

**RELATÓRIO TÉCNICO**

<b>Nº:</b> DP/DEA – 47530/09	<b>Nº DE PÁGINAS:</b> 137	<b>ANEXOS:</b> 0
<b>TÍTULO:</b> Notas técnicas referentes ao projeto NEWAVE (nº 1345) emitidas entre os anos de 2007 e 2009.		
<b>DEPARTAMENTO:</b> DEPARTAMENTO DE OTIMIZAÇÃO ENERGÉTICA E MEIO AMBIENTE – DEA		
<b>ÁREA DE RESPONSABILIDADE:</b> B200		<b>Nº DO PROJETO:</b> 1600
<b>CLIENTE:</b> OPERADOR NACIONAL DO SISTEMA – ONS Rua da Quitanda 196, 14º andar 20091 000, Centro – Rio de Janeiro		
<b>ATENÇÃO:</b> Sr. Alberto Sérgio Kligerman		
<b>RESUMO:</b> Este relatório apresenta as notas técnicas que descrevem os desenvolvimentos metodológicos do programa NEWAVE realizados durante os anos de 2007 e 2009.		
<b>AUTORES:</b> Maria Elvira Piñeiro Maceira - CEPEL Vitor Silva Duarte – CEPEL Débora Dias Jardim Penna – CEPEL Jorge Machado Damázio - CEPEL Michel Pompeu Tcheou – CEPEL Renato Neves Cabral – CEPEL		<b>PALAVRAS-CHAVE:</b> Planejamento da Operação Interligada, Sistemas Hidrotérmicos, Programação Dinâmica Dual Estocástica
		<b>CLASSIFICAÇÃO:</b> CONTROLADO
 <b>GERENTE DO PROJETO</b> <b>NOME:</b> Maria Elvira Pineiro Maceira <b>TEL: (21) 2598-6454 FAX: (21) 2598-6482</b> <b>EMAIL:</b> <a href="mailto:elvira@cepel.br">elvira@cepel.br</a>		 <b>CHEFE DE DEPARTAMENTO</b> <b>NOME:</b> Maria Elvira Piñeiro Maceira <b>TEL: (21) 2598-6454 FAX: (21) 2598-6482</b> <b>EMAIL:</b> <a href="mailto:elvira@cepel.br">elvira@cepel.br</a>
<b>APROVAÇÃO</b>		
 <b>DIRETOR DE PESQUISA, DESENVOLVIMENTO E INOVAÇÃO</b> <b>ROBERTO PEREIRA CALDAS</b>  29/12/09		

## **RELATÓRIO TÉCNICO**

**Notas técnicas referentes ao projeto  
NEWAVE (nº 1345) emitidas  
entre os anos de 2007 e 2009**

## Sumário

1	NOTA TÉCNICA nº 29 - Possibilidade de adoção de simulação forward inicial para determinação dos estados de armazenamento para a primeira recursão backward no processo iterativo do modelo NEWAVE (Abril 2007).....	1
1.1	Introdução .....	1
1.2	Implementação.....	2
1.3	Entrada de dados .....	2
2	NOTA TÉCNICA nº 30 - Simulação final utilizando uma única série sintética com comprimento de 10.000 anos (Março 2007).....	3
2.1	Introdução .....	3
2.2	Modelagem .....	4
2.2.1	Módulo de geração de séries sintéticas.....	4
2.2.2	Módulo de simulação da operação do sistema .....	5
2.3	Entrada de dados .....	5
2.4	Rebatimento nos demais modelos da cadeia .....	5
3	NOTA TÉCNICA nº 31 - Variação temporal da restrição de vazão mínima obrigatória e consideração de penalidades ao não atendimento dessa restrição (Março 2007) .....	6
3.1	Introdução .....	6
3.2	Modelagem .....	7
3.2.1	Variação temporal da energia de vazão mínima.....	7
3.2.2	Penalidade pelo não atendimento à restrição de energia de vazão mínima obrigatória.....	8
3.3	Entrada de dados .....	12
3.4	Rebatimento nos demais modelos da cadeia .....	13
4	NOTA TÉCNICA nº 34 - Refinamento da Discretização de Estados para a Primeira Recursão do Processo Iterativo de Cálculo da Função de Custo Futuro, com	

Fixação de Subconjunto de Estados a Serem Visitados a Cada Recursão (Abril 2007)	
14	
4.1	Introdução ..... 14
4.2	Aperfeiçoamento da discretização inicial de estados para o cálculo da FCF17
4.2.1	Refinamento da discretização inicial dos valores da variável de estado níveis de armazenamento..... 17
4.2.2	Classificação dos cenários de aflúências..... 18
4.2.3	Associação de estados..... 19
4.3	Fixação de subconjunto de estados para todas as iterações..... 20
4.4	Entrada de dados ..... 20
4.5	Rebatimento nos demais modelos da cadeia ..... 21
5	NOTA TÉCNICA nº 35 - Consideração de restrição de disponibilidade de combustível para conjuntos de classes térmicas, da possibilidade de utilização de mais de um combustível por classe térmica e da possibilidade de variação mensal do custo unitário das classes térmicas. (Abril 2007)..... 22
6	Introdução..... 22
6.1	Modelagem ..... 23
6.2	Entrada de dados ..... 25
6.3	Rebatimento nos demais modelos da cadeia ..... 26
7	NOTA TÉCNICA nº 36 - Avaliação do impacto do ajuste na parcela de meta de desvio de água controlável da derivada do corte de Benders na versão 12b do NEWAVE (Setembro 2007)..... 27
8	NOTA TÉCNICA nº 39 - Alteração na indexação do vetor de configuração da parábola de geração hidráulica máxima - Revisão 1 (Setembro 2007) ..... 30
9	NOTA TÉCNICA nº 40a - Análise da nota técnica ONS - Metodologia para consideração do intervalo de confiança para as auto-correlações temporais para uso no modelo PAR(p) (Janeiro 2009)..... 36
9.1	Considerações Iniciais..... 36
9.2	Análise da Metodologia Proposta ..... 41
9.3	Conclusão ..... 43

9.4	Avaliação da nova proposta do ONS.....	44
9.5	Conclusão .....	58
10	Nota Técnica nº 42 - Aplicação de Técnicas de Agregação na Geração de Cenários Hidrológicos para o Planejamento de Médio Prazo (Novembro 2008) .....	60
10.1	Introdução .....	60
10.2	Análise de Conglomerados.....	61
10.2.1	Considerações Gerais .....	61
10.3	Medidas de Similaridade .....	63
10.4	Métodos de Agrupamento .....	65
10.4.1	Métodos Hierárquicos.....	66
10.4.2	Métodos Não Hierárquicos .....	67
10.5	Método K-Means .....	68
10.6	Aplicação no Modelo de Geração de Cenários.....	72
10.7	Cortes de Benders .....	77
10.8	Conclusões .....	79
11	NOTA TÉCNICA nº 43 - Representação da Correlação Cruzada na Geração de Cenários de Afluências para o Planejamento da Operação Energética de Médio Prazo(Setembro 2009).....	81
11.1	Introdução .....	81
11.2	Modelagem Matemática no GEVAZP .....	82
11.3	Estatísticas Médias e Desvio-Padrão .....	84
11.4	Estatística Correlações Cruzadas.....	92
11.5	Compensação da Correlação Cruzada.....	100
11.6	Estudo de Caso.....	109
11.7	Conclusões .....	113
11.8	BIBLIOGRAFIA .....	113
12	NOTA TÉCNICA nº 44 - INTERCÂMBIO MÍNIMO COM VARIÁVEL DE FOLGA (Junho 2009) .....	115

