

APLICAÇÃO DA AMOSTRAGEM POR HIPERCUBO LATINO E AMOSTRAGEM DESCRITIVA NA DEFINIÇÃO DA ÁRVORE DE CENÁRIOS DE AFLUÊNCIAS PARA O PLANEJAMENTO DA OPERAÇÃO ENERGÉTICA

Débora Dias Jardim Penna¹, Maria Elvira Piñeiro Maceira² & Jorge Machado Damázio³

RESUMO --- No modelo atualmente utilizado para o planejamento da operação de médio prazo do Sistema Interligado Nacional, a incerteza referente às afluições é considerada explicitamente no cálculo de valores da função de custo futuro bem como no percorrimento do espaço de estados através da utilização de cenários hidrológicos multivariados. O conjunto de todas as possíveis realizações do processo estocástico de afluições, ao longo de todo horizonte de planejamento, forma uma árvore de cenários. Esta árvore representa todo o universo probabilístico sobre o qual é efetuado o processo de otimização da operação energética. Como a árvore possui uma cardinalidade bastante elevada, torna-se impossível do ponto de vista computacional percorrer completamente a árvore. Portanto, apenas uma porção da árvore (sub-árvore) é percorrida. Atualmente a sub-árvore é definida utilizando amostragem aleatória simples. Neste trabalho é avaliado o uso da amostragem por hipercubo latino e a amostragem descritiva no modelo de geração de cenários hidrológicos multivariados com o intuito de tornar mais robusto os resultados obtidos pela política de operação em relação a variações no número de cenários das simulações forward e backward e em relação a variações da amostra de cenários hidrológicos utilizada.

ABSTRACT --- In the model currently used in the long-term operation planning of the Brazilian Interconnected System, the uncertainty concerning streamflow is considered explicitly in the calculation of the expected cost-to-go function, as well as in the covering of the state space, by the use of multivariate hydrological scenarios. The set of all possible realizations of the streamflow stochastic process throughout the planning horizon forms a scenario tree. This tree represents the entire probabilistic universe on which it is calculated the optimal operation strategies. As the scenario tree of the long-term operation planning problem has a high cardinality, it is impossible to visit the complete tree due to computational effort. Therefore, only a portion of the tree (sub-tree) is covered. Currently the sub-tree is selected using the Monte-Carlo method that uses classical simple random sampling. This paper evaluate change the simple random sampling by latin hypercube sampling or descriptive sampling in the multivariate streamflow scenario generation model in order to obtain more robust results from the operation policy with regard to variations in the number of scenarios of forward and backward simulations, and variations in the sample hydrological scenario.

Palavras-chave: Planejamento da operação energética, geração de árvore de cenários, técnicas de amostragem.

1) Pesquisadora do Centro de Pesquisas de Energia Elétrica, Av. Horácio Macedo, 354 – Cidade Universitária. E-mail: debora@cepel.br

2) Pesquisadora do Centro de Pesquisas de Energia Elétrica e professora da Universidade do Estado do Rio de Janeiro. E-mail: elvira@cepel.br

3) Pesquisador do Centro de Pesquisas de Energia Elétrica e professor da Universidade do Estado do Rio de Janeiro. E-mail: damazio@cepel.br

1. INTRODUÇÃO

O problema de planejamento da operação energética de um sistema de geração de energia elétrica hidrotérmico consiste em determinar metas de geração para as usinas hidrelétricas e termelétricas para cada estágio ao longo do horizonte de estudo, atendendo à demanda de energia elétrica, às restrições operativas das usinas e às restrições elétricas do sistema. A determinação da alocação ótima dos recursos hídricos e térmicos é norteada segundo um determinado critério, como por exemplo, o critério de minimização dos custos globais de produção da energia elétrica.

Devido às incertezas associadas ao problema, ao número de aproveitamentos existentes, a inter-relação entre as decisões tomadas num determinado instante e suas conseqüências futuras, o problema do planejamento da operação energética se torna um problema de difícil solução, principalmente para sistemas de grande porte como o sistema brasileiro.

Por causa do porte e complexidade deste problema é necessária a sua divisão em diferentes etapas. Em cada etapa são utilizados modelos com diferentes graus de detalhamento para a representação do sistema, abrangendo períodos de estudos com horizontes distintos (planejamento da operação de médio prazo, curto prazo e programação da operação diária).

O modelo de planejamento de médio prazo, NEWAVE (Maceira *et al*, 2008), considera um horizonte temporal de 5 a 10 anos com discretização mensal. Para reduzir a dimensionalidade do problema é utilizada uma representação simplificada do parque gerador, através de sistemas equivalentes de energia. A estocasticidade das afluições é representada por um número muito grande de cenários hidrológicos gerados sinteticamente por um modelo estocástico periódico autorregressivo de ordem p . O cálculo da política de operação é baseado na programação dinâmica dual estocástica. No NEWAVE, a política de operação, ou estratégia ótima de operação, é representada pela função de custo futuro e calculada por um processo iterativo para um conjunto de estados (energia armazenada no início do estágio e tendência hidrológica¹). Para cada estado, o corte da função de custo futuro corresponde a uma média calculada para um conjunto de aberturas².

A incerteza referente às afluições é considerada explicitamente no cálculo da função de custo futuro através da utilização de cenários hidrológicos multivariados. O conjunto de todas as possíveis realizações do processo estocástico de afluições, ao longo de todo horizonte de planejamento, forma uma árvore de cenários. Esta árvore representa todo o universo probabilístico sobre o qual será efetuado o processo de otimização da operação energética.

Como a árvore de cenários do problema de planejamento de médio prazo possui uma cardinalidade bastante elevada, igual ao número aberturas elevado ao número de estágios do horizonte de planejamento (normalmente igual 20^{120-1}), torna-se impossível do ponto de vista

¹ O conjunto dos últimos valores observados em uma série hidrológica

computacional percorrer completamente a árvore. Portanto, apenas uma porção da árvore (sub-árvore) é percorrida.

Atualmente a sub-árvore é definida utilizando amostragem aleatória simples (AAS), porém a escolha aleatória da sub-árvore está submetida às flutuações amostrais podendo-se perceber variações nos resultados ao se aumentar o tamanho da amostragem ou a se alterar as sementes dos geradores de números aleatórios utilizadas na amostragem.

O presente trabalho propõe que sejam aplicados métodos que permitam uma forma de amostragem mais eficiente para o modelo de geração de cenários hidrológicos multivariados, utilizado pelo modelo de planejamento da operação de médio prazo. Desta forma, busca-se aprimorar a construção da função de custo futuro com o intuito de tornar mais robusto os resultados obtidos pela política de operação, com relação a variações no número de cenários de simulação forward e backward, e com relação à amostra de cenários hidrológicos utilizada.

Neste trabalho são analisados o método de conjuntos estratificados, Hipercubo Latino (LHC), o método de conjuntos determinísticos, Amostragem Descritiva (AD) e o método de Monte-Carlo clássico. A amostragem LHC e AD basicamente consiste em subdividir o universo de amostras em um certo número de subconjuntos disjuntos e para cada um desses subconjuntos extrair um elemento representativo. Na amostragem LHC, o elemento é escolhido de forma aleatória e na AD é selecionado de forma determinística, como por exemplo o centróide do subconjunto.

Os resultados obtidos pela aplicação das técnicas propostas na representação do processo estocástico das afluições e no modelo de cálculo da estratégia ótima de operação serão apresentados considerando um caso real do Programa Mensal de Operação (PMO) elaborado pelo Operador Nacional do Sistema Elétrico (ONS).

2. MODELO DE GERAÇÃO DE SÉRIES SINTÉTICAS

Séries hidrológicas de intervalo de tempo menor que o ano, tais como séries mensais, têm como característica o comportamento periódico (sazonal) das suas propriedades probabilísticas, como por exemplo, a média, a variância, a assimetria e a estrutura de autocorrelação. 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 (Maceira *et al*, 2005) e são referenciados por modelos PAR(p), onde p é 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 mês.

