68,40	57,60	91,08	46,08	46,44	53,64	42,84	46,80	54,00	70,20	65,16	79,92
1964	57,60	85,68	57,96	75,60	43,20	61,20	61,20	55,08	54,00	64,44	68,40	58,68
1965	59,40	86,40	70,20	57,60	50,40	36,00	72,00	53,28	56,52	50,40	47,16	50,40
1966	68,40	61,20	50,40	76,68	72,00	43,20	79,20	61,56	54,00	101,16	59,40	67,68
1967	87,84	64,08	61,20	54,00	39,96	47,16	45,00	50,04	99,00	107,64	63,00	55,08
1968	79,20	48,60	61,20	53,64	36,36	53,64	56,52	61,20	58,68	92,52	140,04	72,00
1969	71,64	84,60	97,20	59,40	56,88	71,64	75,60	79,92	86,40	101,52	64,80	73,80
1970	54,00	53,64	57,60	54,00	42,12	40,32	40,32	49,68	57,60	93,96	57,60	86,40
1971	90,36	61,20	74,88	79,92	79,56	46,80	65,16	46,80		76,68	72,00	48,96
1974		77,40	46,80	43,20	43,20	82,80	36,00	64,80	79,20	82,80	88,20	66,60
1975	75,60	79,20	86,40	57,60	50,40	46,80	90,00	50,40	64,80	61,20	82,80	81,00
1976	72,72	76,32	68,40	61,92	79,20	72,00	61,20	50,40	63,00	79,20	82,80	82,80
1977	70,20	64,80	59,40	52,20	43,20	57,60	54,00	61,92	99,00	57,60	79,20	72,00
1978	81,00	57,60	39,60	52,20	79,20	61,20	72,00	57,60	73,80	93,60	57,60	72,00
1979	50,40	73,80	75,60	90,00	90,00	46,80	57,60	79,20	63,00	77,40	81,00	57,60
1980	72,00	63,36	51,12	59,40	36,00	64,80	51,12	54,72	84,96	90,00	77,40	78,12
1981	57,96	82,08	74,16	54,36	74,88	49,32	54,72	50,76	75,60	61,56	63,72	61,56
1982	60,48	90,36	59,76	47,16	93,60	46,08	54,00	68,40	72,36	86,76	75,24	68,40
1983	86,40	86,76	83,88	61,20	100,80	54,00	68,40	52,56	93,60	86,40	64,80	68,04
1984	92,16	64,08	77,40	48,24	61,92	51,48	51,12	54,00	75,96	72,00	72,36	99,00
1985	61,20	69,84	65,52	79,20	60,48	46,80	51,84	72,72	79,56	70,92	66,60	72,00
1986	65,88	79,20	68,40	47,88	65,16	45,72	47,52	75,24	57,24	46,44	54,72	114,48
1987	72,00	64,44	72,72	79,20	83,16	69,48	61,20	53,64	96,48	65,16	58,32	68,76
1988	94,32	71,28	50,76	52,92	59,76	58,32	51,12	48,60	54,00	72,72	60,84	55,44
1989	69,12	61,20	65,16	97,20	56,16	71,64	64,08	81,00	78,84	61,20	74,52	74,88
1990	68,40	60,48	68,76	43,92	57,96	73,80	79,20	75,60	92,52	76,32	117,72	73,44
1991	56,52	57,24	44,64	43,56	47,52	58,68	37,80	46,80	66,24	74,88	46,08	104,76
1992	61,56	75,96	57,24	43,56	73,80	46,80	70,20	42,12	52,20	57,24	65,52	54,00
1993	78,84	58,68	70,92	78,12	71,28	41,04	44,28	36,00	70,20	52,56	90,00	73,44
1994	46,08	71,28	44,28	52,56	40,32	48,60	51,84	42,12	61,56	75,60	61,92	65,88
1995	54,00	53,28	64,80	38,52	47,52	81,36	45,72	37,08	69,48	126,72	79,92	66,60
1996	65,52	99,36	61,20	43,92	48,96	61,92	51,12	79,20	76,32	75,24	72,72	89,64
1997	68,40	52,20	61,56	44,28	42,12	43,92	42,48	76,32	65,52	68,40	55,44	63,00
1998	100,08	60,12	70,56	51,48	55,44	53,64	41,40	64,08	49,68	73,08	65,52	73,80
1999	64,08	70,20	39,96	66,96	75,96	40,32	36,00	53,28	61,20	70,92	56,88	82,08
2000	69,12	50,76	54,00	46,08	47,16	57,60	50,76	60,84	65,52	73,44	78,12	58,32

n observações amostrais $(x_1, \dots, x_n)$, para um certo mês, é aleatória, foi utilizado o teste de chorrilho (*run test*), descrito em Siegel (1956), e que segundo Díaz (1999), é de utilidade na detecção de desvios na aleatoriedade de uma sequência de medições quantitativas no tempo, ocasionados por tendência ou periodicidade.

Como passo inicial deste teste, obtém-se o valor da mediana, M_d , dada por:

$$M_d = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{se } n \text{ for ímpar,} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & \text{se } n \text{ for par,} \end{cases}$$

sendo $x_{(\cdot)}$ a $(\cdot)$ –ésima observação da série de dados colocados em ordem ascendente. Em seguida, obtém-se, para os dados sem ordenar, os valores da variável indicadora A_i ($i=1,2,\dots,n$), definida por:

$$A_i = \begin{cases} 1 & \text{se } x_i > M_d \\ 0 & \text{se } x_i < M_d, \end{cases}$$

omitindo os casos em que $x_i = M_d$. Desta forma, são geradas seqüências $(A_1, A_2, \dots, A_n)$ de zeros e uns, para as quais testa-se sua aleatoriedade.

Seja V a variável aleatória número total de seqüências de zeros e uns na série $(A_1, A_2, \dots, A_n)$ e v seu valor observado, n_1 o número de valores de A_i ($A_1, A_2, \dots, A_n$) iguais a 1 e n_2 o número de valores de A_i ($A_1, A_2, \dots, A_n$) iguais a 0. Sob a hipótese de que a seqüência é aleatória, Zar (1999) construiu uma tabela, utilizando os procedimentos descritos por Eisenhart & Swed (1943) e Browlee (1965), que contém os pares de valores críticos $(v_{a,n1,n2})$, para $n_1 \leq 30$ e $n_2 \leq 30$ e um certo nível de significância α . O par de valores críticos define os limites da região de não rejeição da hipótese de aleatoriedade para um determinado número de seqüências (v). Assim, rejeita-se esta hipótese se os valores de v estiverem fora desse intervalo.

3.2.3 Estimação dos parâmetros da distribuição GVE

Seguindo as recomendações de Smith (1985), foi utilizado o método da máxima verossimilhança, descrito a seguir.

Seja $X_1, X_2, \dots, X_n$ uma série de n variáveis aleatórias independentes e identicamente distribuídas, com distribuição GVE e $(x_1, x_2, \dots, x_n)$ uma série de observações. Supondo que há independência entre as observações, obtém-se a função de verossimilhança,

$$L(\mathbf{\hat{\theta}}) = L(\mathbf{m}, \mathbf{s}, \mathbf{x}) = \prod_{i=1}^n f(x_i) = \frac{1}{\mathbf{s}^n} \prod_{i=1}^n \left\{ \left[1 + \mathbf{x} \left(\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) \right]^{\left(\frac{1+\mathbf{x}}{\mathbf{x}} \right)} \right\} \exp \left\{ \sum_{i=1}^n \left\{ - \left[1 + \mathbf{x} \left(\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) \right]^{\frac{1}{\mathbf{x}}} \right\} \right\},$$

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

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

$$\begin{aligned} l(\mathbf{m}, \mathbf{s}, \mathbf{x}) &= \ln [L(\mathbf{m}, \mathbf{s}, \mathbf{x})] \\ &= -n \ln \mathbf{s} - \left(\frac{1+\mathbf{x}}{\mathbf{x}} \right) \sum_{i=1}^n \ln \left[1 + \mathbf{x} \left(\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) \right] - \sum_{i=1}^n \left[1 + \mathbf{x} \left(\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) \right]^{\frac{1}{\mathbf{x}}} \\ &= \sum_{i=1}^n \left\{ -\ln \mathbf{s} - \left(\frac{1+\mathbf{x}}{\mathbf{x}} \right) \ln \left[1 + \mathbf{x} \left(\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) \right] - \left[1 + \mathbf{x} \left(\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) \right]^{\frac{1}{\mathbf{x}}} \right\}, \end{aligned} \quad (5)$$

para $\mathbf{m} - \frac{\mathbf{s}}{\mathbf{x}} > x_{(n)}$ e $\mathbf{x} < 0$ ou $\mathbf{m} - \frac{\mathbf{s}}{\mathbf{x}} < x_{(1)}$ e $\mathbf{x} > 0$. Caso contrário, $l(\mathbf{m}, \mathbf{s}, \mathbf{x})$ não existe.

