

RBE

VOL.1 / N.2

**REVISTA BRASILEIRA
DE ENGENHARIA**

**ASSOCIAÇÃO BRASILEIRA DE HIDROLOGIA E
RECURSOS HÍDRICOS
CADERNO DE RECURSOS HÍDRICOS**



GERAÇÃO DE SÉRIES SINTÉTICAS DE VAZÕES DIÁRIAS
MODELO DIANA

POR

J.Kelman^{1,2}, J.M.Damázio¹, e J.P. da Costa¹

RESUMO -- Este artigo descreve os aspectos metodológicos do modelo DIANA, desenvolvido no CEPEL para a geração multivariada de séries de vazões diárias. É descrita também a metodologia adotada para a avaliação do desempenho das séries geradas pelo modelo. O modelo DIANA foi elaborado para ser aplicado em estudos hidrológicos de cheias que afetam o sistema hidroelétrico brasileiro, seja na determinação de volumes de espera nos reservatórios, seja na análise de segurança das barragens.

É apresentado um caso exemplo para a bacia do Rio Grande.

INTRODUÇÃO

É cada vez mais freqüente a utilização de séries hidrológicas sintéticas no planejamento de sistemas de aproveitamento de recursos hídricos. As equipes de projeto e/ou planejamento dispõem atualmente de um grande número de modelos matemáticos e algoritmos de estimação de parâmetros, muitos deles adaptados a partir de teorias tradicionais de modelagem de séries temporais.

Por exemplo, hoje é possível testar a confiabilidade dos planos de operação e expansão do sistema hidroelétrico brasileiro através do uso de séries sintéticas de vazões mensais afluentes às usinas (Kelman et al 1979). A confiança depositada nesta técnica se baseia na capacidade destes modelos de produzir séries sintéticas estatisticamente indistinguíveis das correspondentes séries históricas de afluições (Costa et al., 1981). Adota-se implicitamente a hipótese de que as leis probabilísticas que regeram o fenômeno no passado serão preservadas no futuro.

Existem estudos hidrológicos que demandam séries sintéticas a intervalos de tempo menores que o mês. Por exemplo, Beard (1968) indica a necessidade de se considerar as variações das afluições dentro do mês no cálculo do volume de espera a ser alocado nos reservatórios para o controle de cheias. O'Connell e Jones (1978) demonstraram a necessidade de séries diárias para o cálculo da confiabilidade de um reservatório de abastecimento de água na Inglaterra. Kelman e Damázio (1983) concluíram que o uso de modelos de geração de vazões diárias pode ser uma abordagem válida na modelagem de vazões extremas e extremamente útil na análise de segurança de barragens. No entanto a experiência tem demonstrado que quanto menor o intervalo de tempo considerado, maiores as dificuldades enfrentadas para a construção de bons modelos de geração. No caso da modelagem de séries de vazões diárias a adaptação de um modelo tradicional de série temporal (exemplo: ARIMA) em geral esbarra nas exageradas assimetrias apresentadas nos registros de vazões diárias, na forte sazonalidade e na diversidade de comportamento dos ramos de ascensão e recessão das hidrógrafas.

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)

Kelman (1976), obteve resultados satisfatórios para rios dos EUA modelando separadamente ascensões e recessões. As ascensões, ordinariamente causadas por fatores externos foram modelados por um processo intermitente utilizado por ele mesmo para representar precipitação diária. As recessões, ordinariamente resultado de efeitos de armazenamento da bacia, foram modeladas por hipotéticos reservatórios lineares.

Neste artigo é descrito o modelo para a geração multivariada de séries de vazões diárias, DIANA, desenvolvido no CEPEL a partir da abordagem proposta por Kelman (1976). É descrita também a metodologia adotada para a avaliação do desempenho das séries geradas pelo modelo DIANA que pode ser aplicada a qualquer outro modelo. O modelo DIANA foi elaborado para ser aplicado em estudos hidrológicos de cheias que afetam o sistema hidroelétrico brasileiro. O modelo é testado na bacia do rio Grande tendo sido gerado concomitantemente 1000 anos de vazões diárias afluentes ao reservatório de Furnas e de vazões diárias incrementais para o trecho entre este reservatório e o reservatório de Marimbondo.

DESCRIÇÃO DO MODELO DIANA

Considera-se que a vazão no dia t da hidrógrafa de um dado posto fluviométrico é dada pela soma de dois componentes:

$$Q(t) = U(t) + O(t); \quad t = 1, 2, \dots \quad (1)$$

Idealmente, $U(t)$ dependeria principalmente de fatores externos (exemplo: precipitações) cuja natureza intermitente ocasionariam na hidrógrafa pulsos irregularmente espaçados no tempo. Por outro lado $O(t)$ representaria o contínuo esvaziamento da água armazenada na bacia hidrográfica.

No entanto só se encontram disponíveis registros de vazão total, $Q(t)$. Algum método é pois necessário para a repartição de $Q(t)$ em seus componentes $U(t)$ e $O(t)$.

A opção adotada no modelo DIANA considera a ocorrência de $u(t)$ positivo sempre que $q(t) > \lambda q(t-1)$, onde $0 \leq \lambda \leq 1$ é uma taxa de recessão característica do posto fluviométrico.

Kelman (1976), adota $\lambda=1$ em qualquer caso. Na realidade espera-se que λ seja inferior a 1 e durante o desenvolvimento de DIANA foram testados diversos valores para λ resultando daí uma regra empírica para sua definição, descrita no item de estimação mais adiante.

Para qualquer λ a opção implica que:

$$u(t) = 0 \quad \text{se} \quad q(t) \leq \lambda q(t-1) \quad (2a)$$

$$u(t) = q(t) - \lambda q(t-1) \quad \text{se} \quad q(t) > \lambda q(t-1) \quad (2b)$$

Nos dias em que atuam fatores externos ($u(t) > 0$) a vazão total é dada rearrumando-se (2b):

$$q(t) = \lambda q(t-1) + u(t), \quad u(t) > 0 \quad (3)$$

e portanto:

$$O(t) = \lambda q(t-1), \quad u(t) > 0 \quad (4)$$

Nos dias em que não atuam fatores externos ($u(t) = 0$) a vazão total é dada

simplesmente por $O(t)$. Kelman (1976), considerou $O(t)$ como a soma das defluências de dois reservatórios lineares, cada parcela dependendo de $q(t-1)$ através de uma equação probabilística. Esta abordagem pode ser simplificada considerando-se apenas um reservatório linear de comportamento estocástico. Assim, quando $u(t)=0$, $O(t)$, e portanto $q(t)$, é uma fração $k(t) \leq \lambda$ da vazão anterior:

$$q(t) = O(t) = k(t) q(t-1), \quad k(t) \leq \lambda; \quad u(t) = 0 \quad (5)$$

Juntando-se as equações (1), (3) e (5) o modelo DIANA pode ser descrito por:

$$q(t) = u(t) + k(t) q(t-1) \quad (6)$$

$$u(t) = 0 \Rightarrow k(t) \leq \lambda \quad (6a)$$

$$u(t) > 0 \Rightarrow k(t) = \lambda \quad (6b)$$

Existem dois aspectos a serem considerados na modelagem do processo $U(t)$:

- (i) definir sua distribuição de probabilidades marginal, levando-se em conta que existe uma probabilidade finita p de que $U(t)$ seja exatamente igual a zero;
- (ii) preservar a dependência entre valores sucessivos de $U(t)$, visto que os fatores externos que influenciam os ramos ascendentes da hidrógrafa podem resultar da ação de fenômenos meteorológicos persistentes sobre a bacia hidrográfica (exemplo: frente fria estacionária).

Kelman (1976), resolve estas duas questões mapeando o processo $U(t)$ no processo $Y(t)$ através de uma transformação do tipo $U(t) = (Y(t))^2$. O processo $Y(t)$ é suposto resultante de uma censura imposta a um processo auto-regressivo de ordem 1 de distribuição normal, em que todas as observações negativas foram "perdidas". Esta representação paramétrica de $U(t)$ permite a estimação dos parâmetros envolvidos (inclusive ρ , a autocorrelação do processo não censurado, e α) através do método da máxima verossimilhança.

A aplicação deste método a rios brasileiros revelou em diversos casos um mau ajuste à distribuição de frequências empíricas de $U(t)$, $F_U(\cdot)$. O modelo DIANA adota a própria distribuição empírica, $F_U(\cdot)$, visto que usualmente dispõe-se de um grande número de observações de $u(t)$. Quanto a questão da dependência entre $U(t-1)$ e $U(t)$, o modelo DIANA, analogamente a Kelman (1976), assume a existência de um processo autoregressivo normal de ordem 1 com censura, cujo coeficiente de autocorrelação ρ deve ser preservado. Este processo é mapeado em $U(t)$ através de uma transformação não-paramétrica que deve preservar a distribuição empírica $F_U(\cdot)$, conforme a seguir exposto.

Seja o processo markoviano $Z(t)$ definido por:

$$z(t) = \rho z(t-1) + \sqrt{1-\rho^2} \varepsilon(t) \quad (7)$$

onde $\varepsilon(t)$ é um ruído normal padrão e ρ a correlação lag-um do processo markoviano.

Seja o processo $Y(t)$ resultado de uma censura imposta ao processo $Z(t)$ definido por:

$$y(t) = z(t) \quad \text{se} \quad z(t) > \beta \quad (8a)$$

$$y(t) = \beta \quad \text{se} \quad z(t) \leq \beta \quad (8b)$$

$$\beta = \Phi^{-1}(p) \quad (8c)$$

onde β define o intervalo de censura $(-\infty, \beta)$, $\Phi(\cdot)$ é a distribuição acumulada normal, e $p = P[u(t)=0]$.

A relação entre $U(t)$ e $Y(t)$ é obtida resolvendo-se $F_U(u(t)) = \Phi(y(t))$.

Para a modelagem da série de $k(t)$'s assumiu-se que seja suficiente simplesmente usar para a geração de valores sintéticos a distribuição empírica de $k(t)$ menor que λ , $F_K(\cdot)$, sempre que $u(t)=0$.

O algoritmo de geração de séries compõe-se então dos seguintes passos:

- 1- Sortear uma vazão inicial $q(0)$ de uma distribuição empírica $F_{Q0}(\cdot)$.
- 2- Sortear um valor inicial $z(0)$ da distribuição normal padrão $\Phi(\cdot)$.
- 3- Fazer $t=1$.
- 4- Sortear um ruído $\epsilon(t)$ da distribuição normal-padrão $\Phi(\cdot)$.
- 5- Calcular $z(t) = \rho z(t-1) + \sqrt{1-\rho^2} \epsilon(t)$ e fazer $y(t) = \max(\beta, z(t))$.
- 6- Obter $u(t)$ resolvendo

$$F_U(u(t)) = \Phi(y(t))$$

A aproximação é necessária porque $F_U(\cdot)$ é uma distribuição discreta e $\Phi(\cdot)$, contínua (este tópico será visto mais detalhadamente no próximo item).

- 7- Se $u(t) > 0$ obter $q(t) = u(t) + \lambda q(t-1)$. Se $u(t)=0$, sortear $k(t)$ de $F_K(\cdot)$ e obter $q(t) = k(t) \cdot q(t-1)$.
- 8- Fazer $t = t+1$ e retornar a 4.

ESTIMAÇÃO

Dada uma série histórica de vazões diárias $q(0), q(1), \dots, q(h)$ e para uma particular taxa de recessão característica λ , é possível obter a série $u(1), u(2), \dots, u(h)$ correspondente através de (2a) e (2b).

Para a obtenção da correspondente série de $y(t)$, a série $u(t)$ é arrumada no vetor: $W = (u(t_1), u(t_2), \dots, u(t_h))$, de tal forma que $u(t_1) \leq u(t_2) \leq \dots \leq u(t_h)$.

Seja m o número de zeros na série $u(t)$. A estimativa de p é dada por $\hat{p} = \frac{m}{h}$.

A função que mapeia $y(t)$ em $u(t)$ é dada por $u(t) = g[\Phi(y(t))]$ de acordo com a figura 1.