O modelo PAR(p_1, p_2, \dots, p_{12}) pode ser descrito matematicamente por:

² Conjunto de afluições utilizada durante a recursão backward

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

onde, Z_t é a série sazonal, m é o índice do período sazonal, μ_m é a média sazonal, σ_m é o desvio-padrão sazonal, 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 e a_t é a série de ruídos independentes com média zero e variância $\sigma_a^{2(m)}$. Um problema desta formulação para a produção de séries sintéticas de ENA é a geração de ENA mensais negativas. Para obter-se um Z_t positivo é necessário que:

$$a_t \geq -\left(\frac{\mu_m}{\sigma_m}\right) - \phi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}}\right) - \dots - \phi_{p_m}^m \left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}}\right) \quad (2)$$

Muitos pesquisadores assumem que os resíduos a_t apresentam distribuição Normal e uma possível não-normalidade pode ser corrigida pela transformação Box-Cox. Como as séries sintéticas produzidas serão utilizadas em modelos que calculam as políticas ótimas de operação, 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 , cujo terceiro parâmetro Δ_t é conhecido como deslocamento. Conseqüentemente, a variável ξ_t tem distribuição Normal com média μ_{ξ_t} e desvio-padrão igual a $\sigma_{\xi_t}^2$:

$$\xi_t = \ln(a_t - \Delta_t) \quad (3)$$

Os parâmetros Δ_t , $\sigma_{\xi_t}^2$ e μ_{ξ_t} são estimados de tal forma a preservar os momentos dos resíduos $\sigma_a^{2(m)}$ e $\mu_a^{(m)}$, através das seguintes relações:

$$\mu_{\xi_t} = \frac{1}{2} \ln\left(\frac{\sigma_a^{2(m)}}{\theta^2 - \theta}\right) \quad \sigma_{\xi_t}^2 = \ln(\theta) \quad \Delta_t = \mu_a^{(m)} - \frac{\sigma_a^{(m)}}{\sqrt{(\theta - 1)}} \quad (4)$$

Fazendo Δ_t igual ao lado direito da expressão (2), podemos escrever:

$$\theta = 1 + \frac{\sigma_a^{2(m)}}{(\mu_a^{(m)} - \Delta_t)^2} \quad (5)$$

$\sigma_{\xi_t}^2$ pode ser diretamente obtido a partir das expressões (5) e (4).

Com o intuito de gerar ENA mensais multivariadas, assume-se que os resíduos espacialmente não correlacionados, ξ_t , podem ser transformados em resíduos espacialmente correlacionados, W_t , da seguinte forma: $W_t = D\xi_t$, onde D é uma matriz quadrada de dimensão igual ao número de subsistemas.

A matriz D é estimada utilizando $DD^T = \hat{U}$, onde \hat{U} é a correlação espacial entre as ENA anuais. A matriz D pode ser estimada adotando-se D como uma matriz triangular inferior, utilizando a decomposição de Cholesky, ou por decomposição espectral.

3. GERAÇÃO DOS CENÁRIOS PARA SIMULAÇÕES FORWARD E BACKWARD

A geração das séries sintéticas para a simulação forward é feita de forma seqüencial, isto é, a cada instante t , Z_t é obtido como uma função dos p_m valores passados e de um sorteio a_t . Sendo assim, para cada instante t é sorteado um conjunto diferente de ruídos. Na Figura 1 é apresentado o esquema de geração para a simulação forward, supondo três séries sintéticas.

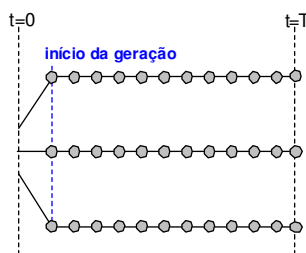


Figura 1 – Geração em Paralelo (Pente) para a simulação Forward

Durante a simulação forward apenas um valor de afluência é utilizado a cada período e série. Já na simulação backward, um número maior de valores de afluência é utilizado a cada período e série. Cada valor de afluência que é utilizado em um determinado período e série é denominado de ‘abertura’. A geração das afluências para as aberturas é feita de forma seqüencial, isto é, a cada instante t e série i , $Z_{t,i}$ é obtido como uma função dos p_m valores passados da série i e de um sorteio a_t . Os valores passados da série i são aqueles valores gerados para a simulação forward, Figura 1.

A cada instante t é sorteado um conjunto diferente de n_{leq} ruídos a_t , onde n_{leq} é o número de aberturas utilizadas no cálculo da política. O conjunto de ruídos sorteado em um instante t é utilizado para gerar as n_{leq} aberturas de todas as séries sintéticas.

Na Figura 2 é mostrada a seqüência de geração para as aberturas da simulação backward.

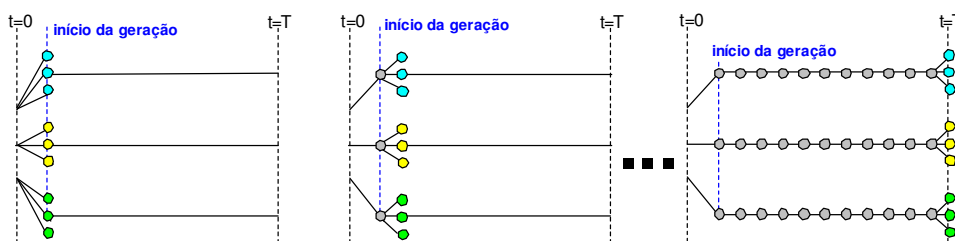


Figura 2: Seqüência de geração para a simulação Backward

Na Figura 3 é apresentado o esquema de geração para a simulação backward, supondo três séries sintéticas com três aberturas.

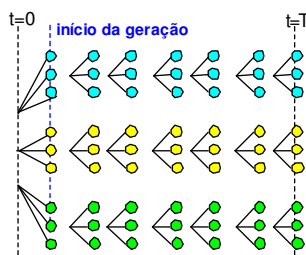


Figura 3: Geração em Paralelo (Pente) para a simulação Backward

4. TÉCNICAS DE AMOSTRAGEM

O método de Monte-Carlo clássico utiliza a amostragem aleatória simples. Em geral, a amostragem aleatória simples faz uso do método da transformada inversa para gerar valores aleatórios para uma determinada distribuição de probabilidades, a partir de valores gerados segundo uma distribuição uniforme [0,1]. As técnicas de amostragem abordadas neste item diferem da tradicional amostragem aleatória simples por terem um controle parcial do processo da amostragem, resultando desta forma em uma redução da variância dos estimadores, ou dito de outra forma, um aumento na precisão dos resultados das simulações.

A amostragem por hipercubo latino foi sugerida por McKay et al (1979), nela o domínio de cada variável aleatória (VA) X_k ($k = 1, \dots, M$) é dividido em N intervalos, ΔX_k ($i=1, 2, \dots, N$), de igual probabilidade $1/N$. O número de intervalos N na amostragem por hipercubo latino deve ser igual ao tamanho da amostra desejada, ou seja, igual ao número total de simulações. Para cada intervalo é amostrado apenas um valor (x_k^i), isto é, este valor será usado em uma e apenas uma simulação.

Os valores amostrados x_k^i , para um valor i qualquer, são obtidos pela resolução de (6):

$$F_k(x_k^i) = \frac{i-1+R_i}{N} \quad i = 1, 2, \dots, N \quad (6)$$

onde R_i representa uma distribuição aleatória uniforme no intervalo [0,1].

A amostragem é realizada utilizando a transformada inversa da função de distribuição de probabilidade em questão, como em (7):

$$x_k^i = F_k^{-1}\left(\frac{i-1+R_i}{N}\right) \quad (7)$$

Após obtidos os N valores para cada variável X_k , esses devem ser emparelhados de forma aleatória com os valores das demais variáveis. Dessa forma, são formados N vetores de dimensão M . A seleção aleatória do i -ésimo valor para cada variável é realizada mediante a permutação aleatória dos inteiros $1, 2, \dots, N$. A amostra dos N pontos da amostragem por hipercubo latino no espaço Euclidiano de M -dimensões contém apenas um ponto em cada intervalo de cada uma das M variáveis.

Deve ser observado ainda que, embora as M variáveis sejam amostradas de forma independente e emparelhadas de forma aleatória, o coeficiente de correlação amostral dos N pares de variáveis, tanto na amostragem aleatória simples quanto na amostragem por hipercubo latino, será em geral diferente de zero devido às flutuações amostrais (Wiss e Jorgensen, 1998).

Suponha que seja necessário gerar uma amostra com um determinado coeficiente de correlação de população. Para obter uma amostra com tal coeficiente de correlação de população, Iman & Conover (1982) propuseram um método para restringir a maneira como as variáveis são emparelhadas. Para tanto utiliza-se uma transformação considerando a matriz de rank.

McKay et al (1979) e Stein (1987) em suas comparações observaram que a amostragem por hipercubo latino é mais vantajosa em relação a outros métodos porque o número de simulações pode ser reduzido consideravelmente para alcançar o mesmo nível de precisão.

