

Centro de Pesquisas de Energia Elétrica - CEPEL



Relatório Técnico

Nº/Ano: 3383/2019 **Nº de Páginas:**24 **Nº de Anexos:** 0

Título:

Avaliação do uso do centroide no processo de agregação da Amostragem Seletiva na geração de cenários de fluências para o planejamento da operação de curto prazo

Departamento:

Departamento de Otimização Energética e Meio Ambiente

Área de Responsabilidade:

B200

Conta de Apropriação:

1798 e 1129

Cliente:

Comissão Permanente para Análise de Metodologias e Programas Computacionais do Setor Elétrico - CPAMP

Resumo:

Neste relatório são avaliadas alternativas para escolha do representante no processo de agregação da Amostragem Seletiva com o objetivo de redução da variabilidade amostral nos estudos de planejamento da operação de curto prazo, atividade em estudo pela Comissão Permanente para Análise de Metodologias e Programas Computacionais do Setor Elétrico (CPAMP), para o ciclo 2018/2019.

Autores:

Felipe Treistman - PUC-Rio
Lilian Chaves Brandão dos Santos - CEPEL
Débora Dias Jardim Penna - CEPEL
Maria Elvira Piñeiro Maceira - CEPEL
Hugo Santarém de Araújo - CEPEL

Palavras-Chave:

Geração de cenários. Amostragem Seletiva. Planejamento da operação de curto prazo

Classificação: CONTROLADO**Gerentes dos Projetos**

Nome: Débora Dias Jardim Penna

Tel.: 21-2598-6450

E-mail: debora@cepel.br

Nome: Lilian Chaves Brandão dos Santos

Tel.: 21-2598-6059

Email: liliancbs@cepel.br

Chefe do Departamento de Otimização Energética e Meio Ambiente

Nome: André Luiz Diniz Souto Lima

Tel.: 21-2598-6046

E-mail: diniz@cepel.br

Aprovação

Raul Balbi Sollero

Diretor de Pesquisa, Desenvolvimento e Inovação

28/10/2019

Centro de Pesquisas de Energia Elétrica - CEPEL www.cepel.br

Sede: Av. Horácio Macedo, 354 - Cidade Universitária - CEP 21941-911 - Rio de Janeiro - RJ - Brasil - Tel.: 21 2598-6000 - Fax: 21 2260-1340

Unidade Adrianópolis: Av. Olinda, 5800 - Adrianópolis - CEP 26053-121 - Nova Iguaçu - RJ - Brasil - Tel.: 21 2666-6200 - Fax: 21 2667-3518

Endereço Postal: CEPEL Caixa Postal 68007 - CEP 21944-970 - Rio de Janeiro - RJ - Brasil / Endereço Eletrônico: cepel@cepel.br

CEPEL

Centro de Pesquisas de Energia Elétrica

PROJETO GEVAZP e DECOMP

Relatório Técnico

**Avaliação do uso do centroide no processo de agregação da
Amostragem Seletiva na geração de cenários de afluências para o
planejamento da operação de curto prazo**

Junho/2019

ÍNDICE

1	Introdução	4
2	Geração de Cenários.....	6
2.1	Amostragem Seletiva	7
2.1.1	Fator de compensação do desvio-padrão.....	8
3	Estudos de caso e métricas de avaliação	11
3.1	Avaliação da qualidade da geração de cenários de afluências.....	11
3.2	Métricas para a avaliação dos resultados do planejamento da operação energética de curto prazo.....	12
4	Resultados.....	13
4.1	Qualidade dos cenários de vazões afluentes	13
4.2	Avaliação dos resultados do planejamento da operação energética de curto prazo	17
5	Conclusões	21
6	Referências.....	22

1 Introdução

A geração de cenários sintéticos de afluências possui um relevante papel para a definição do planejamento da operação energética de curto, médio e longo prazos, representando a principal fonte de incerteza desse planejamento. Durante estudos da Comissão Permanente para Análise de Metodologias e Programas Computacionais do Setor Elétrico (CPAMP) do ciclo 2018/2019, identificou-se um padrão anômalo na evolução temporal do valor esperado do custo marginal da operação (CMO) obtido em uma simulação da operação do sistema com 2.000 séries sintéticas de afluências aos reservatórios equivalentes de energia com a função de custo futuro calculada pelo modelo NEWAVE [1], [2]. Esse comportamento foi avaliado detalhadamente pelo CEPEL no Relatório Técnico intitulado “Análise da inflexão do custo marginal de operação no modelo NEWAVE entre os quarto e quinto estágios temporais ao se adotar a representação de 12 REEs para Sistema Interligado Nacional” de setembro de 2018 [3]. Chegou-se à conclusão que a redução abrupta do CMO foi ocasionada pela geração de cenários de afluência com valores atipicamente superiores, gerados pelo sorteio de ruídos aleatórios com média bastante elevada, cuja probabilidade de ocorrência era bem pequena ($\sim 0.27\%$).

Ainda durante o ciclo de estudos de 2018/2019, a questão da variabilidade amostral foi foco de diversos estudos. Uma das propostas para a sua redução foi a utilização da técnica de reamostragem de cenários *forward* [4], [5]. Observou-se que sua utilização aprimorou a construção da política de operação e reduziu a variabilidade amostral, fixada uma mesma árvore de cenários utilizada na solução do problema de planejamento da operação (árvore completa). Porém, ao realizar estudos variando a árvore de cenários completa, definida pelos cenários da recursão *backward*, observou-se ainda uma variabilidade amostral nos resultados.

A partir destas análises, o CEPEL investigou aprimoramentos no processo de geração de cenários de afluências de forma a minimizar a ocorrência de cenários tão atípicos e a variabilidade amostral. No relatório técnico intitulado “Avaliação de Alternativas para Escolha do Representante no Processo de Agregação da Amostragem Seletiva” [6] foi apresentada uma alternativa para a escolha do representante do processo de agregação do processo de Amostragem Seletiva (AS) [7], implementada no modelo NEWAVE, que resultou em uma redução na variabilidade amostral observada nos resultados do planejamento da operação de médio/longo prazo quando se variava a árvore de cenários para a simulação *backward*.

Este relatório tem como objetivo avaliar o uso da alternativa proposta para o representante do processo de agregação na geração de cenários de vazões feita pelo modelo GEVAZP [8], [9], [10] também para o modelo de planejamento da operação de curto prazo DECOMP [11].

Nesta seção, apresentou-se a motivação para o aprimoramento no processo de AS. Na seção 2 é apresentada a forma atual de utilização da técnica de agregação AS para a geração dos cenários de afluências no modelo de planejamento da operação energética de curto prazo DECOMP, e a proposta apresentada neste relatório. Na seção 3 são apresentados os estudos de caso e os critérios de avaliação das metodologias atual e proposta. Na seção 4, apresentam-se os resultados que confirmam a redução da variabilidade amostral obtida com a nova alternativa proposta neste relatório, tanto na qualidade dos cenários de vazões, quanto nos resultados da política energética de curto prazo e, por fim, a seção 5 traz as conclusões.

2 GERAÇÃO DE CENÁRIOS

O modelo GEVAZP é responsável pela geração de cenários sintéticos multivariados de energia e vazão mensais que são utilizados pelos modelos NEWAVE, SUIISHI [12] e DECOMP. Esses cenários são essenciais para o cálculo da política de operação ótima no problema de otimização multi-estágio e multi-reservatório do SIN.

