

CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ
Rio de Janeiro - RJ - Brasil
ISSN 1413-9022 / v. 21 p. 11 - 35, 2006

GERAÇÃO DE CENÁRIOS SINTÉTICOS DE ENERGIA E VAZÃO PARA O PLANEJAMENTO DA OPERAÇÃO ENERGÉTICA

Maria Elvira Piñeiro Maceira

Universidade do Estado do Rio de Janeiro - Centro de Pesquisas de Energia Elétrica
elvira@cepel.br

Débora Dias Jardim Penna

Centro de Pesquisas de Energia Elétrica
debora@cepel.br

Jorge Machado Damázio

Universidade do Estado do Rio de Janeiro - Centro de Pesquisas de Energia Elétrica
damazio@cepel.br

Resumo

A adoção de critérios probabilísticos no planejamento e operação de sistemas hidrotérmicos exige a modelagem probabilística de aflúências a aproveitamentos ou a subsistemas. O histórico de aflúências compõe uma amostra insuficiente para estimar índices de risco com incertezas aceitáveis. Por outro lado, as características básicas da série histórica podem ser capturadas por modelos estocásticos capazes de produzir séries sintéticas de aflúências diferentes da série histórica mas igualmente prováveis. Este trabalho descreve formulação e testes de validação do modelo auto-regressivo periódico, PAR(p), proposto para ser utilizado nos modelos Planejamento da Operação de Médio e Curto (NEWAVE, SUISHI-O e DECOMP) do Sistema Integrado Nacional.

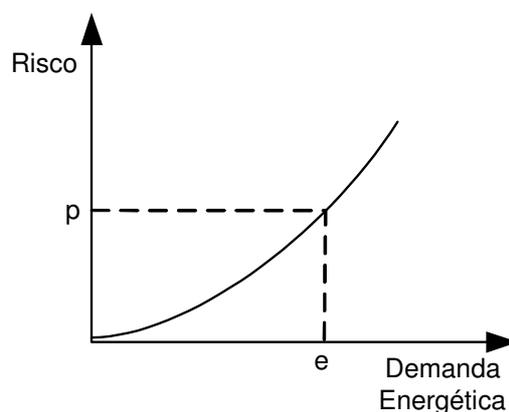
Palavras-chave: *Geração de Séries Sintéticas, Modelos Auto-regressivos, Modelos Estocásticos Multivariados.*

1. Aplicação de séries sintéticas no planejamento e operação de sistemas hidrotérmicos

A adoção de critérios probabilísticos em diversas atividades do planejamento e operação de sistemas hidrotérmicos criou a necessidade da modelagem probabilística de afluências a locais de aproveitamentos hidroelétricos ou a subsistemas. Em estudos energéticos, critérios de suprimento são baseados em índices de risco, estimados a partir da simulação da operação energética do sistema para diversos cenários de afluências aos aproveitamentos hidroelétricos ou subsistemas.

Por exemplo, um grande número de cenários pode ser utilizado para estimar a relação entre demanda energética e risco de não atendimento em um ano qualquer (energia garantida). A figura 1 mostra que um sistema hidroelétrico pode suprir a demanda e , que está associada ao risco p , previamente selecionado.

Figura 1 – Relação Demanda x Risco em um Ano Qualquer



O único cenário disponível na prática, o registro de afluências observado no passado (chamado de série histórica) é, no entanto, insuficiente para compor uma amostra de tamanho necessário para estimar índices de risco com incertezas aceitáveis. Entretanto, as características básicas da série histórica podem ser capturadas por modelos estocásticos capazes de produzir séries sintéticas de afluências, diferentes da série histórica mas igualmente prováveis. Dessa forma, a informação contida na série histórica pode ser mais completamente extraída, permitindo a avaliação de riscos e incertezas pertinentes a um sistema hidroelétrico.

Neste trabalho será descrito o modelo auto-regressivo periódico, PAR(p), proposto para ser utilizado no Modelo Estratégico de Geração Hidrotérmica a

Subsistemas Equivalentes Interligados - NEWAVE, no Modelo de Simulação a Usinas Individualizadas para Subsistemas Hidrotérmicos Interligados – SUIISHI-O e no Modelo de Determinação da Coordenação da Operação a Médio Prazo – DECOMP, todos utilizados pelo Operador Nacional do Sistema Elétrico (site: www.ons.org.br).

2. Geração de séries sintéticas para os modelos de planejamento da operação energética

O sistema de geração brasileiro é predominantemente hidráulico e possui acoplamento temporal e espacial. Estas características tornam o planejamento da operação energética um problema de grande porte e de difícil solução. Devido ao porte e complexidade deste problema é necessária a sua divisão em diversas etapas. Em cada etapa são utilizados modelos com diferentes graus de detalhamento para a representação do sistema, abrangendo períodos de estudos com horizontes distintos (médio prazo, curto prazo e programação diária), Maceira *et al.* (2002).

A aleatoriedade das vazões também é uma característica marcante de sistemas com predominância hidráulica. Essa incerteza é tratada de diferentes formas dependendo da representação utilizada no modelo do sistema gerador. A representação dos possíveis cenários de vazões é diferenciada para cada etapa do processo de planejamento da operação.

2.1. Representação das aflúncias no modelo de planejamento da operação de médio prazo

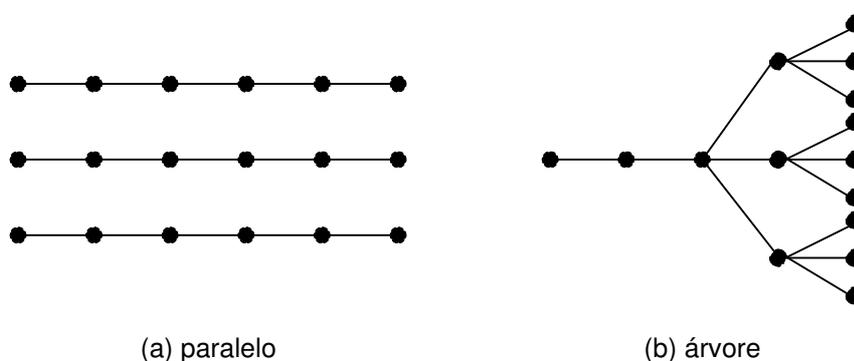
O modelo NEWAVE foi desenvolvido para o planejamento da operação de médio prazo. Ele define para cada mês do período de planejamento, que pode variar de 5 a 10 anos, a alocação ótima dos recursos hídricos e térmicos de forma a minimizar o valor esperado do custo de operação ao longo de todo o período de planejamento, Maceira (1999). O parque hidroelétrico é representado de forma agregada e o cálculo da política de operação é baseado na programação dinâmica dual estocástica. O NEWAVE apresenta quatro módulos básicos: módulo de cálculo dos sistemas equivalentes, módulo de cálculo do modelo estocástico de energias afluentes, módulo de cálculo da política de operação hidrotérmica e módulo de simulação da operação.