A amostragem descritiva está descrita em Saliby (1997). Seu procedimento se assemelha bastante ao procedimento da amostragem por hipercubo latino, a diferença está na escolha dos valores amostrados. O domínio de cada variável aleatória X_k também é dividido em N intervalos de igual probabilidade ($p=1/N$). No caso da amostragem descritiva o ponto x_k^i é escolhido como sendo o centróide do intervalo i . A fórmula usada para a geração do conjunto de valores descritivos, a serem depois permutados aleatoriamente é dada em (8).

$$x_k^i = F_k^{-1}\left(\frac{i-1+0.5}{N}\right) = F_k^{-1}\left(\frac{i-0.5}{N}\right) \quad (8)$$

Em resumo, a amostragem descritiva se baseia numa seleção totalmente determinística e intencional dos valores amostrais das variáveis aleatórias consideradas no problema. Uma vez selecionados, esses valores são permutados aleatoriamente.

5. AVALIAÇÃO DOS CENÁRIOS HIDROLÓGICOS GERADOS

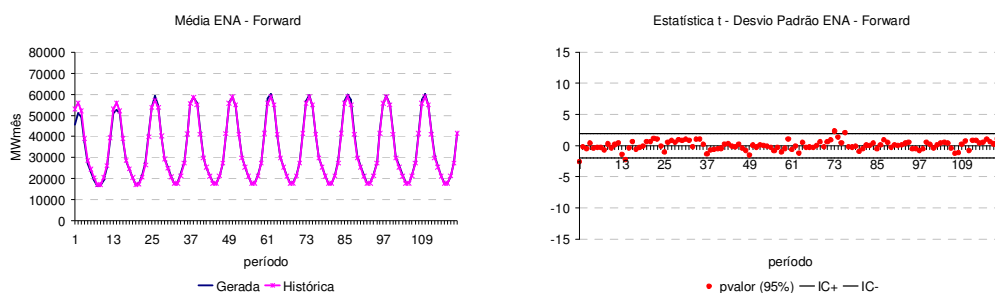
Para verificar a acurácia da geração da árvore de cenários é utilizado o Programa Mensal de Operação de janeiro de 2008. A configuração hidroelétrica utilizada neste estudo é composta por 145 usinas hidroelétricas. Estas usinas são agregadas em quatro subsistemas equivalentes de energia (Sudeste/Centro-Oeste, Sul, Nordeste e Norte). O horizonte de planejamento é de cinco anos divididos em estágios mensais e mais um período de pós-estudo com cinco anos, totalizando 120 períodos. São considerados 200 cenários para simulação forward e 20 aberturas para simulação backward.

5.1 Análise dos Cenários Gerados para Passo Forward

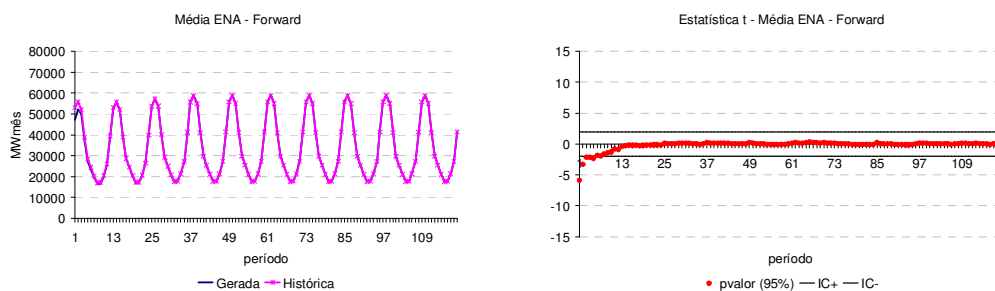
Para a avaliação dos cenários hidrológicos obtidos para o passo forward procede-se uma análise detalhada da média e do desvio-padrão dos cenários hidrológicos. O nível de significância

do teste é de 5% e como parâmetros de população são considerados os valores históricos (testes não condicionados). Na Figura 4 são apresentadas a evolução temporal da média histórica e dos cenários gerados e a estatística t desta estimativa, para o subsistema Sudeste. Os cenários gerados para os primeiros períodos são fortemente condicionados ao passado recente (tendência hidrológica recente).

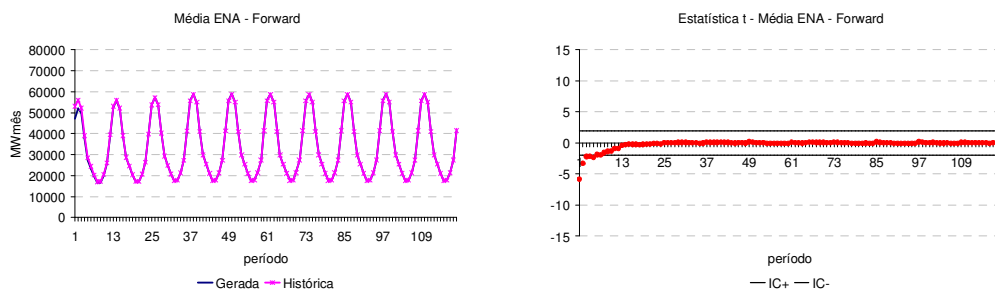
Comparando as opções AAS, LHC e AD, verifica-se que as técnicas de amostragem estratificadas reduzem de sobremaneira a variabilidade da estimativa da média.



(a) Opção AAS



(b) Opção LHC

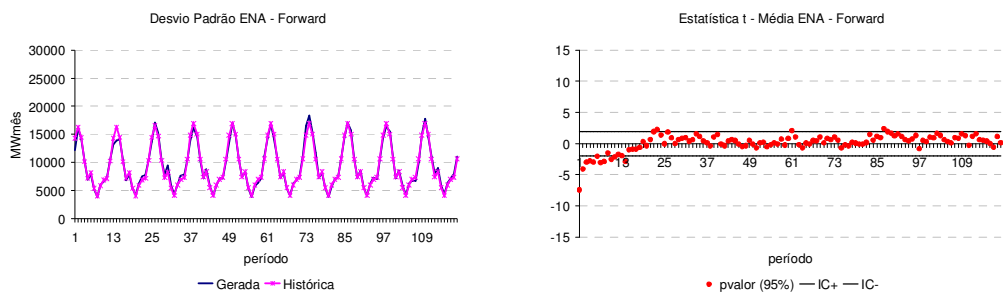


(c) Opção AD

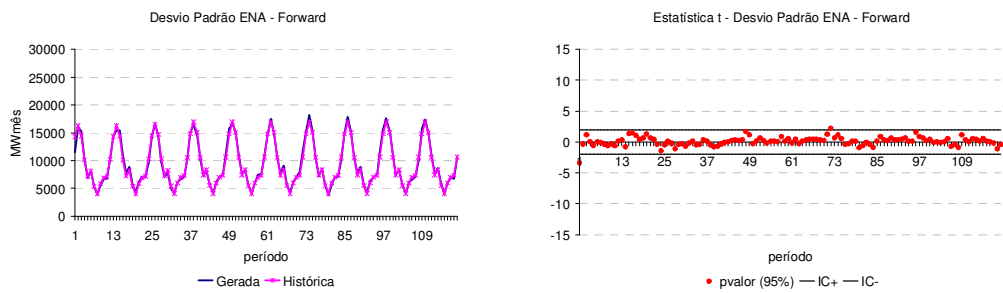
Figura 4: Média Cenários Forward – Sudeste

Na Figura 5 é avaliada a estimativa do desvio-padrão para o subsistema Sudeste, com relação ao desvio-padrão histórico. A estimativa do desvio-padrão nos primeiros meses do estudo está muito condicionada ao passado recente, assim como ocorre com a estimativa da média.

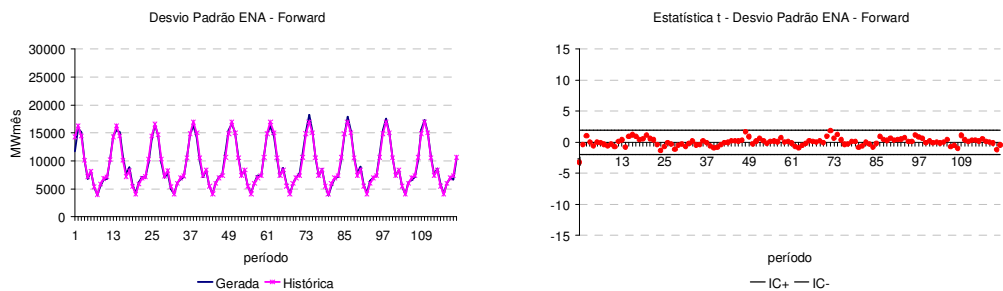
A preservação do desvio histórico é alcançada pela opção AAS, com exceção dos primeiros meses. As opções LHC e AD apresentam uma variabilidade menor para a estimativa do desvio-padrão do que a observada na opção AAS.



