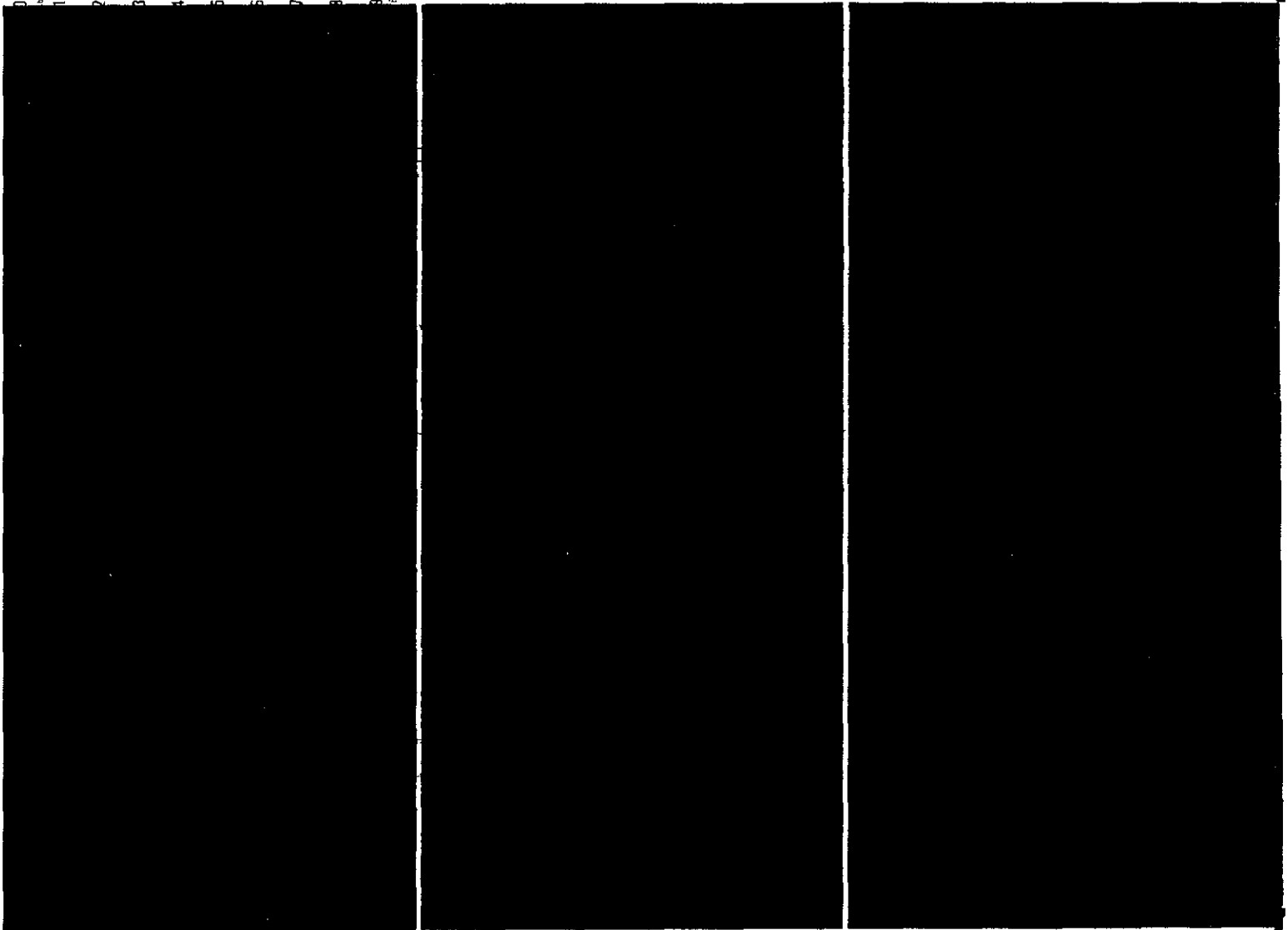


curso de  
**ENGENHARIA  
HIDROLÓGICA**



**hidrologia operacional**

Escola Politécnica da Universidade de São Paulo  
Departamento de Águas e Energia Elétrica  
Associação Brasileira de Hidrologia e Recursos Hídricos

GERAÇÃO DE SÉRIES SINTÉTICAS DE VAZÕES

P

JERSON KELMAN

---

## SUMÁRIO

---

1. GERAÇÃO DE SÉRIES SINTÉTICAS DE VAZÕES	P. 1
1.1. Porque Séries Sintéticas são necessárias	P. 1
1.2. Modelagem de Vazões Anuais	P. 2
1.3. Modelagem de Vazões Mensais	P.23
1.4. Modelagem de Vazões Diárias	P.27
REFERÊNCIAS	P.31

---

## 1. GERAÇÃO DE SÉRIES SINTÉTICAS DE VAZÕES

---

### 1.1 PORQUE SÉRIES SINTÉTICAS SÃO NECESSÁRIAS

---

Admita que se pretenda realizar algum aproveitamento hídrico num local em que se disponha de um registro de vazões. Este registro, chamado de série histórica é usualmente um dado de entrada para um modelo de simulação e/ou otimização que produz como saída algum valor relevante para o projeto. Por exemplo : o volume do reservatório necessário para regularizar uma descarga pré-fixada, ou o mínimo volume afluente num determinado intervalo de tempo, ou a descarga máxima que o vertedor deve ser capaz de dar passagem, ou outra grandeza qualquer. A figura la representa a série histórica de uma seção do rio na qual se pretende construir uma barragem que crie um reservatório de regularização. O "modelo de simulação/otimização" neste caso é simplesmente o diagrama de Rippl, que produz como saída  $x_n$  o volume de reservatório.

Ocorre que a série histórica é apenas uma das possíveis realizações de um processo estocástico; ou seja, pode-se imaginar que a Natureza "sorteou" a série histórica, segundo algum conjunto de leis probabilísticas. Um novo sorteio redundaria numa outra série, diferente da histórica, mas igualmente provável. Neste capítulo serão apresentados modelos que procuram aproximar este comportamento estocástico. Cada um destes modelos permite que artificialmente se façam tantos sorteios quantos forem necessários para o estudo em foco.

Cada sorteio estará associado a uma série de vazões, chamadas de série sintética. Como estas séries serão todas distintas entre si, pode-se obter diversos resultados provenientes de simulações, ao invés de um único resultado que seria obtido caso apenas a série histórica estivesse disponível. Desta forma, a informação contida na série histórica pode ser mais completamente extraída. A figura lb mostra que a série histórica é dado de entrada para o modelo estocástico, que produz tantas séries sintéticas quanto se queira. Como a cada série estocástica estará associado um tamanho de reservatório, é possível inferir a partir da amostra aleatória  $\{x_1, x_2, x_3, \dots\}$  a distribuição de probabilidades de X.

Caso se pretenda definir o volume do reservatório de forma tal que a probabilidade de falha seja de  $\alpha \times 100\%$ , deve-se adotar o valor  $x_\alpha$ , conforme ilustra a

figura 1c. Caso se pretenda definir o volume pelo método "tradicional", isto é  $x_h$ , existe um risco de falha igual a  $\beta$ . Neste exemplo, "falha" significa o não atendimento da descarga que se escolheu para regularizar num horizonte de planejamento de comprimento igual ao da série histórica.

A disponibilidade de muitas séries sintéticas pode dar a impressão de que é possível ter confiança absoluta nos resultados encontrados com o seu emprego. É importante ressaltar, no entanto, que a utilização de série geradas não cria informação alguma. Afinal, o processo de seleção de um modelo e de estimação de seus parâmetros é feito, como se verá adiante, a partir de inferência com a série histórica. Portanto, a utilização do modelo estocástico de vazões deve ser encarado simplesmente como uma maneira eficiente de se manipular a informação existente.

Tanto a série histórica quanto as sintéticas são discretas no tempo, ao contrário da figura 1 que por simplicidade as representou de forma contínua. Usualmente dispõe-se de vazões com intervalo de discretização  $\Delta t$  diário, que servem para definir as séries mensais e anuais. Procura-se em geral utilizar nos estudos de sistemas hídricos o maior  $\Delta t$  que não afete a precisão dos resultados. Isto porque quanto menor for  $\Delta t$  mais difícil é modelar o processo estocástico correspondente. Em outras palavras, em geral é mais fácil representar vazões anuais do que vazões diárias. A seguir serão expostas algumas idéias relacionadas a modelagem estocástica de séries hidrológicas, classificando-se os métodos de acordo com o intervalo de discretização.

## 1.2 MODELAGEM DE VAZÕES ANUAIS

Os estudos do sistema de recursos hídricos não utilizam em geral intervalos de discretização anual. O máximo  $\Delta t$  usualmente utilizado é o mês. No entanto é indispensável que se aborde a questão da modelagem de vazões anuais, tanto por fornecer uma base teórica sobre a qual se irá construir as demais modelagens, como porque uma das alternativas de se produzir séries sintéticas mensais necessita que já estejam disponíveis as séries sintéticas anuais.

### 1.2.1 Estacionariedade

A figura 2 mostra a hidrografa do Rio Paraíba do Sul em Barra do Piraí com intervalo de discretização anual. Como se vê, as vazões anuais (na realidade vazão média ao longo de um ano) flutuam entre 177 e 515 m<sup>3</sup>/s. Diversos pesquisadores tem procurado descobrir regularidades nestas flutuações, que teriam origem nas