Os estimadores de máxima verossimilhança de $\mathbf{m}$, $\mathbf{s}$ e $\mathbf{x}$ são obtidos pela solução do sistema de equações não lineares formado pelas derivadas de primeira ordem da eq. (5), em relação a cada parâmetro, igualadas a zero, isto é, pela solução de:

$$\begin{aligned} \frac{1}{\hat{\mathbf{s}}} \sum_{i=1}^n \left(\frac{1 + \hat{\mathbf{x}} - w_i \frac{1}{\hat{\mathbf{x}}}}{\hat{\mathbf{s}}} \right) &= 0 \\ -\frac{n}{\hat{\mathbf{s}}} + \frac{1}{\hat{\mathbf{s}}^2} \sum_{i=1}^n \left\{ \frac{(x_i - \hat{\mathbf{m}}) \left[(1 + \hat{\mathbf{x}}) - w_i \frac{1}{\hat{\mathbf{x}}} \right]}{w_i} \right\} &= 0 \end{aligned} \quad (6)$$

$$\sum_{i=1}^n \left\{ \left(1 - w_i \frac{1}{\hat{\mathbf{x}}} \right) \left[\frac{1}{\hat{\mathbf{x}}^2} \ln(w_i) - \frac{(x_i - \hat{\mathbf{m}})}{\hat{\mathbf{x}} \hat{\mathbf{s}} w_i} \right] - \frac{(x_i - \hat{\mathbf{m}})}{\hat{\mathbf{s}} w_i} \right\} = 0 ,$$

sendo $w_i = 1 + \hat{\mathbf{x}} \left(\frac{x_i - \hat{\mathbf{m}}}{\hat{\mathbf{s}}} \right)$.

Visto que o sistema de equações (6) não possui solução analítica, utilizou-se o método de Newton-Raphson para obtenção de uma solução numérica, partindo-se de valores iniciais para $\mathbf{m}$, $\mathbf{s}$ e $\mathbf{x}$. Fixando um valor inicial arbitrário $\mathbf{x}_0$ para $\mathbf{x}$, propõem-se como valores iniciais $\mathbf{m}_0$ e $\mathbf{s}_0$ para $\mathbf{m}$ e $\mathbf{s}$, os valores tais que $E(X) = \bar{x}$ e $Var(X) = s^2$, sendo $\bar{x}$ e s^2 , respectivamente, a média e a variância da série de observações. Considerando-se a função densidade de probabilidade dada por (3) obtém-se:

$$E(X) = \mathbf{m} + \frac{\mathbf{s}}{\mathbf{x}} [\Gamma(1 - \mathbf{x}) - 1] ,$$

se $\mathbf{x} < 1$, e

$$Var(X) = \frac{\mathbf{s}^2}{\mathbf{x}^2} [\Gamma(1 - 2\mathbf{x}) - \Gamma^2(1 - \mathbf{x})] ,$$

se $\mathbf{x} < 1/2$ sendo $\Gamma(\cdot)$ a função Gama. De uma forma geral o k -ésimo momento centrado existe se $\mathbf{x} < 1/k$ (Smith, 2001). Obtêm-se assim, as seguintes expressões para os valores iniciais:

$$\mathbf{s}_o = s \sqrt{\frac{\mathbf{x}_o^2}{\Gamma(1-2\mathbf{x}) - \Gamma^2(1-\mathbf{x})}} \quad (7)$$

e

$$\mathbf{m}_o = \bar{x} - \frac{\Gamma(1-\mathbf{x}_o) - 1}{\mathbf{x}_o} \mathbf{s}_o = \bar{x} - \frac{s}{\mathbf{x}_o} [\Gamma(1-\mathbf{x}_o) - 1] \sqrt{\frac{\mathbf{x}_o^2}{\Gamma(1-2\mathbf{x}_o) - \Gamma^2(1-\mathbf{x}_o)}}. \quad (8)$$

Considerando-se, por exemplo, $\mathbf{x}_o = -0,10$ nas equações (7) e (8), obtém-se:

$$\mathbf{s}_o \cong 0,87369 \ s$$

e

$$\mathbf{m}_o \cong \bar{x} - 0,4250 \ s.$$

A seguir será considerado o caso particular da distribuição GVE com $\mathbf{x}$ tendendo a zero, ou seja, a distribuição de Gumbel. Nesse caso o logaritmo da função de verossimilhança é dado por:

$$l(\hat{\theta}) = \sum_{i=1}^n \left\{ -\ln \mathbf{s} - \left(\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) - \exp \left(-\frac{x_i - \mathbf{m}}{\mathbf{s}} \right) \right\}, \quad (9)$$

e os estimadores de máxima verossimilhança de $\mathbf{m}$ e $\mathbf{s}$ são obtidos pela solução do sistema de equações:

$$-\frac{1}{\hat{\mathbf{s}}} \left\{ \left[\sum_{i=1}^n \exp \left(-\frac{x_i - \hat{\mathbf{m}}}{\hat{\mathbf{s}}} \right) \right] - n \right\} = 0$$

$$\frac{1}{\hat{\mathbf{s}}} \left\{ \left[\sum_{i=1}^n \left[\left(\frac{x_i - \hat{\mathbf{m}}}{\hat{\mathbf{s}}} \right) - \left(\frac{x_i - \hat{\mathbf{m}}}{\hat{\mathbf{s}}} \right) \exp \left(-\frac{x_i - \hat{\mathbf{m}}}{\hat{\mathbf{s}}} \right) \right] \right] - n \right\} = 0.$$

Devido ao fato deste sistema de equações não possuir solução analítica utilizou-se o método de Newton Raphson para obtenção de uma solução numérica, tomando-se como valores iniciais $\mathbf{m}_0$ e $\mathbf{s}_0$ para $\mathbf{m}$ e $\mathbf{s}$, as soluções obtidas através do método dos momentos. Para o caso da distribuição de Gumbel, tem-se que:

$$E(X) = \mathbf{m} + g\mathbf{s}$$

e

$$Var(X) = \frac{\mathbf{p}^2 \mathbf{s}^2}{6},$$

sendo γ a constante de Euler, aproximadamente igual a 0,577216.

Logo, os valores iniciais são:

$$\mathbf{m}_o = \bar{x} - \mathbf{g} \frac{\sqrt{6}}{\mathbf{p}} s \cong \bar{x} - 0,45005 s \quad (10)$$

e

$$\mathbf{s}_o = \frac{\sqrt{6}}{\mathbf{p}} s \cong 0,77970 s, \quad (11)$$

que correspondem aos limites de (7) e (8) quando $\mathbf{x}_o$ tende a zero.

3.2.4 Seleção da distribuição de valores extremos

Para testar se as observações seguem uma distribuição de valores extremos tipo I, II ou III, basta testar se $\mathbf{x}$ é igual a zero na distribuição GVE, o que pode ser feito através do teste da razão de verossimilhança modificado (Hosking, 1984), descrito a seguir.

Seja $(x_1, \dots, x_n)$ a série de n observações, $l(\hat{\mathbf{e}}_{GVE})$ e $l(\hat{\mathbf{e}}_G)$ os máximos do logaritmo da função de máxima verossimilhança das distribuições GVE (5) e de Gumbel (9), em que $\hat{\mathbf{e}}_{GVE} = (\hat{\mathbf{m}}, \hat{\mathbf{s}}, \hat{\mathbf{x}})'$ e $\hat{\mathbf{e}}_G = (\hat{\mathbf{m}}, \hat{\mathbf{s}})'$ são os vetores de estimativas de máxima verossimilhança. A estatística de razão de verossimilhança (T_{LR}) é dada por:

$$T_{LR} = -2 [l(\hat{\mathbf{e}}_G) - l(\hat{\mathbf{e}}_{GVE})] = 2 [l(\hat{\mathbf{e}}_{GVE}) - l(\hat{\mathbf{e}}_G)],$$

que tem distribuição assintótica χ^2 com 1 grau de liberdade.

De modo a obter uma aproximação mais acurada à distribuição assintótica dessa estatística, Hosking (1984) sugere a utilização da estatística modificada:

$$T_{LR}^* = \left(1 - \frac{2,8}{n}\right) T_{LR},$$

sendo n o tamanho da amostra.

Assim, para testar a hipótese $H_0: \mathbf{x} = 0$ contra $H_a: \mathbf{x} \neq 0$, compara-se o valor da estatística T_{LR}^* para a amostra com o valor tabelado de $\chi^2_{a,1}$ da distribuição χ^2 com 1 grau de liberdade e um certo nível de significância $\mathbf{a}$. Se $T_{LR}^* \geq \chi^2_{a,1}$, rejeita-se H_0 .

3.2.5 Obtenção da matriz de variâncias e covariâncias de $\hat{\mathbf{e}}$.

Seja $H(\mathbf{e})$ a matriz formada pelas derivadas parciais de segunda ordem de $l(\mathbf{e})$ em relação a $\mathbf{e}$, cujo (i,j) -ésimo elemento é dado por:

$$H_{ij} = - \left. \frac{\partial^2 l(\mathbf{e})}{\partial \mathbf{e}_i \partial \mathbf{e}_j} \right|_{\mathbf{e} = \hat{\mathbf{e}}} \quad i, j = 1, \dots, 3 .$$

sendo $\mathbf{e} = (\mathbf{m}, \mathbf{s}, \mathbf{x})'$ o vetor de parâmetros e $\hat{\mathbf{e}} = (\hat{\mathbf{m}}, \hat{\mathbf{s}}, \hat{\mathbf{x}})'$ o vetor de estimativas. Essa matriz é chamada matriz Hessiana ou matriz de informação observada e tem dimensões $(p \times p)$, em que p é o número de parâmetros do modelo, no caso igual a 3.