(a) Opção AAS



(b) Opção LHC



(c) Opção AD

Figura 5: Desvio-Padrão Cenários Forward – Sudeste

Na próxima análise é aplicado o teste de Kolmogorov-Smirnov para verificar a aderência entre as distribuições histórica e gerada a cada período do horizonte de estudo. O nível de significância adotado é de 5%. Na Figura 6 são apresentados os resultados para os cenários hidrológicos do subsistema Sudeste. As melhores representações da distribuição univariada são obtidas com as opções LHC e AD. A opção AAS apresenta um comportamento satisfatório.

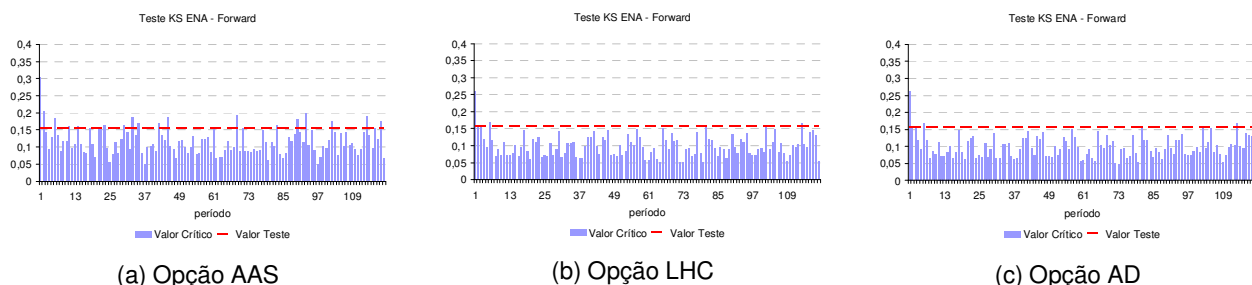


Figura 6: Teste de Aderência Cenários Forward – Sudeste

A preservação da correlação cruzada é um objetivo a ser perseguido pelo processo de geração de cenários sintéticos multivariados. Os valores calculados são comparados com o valor histórico.

Na Figura 7 é apresentada a evolução temporal da correlação cruzada entre os subsistemas

Sudeste e Sul. Observe que há uma variação em torno do valor histórico para a opção AAS. Nas opções LHC e AD pode-se observar uma tendência de redução no valor da correlação cruzada estimada.

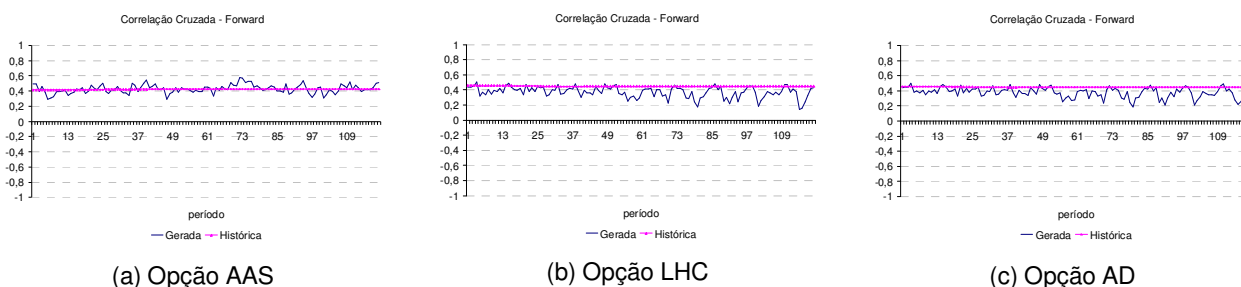


Figura 7: Correlação Cruzada Cenários Forward – Sudeste x Sul

Na Tabela 1 é mostrada a correlação cruzada média ao longo de todo horizonte de estudo para opções analisadas neste item. Conforme o observado nas figuras anteriores, em geral as opções LHC e AD têm dificuldade em preservar o valor da correlação cruzada histórica.

Tabela 1: Correlação Cruzada Média Cenários Forward

OPÇÃO	Sudeste-Sul	Sudeste-Nordeste	Sudeste-Norte	Sul-Nordeste	Sul-Norte	Nordeste-Norte
Histórica	0,4502	0,5709	0,4265	-0,0792	-0,1313	0,6372
AAS	0,4208	0,5594	0,4269	-0,0543	-0,0989	0,6486
LHC	0,3710	0,5342	0,4173	-0,0396	-0,0563	0,5972
AD	0,3779	0,5367	0,4188	-0,0371	-0,0569	0,6004

5.2 Análise dos Cenários Gerados para Passo Backward

São utilizados testes não condicionados para a avaliação da capacidade de preservação das estatísticas média, desvio-padrão e correlação cruzada, e testes para verificação da aderência entre as distribuições gerada e histórica. Para realização dos testes não condicionados, a cada período são considerados todos os cenários backward gerados, independentemente do estado de afluência passada, conforme ilustrado na Figura 8a. Desta forma, neste estudo é considerada uma amostra com 4000 vetores multivariados de ENA (200x20).

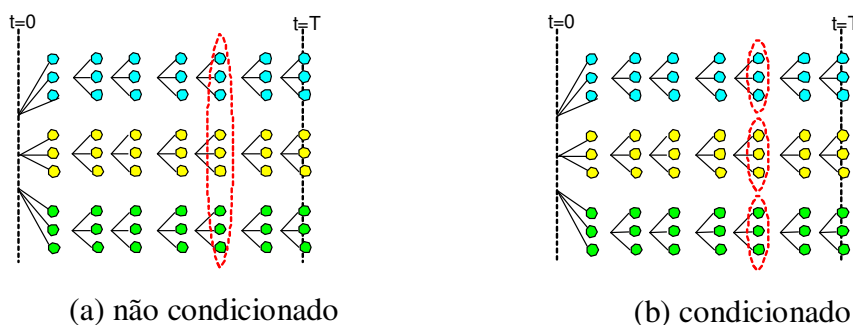


Figura 8 Amostra utilizada nos testes

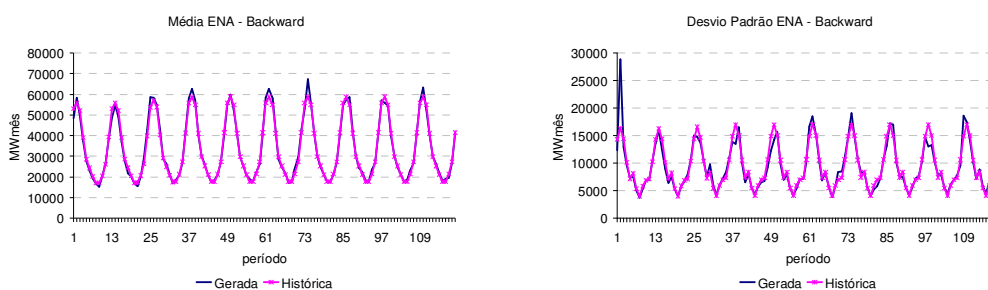
Para a avaliação da qualidade dos cenários gerados para cada conjunto de cenários aberturas são realizados testes condicionados. Neste caso, o teste é aplicado a cada estado de afluência

passada, conforme ilustrado na Figura 8b. Neste caso, são feitos a cada período 200 testes considerando uma amostra de 20 vetores multivariados de ENA.

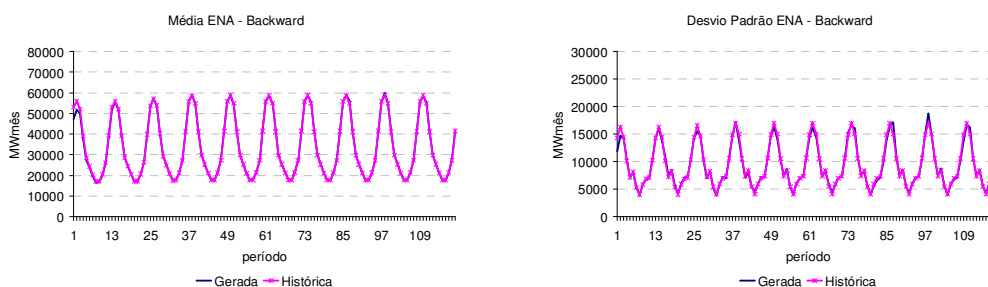
Na primeira coluna da Figura 9 são apresentadas as evoluções temporais da média e do desvio padrão para a série histórica e para os cenários gerados, para o subsistema Sudeste.