12.1	Introdução .....	115
12.2	Modelagem .....	115
12.3	Entrada de dados .....	117
12.4	Rebatimento nos demais modelos da cadeia .....	117
13	NOTA TÉCNICA nº 45 - REPRESENTAÇÃO DE RESTRIÇÕES DE INTERCÂMBIOS COM AGRUPAMENTO LIVRE - Atualização da Nota Técnica nº 21 (Março 2009) .....	118
13.1	Introdução .....	118
13.2	Alterações .....	118
13.3	Entrada de dados .....	119
13.4	Rebatimento nos demais modelos da cadeia .....	120
14	NOTA TÉCNICA nº 46 - Equalização das penalidades de intercâmbio de energia entre os subsistemas (Março 2009) .....	121
15	NOTA TÉCNICA nº 47 - Consideração de penalidades ao não atendimento da restrição de vazão mínima obrigatória (Junho 2009).....	123
15.1	Introdução .....	123
15.2	Modelagem .....	124
15.2.1	Construção dos Cortes .....	126
15.3	Flexibilização da restrição de vazão mínima obrigatória.....	128
16	NOTA TÉCNICA nº 55 - PROCEDIMENTO DE SOLUÇÃO DA BIBLIOTECA DE OTIMIZAÇÃO (Maio 2009).....	131
17	NOTA TÉCNICA nº 56 - Execução do Módulo de Geração de Cenários Sintéticos de Energia Natural Afluente num Ambiente de Processamento Distribuído (Setembro 2009).....	133
17.1	Introdução .....	133
17.2	Módulo de Geração de Cenários de ENA.....	134
17.3	Aplicação da Técnica de Paralelização .....	135
17.4	Referências .....	137



# 1 NOTA TÉCNICA nº 29 - Possibilidade de adoção de simulação forward inicial para determinação dos estados de armazenamento para a primeira recursão backward no processo iterativo do modelo NEWAVE (Abril 2007)

## 1.1 Introdução

No NEWAVE, a programação dinâmica dual estocástica (PDDE) é adotada na solução do problema de alocação ótima de recursos hidrotermicos para o atendimento ao mercado de energia. Esse processo consiste em determinar uma estratégia de operação em que a cada estágio do período de estudo, dado um estado inicial, sejam fornecidas metas de geração para cada subsistema a fim de minimizar o valor esperado do custo total de operação ao longo do horizonte de estudo. Essa estratégia de operação, representada pela função de custo futuro (FCF), é aproximada através de uma função linear por partes, construída iterativamente pelos cortes de Benders. A cada estágio e para cada estado do sistema (nível de armazenamento inicial e afluições nos meses anteriores), o problema de operação hidrotérmica é modelado como um problema de programação linear e as variáveis duais associadas à solução desse problema são utilizadas para a construção dos cortes de Benders. Esse processo recursivo é chamado de *backward*. Para se obter uma estimativa do valor esperado do custo de operação ao longo do horizonte de estudo, considerando-se a FCF construída no processo *backward*, processa-se a simulação da operação do sistema para os diversos cenários de afluições. Esse processo é chamado *forward*.

Na primeira iteração, inicia-se o processo por uma recursão *backward*, em que a seleção de estados de energia armazenada é pré-definida para todos os estágios e para todos os cenários de energia natural afluyente.

Nas demais iterações, os estados de armazenamento inicial a serem utilizados na *backward* são aqueles oriundos da simulação *forward* imediatamente anterior.

Com o objetivo de avaliar o impacto da discretização pré-definida de estados de armazenamento na convergência do processo, foi criada uma versão do programa NEWAVE onde o processo se inicia por uma simulação *forward*, de forma a



obter os estados de armazenamento a serem utilizados na primeira iteração. Essa versão foi baseada versão 12a e chamada de versão 12afrw.

## **1.2 Implementação**

O procedimento de cálculo da função de custo futuro do programa NEWAVE consiste, de forma simplificada, em um processo iterativo iniciando-se por uma simulação recursiva *backward*, seguido de uma simulação *forward*. Esse processo iterativo é interrompido quando o número máximo de iterações é alcançado ou o critério de convergência do processo é satisfeito. Uma recursão *backward* seguida de uma simulação *forward* definem uma iteração. Todas as iterações são numeradas sequencialmente, iniciando-se na iteração 1 e incrementadas de forma unitária. A trajetória de convergência é impressa tanto no relatório de convergência quanto no console do computador.

Na versão 12afrw o processo iterativo inicia-se com uma simulação *forward*, sendo essa primeira iteração denominada de iteração 0 (zero). Nessa iteração, nenhum corte de Benders foi estimado, o que significa que a função de custo futuro é igual a zero para qualquer estado do problema. Essa iteração é impressa tanto no relatório de convergência quanto no console do computador.

As energias armazenadas oriundas dessa simulação serão utilizadas como armazenamento inicial para a simulação *backward* da iteração 1.

## **1.3 Entrada de dados**

Não há alterações nos dados de entrada do programa devido à implementação contida nessa Nota Técnica.

## **2 NOTA TÉCNICA nº 30 - Simulação final utilizando uma única série sintética com comprimento de 10.000 anos (Março 2007)**

### **2.1 Introdução**

O modelo NEWAVE define para cada mês do horizonte de estudo, que pode variar até 45 anos (30 anos de planejamento + 10 anos de estático inicial + 5 anos de estático final), 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 estudo. O programa NEWAVE apresenta quatro módulos básicos:

1. Módulo de cálculo dos subsistemas equivalentes de energia.
2. Módulo de geração de séries sintéticas.
3. Módulo de cálculo da política de operação hidrotérmica mais econômica para os subsistemas equivalentes.
4. Módulo de simulação da operação do sistema.

A simulação da operação é realizada para diversos cenários de seqüências hidrológicas distintas, utilizando a função de custo futuro calculada previamente e adotando o horizonte de estudo informado pelo usuário no arquivo de dados gerais (DGER.DAT). Esse horizonte de estudo, em geral, é o mesmo adotado durante o cálculo da política de operação<sup>1</sup>.

A funcionalidade descrita nesta Nota Técnica consiste em realizar a simulação da operação para uma única série sintética de energia afluyente com comprimento de 10.000 anos para o horizonte de planejamento.

Essa funcionalidade está implementada na versão 12\_SSU, que foi baseada na versão 12 do programa NEWAVE. Nessa versão não é permitido:

- ir direto para simulação final usando uma política de operação previamente calculada;
- simular a operação com séries de energias afluentes históricas;

---

<sup>1</sup> No módulo da simulação da operação pode-se considerar um período estático final diferente daquele utilizado no módulo de cálculo da política de operação.

- utilizar período estático pós na simulação final.

Além disso, no programa auxiliar NWLISTOP é permitido apenas utilizar a opção “Tabelas” (opção 2).

## **2.2 Modelagem**

Para a implementação dessa funcionalidade foram necessárias alterações nos módulos de geração de séries sintéticas, no módulo de simulação da operação e na capacidade do programa. O número máximo permitido de subsistemas reais passou de 10 para 4 e o número máximo de subsistemas fictícios passou de 4 para 1.

### **2.2.1 Módulo de geração de séries sintéticas**

Nesse módulo são realizadas a estimativa dos parâmetros do modelo estocástico de energias afluentes e a geração das séries sintéticas de energias afluentes que são utilizadas no cálculo da política de operação hidrotérmica e na simulação da operação<sup>2</sup>.

