

ESTIMAÇÃO DE VAZÕES EXTREMAS: SÉRIES PARCIAIS OU MÁXIMOS ANUAIS?

POR

J.C.Moreira¹, J.M.Damázio², J.P.da Costa² e J.Kelman^{2,3}

RESUMO -- Este trabalho compara o emprego dos métodos de séries parciais e de máximos anuais para a estimativa de vazões com tempo de recorrência elevados. Para esta comparação é utilizado um modelo matemático gerador de vazões diárias. Foram geradas, com o auxílio do modelo, séries representativas dos registros de vazões médias diárias de dois postos fluviométricos da bacia do rio Doce. Destas séries completas geradas, foram selecionadas para o estudo quatro séries de vazões extremas: séries de máximos anuais e séries parciais contendo um, dois e três picos por ano. A metodologia de avaliação dos métodos consiste em subdividir estas quatro grandes séries de extremos em sub-amostras de diferentes tamanhos e, a partir de cada uma, estimar quantis com diferentes tempos de recorrência. A escolha, em cada caso, recai sobre o método que apresenta o menor erro médio na estimativa destes quantis. Os resultados obtidos revelaram que não se pode a priori optar por séries de máximos anuais ou séries parciais ao contrário do que é sugerido por alguns autores que recomendam o uso de séries parciais para pequenos registros.

INTRODUÇÃO

O objetivo da análise de frequência de cheias é obter uma relação entre uma magnitude de vazão $q(T)$ e o tempo de recorrência T , a ela associado, para uma determinada seção de um curso d'água. Uma forma de tratar o problema é considerar unicamente a série de vazões afluentes à seção em estudo como sendo uma realização de um processo estocástico cujos mecanismos de geração não são tratados um a um, mas sim de forma global. Nesta série são procuradas evidências estatísticas que possam caracterizar tal processo, cuja identificação permitiria resolver o problema. Os métodos que usam este procedimento são chamados métodos de análise direta de frequência de cheias. Constitui o objetivo deste estudo a investigação do comportamento de alguns destes métodos com relação à determinação de vazões com tempos de recorrência muito elevados, como é o caso das vazões de projeto de vertedores. Dada a pequena extensão dos registros usualmente existentes e face ao grande tempo de recorrência de interesse é de fundamental importância extrair-se dos registros a maior quantidade possível de informação.

¹Aluno da Coordenação dos Programas de Pós-Graduação em Engenharia da Universidade Federal do Rio de Janeiro (COPPE/UFRJ), atualmente engenheiro da Engevix S.A., Rio de Janeiro.

²Pesquisador do Centro de Pesquisas de Energia Elétrica, CEPEL, Rio de Janeiro.

³Professor visitante da COPPE/UFRJ.

Seria ideal que na análise direta de frequência de cheias, fossem utilizadas todas as informações contidas nos registros de vazões diárias. Todavia, esta série apresenta uma estrutura de dependência muito complexa. Assim, o procedimento usualmente adotado na prática consiste em extrair-se destes registros de vazões diárias as chamadas séries de vazões extremas, centrando desta forma a atenção nas vazões que apresentam interesse mais direto na análise de extremos.

Existem basicamente dois tipos de séries de vazões extremas: as séries de máximos anuais e as séries parciais. As primeiras são constituídas pelos maiores valores de vazões ocorridas a cada ano enquanto que as últimas compreendem as maiores vazões ocorridas a cada cheia.

A simplicidade dos modelos matemáticos usados na análise de máximos anuais, aliada à facilidade com que estas séries são obtidas, faz com que seu uso seja mais difundido que as séries parciais, cuja análise exige o emprego de modelos matemáticos mais sofisticados. Além disso, a obtenção das séries parciais a partir dos registros de vazões diárias exige uma definição das cheias ocorridas no histórico, o que geralmente dá margem a controvérsias. Devido a esta dificuldade costuma-se definir séries parciais como as séries contendo todos os picos da hidrógrafa cujas magnitudes excedam um determinado valor usualmente chamado de nível de truncamento.