O modelo SUIISHI-O é um modelo de simulação a usinas individualizadas da operação energética de sistemas hidrotérmicos interligados. É capaz de simular

subsistemas hidrotérmicos eletricamente interligados em malha, mas hidráulicamente independentes, Maceira *et al.* (1999). Pode ser acoplado ao modelo de decisão estratégica NEWAVE pela função de custo futuro para cada estágio e pode considerar restrições de operação locais decorrentes do uso múltiplo da água.

Os modelos descritos acima simulam um grande número de séries hidrológicas, calculando assim, índices probabilísticos de desempenho do sistema para cada estágio da simulação. Estes diversos cenários de afluições estão em uma estrutura paralela (pente), figura 2(a), e são gerados pelo modelo GEVAZP.

Figura 2 – Séries em: (a) paralelo ,(b) árvore



2.2. Representação das afluições no modelo de planejamento da operação de curto prazo

O modelo DECOMP foi desenvolvido para aplicação no horizonte de curto prazo. Seu objetivo é determinar as metas de geração de cada usina de um sistema hidrotérmico, sujeito a afluições estocásticas, de forma a minimizar o valor esperado do custo de operação ao longo do período de planejamento, Cepel (2002). A incerteza acerca das vazões afluentes aos diversos aproveitamentos do sistema é representada através de cenários hidrológicos. Os cenários hidrológicos podem ser representados através de uma árvore de afluições, figura 2(b), com probabilidades de ocorrência associadas a cada ramo. Estes cenários são gerados a partir do modelo GEVAZP.

Devido à restrições de tempo computacional, no DECOMP é interessante trabalhar com o menor número de cenários hidrológicos possível e esta limitação não é ideal para o esquema de geração de Monte Carlo. Um número reduzido de séries não é suficiente para caracterizar bem um processo estocástico.

Para representar bem o processo estocástico de vazões no modelo de otimização do planejamento da operação a curto prazo – DECOMP - com um reduzido número de cenários empregou-se a análise de conglomerados. Uma descrição detalhada da metodologia pode ser encontrada em Maceira e Jardim (2001). Aplicadas à um grande número de cenários hidrológicos gerados, as técnicas de agrupamento proporcionam a escolha de um conjunto representativo de cenários. Este conjunto representativo de cenários hidrológicos irá conter toda a informação necessária para representar, dentro de uma dada precisão, o processo estocástico de vazões. Os cenários que fazem parte deste conjunto representativo são obtidos através de agrupamento de cenários semelhantes e possuem características similares aos demais componentes do grupo em que estão localizados. Dessa forma, as correlações temporal e espacial são preservadas.

3. Modelo auto-regressivo periódico

3.1. Descrição do modelo

Séries hidrológicas de intervalo de tempo menor que o ano, tais como séries mensais, têm como característica o comportamento periódico das suas propriedades probabilísticas, como por exemplo a média, a variância, a assimetria e a estrutura de auto-correlação. A análise deste tipo de séries pode ser feita pelo uso de formulações auto-regressivas cujos parâmetros apresentam um comportamento periódico. A esta classe de modelos costuma-se denominar modelos auto-regressivos periódicos, Salas *et al.* (1980). Estes modelos são referenciados por modelos PAR(p), onde p é a ordem do modelo, ou seja, o número de termos auto-regressivos do modelo. Em geral, p é um vetor, $p = (p_1, p_2, \dots, p_{12})$, onde cada elemento fornece a ordem de cada período.

O modelo PAR(p1, p2, ..., p12) pode ser descrito matematicamente por:

$$\left(\frac{Z_t - \mu_m}{\sigma_m} \right) = \phi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \quad (1)$$

onde:

Z_t é uma série sazonal de período s

s é o número de períodos (s = 12 para séries mensais)

t é o índice do tempo, $t = 1, 2, \dots, sN$, função do ano T (T = 1, 2, ..., N) e do período m (m = 1, 2, ..., s)

N é o número de anos

μ_m é a média sazonal de período s

σ_m é desvio-padrão sazonal de período s

ϕ_i^m é o i -ésimo coeficiente auto-regressivo do período m

p_m é a ordem do operador auto-regressivo do período m

a_t série de ruídos independentes com média zero e variância

Seja $\rho^m(k)$ a correlação entre Z_t e Z_{t-k} , de tal forma que t corresponda ao período m :

$$\rho^m(k) = E \left[\left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] \quad (2)$$

O conjunto de funções de autocorrelação $\rho^m(k)$ dos períodos $m = 1, \dots, s$, descrevem a estrutura de dependência temporal da série. Estas funções são dadas por Maceira (1989):

$$\begin{aligned} E \left[\left(\frac{Z_t - \mu_m}{\sigma_m} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] &= \phi_1^m E \left[\left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + \\ &\dots + \phi_{p_m}^m E \left[\left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] + E \left[a_t \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] = \rho^k(k) \end{aligned} \quad (3)$$

Conhecidos os parâmetros de um modelo PAR(p) as funções $\rho^m(k)$ são dadas pela solução de (3) e podem ser expressas por uma combinação de decaimentos exponenciais e/ou ondas senoidais, o que faz com que cada $\rho^m(k)$ tenda a zero à medida que k cresce.

Fixando-se m e variando k de 1 a p_m em (3) obtemos para cada período um conjunto de equações comumente denominado de equações de Yule-Walker. Para um período m qualquer:

$$\begin{bmatrix} 1 & \rho^{m-1}(1) & \rho^{m-1}(2) & \dots & \rho^{m-1}(p_m-1) \\ \rho^{m-1}(1) & 1 & \rho^{m-2}(1) & \dots & \rho^{m-2}(p_m-2) \\ \rho^{m-1}(2) & \rho^{m-2}(1) & 1 & \dots & \rho^{m-3}(p_m-3) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho^{m-1}(p_m-1) & \rho^{m-2}(p_m-2) & \rho^{m-3}(p_m-3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_1^m \\ \phi_2^m \\ \phi_3^m \\ \vdots \\ \phi_{p_m}^m \end{bmatrix} = \begin{bmatrix} \rho^m(1) \\ \rho^m(2) \\ \rho^m(3) \\ \vdots \\ \rho^m(p_m) \end{bmatrix} \quad (4)$$

Chamando o j -ésimo parâmetro auto-regressivo de um processo de ordem k , ϕ_{kk} é o último parâmetro deste processo. As equações de Yule-Walker para cada período m podem ser reescritas da seguinte forma:

$$\begin{bmatrix} 1 & \rho^{m-1}(1) & \rho^{m-1}(2) & \dots & \rho^{m-1}(p_m-1) \\ \rho^{m-1}(1) & 1 & \rho^{m-2}(1) & \dots & \rho^{m-2}(p_m-2) \\ \rho^{m-1}(2) & \rho^{m-2}(1) & 1 & \dots & \rho^{m-3}(p_m-3) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho^{m-1}(k-1) & \rho^{m-2}(k-2) & \rho^{m-3}(k-3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1}^m \\ \phi_{k2}^m \\ \phi_{k3}^m \\ \vdots \\ \phi_{kk}^m \end{bmatrix} = \begin{bmatrix} \rho^m(1) \\ \rho^m(2) \\ \rho^m(3) \\ \vdots \\ \rho^m(k) \end{bmatrix} \quad (5)$$

Ao conjunto de valores ϕ_{kk}^m , $k = 1, 2, \dots$, chamamos de função autocorrelação parcial do período m . O conjunto de funções ϕ_{kk}^m , $m = 1, 2, \dots, s$, é uma outra forma de representar a estrutura de dependência do processo estocástico ao longo do tempo. Em um processo auto-regressivo de ordem p_m , a função de autocorrelação parcial ϕ_{kk}^m será diferente de zero para k menor ou igual a p_m e zero para k maior que p_m .

