

CONFIABILIDADE NA ESTIMAÇÃO DE VAZÕES EXTREMAS

POR

J.P.da Costa¹, J.M.Damazio¹, M.V.F.Pereira¹ e J.Kelman^{1,2}

RESUMO -- Fazendo uso de um longo registro sintético de vazões diárias mostra-se que diversas distribuições de probabilidades estimam com igual precisão vazões máximas com período de retorno da ordem do tamanho da amostra. Entretanto ocorrem diferenças significativas quando se trata de obter vazões com elevado tempo de recorrência. Procedeu-se a uma análise de sensibilidade do desempenho das distribuições, usando como critério o erro médio quadrático e a probabilidade de subestimar a vazão de cheia.

INTRODUÇÃO

A metodologia probabilística de cálculo de vazões de cheia requer o uso de uma distribuição de probabilidades para eventos extremos.

A escolha da distribuição mais adequada a partir de um histórico de vazões é uma tarefa que apresenta enormes dificuldades conceituais: todos os métodos de seleção e validação de modelos estocásticos avaliam a capacidade do modelo de reproduzir eventos "semelhantes" aos registrados no histórico. Em outras palavras, estes métodos são aplicáveis a eventos que correspondem a interpolações da amostra existente e cujo tempo de retorno, portanto, está na ordem de grandeza do comprimento da série histórica. Este é o caso, por exemplo, dos critérios recentemente desenvolvidos para modelagem de vazões mensais (Costa et al., 1981), vazões diárias (Barbosa, 1981) e vazões de cheias (Pinheiro et al., 1981).

Entretanto, no projeto de vertedores, o tempo de retorno de interesse pode chegar a 10.000 anos. É necessário, portanto, extrapolar a amostra existente.

Como é impossível validar qualquer técnica de extrapolação, adotam-se na prática modelos de bom desempenho para tempos de recorrência relativamente curtos, na expectativa de que este desempenho se mantenha para eventos menos frequentes.

¹Pesquisador do Centro de Pesquisas de Energia Elétrica - CEPEL

²Professor Visitante da Coordenação dos Programas de Pós-Graduação em Engenharia da Universidade Federal do Rio de Janeiro (COPPE/UFRJ).

O presente artigo faz uma análise de sensibilidade sobre a capacidade de extrapolação de diversos modelos existentes na literatura. Utilizou-se uma série sintética de vazões diárias longa o suficiente para permitir a estimação de eventos com tempo de retorno elevado. Os diversos modelos são ajustados a partir de subamostras desta série e utilizados na estimação destes eventos extremos. O desempenho dos modelos é então medido pelos erros cometidos na estimação e pela probabilidade de subestimar/superestimar os valores "corretos".

ANÁLISE DE SENSIBILIDADE

O modelo estocástico proposto por Kelman (1980) foi ajustado aos 33 anos de vazões diárias (1931-1963) do posto São José da Barra, situada no Rio Grande. Foram gerados 10.000 anos de vazões diárias, das quais se extraiu as máximas vazões diárias de cada ano.

Slacket al. (1978) e Landwehr et al. (1980) optaram por produzir registros sintéticos de máximas vazões diárias utilizando diretamente as distribuições de probabilidades usualmente empregadas em análise de frequência de cheias. Este procedimento pode levar a séries de máximas pouco plausíveis, pois as distribuições adotadas representam aproximações de um fenômeno natural bastante complexo. A modelagem de vazões diárias, por outro lado, permite aproveitar a totalidade da informação histórica sobre a estrutura sequencial das vazões diárias e, portanto, aproximar de maneira mais adequada este fenômeno. Barbosa (1981) confirma o bom desempenho do modelo de vazões adotado.