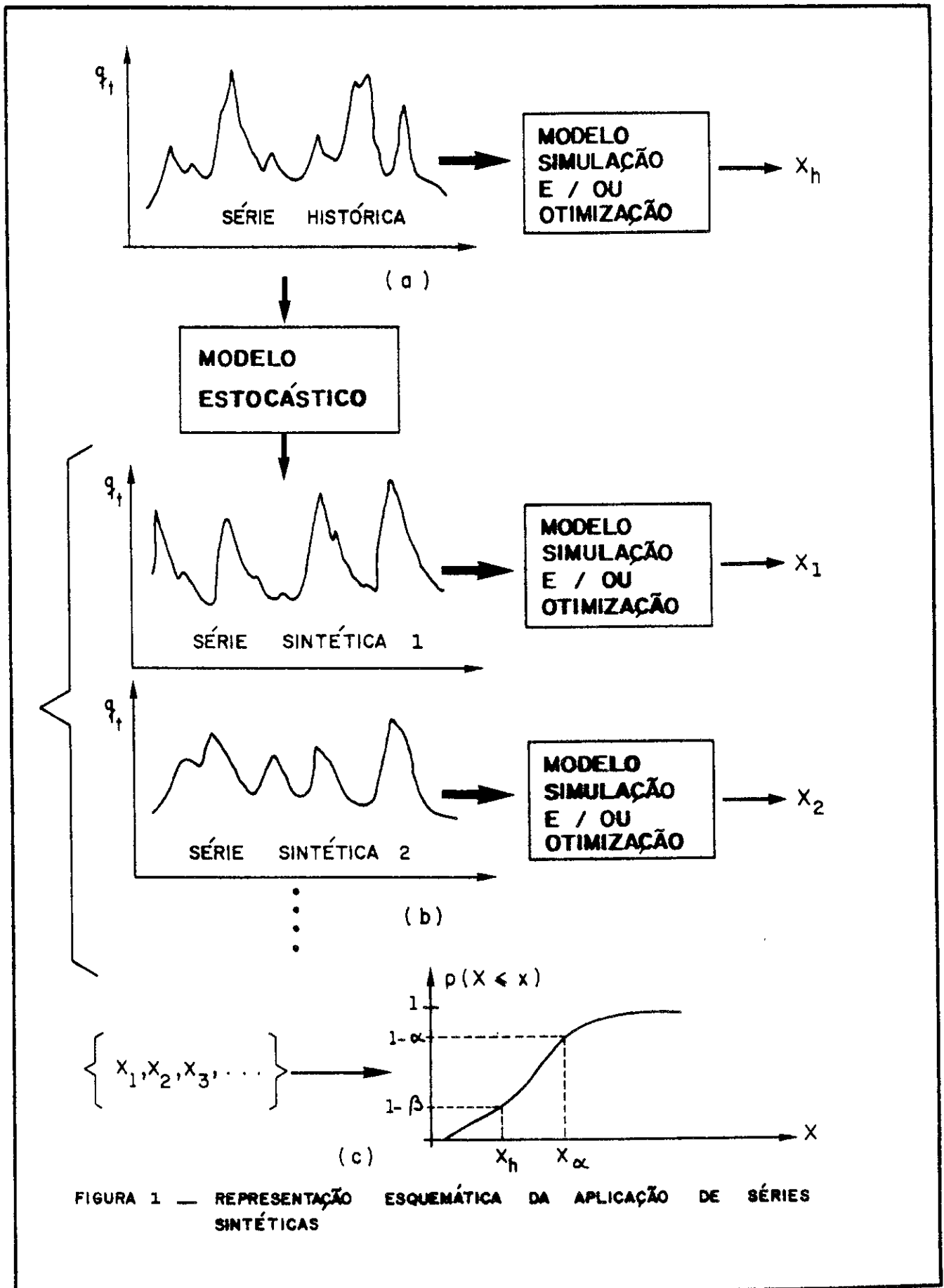


FIGURA 1 — REPRESENTAÇÃO ESQUEMÁTICA DA APLICAÇÃO DE SÉRIES SINTÉTICAS

relações entre as variáveis climáticas hidrológicas e as atividades do sol. A ocorrência das manchas solares em intervalos regulares permitiria prever, segundo estes pesquisadores, ciclos de anos secos e anos úmidos. Outras tentativas de descobrir periodicidades nas vazões anuais estiveram muito em voga na primeira metade do século, quando pesquisadores procuravam ajustar as séries hidrológicas observadas com o emprego das Séries de Fourier. Como se sabe, esta técnica consiste em se ajustar uma série temporal por uma soma de funções periódicas (senos e cossenos) de amplitude e frequência distintas. Para uma particular série hidrológica, algumas destas funções serão mais relevantes do que as outras para a obtenção de um bom ajuste. As correspondentes frequências são chamadas de "periodicidades escondidas", uma vez que não foram identificadas causas físicas que as expliquem. A técnica de se fazer previsões a partir da extrapolação destes sinais periódicos foi bastante utilizada, porém em apenas alguns casos estas previsões tem significado estatístico. Por exemplo, Damazio e Kelman (1981) mostraram ser esta abordagem útil para a previsão das precipitações anuais na cidade de Fortaleza. Em geral, no entanto, o método carece de consistência teórica (Yevjevich, 1979).

Ocorreram no passado notáveis variações climáticas, conforme o atestam as diversas eras glaciais. Existem fatores astronômicos causativos de alterações na distribuição de energia solar sobre a superfície da Terra com periodicidades conhecidas. Só que estas variações se processam numa enorme escala de tempo, absolutamente irrelevante para os estudos de planejamento que se possa fazer para os próximos 50 anos, baseados em dados coletados nos últimos 50 anos (Yevjevich, 1979). No remanescente deste texto a série de vazões anuais será considerada estacionária (as leis probabilística que governam o processo não se alteram com o tempo), ressalvada a intervenção humana sobre a Natureza.

Uma série anual pode ser não estacionária devido à alteração no uso da terra, modificação na seção transversal ou mudança de régua. A melhor maneira de se determinar se uma série é ou não estacionária é através do conhecimento da eventual existência de algum destes fatores. Caso isto não seja possível, é preciso examinar-se a série histórica em busca de algum indício. A inspeção visual da figura 2 não revela irregularidade alguma para a série de vazões anuais em Barra do Piraí. Esta avaliação subjetiva é em geral suficiente. Uma confirmação pode no entanto ser obtida de forma aproximada da seguinte maneira: divide-se a série anual em duas sub-séries conforme mostrado na tabela 1. Para cada sub-série computa-se o valor de média e desvio padrão amostral,  $\hat{\mu}_j$  e  $\hat{\sigma}_j$ ,  $j=1,2$ . Em seguida aplica-se o teste t de student para verificar a hipótese de igualdade entre as duas médias (ver, por exemplo, Gomide, 1976).

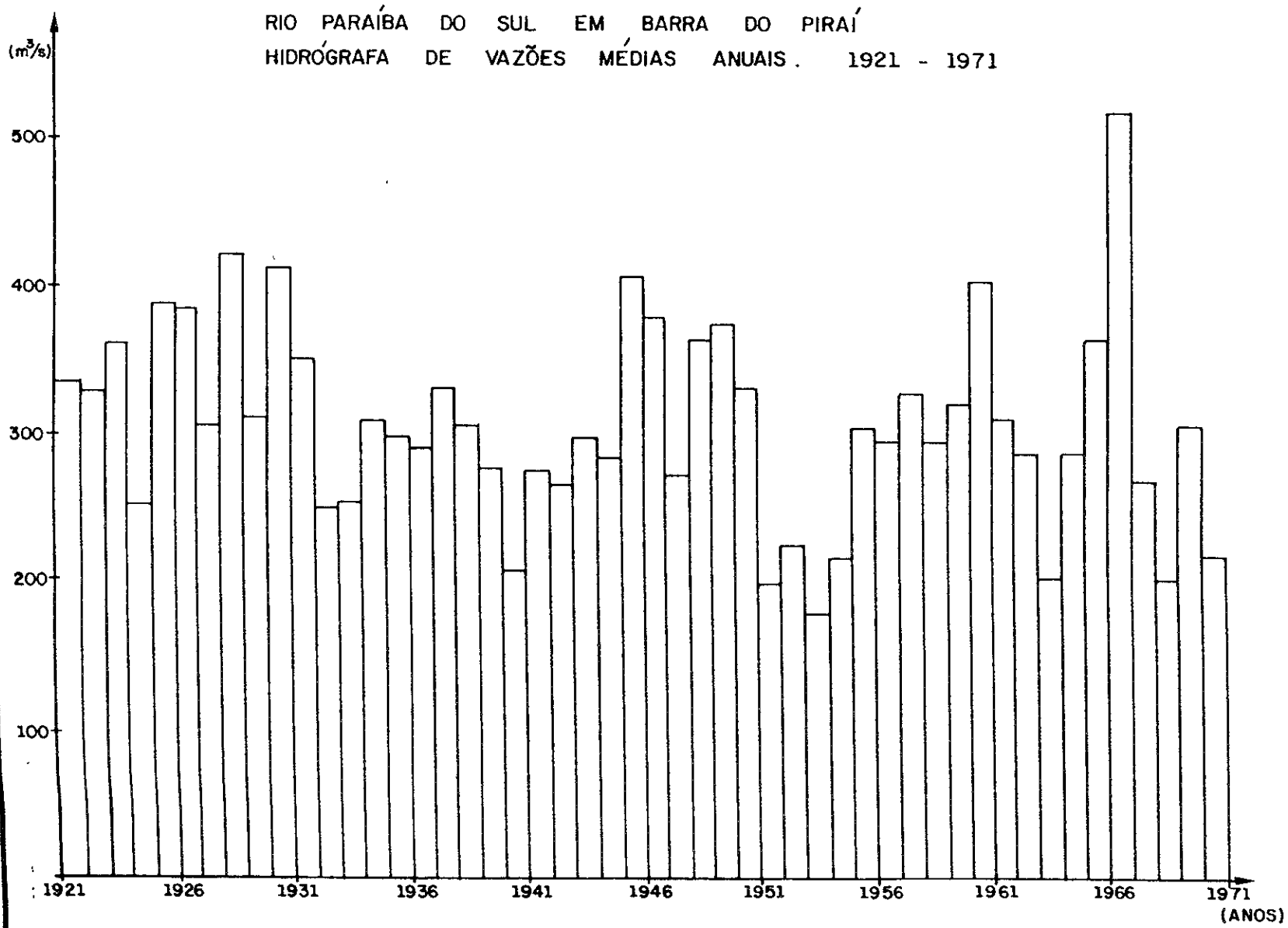


FIGURA 2



Tabela 1 - Série de Vazões Anuais do Rio Paraíba do Sul em Barra do Pirai

SUB-SÉRIE 1		SUB-SÉRIE 2		OBSERVAÇÕES
ANO	VAZÃO (M <sup>3</sup> /S)	ANO	VAZÃO (M <sup>3</sup> /S)	
21 - 22	336	46 - 47	406	1) Média Amostral $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n q_i$
22 - 23	329	47 - 48	380	
23 - 24	361	48 - 49	271	2) Desvio-Padrão Amostral $\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (q_i - \hat{\mu})^2}$
24 - 25	247	49 - 50	361	
25 - 26	387	50 - 51	376	3) Coeficiente de Variação Amostral $\hat{C}_v = \frac{\hat{\sigma}}{\hat{\mu}}$
26 - 27	385	51 - 52	331	
27 - 28	306	52 - 53	197	4) Coeficiente de Assimetria Amostral $\hat{\gamma} = \frac{1}{\hat{\sigma}^3} \cdot \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (q_i - \hat{\mu})^3$
28 - 29	421	53 - 54	223	
29 - 30	313	54 - 55	177	5) Coeficiente de Auto-Correlação Amostral $\hat{\rho} = \frac{\sum_{i=1}^{n-1} (q_i - \hat{\mu})(q_{i+1} - \hat{\mu})}{\sum_{i=1}^n (q_i - \hat{\mu})^2}$
30 - 31	416	55 - 56	214	
31 - 32	355	56 - 57	303	n = Número de anos
32 - 33	253	57 - 58	293	
33 - 34	258	58 - 59	319	
34 - 35	306	59 - 60	282	
35 - 36	299	60 - 61	401	
36 - 37	292	61 - 62	308	
37 - 38	329	62 - 63	286	
38 - 39	304	63 - 64	201	
39 - 40	276	64 - 65	285	
40 - 41	206	65 - 66	362	
41 - 42	274	66 - 67	515	
42 - 43	265	67 - 68	265	
43 - 44	298	68 - 69	200	
44 - 45	273	69 - 70	303	
45 - 46	284	70 - 71	216	

	SUB-SÉRIE 1	SUB-SÉRIE 2	SÉRIE COMPLETA
$\hat{\mu}$	311	299	305
$\hat{\sigma}$	54	81	68
$\hat{C}_v$	0,17	0,27	0,22
$\hat{\gamma}$	0,45	0,63	0,49
$\hat{\rho}$	0,22	0,29	0,26

Para o caso do rio Paraíba do Sul pode-se calcular a estatística  $t$  como:

$$t = (\hat{\mu}_1 - \hat{\mu}_2) \sqrt{\frac{m}{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}} = 0,62$$

onde  $m$  é o tamanho de cada sub-amostra (no caso  $m = 25$ ). Como o valor crítico para o nível de significância de 95% (48 graus de liberdade) é de 2,00, não se pode rejeitar a hipótese da igualdade das duas médias. Um teste análogo poderia ser feito para as duas variâncias,  $\hat{\sigma}_1$  e  $\hat{\sigma}_2$ , empregando-se a estatística  $F$ .