Nesta função, para cada $u(t_j) > 0$ existe um intervalo associado $[d_j, e_j]$ relacionado com a sua posição j no vetor ordenado W tal que:

$$\Phi(d_j) = \frac{j-1}{h}, \quad j \geq m+1 \quad (9a)$$

$$\Phi(e_j) = \frac{j}{h}, \quad j \geq m+1 \quad (9b)$$

Para $u(t_j) = 0$ só existe um valor associado:

$$y(t_j) = \beta, \quad \Phi(\beta) = \frac{m}{h} = \hat{p} \quad (10)$$

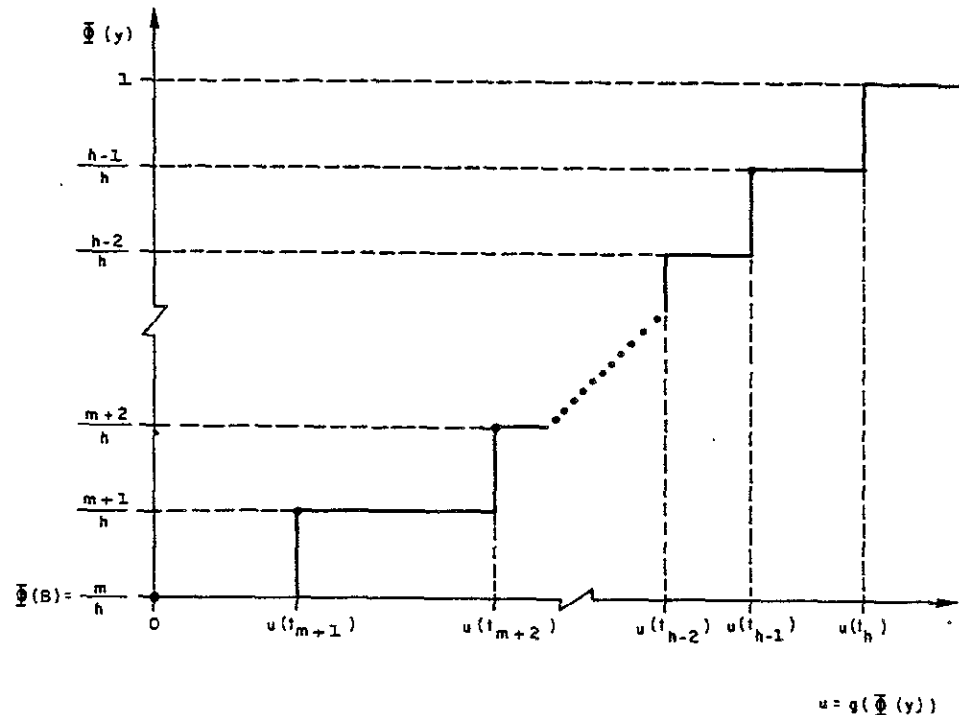


Figura 1. Relação não-paramétrica entre $y(t)$ e $u(t)$.

A função g é útil no passo 6 do esquema de geração para a obtenção de $u(t)$ através da solução da equação $F_{ij}(u(t)) = \Phi(y(t))$, fazendo-se simplesmente $u(t) = g[\Phi(y(t))]$. Por outro lado a função g não é biunívoca e conseqüentemente fica difícil a obtenção das séries históricas dos $y(t)$'s: quando $u(t_j) > 0$, é impossível se obter exatamente o valor de $y(t_j)$ correspondente, conhecendo-se apenas um intervalo viável, $d_j < y(t_j) < e_j$.

A abordagem adotada consiste em se associar a cada $u(t_j) > 0$ o valor média no do intervalo correspondente:

$$\Phi(y(t_j)) = \frac{1}{2} [\Phi(d_j) + \Phi(e_j)] \quad (11)$$

Em resumo, ordenada a série de $u(t)$, é possível, através das equações (9a), (9b), (10) e (11) obter uma realização do processo $Y(t)$. Esta realização pode ser considerada como uma amostra censurada do processo markoviano padrão $Z(t)$ definido por (7) com intervalo de censura $(-\infty, \beta)$. O parâmetro ρ poderia ser estimado por máxima verossimilhança, de acordo com o método proposto por Kelman (1976). De forma mais simples, se os valores nulos de $u(t)$ forem eliminados do registro, a amostra resultante pode ser considerada como uma amostra truncada de $Z(t)$ com nível de truncamento dado por β . A estimativa do parâmetro ρ pode ser feita considerando-se os pares contíguos ($y(t) > \beta$, $y(t+1) > \beta$) resultantes de uma amostra de uma distribuição bivariada normal padrão truncada identicamente nas duas variáveis. Reiger et al. (1971) apresentam uma tabela que relaciona a correlação na distribuição truncada, ρ' , o nível de truncamento, β , e o coeficiente de correlação de população ρ na distribuição não-truncada. No modelo DIANA, β é obtido por (10) e $\hat{\rho}$ usando-se todos os pares contíguos ($y(t) > \beta$, $y(t+1) > \beta$) em:

$$\hat{\rho}' = \frac{\sum (y(t) - \bar{y}_t) (y(t+1) - \bar{y}_{t+1})}{[\sum (y(t) - \bar{y}_t)^2 \cdot \sum (y(t+1) - \bar{y}_{t+1})^2]^{1/2}} \quad (12)$$

$$\bar{y}_t = \frac{\sum y(t)}{n} ; \quad \bar{y}_{t+1} = \frac{\sum y(t+1)}{n} \quad (13)$$

onde n é o número de pares contíguos obtidos. A estimativa de ρ é obtida entrando-se na tabela com $\hat{\rho}'$ e β .

Não existindo nenhuma formulação para o ajuste da taxa de recessão característica, λ , aplicou-se o modelo para diversos valores tendo-se escolhido adotar para λ o quantil de 15% da distribuição empírica de $Q(t)/Q(t-1) < 1$.

EXTENSÃO PARA O CASO MULTIVARIADO

O modelo foi estendido para a geração simultânea de l hidrógrafas usando-se o modelo proposto por Matalas (1967) para a geração multivariada de variáveis normais padronizadas. Este modelo é formulado pela equação:

$$Z(t) = A Z(t-1) + B \varepsilon(t) \quad (14)$$

onde $Z(t)$ é o vetor $[Z_1(t), Z_2(t), \dots, Z_l(t)]^T$ e $Z_i(t)$ é o processo markoviano

correspondente ao posto i , $\varepsilon(t)$ é um vetor com l variáveis normais padronizadas independentes.

As matrizes A e B devem ser escolhidas para representar tanto a persistência de cada hidrógrafa quanto a correlação existente entre as hidrógrafas representada pela matriz de covariância de $Z(t)$, M_0 .