Para desempenhar essa tarefa, o modelo GEVAZP adota em sua modelagem estocástica a família de modelos autorregressivos periódicos de ordem p , PAR(p) [13], [14]. A escolha desse modelo é justificada pela sua capacidade de capturar o comportamento periódico da estrutura de autocorrelação observado comumente em séries de afluições mensais. O modelo PAR(p) pode ser descrito matematicamente pela Equação (2.1):

$$\left(\frac{Z_t - \mu_m}{\sigma_m}\right) = \varphi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}}\right) + \dots + \varphi_{pm}^m \left(\frac{Z_{t-pm} - \mu_{m-pm}}{\sigma_{m-pm}}\right) + a_t \quad (2.1)$$

Onde:

Z_t é a série hidrológica sazonal de período T ;

N é número de anos;

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

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

μ_m é a média sazonal do período s ;

σ_m é o desvio-padrão sazonal do período s ;

φ_i^m é o i -ésimo coeficiente autorregressivo do período m ;

p_m é a ordem do operador de defasagem de período m ;

a_t é a série de ruídos independentes com média zero e variância σ_a^{2m}

Usualmente, assume-se que os ruídos a_t na equação (2.1) possuem distribuição normal e são independentes e identicamente distribuídos. Se for constatada a não normalidade dos ruídos, pode-se aplicar a transformação Box-Cox [15]. No caso da modelagem de afluições feita pelo modelo GEVAZP, emprega-se a série histórica original sem transformação e, conseqüentemente, deve-se modelar ruídos que demonstram uma distribuição assimétrica. A solução adotada passa pelo ajuste de uma distribuição lognormal três parâmetros [16].

O modelo GEVAZP é utilizado de forma a representar a incerteza hidrológica no modelo de planejamento da operação energética de curto prazo, DECOMP, através de uma árvore de cenários de afluições, conforme a Figura 2.1.

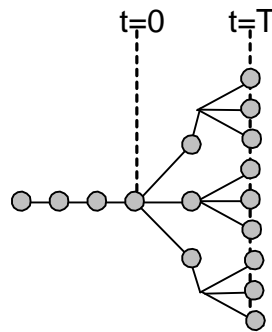


Figura 2.1 – Árvore de cenários de afluências utilizada no DECOMP.

Para que o processo estocástico de afluências seja representado corretamente, a árvore de cenários deve conter muitos ramos, de forma a descrever adequadamente as características do processo representado. Porém, para o modelo DECOMP, devido a restrições de tempo computacional, é interessante trabalhar com um número reduzido de cenários de afluência. O modelo GEVAZP faz uso da técnica de Amostragem Seletiva, através do uso da técnica de agregação K-means [17], para obter um conjunto reduzido de cenários de vazões que seja capaz de conter toda a informação necessária para representar o processo estocástico de vazões afluentes.

2.1 Amostragem Seletiva

Atualmente, para a geração de cenários de vazões aplicada ao modelo DECOMP, primeiro gera-se uma quantidade grande de cenários multivariados de *vazões afluentes*. Posteriormente, aplica-se a técnica de agregação, obtendo-se, assim, o conjunto final de cenários representativos. O número de aberturas da árvore de cenários varia de acordo com o mês. A Figura 2.2 traz um fluxograma do processo utilizado no GEVAZP para a geração de cenários sintéticos de vazões afluentes para o DECOMP.

Neste relatório técnico, propõe-se uma alteração no processo de geração de cenários sintéticos de vazões afluentes para o modelo DECOMP. Assim como proposto para a geração de cenários de afluências para o modelo de planejamento da operação de médio/longo prazo em [6], o uso da técnica de agregação será realizado nos vetores de ruídos normais, sendo o objeto representativo de cada grupo os respectivos centroides. A adoção do centroide ao invés do objeto mais próximo como representante da Amostragem Seletiva apresentou uma maior robustez nos resultados decorrentes do planejamento da operação (geração térmica, geração hidráulica, custo marginal de operação, entre outros), além de gerar conjuntos de vetores de ruídos com menor dispersão na média e desvio-padrão. Deve-se observar que a escolha do centroide como representante só pode ser realizada na agregação dos ruídos. A Figura 2.3 ilustra o processo de escolha dos representantes.

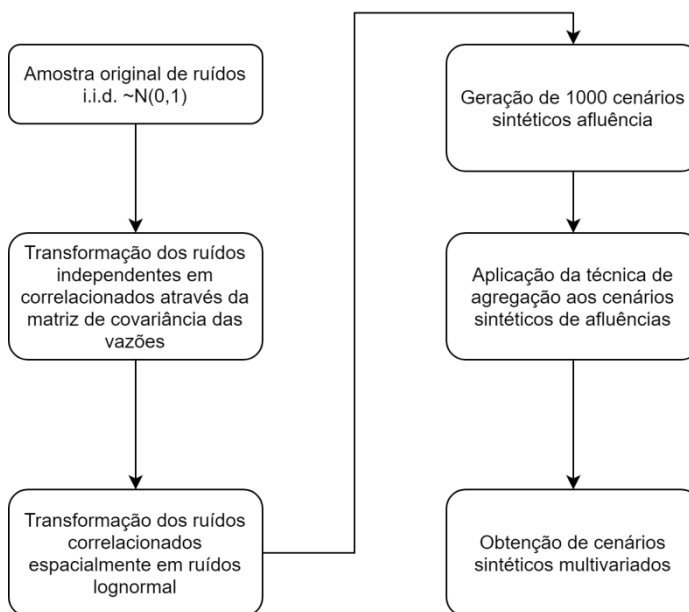


Figura 2.2 – Fluxograma do processo de geração de cenários sintéticos de vazões afluentes para o DECOMP.

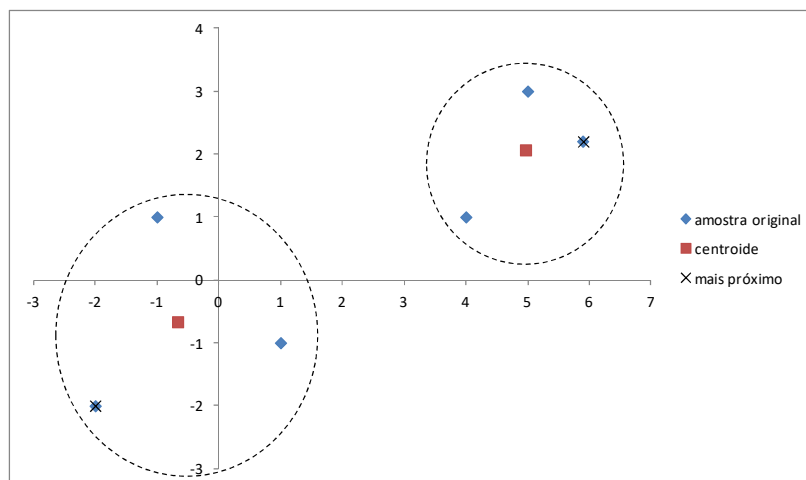


Figura 2.3: Exemplo ilustrativo da escolha do representante.

2.1.1 Fator de compensação do desvio-padrão

A aplicação das técnicas de agregação na amostra original, cujo algoritmo envolve a substituição de agrupamentos de objetos da amostra original por um único representante, resulta na obtenção de amostras agregadas com menor variabilidade do que a correspondente amostra original (a variabilidade interna nos agrupamentos é perdida). Essa degradação da amostra original é indesejada para a correta representação do processo estocástico. Apresenta-se a seguir um procedimento capaz de minimizar essa perda de variabilidade da amostra original [6], [18].

Quando aplicada a técnica de agregação nos vetores de ruídos normais, pode-se aplicar um fator de correção do desvio-padrão de forma a recuperar a variabilidade perdida anteriormente. Dado que o desvio-padrão dos ruídos normais que compõem a amostra original é igual a um, para corrigir a degradação do desvio-padrão da amostra obtida após o processo de agregação, pode ser calculado o seguinte fator de correção:

$$FC = \frac{\sum_{i=1}^{N_{UHE}} \frac{\text{desvio padrão desejado}}{\sigma_i}}{N_{UHE}} = \frac{\sum_{i=1}^{N_{UHE}} \frac{1}{\sigma_i}}{N_{UHE}} \quad (2.2)$$

onde:

σ_i é desvio-padrão dos ruídos agregados para a usina i ;

N_{UHE} é o número total de usinas participantes do processo de agregação;

FC é o fator de correção do desvio-padrão;

Este fator de correção é então aplicado aos vetores de ruídos agregados, tornando o desvio-padrão da amostra agregada aproximadamente igual a um.