A distribuição empírica obtida a partir dos 10.000 máximos fornece bons estimadores de vazões com períodos de retorno de até 1.000 anos. Foram selecionadas sete distribuições teóricas: normal (N), log-Normal com dois parâmetros (LN2), log-Normal com três parâmetros (LN3), Gumbel (G), Gumbel generalizada (GEV), log-Pearson tipo III (LP3) e Wakeby (W). As seis primeiras distribuições são usualmente empregadas na análise de frequência de cheias. A distribuição Wakeby foi incluída porque a partir de 1978 vem crescendo o número de artigos publicados a seu respeito, muitos dos quais mostrando resultados promissores quanto a sua adequação a modelagem de cheias (Houghton, 1978; Landwehr et al., 1978, 1979a, 1979b, 1980; Wallis, 1980, 1981).

Avaliação das incertezas na estimação

As distribuições acima foram inicialmente ajustadas a 50 subamostras de máximas anuais de tamanho 30 e usadas para estimar os mesmos percentis obtidos a partir da distribuição empírica. Para cada tempo de recorrência as distribuições teóricas foram comparadas com base no erro médio quadrático definido por:

$$EMQ_t = E[(Q_t - q_t)^2] = \{q_t - E[Q_t]\}^2 + \text{Var}[Q_t] \quad (1)$$

onde:

Q_t : é a cheia de projeto para o período de retorno t , estimada a partir da distribuição empírica dos 10.000 valores;

\hat{Q}_t : é a cheia de projeto para o período de retorno t , estimada a partir da distribuição empírica.

A figura 1 apresenta para cada distribuição a evolução do erro percentual, definido por $\sqrt{(EMQ_t)}/Q_t \times 100\%$ como função do tempo de retorno t para subamostras de 30 anos. Nota-se que para tempos de recorrência de 30 anos, que corresponde ao tamanho da amostra, os erros percentuais das sete distribuições são próximos. Com o aumento do tempo de recorrência percebe-se uma divergência no caso da distribuição Wakeby e, em grau menor, no caso da log-Pearson tipo III.

O procedimento foi então repetido para 50 subamostras de máximas anuais de tamanhos 10 e 50 anos. O comportamento encontrado foi semelhante, como pode ser verificado nas figuras 2 e 3. É interessante observar que o ponto de divergência continuou a ser aproximadamente igual ao comprimento das respectivas subamostras: na figura 2, 10 anos e na figura 3, 50 anos.

Riscos de subdimensionamento

Landwehr et al. (1980) observam que o erro médio quadrático como critério de escolha entre estimadores alternativos de Q_t equivale a funções de penalidade quadráticas e simétricas. Nos casos em que os prejuízos devidos ao superdimensionamento forem diferentes daqueles devidos ao subdimensionamento, é de interesse o conhecimento da probabilidade de subdimensionamento, isto é, $P[\hat{Q}_t < Q_t]$. A figura 4 apresenta, para cada distribuição, a evolução da probabilidade de subdimensionamento como função de t , para subamostras de 30 anos.

Pode-se observar que, para as distribuições N, GEV e LN3, a frequência de subdimensionamento excede 50%. As figuras 5 e 6 mostram que comportamento semelhante ocorre para subamostras de tamanho 10 e 50 anos.

CONCLUSÕES

Os resultados indicam que para o posto de São José da Barra é indiferente a distribuição utilizada na interpolação. Por outro lado, resultado significativamente distinto foi obtido na extrapolção. Considerando como critério para a escolha da distribuição apenas o erro percentual, deve-se descartar as distribuições W e LP3, cujos desempenhos foram nitidamente inferiores. Entretanto caso se considere o superdimensionamento preferível ao subdimensionamento deve-se usar as distribuições LN2 ou G. Isto porque, apesar das distribuições N, GEV e LN3 conduzirem a baixos erros percentuais, apresentam elevada probabilidade de subdimensionamento.