Vale ressaltar que os valores não condicionados de média e desvio-padrão dos cenários backward são influenciados tanto pela qualidade da geração dos cenários backward, propriamente ditos, quanto pela geração dos cenários forward.

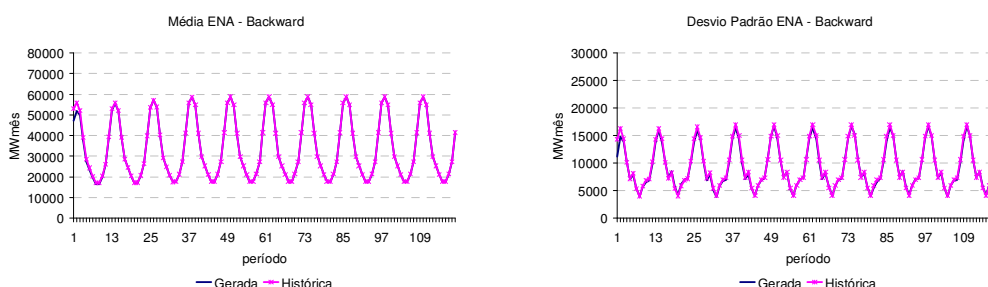
Comparando as opções AAS, LHC e AD, verifica-se que as técnicas de amostragem estratificadas reduzem bastante a variabilidade dos valores médios. A opção AAS apresenta um pouco de dificuldade em preservar o desvio-padrão histórico, porém preservação do desvio histórico é alcançada quase com perfeição pelas opções LHC e AD, com exceção dos primeiros meses.



(a) Opção AAS



(b) Opção LHC



(c) Opção AD

Figura 9: Média Cenários Backward – Sudeste

O próximo ponto a ser analisado é a distribuição univariada da amostra de cenários aplicando-se o teste Kolmogorov-Smirnov. Na Figura 10 são apresentados os resultados do teste KS para os cenários hidrológicos do subsistema Sudeste. As melhores representações da distribuição univariada do subsistema Sudeste são obtidas com as opções LHC e AD. A opção AAS apresenta um comportamento satisfatório.

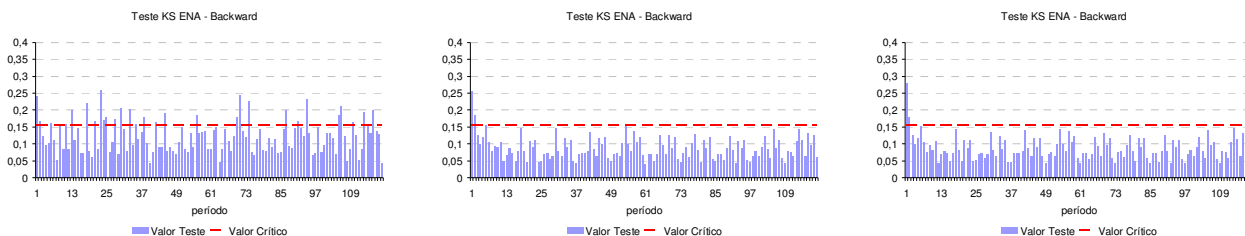
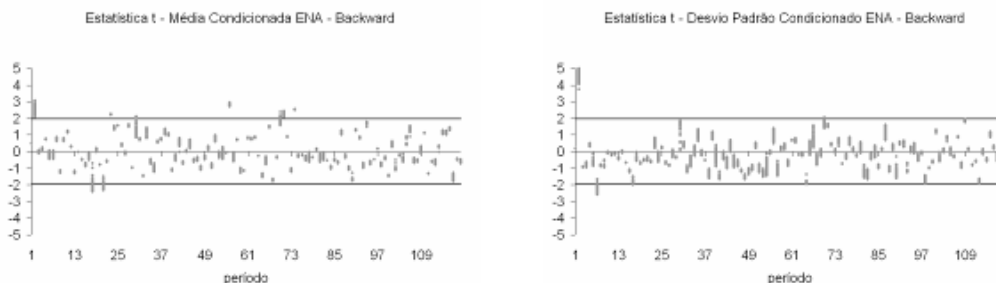


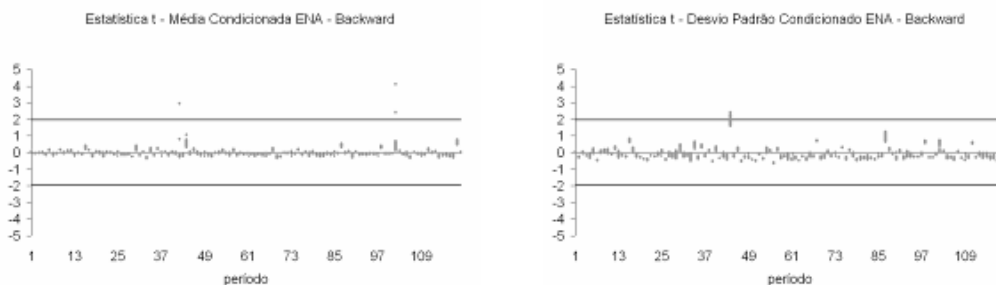
Figura 10: Teste de Aderência Cenários Backward – Sudeste

A seguir são realizados os testes condicionados para média e desvio padrão dos cenários gerados para o passo backward.

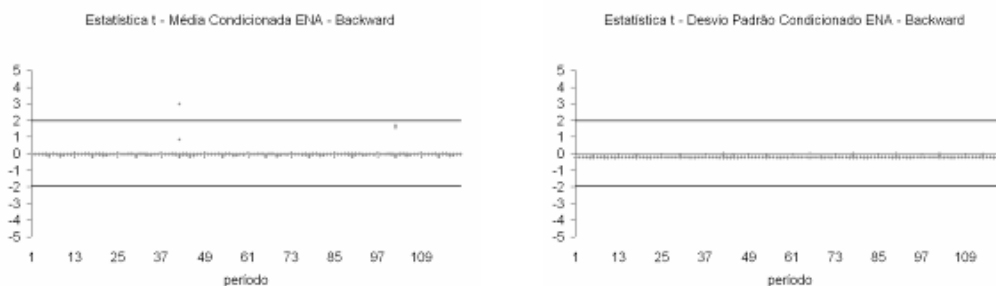
Como pode ser constatado na Figura 11, a opção AAS apresenta uma maior variabilidade com relação aos valores de média e desvio-padrão. Esta variabilidade tende a diminuir quando o tamanho da amostra aumenta, mas se o número de cenários backward for grande pode-se tornar inviável do ponto de vista computacional a resolução do problema de planejamento da operação energética.



(a) Opção AAS



(b) Opção LHC

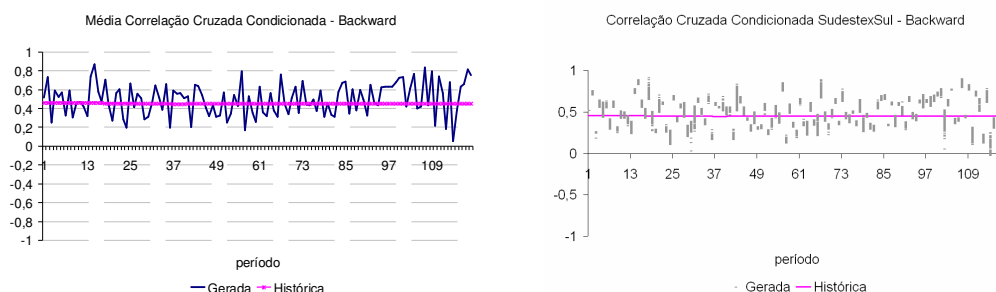


(c) Opção AD

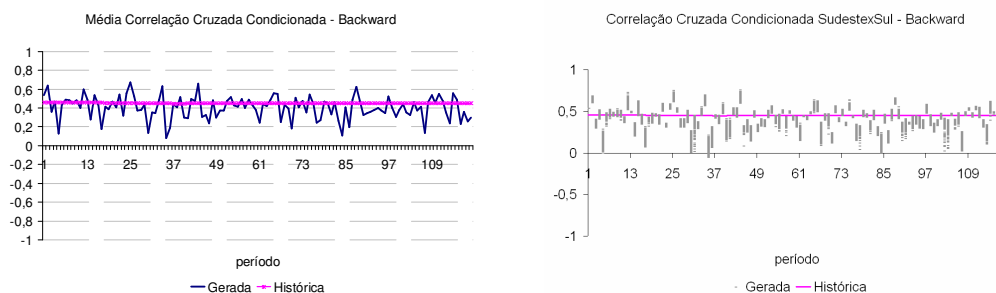
Figura 11: Estatística t Média e Desvio-Padrão – Cenários Backward - Sudeste

A seguir é avaliada a preservação da correlação cruzada para cada conjunto de cenários backward. Na primeira coluna da Figura 12 é apresentada a evolução temporal da média da correlação cruzada entre os subsistemas Sudeste e Sul. Esta média é calculada a partir dos 200 valores obtidos de correlação cruzada para cada período, que são ilustrados na segunda coluna desta figura.