O procedimento de identificação e estimação dos parâmetros do modelo estocástico não foi modificado. Para todo período do horizonte de estudo informado pelo usuário é ajustado um modelo autoregressivo periódico.

O processo de geração das séries sintéticas de energias afluentes que são utilizadas no cálculo da política de operação hidrotérmica permanece inalterado.

Para a simulação da operação é gerada apenas uma única série sintética de energia afluente com 10.000 anos de horizonte de planejamento e mais 10 anos se houver período estático inicial. Os modelos estocásticos utilizados para a geração da série, durante o horizonte de planejamento, se repetirão a cada ciclo de dimensão igual ao número de anos do horizonte de planejamento originalmente fornecido pelo usuário no arquivo de dados gerais.

O formato de gravação do arquivo temporário, que contém a série sintética gerada para a simulação da operação, foi alterado de forma a reduzir o espaço alocado no disco rígido.

---

<sup>2</sup> As séries de energias afluentes sintéticas empregadas no módulo do cálculo da política de operação e no módulo da simulação final da operação são distintas.

## **2.2.2 Módulo de simulação da operação do sistema**

Nesse módulo é realizada a simulação da operação do sistema ao longo do horizonte de estudo e são calculados índices de desempenho, tais como a média dos custos de operação, dos custos marginais de operação, o risco de déficit e os valores médios de intercâmbio de energia.

Esse módulo foi alterado para permitir a simulação de apenas uma série sintética de energia natural afluyente com horizonte de planejamento de 10.000 e mais 10 anos se houver período estático inicial. Para tanto, foi elaborada uma nova rotina de cálculo da simulação da operação.

Durante a simulação do novo período de planejamento (10.000 anos), as funções de custo futuro e demais valores que são indexados por período serão repetidos a cada ciclo de dimensão igual ao número de anos do horizonte de planejamento fornecido pelo usuário no arquivo de dados gerais. Não há modificação na indexação das FCFs e demais variáveis durante o período estático inicial.

Vale a pena lembrar que nessa versão não é permitido rodar a simulação da operação com período estático final.

## **2.3 Entrada de dados**

Para implementação dessa nova funcionalidade não foi necessário alterar os dados de entrada do programa NEWAVE. O número de séries e o número de anos do período de planejamento a serem considerados na simulação da operação são parâmetros fixos iguais a 1 série e 10.000 anos, respectivamente.

Para adequar o programa auxiliar NWLISTOP (opção 2: Tabelas) à nova funcionalidade, foi alterado o formato de impressão das tabelas de saída.

## **2.4 Rebatimento nos demais modelos da cadeia**

A versão 6.5 do programa SUIISHI-O e a versão 13 do programa DECOMP não se encontram compatível com essa nova funcionalidade pois o número máximo de subsistemas do programa NEWAVE foi alterado. Esse parâmetro é usado na especificação do formato do arquivo que armazena a função de custo futuro.

### **3 NOTA TÉCNICA nº 31 - Variação temporal da restrição de vazão mínima obrigatória e consideração de penalidades ao não atendimento dessa restrição (Março 2007)**

#### **3.1 Introdução**

A vazão mínima de um aproveitamento é uma vazão defluente obrigatória, turbinada ou vertida, que tem por finalidade manter um nível mínimo de um rio para fins de navegação, abastecimento, controle de poluição, irrigação e lazer, entre outros. No NEWAVE, onde o sistema hidrotérmico tem uma representação agregada, é calculada uma energia de vazão mínima correspondente a essa defluência obrigatória. Essa energia é considerada no problema de despacho hidrotérmico como uma geração hidráulica obrigatória.

Na versão 12.19 do NEWAVE, descrita por essa nota técnica, foi implementada a possibilidade do usuário informar o valor dessa restrição ao programa, por usina e em  $m^3/s$ , para cada mês do horizonte de planejamento, além dos meses que compõem o período estático inicial e o período estático final. Anteriormente, não era possível a variação temporal dessa restrição.

Uma outra funcionalidade descrita nessa nota técnica é a possibilidade de consideração de uma penalidade associada ao não atendimento dos requisitos de energia de vazão mínima obrigatória. Essa penalidade é fornecida pelo usuário através de um arquivo específico para tal finalidade (e.g. PENALID.DAT). Como consequência, foram incluídas nos problemas de otimização restrições que representam o atendimento à meta de energia de vazão mínima obrigatória e variáveis de folga, representando o não atendimento desse requisito. Além disso, adaptações no cálculo dos cortes de Benders fizeram-se necessárias, uma vez que essa meta é uma função da energia armazenada inicial, que é um estado do sistema. Essa nova implementação permite a consideração dessa restrição no cálculo da política ótima de operação. Nas versões anteriores, a energia de vazão mínima era considerada como sendo controlável durante o cálculo da política de operação.

## 3.2 Modelagem

### 3.2.1 Variação temporal da energia de vazão mínima

A energia de vazão mínima independe da série hidrológica considerada, dependendo tão somente da configuração do sistema. Seu valor máximo é calculado multiplicando-se a defluência mínima obrigatória de cada usina com reservatório pela soma da produtividade associada à altura queda líquida máxima, e as de todas as usinas fio d'água existentes entre o reservatório e o próximo reservatório a jusante. Desta forma, o valor de energia de vazão mínima, no mês  $k$ , é dada por:

$$EVMIN_{s,t} = c \sum_{i \in R} Q_{\min,i,t} (\rho_i H_i + \sum_{j \in F} \rho_j heq_j) \quad (1)$$

onde  $EVMIN_{s,t}$  é a energia de vazão mínima do subsistema  $s$  no mês  $t$  considerando todos os reservatórios na altura equivalente,  $Q_{\min,i,t}$  é a defluência mínima obrigatória do reservatório  $i$  no mês  $t$ ,  $c$  é uma constante de conversão de unidades,  $\rho_i$  é a produtividade específica da usina  $i$ ,  $H_i$  é a altura de queda equivalente para um reservatório  $i$ ,  $heq_j$  é a altura de queda das usinas a fio d'água,  $R$  é o conjunto dos reservatórios do subsistema  $s$  na configuração  $w$ ,  $F$  é o conjunto de usinas a fio d'água do subsistema  $s$  na configuração  $w$ .

Os valores máximos, médios e mínimos da energia de vazão mínima, respectivamente  $EVMIN_{max}$ ,  $EVMIN_{med}$  e  $EVMIN_{min}$ , são obtidos substituindo-se a altura de queda líquida equivalente na equação (1) pelas alturas de queda correspondentes a um armazenamento máximo ( $H_{max,i}$ ), de metade do volume útil ( $H_{med,i}$ ) e ao nível mínimo operativo ( $H_{min,i}$ ).

A partir destes três pontos,  $(0, EVMIN_{min})$ ,  $(EA_{med}, EVMIN_{med})$  e  $(EA_{max}, EVMIN_{max})$ , ajusta-se um polinômio de segundo grau, a partir do qual obtém-se a meta de energia de vazão mínima,  $M_{EVMIN}$ , em função da energia armazenada no mês.

Vale a pena ressaltar que o usuário pode informar o valor da vazão mínima por usina e para cada mês do horizonte de planejamento, além dos meses que compõem o período estático inicial e o período estático final.