O algoritmo de geração de l séries multivariadas compõe-se então dos seguintes passos:

- 1 - Sortear l vazões iniciais $q_i(0)$, $i=1, \dots, l$ de uma distribuição empírica multivariada $F_{Q0}(\cdot)$.
- 2 - Sortear l valores iniciais $z_i(0)$, $i=1, \dots, l$ da distribuição normal padrão multivariada com covariância M_0 e arrumá-los no vetor $Z(0) = [z_1(0), z_2(0), \dots, z_l(0)]^T$. Este sorteio é feito sorteando-se um vetor normal padrão multivariado independente $\varepsilon(0)$ que é pré-multiplicado por uma matriz iniciadora, C , solução de $CC^T = M_0$.
- 3 - Fazer $t=1$.
- 4 - Sortear l ruídos normais independentes da distribuição normal padrão multivariada independente e arrumá-los no vetor $\varepsilon(t) = (\varepsilon_1(t), \varepsilon_2(t), \dots, \varepsilon_l(t))^T$.
- 5 - Calcular $Z(t) = A Z(t-1) + B \varepsilon(t)$.
- 6 - Para cada série fazer $y_i(t) = \text{Max}(\beta_i, z_i(t))$, $i=1, 2, \dots, l$, e obter $u_i(t)$ correspondente fazendo $u_i(t) = g_i[\Phi(y_i(t))]$.
- 7 - Se $u_i(t) > 0$, obter $q_i(t) = u_i(t) + \lambda_i q_i(t-1)$. Se $u_i(t) = 0$, sortear $k_i(t)$ de $F_{k_i}(\cdot)$ e obter $q_i(t) = k_i(t) q_i(t-1)$.
- 8 - Fazer $t=t+1$, e retornar a 4.

ESTIMAÇÃO DAS MATRIZES A e B

De acordo com Matalas (1967) o esquema de geração multivariado definido na equação (14) deve considerar não só os coeficientes de correlação temporal em cada série, ρ_{ij} , como as correlações cruzadas entre as séries $\{z_i(t)\}$ $i=1, \dots, l$ representadas pelas correlações cruzadas lag-zero $\rho_{i,j}(0)$ definida por:

$$\rho_{i,j}(0) = E[z_i(t) \cdot z_j(t)] \quad \forall i, j \quad (15)$$

Para a estimação dos valores de $\rho_{i,j}(0)$, $i \neq j$; pode-se considerar o conjunto dos pares ($y_i(t) > \beta_i$, $y_j(t) > \beta_j$) como uma amostra aleatória de uma distribuição normal bivariada padrão com correlação $\rho_{i,j}(0)$ a estimar e truncada em β_i e β_j . No anexo apresenta-se o método de estimação adotado no modelo DIANA.

O cálculo das matrizes A e B correspondentes as estimativas de $\rho_{i,j}(0)$, $i=1, \dots, l$; $j=1, \dots, l$ é feita considerando-se as matrizes $M_0 = E[Z_t Z_t^T]$ e $M_1 = E[Z_{t+1} Z_t^T]$. M_0 é uma matriz de dimensão $l \times l$ cujos elementos são dados por:

$$m_0(i, j) = 1; \quad i=j \quad (16a)$$

$$m_0(i, j) = \rho_{i,j}(0); \quad i \neq j \quad (16b)$$

M_1 é uma matriz de dimensão $l \times l$ cujos elementos são dados por:

$$m_1(i,j) = \rho_i; \quad i=j \quad (17a)$$

$$m_1(i,j) = \rho_{i,j}(0) \rho_i; \quad i \neq j \quad (17b)$$

Nota-se que o uso da equação (17b) dispensa a estimativa de $\rho_{i,j}(1)$.

A matriz A é dada por:

$$A = M_1 M_0^{-1} \quad (18)$$

A matriz B é dada por:

$$BB^T = M_0 - M_1 M_0^{-1} M_1^T \quad (19)$$

A equação (19) não tem solução única e conforme Matalas (1967) qualquer uma das soluções adotadas pode ser usada. O modelo DIANA adota a solução de componentes principais (Matalas, 1967). A solução das equações (18) e (19) impõe restrições nos valores dos parâmetros ρ_i e $\rho_{i,j}(0)$, $i=1, \dots, l$; $j=1, \dots, l$ de forma a garantir que as matrizes M_0 e $M_0 - M_1 M_0^{-1} M_1^T$ sejam positivas definidas. Estas duas condições podem ser garantidas se a matriz V de dimensão $(2l \times 2l)$ e definida por:

$$V = E \begin{bmatrix} Z_{t+1} \\ Z_t \end{bmatrix} \begin{bmatrix} Z_{t+1}^T & Z_t^T \end{bmatrix} \quad (20)$$

for positiva definida. No modelo DIANA se o conjunto ρ_i e $\rho_{i,j}(0)$ $i=1, \dots, l$, $j=1, \dots, l$ obtido provocar uma matriz V não positiva definida esta matriz é substituída a pequenas perturbações até se obter uma matriz positiva definida. O mesmo método é sugerido em Fiering (1968).

AVALIAÇÃO DO DESEMPENHO DO MODELO

A aplicabilidade em estudos hidrológicos das séries geradas pelo modelo DIANA é aferida pela comparação entre propriedades relevantes ao estudo (neste caso, estudo de cheias) das séries históricas e as correspondentes propriedades das séries geradas.

Por exemplo, verifica-se se a distribuição de frequências empíricas históricas de vazões máximas anuais pode ser considerada como gerada pelo modelo, através do teste χ^2 usando-se, a distribuição de frequência das vazões máximas anuais geradas como população. São comparadas graficamente a sazonalidade das médias e desvio padrão das vazões diárias históricas e geradas. Apresenta-se ainda gráficos de hidrógrafas históricas e geradas "equivalentes" (por exemplo de mesma vazão máxima).

Outra abordagem, adotada também em Costa et al. (1981) consiste em calcular para a série histórica e para diversas séries geradas com o mesmo comprimento do histórico, estatísticas específicas ou índices. Resulta daí uma distribuição de probabilidades para os índices e coerência entre as séries geradas pelo modelo e a série histórica pode ser medida pela posição do índice histórico nesta distribuição.

Um índice I qualquer é definido como uma função de uma série de vazões, $I=f(Q(1), Q(2), \dots, Q(h))$. Por exemplo, $I=\text{Max}(Q(1), \dots, Q(h))$. Aplicando-se a fun

ção à série de vazões históricas obtêm-se o índice histórico i_{hist} . Com n séries geradas com o mesmo comprimento do histórico obtêm-se n índices gerados i_j , $j=1, \dots, n$. O desempenho do modelo é então medido pela proporção de índices gerados maiores do que o índice histórico, $P\{i_j > i_{hist}\}$. Se esta proporção for muito pequena ou grande, isto é uma indicação de que a observação histórica é atípica para o modelo. Em geral, deseja-se testar m índices, I_i , $i=1, \dots, m$, arrumados no vetor

$$R = [I_1, I_2, \dots, I_m]^T \quad (21)$$

$$I_i = f_i(Q(1), \dots, Q(h)) \quad (22)$$

O vetor obtido com o histórico é denotado R_{hist} e existem n vetores R_j , $j=1, \dots, n$ cada um obtido com uma série gerada. Define-se portanto uma distribuição de probabilidades multivariada.