Ao trocar a matriz $H(\mathbf{e})$ pela matriz de valores esperados de derivadas parciais de segunda ordem de $l(\mathbf{e})$ em relação a $\mathbf{e}$, obtém-se a matriz de informação de Fisher, denotada por $I(\mathbf{e})$, cujo (i,j) -ésimo elemento é dado por:

$$I_{ij} = E_x \left[- \left. \frac{\partial^2 l(\mathbf{e})}{\partial \mathbf{e}_i \partial \mathbf{e}_j} \right|_{\mathbf{e} = \hat{\mathbf{e}}} \right] \quad i, j = 1, \dots, 3 . \quad (12)$$

A matriz $I(\mathbf{e})$ tem um papel importante na estimação de máxima verossimilhança, uma vez que a inversa de $I(\mathbf{e})$ é a matriz de variâncias e covariâncias das estimativas de máxima verossimilhança dos parâmetros. O i -ésimo elemento da diagonal de $I(\mathbf{e})^{-1}$ é a variância de $\hat{\mathbf{e}}_i$ e o (i,j) -ésimo elemento de $I(\mathbf{e})^{-1}$, $i \neq j$, a covariância entre $\hat{\mathbf{e}}_i$ e $\hat{\mathbf{e}}_j$.

3.2.6 Diagnóstico do ajuste

Seja $x_{(1)} < x_{(2)} < \dots < x_{(n)}$ a série de dados observados colocados em ordem ascendente. A função de distribuição acumulada empírica de X , para $X = x_{(i)}$, $i = 1, 2, \dots, n$, é dada pela fórmula de Cunnane (Hesal, 1995):

$$\hat{F}(x_{(i)}) = \frac{i - 0,4}{n + 0,2} \quad i = 1, \dots, n . \quad (13)$$

Essa função pode ser utilizada na construção dos gráficos probabilidade-probabilidade (*PP-plot*, em inglês), com pontos dados pelas coordenadas:

$$\left[\hat{F}(x_{(i)}), F(x_{(i)}) \right]_{\mathbf{e} = \hat{\mathbf{e}}} \quad i, j = 1, \dots, n ,$$

sendo $F(\cdot)$ a função de distribuição acumulada GVE, apresentada na eq. (1). Sob a hipótese de os dados apresentarem essa distribuição, os pontos estarão alinhados na reta que passa pelos pontos (0;0) e (1;1). Assim, uma forma de interpretar tal gráfico é observar o quão distantes esses pontos estão dessa reta. Quanto mais distantes, menos adequada é a distribuição.

Outra forma de avaliar graficamente o ajuste da distribuição é a utilização dos gráficos quantil-quantil (*QQ-plot*, em inglês) formado pelos pontos de coordenadas:

$$\left[F^{-1}\left(\frac{i-0,4}{n+0,2}\right) \middle| \hat{\mathbf{e}} = \hat{\mathbf{e}}, x_{(i)} \right] \quad i=1, \dots, n,$$

sendo $F^{-1}(\cdot)$ a função inversa de $F(\cdot)$.

Também, neste caso, sob a hipótese de os dados apresentarem distribuição GVE, os pontos estarão alinhados em uma reta. Quanto mais afastados de uma reta, menos adequada é a distribuição proposta. Dada a facilidade adicional de se detectarem eventuais pontos atípicos, o gráfico quantil-quantil foi utilizado no presente trabalho.

Como forma de testar a hipótese (H_o) da distribuição dos dados ser GVE, foi utilizado o teste de Kolmogorov-Smirnov, resumido nos seguintes passos:

- i) colocar a séries de dados em ordem ascendente,
- ii) obter os valores de probabilidade da distribuição teórica $F(x_{(i)})$ através da substituição das estimativas de máxima verossimilhança de $\mathbf{m}$, $\mathbf{s}$ e $\mathbf{x}$ na eq. (1) e os valores de probabilidade da distribuição empírica (13),
- iii) calcular a estatística D através da seguinte expressão:

$$D = \sup \left| F(x_{(i)}) - \hat{F}(x_{(i)}) \right|, \quad i=1, \dots, n,$$

- iv) rejeitar a hipótese (H_o) se $D \geq D_{n,\alpha}$, sendo $D_{n,\alpha}$ o valor crítico para valores de n e um determinado nível de significância (α). Os valores de $D_{n,\alpha}$ podem ser consultados na tabela construída por Massey (1951).

No caso de se desejar testar a hipótese da distribuição dos dados ser de Gumbel, os valores de $F(x_{(i)})$, no passo (ii), são obtidos através da substituição das estimativas de máxima verossimilhança de $\hat{m}$ e $\hat{s}$ em (2).

3.2.7 Obtenção das probabilidades de ocorrência de valores extremos de velocidade de vento acima de valores pré-estabelecidos

Para obter as probabilidades de ocorrência mensal de valores extremos de velocidade de vento acima de x km h⁻¹, utilizou-se a seguinte expressão:

$$P(X > x) = 1 - F(x) \Big|_{\hat{e}=\hat{e}} = 1 - \exp \left\{ - \left[1 + \hat{x} \left(\frac{x - \hat{m}}{\hat{s}} \right) \right]^{-\frac{1}{\hat{x}}} \right\},$$

cujo limite para $\hat{x}$ tendendo a zero é dado por:

$$P(X > x) = 1 - \exp \left\{ - \exp \left(- \frac{x - \hat{m}}{\hat{s}} \right) \right\}.$$

3.2.8 Estimação do período de retorno

Seja A um evento e T o tempo aleatório entre ocorrências consecutivas de eventos A . O valor médio τ da variável T é denominado período de retorno do evento A . No caso em estudo, A é o evento: "vento máximo excede um determinado valor x " cuja probabilidade $P(A)$ é dada por $1 - F(x)$. Assim, o período de retorno para esse evento, é dado por:

$$t = \frac{1}{P(A)} = \frac{1}{1 - F(x)},$$

sendo τ geralmente expresso em anos.

3.2.9 Estimativas dos níveis de retorno da distribuição GVE

O nível de retorno (x_p), associado ao período de retorno τ , é obtido a partir da solução da equação:

$$\int_{-\infty}^{x_p} f(\hat{e}) dx = 1 - p,$$

para $p = 1/\tau$, ou seja,

$$F(x_p) = (1 - p) . \quad (14)$$

Ao inverter (14) chega-se à solução:

$$x_p = F^{-1}(1 - p) = \mathbf{m} - \frac{\mathbf{S}}{\mathbf{x}} \left\{ 1 - [-\ln(1 - p)]^{-\mathbf{x}} \right\} , \quad (15)$$

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

$$x_p = F^{-1}(1 - p) = \mathbf{m} - \mathbf{S} \left\{ \ln [-\ln(1 - p)] \right\} . \quad (16)$$

A estimativa $\hat{x}_p$ do nível de retorno x_p para períodos de retorno $\tau = 1/p$, é obtida por substituição das estimativas de máxima verossimilhança de $\mathbf{m}$, $\mathbf{S}$ e $\mathbf{x}$ na eq. (15) e de $\mathbf{m}$ e $\mathbf{S}$ na eq. (16).

3.2.10 Obtenção dos intervalos de confiança para os níveis de retorno

Além das estimativas pontuais, foram construídos os intervalos de confiança (I.C.) com coeficiente de confiança de $(1-\alpha)100\%$ para os níveis de retorno x_p utilizando inicialmente o método delta e posteriormente, o método baseado na estatística de razão de verossimilhança, apresentado em Walshaw (1994).

De acordo com o método delta, descrito em Grigoriu (1984), o intervalo de confiança para x_p com $(1-\alpha)100\%$ de confiança é dado por:

$$\text{I.C.}(x_p) = \hat{x}_p \pm z_{\alpha/2} \sqrt{\text{Var}(x_p)} ,$$

sendo α o nível de significância, $z_{\alpha/2}$ o valor tal que $P(|Z| < z_{\alpha/2}) = 1-\alpha$ e Z , uma variável com distribuição normal padronizada e $\text{Var}(x_p)$, a variância associada ao nível de retorno x_p .

Este método baseia-se no fato de a distribuição de $\hat{\mathbf{e}} = (\hat{\mathbf{m}}, \hat{\mathbf{S}}, \hat{\mathbf{x}})'$ ser assintoticamente normal com média $\mathbf{e} = (\mathbf{m}, \mathbf{S}, \mathbf{x})'$ e matriz de variâncias e covariâncias dada por $I(\mathbf{e})^{-1}$. Como a eq. (15) é uma função não linear de $\mathbf{m}$, $\mathbf{S}$ e $\mathbf{x}$, pode-se linearizá-la através de expansão de primeira ordem em série de Taylor em torno do ponto inicial correspondente ao vetor de estimativas dos parâmetros $\hat{\mathbf{m}}, \hat{\mathbf{S}}$ e $\hat{\mathbf{x}}$, ou seja,

$$x_p \equiv \hat{x}_p + \frac{\partial x_p}{\partial \mathbf{m}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} (\mathbf{m} - \hat{\mathbf{m}}) + \frac{\partial x_p}{\partial \mathbf{S}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} (\mathbf{S} - \hat{\mathbf{S}}) + \frac{\partial x_p}{\partial \mathbf{x}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} (\mathbf{x} - \hat{\mathbf{x}}).$$