A Figura 2.4 traz o fluxograma do processo de geração de cenários de aflúências proposto neste relatório.

Durante a geração da árvore de aflúências, a agregação dos ruídos é realizada apenas uma vez por período. A partir do segundo período de geração, os ruídos agregados são repetidos para todos os ramos da árvore, assim como demonstrado na Figura 2.5.

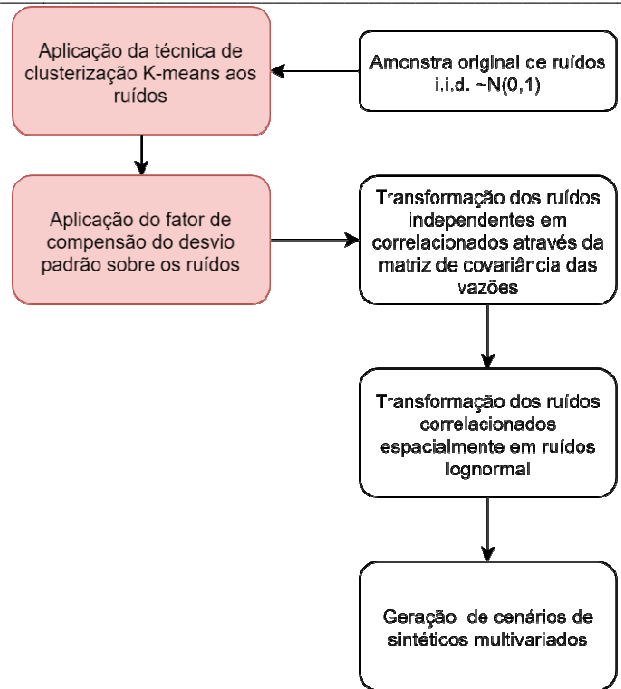


Figura 2.4- Fluxograma do processo de geração de cenários de afluências proposto neste relatório.

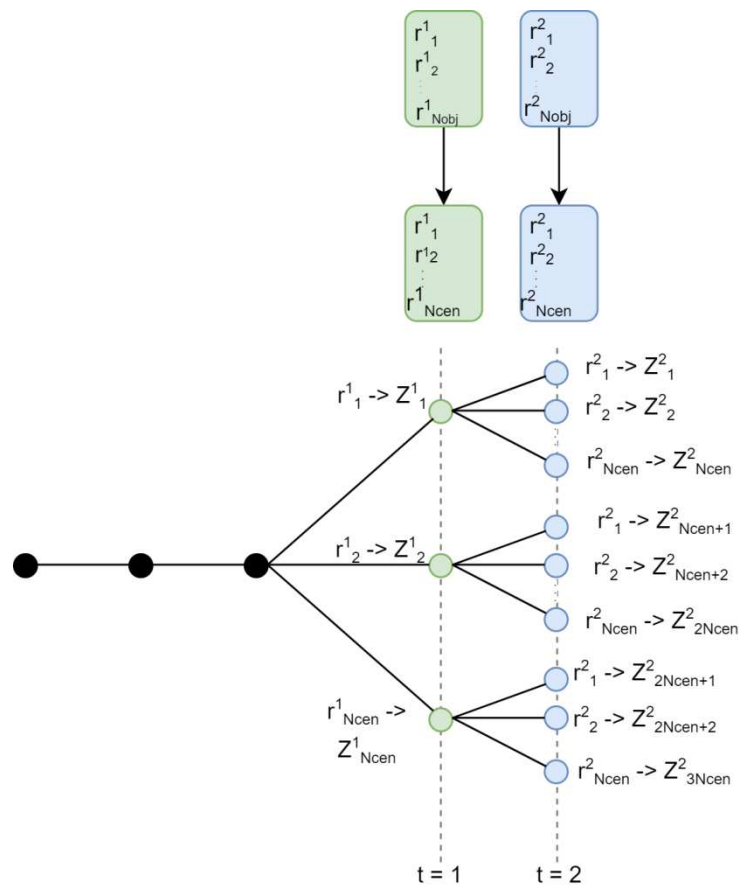


Figura 2.5 – Construção da árvore de cenários com agregação dos ruídos

3 ESTUDOS DE CASO E MÉTRICAS DE AVALIAÇÃO

3.1 Avaliação da qualidade da geração de cenários de afluências

Para a análise dos resultados da geração de cenários de vazões afluentes, selecionaram-se dois PMOs, um representante do período úmido brasileiro (fevereiro de 2019) e o segundo alocado no período seco (junho de 2016). Para os dois PMOs, são ajustados modelos PAR(p) para 79 e 76 usinas hidroelétricas (UHEs), respectivamente. O número de aberturas da árvore de cenários é de 143 para fevereiro e 513 para junho. Em cada um deles, foram feitas as seguintes variações de casos:

1. Geração de cenários utilizando a agregação nas vazões, com o objeto representante sendo o objeto mais próximo – amostra original de 1000 objetos (método utilizado oficialmente);
2. Geração de cenários utilizando agregação nos vetores de ruídos normais independentes, com o objeto representante sendo o centroide do grupo – amostra original de 1.000 objetos;
3. Geração de cenários utilizando agregação nos vetores de ruídos normais independentes, com o objeto representante sendo o centroide do grupo – amostra original de 5.000 objetos;
4. Geração de cenários utilizando agregação nos vetores de ruídos normais independentes, com o objeto representante sendo o centroide do grupo – amostra original de 10.000 objetos;
5. Geração de cenários utilizando agregação nos vetores de ruídos normais independentes, com o objeto representante sendo o centroide do grupo – amostra original de 50.000 objetos;
6. Geração de cenários utilizando agregação nos vetores de ruídos normais independentes, com o objeto representante sendo o centroide do grupo – amostra original de 100000 objetos.

De forma a se avaliar a robustez dos resultados, para cada uma das opções, foram feitas rodadas com a utilização de dez sementes diferentes para o processo de geração de ruídos pseudoaleatórios.

Dado que a cada UHE é ajustado um modelo PAR(p) específico, e que a geração de cenários de afluências é realizada condicionada ao passado recente, podem ser calculados quais são os valores teóricos de média e desvio-padrão de cada UHE, dado pelas Equações (3.1) e (3.2), respectivamente.

$$E(Z_t|Z_{t-1}) = \left[\varphi_1^m \left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \varphi_{pm}^m \left(\frac{Z_{t-pm} - \mu_{m-pm}}{\sigma_{m-pm}} \right) + \mu_m \right] \quad (3.1)$$

$$\sigma_{Z_t|Z_{t-}} = \sigma_m \sigma_a^m = \sigma_m \sqrt{1 - \varphi_1^m \rho^m(1) - \varphi_2^m \rho^m(2) - \dots - \varphi_p^m \rho^m(p^m)} \quad (3.2)$$

onde:

$E(Z_t|Z_{t-})$ é a média condicionada do modelo autorregressivo ajustado;

$\sigma_{Z_t|Z_{t-}}$ é o desvio-padrão condicionado do modelo autorregressivo ajustado.

De forma a avaliar os resultados encontrados, define-se o índice de similaridade definido como a razão das estatísticas agregadas pelos respectivos valores teóricos conforme a equação (3.3):

$$\omega = \frac{\hat{\theta}_{ag}}{\theta_{vt}} \quad (3.3)$$

onde $\hat{\theta}_{ag}$ é o valor da estatística (médio ou desvio-padrão) após o processo de agregação e θ_{vt} o valor teórico das estatísticas. Quanto mais próximo do valor unitário, melhor o processo de geração/agregação. Para cada das seis opções de geração de cenários, em cada PMO, é apresentado um único gráfico do tipo *box-plot*, que contém o índice de similaridade de todas as UHEs envolvidas nos casos de PMO avaliados neste relatório.