Para cada vetor R_j e para R_{hist} define-se a distância ao centro da distribuição pelo escalar positivo:

$$d_j = (R_j - \mu)^T \Sigma^{-1} (R_j - \mu) \quad (23)$$

$$d_{hist} = (R_{hist} - \mu)^T \Sigma^{-1} (R_{hist} - \mu) \quad (24)$$

$$\mu = \frac{1}{n} \sum_{j=1}^n R_j \quad (25)$$

$$\Sigma = \frac{1}{n} \sum_{j=1}^n (R_j - \mu) (R_j - \mu)^T \quad (26)$$

A variável d_{hist} é então comparada com as distâncias d_j através de duas medidas de desempenho: a proporção de distâncias geradas maiores que a distância histórica, $P\{d_j > d_{hist}\}$, e a distância padronizada (nota) definida por:

$$\gamma = \frac{d_{hist} - m}{\sigma} \quad (27)$$

$$\sigma^2 = \frac{1}{n} \sum_{j=1}^n (d_j - m)^2 \quad (28)$$

Quanto maior $P\{d_j > d_{hist}\}$ e quanto menor γ melhor o desempenho do modelo.

CASO-EXEMPLO: APLICAÇÃO NO RIO GRANDE

O modelo DIANA foi utilizado na geração de séries sintéticas bivariadas de vazões diárias afluentes ao reservatório de Furnas no rio Grande (área de drenagem de 52.300km²) e vazões diárias incrementais para o trecho entre este reservatório e o reservatório de Marimondo (118.488km²) no mesmo rio. O modelo foi ajustado a partir das séries históricas dos postos de São José da Barra

(correspondente ao reservatório de Furnas, 30 anos hidrológicos de observação (31-32 a 60-61) e Porto José Américo (correspondente ao reservatório de Marimbondo, 25 anos hidrológicos (35-36, 59-60)). Foi modelado apenas o período de 1º de outubro a 30 de abril (212 dias). De forma a melhor levar em conta os efeitos sazonais, este período foi dividido em 14 estações de 14 dias seguidas de uma estação de 15 dias ao final (o 1º dia é tratado separadamente pelo modelo). As vazões incrementais foram calculadas por simples diferença adotando-se apenas o defasamento de três dias. Alguns ajustes foram necessários para evitar vazões incrementais negativas.

Para a série de S.José da Barra obteve-se a recessão característica de 0,91 e para a série de incrementais a Marimbondo, 0,89. Em nenhuma estação foi necessário adotar o método das perturbações para se obter matrizes positiva de finidas. O modelo foi então utilizado para a geração bivariada de 1000 anos. A figura 2 exemplifica hidrógrafas observada e sintética fornecida pelo modelo para a série de Furnas.

As figuras 3 e 4 comparam as distribuições de frequência de máximos anuais obtidas nas séries observadas (ou calculadas, no caso de vazões incrementais) e em 1000 anos de séries sintéticas. O bom ajuste é confirmado pelas estatísticas de qui-quadrado obtidas com seis intervalos (2,75 para S.José da Barra e 2,83 para Marimbondo). A figura 5 apresenta as curvas sazonais de média e desvio padrão diário das séries observadas e sintéticas para a série incremental a Marimbondo, resultados semelhantes são obtidos para a série de Furnas.

Para a aplicação da metodologia de avaliação proposta foram considerados isoladamente 10 índices definidos de acordo com a tabela 1.

ÍNDICE	DEFINIÇÃO
1	Média da vazão diária
2	Desvio-padrão da vazão diária
3	Assimetria da vazão diária
4	Curtose da vazão diária
5	Média da vazão máxima anual
6	Desvio-padrão da vazão máxima anual
7	Assimetria da vazão máxima anual
8	Curtose da vazão máxima anual
9	Recorde de vazão máxima
10	Volume de espera para vazão de restrição igual a média da vazão máxima anual

Tabela 1. Definição dos Índices

Os índices de 1 a 4 refletem a distribuição de probabilidades da vazão diária e os índices de 5 a 9 a distribuição de probabilidades da vazão máxima anual. O índice 10 reflete a distribuição de probabilidades dos volumes. Os índices 1, 5, 9 e 10 foram escolhidos para análises multivariadas.