### 1.2.2 Distribuição de Probabilidades

As vazões do Rio Paraíba do Sul podem ser ordenados de forma a produzir a distribuição acumulada empírica, mostrada na figura 3. Para a  $i$ -ésima vazão  $q_i$  (em ordem crescente) associa-se a ordenada dada pela posição de locação. No exemplo adotou-se a posição  $\frac{i}{n+1}$ , mas outras opções poderiam ser selecionadas. Um primeiro gerador de séries sintéticas de vazões anuais poderia ser concebido tomando-se partido do seguinte teorema: se  $Q$  é uma variável aleatória qualquer e  $F_Q(q)$  sua distribuição de probabilidades acumulada  $F_Q(q) = \text{Prob}(Q < q)$ , então  $U = F_Q(Q)$  tem distribuição de probabilidades uniforme (convém observar que se adota a convenção de utilizar letras maiúsculas para denominar variáveis aleatórias e minúsculas para denominar as observações das variáveis aleatórias). Números aleatórios com distribuição uniforme seriam sorteados,  $U_1, U_2, U_3, \dots$ . Para cada valor  $u_j$  se associa uma vazão anual  $q_j$ , conforme figura 3, formando-se desta forma a série  $q_1, q_2, q_3, \dots$ , que poderá ser tão longa quanto se queira. Alguma hipótese será necessária para extrapolar a distribuição acumulada empírica, uma vez que sempre é possível sortear-se um valor para  $U$  maior do que  $n/n+1$ . Alternativamente, pode-se ajustar uma expressão analítica à distribuição acumulada empírica, garantindo-se uma extrapolação automática. Isto equivale a se adotar uma particular distribuição de probabilidades para vazão anual. Diversas distribuições tem sido sugeridas na literatura, sem que nenhuma possa ser indicada para aplicação generalizada. Markovic (1965) fez um estudo com 446 séries de vazões anuais comparando as diversas distribuições com base na bondade de ajuste medida pelo qui-quadrado. Sua conclusão foi de que "todas as distribuições são aplicáveis e nenhuma tem ajuste melhor do que outra em relação a distribuição empírica de precipitação anual ou vazão anual".

Na União Soviética utiliza-se a distribuição gamma de três parâmetros que pode ser justificada com base num estudo chuva-defluvio. (Klimes, 1978). . . Apesar

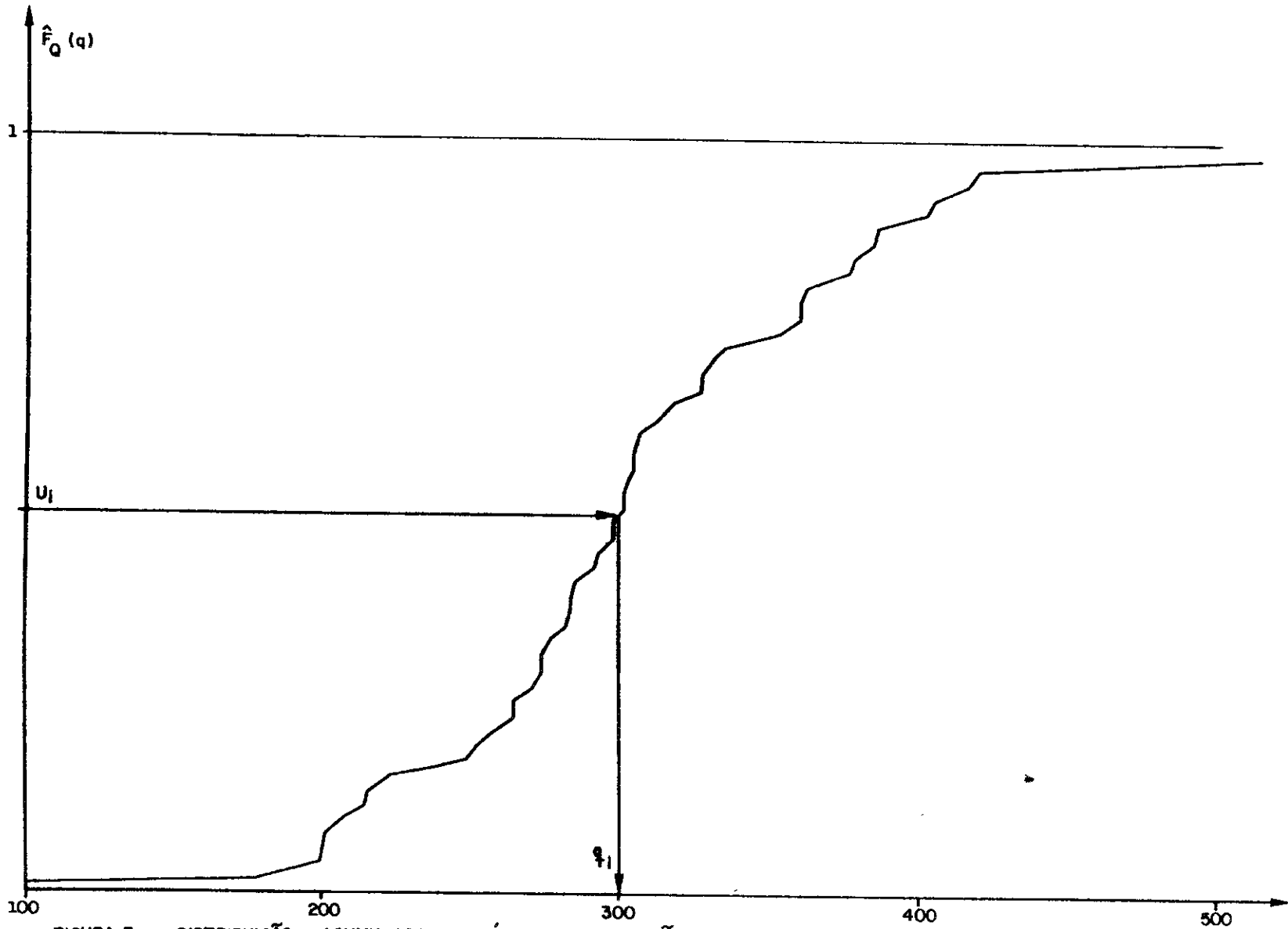


FIGURA 3 — DISTRIBUIÇÃO ACUMULADA EMPÍRICA DAS VAZÕES ANUAIS DO RIO PARAIBA DO SUL EM BARRA DO PIRAI

disto, apenas três distribuições, que aparecem na tabela 2, serão consideradas neste texto. As razões para esta limitação são de ordem operacional : quando o problema for multivariado, isto é, modelagem de diversas séries de vazões, será mais fácil lidar com variáveis que possam ser transformadas em variáveis normalmente distribuídas.

A distribuição normal é frequentemente adotada por ser bastante conhecida e porque vazão anual é o resultado do somatório de um grande número de variáveis aleatórias (365 vazões diárias)<sup>5</sup>. Na realidade a aplicabilidade do teorema do limite central para caso em pauta é bastante questionável, visto que as parcelas não são independentes entre si, nem são igualmente distribuídas. A distribuição normal apresenta o inconveniente de ser definida para valores negativos de  $q$ ; ou seja, existe uma probabilidade de que seja gerada uma vazão anual negativa. No caso do Rio Paraíba do Sul esta probabilidade é menor do que 0,000005; ou seja, pode-se esperar apenas um valor negativo em 200000 anos. Os parâmetros  $\mu$  e  $\sigma$  que aparecem na tabela 2 são estimados respectivamente pela média e desvio padrão amostrais  $\hat{\mu}$  e  $\hat{\sigma}$ , cujas expressões aparecem na tabela 1. O sorteio de números aleatórios normalmente distribuídos é tarefa simples e usualmente rotinas para este fim estão disponíveis em qualquer sistema computacional.

A série sintética  $q_1^*, q_2^*, q_3^*, \dots, q_m^*$  assim obtida será tal que  $\lim_{m \rightarrow \infty} \hat{\mu}^* = \mu$  e  $\lim_{m \rightarrow \infty} \hat{\sigma}^* = \sigma$ . Isto é, a média e desvio padrão amostral da série sintética tenderá aos correspondentes valores da série histórica, a medida que o tamanho da série sintética cresça. Diz-se que este modelo preserva estes dois momentos. Isto não é o mesmo que afirmar que o modelo irá reproduzir a média e o desvio padrão populacionais,  $\mu$  e  $\sigma$ . Na realidade estes valores serão sempre desconhecidos.

A distribuição log-normal de dois parâmetros é definida unicamente para valores positivos de  $q$ . Portanto, é impossível gerar-se uma vazão anual negativa. Se  $Q$  tem distribuição log-normal de dois parâmetros, então  $\log Q$  tem distribuição normal; ou seja, a distribuição log-normal de dois parâmetros é indicado para a hipótese de que a vazão anual seja o resultado de um grande número de fatores que in

---

5

O teorema do limite central garante que, sob condições bastante gerais, a soma de um grande número de variáveis aleatórias independentes e igualmente distribuídas tem distribuição normal, não importa qual seja a distribuição de probabilidades de cada parcela.

Tabela 2 - Distribuições de Probabilidades

1) NORMAL (N)  $-\infty < q < +\infty$

$$f_Q(q) = \frac{1}{\sqrt{2\pi} \sigma} \cdot \exp \left[ -\frac{1}{2} \left( \frac{q - \mu}{\sigma} \right)^2 \right]$$

$$\text{MÉDIA} = \mu$$

$$\text{VARIÂNCIA} = \sigma^2$$

2) LOG-NORMAL DE 2 PARÂMETROS (LN2),  $q > 0$

$$f_Q(q) = \frac{1}{\sqrt{2\pi} \sigma_n} \cdot \exp \left[ -\frac{1}{2} \left( \frac{\ln(q) - \mu_n}{\sigma_n} \right)^2 \right]$$

$$\text{MÉDIA} = \exp \left[ \frac{\sigma_n^2}{2} + \mu_n \right] = E(Q)$$

$$\text{VARIÂNCIA} = \exp [ 2 (\sigma_n^2 + \mu_n) ] - \exp [ \sigma_n^2 + 2\mu_n ] = \text{var}(Q)$$

onde

$$\mu_n = E (\ln(q))$$

$$\sigma_n^2 = E [ (\ln(q) - E (\ln(q)))^2 ]$$

3) LOG-NORMAL DE 3 PARÂMETROS (LN3)  $q > a$

$$f_Q(q) = \frac{1}{\sqrt{2\pi} \sigma_n} \exp \left[ -\frac{1}{2} \left( \frac{\ln(q-a) - \mu_n}{\sigma_n} \right)^2 \right]$$

$$\text{MÉDIA} = a + \exp \left[ \frac{\sigma_n^2}{2} - \mu_n \right]$$

$$\text{VARIÂNCIA} = \exp [ 2(\sigma_n^2 + \mu_n) ] - \exp [ \sigma_n^2 + 2\mu_n ]$$

onde

$$\mu_n = E [ \ln(q-a) ]$$

$$\sigma_n^2 = E [ (\ln(q-a) - \mu_n)^2 ]$$

teragem de forma multiplicativa. Uma possível maneira de se estimar os parâmetros desta distribuição seria repetir o procedimento adotado no caso da distribuição normal para a série  $\log q_1, \log q_2, \log q_3, \dots$ . Desta forma, os parâmetros seriam estimados como

$$\hat{\mu}_n = \frac{1}{n} \sum_{t=1}^n \log q_t \quad e \quad \hat{\sigma}_n = \sqrt{\frac{1}{n-1} \sum_{t=1}^n (\log q_t - \mu_n)^2}.$$

Números aleatórios normalmente distribuídos seriam gerados,  $\log q_1^*, \log q_2^*, \dots$  que quando exponenciados resultariam na série sintética  $q_1^*, q_2^*, q_3^*, \dots$ . Ocorre que neste caso  $\lim_{m \rightarrow \infty} \hat{\mu}^* \neq \hat{\mu}$  e  $\lim_{m \rightarrow \infty} \hat{\sigma}^* \neq \hat{\sigma}$ , ou seja, a preservação dos dois primeiros momentos na série dos logaritmos não garante a preservação destes momentos na série original. As relações entre estes dois conjuntos de parâmetros são expressas pelas equações que aparecem na tabela 2, que quando explicitadas para  $\mu_n$  e  $\sigma_n$  resultam nas seguintes equações:

$$\sigma_n^2 = \ln(\psi) \tag{1a}$$

$$\mu_n = \frac{1}{2} \ln(\sigma^2 / (\psi^2 - \psi)) \tag{1b}$$

$$\psi = 1 + (\sigma/\mu)^2 \tag{1c}$$

Como usualmente o que se pretende é preservar os momentos da série histórica, e não dos logaritmos, o que se faz é resolver o sistema de equações (1), substituindo-se os valores de  $\mu$  e  $\sigma$  respectivamente por  $\hat{\mu}$  e  $\hat{\sigma}$ . É comum no entanto que ambas as abordagens de estimação de  $\mu_n$  e  $\sigma_n$  resultem em valores numéricos muito próximos. No caso do Rio Paraíba do Sul em Barra do Piraí, o emprego da equação (1) resulta em  $\mu_n = 5.6957$  e  $\sigma_n = 0.2211$ , ao passo que a estimação com a série dos logaritmos  $\mu_n = 5.6959$  e  $\sigma_n = 0.2258$ .