3.2 Métricas para a avaliação dos resultados do planejamento da operação energética de curto prazo

A avaliação de resultados obtidos pelo modelo DECOMP será baseada na análise do CMO e geração térmica do SIN da primeira semana de estudo. Para isso, foram rodadas todas as revisões dos PMOs desde janeiro de 2014 até fevereiro de 2019. Conforme será apresentado na próxima seção, a geração de cenários com agregação nos ruídos com amostra original de 50mil objetos apresentou o melhor custo/benefício em termos de qualidade dos cenários obtidos e tempo computacional, sendo esta a proposta a ser avaliada. Sendo assim, serão comparados os resultados obtidos com a metodologia atualmente empregada (agregação nas vazões) e a metodologia proposta para dez sementes diferentes em cada um dos PMOs.

Deve-se observar que não há nenhuma alteração metodológica ou de implementação no modelo DECOMP para a utilização dos conjuntos de cenários de afluências fornecidos com a metodologia proposta. Em todos os casos analisados, a única diferença nos dados de entrada são os conjuntos de cenários.

4 RESULTADOS

4.1 Qualidade dos cenários de vazões afluentes

A Figura 4.1 apresenta os resultados obtidos para a revisão 2 do PMO de fevereiro de 2019. Nas Figuras 4.1(a) e 4.1(b) expõem-se os *boxplots* do índice de similaridade das médias para as sementes 0 e 1. Cada *boxplot* contém os valores do índice de similaridade para todas as UHEs avaliadas no PMO em questão. É possível observar que, para a reprodução da média teórica, não há aprimoramento nos resultados ao se passar do processo de agregação em vazões para a agregação em ruídos, com uma amostra original de mil objetos. Porém, nota-se que à medida que se aumenta o tamanho da amostra original, os índices de similaridade tendem a ficar mais próximos do valor unitário, e com menor dispersão. Observa-se que, a partir da utilização de uma amostra original de 5mil há uma melhor reprodução das médias teóricas. A partir de 50mil objetos na amostra original, observa-se que não há uma grande alteração de resultados. Observa-se também que há pouca variabilidade dos resultados entre as sementes.

Nas Figuras 4.1(c) e 4.1(d) apresentam-se os resultados para o desvio-padrão. Pode-se observar que o primeiro *boxplot* tem sua mediana (linha vermelha) abaixo de 0,9 em quase todas as sementes, o que denota a perda da variabilidade dos cenários de afluências gerados. No caso da agregação dos cenários nas vazões, essa degradação não consegue ser recuperada, e os cenários tendem a possuir desvio-padrão menor do que o teórico. Ao se passar para a agregação nos ruídos com centroide, a perda do desvio-padrão é compensada por um fator de correção. Esse fato pode ser observado pela mediana próxima ao valor unitário a partir do segundo *boxplot*. Novamente, à medida que se aumento o tamanho da amostra original, os gráficos ficam um pouco menos dispersos, e também menos variáveis entre as diferentes sementes.

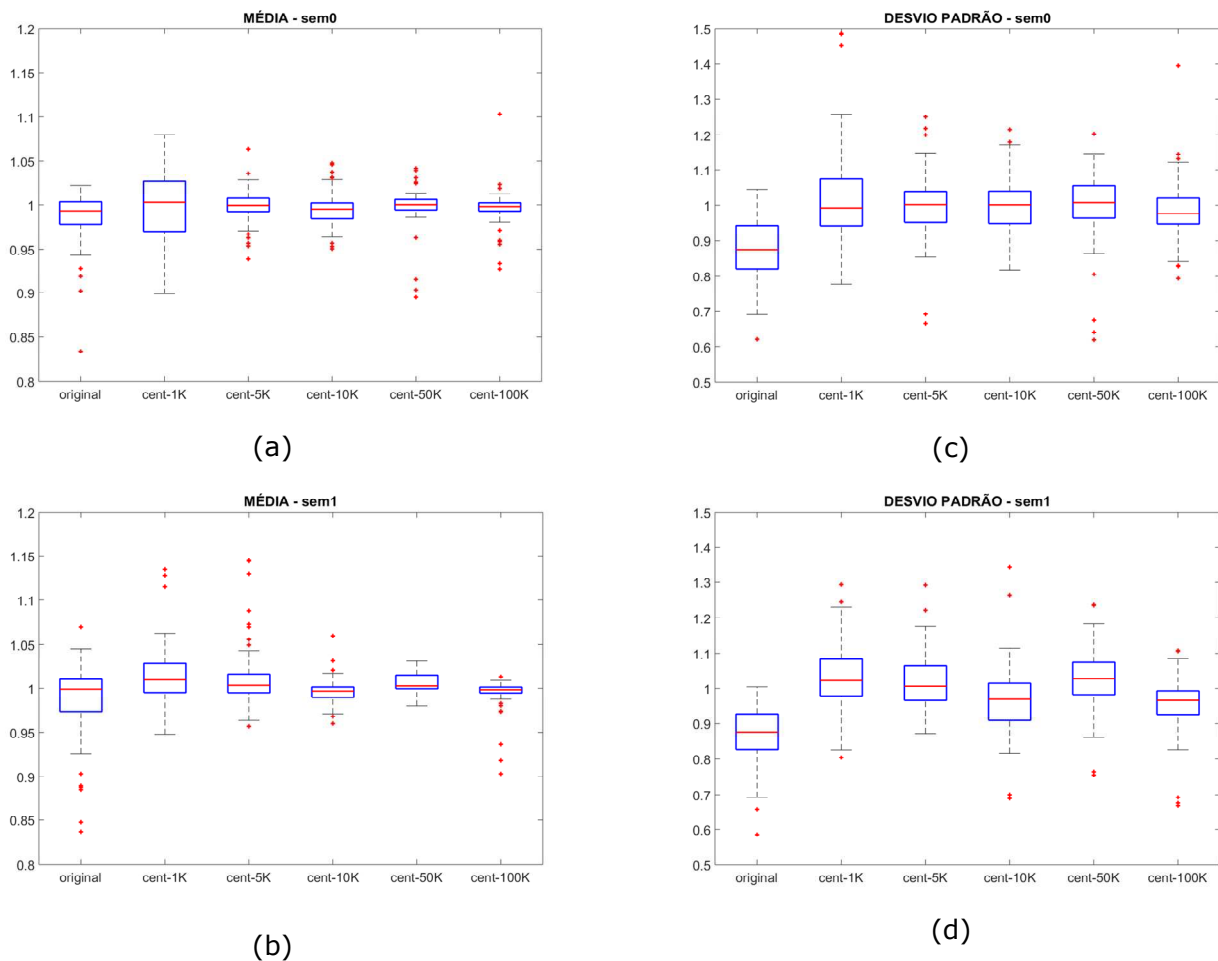
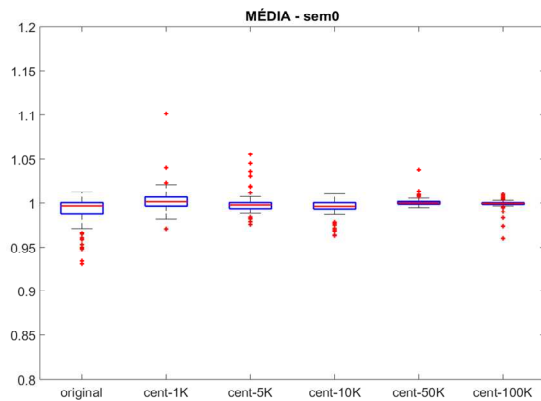
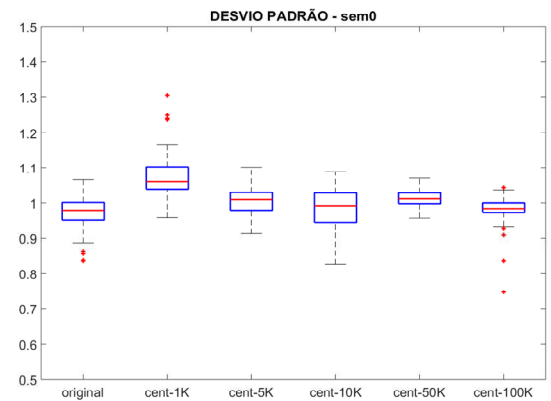


Figura 4.1 – Resultados do índice de similaridade ω para a média (a) e (b) e para o desvio-padrão (c) e (d) para a revisão 2 do PMO de fevereiro de 2019.