A expressão de a_t , válida para qualquer período m , é dada por:

$$\sigma_a^{2(m)} = 1 - \phi_1^m \rho^m(1) - \phi_2^m \rho^m(2) - \dots - \phi_{p_m}^m \rho^m(p_m) \quad (6)$$

3.2. Ajuste do modelo

Box e Jenkins, Box e Jenkins (1970), sugeriram uma metodologia bastante elaborada para ajuste de modelos estocásticos da família ARIMA a séries temporais, que pode ser estendida para modelos da família PAR(p). Nesta metodologia a estratégia de seleção do modelo é dividida em três etapas. A primeira etapa, denominada por Box e Jenkins de *identificação do modelo*, consiste em escolher, tentativamente, a ordem do

modelo baseando-se em estimativas das funções $\rho^m(k)$ e $\hat{\phi}_{kk}^m$ obtidas a partir da série amostral. Na modelagem auto-regressiva periódica isso consiste em escolher o vetor p . A segunda etapa refere-se à **estimação do modelo**, ou seja, estimação dos seus parâmetros, sendo em geral recomendado o uso de estimadores de máxima verossimilhança ou suas aproximações. A terceira etapa diz respeito à **verificação do modelo**, isto é, verificar através de testes estatísticos se as hipóteses assumidas durante as etapas anteriores são atendidas. Se as hipóteses não são verificadas deve-se retornar à primeira etapa até que os resultados sejam satisfatórios.

Ressalta-se que esta estratégia pode muitas vezes resultar em mais de um modelo capaz de descrever o processo estocástico em estudo. Nestes casos a seleção do modelo mais adequado pode ser feita submetendo-os a testes de aplicação similares aos usados em Kelman e Maceira (1992) e em Oliveira *et al.* (1985).

3.2.1. Identificação

A identificação do modelo consiste em determinar as ordens mais apropriadas dos operadores autoregressivos de cada período p_m , $m = 1, \dots, s$. Isto pode ser feito obtendo-se estimativas $\hat{\phi}_{kk}^m$, $k = 1, \dots, N/4$ e substituindo em (5) as autocorrelações pelos respectivos valores amostrais. Se a ordem do operador autoregressivo de um período qualquer m é p_m então $\hat{\phi}_{kk}^m$ para $k > p_m$ tem distribuição aproximadamente Normal com média zero e variância N^{-1} , Noakes *et al.* (1983 e 1985). Para cada período m procura-se a maior ordem i tal que todas as estimativas $\hat{\phi}_{kk}^m$ para $k > i$ não sejam mais significativas.

3.2.2. Estimação

Após a etapa de identificação é necessário obter estimativas para os diversos parâmetros do modelo. Para modelos auto-regressivos os estimadores de momento são em geral bastante eficientes, Harvey (1981).

Os parâmetros ϕ_i^m , $i = 1, \dots, p_m$, são estimados substituindo-se em (4) $\rho^{m-j}(k)$, $j = 0, \dots, (p_m - 1)$, $k = 1, \dots, p_m$, por suas estimativas amostrais.

Observa-se que os parâmetros do modelo para o m -ésimo período podem ser estimados de maneira independente dos parâmetros de qualquer outro período.

Cada um dos m sistemas resultantes pode ser resolvido por Decomposição de Cholesky.

Finalmente, as estimativas de $\sigma_a^{2(m)}$ podem ser obtidos usando-se a expressão (6).

3.3. Vazões incrementais negativas

Um problema freqüente em sistemas com reservatórios em cascata, onde as vazões incrementais podem ser muito pequenas, é a geração de vazões mensais negativas. Para obter-se um Z_t positivo é necessário que:

$$Z_t = \mu_m + \phi_1^m \sigma_m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \sigma_m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + \sigma_m a_t > 0 \quad (7)$$

ou

$$a_t > - \left(\frac{\mu_m}{\sigma_m} \right) - \phi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) - \dots - \phi_{p_m}^m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \quad (8)$$

Muitos pesquisadores assumem que os resíduos a_t apresentam distribuição Normal e uma possível não-normalidade pode ser corrigida pela transformação Box-Cox, Box e Cox (1964). Como as séries sintéticas produzidas serão utilizadas em modelos que calculam as estratégias ótimas de operação de um sistema multi-reservatórios, baseados em programação dinâmica dual estocástica, o modelo de geração de séries sintéticas deve ser aplicado diretamente à série temporal original e deve ser capaz de lidar com resíduos que apresentam um forte coeficiente de assimetria.

A solução adotada foi ajustar uma distribuição Lognormal com três parâmetros aos resíduos mensais a_t , Maceira e Mercio (1997). Consequentemente, a variável ξ_t tem distribuição Normal com média zero e desvio-padrão igual a $\sigma_\xi^{2(m)}$:

$$\xi_t = \ln (a_t - \Delta) \quad (9)$$

Os parâmetros Δ e $\sigma_\xi^{2(m)}$ são estimados de tal forma a preservar os momentos dos resíduos, conforme Charbeneau (1978).

3.4. Correlação espacial

A fim de se gerar vazões mensais multivariadas assume-se que os resíduos espacialmente não correlacionados, ξ_t , podem ser transformados em resíduos espacialmente correlacionados, W_t , da seguinte forma:

$$W_t = D\xi_t \quad (10)$$

onde D é uma matriz quadrada de dimensão igual ao número de usinas hidroelétricas.

A matriz D pode ser estimada por:

$$DD^T = \hat{U} \quad (11)$$

onde \hat{U} pode ser, em princípio, a estimativa da matriz de covariâncias $[E(\hat{\xi}_t, \hat{\xi}_t^T)]$. Por exemplo, o elemento u_{ij} é a covariância espacial lag zero entre os resíduos das usinas i e j . Na prática, o comportamento dos resíduos não segue o comportamento das vazões: os resíduos não são espacialmente correlacionados. No entanto, a fim de se preservar as dependências espaciais entre as usinas utilizou-se a correlação espacial entre as vazões em substituição à correlação espacial entre os resíduos. A matriz D pode ser estimada adotando-se D como uma matriz triangular inferior ou por decomposição espectral.