A figura 4 foi apresentada por Klemes (1978) com base em dados coletados por Yevjevich (1963). Ela mostra a relação entre os coeficientes de assimetria ( $\gamma$ ) e de variação ( $\sigma/\mu$ ) para 140 rios espalhados pelo mundo. Como se pode notar, a maior parte dos pontos ficam localizada entre o eixo vertical (assimetria nula; possivelmente distribuição normal) e a reta  $\gamma = (\frac{\sigma}{\mu})^3 + 3(\frac{\sigma}{\mu})$ , que é a relação pertinente à distribuição log-normal de dois parâmetros. Uma possível explicação para este fato foi dada originalmente por Chow (1954) e repetida por Klemes (1978). Para estes autores, os fenômenos hidrológicos podem ser interpretados como sendo o

resultado de um conjunto não enumerável de causas físicas que interagem tanto de forma multiplicativa como aditiva. Uma vez que os efeitos multiplicativos estão associados à distribuição log-normal de dois parâmetros (LN2) enquanto que os efeitos aditivos estão associados à distribuição normal (N), pode-se esperar que a verdadeira distribuição esteja situada entre os dois tipos limites (N e LN2). A distribuição log-normal de três parâmetros (LN3) (tabela 2) pode satisfazer esta propriedade : quando o parâmetro a (deslocamento) tende para  $-\infty$  a LN3 tende para N ; quando  $a=0$  a distribuição LN3 equivale a LN2.

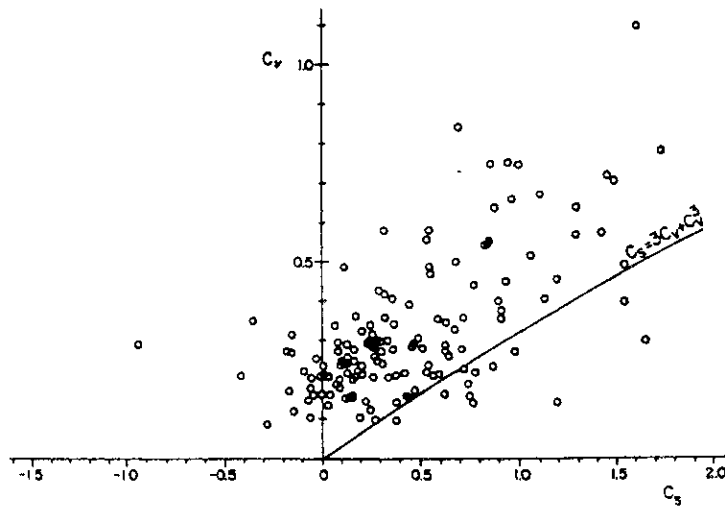


Figura 4 - Relação entre o Coeficiente de Variação e o Coeficiente de Assimetria para Séries de Vazões Anuais de 140 rios de todo o Mundo.

A distribuição LN3 é definida para valores de  $q > a$ . Se  $a < 0$ , existe uma probabilidade de que seja gerada alguma vazão anual negativa. A estimação dos parâmetros pode ser feita tanto baseado em valores amostrais da média, mediana e desvio-padrão como baseado média, desvio padrão e assimetria. Burges, Letteumaier e Bates (1975) demonstraram a superioridade da segunda alternativa. Portanto com o propósito de preservar os momentos da série histórica, e não de série de logaritmos, os três parâmetros  $\mu_n$ ,  $\sigma_n$  e  $a$  devem ser estimados resolvendo-se as equações abaixo onde  $\mu$ ,  $\sigma$  e  $\gamma$  devem ser substituídos pelos respectivos valores amostrais,  $\hat{\mu}$ ,  $\hat{\sigma}$  e  $\hat{\gamma}$ .

$$\phi = \left[ \left( 1 + \frac{\gamma^2}{2} \right) + \left( \gamma^2 + \frac{\gamma^4}{4} \right)^{1/2} \right]^{1/3} + \left[ \left( 1 + \frac{\gamma^2}{2} \right) - \left( \gamma^2 + \frac{\gamma^4}{4} \right)^{1/2} \right]^{1/3} - 1 \quad (2a)$$

para  $\gamma > 0$

$$\sigma_n^2 = \ln(\phi) \quad (2b)$$

$$\mu_n = \frac{1}{2} \ln (\sigma^2 / (\phi^2 - \phi)) \quad (2c)$$

$$a = \mu - [\sigma^2 / (\phi - 1)]^{1/2} \quad (2d)$$

A figura 5 mostra as três distribuições acumuladas de probabilidades, quando ajustadas para as vazões do rio Paraíba do Sul. Como se vê todas as três parecem aderir convenientemente à distribuição acumulada empírica.

### 1.2.3 Persistência

A figura 6 mostra o gráfico dos pares de pontos  $(q_t, q_{t+1})$  do Rio Paraíba do Sul, onde  $q_t$  é a vazão anual do ano  $t$ . Pode-se perceber que se o ano  $t$  foi seco ( $q_t$  pequeno) o ano  $t+1$  tenderá também a ser seco. E vice-versa. A este fenômeno dá-se o nome de persistência ou de dependência temporal. O modelo visto na seção anterior não poderia ser adequado porque adota a hipótese de aleatoriedade para a série anual do rio Paraíba do Sul. Isto é, adota a hipótese de que a distribuição de probabilidades de  $Q_{t+1}$  independe dos valores observados no passado  $\{q_t, q_{t-1}, q_{t-2}, \dots\}$ . Na realidade a figura 6 mostra que talvez esta hipótese se já falsa, uma vez que  $Q_{t+1}$  parece depender pelo menos de  $q_t$ . Com base nesta constatação, pode-se sugerir o seguinte processo estocástico (em contraposição ao processo aleatório) para modelar as séries anuais:

$$Q_{t+1} = \mu + \rho (Q_t - \mu) + \sigma \sqrt{1 - \rho^2} Z_{t+1} \quad (3)$$

onde

$$E(Z_{t+1}) = 0, \quad \text{var}(Z_{t+1}) = 1;$$

$$Q_t \text{ e } Z_{t+1} \text{ são independentes.}$$

$$\text{e } Z_t \text{ e } Z_\tau \text{ são independentes, } t \neq \tau$$

A aplicação do operador  $E(\cdot)$ , valor esperado, em ambos os lados da equação 3 resulta em

$$E(Q_{t+1}) = \mu + \rho (E(Q_t) - \mu) + \sigma \sqrt{1 - \rho^2} E(Z_{t+1}) = \rho E(Q_t) + \mu(1 - \rho)$$