Logo, quando o parâmetro $\mathbf{x}$ na distribuição GVE é diferente de zero, a variância do nível de retorno x_p pode ser aproximada por:

$$\begin{aligned} Var(x_p) &\equiv \begin{bmatrix} \frac{\partial x_p}{\partial \mathbf{m}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} & \frac{\partial x_p}{\partial \mathbf{S}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} & \frac{\partial x_p}{\partial \mathbf{x}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \end{bmatrix} \begin{bmatrix} Var(\mathbf{m}) & Cov(\mathbf{m}, \mathbf{S}) & Cov(\mathbf{m}, \mathbf{x}) \\ Cov(\mathbf{m}, \mathbf{S}) & Var(\mathbf{S}) & Cov(\mathbf{S}, \mathbf{x}) \\ Cov(\mathbf{m}, \mathbf{x}) & Cov(\mathbf{S}, \mathbf{x}) & Var(\mathbf{x}) \end{bmatrix} \begin{bmatrix} \frac{\partial x_p}{\partial \mathbf{m}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \\ \frac{\partial x_p}{\partial \mathbf{S}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \\ \frac{\partial x_p}{\partial \mathbf{x}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \end{bmatrix} \\ &\equiv \left(\frac{\partial x_p}{\partial \mathbf{m}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \right)^2 Var(\mathbf{m}) + \left(\frac{\partial x_p}{\partial \mathbf{S}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \right)^2 Var(\mathbf{S}) + \left(\frac{\partial x_p}{\partial \mathbf{x}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \right)^2 Var(\mathbf{x}) + 2 \frac{\partial x_p}{\partial \mathbf{m}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \frac{\partial x_p}{\partial \mathbf{S}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} Cov(\mathbf{m}, \mathbf{S}) + \\ &2 \frac{\partial x_p}{\partial \mathbf{m}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \frac{\partial x_p}{\partial \mathbf{x}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} Cov(\mathbf{m}, \mathbf{x}) + 2 \frac{\partial x_p}{\partial \mathbf{S}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \frac{\partial x_p}{\partial \mathbf{x}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} Cov(\mathbf{S}, \mathbf{x}), \end{aligned}$$

em que,

$$\frac{\partial x_p}{\partial \mathbf{m}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} = 1,$$

$$\frac{\partial x_p}{\partial \mathbf{S}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} = -\frac{1}{\hat{\mathbf{x}}} \left\{ 1 - [-\ln(1-p)]^{-\hat{\mathbf{x}}} \right\} = \frac{\hat{x}_p - \hat{\mathbf{m}}}{\hat{\mathbf{S}}},$$

e

$$\begin{aligned} \frac{\partial x_p}{\partial \mathbf{x}} \bigg|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} &= \frac{\hat{\mathbf{S}}}{\hat{\mathbf{x}}^2} \left\{ 1 - [-\ln(1-p)]^{-\hat{\mathbf{x}}} \right\} - \frac{\hat{\mathbf{S}}}{\hat{\mathbf{x}}} \left\{ [-\ln(1-p)]^{-\hat{\mathbf{x}}} \ln[-\ln(1-p)] \right\} \\ &= -\left(\frac{\hat{x}_p - \hat{\mathbf{m}}}{\hat{\mathbf{x}}} \right) - \left[(\hat{x}_p - \hat{\mathbf{m}}) + \frac{\hat{\mathbf{S}}}{\hat{\mathbf{x}}} \right] \ln[-\ln(1-p)]. \end{aligned}$$

Por sua vez, para o caso em que $\mathbf{x}$ tende a zero, tem-se:

$$\begin{aligned} Var(x_p) &\cong \begin{bmatrix} \left. \frac{\partial x_p}{\partial \mathbf{m}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} & \left. \frac{\partial x_p}{\partial \mathbf{s}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \end{bmatrix} \begin{bmatrix} Var(\mathbf{m}) & Cov(\mathbf{m}, \mathbf{s}) \\ Cov(\mathbf{m}, \mathbf{s}) & Var(\mathbf{s}) \end{bmatrix} \begin{bmatrix} \left. \frac{\partial x_p}{\partial \mathbf{m}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \\ \left. \frac{\partial x_p}{\partial \mathbf{s}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \end{bmatrix} \\ &\cong \left(\left. \frac{\partial x_p}{\partial \mathbf{m}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \right)^2 Var(\mathbf{m}) + 2 \left. \frac{\partial x_p}{\partial \mathbf{m}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \left. \frac{\partial x_p}{\partial \mathbf{s}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} Cov(\mathbf{m}, \mathbf{s}) + \left(\left. \frac{\partial x_p}{\partial \mathbf{s}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} \right)^2 Var(\mathbf{s}), \end{aligned}$$

em que,

$$\left. \frac{\partial x_p}{\partial \mathbf{m}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} = 1,$$

e

$$\left. \frac{\partial x_p}{\partial \mathbf{s}} \right|_{\hat{\mathbf{e}}=\hat{\mathbf{e}}} = -\{\ln[-\ln(1-p)]\} = \frac{\hat{x}_p - \hat{\mathbf{m}}}{\hat{\mathbf{s}}}.$$

Logo,

$$Var(x_p) \cong Var(\mathbf{m}) + 2 \frac{\hat{x}_p - \hat{\mathbf{m}}}{\hat{\mathbf{s}}} Cov(\mathbf{m}, \mathbf{s}) + \left(\frac{\hat{x}_p - \hat{\mathbf{m}}}{\hat{\mathbf{s}}} \right)^2 Var(\mathbf{s}).$$

Como foi visto anteriormente, as variâncias e covariâncias das estimativas dos parâmetros $\mathbf{m}$, $\mathbf{s}$ e $\mathbf{x}$, podem ser obtidas através da matriz de variâncias e covariâncias dada pelo inverso da matriz de informação de Fisher (12).

Apesar da facilidade de implementação desse método, intervalos de confiança mais acurados para os níveis de retorno (x_p), segundo Walshaw (1994), podem ser obtidos através do método baseado na estatística de razão de verossimilhança (T_{LR}), dada por:

$$T_{LR} = 2 [l(\hat{\mathbf{e}}) - l(\hat{\mathbf{e}})],$$

que tem distribuição assintótica χ^2 com p graus de liberdade, sendo p o número de parâmetros.

Segundo esse método, a região de 100 (1- α) % de confiança para o vetor θ de parâmetros é dada pela desigualdade:

$$T_{LR} \leq \chi^2_{p,\alpha} , \quad (17)$$

sendo $\chi^2_{p,\alpha}$ o valor tabelado da distribuição χ^2 com p graus de liberdade e um certo nível de significância α . A Figura 2 apresenta, por exemplo, a região de 95% de confiança para os parâmetros m e s da distribuição de Gumbel, ajustada à série de dados de velocidade máxima de vento do mês de janeiro em Piracicaba.

Figura 2 – Região de 95% de confiança para os parâmetros da distribuição de Gumbel.

Para obter o intervalo de confiança para x_p deve-se escolher vários pontos θ que estejam compreendidos na região definida pela desigualdade (17) e baseando-se neles, gerar vários valores de x_p através das equações (15) ou (16). O maior e o menor destes valores definem os limites superior e inferior, respectivamente, do intervalo de 100 (1- α) % de confiança para o nível de retorno x_p .

No caso da distribuição de Gumbel, o nível de retorno é uma função polinomial de primeiro grau de m e s , que é dada por:

$$x_p = m + k s , \quad (18)$$

em que $k = 1,4999, 2,2504, 3,9019$ e $4,6001$, para os períodos de retorno de 5, 10, 50 e 100 anos, respectivamente.

Logo, o máximo e o mínimo de x_p são obtidos com as estimativas $(\mathbf{m}, \mathbf{s})$ pertencentes aos limites da região de confiança. O interesse, então, é maximizar a eq. (18) sujeita à restrição $T_{LR} = \chi^2_{p,\alpha}$. Considerando-se que a estatística T_{LR} é uma função dos parâmetros do modelo, essa restrição pode ser escrita da seguinte forma:

$$h(\mathbf{m}, \mathbf{s}) = \chi^2_{p,\alpha},$$

que é equivalente a:

$$h(\mathbf{m}, \mathbf{s}) - \chi^2_{p,\alpha} = 0,$$

sendo $h(\mathbf{m}, \mathbf{s}) = T_{LR}$.

Com esta informação, construi-se uma função auxiliar $g(\mathbf{m}, \mathbf{s}, \lambda)$ definida por:

$$g(\mathbf{m}, \mathbf{s}, \lambda) = (\mathbf{m} + k \mathbf{s}) + \lambda [h(\mathbf{m}, \mathbf{s}) - \chi^2_{p,\alpha}]. \quad (19)$$

A nova variável λ denomina-se multiplicador de Lagrange. Segundo o método dos multiplicadores de Lagrange, se $(\mathbf{m}^o, \mathbf{s}^o, \lambda^o)$ é um ponto crítico da eq. (19), então $(\mathbf{m}^o, \mathbf{s}^o)$ é também um ponto crítico da eq. (18) sujeita à restrição: $h(\mathbf{m}, \mathbf{s}) - \chi^2_{p,\alpha} = 0$.