4. Aplicação das técnicas de agregação ao modelo de geração de séries de vazões

No modelo de planejamento da operação de curto prazo (modelo DECOMP), a incerteza acerca das vazões afluentes às diversas usinas hidrelétricas do sistema é representada por cenários hidrológicos através de uma árvore de afluências, ilustrada na figura 2(b), com probabilidades de ocorrência associadas a cada ramo. Admite-se que as probabilidades de ocorrência são as mesmas para cada ramo, isto é, os cenários são equiprováveis. Para que o processo estocástico de afluências seja bem representado, a árvore de afluências deve conter muitos ramos, pois poucos cenários hidrológicos não conseguem descrever adequadamente as características do processo estocástico de afluências. Porém a árvore não pode conter muitos ramos uma vez que isto tornaria a solução do problema de planejamento de curto prazo inviável por necessitar de muito esforço computacional. A restrição de tempo computacional pode ser contornada utilizando-se Análise de Conglomerados. O principal objetivo da aplicação das técnicas

de agregação ao modelo GEVAZP é a redução do número de cenários hidrológicos através da escolha de um conjunto representativo de cenários hidrológicos. Os cenários são obtidos através de agrupamento de cenários semelhantes, possuem características similares aos demais componentes do grupo em que estão localizados e têm pesos diferentes, de acordo com o número de cenários contidos no grupo, Jardim *et al.* (2001).

4.1. Método de agrupamento utilizado

A cada estágio do horizonte de estudo e a cada ramo da árvore de aflúências são gerados N cenários hidrológicos distintos, por exemplo $N=1000$, e equiprováveis para o processo de agregação. O número de aberturas em cada estágio é o número de agrupamentos obtidos a partir dos N cenários, onde N é um dado de entrada. O método escolhido foi o método de agrupamento não hierárquico K-Means.

A métrica utilizada para calcular a similaridade entre os objetos (cenários hidrológicos multivariados) e os centróides dos grupos é a distância Mahalanobis. Esta métrica é definida como em (12).

$$d_{i,j} = \sqrt{(Z_i - Z_j)' \Sigma^{-1} (Z_i - Z_j)} \quad (12)$$

onde Σ^{-1} é a inversa da matriz de covariância e Z_i é um vetor de vazões multivariadas (cenário hidrológico i)

Quando a matriz de covariância for igual a matriz identidade, isto é a correlação entre as variáveis for nula, a distância de Mahalanobis é equivalente a distância Euclidiana. Além de ponderar pela variabilidade de cada uma das variáveis, esta medida de distância considera também o grau de correlação entre elas. Utilizando esta métrica pretende-se preservar além da média e do desvio padrão, a correlação espacial entre as usinas hidrelétricas.

As variáveis que descrevem os objetos são padronizadas pela média e desvio padrão mensal. Antes da padronização dos dados, é aplicada a transformação logarítmica nos vetores de vazões. Geralmente, os objetos mais distantes são representantes de cenários com vazões elevadas e, quando não se aplica a transformação logarítmica nos dados, estes objetos tendem a se isolar em um grupo. Por outro lado, cenários com baixas aflúências têm pouca dispersão entre si e tendem a se localizar em um mesmo grupo. Com a transformação logarítmica, a formação de alguns grupos com poucos cenários e outros com muitos é evitada. Desta maneira, tantos os cenários favoráveis

(úmidos) quanto os cenários críticos (secos) são representados de forma mais equilibrada.

O processo de agregação é inicializado através do sorteio de sementes aleatórias para representar os centróides dos grupos. Estes pontos iniciais são objetos do conjunto de entrada escolhidos aleatoriamente. Logo, são cenários hidrológicos do conjunto original gerado sinteticamente. Desta maneira, pode-se garantir que nenhum grupo ficará vazio. Nos passos seguintes até a convergência do processo de agregação, o centróide dos grupos (C_k) será o ponto médio destes grupos, isto é:

$$C_k = \frac{1}{NO_k} \sum_{i=1}^{NO_k} Z_i \quad (13)$$

onde NO_k é o número de cenários hidrológicos pertencentes ao grupo k .

Após a convergência do processo, o centróide dos grupos será o objeto mais próximo do ponto médio deste grupo.

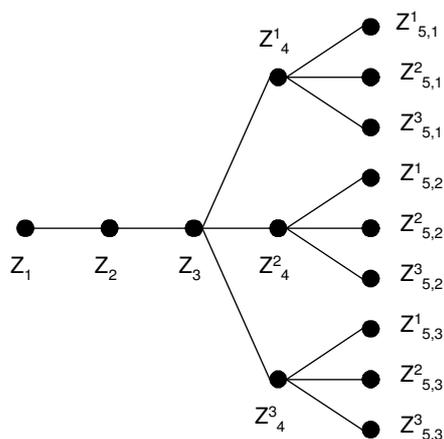
Cada cenário representante tem associado a ele uma probabilidade de ocorrência. A probabilidade P_k associada ao grupo k é calculada como a razão entre o número de cenários hidrológicos pertencentes ao grupo k e o número de cenários equiprováveis de vazões multivariadas gerados antes do processo de agregação (N).

4.2. Estrutura do novo processo de geração de cenários sintéticos em árvore

Supondo-se, por exemplo, que a cada estágio um nó dê origem à três novos nós na estrutura da árvore, e que para todos os meses o processo estocástico de afluições seja representado por um modelo auto-regressivo de ordem três. No primeiro estágio serão gerados N ($N=1000$, por exemplo) cenários equiprováveis de vazões multivariadas a partir de uma única seqüência de vazões multivariadas conhecidas (Z_3, Z_2, Z_1). A seguir, o processo de agregação é aplicado resultando em três cenários não equiprováveis de vazões multivariadas (Z^1_4, Z^2_4, Z^3_4). No segundo estágio, para a seqüência (Z^1_4, Z_3, Z_2), são gerados N cenários equiprováveis de vazões multivariadas. Do processo de agregação três cenários não equiprováveis de vazões multivariadas são produzidos ($Z^1_{5,1}, Z^2_{5,1}, Z^3_{5,1}$). Ainda neste estágio, a partir da seqüência (Z^2_4, Z_3, Z_2) são gerados N cenários equiprováveis de vazões multivariadas. Da mesma forma, o processo de agregação é aplicado e três cenários não equiprováveis de vazões multivariadas são obtidos ($Z^1_{5,2}, Z^2_{5,2}, Z^3_{5,2}$). Por último, a partir da seqüência ($Z^3_4, Z_3,$

Z_2), são gerados N cenários de vazões multivariadas. O processo de agregação é novamente aplicado resultando em três cenários não equiprováveis de vazões multivariadas ($Z_{5,3}^1, Z_{5,3}^2, Z_{5,3}^3$). A árvore de afluições gerada é mostrada na figura 3.