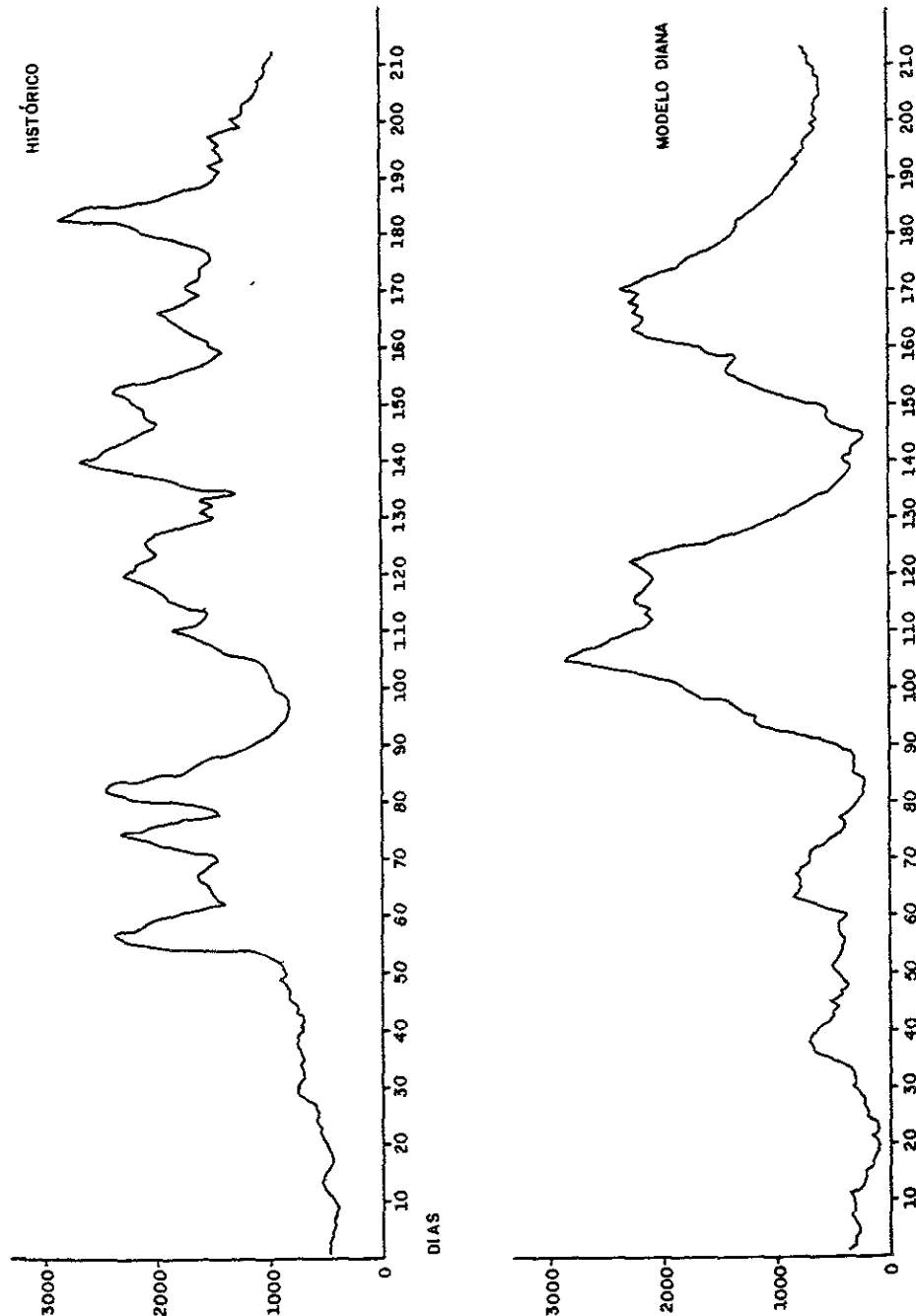


Figura 2. Exemplo de Hidrógrafas Históricas e Sintéticas de mesmo pico. Reservatório de Furnas.

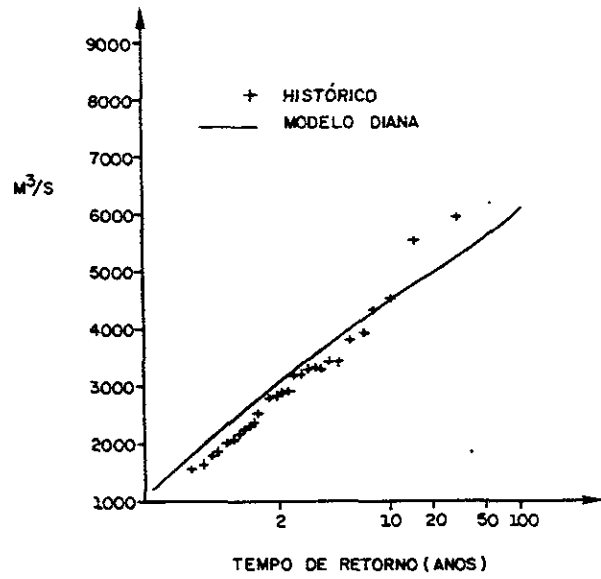


Figura 3. Vazões Extremas. Reservatório de Furnas.

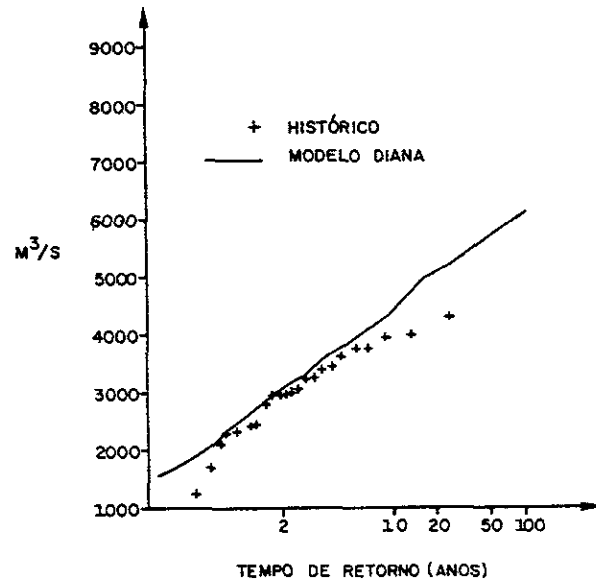


Figura 4. Vazões Extremas Área Incremental entre os Reservatórios de Furnas e Marimbondo.