Desta maneira, basta encontrar os pontos críticos $(\mathbf{m}^o, \mathbf{s}^o)$ da função auxiliar $g(\mathbf{m}, \mathbf{s}, \lambda)$ que são dados pela solução do sistema de equações formado pelas derivadas de primeira ordem da eq. (19), em relação a $\mathbf{m}, \mathbf{s}$ e λ , respectivamente, igualadas a zero, isto é, pela solução de:

$$1 + I^o \left[\frac{\partial h(\mathbf{m}, \mathbf{s})}{\partial \mathbf{m}} \right]_{\mathbf{m}=\mathbf{m}^o, \mathbf{s}=\mathbf{s}^o} = 0$$

$$k + I^o \left[\frac{\partial h(\mathbf{m}, \mathbf{s})}{\partial \mathbf{s}} \right]_{\mathbf{m}=\mathbf{m}^o, \mathbf{s}=\mathbf{s}^o} = 0$$

$$h(\mathbf{m}^o, \mathbf{s}^o) - \chi^2_{p,\alpha} = 0.$$

Uma vez obtida a solução desse sistema de equações, o que pode ser feito através de métodos numéricos, substituem-se $\mathbf{m}^o$ e $\mathbf{s}^o$ no lugar de $\mathbf{m}$ e $\mathbf{s}$, respectivamente, na eq.(18), obtendo-se os limites do intervalo de confiança para x_p .

4 RESULTADOS E DISCUSSÃO

Como primeira etapa, foi realizada uma análise exploratória da variável aleatória velocidade máxima de ventos (km h^{-1}) para cada um dos meses do ano cujas estatísticas são apresentadas na Tabela 2.

Tabela 2. Estatísticas descritivas da variável aleatória velocidade máxima mensal (km h^{-1}) de vento, nos períodos de 1956 a 1971 e de 1974 a 2000, em Piracicaba, SP

Mês	Média	Mediana	Variância	Desvio Padrão	Amplitude Interquartílica	Coefficiente de Assimetria	Coefficiente de variação (%)
Janeiro	68,84	68,40	158,79	12,60	14,76	0,60	18,30
Fevereiro	67,98	64,80	140,63	11,86	16,20	0,61	17,40
Março	63,73	61,56	181,26	13,46	16,92	0,39	21,10
Abril	57,93	54,00	207,84	14,42	20,52	0,96	24,90
Maio	58,91	55,44	279,53	16,72	28,80	0,70	28,40
Junho	57,09	53,64	234,47	15,31	22,68	1,15	26,80
Julho	54,89	51,84	167,98	12,96	19,08	0,67	23,60
Agosto	57,74	54,72	158,88	12,60	19,08	0,36	21,80
Setembro	70,76	67,86	257,34	16,04	22,86	0,45	22,70
Outubro	74,62	73,00	281,25	16,77	25,20	0,79	22,50
Novembro	72,70	69,12	334,79	18,30	19,08	1,61	25,20
Dezembro	70,67	68,40	190,61	13,81	14,04	1,11	19,50

Observa-se na Tabela 2, que os meses de setembro a dezembro apresentam, em média, os valores mais altos de velocidade máxima de vento, sendo outubro, o mês que mostra, em média, o maior valor ($74,62 \text{ km h}^{-1}$). Além dos meses citados anteriormente, também apresentam em média, valores altos (acima de 60 km h^{-1}), janeiro, fevereiro e março. Segundo Vianello & Alves (1991), isto é devido ao aquecimento diurno ser maior nesses períodos do ano, provocando a penetração de linhas de instabilidade, com mudanças bruscas no regime de vento, que pode passar de calmo ($<1 \text{ km h}^{-1}$) a vendaval ($88 \text{ a } 101 \text{ km h}^{-1}$).

Figura 3 - Gráfico de radar para representar as medidas de tendência central média e mediana, da variável velocidade máxima de vento (km h^{-1}) em cada um dos meses do ano.

A Figura 3 apresenta a média e a mediana em cada um dos meses do ano, onde se nota um acentuado decréscimo dessas medidas de fevereiro a julho, sendo que, neste último mês, se registram os menores valores (média= $54,89 \text{ km h}^{-1}$ e mediana= $51,84 \text{ km h}^{-1}$). Essa figura mostra, ainda, que a mediana é sistematicamente menor do que a média, o que sugere que as distribuições sejam assimétricas à direita, fato reforçado pelos valores positivos dos coeficientes de assimetria apresentados na Tabela 2. Segundo Tubelis & Nascimento (1984) e Vianello & Alves (1991), este padrão de variação da velocidade máxima ao longo do ano poder ser considerado como típico da região sudeste do Brasil.

A Figura 4 apresenta os gráficos de caixa (*box plot*) para a variável velocidade máxima de vento, para cada mês do ano, que sugerem a presença de alguns valores aparentemente atípicos (representados pelos símbolos $\circ$ e $*$), principalmente nos meses de janeiro, março, junho, outubro, novembro e dezembro. Como esses valores podem estar influenciando as medidas de dispersão, variância e desvio padrão, são apresentados, na Tabela 2, os valores de amplitude interquartílica, cujos maiores valores são observados em maio e outubro. Nota-se assim, que apesar de maio apresentar um dos menores valores de média e mediana, possui uma das maiores dispersões, o que pode ser visualizado na Figura 4 e quantificado através do coeficiente de variação, apresentado na Tabela 2.

Figura 4 - Gráficos de caixa (*box plot*) para a variável velocidade máxima de vento para cada um dos meses do ano.

Para verificar a pressuposição de independência dos dados através do teste de chorrilho, foram calculados, para cada série mensal de velocidade máxima de vento, os números de valores menores do que a mediana (n_1) e maiores do que a mediana (n_2), e o número total de seqüências v , apresentados na Tabela 3. Comparando os valores de v com os valores críticos ($v_{0,05,n1,n2}$), para um nível de 5% de significância, apresentados na Tabela 3, nota-se que todos se encontram dentro dos limites da região de aceitação da hipótese de aleatoriedade. Conclui-se, então, que não há evidências para assumir que as seqüências de medições sejam dependentes. Assim,

segundo Sharma et al. (1999), o cumprimento desta pressuposição garante a obtenção de inferências estatísticas satisfatórias a partir dos modelos probabilísticos de valores extremos.

Tabela 3. Números totais de valores menores (n_1) e maiores (n_2) do que a mediana, número total de seqüências (v) e valores críticos ($v_{0,05;n_1;n_2}$) segundo cada mês do ano, para uso no teste de chorrilho.

Mês	n_1	n_2	v	Valores críticos ($v_{0,05;n_1;n_2}$)	
				Inferior	Superior
Janeiro	20	18	18	13	27
Fevereiro	21	21	27	15	29
Março	20	21	27	14	29
Abril	21	20	26	14	29
Maio	21	21	20	15	29
Junho	20	20	22	14	28
Julho	21	20	18	14	29
Agosto	21	21	25	15	29
Setembro	21	21	22	15	29
Outubro	21	21	16	15	29
Novembro	21	21	20	15	29
Dezembro	20	21	21	14	29

Como uma segunda etapa, foi ajustada a distribuição GVE, cujas estimativas dos parâmetros $\mathbf{m}$, $\mathbf{S}$ e $\mathbf{x}$, obtidas através do método de máxima verossimilhança, e suas respectivas variâncias e covariâncias estimadas, para cada um dos meses do ano, são apresentadas na Tabela 4.

Tabela 4. Estimativas dos parâmetros da distribuição generalizada de valores extremos e respectivas variâncias e covariâncias estimadas, para cada um dos meses do ano.

Mês	$\hat{m}$	$\hat{s}$	$\hat{x}$	$V\hat{a}r(\hat{m})$	$V\hat{a}r(\hat{s})$	$V\hat{a}r(\hat{x})$	$C\hat{o}v(\hat{m},\hat{s})$	$C\hat{o}v(\hat{m},\hat{x})$	$C\hat{o}v(\hat{s},\hat{x})$
Janeiro	63,53	10,77	-0,10	3,52	1,79	0,01	0,71	-0,09	-0,07
Fevereiro	62,74	9,74	-0,05	2,96	1,61	0,02	0,83	-0,10	-0,08
Março	58,41	12,28	-0,17	4,40	2,19	0,01	0,58	-0,09	-0,08
Abril	50,39	9,27	0,22	2,04	1,90	0,03	1,45	-0,11	-0,06
Mai	50,37	11,88	0,14	4,95	3,17	0,03	2,38	-0,20	-0,13
Junho	49,39	10,08	0,17	3,28	2,09	0,02	1,53	-0,11	-0,06
Julho	49,13	10,55	-0,04	3,46	1,88	0,02	0,98	-0,11	-0,08
Agosto	52,65	11,17	-0,15	4,07	2,26	0,02	0,95	-0,16	-0,14
Setembro	63,57	13,20	-0,05	6,25	3,69	0,03	2,23	-0,24	-0,20
Outubro	67,36	13,90	-0,06	5,68	2,91	0,01	1,30	-0,10	-0,07
Novembro	64,51	12,68	0,07	4,65	2,51	0,01	1,48	-0,07	-0,03
Dezembro	64,54	10,67	-0,002	3,34	1,74	0,01	0,91	-0,07	-0,04

Analisando-se a Tabela 4, observa-se que as estimativas pontuais do parâmetro de forma ($\hat{x}$) com valores menores do que zero, que correspondem à distribuição de Weibull, é o caso mais freqüente (em 8 dos 12 meses). Esta distribuição, segundo Holmes & Moriarty (1999), é a mais apropriada para representar fenômenos ambientais, como a velocidade máxima de ventos, devido ao fato de possuir uma cauda superior com limite finito. No entanto, para os meses de abril, maio e junho e novembro, as estimativas pontuais do parâmetro de forma são maiores do que zero, correspondendo à distribuição de Fréchet. Esta distribuição, segundo os autores, não é adequada para estudar o comportamento da velocidade máxima do vento, pois apresenta cauda superior com limite infinito, que conduz a previsões ilimitadas de níveis de retorno com os incrementos nos períodos de retorno. Comentam, no entanto, que essa distribuição pode surgir devido a velocidades originadas por diferentes tipos de ventos, quanto ao seu mecanismo de origem, ou a possíveis erros na amostragem para amostras pequenas. Sugerem, nesse caso, que se opte pela distribuição de Gumbel, que apesar de apresentar cauda superior com limite infinito, levam a previsões de níveis de retorno inferiores aos obtidos quando se utiliza a distribuição de Fréchet.