Figura 3 – Caso Exemplo (Árvore Completa)



5. Critérios para testar a adequação do modelo

5.1. Testes para séries geradas em paralelo (pente)

5.1.1. Verificação das estatísticas anuais e mensais

Qualquer característica do processo que tenha sido usada na definição do modelo estocástico de afluições não pode ser utilizada para testar a adequação do modelo. Por exemplo, se as médias sazonais são parâmetros do modelo que está sendo analisado, é comum comparar as médias históricas e geradas: já se sabe "a priori" que elas serão preservadas. Na prática, os momentos são verificados na tentativa de achar inconsistências ou desvios.

5.1.2. Avaliação do desempenho do modelo

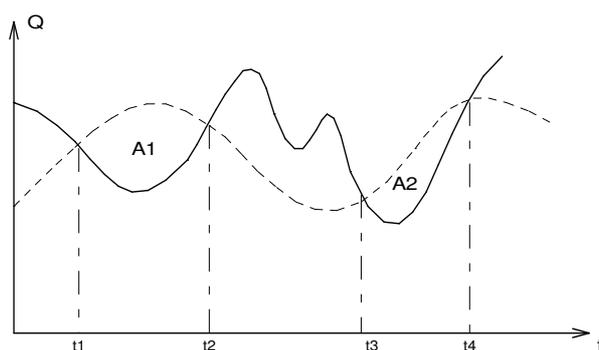
Um modelo de geração de séries sintéticas deve preservar as principais características da série observada. Em termos práticos, a utilidade de um modelo pode ser aferida por sua capacidade de reproduzir distribuições de probabilidade de variáveis aleatórias relevantes ao processo. Estas variáveis devem naturalmente refletir os requisitos da aplicação proposta. No caso do planejamento da expansão de um sistema de geração, a representação de períodos críticos, as correlações serial e espacial entre as

vazões afluentes aos diversos reservatórios, curvas de regularização para diversos reservatórios etc., devem de alguma forma ser levados em conta nesta avaliação, Kelman *et al.* (1977).

Em geral, aceita-se que as variáveis aleatórias selecionadas devem estar relacionadas com o conceito de **sequência**: uma sequência negativa é o período de tempo em que as vazões afluentes estão continuamente abaixo de valores pré-determinados, por exemplo, as médias mensais, precedidos e sucedidos por valores acima destes limites.

A figura 4 ilustra o conceito de sequência e das variáveis associadas. A linha contínua representa afluências e a linha pontilhada um limite pré-estabelecido. Os intervalos $(t_1 - t_2)$ e $(t_3 - t_4)$ correspondem às sequências negativas, isto é, as vazões estão abaixo dos limites.

Figura 4 - Sequência Negativa



As variáveis associadas são:

- Comprimento de sequência: $C = t_2 - t_1$, corresponde ao comprimento dos intervalos $(t_2 - t_1)$, $(t_4 - t_3)$, na figura 4;

- Soma de sequência: $S = \sum_{i=t_1}^{t_2} (Z_i - \hat{\mu}_i)$, corresponde à área abaixo do limite durante a sequência. Na figura 4, as somas de sequências são as áreas A_1 e A_2 .;

- Intensidade de sequência: $I = \frac{S}{C}$, corresponde ao valor médio abaixo do limite, isto é, a soma de sequência dividida pelo respectivo comprimento de sequência.

Repetindo estes cálculos para cada sequência negativa é possível obter da série histórica amostras de três variáveis aleatórias: soma de sequência negativa, comprimento de sequência negativa e intensidade de sequência negativa.

O mesmo procedimento é realizado para a série sintética, onde o limite pré-determinado é o mesmo usado na série histórica, e obtemos outras três amostras. Podemos, então, testar a hipótese de que as duas amostras são provenientes de distribuições idênticas. Seja n_1 o número de observações da variável aleatória X , obtida a partir da série histórica e n_2 o número de observações da variável aleatória Y , obtida a partir da série gerada. Caso essas variáveis aleatórias sejam discretas (ex.: comprimento de sequência) utilizamos o teste de igualdade de duas distribuições multinomiais Mood *et al.* (1974).

Caso as variáveis aleatórias em estudo sejam contínuas, por exemplo, a soma e intensidade de sequência, o teste de Smirnov para duas amostras é bastante adequado. Descrição detalhada é fornecida por Bradley (1968).

Outra variável aleatória que pode ser considerada na análise de desempenho é a de *somas parciais* que pode ser definida como a sequência $\{S_t\}$, $t = 1, 2, \dots, 12N$, onde:

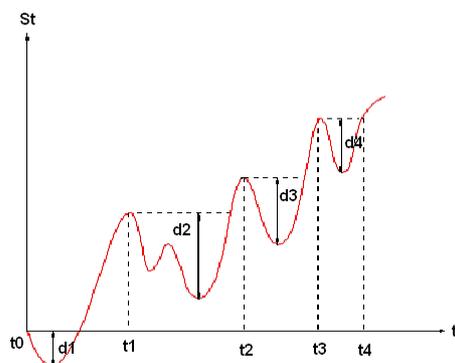
$$S_t = \sum_{j=1}^t (Z_j - \beta\bar{\mu}) \quad (14)$$

onde β é chamado de nível de regularização.

Neste caso a sequência de somas parciais é dada por (15) e está ilustrada na figura 5.

$$S_t = \sum_{j=1}^t Z_j - \beta t\bar{\mu} \quad (15)$$

Figura 5 - Volume de regularização (Déficit)



O período t_0 a t_1 só transcorrerá sem ocorrência de deficit caso esteja disponível uma capacidade de armazenamento igual ou maior que $d_1 \cdot \Delta t$, onde Δt é o intervalo de discretização. Analogamente para o intervalo (t_1, t_2) é necessário uma capacidade de armazenamento no mínimo igual a $d_2 \cdot \Delta t$ e assim sucessivamente. Portanto, as grandezas d_1, d_2, \dots , que são observações de uma variável aleatória chamada deficit, tem grande interesse nos estudos de regularização das aflúências.

Um modelo orientado para aplicação no setor energético deve estar apto a produzir séries sintéticas que revelem uma distribuição amostral para a variável aleatória *déficit* estatisticamente indistinguível da distribuição amostral obtida a partir da série histórica. Verificou-se, entretanto, que os índices mais relevantes para o planejamento estavam em geral associados a valores extremos das distribuições. O primeiro índice do tipo "máximo" a ser definido foi o chamado máximo déficit, que corresponde ao volume do reservatório capaz de regularizar uma vazão especificada.

Para estudo das variáveis do tipo "máximo", o procedimento é o seguinte: divide-se a série gerada em M segmentos de comprimento igual ao da série histórica e calcula-se o valor do índice para cada segmento, obtendo-se assim uma distribuição de frequências.