Tem sido desenvolvidos estudos (NERC, 1975; Yevjevich e Taesombut, 1976) com o objetivo de comparar o uso das séries de máximos anuais com o uso das séries parciais. Uma aparente vantagem dos métodos de séries parciais é a possibilidade de se manipularem amostras maiores. Entretanto, existem dúvidas sobre a qualidade destas amostras, principalmente porque dois picos de cheias consecutivos podem guardar entre si um elevado grau de dependência o que anularia a aparente vantagem. Além disso, os métodos de máximos anuais consideram uma informação adicional: cada elemento da amostra é a maior vazão ocorrida em um ano (NERC, 1975). O que acontece é que a realização destes estudos comparativos esbarram em algumas dificuldades. Por um lado, as comparações feitas em bases analíticas são limitadas por inúmeras hipóteses simplificadoras enquanto que as comparações diretas necessitam de um número muito grande de hidrógrafas.

De acordo com Damázio et al. (1983), a geração de um número elevado de hidrógrafas sintéticas pode fornecer uma base objetiva para a comparação direta entre o uso de séries parciais e de séries de máximos anuais na análise de vazões extremas. Estas hidrógrafas sintéticas constituiriam um cenário perfeitamente conhecido, e os desempenhos dos dois métodos poderiam ser avaliados e comparados. As conclusões fornecidas por um estudo desta natureza podem ser aplicáveis, desde que os cenários criados apresentem características semelhantes à realidade que se tenta representar.

Neste artigo são geradas séries de vazões diárias sintéticas representativas de dois postos da bacia do rio Doce com o objetivo de comparar os métodos de máximos anuais e séries parciais. Para a geração das séries sintéticas usou-se o modelo DIANA (Kelman et al. (1983)).

METODOLOGIA DE AVALIAÇÃO

Inicialmente foram geradas séries sintéticas de vazões diárias com 10000 anos, para que se pudesse determinar empiricamente a distribuição de probabilidades dos máximos anuais, definindo-se assim os valores de população $q(T)$ para vazões com tempos de retorno de até 1000 anos. Mil destas séries são divididas em subamostras contendo N anos, sendo extraídas de cada sub-amostra as séries de vazões extremas a serem analisadas: séries de máximos anuais e séries parciais com diferentes taxas anuais de pico.

A cada uma destas séries de extremos foram ajustadas as distribuições de probabilidades de interesse, sendo então estimadas, a partir do ajuste obtido, as vazões para diferentes tempos de retorno. A repetição deste procedimento para todas as sub-amostras fornece um conjunto de estimativas de $q(T)$ para cada distribuição ajustada a cada tipo de série de extremos. Os índices de desempenho de cada modelo são os mesmos que aqueles adotados em Damázio et al. (1983), quais sejam: erro médio absoluto, tendência e probabilidades de sub e superestimação.

Para se avaliar a influência do tamanho dos registros na precisão das estimativas, estes índices são levantados para diversos tamanhos de sub-amostra. O método que fornece os menores valores para o erro médio absoluto é considerado o melhor, sendo a tendência e probabilidades de sub e superdimensionamento resultados auxiliares na avaliação.

APLICAÇÃO

O modelo DIANA foi calibrado para os registros de vazões médias diárias dos postos fluviométricos de Colatina no rio Doce ($Ad = 76615 \text{ Km}^2$) e de São Sebastião da Encruzilhada no rio Manhuaçu ($Ad = 8454 \text{ Km}^2$), também na bacia do rio Doce. Foram considerados na modelagem apenas os dias correspondentes ao período chuvoso da bacia do rio Doce (19 de outubro a 30 de abril).

A tabela 1 apresenta algumas características das populações de máximos anuais geradas. A metodologia foi aplicada a quatro diferentes tipos de séries: séries de máximos anuais (SMA) e séries parciais com taxas anuais de um pico por ano, (SP1) dois picos por ano, (SP2) e três picos por ano (SP3).

As séries parciais SP1, SP2 e SP3, foram obtidas fixando-se um nível de truncamento e considerando como pico a máxima vazão de cada período de vazões acima deste nível. É computacionalmente muito trabalhoso procurar, para cada sub-amostra, o nível de trunca-

mento que fornece exatamente um, dois ou três picos por ano. Assim sendo, foi fixado como nível de truncamento a metade da média dos máximos anuais de cada sub-amostra e a partir deste nível defini da uma série parcial básica. As séries SP1, SP2 e SP3 foram então obtidas selecionando-se os N, 2N e 3N maiores valores das séries parciais básicas de cada sub-amostra de N anos.