Os erros percentuais correspondentes ao período de retorno de 1.000 anos (considerando apenas os resultados das distribuições LN2 e G) se situam na faixa de 22%, 17% e 15% respectivamente para tamanhos de amostra de 10, 30 e 50 anos. Convém lembrar que estes valores são estimativas dos erros que se pode cometer em postos similares quando se faz uso apenas de informações locais. Uma alternativa para diminuir os erros percentuais consiste na regionalização. A disponibilidade de longos registros de vazões sintéticas também permite testar métodos de regionalização, criando uma região fictícia, onde cada posto corresponde a uma subamostra gerada pelo modelo. Este assunto está sendo presentemente investigado.

REFERÊNCIAS

- BARBOSA, P. J., (1981), "Comparação de Modelos Estocásticos para Vazões Diárias", tese de mestrado, COPPE/UFRJ.
- COSTA, C.G., OLIVEIRA, G.C., PEREIRA, M.V.F., KELMAN, J., (1981), "Seleção de Modelo de Afluências Mensais para Planejamento Energético", VI Seminário Nacional de Produção e Transmissão de Energia Elétrica, Camboriú, Santa Catarina.
- HOUGHTON, J.C., (1978), "Birth of a Parent: The Wakeby Distribution for Modelling Flood Flows", Water Resources Research, 14(6), 1111-1115.
- KELMAN, J., (1980), "A Stochastic Model for Daily Streamflow", Journal of Hydrology, 47: 235-249.
- LANDWEHR, J.M., MATALAS, N.C., WALLIS, J.R., (1978), "Some Comparisons of Flood Statistics in Real and Log Space", Water Resources Research, 14(5), 902-920.
- _____, (1979a), "Estimation of Parameters and Quantiles of Wakeby Distributions. Part 1: Known Lower Bounds", Water Resources Research, 15(6), 1361-1372.
- _____, (1979b), "Estimation of Parameters and Quantiles of Wakeby Distributions. Part 2: Unknown Lower Bounds", Water Resources Research, 15(6), 1373-1379.
- _____, (1980), "Quantile Estimation with more or less Floodlike Distributions", Water Resources Research, 16(3), 547-555.
- PINHEIRO, M.C., DAMAZIO, J.M., KELMAN, J., (1981), "Análise de Frequência de Cheias para as Regiões Sul e Sudeste do Brasil", IV Simpósio Brasileiro de Hidrologia e Recursos Hídricos, 15 a 19 de Novembro, Fortaleza, Ceará.
- SLACK, J.R., WALLIS, J.R., MATALAS, N.C., (1975), "On the Value of Information to Flood Frequency Analysis", Water Resources Research, 11(5), 629-647.

WALLIS, J.R., (1980), "Risk and Uncertainties in the Evaluation of Flood Events for the Design of Hydrologic Structures", Seminar on Extreme Hydrological Events: Floods and Droughts, Erice, Italy, March 22-25.

(1981), "Hydrologic Problems Associated with Oilshale Development", to be presented at the IFIP working conference on September 28-30,

ABSTRACT

It is shown, with the aid of a long synthetic daily flow series, that although flood records can be fitted over the range of usually observed data ($N = 10, 30, 50$) with the same accuracy by several types of probabilities distributions, significantly different results are obtained when one tries to extrapolate. The mean squared error and the probability of underdesign is used to evaluate the performance of several distributions.

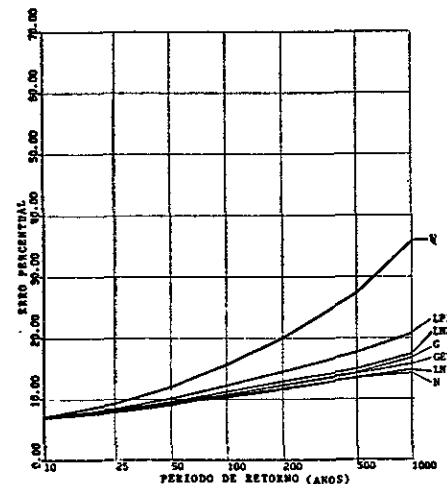


Figura 1. Erro percentual das distribuições teóricas em função do período de retorno. (Tamanho de cada subamostra: 30 anos)