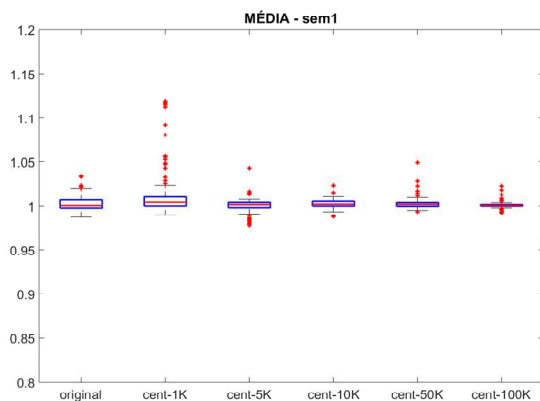
A Figura 4.2 traz os resultados obtidos para o PMO de junho de 2016 para duas primeiras diferentes. Nas Figuras 4.2(a) e 4.2(b) apresentam-se os *boxplots* referentes à média e ao desvio-padrão respectivamente. Primeiramente, nota-se que os resultados tanto para a média quanto para o desvio-padrão nesse PMO são mais concentrados em torno do valor unitário, quando comparados com os resultados do PMO de fevereiro. Esse fato pode estar relacionado com a maior quantidade de aberturas da árvore de cenários no PMO de junho. Tanto para as médias quanto para o desvio-padrão, é possível observar que os resultados tendem a se concentrar próximo ao valor unitário quanto maior o tamanho da amostra original. Novamente, a partir de 50mil não há um aprimoramento relevante no resultado comparado com o caso de 100mil objetos.



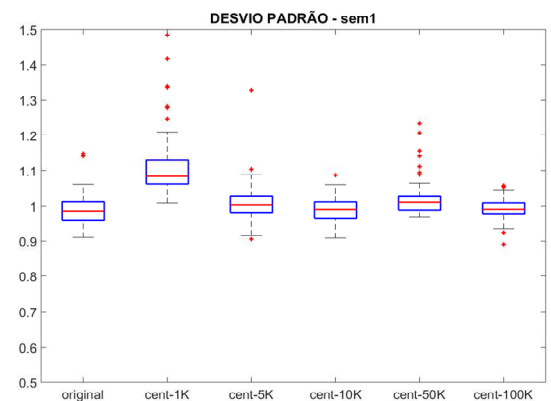
(a)



(c)



(b)



(d)

Figura 4.2 – Resultados do índice de similaridade ω para a média (a) e (b) e para o desvio-padrão (c) e (d) para o PMO de junho de 2016.

Deve-se levar em conta, que com o aumento da amostra original, o tempo computacional tende a crescer. A Tabela 4.1 mostra o tempo computacional médio para a execução de cada variação do PMO de fevereiro de 2019. Pode-se observar que ao passar de 50mil objetos para 100mil o tempo computacional tem um incremento considerável. Dada a qualidade dos resultados obtidos com a amostra original de 50mil e levando em consideração o tempo computacional necessário para a execução dos casos, elegeu-se esse valor para as próximas análises.

Tabela 4.1 – Tempo computacional da execução variando o tamanho da amostra original para o PMO de fevereiro de 2019.

Agregação	Tempo (s)
Geração com agregação tradicional: agregação nas vazões com mais próximo – amostra original = 1000 objetos (método usado oficialmente):	6
Geração com agregação pelo centroide nos ruídos e 1mil objetos	4
Geração com agregação pelo centroide nos ruídos e 5mil objetos	6
Geração com agregação pelo centroide nos ruídos e 10mil objetos	7
Geração com agregação pelo centroide nos ruídos e 50mil objetos	31
Geração com agregação pelo centroide nos ruídos e 100mil objetos	104

A seguir, a Figura 4.3 apresenta a distribuição acumulada de frequência dos cenários de Energia Natural Afluente (ENA) do subsistema Sudeste, calculadas a partir das vazões geradas para o PMO de fevereiro de 2019. A Figura 4.3(a) apresenta os resultados obtidos com a agregação nas vazões para as 10 sementes, e a Figura 4.3(b) os resultados com a utilização da agregação nos ruídos, com o centroide como representante, e amostra original de 50mil objetos. É possível observar que há uma maior dispersão das curvas no gráfico (a) em comparação com o (b). Pode-se então afirmar que a metodologia proposta diminuiu a variabilidade amostral apresentada em (a).

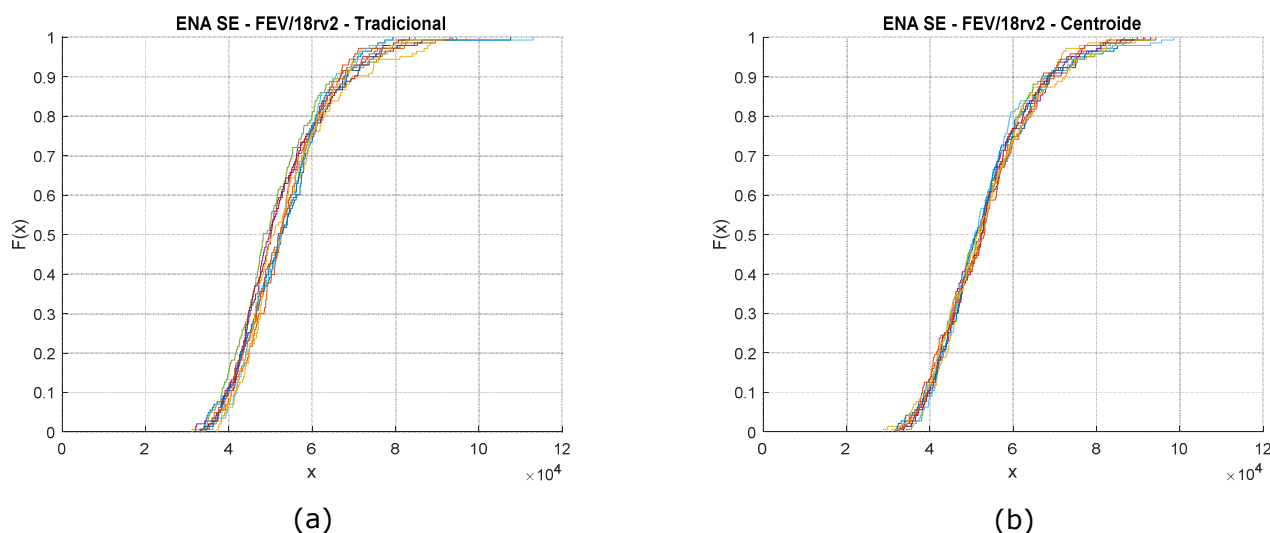


Figura 4.3 – Distribuição de frequência acumulada das 10 sementes geradas utilizando a metodologia oficial (a) e a proposta (b) para o PMO de fevereiro de 2019.

A Figura 4.4 expõe as distribuições de frequência das ENAs do subsistema sudeste, obtidas para o PMO de junho de 2016 com a metodologia oficial (a) e a proposta (b). Nota-se que por ser um mês seco, e possuir um número maior de aberturas da árvore de cenários, as distribuições obtidas pelo método (a) na Figura 4.4 são menores que a apresentada na Figura 4.3(a). Ainda assim, a Figura

4.4(b) apresenta uma menor dispersão das distribuições, demonstrando mais uma vez o aprimoramento na qualidade dos cenários com relação à redução da variabilidade amostral.