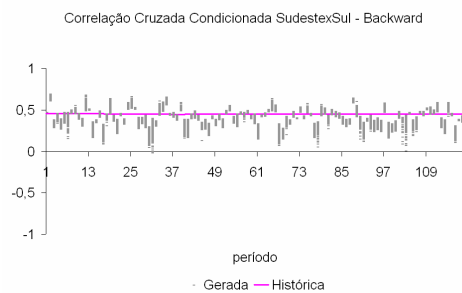
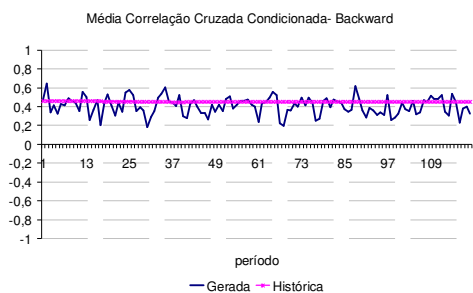
Observe que há uma grande variação em torno do valor histórico para a opção AAS. A variabilidade bastante acentuada verificada para estas opções está relacionada ao tamanho da amostra utilizada. Como pode ser observado na Figura 12 as opções LHC e AD também apresentam uma variabilidade com relação ao valor histórico. Além disso, apresentam uma tendência de redução no valor da correlação cruzada estimada.



(a) Opção AAS



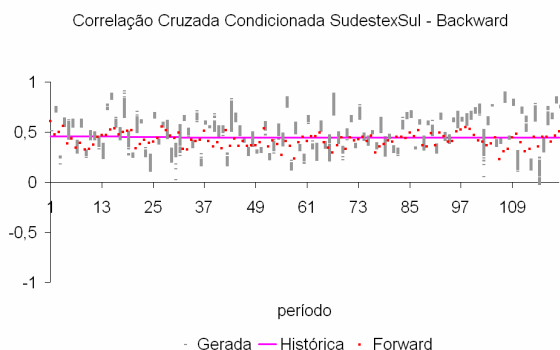
(b) Opção LHC



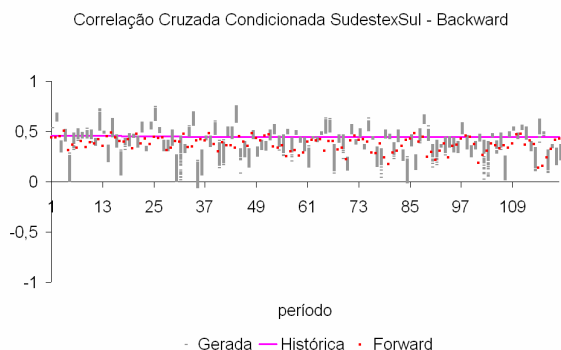
(c) Opção AD

Figura 12: Correlação Cruzada Cenários Backward – Sudeste x Sul

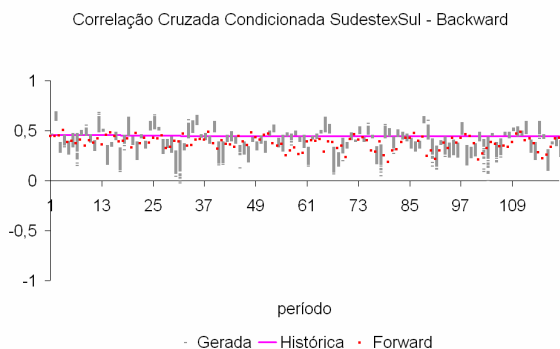
É desejável que os cenários dos passos forward e backward representem a correlação cruzada de forma compatível e que ambas se aproximem do valor observado no histórico. Na Figura 13 são apresentadas as correlações cruzadas de cada conjunto cenário backward e a correlação cruzada dos cenários forward entre os subsistemas Sudeste e Sul. Pode-se observar uma grande variabilidade na correlação da backward, porém uma menor variação na correlação do passo forward, com relação ao valor histórico.



(a) AAS



(b) LHC



(c) AD

Figura 13: CC Backward x Forward – Sudeste x Sul

Na Tabela 2 é mostrada a correlação cruzada média não condicionada, ao longo de todo horizonte de estudo para opções. Da mesma forma como observado anteriormente, as opções LHC e AD têm dificuldade em preservar o valor da correlação cruzada histórica.

Tabela 2: Correlação Cruzada Média Cenários Backward – Não Condicionado

OPÇÃO	Sudeste-Sul	Sudeste-Nordeste	Sudeste-Norte	Sul-Nordeste	Sul-Norte	Nordeste-Norte
Histórico	0,4502	0,5709	0,4265	-0,0792	-0,1313	0,6372
AAS	0,4306	0,553	0,4151	-0,0586	-0,107	0,6416
LHC	0,362	0,5216	0,4009	-0,0525	-0,0753	0,5794
AD	0,3684	0,5249	0,4074	-0,0534	-0,0765	0,592

Quando os cenários backward e forward não são representados da mesma forma, por exemplo, com relação à correlação cruzada, podem ocorrer distorções durante a solução do problema de planejamento da operação.

Na Figura 14 é ilustrada uma comparação entre as correlações cruzada dos cenários forward e a média das correlações cruzadas condicionadas dos cenários backward. A linha pontilhada indica uma reta com ângulo igual 45° e passando pela origem, isto é, quanto mais próximos os pontos estiverem desta linha maior a proximidade da representação da correlação cruzada nos cenários backward e forward. Note que em geral todas as opções, exceto as opções LHC e AD, representam da mesma forma a correlação cruzada nos cenários da backward e forward.

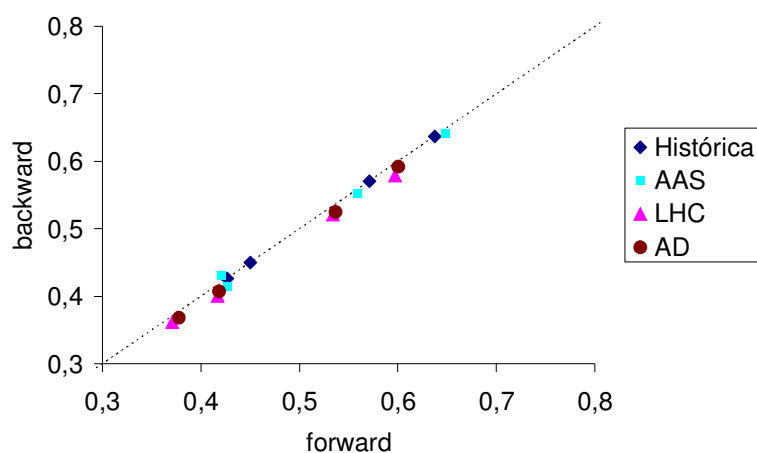


Figura 14: Comparação entre CC Forward e Backward

6. AVALIAÇÃO DOS RESULTADOS DO PROBLEMA DE PLANEJAMENTO DA OPERAÇÃO ENERGÉTICA

Para o cálculo da política ótima de operação são utilizados 200 cenários para o passo forward e para cada período são consideradas 20 aberturas (cenários backward). Após o cálculo da política mensal, é realizada uma simulação final utilizando-se 2000 cenários sintéticos de energia natural afluentes para obtenção dos índices de desempenho do sistema. O conjunto de cenários utilizado na simulação final é gerado com amostragem aleatória simples e tem um horizonte de cinco anos. Este conjunto é utilizado na simulação final de todas as opções analisadas. É considerado neste estudo o PMO FEV/07.