### 3.2.2 Penalidade pelo não atendimento à restrição de energia de vazão mínima obrigatória

A outra funcionalidade implementada na versão 12.19 do NEWAVE caracteriza-se pela criação de uma penalização à violação dos requisitos de atendimento à vazão mínima obrigatória, através da inclusão de uma variável de folga nas restrições de atendimento à demanda e balanço hídrico, penalizada na função objetivo, de forma a evitar que diferentes condições iniciais possam levar a um mesmo custo de operação, conforme colocado anteriormente. Além disso, a penalização na função objetivo permite que a função de custo futuro possa vislumbrar um possível não atendimento à vazão mínima obrigatória no futuro, de forma que a violação à descarga mínima obrigatória possa ser minimizada.

#### 3.2.2.1 Formulação do Problema de Despacho de Operação

Para um dado estado em um instante  $t$ , a meta de atendimento à energia de vazão mínima obrigatória,  $M_{EVMIN}(EA_t)$ , que é função da energia armazenada, pode ser expressa como a soma da energia efetivamente atendida ( $evmin$ ) mais uma folga ( $\delta_{EVMIN}$ ), representando a parcela da meta que não foi atendida. Isso é expresso pela equação (2).

$$M_{EVMIN}(EA_t) = evmin + \delta_{EVMIN} \quad (2)$$

A seguir serão apresentadas as restrições do problema de despacho hidrotérmico do NEWAVE afetadas por essa implementação, além da função objetivo do problema. Entretanto, esse problema é apresentado aqui de forma simplificada, omitindo-se diversas restrições e variáveis, visando facilitar o entendimento da nova implementação.

#### Função objetivo:

A função objetivo do problema, passa a considerar a parcela de vazão mínima eventualmente não atendida,  $\delta_{EVMIN}$ , penalizada por  $\psi_{EVMIN}$ . O valor de  $\psi_{EVMIN}$  é definido pelo usuário. Logo, a função objetivo do problema passa a ser:

$$z_t = \min \left[ \sum^{NSIS} \left( \sum^{NPMC} \left[ \sum^{TCLISIS} \psi_T \cdot g_T + \sum^{NPDF} \psi_D \cdot def \right] + \psi_{EVMIN} \cdot \delta_{EVMIN} \right) + \frac{1}{1+\beta} \alpha_{t+1} \right] \quad (1)$$

onde  $NSIS$  é o número de subsistemas,  $NPMC$  é o número de patamares de mercado,  $TCLISIS$  é o número de classes térmicas de um determinado subsistema,  $NPDF$  é o número de patamares de déficit,  $g_T$  é a geração térmica de uma determinada classe,  $def$  é o déficit em um subsistema,  $\beta$  é a taxa que traz o custo futuro a valor presente e  $\alpha_{t+1}$  é o valor do custo futuro.

### Equações de balanço hídrico – EBH:

Conforme dito acima, uma folga para o atendimento de vazão mínima obrigatória,  $\delta_{EVMIN}$ , deve ser incluída na equação de balanço hídrico. Considerando essa folga, tem-se, de uma forma simplificada:

$$ea_{t+1} + \sum^{NPMC} g_H + evert - \delta_{EVMIN} = EA_t + EC_t - M_{EVMIN}(EA_t) - EVAP(EA_t) - EMORTO \quad (3)$$

onde  $ea_{t+1}$  é a energia armazenada no final do mês  $t$ ,  $g_H$  é a geração hidráulica controlável,  $evert$  é o vertimento,  $EA_t$  é a energia armazenada no início do mês  $t$ ,  $EC_t$  é a energia controlável,  $EVAP(EA_t)$  é a evaporação e  $EMORTO$  é a energia de enchimento de volume morto.

Substituindo (2) em (3):

$$ea_{t+1} + \sum^{NPMC} g_H + evert + evmin = EA_t + EC_t - EVAP(EA_t) - EMORTO \quad (4)$$

### Equações de atendimento à carga própria – EAD:

Assim como nas equações de balanço hídrico, nas de atendimento à demanda também será incluída a folga para energia de vazão mínima. Deve-se observar porém que, uma vez que o estudo pode considerar o mercado discretizado em mais de um patamar de carga, haverá, para cada subsistema, uma restrição de atendimento à demanda para cada patamar. Nesse caso, para cada restrição, tanto a meta de vazão mínima quanto a folga devem ser multiplicadas pelo número de horas do patamar, expresso na variável  $FPENG$ . Considerando a variável de folga, tem-se:



$$g_H + \sum^{TCLISIS} g_T + \sum^{NPDF} def + \sum i_{recebido} - \sum i_{exportado} - exc - FPENG \cdot \delta_{EVMIN} = \quad (5)$$

$$[DEMLIQ - EFIO - M_{EVMIN}(EA_t)] \cdot FPENG$$

Substituindo (2) em (5) chega-se a:

$$g_H + \sum^{TCLISIS} gt + \sum^{NPDF} def + \sum i_{recebido} - \sum i_{exportado} - exc + FPENG \cdot evmin = \quad (6)$$

$$= [DEMLIQ - EFIO] \cdot FPENG$$

### Restrições de geração hidráulica máxima controlável – EEH:

A restrição de geração hidráulica máxima tem por objetivo limitar a geração hidráulica controlável ( $g_H$ ). A energia máxima que pode ser gerada pelos reservatórios é dada pela folga entre a geração hidráulica máxima e a geração de usinas a fio d'água mais energia de vazão mínima. Pela proposta do trabalho, a variável de folga para a energia de vazão mínima também deve ser incluída nesta restrição. Considerando  $\delta_{EVMIN}$ , tem-se:

$$g_H - \delta_{EVMIN} \leq [\overline{GH}(EA_t) - EFIO - M_{EVMIN}(EA_t)] \cdot FPENG \quad (7)$$

Substituindo (2) em (7):

$$g_H + FPENG \cdot evmin \leq [\overline{GH}(EA_t) - EFIO] \cdot FPENG \quad (8)$$

### 3.2.2.2 Construção dos Cortes

A introdução de uma nova restrição e novas variáveis ao problema de planejamento de operação causa alterações no cálculo dos cortes de Benders. A restrição (2), que representa a meta de geração de vazão mínima, tem a variável  $M_{EVMIN}$ , que é função do armazenamento inicial. Por outro lado, tal variável não figura mais nas restrições de balanço hídrico, atendimento a demanda e geração hidráulica máxima.

O coeficiente do corte de Benders associado ao estado armazenamento no início do período  $t$  é dado pela equação (9).

$$\begin{aligned} \pi_v = & \sum^{NSIS} \eta_{BH} \frac{\partial}{\partial EA_t} EBH + \sum^{NPMC} \sum^{NSIS} \eta_{AD} \frac{\partial}{\partial EA_t} EAD + \sum^{NCOR} \eta_{COR} \frac{\partial}{\partial EA_t} ECOR + \\ & + \sum^{NPMC} \sum^{NSIS} \eta_{EH} \frac{\partial}{\partial EA_t} EEH + \sum^{NSIS} \eta_{EV} \frac{\partial}{\partial EA_t} EMAVM \end{aligned} \quad (9)$$

Sejam as equações a seguir que representam as parábolas do fator de correção da energia controlável, da meta de energia de vazão mínima, energia evaporada e geração hidráulica máxima:

$$FC(EA_t) = a_{EC} EA_t^2 + b_{EC} EA_t + c_{EC} \quad (10)$$

$$M_{EVMIN}(EA_t) = a_{EZ} EA_t^2 + b_{EZ} EA_t + c_{EZ} \quad (11)$$

$$EVAP(EA_t) = a_{EV} EA_t^2 + b_{EV} EA_t + c_{EV} \quad (12)$$

$$\overline{GH}(EA_t) = a_{EH} EA_t^2 + b_{EH} EA_t + c_{EH} \quad (13)$$

Além disso:

$$EC_t = FC(EA_t) \cdot \gamma \cdot EAF_t \quad (14)$$

onde  $FC(EA_t)$  é o fator de correção de energia controlável,  $\gamma$  é a parcela da energia natural afluyente que pode ser controlada,  $EAF_t$  é a energia natural afluyente.