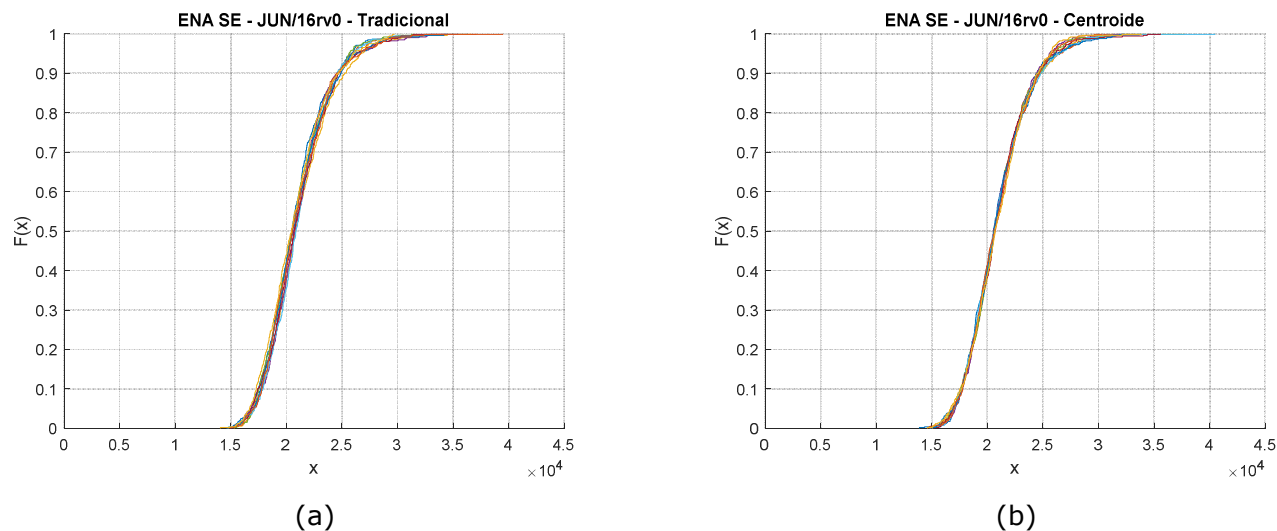
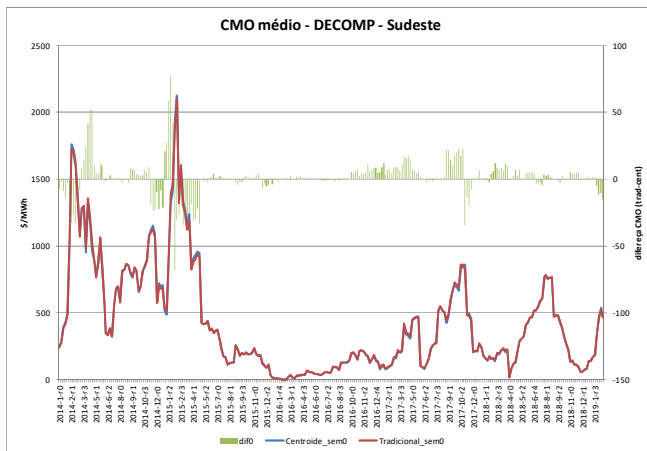


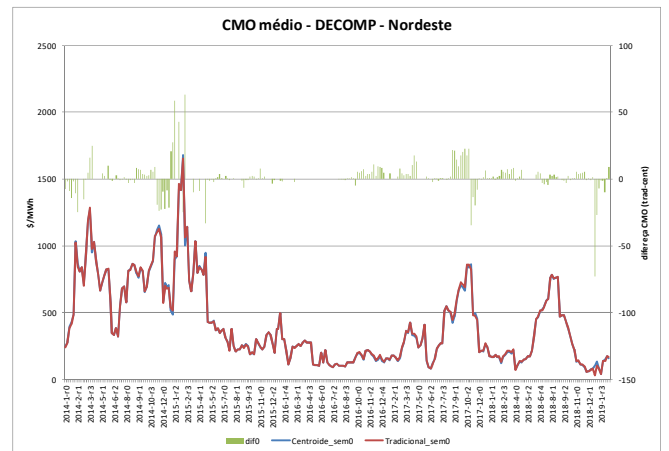
Figura 4.4 – Distribuição de frequência acumulada das 10 sementes geradas utilizando a metodologia oficial (a) e a proposta (b) para o PMO de junho de 2016.

4.2 Avaliação dos resultados do planejamento da operação energética de curto prazo

Foram rodados todos os PMOs desde janeiro de 2014 até fevereiro de 2019, com a metodologia utilizada oficialmente para a geração de cenários, e a metodologia proposta. É importante frisar que não houve alteração nos dados de entrada de cada caso, com exceção dos cenários de vazões afluentes. Para cada revisão, foram feitas gerações de cenários a partir de 10 sementes diferentes, com o intuito de se observa a variabilidade dos resultados, principalmente de CMO. A Figura 4.5 apresenta a evolução do CMO do primeiro estágio para o subsistema sudeste (a) e nordeste (b) no período analisado, apenas para a semente original. A linha vermelha representa os resultados obtidos com a metodologia tradicional, e a linha azul os obtidos com a metodologia proposta. As barras em verde referem-se às diferenças entre as duas curvas. Pode-se observar que, apesar de haver algumas pequenas diferenças entre as duas alternativas, não há alteração significativa da trajetória das curvas.



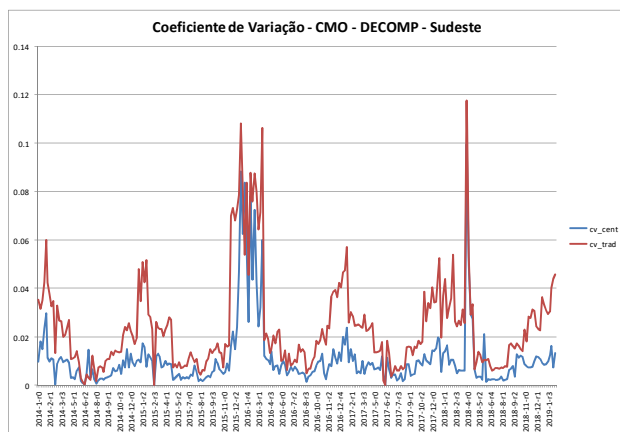
(a)



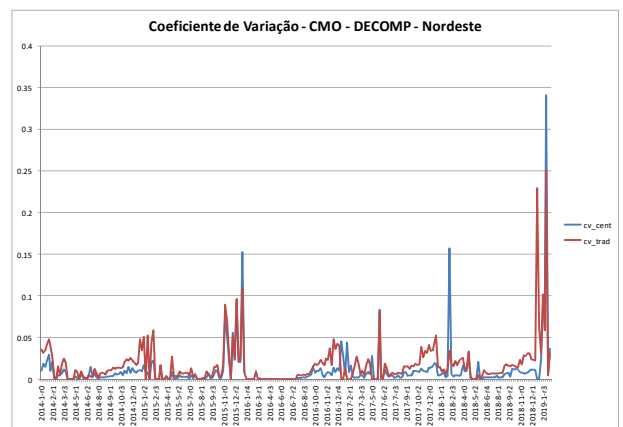
(b)

Figura 4.5 – Evolução do CMO do primeiro estágio obtido pelo método tradicional (linha vermelha) e com a metodologia proposta (linha azul) para a semente 0, e suas diferenças (barras em verde) para o subsistema sudeste (a) e nordeste (b).

A Figura 4.6 apresenta o coeficiente de variação do CMO das 10 sementes, calculados para todas as revisões, para os subsistemas sudeste (a) e nordeste (b). É possível observar que na maior parte dos casos o coeficiente de variação do CMO é menor nos casos com a metodologia proposta, para os dois subsistemas. Ou seja, alterando-se as sementes, os CMOs obtidos na metodologia proposta possuem uma menor variabilidade.



(a)



(b)

Figura 4.6 – Coeficiente de variação do CMO do primeiro estágio para os 10 conjuntos de cenários obtidos pelo método tradicional (linha em vermelho) e com a metodologia proposta (linha azul) para o subsistema sudeste (a) e nordeste (b).