como o processo é estacionário  $E(Q_{t+1}) = E(Q_t)$  e portanto



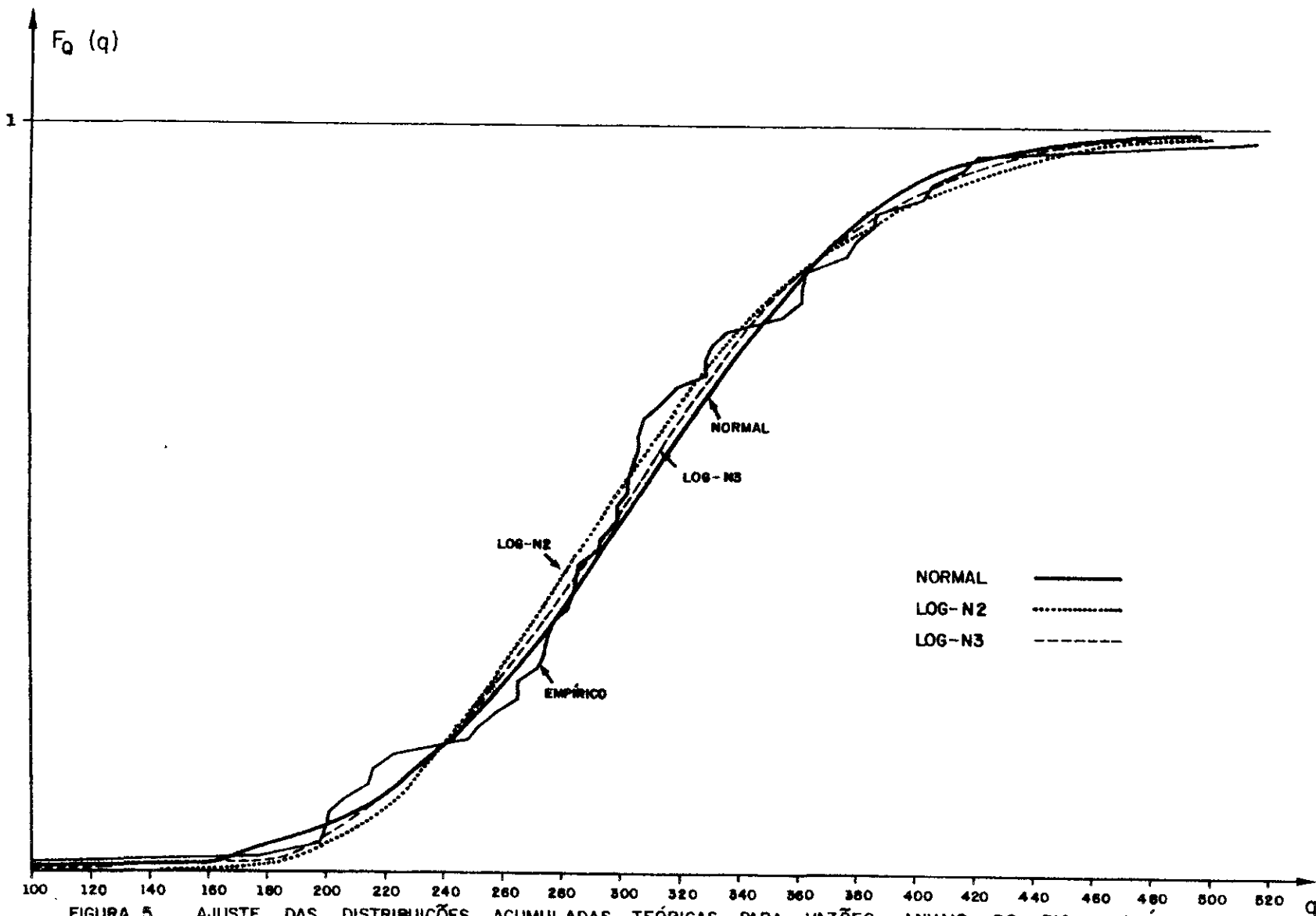


FIGURA 5 — AJUSTE DAS DISTRIBUIÇÕES ACUMULADAS TEÓRICAS PARA VAZÕES ANUAIS DO RIO PARAÍBA DO SUL EM BARRA DO PIRAI

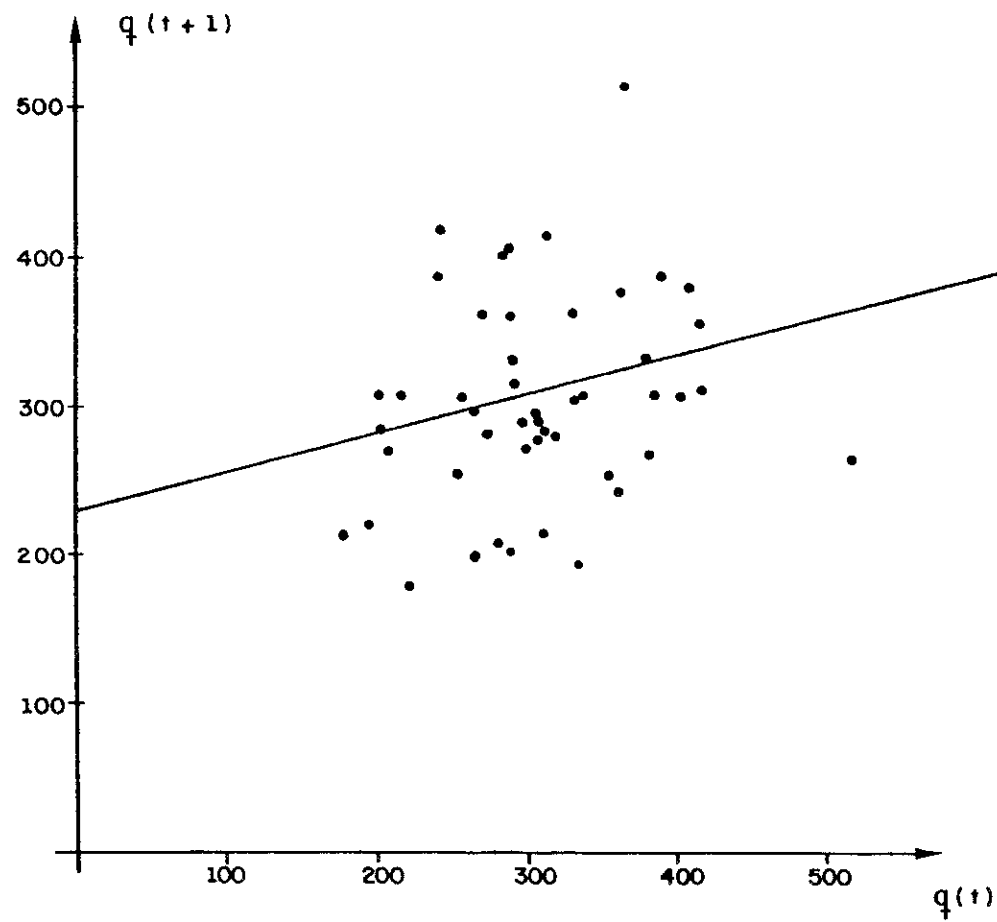


FIGURA 6 — RELAÇÃO ENTRE VAZÕES DE ANOS CONSECUTIVOS PARA RIO PARAÍBA DO SUL EM BARRA DO PIRAI

$$E(Q_{t+1}) = \mu \quad (4)$$

Analogamente a aplicação do operador  $\text{var}(\cdot)$ , variância, resulta em

$$\text{var}(Q_{t+1}) = \rho^2 \text{var}(Q_t) + \sigma^2 (1 - \rho^2) \text{var}(Z_{t+1}) + 2\rho\sigma \sqrt{1 - \rho^2} \text{cov}(Q_t, Z_{t+1})$$

onde  $\text{cov}(Q_t, Z_{t+1})$  significa covariância entre  $Q_t$  e  $Z_{t+1}$ , que é nula porque estas duas variáveis aleatórias são independentes. Portanto, lançando-se mão que  $\text{var}(Q_{t+1}) = \text{var}(Q_t)$  tem-se que

$$\text{var}(Q_{t+1}) = \sigma^2 \quad (5)$$

A regressão de  $Q_{t+1}$  em  $q_t$ , isto é, a expressão do valor esperado da vazão no ano  $t+1$  conhecida a vazão no ano  $t$  é dada por

$$E(Q_{t+1} | Q_t = q_t) = E(Q_{t+1} | q_t) = \mu + \rho(q_t - \mu) + \sigma \sqrt{1 - \rho^2} E(Z_{t+1}) = \mu + \rho(q_t - \mu) \quad (6)$$

A variância condicionada, as vezes também chamada de variância não explicada pela regressão, ou variância residual, é dada por:

$$\text{var}(Q_{t+1} | q_t) = \sigma^2(1 - \rho^2) \quad \text{var}(Z_{t+1}) = \sigma^2 (1 - \rho^2) \quad (7)$$

Multiplicando ambos os lados da equação (3) por  $Q_t$  e aplicando-se o operador  $E(\cdot)$  resulta

$$E(Q_{t+1} Q_t) = \mu E(Q_t) + \rho E(Q_t^2) - \mu E(Q_t) + \sigma \sqrt{1 - \rho^2} E(Z_{t+1} Q_t)$$

Utilizando o fato de que  $\text{var}(Q_t) = E(Q_t^2) - E^2(Q_t)$ , as relações (4) e (5) e a independência entre  $Q_t$  e  $Z_{t+1}$ , resulta

$$E(Q_{t+1} Q_t) = \mu^2 + \rho(\sigma^2 + \mu^2 - \mu^2) + \sigma \sqrt{1 - \rho^2} \mu E(Z_{t+1}) = \mu^2 + \rho\sigma^2$$

Logo a correlação entre  $Q_{t+1}$  e  $Q_t$ , também chamada de auto-correlação, é dada por

$$\text{corr}(Q_{t+1} Q_t) = \frac{E(Q_{t+1} Q_t) - E(Q_t) E(Q_{t+1})}{\sqrt{\text{var}(Q_t) \text{var}(Q_{t+1})}} = \frac{\mu^2 + \rho\sigma^2 - \mu^2}{\sigma^2} = \rho \quad (8)$$

O modelo expresso pela equação 3 implica que a vazão num ano qualquer  $Q_{t+1}$ , é uma variável aleatória cuja distribuição de probabilidades marginal tem média  $\mu$  e desvio padrão  $\sigma$  (ambos os valores de população). Entretanto quando a vazão no ano imediatamente anterior  $q_t$  é conhecida, então  $Q_{t+1}$  continua sendo uma variável aleatória, só que com média situada na reta que aparece na figura 5 e desvio padrão (medida de dispersão) menor, dado por  $\sigma \sqrt{1 - \rho^2}$  (o coeficiente de auto-correlação situa-se entre -1 e +1). Isto é, a distribuição de probabilidades condicionada tem média  $\mu + \rho(q_t - \mu)$  e desvio padrão  $\sigma \sqrt{1 - \rho^2}$ . Convém observar que se  $\rho = 0$  não há persistência alguma e o processo estocástico é na realidade aleatório, que foi assunto examinado na seção anterior. Por outro lado se  $\rho = 1$  a variância residual é nula, o que significa que todos os pontos na figura 5 cairiam rigorosamente sobre a reta de regressão. Ou seja, a série anual  $q_1, q_2, q_3, \dots$  seria formada por valores iguais.

Os parâmetros do modelo  $\mu, \sigma$  e  $\rho$  podem ser estimados a partir da média, desvio padrão e auto-correlações amostrais,  $\hat{\mu}, \hat{\sigma}$  e  $\hat{\rho}$ , cujas expressões se encontram na tabela 1. A geração de uma série sintética se faz da seguinte maneira:

- a) escolhe-se um valor para  $q_1$  e faz-se  $t=0$
- b)  $t = t+1$
- c) sorteia-se um número aleatório  $Z_{t+1}$  de média zero e desvio padrão unitário
- d)  $q_{t+1} = \hat{\mu} + \hat{\rho} (q_t - \hat{\mu}) + \hat{\sigma} \sqrt{1 - \hat{\rho}^2} z_{t+1}$
- e) vá para (b)

Se a distribuição de probabilidade de  $Z_{t+1}$  for normal padrão, então tanto a distribuição marginal de  $Q_{t+1}$  como a distribuição condicionada de  $Q_{t+1}$  conhecido  $q_t$  serão também normais. Desta forma é possível ampliar-se um dos modelos descritos na seção anterior - quando se assumia que a série histórica anual é uma realização de um processo aleatório, com distribuição marginal normal - para o modelo descrito acima, em que se conserva a mesma hipótese quanto à distribuição marginal, mas se inclui o efeito da persistência.

Quanto ao valor inicial  $q_1$ , há duas possibilidades :

- (i) caso se pretenda produzir diversas séries sintéticas para fazer simulações com relação ao futuro imediato,  $q_1$  deve ser igual à última

observação da série histórica;

- (ii) caso o estudo seja sobre o futuro longínquo,  $q_1$  deverá ser sorteado de uma distribuição normal de média  $\hat{\mu}$  e desvio padrão  $\hat{\sigma}$ , para cada série sintética.

Caso se pretenda adotar um modelo que considere a persistência e simultaneamente produza vazões com distribuição marginal de probabilidades LN3, modifica-se a equação (3) da seguinte forma:

$$\log(Q_{t+1} - a) = \mu_n + \rho_n (\log(Q_t - a) - \mu_n) + \sigma_n^2 \sqrt{1 - \rho_n^2} Z_{t+1} \quad (9)$$

onde

$Z_{t+1}$  é uma variável aleatória com distribuição normal padrão.

A estimação de  $\rho_n$  pode ser feita calculando-se o coeficiente de auto-correlação amostral, segundo expressão contida na tabela 1, para a série  $\{\log(q_1 - \hat{a}), \log(q_2 - \hat{a}), \dots, \log(q_m - \hat{a})\}$ , sendo  $\hat{a}$  determinado pela resolução do sistema de equações (2). Neste caso, analogamente ao que foi visto no item anterior,  $\lim_{m \rightarrow \infty} \hat{\rho}^* \neq \hat{\rho}$ . Isto é, as séries sintéticas não preservarão o coeficiente de auto-correlação amostral da série histórica. Alternativamente pode-se obter esta propriedade estimando-se  $\rho_n$  por

$$\rho_n = \frac{\ln(1 + \rho(\phi - 1))}{\ln \phi} \quad (10)$$

onde  $\phi$  e  $\rho$  devem ser substituídos pelos respectivos valores amostrais.