A cada iteração do algoritmo de solução da PDDE são calculados os limites superior (ZSUP) e inferior (ZINF) da solução ótima. O limite inferior é obtido na solução do problema de primeiro período, e o limite superior é estimado a partir dos resultados de todos os períodos e cenários visitados durante a simulação forward. A incerteza em torno da estimativa de ZUP é utilizada como critério de convergência estatístico para o algoritmo.

São realizadas sensibilidades com relação a variação de alguns parâmetros associados à árvore de cenários, como variação no número de cenários considerados nos passos forward e backward e variação da amostra de ruídos utilizada para geração dos cenários de aflúncias (variação da semente inicial do processo de geração dos ruídos aleatórios).

Primeiramente são avaliados os resultados do modelo NEWAVE quando são consideradas diferentes amostras de ruídos (variação de semente). Na Figura 15 são apresentados os resultado da última iteração. Pode-se observar que o valor de ZSUP para os casos que utilizam LHC e AD é maior do que o observado com o método AAS, assim como sua a variância, representada pelo intervalo de confiança. Com relação ao valor de ZINF, este é mais estável e tem a mesma magnitude observada nos caso AAS. Vale observar que todos os casos LHC e AD atingem o número máximo de iterações sem que o critério de convergência estatístico tenha sido alcançado. O mesmo ocorre com alguns casos AAS (casos marcados com *).

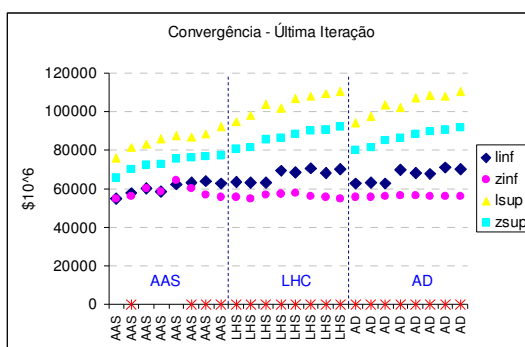
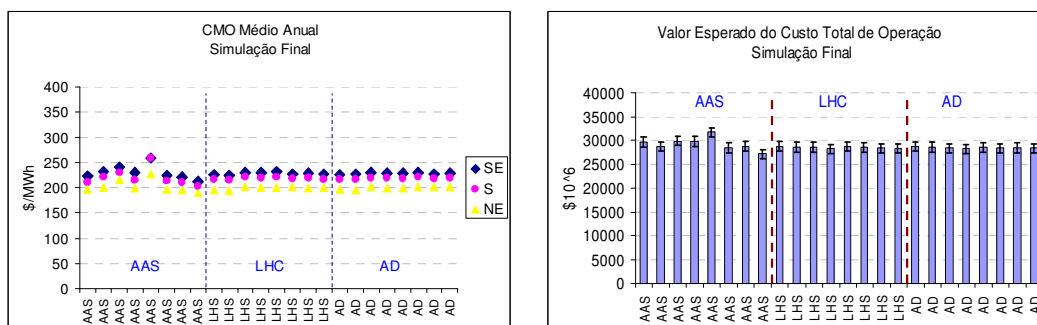


Figura 15: Convergência – Última Iteração – Variação Amostra

Os resultados de CMO, COPER, risco e EENS, apresentados na Figura 16 também apresentam uma maior estabilidade com relação aos mesmos resultados observados com o método AAS.



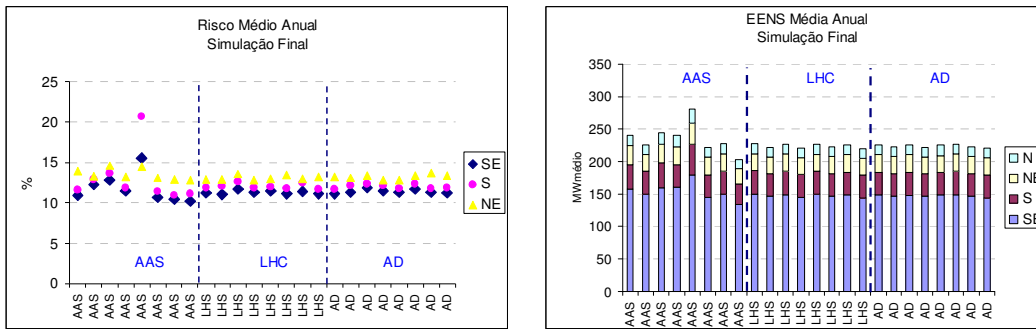


Figura 16: Resultados- Variação Amostra

O fato de ZINF e ZSUP estar convergindo para valores bem diferentes, pode está sendo causado pela dificuldade que os métodos LHC e AD têm de reproduzir de forma compatível a correlação cruzada das amostras forward e backward. Quando os cenários backward e forward não estão compatíveis estatisticamente há sinalizações distintas durante a etapa de operação do sistema (etapa relacionada com ZSUP) e a etapa de construção da FCF (etapa associada ao ZINF). Por exemplo, a FCF pode estar sendo construída enxergando cenários mais favoráveis e, portanto sinalizando um custo futuro baixo e indicando o uso da água armazenada nos reservatórios. Porém, quando o sistema é realmente operado, os cenários visitados podem ser mais críticos levando a corte de carga futuros, desta forma elevando o custo total de operação. Assim, os valores de ZINF, calculado através da FCF, e o valor de ZSUP, calculado durante a etapa de operação, podem ser bem distintos.

Com relação à variação do número de cenários backward, pode-se observar na Figura 17 que não houve variações nos resultados de ZINF e ZSUP obtidos na última iteração. Porém, pode-se verificar novamente que o valor de ZSUP para os métodos de amostragem LHC e AD estabilizam em um patamar bem superior ao valor observado com AAS. Na Figura 18 são apresentados valores médios anuais obtidos para CMO, risco e EENS, além do custo total de operação da simulação final. Esses resultados são se alteram nos casos LHC e AD e nos casos AAS apresentam uma pequena variação.

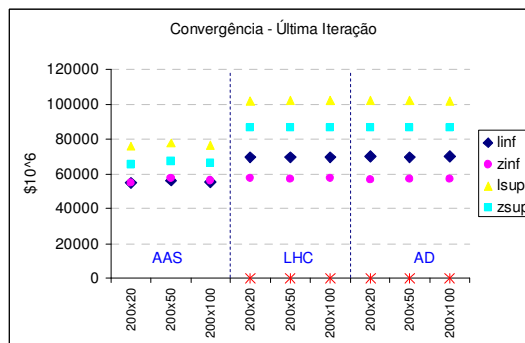


Figura 17: Convergência – Última Iteração – Variação Backward

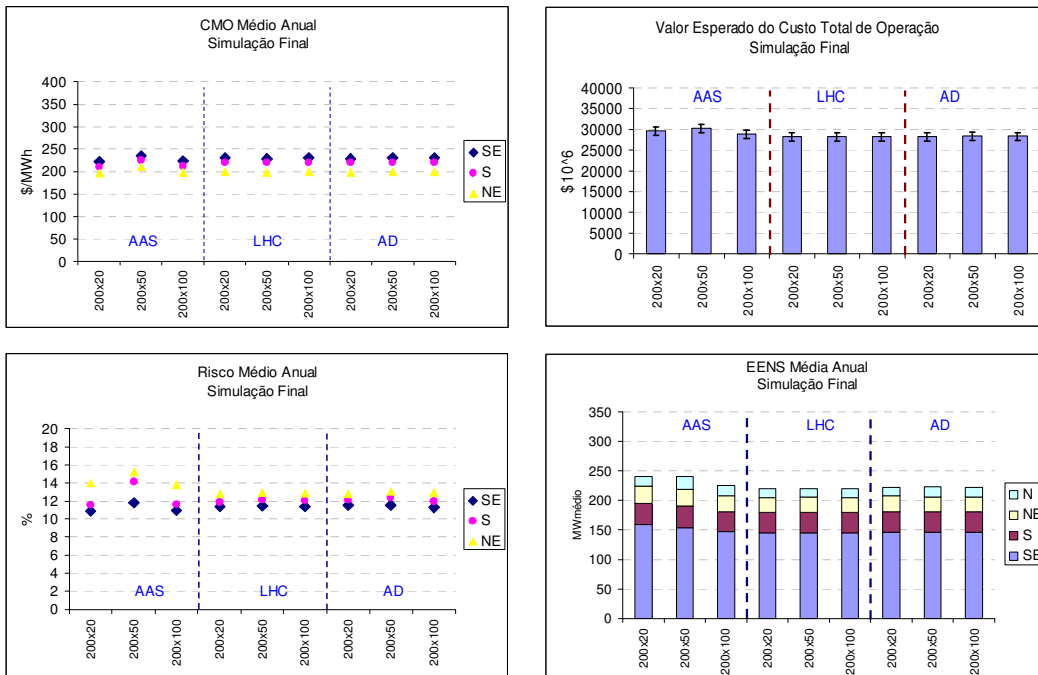


Figura 18: Resultados- Variação Backward

Na Figuras 19 e 20 são mostrados os resultados de uma sensibilidade com relação ao número de cenários forward. Da mesma forma como observado anteriormente, os resultados não apresentam variação com a mudança do número de cenários forward, porém há um acréscimo no valor de ZUP para os casos LHC e AD.