Considerando que o parâmetro de forma ξ define o tipo de distribuição de valores extremos a utilizar, são apresentados na Tabela 5 os intervalos de 95% de confiança para ξ , baseados na aproximação normal, em cada um dos meses do ano. A Tabela 5 mostra, ainda, os valores da estatística de razão de verossimilhança modificada (T_{LR}^*).

Tabela 5. Intervalos de 95 % de confiança para o parâmetro de forma (ξ) e valores da estatística de razão de verossimilhança modificada (T_{LR}^*) para cada um dos meses do ano.

Mês	Limites de 95 % de confiança para ξ		T_{LR}^*
	Inferior	Superior	
Janeiro	-0,30	0,10	0,6331
Fevereiro	-0,33	0,23	0,1259
Março	-0,37	0,03	1,9727
Abril	-0,12	0,56	1,9650
Maiο	-0,20	0,48	0,5999
Junho	-0,11	0,45	1,4555
Julho	-0,32	0,24	0,0850
Agosto	-0,43	0,13	0,9477
Setembro	-0,34	0,29	0,0662
Outubro	-0,26	0,14	0,2731
Novembro	-0,13	0,27	0,4445
Dezembro	-0,20	0,19	0,0004

Comparando-se os valores da estatística T_{LR}^* apresentados na Tabela 5, com o valor tabelado de χ^2 com um grau de liberdade e nível de 5 % de significância, dado por $\chi^2_{1;0,05} = 3,84$, conclui-se que a distribuição de Gumbel é a mais adequada para modelar os dados de velocidade máxima de vento considerados. Essa conclusão é reforçada pelo fato de o valor nulo de ξ , que corresponde à distribuição de Gumbel, estar compreendido dentro dos limites do intervalo de confiança para ξ .

Na Tabela 6, são apresentadas as estimativas dos parâmetros m e s da distribuição de Gumbel obtidas através do método de máxima verossimilhança para cada um dos meses do ano, e suas respectivas variâncias e covariâncias estimadas.

Tabela 6. Estimativas dos parâmetros $\hat{m}$ e $\hat{s}$ da distribuição de Gumbel e correspondentes variâncias e covariâncias estimadas.

Mês	$\hat{m}$	$\hat{s}$	$V\hat{a}r(\hat{m})$	$V\hat{a}r(\hat{s})$	$C\hat{o}v(\hat{m}, \hat{s})$
Janeiro	62,97	10,49	2,92	1,54	0,68
Fevereiro	62,48	9,57	2,37	1,32	0,56
Março	57,31	11,82	3,64	1,86	0,85
Abril	51,53	10,33	2,73	1,70	0,65
Mai	51,27	12,68	4,13	2,51	0,99
Junho	50,37	10,96	3,07	1,88	0,73
Julho	48,91	10,40	2,80	1,56	0,66
Agosto	51,76	10,63	2,93	1,58	0,69
Setembro	63,24	12,95	4,44	2,55	1,07
Outubro	66,91	13,66	4,95	2,60	1,19
Novembro	64,98	12,97	4,31	2,43	0,98
Dezembro	64,52	10,66	2,93	1,60	0,68

Convém observar que durante o procedimento de estimação dos parâmetros das distribuições GVE e de Gumbel, houve sempre uma rápida convergência através do método de Newton-Raphson, proporcionada pela boa seleção dos valores iniciais dos parâmetros, obtidos através de (10) e (11).

Para verificar a qualidade do ajuste da distribuição de Gumbel, inicialmente foram construídos os gráficos quantil-quantil, apresentados na Figura 5, que sugerem o bom ajuste da distribuição de Gumbel aos dados de velocidade máxima de vento, para todos os meses do ano.

Figura 5 – Gráficos quantil-quantil para diagnóstico do ajuste da distribuição de Gumbel aos dados de velocidade máxima de vento para os doze meses do ano.

Em seqüência, para verificar a qualidade do ajuste da distribuição de Gumbel, aplicou-se o teste de Kolmogorov-Smirnov, com um nível de significância de 5%, cujas diferenças máximas absolutas observadas entre os valores de probabilidade das funções de probabilidade acumulada empírica e de Gumbel (teórica), assim como a estatística $D_{n,\alpha}$, para cada mês, são apresentados na Tabela 7. Segundo este teste, a distribuição de Gumbel ajusta-se bem aos dados em todos os meses do ano, concordando com as conclusões obtidas a partir da análise dos gráficos quantil-quantil.

Tabela 7. Resultados do teste de Kolmogorov-Smirnov para verificação da qualidade do ajuste da distribuição de Gumbel aos dados de velocidade máxima de ventos (km h^{-1}).

Mês	Número de dados (n)	Diferença máxima Absoluta (D)	$D_{n, 0,05}$
Janeiro	42	0,05	0,21
Fevereiro	43	0,07	0,20
Março	43	0,07	0,20
Abril	43	0,11	0,20
Maiο	43	0,11	0,20
Junho	43	0,12	0,20
Julho	43	0,09	0,20
Agosto	43	0,08	0,20
Setembro	42	0,06	0,21
Outubro	43	0,07	0,20
Novembro	43	0,07	0,20
Dezembro	43	0,06	0,20

Estes resultados, no entanto, concordam parcialmente com os obtidos por Angelocci et al. (1995) para a mesma localidade e registros provenientes da mesma estação meteorológica. Estes autores concluíram que a distribuição de Gumbel não teve um bom ajuste para os meses de fevereiro, abril e novembro. Esta divergência nas conclusões pode ser explicada pelo fato da série de dados utilizada por estes autores ter sido pequena (20 anos).

Na Tabela 8, apresentam-se as probabilidades de ocorrência de velocidades máximas de vento acima de 40, 50, 60, 70, 80, 90 e 100 km h⁻¹, para todos os meses do ano. Observa-se nesta tabela, que no período de setembro a março, registram-se rajadas com velocidades acima de 60 km h⁻¹, classificadas como ventos muito fortes segundo a escala de medida de intensidade dos ventos proposta por Beaufort (National Weather Service, 2002), com probabilidade de ocorrência maior do que 0,50 (valor considerado alto, segundo Angelocci et al., 1995). Essa tabela mostra, ainda, que a probabilidade de ocorrência de ventos com velocidades superiores a 100 km h⁻¹ é maior nos meses de setembro, outubro e novembro, em comparação com o resto de meses do ano. Estes ventos são classificados como tormentas violentas ou tempestades, com grau 11 na escala de Beaufort (NWS, 2002), e têm grande importância, já que podem causar danos estruturais consideráveis e arrancamento de árvores.

Tabela 8. Probabilidades de ocorrência de rajadas máximas mensais de vento com velocidade acima de 40, 50, 60, 70, 80, 90 e 100 km h⁻¹, a 10 m acima do nível do solo, para os 12 meses do ano, em Piracicaba, SP.

Mês	Velocidade (km h ⁻¹)						
	>40	>50	>60	>70	>80	>90	>100
Janeiro	0,99987	0,96799	0,73477	0,40056	0,17909	0,07327	0,02892
Fevereiro	0,99997	0,97485	0,72625	0,36601	0,14812	0,05483	0,01964
Março	0,98676	0,84369	0,54907	0,28950	0,13642	0,06100	0,02665
Abril	0,95282	0,68653	0,35640	0,15414	0,06161	0,02387	0,00913
Mai	0,91223	0,66895	0,39486	0,20405	0,09850	0,04602	0,02118
Junho	0,92386	0,64442	0,33978	0,15355	0,06475	0,02652	0,01073
Julho	0,90508	0,59357	0,29125	0,12334	0,04909	0,01906	0,00733
Agosto	0,95145	0,69285	0,36910	0,16451	0,06773	0,02700	0,01062
Setembro	0,99756	0,93792	0,72305	0,44738	0,23965	0,11888	0,05679
Outubro	0,99923	0,96821	0,80954	0,54952	0,31851	0,16840	0,08486
Novembro	0,99895	0,95814	0,76953	0,49275	0,26942	0,13513	0,06494
Dezembro	0,99995	0,97988	0,78319	0,45024	0,20874	0,08756	0,03522

Os períodos de retorno estimados (expressos em anos) para o maior valor de velocidade máxima de vento (x_n) registrado em cada um dos meses do ano, são apresentados na Tabela 9.

Observa-se nesta tabela que janeiro e maio, assim como março e abril, apresentam valores idênticos de x_n , com períodos de retorno diferentes, o que pode ser explicado devido ao fato de o comportamento da distribuição de Gumbel ser distinto para os diferentes meses do ano.

Tabela 9. Períodos de retorno estimados para os maiores valores de velocidade máxima de vento (km h^{-1}) registrados em cada um dos meses do ano, nos períodos de 1956 a 1971 e de 1974 a 2000.