O modelo expresso pela equação (3) é chamado de auto-regressivo linear de ordem 1 — AR(1). O termo deriva do fato de que  $E(Q_{t+1} | Q_t = q_t)$  é uma função linear de  $q_t$  e somente de  $q_t$  (se também dependesse de  $q_{t-1}$ ,  $q_{t-2}$ , ... e de  $q_{t-k+1}$ , seria de ordem  $k$ ). Pode-se demonstrar que de acordo com a equação (3)

$$\text{corr}(Q_{t+1}, Q_{t-k+1}) = \rho^k \quad (11)$$

para  $k = 1, 2, \dots$

A figura 7 mostra o gráfico de  $\text{corr}(Q_{t+1}, Q_{t-k+1})$  em função de  $k$ , chamado de correlograma. A curva teórica é obtida da equação (11) e a amostral pela expressão

$$\hat{\rho}_k = \frac{1}{n-k} \frac{\sum_{t=1}^{n-k} (q_t - \hat{\mu})(q_{t+k} - \hat{\mu})}{\sum_{t=1}^n (q_t - \hat{\mu})^2} \quad k = 1, 2, \dots \quad (12)$$

Convém observar que quando  $k=1$  a equação (12) já foi utilizada na tabela 1. A comparação entre a curva teórica e a amostral permite que seja verificado, em termos probabilísticos, se o modelo expresso pela equação (3) é de fato uma boa representação do processo estocástico em tela. Existem maneiras objetivas de se fazer esta comparação, assim como existem inúmeros procedimentos para se propor modelos alternativos. Estas questões extrapolam o escopo do presente trabalho. O leitor interessado deve consultar, por exemplo Yevjevich (1972) e/ou Salas et al (1981).

O modelo expresso pela equação (3) é também chamado de "markoviano": ter informação sobre o valor de  $q_t$  é tudo o que é necessário para se fazer prognósticos quanto a  $Q_{t+1}$ , sendo inútil conhecer-se também  $q_{t-1}$ ,  $q_{t-2}$ , .... Colocado de outra maneira, pode-se demonstrar que

$$\text{corr}(Q_{t+1}, Q_{t-k} | Q_t = q_t) = 0 \quad (13)$$

para  $k = 1, 2, \dots$

#### 1.2.4 Modelagem Multivariada

Dois aproveitamentos hídricos próximos são em geral atingidos simultaneamente por baixas (ou altas) vazões ocasionando secas (ou cheias) regionais. Pode-se modelar este fenômeno através do emprego de equações que garantam a preservação das correlações amostrais, correspondentes a vazões de mesmo ano em locais distintos (Matalas, 1967). Isto é, cuidar para que os valores de  $\text{corr}(Q_t(i), Q_t(j))$  nas séries históricas e nas sintéticas sejam estatisticamente indistinguíveis para  $\forall i, j$ , onde  $i$  e  $j$  são os índices dos locais de aproveitamentos hídricos. De forma mais simples, pode-se preservar apenas as correlações entre pares de "ruídos",  $\text{corr}(Z_t(i), Z_t(j)) = E(Z_t(i), Z_t(j))$  (o ruído de cada local é padronizado).

Considere o seguinte modelo

$$\tilde{Z}_t = F \tilde{Y}_t \quad (14)$$

$\dots = \phi_1 z + \phi_2$   
 $\dots = \phi_0$

onde  $\tilde{Z}_t^T = [Z_t(1), Z_t(2), \dots, Z_t(k)]$  e  $\tilde{Y}_t$  é um vetor com k variáveis normais padronizadas independentes e B é uma matriz de parâmetros de dimensão kxk k é o número de locais.

Pode-se provar que

$$\text{cov}(\tilde{Z}_t) = F \text{cov}(\tilde{Y}_t) F^T$$

e como  $\text{cov}(\tilde{Y}_t)$  é a matriz identidade,

$$\text{cov}(\tilde{Z}_t) = FF^T \quad (15)$$

As equações 14 e 15 revelam que é possível produzir um vetor de variáveis aleatórias associado a uma determinada matriz de covariância a partir da combinação linear de variáveis aleatórias independentes entre si, desde que a matriz F atenda à equação 15.

A estimação de  $\text{cov}(\tilde{Z}_t)$  pode ser feita obtendo-se a série histórica de ruídos  $\{z_1(i), z_2(i), \dots, z_t(i), \dots, z_t(k)\}$  a partir da resolução da equação (3) (ou (9)) onde os valores de  $Q_{t+1}$  e  $Q_t$  são substituídos pelos valores observados na série histórica de vazões e os parâmetros são substituídos pelas suas respectivas estimativas numéricas. A estimativa do elemento (i,j) de  $\text{cov}(\tilde{Z}_t)$  é neste caso

$$\widehat{\text{cov}(\tilde{Z}_t)}_{i,j} = \frac{1}{n} \sum_{t=1}^n z_t(i) z_t(j) \quad (16)$$

A diagonal de  $\text{cov}(\tilde{Z}_t)$  deve ser unitária, na medida que  $\text{var}(Z_t(i)) = 1, \forall i$ .

A matriz F é estimada substituindo-se o lado esquerdo da equação 15 pela estimativa da matriz, expressa em (16). Existem no entanto muitas matrizes F que satisfazem a equação (15). Uma maneira de se encontrar uma destas soluções é de se arbitrar que F seja triangular, o que resulta num sistema linear de fácil solução (Young, 1968).

Em resumo, a geração multivariada de vazões anuais pode ser obtida produzindo-se pela equação (14) ruídos para os k locais de interesse. Individualmente, para cada local, a persistência é garantida utilizando-se estes ruídos na equação (3) (ou (9)).

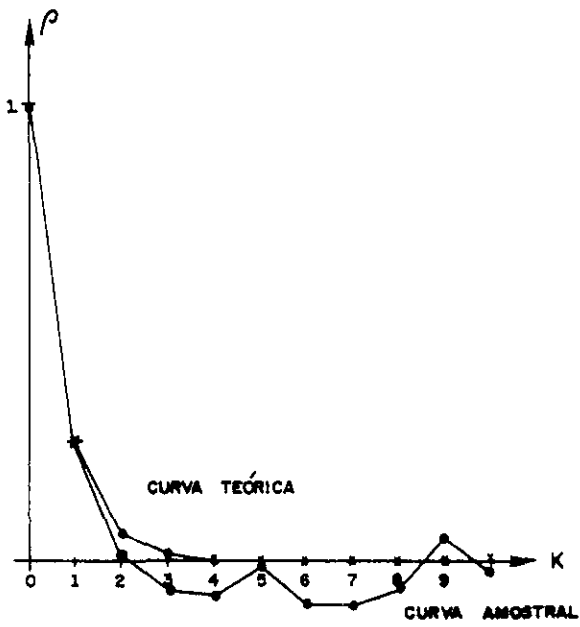


FIG. 7 — CORRELOGRAMO PARA VAZÕES ANUAIS DO RIO PARAÍBA DO SUL EM BARRA DO PIRAI

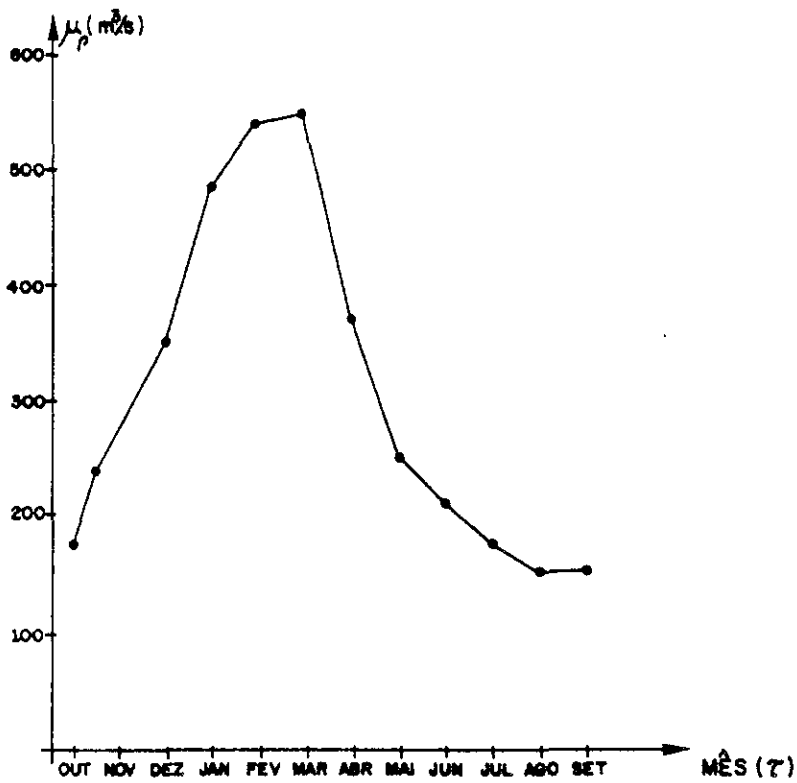


FIG. 8 — VARIACÃO SAZONAL DA MÉDIA DAS VAZÕES MENSASIS DO RIO PARAÍBA DO SUL EM BARRA DO PIRAI



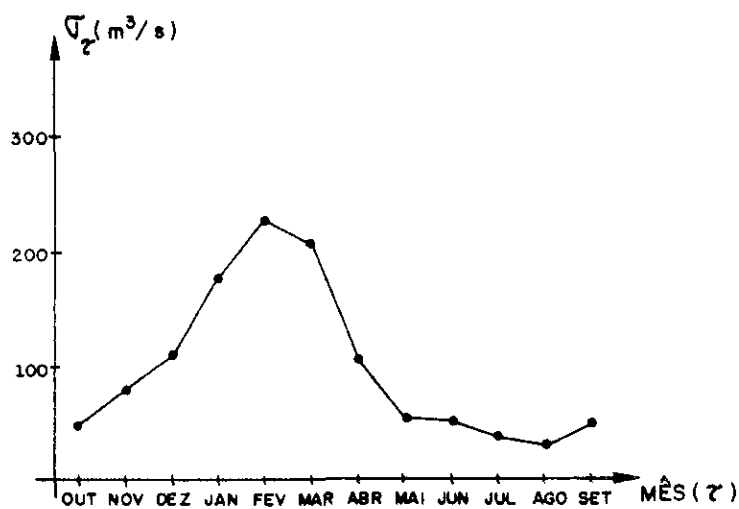


FIG 9 — VARIACÃO SAZONAL DO DESVIO PADRÃO DAS VAZÕES MENSIS DO RIO PARAÍBA DO SUL EM BARRA DO PIRAI

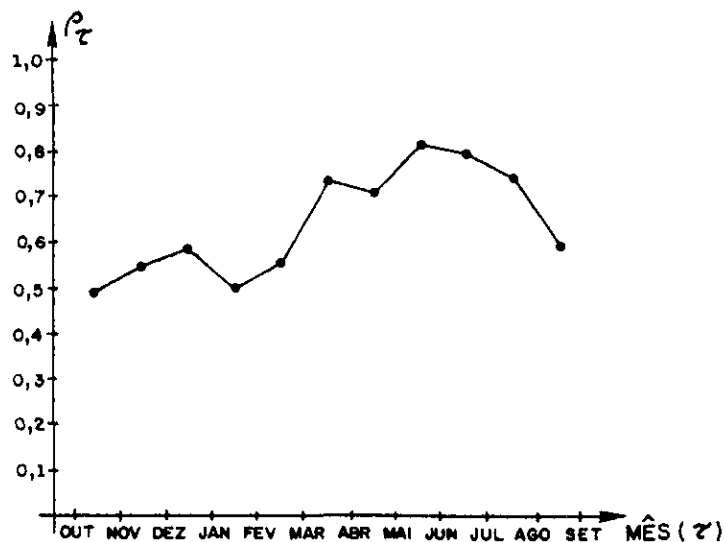


FIG 10 — VARIACÃO SAZONAL DO COEFICIENTE DE CORRELAÇÃO DE VAZÕES MESES CONSECUTIVOS DO RIO PARAÍBA DO SUL EM BARRA DO PIRAI

### 1.3 Modelagem de Vazões Mensais

Os modelos de simulação e/ou otimização empregados para o planejamento e operação de sistemas de recursos hídricos utilizam com frequência séries de vazões com intervalo de discretização mensal.