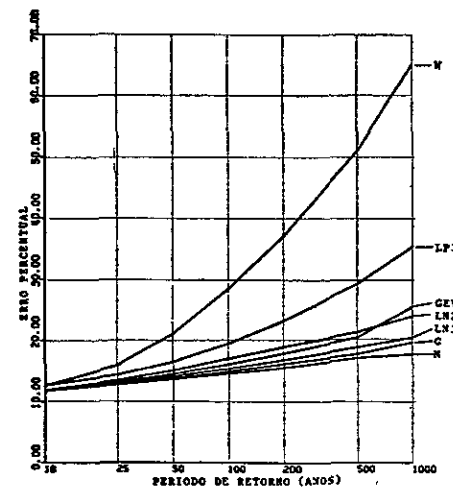


Figura 2. Erro percentual das distribuições teóricas em função do período de retorno. (Tamanho de cada subamostra: 10 anos)

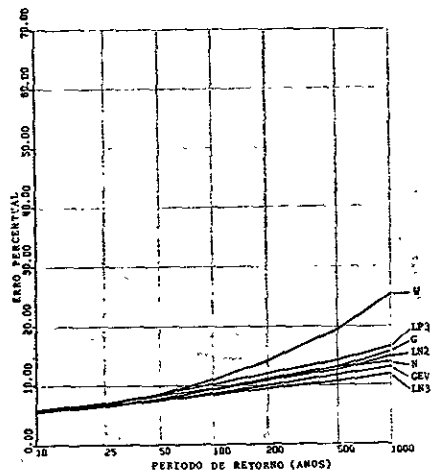


Figura 3. Erro percentual das distribuições teóricas em função do período de retorno. (Tamanho de cada subamostra: 50 anos)

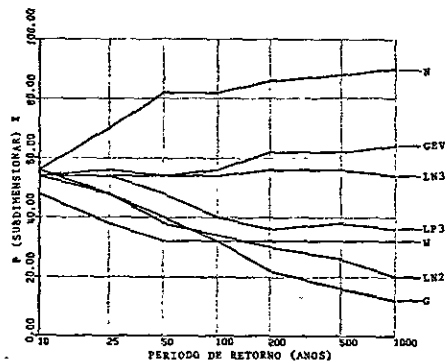


Figura 4. Probabilidade de subdimensionamento das distribuições teóricas em função do período de retorno, (Tamanho de cada subamostra: 30 anos)

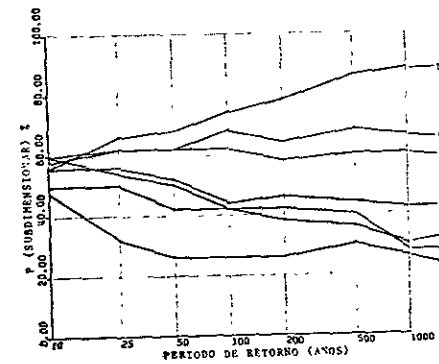


Figura 5. Probabilidade de subdimensionamento das distribuições teóricas em função do período de retorno. (Tamanho de cada subamostra: 10 anos)

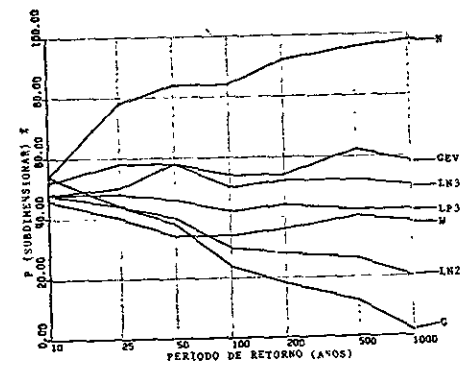


Figura 6. Probabilidade de subdimensionamento das distribuições teóricas em função do período de retorno. (Tamanho de cada subamostra: 50 anos)