Substituindo (10) e (12) em (4):

$$\begin{aligned} ea_{t+1} + \sum^{NPMC} g_H + evert + evmin = & EA_t + (a_{EC} EA_t^2 + b_{EC} EA_t + c_{EC}) \cdot \gamma \cdot EAF_t - \\ & - (a_{EV} EA_t^2 + b_{EV} EA_t + c_{EV}) - EMORTO \end{aligned} \quad (15)$$

A derivada de (15) em relação à  $EA_t$  é dada por:

$$\frac{\partial}{\partial EA_t} EBH = 1 + \gamma \cdot EAF_t (2a_{EC} EA_t + b_{EC}) - (2a_{EV} EA_t + b_{EV}) \quad (16)$$

A equação de atendimento à demanda, não possui nenhum termo em função de  $EA_t$ . Logo,

$$\frac{\partial}{\partial EA_t} EAD = 0 \quad (17)$$

Substituindo (13) em (8), a restrição de geração hidráulica máxima pode ser expressa como:

$$g_H + FPENG \cdot evmin \leq \left[ (a_{EH} EA_t^2 + b_{EH} EA_t + c_{EH}) - (1 - \gamma) \cdot EAF_t \right] \cdot FPENG \quad (18)$$

Derivando (18) em relação à  $EA_t$ :

$$\frac{\partial}{\partial EA_t} EEH = (2a_{EH} EA_t + b_{EH}) \cdot FPENG \quad (19)$$

Finalmente, a nova restrição de atendimento aos requisitos de energia de vazão mínima obrigatória:

$$evmin + \delta_{evmin} = a_{EZ} EA_t^2 + b_{EZ} EA_t + c_{EZ} \quad (20)$$

A derivada de (20) em relação ao armazenamento inicial é dada por:

$$\frac{\partial}{\partial EA_t} EMAVM = (2a_{EV} EA_t + b_{EV}) \quad (21)$$

Finalmente, substituindo-se (16), (17), (19) e (21) em (9), é obtida para um subsistema  $i$  o seguinte coeficiente da equação de cortes de Benders:

$$\begin{aligned} \pi_V = & \eta_{BH} \left[ 1 + \gamma \cdot EAF_t (2a_{EC} EA_t + b_{EC}) - (2a_{EV} EA_t + b_{EV}) \right] + \\ & + \eta_{EH} (2a_{EH} EA_t + b_{EH}) + \\ & + \eta_{EV} (2a_{EV} EA_t + b_{EV}) \end{aligned} \quad (IV.2)$$

Quanto à derivada do problema em relação às energias afluentes passadas, não há qualquer modificação, uma vez que a meta de geração de vazão não é função desses estados.

### 3.3 Entrada de dados

Para implementação dessa nova funcionalidade foi incluída uma nova palavra-chave no arquivo de alterações cadastrais que permite ao usuário informar, sempre que necessário, o novo valor da restrição de vazão mínima obrigatória da usina. Essa palavra chave é VAZMINT.

Além disso, no arquivo destinado às penalidades associadas a violação de restrições, foi adicionada a palavra chave EVMIN, a partir de onde o usuário pode informar qual é o valor associado à violação da restrição de energia de vazão mínima obrigatória., em u.m./MWh.

### **3.4 Rebatimento nos demais modelos da cadeia**

A versão 6.5 do programa SUISHI-O não se encontra compatível com esta nova funcionalidade do NEWAVE, posto que não permite a variação temporal da restrição de energia de vazão mínima obrigatória, além de não penalizar na função objetivo do problema de otimização a violação desse requisito.

Na versão 13 do programa DECOMP é permitida a variação temporal da restrição de energia de vazão mínima obrigatória. A violação dessa restrição é penalizada na função objetivo do problema de otimização porém, a penalidade utilizada não é definida pelo usuário. Este valor é um parâmetro fixo dentro do programa.

## 4 NOTA TÉCNICA nº 34 - Refinamento da Discretização de Estados para a Primeira Recursão do Processo Iterativo de Cálculo da Função de Custo Futuro, com Fixação de Subconjunto de Estados a Serem Visitados a Cada Recursão (Abril 2007)

### 4.1 Introdução

O cálculo da função de custo futuro (FCF) do modelo NEWAVE é realizado através de um algoritmo de programação dinâmica dual estocástica (PDDE). As variáveis de estado utilizadas no problema de planejamento da operação de médio prazo são: tendência hidrológica (representada pela energia natural afluyente aos subsistemas nos meses passados), sorteada aleatoriamente, e nível de armazenamento dos subsistemas no início do mês.

O processo iterativo de cálculo desta função é iniciado com uma recursão *backward*. Nesta primeira recursão, é adotada uma discretização inicial para a variável de estado nível de armazenamento, que é dividida em faixas de igual percentual da energia armazenável máxima para todos os subsistemas e para todos os estágios. Normalmente, este tipo de discretização inicial é chamado de discretização em paralelo.

Denomina-se como estado, para o modelo NEWAVE, a associação entre valores de nível de armazenamento e valores de tendências hidrológicas em um determinado mês.

Na primeira iteração da PDDE, e na primeira associação dessas variáveis de estado o valor do nível de armazenamento é:

$$EAR_{inicial\_1^a\_comb.} = \max \left( EAR_{Mn\_oper}, \left( \frac{1}{n_{sim}} \right) * EAR_{Mx} \right) \quad (1)$$

onde  $EAR_{Mx}$  é a energia armazenável máxima,  $EAR_{Mn\_oper}$  é a restrição de energia armazenada mínima e  $n_{sim}$  é o número de estados associação de tendência hidrológica e nível de armazenamento.

Portanto, se o número de estados associação de tendência hidrológica e nível de armazenamento é igual a 200, o valor do nível de armazenamento correspondente à primeira associação das variáveis de estado é 0.5% da energia armazenável máxima, caso a restrição de energia armazenada mínima, correspondente ao volume mínimo operativo (VMINT) das usinas deste subsistema, seja inferior a este valor.

Assim, generalizando:

$$EAR_{inicial\_n - \acute{e}sima\_comb.} = \max \left( EARM_{n\_oper}, \left( \frac{n}{nsim} \right) * EARM_x \right) \quad (2)$$

Estes valores são utilizados na primeira recursão *backward* do algoritmo de PDDE.

A associação para cada estado dos valores das variáveis de estado nível de armazenamento e tendência hidrológica é feita de maneira aleatória.

Os valores que representam a variável de estado tendência hidrológica não se alteram de iteração para iteração do algoritmo de programação dinâmica dual estocástica. Já os valores da variável de estado nível de armazenamento são atualizados a cada iteração do algoritmo PDDE, sendo que os valores a serem considerados na próxima recursão *backward* são aqueles obtidos ou visitados na simulação *forward* prévia. Esta metodologia proporciona que um número maior de níveis de armazenamento seja visitado.

Até o presente momento, todas as versões do programa NEWAVE adotam o procedimento descrito acima para definir o conjunto de estados que serão visitados durante a primeira recursão *backward* do processo iterativo de cálculo da FCF.