Mês	Maior valor registrado de velocidade máxima de vento (km h^{-1}) (x_n)	Período de retorno (anos)
Janeiro	100,08	35
Fevereiro	99,36	48
Março	97,20	30
Abril	97,20	84
Maio	100,08	48
Junho	108,00	193
Julho	90,00	52
Agosto	81,00	16
Setembro	102,96	22
Outubro	126,72	80
Novembro	140,04	327
Dezembro	114,48	109

As estimativas dos níveis de retorno mensais e intervalos de confiança associados aos períodos de retorno 5, 10, 50 e 100 anos obtidos pelo método delta são apresentados na tabela 10, e os obtidos através da estatística de razão de verossimilhança, na Tabela 11.

Tabela 10. Níveis de retorno (km h^{-1}) estimados e limites inferior (LI) e superior (LS) de seus respectivos intervalos de 95 % de confiança, para os períodos de retorno 5, 10, 50 e 100 anos, obtidos através do método delta.

Mês	Período de retorno (anos)											
	5 anos			10 anos			50 anos			100 anos		
	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS
Janeiro	73,02	78,71	84,40	79,31	86,59	93,86	92,88	103,92	114,95	98,57	111,24	123,92
Fevereiro	71,64	76,83	82,03	77,35	84,02	90,69	89,67	99,82	109,98	94,84	106,51	118,18
Março	68,73	75,04	81,35	75,85	83,91	91,97	91,25	103,43	115,61	97,71	111,69	125,66
Abril	61,31	67,03	72,75	67,38	74,78	82,19	80,47	91,85	103,23	85,95	99,06	112,17
Maio	63,28	70,29	77,29	70,75	79,80	88,84	86,87	100,73	114,60	93,62	109,59	125,55
Junho	60,76	66,80	72,84	67,22	75,03	82,84	81,14	93,13	105,12	86,97	100,78	114,59
Julho	58,87	64,51	70,15	65,07	72,32	79,56	78,46	89,49	100,53	84,07	96,76	109,45
Agosto	61,96	67,70	73,44	68,32	75,68	83,03	82,06	93,23	104,39	87,82	100,64	113,47
Setembro	75,49	82,66	89,83	83,15	92,37	101,60	99,68	113,76	127,84	106,61	122,80	138,99
Outubro	79,96	87,40	94,83	88,14	97,64	107,15	105,82	120,20	134,59	113,22	129,74	146,26
Novembro	77,44	84,43	91,42	85,17	94,16	103,14	101,86	115,57	129,29	108,86	124,63	140,40
Dezembro	74,77	80,51	86,25	81,15	88,51	95,87	94,92	106,12	117,32	100,69	113,56	126,43

Tabela 11. Níveis de retorno (km h^{-1}) estimados e limites inferior (LI) e superior (LS) de seus respectivos intervalos de 95 % de confiança, para os períodos de retorno 5, 10, 50 e 100 anos obtidos através da estatística de razão de verossimilhança.

Mês	Período de retorno (anos)											
	5 anos			10 anos			50 anos			100 anos		
	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS	LI	$\hat{x}_p$	LS
Janeiro	72,54	78,71	87,24	79,84	86,59	97,68	92,36	103,92	121,00	98,04	111,24	130,97
Fevereiro	71,23	76,83	84,62	76,94	84,02	94,17	89,21	99,82	115,50	94,36	106,51	124,61
Março	68,17	75,04	84,43	75,30	83,91	96,09	90,62	103,43	122,15	97,05	111,69	133,28
Abril	60,93	67,03	75,69	67,00	74,78	86,15	80,04	91,85	109,53	85,51	99,06	119,51
Maio	62,80	70,29	80,85	70,27	79,80	93,64	86,30	100,73	122,20	93,03	109,59	134,40
Junho	60,55	66,80	75,77	66,81	75,03	87,01	80,67	93,13	111,74	86,49	100,78	122,31
Julho	58,42	64,51	72,97	64,62	72,32	83,35	77,95	89,49	106,53	83,55	96,76	116,44
Agosto	61,49	67,70	76,28	67,86	75,68	86,84	81,53	93,23	110,43	87,27	100,64	120,51
Setembro	74,96	82,66	93,45	82,62	92,37	106,46	99,07	113,76	135,55	105,98	122,80	147,97
Outubro	79,44	87,40	98,38	87,63	97,64	111,94	105,23	120,20	142,24	112,62	129,74	155,19
Novembro	76,90	84,43	94,95	84,65	94,16	107,91	101,29	115,57	136,87	108,27	124,63	149,24
Dezembro	74,31	80,51	89,11	80,69	88,51	99,72	94,40	106,12	123,43	100,16	113,56	133,56

Analisando a Tabela 10, nota-se que o método delta gera intervalos de confiança sempre simétricos, em torno à estimativa pontual $\hat{x}_p$ e de menor amplitude, quando comparados com os obtidos através da estatística de razão de verossimilhança (Tabela 11), assimétricos à direita. Os gráficos de radar apresentados nas Figuras 6 e 7 ilustram esta comparação.

(a)

(b)

Período de retorno de 5 anos

(a)

(b)

Período de retorno de 10 anos

Figura 6 – Gráficos de radar para representar a estimativa pontual do nível de retorno (km h^{-1}) para períodos de retorno de 5 e 10 anos, em cada um dos meses do ano, e seus respectivos intervalos de confiança obtidos através dos métodos delta (a) e da estatística de razão de verossimilhança.

(a)

(b)

Período de retorno de 50 anos

(a)

(b)

Período de retorno de 100 anos

Figura 7 – Gráficos de radar para representar a estimativa pontual do nível de retorno (km h^{-1}) para períodos de retorno de 50 e 100 anos, em cada um dos meses do ano, e seus respectivos intervalos de confiança obtidos através dos métodos delta (a) e da estatística de razão de verossimilhança.

5 CONCLUSÕES

Diante dos resultados obtidos pode-se concluir que:

- a) Na parte metodológica referente à obtenção dos parâmetros das distribuições generalizada dos valores extremos e de Gumbel, o método da máxima verossimilhança mostrou ser satisfatório.
- b) A distribuição generalizada de valores extremos com parâmetro $x=0$, que corresponde à distribuição de valores extremos tipo I ou de Gumbel é adequada para estudar o comportamento da velocidade máxima de vento em todos os meses do ano, em Piracicaba.
- c) As maiores velocidades máximas de vento registram-se nos meses de setembro a dezembro. Ventos com velocidades acima de 60 km h^{-1} , considerados como muito fortes, e com probabilidade de ocorrência superior a 0,5, apresentam-se também neste período do ano.

REFERÊNCIAS BIBLIOGRÁFICAS

ABILD, J.; ANDERSEN, E.Y.; ROSBJERG, D. The climate of extreme winds at the Great Belt of Denmark. **Journal of Wind Engineering and Industrial Aerodynamics**, v.41, p.521-532, 1992.

ANGELOCCI, L.R.; WIENDL, F.W.; ARRUDA, H.V. Probabilidades mensais de ocorrência de rajadas de vento na região de Piracicaba, SP. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 9., Campina Grande, 1995. **Anais**. Campina Grande: Sociedade Brasileira de Agrometeorologia; UFPB, 1995. p.498-500.

BANZATO, D.A. Estimativa das precipitações máximas prováveis no Estado de São Paulo. Jaboticabal, 1973. 119p. Tese (Doutorado) – Faculdade de Ciências Agrárias, Universidade Estadual Paulista “Julio de Mesquita Filho”.

BEIRLANT, J.; MATTHYS, G. **The generalized extreme value distribution and statistics of maxima**. <http://www.kuleuven.ac.be/ucs/research/publi.htm> (30 abr. 2001)

BLESSMANN, J. **Acidentes causados pelo vento**. Porto Alegre: Editora da UFRGS, 1986. 81p.

BRABSON, B. B.; PALUTIKOF, J. P. Tests of the generalized Pareto distribution for predicting extreme wind speeds. **Journal of Applied Meteorology**, v.39, p.1627-1640, 2000.

BROWNLEE, K.A. **Statistical theory and methodology in science and engineering**. 2.ed. New York: John Wiley, 1965. 590p.

- CAMARGO, M.B.P.; ORTOLANI, A.A.; ARRUDA, H.V. Ocorrência mensal de rajadas máximas diárias de vento em Campinas. **Bragantia**, v.53, p.107-112, 1994.
- CARVALHO, A.A. de Estudo sobre a ocorrência de aguaceiros e estiagens em Presidente Prudente, Avaré e Itapeva. Piracicaba, 1982. 76p. Dissertação (Mestrado) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- CASTILLO, E. **Extreme value theory in engineering**. San Diego: Academic Press, 1988. 388p.
- CHOWDHURY, J.U.; STEDINGER, J.R.; LU, L.H. Goodness of fit for regional generalized extreme value flood distribution. **Water Resources Research**, v.27, p.1765-1776, 1991.
- COLES, S. **Extreme value theory and applications**. In: REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 44.; SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 8., Botucatu, 1999. **Minicurso**. Botucatu: UNESP, Departamento de Bioestatística, 1999. p.1-10.
- CORONADO, M. **Extreme value theory (EVT) for risk managers pitfall and opportunities in the use of EVT in measuring EVT**.
<http://www.gloriamundi.org/var/pub/coronadoEVT.pdf> (10 dez. 2000)
- DÍAZ, J.F. **Introducción a los métodos no paramétricos**. Xalapa: Universidad Veracruzana, Facultad de Estadística e Informática, 1999. 134p.
- EISENHART, C.; SWED, F. Tables for testing randomness of grouping in a sequence of alternatives. **Annals of Mathematical Statistics**, v.14, p.66-87, 1943.
- FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distributions of the largest or smallest member of a sample. **Proceedings of the Cambridge Philosophical Society**, v.24, p.180-190, 1928.