Foram escolhidas dez distribuições de probabilidades teóricas: normal (N), log-normal de dois parâmetros (LN2), log-normal de três parâmetros (LN3), exponencial de dois parâmetros (EXP), gama de dois parâmetros (GAM2) gama de três parâmetros ou Pearson tipo 3 (P3), log-Pearson tipo três (LP3), Gumbel (GU), generalizada de extremos (GEV) e Wakeby (WAK). Estas distribuições foram ajustadas pelo método dos momentos, excessão feita à distribuição Wakeby, ajustada pelo método dos momentos ponderados pela probabilidade (Landwehr et al. (1979)) e da distribuição exponencial, ajustada tanto pelo método dos momentos quanto pelo método da máxima verossimilhança. Moreira (1983), detalha os métodos de estimação de parâmetros e as expressões usadas para a estimação dos quantis para cada distribuição e tipo de série.

A influência do tamanho das amostras foi considerada tomando se sub-amostras com N= 5, 10, 15, 20, 25, 30, 40, 50 e 60 anos. Neste trabalho foram analisadas, as previsões de vazões com os seguintes tempos de recorrência: 100, 250, 500 e 1000 anos.

A tabela 2 lista para a série de Colatina e para cada tempo de retorno e tamanho de amostra, as distribuições que apresentaram o menor erro médio absoluto na estimação dos quantis, tanto no caso das séries de máximos anuais como de séries parciais. É possível verificar, que para as séries parciais a distribuição exponencial foi a que apresentou os menores erros ao passo que para as séries de máximos anuais a distribuição de Gumbel foi a que apresentou melhor desempenho. Pode-se notar ainda que os métodos de séries parciais (exceto quando aplicados as séries SP1) forneceram resultados melhores que os métodos de máximos anuais.

Para verificar a sensibilidade dos resultados ao modelo estocástico e também ao número de séries sintéticas geradas, foi utilizado um outro modelo estocástico, de concepção mais simples (Kelman e Damázio, 1983) que depois de ajustado aos dados de Colatina foi empregado para gerar 100000 anos de vazões diárias. O processamento desta informação confirmou os resultados acima apresentadas (Moreira, 1983).

Para o posto de São Sebastião da Encruzilhada o método de máximos anuais apresentou resultados bem superiores aos dos métodos de séries parciais (tabela 3). A distribuição exponencial neste caso apresentou os menores erros tanto para as séries de máximos anuais quanto para séries parciais. Damázio et al. (1983) concluiu que quando a assimetria de população dos máximos anuais é próxima ou superior a dois, deve-se usar a distribuição exponencial para modelagem das séries de máximos anuais. Este resultado é aqui confirmado (tabela 1). Outra característica interessante dos resultados obtidos neste caso é que para registros mais longos as distribuições LP3 e GEV, que são de três parâmetros, forneceram erros absolutos mais baixos que as de dois parâmetros.

CONCLUSÕES

Foram criados dois cenários constituídos por um grande número de séries sintéticas de vazões diárias, a partir dos quais pode ser feita uma comparação do desempenho dos métodos de máximos anuais e de séries parciais no tocante à determinação de vazões com tempos de recorrência elevados. Os resultados obtidos revelaram que não se pode a priori optar por séries de máximos anuais ou séries parciais ao contrário do que é sugerido por alguns autores que recomendam o uso de séries parciais para pequenos registros (NERC, 1975).

AGRADECIMENTOS

Este artigo resulta das atividades de pesquisa do projeto CEPEL 7275 que recebe o apoio do Departamento de Estudos Energéticos da ELETROBRÁS.

REFERÊNCIAS

- DAMÁZIO, J.M., MOREIRA, J.C., COSTA, J.P. e KELMAN, J.; Seleção de métodos para estimação de vazões com tempos de retorno elevados, V Simpósio Brasileiro de Hidrologia e Recursos Hídricos, Blumenau, 1983.
- LANDWEHR, J.M., MATALAS, N.C. e WALLIS, J.R.; Estimation of parameters and quantiles of Wakeby distribution, part II: Unknown Lower Bounds, Water Resources Research, vol. 15, nº 6 1979.
- MOREIRA, J.C., Estimação de vazões extremas a partir de séries parciais e de máximos anuais, tese de mestrado a ser submetida a COPPE/UFRJ, 1983.
- National Environmental Research Council (NERC), Flood Studies Report, London, 1975.
- KELMAN, J., DAMÁZIO, J.M. e COSTA, J.P.; Geração de séries sintéticas de vazões diárias. Modelo DIANA, artigo submetido a Revista Brasileira de Engenharia, Caderno de Recursos Hídricos, 1983.
- KELMAN, J., DAMÁZIO, J.M., Synthetic Hydrology and Spillway Design, XX Congresso da IAHR, Moscou, 1983.
- YEVJEVICH, V. e TAESOMBUT V., Information on flood peaks in daily flow series, Proc. Int. Symp. on Risk and Reliability in Water Resources, Ontario, Canada, 1979.