Por fim, a Figura 4.7 expõe a diferença entre o CMO máximo e mínimo obtido entre as sementes de cada PMO, para o subsistema sudeste (a) e nordeste (b). pode-se observar que a metodologia proposta possui, na maioria dos casos, uma diferença menor entre os CMOs obtidos com seus conjuntos de cenários de afluências.

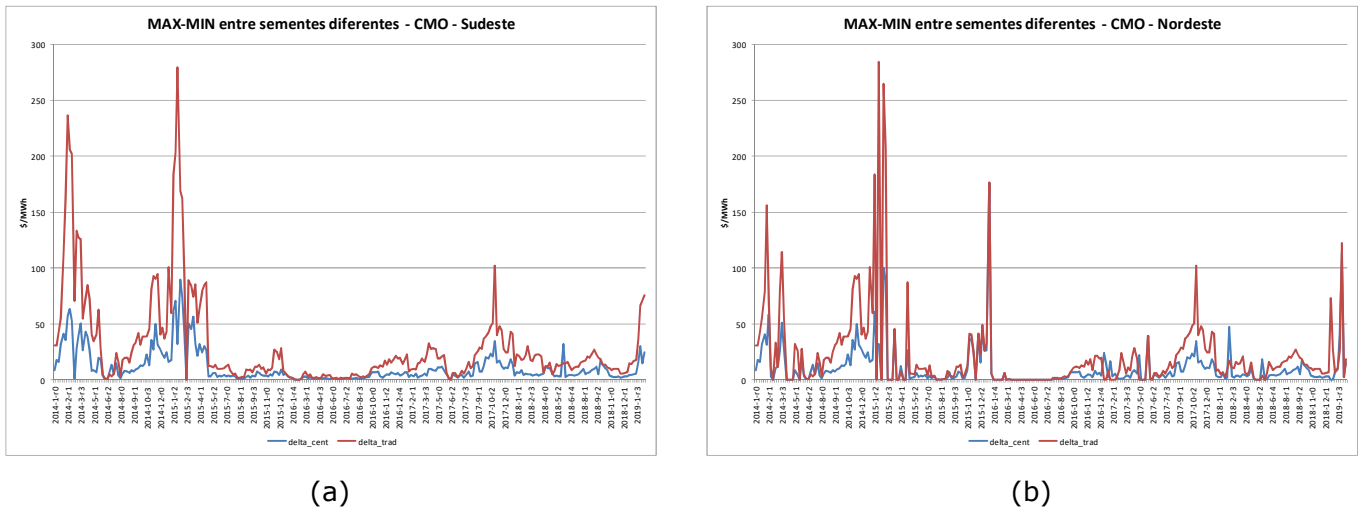


Figura 4.7 – Diferença do CMO máximo e mínimo de primeiro estágio para os 10 conjuntos de cenários obtidos pelo método tradicional (linha em vermelho) e com a metodologia proposta (linha azul) para o subsistema sudeste (a) e nordeste (b).

Para avaliar o impacto da utilização do centroide na operação energética de curto prazo, será feita uma análise da geração térmica e hidráulica do SIN na primeira semana. Na Figura 4.8 apresenta-se a evolução da geração térmica (a) e da geração hidráulica (b) para a primeira semana de planejamento para o SIN, obtidas com as duas opções de geração de cenários, e a diferença entre elas, para a semente original. É possível observar que ambas as curvas apresentam o mesmo perfil, com diferenças nulas na maior parte do tempo.

A Figura 4.8(b) expõe o coeficiente de variação da geração térmica do SIN obtido com as 10 sementes. Nota-se que, novamente, a utilização do centroide causa uma redução da variabilidade dos resultados entre as 10 sementes avaliadas.

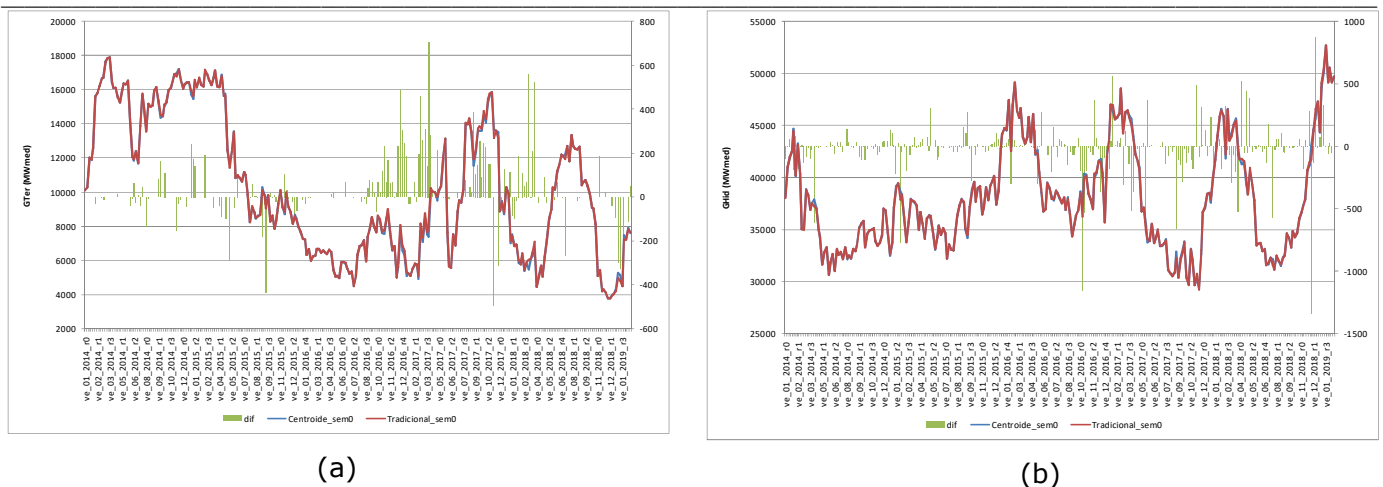


Figura 4.8 – (a) Evolução da geração térmica (MWmed) da primeira semana para o SIN, com a opção centroide (linha azul) e tradicional (linha vermelha) e suas diferenças (barras verdes) para a semente original. (b) Evolução do coeficiente de variação da geração térmica do SIN na primeira, obtido com 10 sementes.

A Figura 4.9 expõe o coeficiente de variação da geração térmica (a) e da hidráulica (b) do SIN obtido com as 10 sementes. Nota-se que, novamente, a utilização do centroide causa uma redução da variabilidade dos resultados entre as 10 sementes avaliadas, em ambas as variáveis.

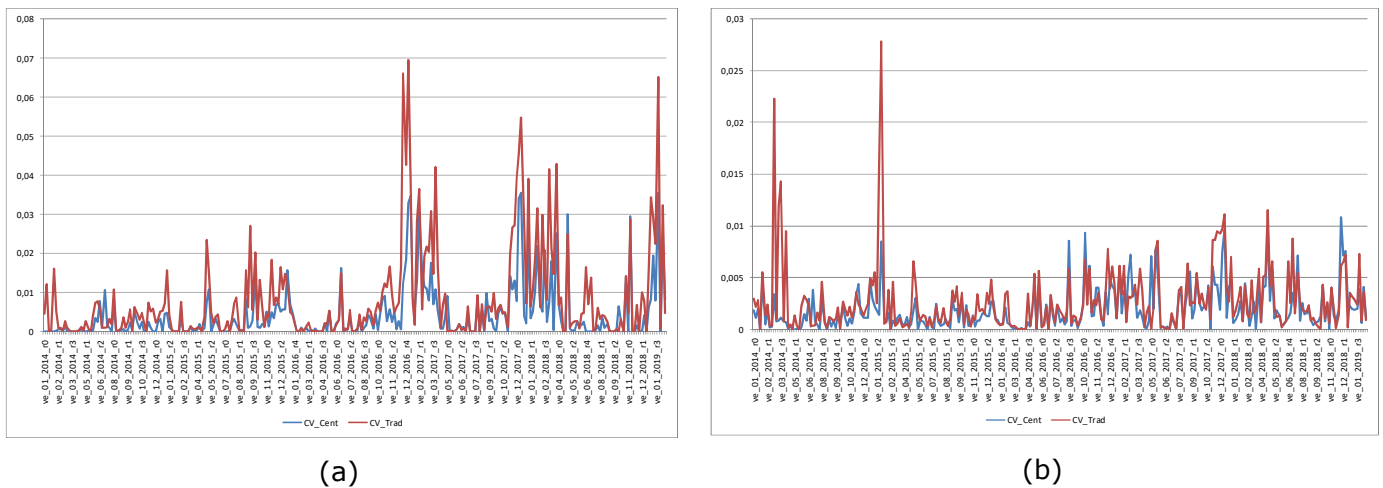


Figura 4.9 - Coeficiente de variação da geração térmica (a) e hidráulica (b) do primeiro estágio para os 10 conjuntos de cenários obtidos pelo método tradicional (linha em vermelho) e com a metodologia proposta (linha azul), para o SIN.