- GALAMBOS, J.; MACRI, N. Classical extreme value model and prediction of extreme wind. **Journal of Structural Engineering**, v.125, p.792-794, 1999.
- GNEDENKO, B. V. Sur la distribution limite du terme maximum d'une série aléatoire. **Annales des Mathématiques**, v.44, p.423-453, 1943.
- GRIGORIU, M. Estimates of extreme wind from short records. **Journal of Structural Engineering**, v.110, p.1467-1483, 1984.
- GUSELLA, V. Estimation of extreme winds from short-term records. **Journal of Structural Engineering**, v.117, p.375-390, 1991.
- HELSEL, D. R.; HIRSCH, R.M. **Statistical methods in water resources**. Amsterdam: Elsevier Science, 1995. 529p.
- HOLMES, J.D.; MORIARTY, W.W. Application of the generalized Pareto distribution to extreme value analysis in wind engineering. **Journal of Wind Engineering and Industrial Aerodynamics**, v.83, p.1-10, 1999.
- HOSKING, J. R. M. Testing whether the shape parameter is zero in the generalized extreme-value distribution. **Biometrika**, v.71, p.367-374, 1984.
- HOSKING, J. R. M. Algorithm AS 215: Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. **Journal of the Royal Statistical Society. Applied Statistics**, v.34, p.301-310, 1985.
- HOSKING, J. R. M.; WALLIS, J. R. The effect of intensity dependence on regional frequency analysis. **Water Resources Research**, v.24, p.588-590, 1988.
- HOSKING, J. R. M.; WALLIS, J. R.; WOOD, E. F. An appraisal of the regional flood frequency procedure in the UK Flood Studies Regionals. **Hydrology Science Journal**, v.30, p.85-109, 1985a.

- HOSKING, J. R. M.; WALLIS, J. R.; WOOD, E. F. Estimation of the generalized extreme-value distribution by method of probability-weighted moments. **Technometrics**, v.27, p.251-261, 1985b.
- JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, v.81, p.158-171, Apr. 1955.
- KRISHNAM, A.; KUSHWAHA, R.S. Comparison of Gumbel and Jenkinson's methods of estimating probable maximum rainfall intensities over India for different return periods. **Journal of the Indian Geophysical Union**, v.4, p.50-70, 1978.
- KÜCHENHOFF, H; THAMERUS, M. Extreme value analysis of Munich air pollution data. **Environmental and Ecological Statistics**, v.3, p.127-141, 1996.
- LETTENMAIER, D.P.; WALLIS, J.R.; WOOD, E.F. Effect of regional heterogeneity on flood frequency estimation. **Water Resources Research**, v.23, p.313-323, 1987.
- LONGIN, F.M. From value at risk to stress testing: the extreme value approach. **Journal of Banking & Finance**, v.24, p.1097-1130, 2000.
- MACLEOD, A. J. AS 276 –A remark on algorithm AS 215: Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. **Journal of the Royal Statistical Society. Applied Statistics**, v.38, p.198-199, 1989.
- MADSEN, H.; PEARSON, C.P.; ROSBJERG, D. Comparison of annual maximum series and partial duration serie methods for modeling hydrologic events. **Water Resources Research**, v.33, p.759-770, 1997.
- MARTINS, D. O comportamento dos ventos na região de Botucatu, SP. In: CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA, 22., Ilhéus, 1993. **Anais**. Ilhéus: Sociedade Brasileira de Engenharia Agrícola, 1993. p.815-825.

- MARTINS, E.S.; STEDINGER, J.R. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. **Water Resources Research**, v.36, p.737-744, 2000.
- MASSEY, F.R. The Kolmogorov-Smirnov test for goodness of fit. **Journal of American Statistical Association**, v.46, p.68-78, 1951.
- MEDICI, E.V.; ABRAMOVITZ, L.; COSTA, L.C. da; et al. Teoria dos valores extremos: uma aplicação ao meio ambiente. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA, 14., Caxambu, 2000. **Resumos**. Caxambu: Associação Brasileira de Estatística, 2000. p.315-316.
- NATIONAL WEATHER SERVICE. **The beaufort wind force scale**.
<http://www.crh.noaa.gov/lot/webpage/beaufort/> (20 fev. 2002)
- OCCHIPINTI, A.G.; SANTOS, P.M. **Análise das máximas intensidades de chuva na cidade de São Paulo**. São Paulo: USP, Instituto Astronômico e Geofísico, 1965. 41p.
- ORTOLANI, A.A. Agroclimatologia e o cultivo da seringueira. In: SIMPÓSIO SOBRE A CULTURA DA SERINGUEIRA NO ESTADO DE SÃO PAULO, 1., Piracicaba, 1986. **Trabalhos apresentados**. Campinas: Fundação Cargill, 1986. p.11-32.
- OTTEN, A.; MONTFORT, A.J. Maximum-likelihood estimation of the general extreme-value distribution parameters. **Journal of Hydrology**, v.47, p.187-192, 1980.
- PIEGORSCH, W. W.; SMITH, E. P.; EDWARDS, D.; et al. Statistical advances in environmental science. **Statistical Science**, v.13, p.186-208, 1998.
- PIRES, R.F.; ANGELOCCI, L.R.; ALFONSI, R.R.; et al. Probabilidades mensais de ocorrência de rajadas de rajadas máximas de vento nas localidades de Ribeirão Preto e Jaú, SP. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 11., Florianópolis, 1999. **Anais**. Florianópolis: Sociedade Brasileira de Agrometeorologia; UFSC, 1999. p.320.

- PRESCOTT, P.; WALDEN, A. T. Maximum likelihood estimation of the parameters of the generalized extreme value distribution. **Biometrika**, v.67, p.723-724, 1980.
- PRESCOTT, P.; WALDEN, A. T. Maximum likelihood estimation of the parameters of the generalized extreme value distribution from censored samples. **Journal of Statistical Computation and Simulation**, v.6, p.241-250, 1983.
- RAYNAL, J. A.; Sobre el uso del dominio de atracción para la identificación de valores extremos para máximos. **Ingeniería Hidráulica en México**, v.12, p.57-62, 1997.
- ROSS, W.H. A peaks over threshold analysis of extreme wind speeds. **Canadian Journal of Statistic**, v.15, p.328-335, 1987.
- SHARMA, P.; KHARE, M.; CHAKRABARTI, S. P. Application of extreme value theory for predicting violations of air quality standards for an urban road intersection. **Transportation Research**, v.23, p.133-139, 1999.
- SIEGEL, S. **Nonparametric statistics for behavioral sciences**. New York: McGraw-Hill, 1956. 312 p.
- SILVA, M. M.; ASSIS, S.V. Probabilidades de ocorrência de estiagem no Triangulo Mineiro e sua influência na cultura do milho. **Boletim da Sociedade Brasileira de Meteorologia**, v.62, p.43-56, 1979.
- SIMIUI, E.; FILLIBEN, J. Probability distribution of extreme wind speeds. **Journal of Structural Engineering**, v.102, p.1861-1877, 1976.
- SIMIUI, E.; HECKERT, N.A. Extreme wind distribution tails: A peak over threshold approach. **Journal of Structural Engineering**, v.122, p.539-547, 1996.
- SMITH, R. L. Maximum likelihood estimation in a class of nonregular cases. **Biometrika**, v.72, p.67-92, 1985.

SMITH, R. L. **Environmental statistics:** extreme value theory section.

<http://www.stat.unc.edu/postscript/rs/envstat/env.html> (15 oct. 2001)

THAS, O.; VANROLLEGHEM, P.; KOPS, B. et al. Extreme value statistics: potencial benefits in water quality management . **Water Science Technology**, v.36, p.133-140, 1997.

TIAGO DE OLIVEIRA, J. Statistical choice of univariate extreme models. **Statistical Distributions in Scientific Work**, v.6, p.367-387, 1981.

TUBELIS, A.; NASCIMENTO, F.J. do **Meteorologia descritiva:** fundamentos e aplicações brasileiras. São Paulo: Nobel, 1984. cap.7, p.145-165: Vento.

VAN MONTFORT, M.A.J. On testing that the distribution of extreme is type I when type II is the alternative. **Journal of Hydrology**, v.11, p.421-427, 1978.

VIANELLI, R.L.; ALVES, A.R. **Meteorologia básica e aplicações.** Viçosa: UFV, 1991. 449p.

WAGNER, S.C. ; BERNARDES, L.R.M.; CORREA, A. et al. Estudo da velocidade e direção predominante dos ventos no estado de Paraná. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 5., Belém, 1987. **Anais.** Belém: Sociedade Brasileira de Agrometeorologia; CPATU, 1987. p.269-273.

WALLIS, J.R.; WOOD, E.F. Relative accuracy of log Pearson III procedures. **Journal of Hydraulic Engineering**, v.111, p.1043-1056, 1985.

WALSHAW, D. Getting the most from your extreme wind data: a step by step guide. **Journal Resource of National Standard Technology**, v.99, p.399-411, 1994.

ZAR, J.H. **Bioestatistical analysis.** 4.ed. New Jersey: Prentice Hall, 1999. cap.25, p.583-587: Testing for randomness.