Como a série histórica possui apenas um valor, não faz sentido falar de aderências de distribuições, mas sim da "tipicidade" do valor histórico em relação à distribuição dos valores gerados. Em outras palavras, deseja-se saber a probabilidade da amostra histórica ser sorteada dado que o modelo de geração escolhido é verdadeiro .

Em termos univariados, o desempenho do modelo pode ser medido pela proporção de índices gerados maiores ou menores do que o índice histórico. Se esta proporção for muito pequena, isto é uma indicação de que a observação histórica é atípica para o modelo considerado.

Esta análise pode ser realizada para os seguintes índices: máximo comprimento de seqüência, máxima soma de seqüência, máxima intensidade de seqüência, máximo déficit para um determinado nível de regularização, comprimento do período crítico e vazão média de período crítico. Maiores detalhes podem ser encontrados em Maceira (1989).

5.2. Testes para séries geradas em árvore

Alguns testes foram implementados para verificar se os cenários hidrológicos resultantes do processo de agregação conseguem manter as mesmas características do conjunto de cenários gerados originalmente e, se os cenários gerados estão coerentes com o modelo ajustado. Foram realizados testes para avaliar os momentos mensais condicionados de primeira e segunda ordem para os valores gerados, os valores de média, desvio padrão e correlação espacial entre as usinas hidrelétricas resultantes dos cenários gerados e agregados e a aderência entre as distribuições marginais das vazões geradas e agregadas.

5.2.1. Teste de momentos condicionados de ordem um e dois

Os valores teóricos correspondentes à média e ao desvio-padrão condicionados podem ser derivados dos modelos estocásticos mensais ajustados às usinas hidrelétricas. Estes valores devem ser utilizados como referência de comparação para avaliar os valores de média e desvio-padrão estimados a partir dos N cenários hidrológicos gerados. Através deste teste, será possível verificar se os cenários gerados são coerentes com o modelo estocástico usado para gerar a árvore.

A média de Z_t condicionada aos valores passados $Z_{t-} = \{Z_{t-1}, Z_{t-2}, \dots, Z_{t-p_m}\}$ é dada por:

$$E(Z_t | Z_{t-}) = \left[\phi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \right] \sigma_m + \mu_m \quad (16)$$

A variância de Z_t condicionada aos valores passados Z_{t-} é dada por:

$$\text{VAR}(Z_t | Z_{t-}) = \sigma_m^2 \sigma_a^{2(m)} = \sigma_m^2 \left[1 - \phi_1^m \rho^m(1) - \phi_2^m \rho^m(2) - \dots - \phi_{p_m}^m \rho^m(p_m) \right] \quad (17)$$

Parâmetros das distribuições amostrais

Em Yevjevich, (1972) são apresentados os parâmetros da distribuição amostral para a média e o desvio-padrão. Estes parâmetros são necessários para verificar se os momentos condicionados de primeira e segunda ordem obtidos a partir dos N cenários gerados são condizentes com os valores teóricos calculados para estes momentos. O valor esperado da média amostral, $E(\bar{x})$, é igual à média da população (16) e a variância da média amostral, $\text{VAR}(\bar{x})$, é igual à variância da população (17) dividida pelo número de elementos da amostra. Pelo teorema do limite central pode-se dizer que a

média tem distribuição Normal. Então o intervalo de confiança de 95%, calculado para a média amostral estimada a partir dos N cenários hidrológicos gerados, para uma dada usina hidrelétrica, é dado por:

$$IC_{\mu} = E(Z_t | Z_{t-}) \pm 1,96 \frac{\sigma_{Z_t|Z_{t-}}}{\sqrt{N}} \quad (18)$$

O valor esperado do desvio padrão amostral, $E(s)$, é igual ao desvio padrão da população, $\sigma_{Z_t|Z_{t-}}$. A variância do desvio padrão amostral, $VAR(s)$, é definida em (19). A definição do momento central de quarta ordem de população correspondente à distribuição Lognormal pode ser encontrada em Yevjevich, (1972).

$$VAR(s) = \frac{\mu_4 - \mu_2^2}{4\mu_2 N} = \frac{\mu_2(\eta^8 + 6\eta^6 + 15\eta^4 + 16\eta^2 + 2)}{4N} \quad (19)$$

onde μ_2 e μ_4 são os momentos centrados de segunda e quarta ordem da população e $\eta = \sqrt{e^{\sigma_n^2} - 1}$ é o coeficiente de variação da população e σ_n^2 é o parâmetro de forma da Distribuição Lognormal 3 parâmetros ajustada para a geração dos cenários de afluências.

Para estabelecer um intervalo de confiança de 95%, para a estimativa do desvio padrão, foi calculada uma tabela com valores-padrão, VP.

$$IC_{\sigma} = \sigma_{Z_t|Z_{t-}} + VP \frac{\sigma_{Z_t|Z_{t-}} \sqrt{\eta^8 + 6\eta^6 + 15\eta^4 + 16\eta^2 + 2}}{\sqrt{4N}} \quad (20)$$

5.2.2. Teste de aderência

Testes de aderência, também conhecidos como testes de adequação de ajuste ou testes de adaptabilidade, têm como objetivo determinar se uma certa distribuição postulada é razoável na presença de dados coletados (amostra), isto é, diz a respeito ao grau de concordância entre a distribuição da amostra e da população da qual esta foi extraída. Estes testes pressupõem que as amostras sejam independentes, o que não ocorre com os cenários gerados e agregados. Estas amostras são dependentes pois os cenários agregados são obtidos a partir dos cenários gerados. A utilização destes testes, para comparar as distribuições de probabilidades obtidas através dos cenários gerados e agregados, não é feita de maneira formal porém fornece uma indicação sobre o grau de

adequação entre estas duas distribuições. Pode-se citar os testes de Cramér–von Mises e o teste de Kolmogorov-Smirnov.

5.2.3. Comparação das estatísticas geradas e agregadas

O objetivo desta comparação é testar se as estatísticas média, desvio padrão e correlação espacial obtidas para cada usina hidrelétrica a partir dos cenários agregados estão de acordo com as mesmas estatísticas obtidas a partir dos N cenários gerados. Testes de hipóteses formais não podem ser usados neste caso, pois o conjunto dos cenários agregados advém dos cenários gerados, o que faz com que estas amostras sejam dependentes.

Este teste é baseado no ajuste de uma reta aos valores das estatísticas de cada usina hidrelétrica, dos cenários gerados e agregados, através de um modelo de regressão linear de equação única, (21). A reta ajustada deve sempre passar pela origem dos eixos.

$$X_{AG} = \beta X_{GER} \quad (21)$$

onde $X_{AG}^t = [x_{AG}^1 \quad x_{AG}^2 \quad x_{AG}^3 \quad \dots \quad x_{AG}^p]$ e $X_{GER}^t = [x_{GER}^1 \quad x_{GER}^2 \quad x_{GER}^3 \quad \dots \quad x_{GER}^p]$ são as estatísticas (média ou desvio padrão ou correlação cruzada) obtidas a partir dos cenários hidrológicos agregados e gerados, respectivamente, para cada uma das P usinas hidrelétricas consideradas na configuração do sistema.