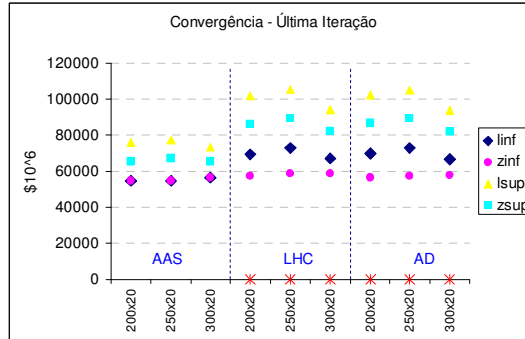
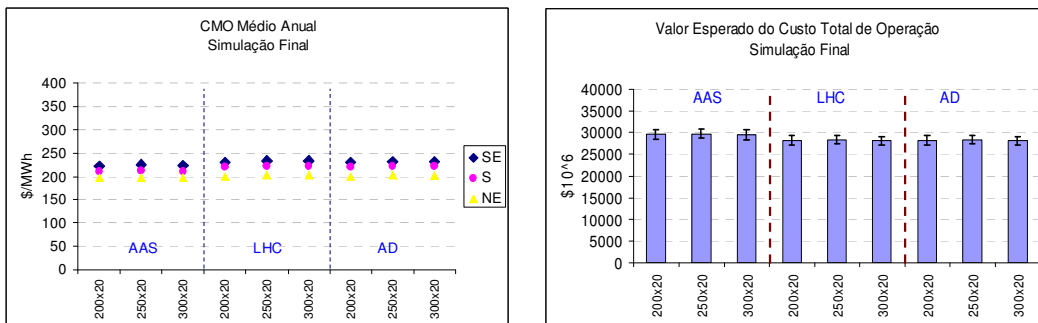


Figura 19: Convergência – Última Iteração – Variação Forward



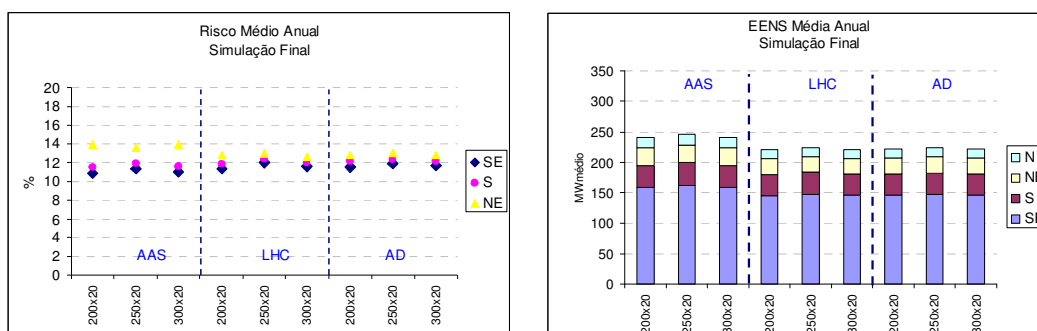


Figura 20: Resultados- Variação Forward

Os casos que utilizam os métodos de amostragem LHC e AD apresentaram um valor de ZSUP bem maior do que observado com o método AAS. Este fato pode estar associado ao fato destes métodos não reproduzirem a correlação cruzada da mesma forma nos cenários forward e backward.

7. Conclusões

Neste trabalho foi avaliado o uso de outros métodos de amostragem com o objetivo de reduzir a variabilidade causada naturalmente pela AAS. Desta maneira, pretende-se melhorar a qualidade da função de custo futuro que é construída pelo NEWAVE e repassada aos demais modelos da cadeia de planejamento energético e produzir resultados mais robustos de forma que pequenas variações na árvore de afluências não ocasionem grandes variações na solução do problema.

Os métodos de amostragem avaliados foram aplicados no modelo de geração de cenários sintéticos de energia natural afluyente. A geração dos cenários sintéticos é feita empregando-se modelos autorregressivos periódicos para modelar o processo estocástico das afluências para cada subsistema equivalente de energia. Os cenários sintéticos são gerados de forma multivariada e tomando-se o cuidado para que não sejam gerados valores negativos para as afluências, para tanto, é ajustada uma distribuição lognormal três parâmetros à parcela aleatória do modelo PAR(p).

A opção AAS, que utiliza amostragem aleatória simples para geração das amostras ruídos para os passos forward e backward de forma independente, apresentou variabilidade nos resultados quando a amostra de ruídos é alterada. Neste trabalho foi investigada a substituição do método de amostragem aleatória simples por métodos de amostragem estratificada, tais como o LHC e AD, com o intuito de reduzir a variação nos resultados causada por alterações nas amostras de ruídos. Como era de se esperar, a variabilidade com relação aos cenários hidrológicos são bem pequenas. Porém, esses métodos de amostragem apresentaram grande dificuldade em reproduzir uma característica muito importante da série histórica: a correlação cruzada. Há uma tendência de geração de valores mais baixos de correlação cruzada para os cenários da backward em comparação com os cenários do passo forward, causando impactos na construção da FCF e tornando o valor ZSUP mais elevado.

Um ponto que deve ser investigado futuramente é a redução da degradação observada na representação da correlação cruzada dos cenários gerados pelos métodos LHC e AD. Uma possível solução é calcular uma heurística capaz de melhorar a reprodução destas correlações, assim como foi proposto em Damázio et al (2009).

BIBLIOGRAFIA

DAMÁZIO, J.M.; MACEIRA, M.E.P; COSTA, F.S.; PENNA, D.D.J.; DUARTE, V.S. (2009) “*Representação da Correlação Cruzada na Geração de Cenários de Afluências para o Planejamento da Operação Energética de Médio Prazo*”. XI Simpósio de Especialistas em Planejamento da Operação e Expansão Elétrica - SEPOPE, Belém, Brasil.

IMAN, R.L., CONOVER, W.J., (1982), “*A Distribution-free Approach to Inducing Rank Correlation among Input Variables*”, Communications in Statistics B 11(3): 311-334.

MACEIRA, M. E. P.; DUARTE, V. S.; PENNA, D.D.J.; MORAES, L.A. M.; MELO, A.C.G., “*Ten Years of Application of Stochastic Dual Dynamic Programming in Official and Agent Studies in Brazil Description of the NEWAVE*”. In: Power Systems Computation Conference, (2008), Glasgow.

MACEIRA, M. E. P.; PENNA, D.D.J.; DAMÁZIO, J.M. (2005). “*Geração de Cenários Sintéticos de Energia e Vazão para o Planejamento da Operação Energética*” In: XVI Simpósio Brasileiro de Recursos Hídricos, João Pessoa, Paraíba, Brasil, pp. 3-12.

McKAY, M.D., BECKMAN, R.J., CONOVER, W.J., 1979, “*A Comparision of Three Methods for Selecting Values of Input Variables in the Analysis of Output form a Computer Code*”, Technometrics 21(2): 239-245.

SALIBY,E., (1997), “*Descriptive Sampling: An Improvement Over Latin Hypercube Sampling*”, In: Proceedings of the 1997 Winter Simulation Conference, pp. 230-233.

STEIN, M., (1987), “*Large Sample Properties of Simulations Using Latin Hypercube Sampling*”, Technometrics 29(2): 143-151.

WISS, G.D., JORGENSEN, K.H., (1998). “*A User’s Guide to LHS: Sandia’s Latin Hypercube Sampling Software*”, Sandia National Laboratories.