5 CONCLUSÕES

Recentemente, avaliou-se no modelo NEWAVE uma alternativa para a escolha dos objetos representantes do processo de agregação dos ruídos (uso do centroide), que foi capaz de diminuir a variabilidade amostral dos resultados tanto para os cenários de ENAs, quanto para a política energética. Este relatório técnico teve como objetivo realizar uma análise dos resultados obtidos na geração de cenários de vazões pelo modelo GEVAZP e na política energética de curto prazo obtida pelo modelo DECOMP, com o uso da mesma metodologia.

A consideração do centroide como representante do processo de agregação nos ruídos, em conjunto com o aumento do tamanho da amostra original, reduziu a variabilidade amostral e aumentou a qualidade dos cenários gerados de afluência. Ao contrário do que ocorre na agregação aplicada diretamente nos cenários de vazões, não há uma degradação do desvio-padrão dos cenários de afluências, pois esta é corrigida através de um fator de compensação.

Analisando o custo/benefício da melhoria de representatividade nos cenários de afluências e o tempo computacional gasto na geração dos cenários, escolheu-se o tamanho da amostra original de 50mil. Definido esse parâmetro, foi feita a comparação dos CMOs e da geração térmica, ambas as variáveis para o primeiro estágio do planejamento da operação de curto prazo. Os resultados foram obtidos para o período entre janeiro de 2014 e fevereiro de 2019 com a metodologia oficial e a proposta para a geração de cenários, utilizando-se 10 conjuntos de cenários distintos (variação de semente). Foi observada uma redução na variabilidade tanto para o CMO quanto para a geração térmica, obtidos pelo DECOMP, variando-se as sementes. Os valores das variáveis analisadas para o conjunto original de cenários de cada metodologia ficaram próximos.

Desta forma, assim como foi feito no módulo de geração de cenários de afluências do modelo NEWAVE, recomenda-se utilizar o centroide como representante dos grupos e, para tal, aplicar a agregação nos ruídos. Recomenda-se também uma amostra original de 50mil objetos para o processo de agregação.

6 REFERÊNCIAS

- [1] M.E.P. Maceira, V.S. Duarte, D.D.J. Penna, L.A.M. Moraes, A.C.G. Melo, "Ten years of application of stochastic dual dynamic Programming in official and agent studies in Brazil – Description of the NEWAVE program", 16th PSCC – Power Systems Computation Conference, Glasgow, Scotland, 2008.
- [2] M.E.P. Maceira, D.D.J. Penna, A.L. Diniz, R.J. Pinto, A.C.G. Melo, C.V. Vasconcellos, C.B. Cruz, "Twenty years of application of stochastic dual dynamic Programming in official and agent studies in Brazil – Main features and improvements on the NEWAVE model", 20th PSCC – Power Systems Computation Conference, Ireland, 2018.
- [3] D.D.J. Penna, M.E.P. Maceira, J.M. Damázio, F. Treistman, H.S. Araújo, A.C.G. Melo, "Análise da inflexão do custo marginal de operação no modelo NEWAVE entre os quarto e quinto estágios temporais ao se adotar a representação de 12 REEs para Sistema Interligado Nacional", Relatório Técnico CEPEL no 10783/2018, Setembro de 2018.
- [4] M.E.P. Maceira, D.D.J. Penna, C.B. Cruz, A.L. Diniz, A.C.G. Melo, F. Treistman, "Técnicas de Reamostragem de Cenários Hidrológicos para a simulação forward da PDDE na Definição da Estratégia de Operação Energética de Longo/Médio Prazos", Relatório Técnico Cepel 6606/2018, Junho 2018.
- [5] D.D.J. Penna, M.E.P. Maceira, A.L. Diniz, A.C.G. Melo, C.B. Cruz, "Avanços na Reamostragem de Cenários Hidrológicos Aplicada ao Planejamento da Operação Energética de Médio Prazo", XIV SEPOPE, Recife –PE, Set-Out 2018.
- [6] D.D.J. Penna, F. Treistman, M.E.P. Maceira, "Avaliação de alternativas para escolha do representante no processo de agregação da Amostragem Seletiva", Relatório Técnico CEPEL 13612/2018, Dezembro 2018.
- [7] D.D.J. Penna, M.E.P. Maceira, J.M. Damázio, "Selective sampling applied to long-term hydrothermal generation planning", 17th PSCC - Power Systems Computation Conference, Stockholm, Sweden, August 2011.
- [8] M.E.P. Maceira, C.V. Bezerra, "Stochastic Streamflow model for Hydroelectric Systems", In: Proceedings of 5th International Conference on Probabilistic Methods Applied to Power Systems, pp. 305-310, Vancouver, Canada, September 1997.
- [9] M.E.P. MACEIRA; D.D.J. PENNA; J.M. DAMÁZIO, "Geração de cenários sintéticos de energia e vazão para o planejamento da operação energética". Cadernos do IME – Série Estatística/UERJ, 21(2), 11 p, 2006.
- [10] D.D.J. Penna, M.E.P. Maceira, J.M. Damázio, F. Treistman, H.S. Araújo, Manual de Referência do modelo GEVAZP, Relatório Técnico CEPEL – nº 27155/2017, 2017.
- [11] A.L. Diniz, F.S. Costa, M.E.P. Maceira, T.N. Santos, L.C. Brandão, R.N. Cabral, "Short/Mid-Term Hydrothermal Dispatch and Spot Pricing for Large-Scale Systems - the Case of Brazil", 20th Power Systems Computation Conference, Dublin, Ireland, June 2018.
- [12] M.E.P. Maceira; F.R.S. Batista; A.C.G. Melo; L.F.E.C. Silva; R. Olasagasti; L.G.B. Marzano, "A Probabilistic Approach to Define the Amount of Energy to be Traded in Hydro Dominated Interconnected Systems", 20th PSCC – Power Systems Computation Conference, Ireland, 2018.
- [13] J.D. Salas, J.W. Delleur, V. Yevjevich, W.L. Lane, "Applied Modeling of Hydrologic Time Series", Water Resources Publications, 1980.

[14] W.H. Hipel, A.I. Mcleod, "Time series modeling of water resources and environmental systems". Elsevier, 1994.

[15] G.E.P.Box, D.R.Cox, "An Analysis of Transformations", Journal of The Royal Statistical Society, A127, pp. 211-252, 1964.

[16] R. J. Charbeneau, "Comparison of the two and three parameter lognormal distributions used in streamflow synthesis", Water Resources Research, Vol. 14, No. 1, pp. 149-150, 1978.

[17] J.Hartigan, M. Wong, "A K-Means Clustering Algorithm", Applied Statistics, vol.28, no. 1, pp. 100-108, 1979.

[18] F. Treistman, D.D.J. Penna, H.S. Araújo, M.E.P. Maceira, J.M. Damázio, "Impacto da Aplicação da Técnica de Agregação na Construção da Árvore de Cenários para o Planejamento da Operação de Curto Prazo", XXII Simpósio Brasileiro de Recursos Hídricos – SBRH, Florianópolis, 2017.