POSTO USADO PARA CALIBRAÇÃO	MÉDIA (m³/s)	COEFICIENTE DE VARIAÇÃO	ASSIMETRIA	CURTOSE	VAZÃO MÁXIMA ANUAL COM TEMPO DE RECORRÊNCIA			
					100	250	500	1000
56990000	585	0,53	2,21	11,9	1744	2064	2314	2720
56994500	4384	0,38	1,13	4,9	9682	10761	11418	11966

Tabela 1 - Estatísticas dos Máximos Anuais das Séries Geradas
56990000 - São Sebastião da Encruzilhada
56994500 - Colatina

N	TR	SMA				SP1				SP2				SP3			
		100	250	500	1000	100	250	500	1000	100	250	500	1000	100	250	500	1000
5		GU 0,24	GU 0,25	GU 0,25	GU 0,26	EXP* 0,24	EXP 0,25	EXP 0,26	GU 0,26	EXP* 0,21	EXP* 0,22	EXP 0,22	GU 0,22	EXP* 0,17	EXP* 0,17	EXP* 0,18	EXP* 0,19
10		GU 0,18	GU 0,19	GU 0,20	GU 0,20	EXP* 0,17	EXP* 0,18	EXP 0,18	EXP 0,19	EXP* 0,14	EXP* 0,15	GU 0,15	GU 0,15	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,13
15		GU 0,16	GU 0,17	GU 0,18	GU 0,19	EXP* 0,14	EXP* 0,15	EXP 0,16	EXP 0,17	EXP* 0,11	EXP* 0,11	EXP* 0,12	EXP* 0,14	EXP* 0,09	EXP* 0,09	EXP* 0,09	EXP* 0,10
20		GU 0,14	GU 0,15	GU 0,15	GU 0,16	EXP* 0,11	EXP* 0,12	EXP* 0,14	GU 0,14	EXP* 0,09	EXP* 0,10	EXP* 0,11	GU 0,11	EXP* 0,08	EXP* 0,08	EXP* 0,08	EXP* 0,08
25		GU 0,14	GU 0,14	GU 0,14	GU 0,15	EXP* 0,12	EXP* 0,13	EXP 0,14	GU 0,14	EXP* 0,09	EXP* 0,10	EXP 0,11	GU 0,11	EXP* 0,07	EXP* 0,07	EXP* 0,07	EXP* 0,08
30		GU 0,12	GU 0,12	GU 0,13	GU 0,14	EXP* 0,09	EXP* 0,10	EXP* 0,11	EXP 0,11	EXP 0,08	EXP 0,09	EXP 0,10	GU 0,10	EXP* 0,07	EXP* 0,07	EXP* 0,07	EXP* 0,08
40		GU 0,10	GU 0,11	GU 0,11	GU 0,12	EXP* 0,08	EXP* 0,09	EXP* 0,10	EXP 0,11	EXP* 0,07	EXP* 0,08	EXP* 0,09	GU 0,09	EXP* 0,06	EXP* 0,06	EXP* 0,06	EXP* 0,07
50		GU 0,11	GU 0,11	GU 0,12	GU 0,13	EXP* 0,08	EXP* 0,09	EXP 0,10	EXP 0,11	EXP* 0,06	EXP* 0,06	EXP* 0,08	GU 0,09	EXP* 0,05	EXP* 0,05	EXP* 0,05	EXP* 0,06
60		GU 0,08	GU 0,09	GU 0,09	GU 0,09	EXP* 0,07	EXP* 0,08	EXP 0,09	EXP 0,10	EXP* 0,06	EXP 0,06	EXP 0,08	GU 0,09	EXP 0,05	EXP* 0,05	EXP* 0,05	EXP* 0,07

Tabela 2 - Erro Médio Absoluto da Distribuição de melhor Desempenho Colatina
(* Método da Máxima Verossimilhança)

ABSTRACT

This paper compares the use of partial-duration and maximum annual series for the estimation of flows with high return periods. For this comparison it is used a synthetic-daily streamflow generation model. The model is used to generate daily streamflow series that resemble recorded series of two gauges in Doce basin. From each of these series it is extracted four extreme series: maximum annual series and partial-duration series with one, two and three peaks per year. The comparison is made by splitting these series in a number of samples with the same length. Each sample is used to estimate the extreme flows and the mean absolute error of each type of series can be calculated.