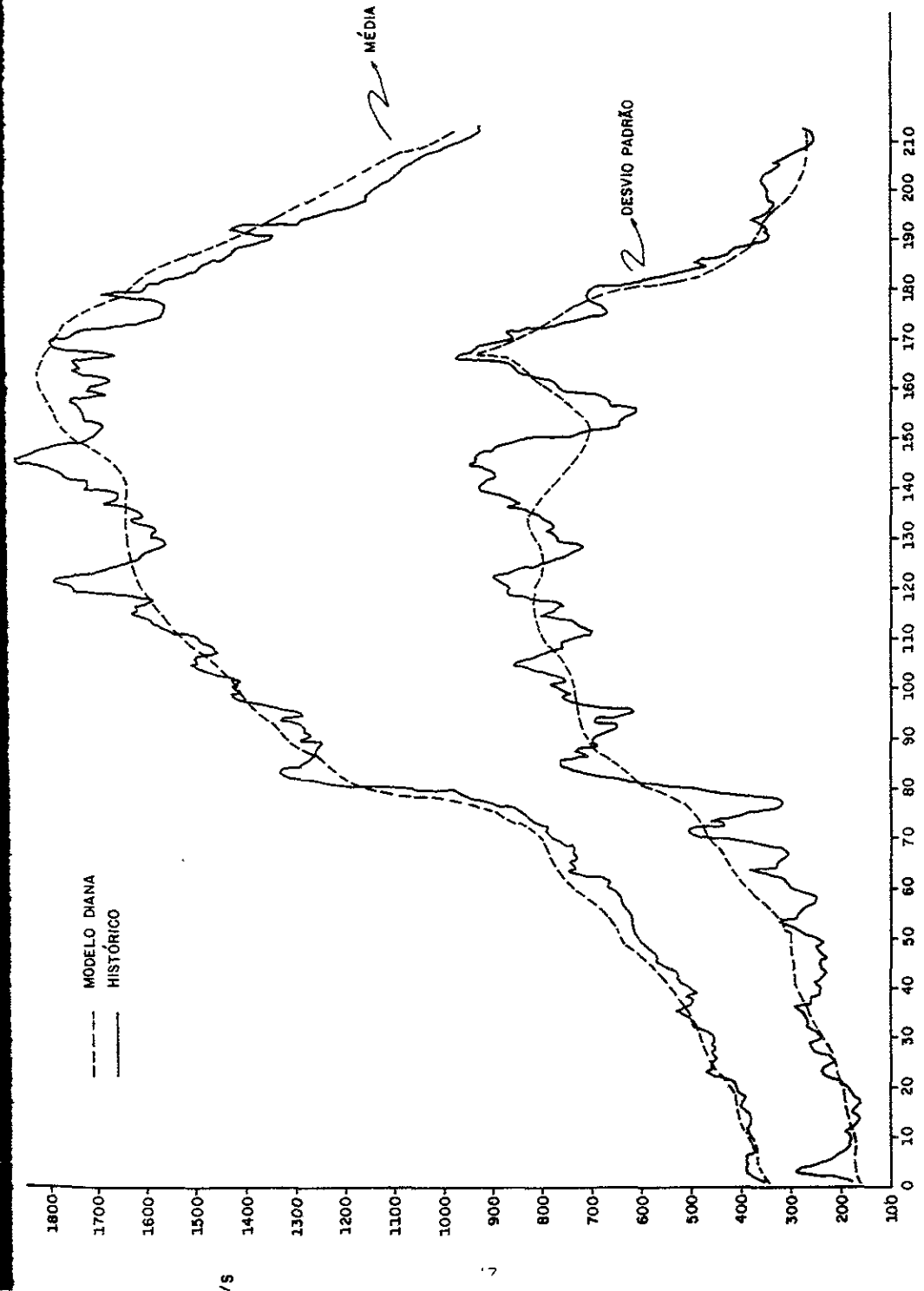


Figura 5. Sazonalidade Para Média e Desvio Padrão de Vazões Incrementais Diárias no Trecho entre os Reservatórios de Furnas e Marimbondo.

Tabela 3. Avaliação do Desempenho do Modelo DIANA para Area Incremental entre Furnas e Marimbondo.

χ² = 0,28340E+01

0.00	-	1724,29	1.0	0.3
1724,29	-	2152,57	2.0	2.0
2152,57	-	2580,86	5.0	4.3
2580,86	-	3009,14	6.0	5.4
3009,14	-	3437,43	4.0	4.2
3437,43	-	99999,00	7.0	8.9

TESTE QUI-QUADRADO PARA O MÁXIMO ANUAL

NBL0COS = 40
DHIST = 0,2819215E+01
P (D>DHIST) = 0,60
NOTA = -0,4559093E+00

ESTATIS. SINT:	1157,11	763,07	1,26	5,61	3231,15	947,13	1,20	5,52	8558,00	0,687E+05
ESTATIS. DIST:	1150,55	769,43	1,00	3,59	2967,32	718,29	-0,25	2,63	4294,00	0,154E+05
MÍNIMOS:	1085,91	644,50	0,74	3,23	2747,04	497,91	-0,22	1,67	3731,00	0,435E+04
MÉDIAS:	1157,11	760,47	1,20	5,26	3231,15	893,89	0,74	3,51	5528,60	0,233E+05
MÁXIMOS:	1234,74	874,59	1,76	9,09	3789,68	1440,30	2,60	11,52	8558,00	0,687E+05
P(SINT>HIST):	0,6000	0,3750	0,7500	0,9000	0,9000	0,8250	1,0000	0,6250	0,8500	0,6750

HISTÓRICO 25 ANOS

IND1	IND2	IND3	IND4	IND5	IND6	IND7	IND8	IND9	IND10
------	------	------	------	------	------	------	------	------	-------

Tabela 2. Avaliação do Desempenho do Modelo DIANA para São José da Barra.

χ² = 0,27469E+01

0.00	-	2210,00	6.0	3.8
2210,00	-	2833,00	8.0	8.0
2833,00	-	3456,00	9.0	7.8
3456,00	-	4079,00	3.0	5.3
4079,00	-	4702,00	2.0	3.0
4702,00	-	99999,00	2.0	2.1

TESTE QUI-QUADRADO PARA O MÁXIMO ANUAL

NBL0COS = 33
DHIST = 0,3322417E+01
P (D>DHIST) = 0,39
NOTA = -0,2460643E+00

ESTATIS. SINT:	1149,89	792,30	1,46	5,91	3209,23	954,92	0,80	3,85	7245,00	0,730E+05
ESTATIS. DIST:	1146,34	741,43	1,52	6,56	3036,57	1046,08	1,03	3,87	5948,00	0,321E+05
MÍNIMOS:	1027,29	675,42	1,03	3,67	2700,67	628,53	-0,27	1,88	4120,00	0,929E+04
MÉDIAS:	1150,08	789,24	1,42	5,69	3210,52	922,33	0,62	3,29	5584,73	0,339E+05
MÁXIMOS:	1250,29	914,59	1,99	9,63	3573,03	1212,08	1,76	6,20	7245,00	0,730E+05
P(SINT>HIST):	0,4848	0,7879	0,2424	0,1818	0,7879	0,1818	0,0909	0,2424	0,3333	0,4545

HISTÓRICO 30 ANOS

IND1	IND2	IND3	IND4	IND5	IND6	IND7	IND8	IND9	IND10
------	------	------	------	------	------	------	------	------	-------

Na tabela 2 apresenta-se o comportamento do modelo para São José da Barra mostrando os valores de cada índice obtido nos 1000 anos sintéticos e nos 30 anos históricos. A seguir os 1000 anos sintéticos foram divididos em 33 blocos (NBLOCO=33), cada bloco com 30 anos. Os 10 índices foram calculados para cada bloco e na tabela 2 aparecem os valores mínimos, médios e máximos obtidos e a proporção de valores maiores que o valor obtido na série histórica. Estes valores revelam apenas para o índice 7, assimetria da vazão máxima, alguma discordância significativa entre as séries geradas pelo modelo e a série histórica. Escolhendo-se os índices 1, 5, 9 e 10 para uma análise multivariada, 39% das distâncias geradas são maiores que a distância histórica e a distância padronizada correspondente é negativa (-0,246) indicando um bom ajuste do modelo DIANA aos dados de vazões diárias afluentes ao posto de São José da Barra.

Os resultados encontrados para a avaliação das séries sintéticas de vazões diárias incrementais para o trecho entre os reservatórios de Furnas e Marimbondo aparecem na tabela 3. Na análise individual de cada índice verifica-se também a não-tipicidade apenas da assimetria das vazões máximas do registro histórico. A análise multivariada dos índices 1, 5, 9 e 10 encontra 60% das distâncias geradas maiores que a distância histórica e a distância padronizada correspondente novamente é negativa (-0,46). Pode-se concluir pelo bom desempenho do modelo DIANA quando aplicado as séries de vazões diárias afluentes ao reservatório de Furnas e incrementais entre este reservatório e o reservatório de Marimbondo.