De acordo com estas estatísticas é traçado um diagrama de dispersão e então, a reta de regressão é obtida. Se os cenários agregados estão representando satisfatoriamente os cenários gerados, espera-se que o coeficiente angular β da reta ajustada esteja próximo de 1 e que o coeficiente de determinação R^2 , que mede o grau de ajuste da reta, esteja próximo de 100%.

6. Estudo de caso

A aplicação da metodologia é ilustrada utilizando-se uma configuração hidrelétrica baseado no Plano Mensal de Operação de Maio/2005, composta por 94 usinas hidrelétricas localizadas ao longo de diversas bacias. Os dados utilizados estão disponíveis no site do ONS. Apenas dois reservatórios são analisados nesse trabalho: Água Vermelha (Rio Grande) e Três Marias (Rio São Francisco). Usando o algoritmo descrito previamente foram geradas 200 séries sintéticas de vazões afluentes com 70 anos cada. O histórico de vazões naturais afluentes compreende o período 1931 a 2003.

Na figura 6 são mostrados média e desvio-padrão mensais para os aproveitamentos listados acima, estimados a partir do registro histórico e das séries sintéticas de vazões naturais afluentes. A figura 7 mostra o coeficiente de correlação lag 1 mensal (a) e a correlação espacial entre as usinas Água Vermelha e Três Marias (b). Observa-se que o modelo ajustado consegue preservar de forma aproximada os momentos observados no histórico.

Figura 6 – Média e desvio padrão mensal (a)Água Vermelha, (b)Três Marias

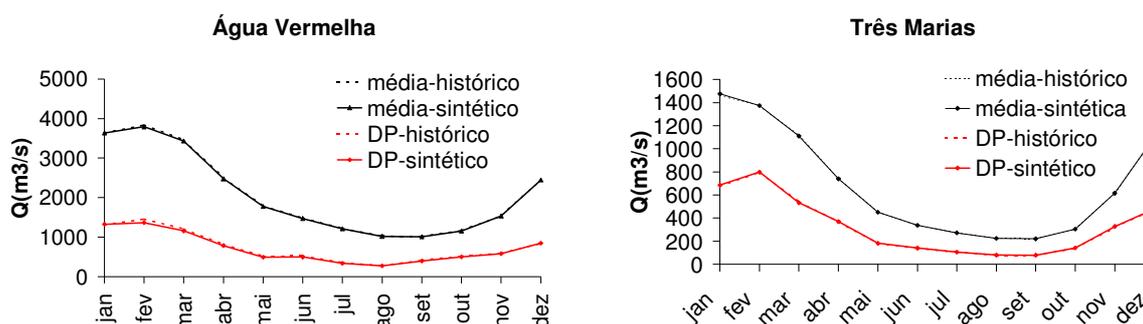
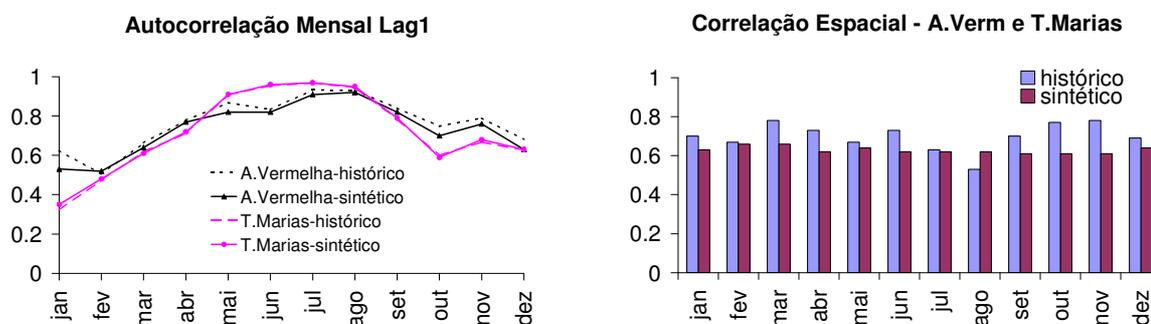


Figura 7 – Correlação mensal (a) lag1, (b)espacial



A tabela 1 mostra a correlação anual lag 1 estimada a partir da amostra sintética e do registro histórico. A performance do modelo pode ser considerada adequada em todos os aspectos analisados até o momento.

Tabela 1 – Autocorrelação anual lag 1

Autocorrelação Anual	Água Vermelha	Três Marias
Histórico	0.29	0.24
PAR(p)	0.27	0.17

A validação do modelo utiliza testes de aderência para as distribuições das variáveis soma de seqüência negativa, comprimento de seqüência negativa e intensidade de seqüência negativa. Os resultados para as usinas de Água Vermelha e Três Marias são apresentados nas tabelas 2 a 4. Nessas tabelas são mostrados os resultados obtidos nos testes de aderência assim como o valor crítico associado (entre parêntesis). Para as distribuições das variáveis soma de seqüência negativa e intensidade de seqüência negativa o teste de Kolmogorov-Smirnov foi utilizado, e para a variável comprimento de seqüência negativa foi aplicado um teste multinomial. Observa-se que o modelo é adequado com relação à todas variáveis analisadas.

Tabela 2 – Testes de Aderência para Soma de Seqüência Negativa (109 m3)

Usina	Fonte de dados	Média	Desvio Padrão	Valor Crítico
Água Vermelha	Histórico	2.16	4.93	(0.129)
	PAR(p)	2.62	4.60	0.112
Três Marias	Histórico	1.05	1.42	(0.134)
	PAR(p)	1.20	1.84	0.064

Tabela 3 – Testes de Aderência para Comprimento de Seqüência Negativa (meses)

Usina	Fonte de dados	Média	Desvio Padrão	Valor Crítico
Água Vermelha	Histórico	4.3	5.9	(11.100)
	PAR(p)	4.8	5.9	8.206
Três Marias	Histórico	4.8	5.0	(14.100)

Tabela 4 – Testes de Aderência para Intensidade de Seqüência Negativa (109 m3 / mês)

Usina	Fonte de dados	Média	Desvio Padrão	Valor Crítico
Água Vermelha	Histórico	0.37	0.28	(0.129)
	PAR(p)	0.39	0.31	0.071
Três Marias	Histórico	0.19	0.15	(0.134)
	PAR(p)	0.19	0.15	0.057

Tabela 5 – Máximo Déficit para Água Vermelha (109 m3)

Vazão Defl. Regulada (% média)	Fonte de dados	Máximo Déficit		Nº de seq. com valores abaixo do histórico
		Média	Desvio Padrão	
70%	Histórico	14.19		
	PAR(p)	11.32	3.59	164
75%	Histórico	19.74		
	PAR(p)	14.77	4.71	169
80%	Histórico	25.38		
	PAR(p)	19.43	6.31	166