Ao contrário do processo anual, vazões mensais não podem mais ser consideradas como uma realização de um processo estacionário. A figura 8 mostra a evolução da média das vazões, para cada mês do ano, do Rio Paraíba do Sul em Barra do Piraí. Um simples exame comprova o caráter sazonal desta estatística. O mesmo se dá com o desvio-padrão e coeficiente de auto-correlação, respectivamente nas figuras 9 e 10.

O modelo estacionário expresso pela equação (3) (ou (9)), não pode mais ser empregado, no sentido que para vazões mensais não é válido, por exemplo, a igualdade  $E(Q_t) = E(Q_{t+1})$ . Uma adaptação da equação (3) pode, no entanto, ser conseguida através do chamado modelo de Thomas-Fiering:

$$X_{\tau+1,t} = \mu_{\tau+1} + \frac{\rho_{\tau} \sigma_{\tau+1}}{\sigma_{\tau}} (X_{\tau,t} - \mu_{\tau}) + \sigma_{\tau+1} \sqrt{1 - \rho_{\tau}^2} Z_{\tau+1,t} \quad (17)$$

onde

$X_{\tau,t}$  é a vazão no mês  $\tau$ , ano  $t$

$\tau = 1, 2, \dots, 12$  é o índice do mês

$t = 1, 2, \dots$  é o índice de ano

Adota-se a convenção "circular" de que se  $\tau = 12$  então

$$X_{\tau+1,t} = X_{1,t+1}, \mu_{\tau+1} = \mu_1 \text{ e } \sigma_{\tau+1} = \sigma_1$$

A manipulação da equação (17), nos mesmos moldes da discussão efetuada após a apresentação da equação (3), revelaria uma série de resultados que generalizam as equações (4) a (8) e (23):

$$E(X_{\tau,t}) = \mu_{\tau} \quad (18)$$

$$\text{var}(X_{\tau,t}) = \sigma_{\tau}^2 \quad (19)$$

$$E(X_{\tau+1,t} | X_{\tau,t} = q_{\tau,t}) = \mu_{\tau+1} + \frac{\rho_{\tau} \sigma_{\tau+1}}{\sigma_{\tau}} (q_{\tau,t} - \mu_{\tau}) \quad (20)$$

$$\text{var}(X_{\tau+1,t} | X_{\tau,t} = q_{\tau,t}) = \sigma_{\tau+1}^2 (1 - \rho_{\tau}^2) \quad (21)$$

$$\text{corr}(X_{\tau+1,t}, X_{\tau,t}) = \rho_{\tau} \quad (22)$$

$$\text{corr}(X_{\tau+1,t}, X_{\alpha,t-\ell}) = \underbrace{\rho_{\tau} \rho_{\tau-1} \cdots \rho_1 \rho_{12} \rho_{11} \rho_{10} \cdots}_{\text{produto de K fatores}} \quad (23)$$

onde

$$K = 12\ell + \tau - \alpha + 1$$

A extensão do modelo expresso pela equação (17) para o caso multivariado poderia ser conseguida, de forma aproximada, pela preservação das correlações entre pares de ruídos  $(Z_{\tau,t}(i), Z_{\tau,t}(j))$  correspondentes aos locais  $i$  e  $j$ , conforme explicado no item 1.2.4.

A figura 11 mostra o gráfico de  $(\text{corr}(X_{\tau+1,t}, X_{\alpha,t-\ell}))$  calculado pela equação (23), para  $\tau+1$  equivalente ao mês de setembro, e a correspondente curva amostral. Como se vê, o ajuste deixa muito a desejar: a curva teórica decai muito mais rapidamente do que a amostral. Pode-se mostrar que uma das consequências deste fenômeno é de que a auto-correlação das vazões anuais sintéticas, produzidas a partir de séries mensais geradas pela equação (17), tenderá a ser inferior a auto-correlação amostral calculada a partir da série histórica de vazões anuais. O efeito disto pode ser grave quando se modela sistemas de recursos hídricos com reservatórios capazes de exercer uma regularização plurianual. Uma possível solução para este problema é a de se empregar um modelo auto-regressivo sazonal de ordem maior do que 1. Alternativamente, pode-se procurar um método que gere inicialmente as vazões anuais e depois desagregue cada valor anual em componentes mensais.

O modelo de desagregação proposto por Mejia e Rousselle (1976), estendendo uma abordagem originalmente proposta por Valência e Schaake (1973), é definido pela regressão das vazões mensais de um ano qualquer na vazão anual do mesmo ano e em algumas vazões mensais do ano anterior. Formalmente

$$\begin{bmatrix} X_{1,t}^{-\mu} \\ X_{2,t}^{-\mu} \\ \vdots \\ X_{12,t}^{-\mu} \end{bmatrix} = \begin{matrix} A \\ (12 \times 1) \end{matrix} (Q_t^{-\mu}) + \begin{matrix} B \\ (12 \times p) \end{matrix} \begin{bmatrix} X_{12,t-1}^{-\mu} \\ X_{11,t-1}^{-\mu} \\ \vdots \\ X_{12-p+1,t-1}^{-\mu} \end{bmatrix} + \begin{matrix} C \\ 12 \times 11 \end{matrix} \begin{bmatrix} V_{1,t} \\ V_{2,t} \\ \vdots \\ V_{11,t} \end{bmatrix} \quad (24a)$$

onde  $\tilde{V}_t$  é um vetor de variáveis aleatórias independentes entre si, de média zero e

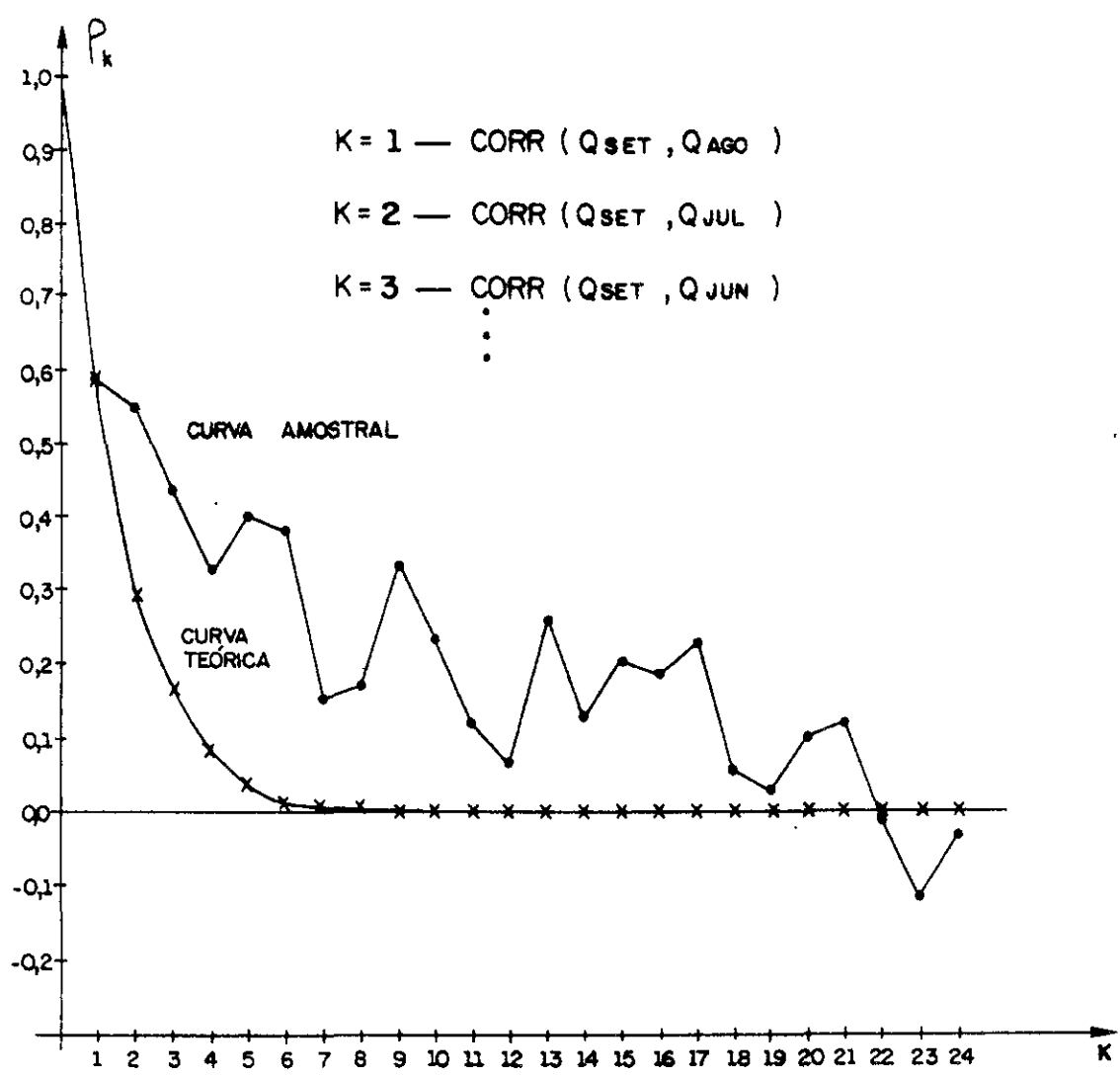


FIGURA 11 — CORRELAÇÃO ENTRE VAZÕES MENSAIS DO RIO PARAÍBA DO SUL  
 EM BARRA DO PIRAI

desvio padrão unitário A, B e C são matrizes de dimensão respectivamente igual a (12x1), (12xp) e (12x11).

Por simplicidade de notação sejam as matrizes aleatórias que aparecem na equação (24a) re-batizadas de forma seguinte:

$$\begin{matrix} (12 \times 1) \\ Y_t \end{matrix} = \begin{matrix} (12 \times 1) \\ A \end{matrix} \begin{matrix} (1 \times 1) \\ W_t \end{matrix} + \begin{matrix} (1 \times 1) \\ BZ_t \end{matrix} + \begin{matrix} (11 \times 1) \\ C \end{matrix} \begin{matrix} (12 \times 1) \\ V_t \end{matrix} \quad (24b)$$

As matrizes parâmetros A, B e C devem ser estimadas de tal maneira que sejam preservadas as estimativas de

$$\text{cov} \begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} S_{yy} & S_{yz} \\ S_{zy} & S_{zz} \end{bmatrix} \quad (25)$$

Ademais, deseja-se que  $W_t = \frac{1}{12} \sum_{\tau=1}^{12} Y_{\tau,t}$ , isto é, que a média do processo mensal seja sempre igual ao processo anual. Este requisito é importante porque o processo  $W_t$  já foi gerado numa etapa anterior e não pode ser modificado.

Os parâmetros A, B e C podem ser calculados através do seguinte procedimento:

- Defina-se  $D = [ \frac{1}{12}, \frac{1}{12}, \dots, \frac{1}{12} ]$  como um vetor constante de 12 componentes. Desta forma:

$$W_t = DY_t \quad (26)$$

Como  $\text{cov}(DY) = D \text{cov}(Y)D^T$ , obtem-se que

$$\text{var}(W_t) = D S_{YY} D^T \quad (27)$$

Multiplicando-se (24b) por  $W_t$  e tomando o valor esperado, obtem-se

$$DS_{YY} = AD S_{YY} D^T + B S_{ZY} D^T \quad (28)$$

Multiplicando-se (24b) por  $Z_t$  e tomando o valor esperado, obtem-se

$$S_{YZ} = ADS_{YZ} + BS_{ZZ}$$

As equações (27) e (28) podem ser resolvidas para A e B:

$$A = (S_{YY} - S_{YZ} (S_{ZZ})^{-1} S_{ZY}) D^T [D(S_{YY} - S_{YZ} (S_{ZZ})^{-1} S_{ZY}) D^T]^{-1} \quad (29)$$

$$B = (I - AD) S_{YZ} (S_{ZZ})^{-1} \quad (30)$$

Multiplicando-se (24b) por  $Y_t^T$  e tomando o valor esperado, obtém-se

$$S_{YY} = AD(S_{YY} D^T A^T + S_{YZ} B^T) + B(S_{ZY} D^T A^T + S_{ZZ} B^T) + CC^T \quad (31)$$

A equação (31) é resolvida substituindo os valores de A e B já calculados. A equação resultante é do tipo vista no item 1.2.4 e admite solução análoga. Alternativamente, uma das soluções para C é dada por

$$C = GH^{1/2} \quad (32)$$

onde

G é a matriz de autovalores de  $CC^T$  e

H é a matriz diagonal de autovalores de  $CC^T$ .