Estudos recentes realizados pelo ONS<sup>3</sup> e pelo CEPEL<sup>4</sup> levantaram a necessidade de possibilitar um refinamento da discretização da variável de estado nível de armazenamento e também, uma associação entre estados de tendência

---

<sup>3</sup> Nota Técnica ONS nº 113/2006

<sup>4</sup> Penna, D.D.J.; Maceira, M.E.P.; Duarte, V.S.; Francisco, R.N.; 2006, “Aperfeiçoamento da Política de Operação no Planejamento do Sistema Interligado Brasileiro Contemplando Sistemas de Pequeno Porte”, X Simpósio de Especialistas em Planejamento da Operação e Expansão Elétrica, Florianópolis, Brasil.

hidrológica, selecionadas segundo um critério de severidade, e estados de níveis de armazenamento previamente selecionados pelo usuário.

A primeira abordagem foi a adoção no programa NEWAVE de refinamento da discretização da variável de estado nível de armazenamento, bem como uma associação entre tendências hidrológicas, selecionadas segundo um critério de severidade, e níveis de armazenamento previamente selecionados pelo usuário.

Outra abordagem, proposta pelo ONS, foi a fixação dos estados para uma faixa de armazenamento definida pelo usuário, para todas as recursões *backward* do algoritmo de PDDE<sup>5</sup>.

Os resultados indicam que as metodologias abordadas em ambas as implementações podem ser aperfeiçoadas e combinadas, considerando-se:

- o refinamento da discretização da variável de estado nível de armazenamento variando mensalmente;
- os estados resultantes da associação entre tendências hidrológicas, selecionadas segundo um critério de severidade, e níveis de armazenamento previamente selecionados pelo usuário;
- a fixação de um subconjunto destes estados para todas as iterações do algoritmo de PDDE.

Esta Nota Técnica tem por objetivo descrever os procedimentos utilizados para definir o conjunto de estados que serão visitados durante a primeira recursão *backward* e para fixar um subconjunto de associações de variáveis de estados que serão visitados em todas as iterações do algoritmo de PDDE.

As funcionalidades descritas nesta Nota Técnica foram implementadas a partir das versões 12c e 12.11k do programa NEWAVE, gerando desta forma, as versões 12c\_RFX e 12.11k\_RFX.

---

<sup>5</sup> Nota Técnica CEPEL nº 19/2006

## 4.2 Aperfeiçoamento da discretização inicial de estados para o cálculo da FCF

### 4.2.1 Refinamento da discretização inicial dos valores da variável de estado níveis de armazenamento

Foi implementada a possibilidade de flexibilização da discretização inicial dos estados de armazenamento do subsistema escolhido, variável por mês, para que uma faixa de armazenamento possa ser mais bem detalhada, conforme indicado na Figura 1. Esta faixa é definida pelo usuário e nesta Nota Técnica será chamada de faixa intermediária.

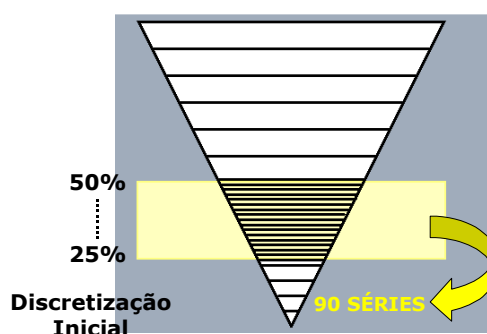


Figura 1 – Discretização Inicial

Os demais subsistemas permanecem adotando a discretização inicial em paralelo, descrita anteriormente.

Também é possível alterar o detalhamento da faixa de armazenamento inferior àquela selecionada pelo usuário (faixa intermediária). Já a discretização da faixa de armazenamento superior à faixa intermediária permanece adotando a discretização em faixas de igual percentual de energia armazenável máxima, calculada a partir dos parâmetros fornecidos pelo usuário para as faixas intermediária e inferior.

Deve-se dar atenção especial para os níveis de armazenamento inferiores e intermediários na construção da função de custo futuro, porque são níveis para os quais há possibilidade de ocorrência de déficits futuros de energia. A FCF definirá, principalmente para estes estados, a estratégia de operação que minimiza em média a ocorrência de déficits futuros de energia. Os níveis de armazenamento superiores podem estar associados a uma representação menos detalhada, pois a possibilidade de vertimento é crescente e a possibilidade de déficit é muito pequena ou inexistente.



Logo, se a faixa intermediária for escolhida para ter uma representação mais refinada, então a faixa superior é que deve ter uma representação menos detalhada.

Vale a pena lembrar que nesta implementação não será alterado o número de estados original (nsim).

Por exemplo, suponha que:

- os limites da faixa intermediária de discretização dos valores da variável de estado nível de armazenamento são 25% e 50% de EARMx, com associação à 100 estados;
- à faixa inferior foram associados 50 estados; e
- o número de estados é 200 (nsim=200).

A faixa inferior de discretização dos valores da variável de estado nível de armazenamento irá de 0% a 25% de EARMx, com 50 estados associados, que serão discretizados em faixas de 0,5% de EARMx.

A faixa intermediária de discretização dos valores da variável de estado nível de armazenamento irá de 25% a 50% de EARMx, com 100 estados associados, que serão discretizados em faixas de 0,25% de EARMx.

Finalmente, a faixa superior de discretização dos valores da variável de estado nível de armazenamento irá de 50% a 100% de EARMx, com 50 estados associados, que serão discretizados em faixas de 1% de EARMx.

#### 4.2.2 Classificação dos cenários de afluências

Os cenários de afluências da simulação *forward* podem ser classificados segundo algum critério de severidade especificado.

Nesta implementação foi adotado como critério de severidade o somatório, para cada tendência hidrológica, das energias naturais afluentes (ENAs) padronizadas, para cada mês de um período definido pelo usuário, para todos os anos do período de estudo, para o subsistema escolhido.

$$ENAp_{i,m} = \frac{ENA_{i,m} - \mu_m}{\sigma_m} \quad (3)$$

onde  $ENAp_{i,m}$  é a energia natural afluente padronizada correspondente ao período m do i-ésimo cenário de afluências,  $ENA_{i,m}$  é a energia natural afluente do período m do

i-ésimo cenário de afluições,  $\mu_m$  e  $\sigma_m$  são a média e o desvio padrão históricos do mês associado ao período  $m$ , respectivamente.

A adoção das ENAs padronizadas permitirá a comparação de ENAs de diferentes subsistemas. Quando aplicado a apenas um subsistema, a adoção da ENA padronizada permite a associação com a modelagem do modelo de geração de cenários – modelo PARp.

### 4.2.3 Associação de estados

Nesta implementação é permitido ao usuário definir associações entre estados de armazenamento especificados no item 2.1 e tendências hidrológicas, classificados no item 2.2, a fim de compor os estados para cálculo da FCF quanto ao percentual de armazenamento e à severidade dos cenários de afluições.

Para tanto são mantidas as tendências hidrológicas fixas e através de dados de entrada, o usuário pode associar o número da tendência hidrológica ao número do estado de armazenamento, que foi definido de acordo com item 2.1.

A associação de estados para a primeira iteração pode ser obtida pela permutação do estado de armazenamento de todos os subsistemas ou somente do subsistema selecionado no item 2.1. Esta escolha é feita pelo usuário através de uma opção no arquivo de entrada de dados da nova funcionalidade.

As tendências hidrológicas selecionadas para associação não poderão ser selecionadas para uma nova associação.