\

Tabela 6 – Máximo Déficit para Três Marias (109 m3)

Vazão Defl. Regulada (% média)	Fonte de dados	Máximo Déficit		Nº de seq. com valores abaixo do histórico
		Média	Desvio Padrão	
70%	Histórico	4.3		
	PAR(p)	5.74	1.70	33
75%	Histórico	5.2		
	PAR(p)	7.35	2.25	30
80%	Histórico	7.04		
	PAR(p)	9.45	3.00	40

Nas tabelas 5 e 6 são apresentados os resultados do cálculo da variável máximo déficit para Água Vermelha e Três Marias. Pode ser visto que as proporções de máximo déficit produzidas pelo modelo nos dois casos são adequadas. O modelo conseguiu produzir secas mais severas do que aquelas observadas no registro histórico

7. Conclusão

Este artigo descreve a formulação e os procedimentos de construção e validação do modelo auto-regressivo periódico PARP(p) multivariado, desenvolvido para aplicações na modelagem probabilística das afluências no planejamento da operação do sistema interligado de geração elétrica brasileiro.

O modelo é aplicado diretamente na série de vazões afluentes evitando assim o uso de transformações especiais, tal como a transformação Box-Cox. Esta característica permite que o modelo seja usado em algoritmos utilizados para calcular a estratégia

ótima de operação de sistemas hidrotérmicos, baseados em programação dinâmica dual estocástica. A geração de vazões afluentes negativas também é evitada.

A metodologia para verificação da adequação do modelo também foi descrita e aplicada em um caso estudo. Os resultados indicaram que a performance do modelo foi muito boa, pois esse conseguiu reproduzir características importantes da série histórica de aflúências e, além disso, produziu secas mais severas do que aquelas encontradas no registro histórico.

Referências

- BOX, G.E.P., COX, D.R. "An Analysis of Transformations". **Journal of The Royal Statistical Society**, A127, pp. 211-252, 1964.
- BRADLEY, J. **Distribution Free Statistical Tests**, Prentice Hall Inc, 1968.
- BOX, G.E.P., JENKINS, G.M. **Time Series Analysis - Forecasting and Control**, Holden-Day, 1970.
- CEPEL (2002). "Modelo DECOMP - Manual de Referência Versão 11.0", **Relatório Técnico CEPEL** nº DPD/PEN- 15053/03.
- CHARBENEAU, R. Comparison of the two and three parameter lognormal distributions used in streamflow synthesis. **Water Resources Research**, Vol. 14, No. 1, pp. 149-150, 1978.
- HARVEY, A.C. **Time Series Models**, P. Allan, 1981.
- JARDIM, D.L.D.D.; MACEIRA, M.E.P.; FALCÃO, D.M. Stochastic Streamflow Model for Hydroelectric Systems Using Clustering Techniques. In **Anais do IEEE Porto Power Tech Conference**, vol.3, Porto, Portugal, Sep, 2001.
- KELMAN, J., PEREIRA, M.V.F. Critérios de Avaliação para Modelos de Séries Hidrológicas. In **Anais do IV Seminário Nacional de Produção e Transmissão de Energia Elétrica**, 1977.
- KELMAN, J., OLIVEIRA, G.C., MACEIRA, M.E.P. (1992). Modelo de Séries Hidrológicas (MSH) - Manual de Metodologia", **Relatório Técnico CEPEL** no 242/92.
- MACEIRA, M.E.P. (1989). "Operação Ótima de Reservatórios com Previsão de Afluências", Tese de Mestrado, COPPE/UFRJ, março.
- MACEIRA, M.E.P., MERCIO, C.M.V.B. Stochastic Streamflow Model for Hydroelectric Systems. In **Anais do V International Conference PMAPS - Probabilistic Methods Applied to Power Systems**, Vancouver, Canada, 1997.
- MACEIRA, M.E.P.; COSTA, A .P.C., MARCATO, A .L.M. (1999). "Newave – Especificação Funcional", **Relatório Técnico CEPEL** nº DPP/PEL – 349/99.
- MACEIRA, M.E.P.; GHIRARDI, A .O., SIMBA, J.C.C (1999). "Modelo de Simulação a Usinas Individualizadas para Subsistemas Hidrotérmicos Interligados (Suishi-O) - Manual do Usuário", **Relatório Técnico CEPEL** nº DPP/PEL – 063/99.
- MACEIRA, M.E.P.; JARDIM, D.L.D.D. (2001). "Manual de Referência de Análise de Conglomerados", **Relatório Técnico CEPEL** nº DPP/PEN – 109/2001.

MACEIRA, M.E.P.; TERRY, L.A.; COSTA, F.S.; DAMÁZIO, J.M.; MELO, A.C.G. Chain of Optimization Models for Setting the Energy Dispatch and Spot Price in the Brazilian System. In **Anais do XIV Power Systems Computation Conference**, session 43, paper 1, Sevilla, Spain, Jun, 2002.

MOOD, A.M., GRAYBILL, F.A., BOES, D.C. **Introduction to the Theory of Statistics**, McGraw-Hill Book Company, Third Edition, 1974.

NOAKES, D.J., MCLEOD, A.I., HIPEL, K.W. (1983). "Forecasting Seasonal Hydrological Time Series", **Technical Report**, Dept. of Statistical and Actuarial Sciences, University of Waterloo, Waterloo, Ontario, Canada.

NOAKES, D.J., MCLEOD, A.I., HIPEL, K.W. Forecasting Monthly Riverflow Time Series. **International Journal of Forecasting**, 1, pp. 179-190, 1985.

OLIVEIRA, G.C., COSTA, J.P., DAMÁZIO, J.M., KELMAN, J. Multivariate Weekly Streamflow Forecasting Model, in **Anais do IFAC Conference on Systems Analysis Applied to Water and Related Land Resources**, Lisbon, Portugal, 1985.

SALAS, J.D., DELLEUR, J.W., YEVJEVICH, V., LANE, W.L. **Applied Modeling of Hydrologic Time Series**, Water Resources Publications, 1980.

YEVJEVICH, V. **Probability and Statistics in Hydrology**, Fort Collins, Water Resources Publications, 1972.

SYNTHETIC STREAMFLOW GENERATION FOR OPERATION PLANNING MODELS

Abstract

The adoption of probabilistic criteria in planning and operation of hydrothermal systems activities has created the necessity of probabilistic modeling of hydrological inflow to hydroelectric plants and subsystems. The only in available scenario, that is the historical series unfortunately is not long enough to estimate risks. Otherwise, the basic characteristics of the historical series can be used to build stochastic models that are able to produce synthetic inflow sequences that are different from the historical one but equally probable. This paper describes the formulation and the validation tests of the periodic autorregressive model, $Par(p)$, used in the Medium and Short Term Operation Planning models of the brazilian interconnected system..

Key-words: *Monthly Streamflow Model, Stochastic Modeling; Multivariate Periodic Autoregressive Models.*