Pode-se demonstrar que

$$DA = I \quad (33)$$

$$DB = 0 \quad (34)$$

$$DC = 0 \quad (35)$$

Multiplicando-se a equação (24b) por D e aplicando (33), (34) e (35), é fácil mostrar que a equação (26) é verdadeira.

Como o vetor  $V_t$  é independente, pode-se sortear amostras a partir de uma distribuição normal padrão ou então sortear pontos das amostras históricas de cada componente de  $V_t$  (Kelman et al 1982).

## 1.4 Modelagem de Vazões Diárias

### 1.4.1 Modelagem Direta

Séries sintéticas com intervalo de discretização diário são em geral utilizadas para estudos relacionados a controle de cheias.

Uma primeira tentativa para modelar estas séries poderia ser feita utili-

zando-se qualquer das formulações apresentadas para modelar as vazões mensais. Isto é, vazões diárias seriam representadas pela equação (17) ou (24), só que o índice  $\tau$  variaria de 1 a 365 em vez de 1 a 12. Esta abordagem, no entanto, não consegue reproduzir seja as exageradas assimetrias apresentadas pelos registros diários, seja a diversidade de comportamento dos ramos de ascensão e recessão das hidrógrafas. Em consequência, existe, na literatura diversas propostas de modelos especificamente desenvolvidos para  $\Delta t = 1$  dia. Estes modelos são em geral complicados e sua descrição tornaria este texto excessivamente longo. O leitor interessado deve consultar, entre outros, Barbosa (1981), Kelman (1980), Yakowitz (1979), O'Connell (1978), Weiss (1977), Treiber e Plate (1975) e Quimpo (1967).

#### 1.4.2 Modelagem Indireta

Uma outra forma de obter séries sintéticas de vazão com intervalo de discretização diário é de gerar séries estocásticas de precipitações, que servem como dados de entrada para um modelo determinístico chuva x vazão. Esta opção pode ser vantajosa nas regiões em que se disponha de um registro pluviométrico de melhor qualidade do que o fluviométrico, o que é frequentemente o caso.

Diversos autores têm estudado precipitação diária, em geral no intuito de derivar resultados analíticos que permitam a definição da distribuição de probabilidades de variáveis relevantes — máxima precipitação ou precipitação acumulada — sem que seja preciso gerar séries sintéticas. Como o objetivo deste texto é de fornecer subsídios para a produção de séries sintéticas de vazão, estes estudos não serão considerados. A seguir apresenta-se uma formulação bastante simples, suficiente para produzir séries sintéticas estacionárias correspondentes a um posto fluviométrico. Kelman (1976) generalizou o procedimento, permitindo a modelagem multivariada e não-estacionária.

Registros diários de precipitação têm um grande número de zeros. Eles podem ser vistos como realizações de um processo estocástico intermitente. O processo poderia ser modelado ajustando-se uma distribuição de probabilidades míxta: uma massa de probabilidade  $p$  ( $0 \leq p \leq 1$ ) correspondente à precipitação nula ( $u = 0$ ), e uma distribuição contínua  $(1-p) f_U(u)$  para  $u > 0$ . Ocorre, no entanto, que esta abordagem desconsidera uma característica relevante da série de precipitação diária: dias chuvosos tendem a ocorrer aos grupos, principalmente em regiões onde a maior parte da precipitação tem origem na chegada de frentes, que por vezes estacionam por dias seguidos. Formalizando, seja a variável aleatória  $S_t$ , definida como:

$$\begin{aligned}
 S_t &= 1, \text{ dia chuvoso} \\
 S_t &= 0, \text{ dia não chuvoso}
 \end{aligned}
 \tag{36}$$

Pode-se definir a cadeia de Markov de dois estados pela matriz de transição

$$\begin{bmatrix}
 1 - p_{01} & p_{01} \\
 1 - p_{11} & p_{11}
 \end{bmatrix}
 \tag{37}$$

onde

$$p_{01} = p[S_{t+1} = 1 \mid S_t = 0]$$

$$p_{11} = p[S_{t+1} = 1 \mid S_t = 1]$$

Espera-se, devido à persistência de dias chuvosos, que  $p_{11} > p_{01}$ .

Para os dias  $t$  em que  $S_t = 1$ , sabe-se que existe alguma precipitação. Sua magnitude deverá ser sorteada da distribuição de probabilidades da variável aleatória  $R$  — altura de precipitação diária ( $F_R(r)$ ). Esta distribuição de probabilidades poderá ser a empírica ou então alguma distribuição paramétrica ajustada aos dados amostrais. A distribuição empírica só deverá ser utilizada quando o "polígono de influência" do posto fluviométrico corresponder a uma área de drenagem com tempo de concentração de alguns dias. Neste caso, cheias mais severas do que as registradas no histórico poderão ser produzidas pela simples persistência da precipitação, sem que haja dia algum no período chuvoso em que a precipitação sintética ultrapasse a máxima do histórico. Em bacias mensais, no entanto, cheias sintéticas mais extremas do que as observadas no passado estarão associadas necessariamente com extrapolações de precipitação diária, o que impõe o ajuste de alguma função analítica à correspondente distribuição de probabilidades.

Na fase de estimação, deve-se selecionar um período suficientemente curto para que se possa aceitar a hipótese de estacionariedade do processo estocástico "precipitação diária", por exemplo o mês mais molhado do ano. As probabilidades  $p_{01}$  e  $p_{11}$  são estimadas por:

$$\hat{p} = \frac{n_0 \text{ dias não chuvosos}}{n_0 \text{ dias chuvosos} + n_1 \text{ dias chuvosos}}$$



$$\hat{p}_{01} = \frac{\text{n}^\circ \text{ de dias n\~{a}o chuvosos seguidos por dias chuvosos}}{\text{n}^\circ \text{ dias n\~{a}o chuvosos}}$$

e

$$\hat{p}_{11} = \frac{\text{n}^\circ \text{ de dias chuvosos seguidos por dias chuvosos}}{\text{n}^\circ \text{ dias chuvosos}}$$

O algoritmo de geraç~{a}o pode ser descrito como :

a) Sorteie um n~{u}mero uniforme U

$$\text{b) } U \begin{cases} \leq \hat{p} & \rightarrow S_o = 0 \\ > \hat{p} & \rightarrow S_o = 1 \end{cases}$$

c)  $t = 0$

d) Sorteie um n~{u}mero uniforme U

$$\text{- Se } S_t = 0 \text{ ent\~{a}o } \quad U \begin{cases} \leq \hat{p}_{01} & \rightarrow S_{t+1} = 1 \\ > \hat{p}_{01} & \rightarrow S_{t+1} = 0 \end{cases}$$

$$\text{- Se } S_t = 1 \text{ ent\~{a}o } \quad U \begin{cases} \leq \hat{p}_{11} & \rightarrow S_{t+1} = 1 \\ > \hat{p}_{11} & \rightarrow S_{t+1} = 0 \end{cases}$$

e) Se  $S_{t+1} = 1 \rightarrow$  sorteie R de  $f_R(r)$

f)  $t = t+1$

g) v~{a} para d.

---

## REFERÊNCIAS

---

1. BARBOSA, P.R.D.J. (1981) - Modelagem Estocástica de Vazões Diárias, Tese de Mestrado, COPPE-UFRJ.
2. BURGESS, S.J., LETTENMAIER, D.P. e BATER, C.L. - Properties of the Three - Parameter Log Normal Probability Distribution, Water Resources Research, Vol. 11, No. 2, 1975.
3. CHOW, V.T. (1954) - The log-probability Law and its Engineering Applications, Proc. Am. Soc. Civ. Eng. 80, Paper No. 536.
4. DAMAZIO, J.M. e KELMAN, J. (1981) - Previsão de Total Anual de Precipitação, IV Simpósio Brasileiro de Hidrologia e Recursos Hídricos, Fortaleza.
5. GOMIDE, F.L.S. (1976) - Noções de Estatística e Probabilidades, em Hidrologia Básica por N.L.S. Pinto e outros, Editora Edgard Blücher Ltda.
6. KELMAN, J. (1977) - Stochastic Modeling of Hydrologic Intermittent Daily Processes, Colorado State University Hydrology Paper No. 89.
7. KELMAN, J. (1980) - A Stochastic Model for Daily Streamflow, Journal of Hydrology, 47.
8. KELMAN, J., OLIVEIRA, G.C., PEREIRA, M.V.F. e COSTA, C.C.G. (1982) - Modelo de Séries Hidrológicas - Manual de Metodologia, Relatório Técnico CEPEL No. 113/81-A, Revisão 1.
9. KLEMES, V. (1978) - Physically Based Stochastic Hydrologic Analysis, em Advances of Hydrosience No. 11, Academic Press.
10. MARKOVIC, R.D. (1965) - Probability Functions of Best Fit to Distributions of Annual Precipitation and Runoff, Colorado State University Hydrology Paper No. 8.
11. MEJIA, J.M. e ROUSSELLE, J. (1976) - Disaggregation Models in Hydrology Revisted, Water Resources Research, Vol. 12, No. 2.

12. O'CONNEL, P.E. e JONES, D.A. - Some Experience With the Development of Models for the Stochastic Simulation of Daily Flows. Risk and reliability Symposium, Waterloo, Canada, 1978.
13. QUIMPO, R.G. - Stochastic Model of Daily River Flow Sequences, Hydrology Papers, No. 18, Colorado State University, Fort Collins, Colorado, 1967.
14. SALAS, J.D., DELLEUR, J.W., YEVJEVICH, V. e LANE, W.L. (1981) - Applied Modeling of Hydrologic Time Series, Water Resources Publications.
15. TREIBER, B. e PLATE, E.J. - A Stochastic Model for the Simulation of Daily Flows. Symposium and Workshop on the Application of Mathematical Models in Hydrology and Water Resources Systems, 1975. Bratislava, Czechoslovakia.
16. VALENCIA, D. e SCHAAKE, J.C. Jr. (1973) - Disaggregation processes in Stochastic Hydrology, Water Resources Research, Vol. 9, No. 3.
17. WEISS, G. - Shot Noise Models for the Generation of Synthetic Streamflow Data. Water Resources Research, Vol. 13, No. 1, 1977.
18. YAKOWITZ, S.J. - A Nonparametric Markov Model for Daily River Flow, Water Resources Research, Vol. 15, No. 5, 1979.
19. YEVJEVICH, V. (1963) - Fluctuations of Wet and Dry Years, Colorado State University Hydrology Paper No. 1.
20. YEVJEVICH, V. (1972) - Stochastic Processes in Hydrology, Water Resources Publications.
21. YEVJEVICH, V. (1979) - Viabilidade de Previsão de Secas a Longo Prazo, Rev. Hidrol. Rec. Hdr., No. 2