AGRADECIMENTOS

Este artigo resulta das atividades de pesquisa dos projetos CEPTEL-7282, e CEPTEL-7275, que respectivamente recebem o apoio dos Departamentos de Operações e de Estudos Energéticos da ELETROBRÁS.

REFERÊNCIAS

- Beard, L.R.; "Simulation of Daily Streamflow" Hydrology Engineering Center, U.S. Army, Technical Paper No.6, 1968.
- Costa, C.C.; Oliveira, G.C.; Pereira, M.V.F.; Kelman, J.; "Seleção de Modelo de Afluências Mensais para Planejamento Energético", VI Seminário Nacional de Produção e Transmissão de Energia Elétrica, Camboriú, 1981.
- Fiering, M.B.; "Schemes for Handling Inconsistent Matrices", Water Resources Research, Vol. 4, No. 2, pp. 291-297, 1968.
- Kelman, J.; "Stochastic Modeling of Hydrologic Intermitent Processes", Colorado State University, Fort Collins, Colorado, Hydrology Paper No.89, 1976.
- Kelman, J.; Damázio, J.M.; "Synthetic Hydrology and Spillway Design", International Association for Hydraulic Research, XX Congresso, Moscou, 1983.
- Kelman, J.; Gomes, A.B.M.; Pinheiro, S.F.; Pereira, M.V.F.; "Revisão do Conceito de Energia Firme Através do Uso de Séries Hidrológicas Sintéticas", V Seminário Nacional de Produção e Transmissão de Energia Elétrica, Recife, 1979.
- Matalas, N.C.; "Mathematical Assessment of Synthetic Hydrology", Water Resources Research, Vol. 3, No.4, pp. 973-945, 1967.

O'Connell, P.E.; Jones, D.A.; "Some Experience with the Development of Models for the Stochastic Simulation of Daily Flows", Risk and Reliability Symposium, Waterloo, Canadá, June 26-28, 1978.

Reigier, M.A.; Harridan, M.A.; "Correlation in a Bivariate Normal Distribution With Truncation in both Variables", Australian Journal of Statistic, vol. 13, No.2, pp. 77-82, 1971.

ANEXO 1

ESTIMACÃO DA CORRELAÇÃO DE UMA DISTRIBUIÇÃO NORMAL PADRÃO BIVARIADA TRUNCADA COM NÍVEIS DE TRUNCAMENTO CONHECIDOS

Seja uma distribuição normal bivariada padrão com correlação ρ e truncada em $x=h$ e $y=k$. Segundo Rosenbaum (1961) vale a seguinte equação:

$$(h+k) \lambda^2 - \{(h+k)m_{11} - h k(m_{10} + m_{01})\} \rho - (h+k) - h k(m_{10} + m_{01}) + k m_{20} + h m_{02} = 0 \quad (A.1)$$

onde:

$$m_{10} = \int_h^\infty \int_k^\infty \frac{x^i \exp[-\frac{1}{2}(x^2 - 2\rho xy + y^2)/(1 - \rho^2)]}{2\pi\sqrt{1 - \rho^2} L(h,k,\rho)} dx dy$$

$$m_{01} = \int_h^\infty \int_k^\infty \frac{y^i \exp[-\frac{1}{2}(x^2 - 2\rho xy + y^2)/(1 - \rho^2)]}{2\pi\sqrt{1 - \rho^2} L(h,k,\rho)} dx dy$$

$$m_{11} = \int_h^\infty \int_k^\infty \frac{xy \exp[-\frac{1}{2}(x^2 - 2\rho xy + y^2)/(1 - \rho^2)]}{2\pi\sqrt{1 - \rho^2} L(h,k,\rho)} dx dy$$

$$L(h,k,\rho) = \int_h^\infty \int_k^\infty \frac{\exp[-\frac{1}{2}(x^2 - 2\rho xy + y^2)/(1 - \rho^2)]}{2\pi\sqrt{1 - \rho^2}} dx dy$$

Além disso, sabe-se que:

$$m_{10} L(h,k,\rho) = Z(h) Q\left[\frac{k - \rho h}{\sqrt{1 - \rho^2}}\right] + \rho Z(k) Q\left[\frac{h - \rho k}{\sqrt{1 - \rho^2}}\right] \quad (A.2)$$

$$m_{01}L(h, k, \rho) = \rho Z(h)Q \left[\frac{k - \rho h}{\sqrt{1 - \rho^2}} \right] + Z(k)Q \left[\frac{h - \rho k}{\sqrt{1 - \rho^2}} \right] \quad (A.3)$$

$$m_{20}L(h, k, \rho) = L(h, k, \rho) + hZ(h)Q \left[\frac{k - \rho h}{\sqrt{1 - \rho^2}} \right] + \rho^2 kZ(k)Q \left[\frac{h - \rho k}{\sqrt{1 - \rho^2}} \right] + \frac{\rho\sqrt{1 - \rho^2}}{\sqrt{2\pi}} Z \left[\frac{\sqrt{(h^2 - 2\rho hk + k^2)}}{(1 - \rho^2)} \right] \quad (A.4)$$

$$m_{02}L(h, k, \rho) = L(h, k, \rho) + \rho^2 hZ(h)Q \left[\frac{k - \rho h}{\sqrt{1 - \rho^2}} \right] + kZ(k)Q \left[\frac{h - \rho k}{\sqrt{1 - \rho^2}} \right] + \frac{\rho\sqrt{1 - \rho^2}}{\sqrt{2\pi}} Z \left[\frac{\sqrt{(h^2 - 2\rho hk + k^2)}}{1 - \rho^2} \right] \quad (A.5)$$

onde

$$Z(x) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2} \right)$$

$$Q(x) = \int_x^\infty Z(t) dt$$

O processo adotado para a estimação de ρ foi o seguinte:

a) Obtenha m_{11} a partir da amostra $\left(\frac{\sum x_i y_i}{n} \right)$

b) Arbitre ρ

c) Use as equações A-2 a A5 para obter m_{10}, m_{01}, m_{20} e m_{02} .

d) Obtenha ρ_1 e ρ_2 a partir da equação A.1 e faça $\rho^* = \rho_1$ se $|\rho - \rho_1| < |\rho - \rho_2|$ ou $\rho^* = \rho_2$ se $|\rho - \rho_2| \leq |\rho - \rho_1|$.

e) Se $\rho^* \neq \rho$ faça $\rho = \rho^*$ e volte para (c), se $\rho^* = \rho$ fim.

Observação: Pode ocorrer que o discriminante da equação A.1 seja menor que zero ou que $|\rho_i| > 0 \forall i$. Nestes casos, sugere-se usar outro ponto inicial para ρ (exemplo: $\rho/2$).