The obtained results showed that it is difficult to "a priori" point to the best type of series.

N	TR	SMA				SPI				SP2				SP3			
		100	250	500	1000	100	250	500	1000	100	250	500	1000	100	250	500	1000
5	EXP* 0,29	EXP* 0,30	EXP* 0,30	EXP* 0,31	EXP* 0,28	EXP* 0,29	EXP* 0,30	EXP* 0,33	EXP* 0,26	EXP* 0,28	EXP* 0,29	EXP* 0,33	EXP* 0,28	EXP* 0,30	EXP* 0,30	EXP* 0,31	EXP* 0,35
10	EXP* 0,20	EXP* 0,21	EXP* 0,21	EXP* 0,22	EXP* 0,21	EXP* 0,22	EXP* 0,23	EXP* 0,27	EXP* 0,23	EXP* 0,25	EXP* 0,26	EXP* 0,30	EXP* 0,24	EXP* 0,26	EXP* 0,27	EXP* 0,27	EXP* 0,31
15	EXP* 0,17	EXP* 0,18	EXP* 0,18	EXP* 0,17	EXP* 0,17	EXP* 0,19	EXP* 0,20	EXP* 0,25	EXP* 0,21	EXP* 0,23	EXP* 0,24	EXP* 0,28	EXP* 0,21	EXP* 0,23	EXP* 0,25	EXP* 0,29	EXP* 0,29
20	EXP* 0,15	EXP* 0,15	EXP* 0,15	EXP* 0,14	EXP* 0,17	EXP* 0,18	EXP* 0,20	EXP* 0,25	EXP* 0,19	EXP* 0,21	EXP* 0,23	EXP* 0,27	EXP* 0,20	EXP* 0,22	EXP* 0,24	EXP* 0,27	EXP* 0,27
25	EXP* 0,14	EXP* 0,14	EXP* 0,14	EXP* 0,13	EXP* 0,14	EXP* 0,17	EXP* 0,18	EXP* 0,23	EXP* 0,17	EXP* 0,19	EXP* 0,21	EXP* 0,26	EXP* 0,18	EXP* 0,21	EXP* 0,22	EXP* 0,25	EXP* 0,25
30	EXP* 0,13	EXP* 0,13	EXP* 0,13	EXP* 0,12	EXP* 0,13	EXP* 0,16	EXP* 0,17	EXP* 0,24	EXP* 0,16	EXP* 0,18	EXP* 0,20	EXP* 0,25	EXP* 0,17	EXP* 0,20	EXP* 0,21	EXP* 0,24	EXP* 0,24
40	EXP* 0,11	EXP* 0,11	EXP* 0,11	EXP* 0,09	EXP* 0,13	EXP* 0,16	EXP* 0,18	EXP* 0,23	EXP* 0,13	EXP* 0,17	EXP* 0,19	EXP* 0,24	EXP* 0,15	EXP* 0,19	EXP* 0,20	EXP* 0,23	EXP* 0,23
50	EXP* 0,12	EXP* 0,12	EXP* 0,11	EXP* 0,09	EXP* 0,14	EXP* 0,16	EXP* 0,18	EXP* 0,24	EXP* 0,12	EXP* 0,18	EXP* 0,20	EXP* 0,25	EXP* 0,15	EXP* 0,18	EXP* 0,19	EXP* 0,22	EXP* 0,22
60	EXP* 0,10	EXP* 0,10	EXP* 0,10	EXP* 0,08	EXP* 0,13	EXP* 0,16	EXP* 0,18	EXP* 0,22	EXP* 0,11	EXP* 0,16	EXP* 0,18	EXP* 0,22	EXP* 0,12	EXP* 0,16	EXP* 0,16	EXP* 0,19	EXP* 0,19

Tabela 3 - Erro Médio Absoluto da Distribuição de melhor Desempenho. S.S.da Encruzilhada (*Método da Máxima Verossimilhança)