É importante ressaltar que:

- o estado de armazenamento inicial corresponde a um processo de decisão de operação;
- as afluições correspondem a um processo hidrológico e;
- a associação de estados de armazenamento escolhidos e das afluições selecionadas segundo um critério de severidade, como o indicado em 2.2, não altera a equiprobabilidade da geração sorteio de cenários de tendências hidrológicas.

Para facilitar o procedimento de associação de estados será impresso no relatório de convergência do programa NEWAVE um conjunto de tabelas, uma para cada subsistema, com informações sobre os estados de armazenamento inicial e de tendência hidrológica. Um novo conjunto de tabelas será impresso após a definição da

associação. Além do relatório de convergência, as informações pré e pós-associação serão impressas em um arquivo auxiliar do tipo CSV.

### **4.3 Fixação de subconjunto de estados para todas as iterações**

Também será permitido ao usuário selecionar individualmente os estados a serem fixados em todas as recursões *backward*. Os estados serão escolhidos pelo seu número, que corresponde à tendência hidrológica e nível de armazenamento do subsistema escolhido.

O valor da variável de estado nível de armazenamento para o subsistema selecionado poderá ser fixado para alguns meses do ano, válido para todos os anos do horizonte de estudo.

O valor da variável de estado nível de armazenamento será fixado apenas para o subsistema selecionado. Para os demais subsistemas os armazenamentos permanecem aqueles atingidos na última simulação *forward*.

Como a geração de cenários de afluência que irão compor cada tendência hidrológica é feita de forma multivariada, a tendência hidrológica será fixada para todos os subsistemas.

É importante frisar que o número máximo de estados a serem fixados para todas as iterações é igual a 10% do número total de estados (nsim).

### **4.4 Entrada de dados**

Para implementação dessa nova funcionalidade foi criado um novo arquivo de dados (DISCR.DAT), onde serão fornecidos os parâmetros relativos à customização da variável armazenamento no início do mês, à classificação dos cenários de afluência, à associação de estados, à permutação dos estados de armazenamento e à fixação de estados. Um detalhamento maior deste novo arquivo pode ser encontrado no Manual do Usuário das versões 12c\_RFX e 12.11k\_RFX.

Adicionalmente, dois arquivos de entrada tiveram seu formato alterado: Dados Gerais (DGER.DAT) e Lista de arquivos (ARQUIVOS.DAT). No arquivo DGER.DAT poderá optar-se pela consideração ou não da nova funcionalidade. No arquivo ARQUIVOS.DAT, deve ser incluído um novo registro referente ao nome do arquivo com os parâmetros a serem utilizados nessa nova funcionalidade. Esse registro é de existência obrigatória.

#### **4.5 *Rebatimento nos demais modelos da cadeia***

A implementação dessa nova funcionalidade no modelo NEWAVE não tem rebatimento nos demais modelos da cadeia energética.

## **5 NOTA TÉCNICA nº 35 - Consideração de restrição de disponibilidade de combustível para conjuntos de classes térmicas, da possibilidade de utilização de mais de um combustível por classe térmica e da possibilidade de variação mensal do custo unitário das classes térmicas. (Abril 2007)**

### **6 Introdução**

Nesta Nota Técnica é descrita a modelagem da consideração de restrição de disponibilidade de combustível para conjuntos de classes térmicas e da possibilidade de utilização de mais de um combustível por classe térmica, no problema de otimização energética de sistemas hidrotérmicos de médio prazo. Essa nova funcionalidade foi implementada na versão 12.20 do NEWAVE, que é uma continuação da versão 12.19. Nessa versão também há a possibilidade de variação mensal de alguns atributos das termelétricas, por classe térmica e por combustível. Essas características são: relação de geração máxima em relação ao combustível de referência, restrição de geração térmica mínima por combustível, custo unitário e consumo específico. Todos esses dados são informados pelo usuário através de um arquivo de dados específico, conforme será visto no item 3.

Na Figura 1 há uma ilustração de um exemplo prático. Pode-se observar que a usina UTE1 pode gerar tanto com carvão quanto com gás natural. A usina UTE2 pode gerar tanto com gás natural quanto com diesel. A escolha do montante de geração elétrica com cada combustível é feita em função do custo de cada combustível, de suas disponibilidades, do rendimento associado a eles e da operação do resto do sistema, levando-se em conta os custos operativos das demais usinas termelétricas, os custos associados ao não atendimento do mercado (déficit) e também da função de custo futuro.

A partir dessa modelagem, pode-se representar, além das restrições dos combustíveis, expansões de gasodutos, por exemplo.

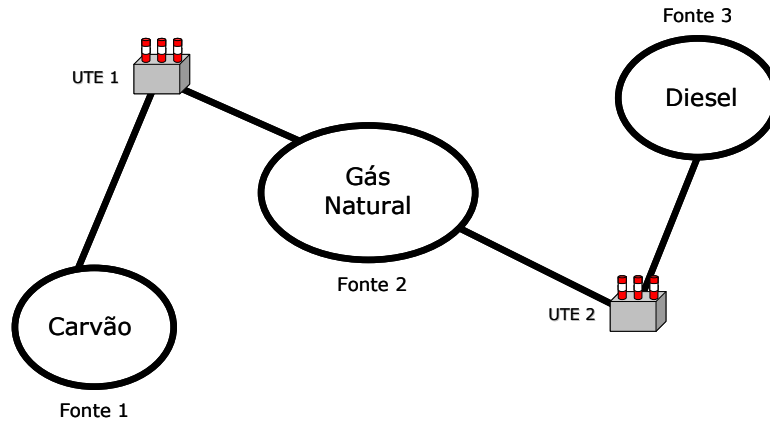


Figura 1 – Exemplo de usinas termelétricas multi-combustível.

## 6.1 Modelagem

A partir da versão 12.20 do NEWAVE, é possível a representação de restrições de limite de combustível e da possibilidade de uma determinada usina termelétrica gerar com mais de um combustível.

O problema de despacho hidrotérmico contemplando a funcionalidade descrita nesta Nota Técnica para um estágio, pode ser representado pelo seguinte problema de otimização simplificado:

$$\min \sum_{def} c_{def} g_{def} + \sum_{iclt} \sum_{icomb} c_{icomb}^{iper} g_{icomb}^{iclt,iper} + penalidades + \frac{1}{1 + txmens} \alpha(\cdot)$$

s.a:

$$\left\{ \begin{array}{l} \text{atendimento à demanda;} \\ \text{balanço hídrico;} \\ \text{função de custo futuro;} \\ g_{icomb}^{iclt,iper} \geq GTMIN_{icomb}^{iclt,iper}, \quad iclt = 1, \dots, nclt, \quad icomb = 1, \dots, numcomb(iclt); \\ \sum_{icomb=1}^{numcomb(iclt)} g_{icomb}^{iclt,iper} \geq GTMIN^{iclt,iper}, \quad iclt = 1, \dots, nclt; \\ \sum_{icomb=1}^{numcomb(iclt)} k_{icomb}^{iclt} g_{icomb}^{iclt,iper} \leq GTMAXREF^{iclt,iper}, \quad iclt = 1, \dots, nclt; \\ \sum_{iclt \in \Omega_{ifnt}} \beta_{icomb}^{iclt,iper} g_{icomb}^{iclt,iper} \leq DISPMAX_{ifnt}^{iper}, \quad ifnt = 1, \dots, nfnt, \end{array} \right.$$

onde:

- *ifnt* representa a fonte de combustível para um determinado conjunto de classes térmicas;

- $\Omega_{ifnt}$  representa o conjunto de classes térmicas associadas à fonte de combustível  $ifnt$ ;
- $C_{def}, G_{def}$  representam o custo unitário e o montante de energia não suprida, respectivamente;
- $c_{icomb}^{iclt, iper}$  representa o custo unitário da classe  $iclt$ , para geração utilizando combustível  $icomb$  no período  $iper$ ;
- $g_{icomb}^{iclt, iper}$  representa a geração térmica da classe  $iclt$  usando o combustível  $icomb$ , no período  $iper$ ;
- $GTMIN_{icomb}^{iclt, iper}$  representa a geração térmica mínima da classe  $iclt$  utilizando o combustível  $icomb$ , no período  $iper$ ;
- $GTMIN^{iclt, iper}$  representa a geração térmica mínima total da classe  $iclt$ , no período  $iper$ ;
- $k_{icomb}^{iclt}$  representa o fator de conversão da geração térmica máxima da classe  $iclt$  com o combustível  $icomb$  em relação ao combustível de referência desta mesma classe;
- $GTMAXREF^{iclt, iper}$  representa a geração térmica máxima total da classe  $iclt$ , no período  $iper$ , em relação ao combustível de referência;
- $\beta_{icomb}^{iclt, iper}$  representa o consumo específico para a classe  $iclt$  do combustível  $comb$ , relativo à fonte  $ifnt$ , no período  $iper$ ;
- $DISPMAX_{ifnt}^{iper}$  representa a disponibilidade máxima, em unidade de volume, da fonte  $ifnt$  no período  $iper$ ;
- *penalidades* agrega o custo pelo não-atendimento à todas as restrições não citadas anteriormente que possuam variáveis de folga associadas.

A consistência de dados deve garantir que a restrição de geração térmica mínima de cada classe térmica seja atendida, considerando-se as restrições de capacidade máxima de fornecimento das fontes de combustíveis e de geração mínima por usina e a geração mínima por combustível, para cada fonte associada a esta classe térmica. Esta consistência deve ser feita para cada um dos meses do período de estudo. Para cada período do horizonte de planejamento é resolvido um problema de otimização, onde deseja-se minimizar o somatório das variáveis de folga,  $s^{ifnt, iper}$ .

Este problema tem solução nula apenas quando as restrições de disponibilidade de combustível não representam inviabilidade das restrições de gerações térmicas mínimas, representando a consistência de dados desejada.

$$\min_s \sum_{ict} s^{ifnt,iper}$$

$$s.a. \left\{ \begin{array}{l} \sum_{ict \in \Omega} \beta_{icomb}^{ict,iper} g_{comb}^{ict,iper} - s^{ifnt,iper} \leq DISPMAX^{ifnt} \\ g_{comb}^{ict,iper} \geq GTMIN_{comb}^{ict,iper} \\ \sum_{icomb} g_{comb}^{ict,iper} \geq GTMIN^{ict,iper} \\ g_{comb}^{ict,iper}, s^{ifnt,iper} \geq 0 \end{array} \right.$$

## 6.2 Entrada de dados

Para contemplar as funcionalidades descritas nessa Nota Técnica, o arquivo de classes térmicas (e.g. clast.dat) foi reformulado. Esse arquivo foi dividido em quatro blocos:

1. Cadastro das fontes térmicas;
2. Definição de classes e associação entre classes e fontes;
3. Restrição de disponibilidade das fontes; e
4. Variação temporal das características das classes/fontes.

Através desses blocos, é possível ao usuário informar os seguintes dados:

- cadastro das fontes de combustíveis e sua disponibilidade;
- definição das classes térmicas e as suas associações às diversas fontes de combustíveis;
- o combustível referencial de cada classe térmica;
- o custo unitário de geração térmica, em R\$/MWh, para cada tipo de combustível utilizado;
- o fator de capacidade máxima para cada combustível, além do referencial, que a classe pode utilizar;
- o consumo específico dos diversos tipos de combustível, em unidade de volume/MWh;
- a restrição de geração térmica mínima relacionada a cada combustível da classe térmica;



- a disponibilidade máxima do combustível de cada fonte, em unidades de volume.

Vale ressaltar que a restrição de geração térmica mínima relacionada a cada combustível da classe térmica é um dado diferente da geração térmica mínima definida para cada usina térmica, que deve continuar sendo fornecida no arquivo de cadastro de usinas térmicas (e.g. term.dat).

Nesse arquivo, também é possível ao usuário definir uma classe térmica que utilize apenas um combustível, sem restrição de capacidade de fornecimento.

A todos os registros que possuem variação temporal, é permitida a declaração explícita para os períodos estáticos inicial e final. Nesse caso é informado o mês e a palavra “PRE” e/ou “POS” no campo relativo ao ano. Caso não sejam definidos estes valores, serão utilizados os valores do primeiro e último períodos do horizonte de planejamento.

### **6.3 *Rebatimento nos demais modelos da cadeia***

As versões 6.5 do programa SUIHI-O e 13 do programa DECOMP não se encontram compatíveis com esta nova funcionalidade.

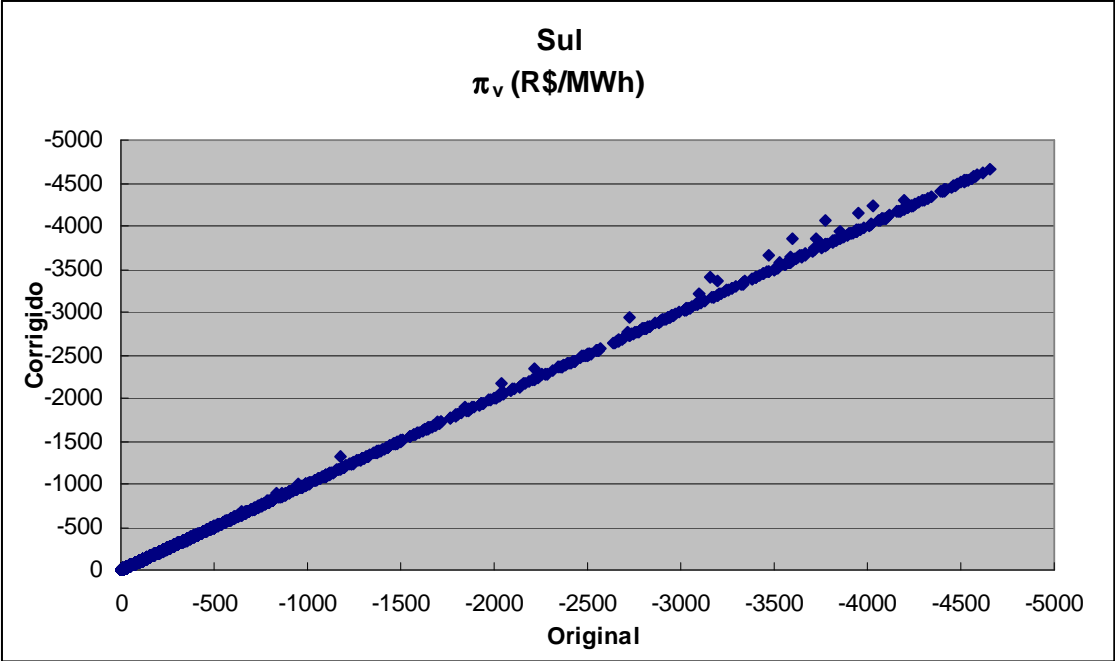
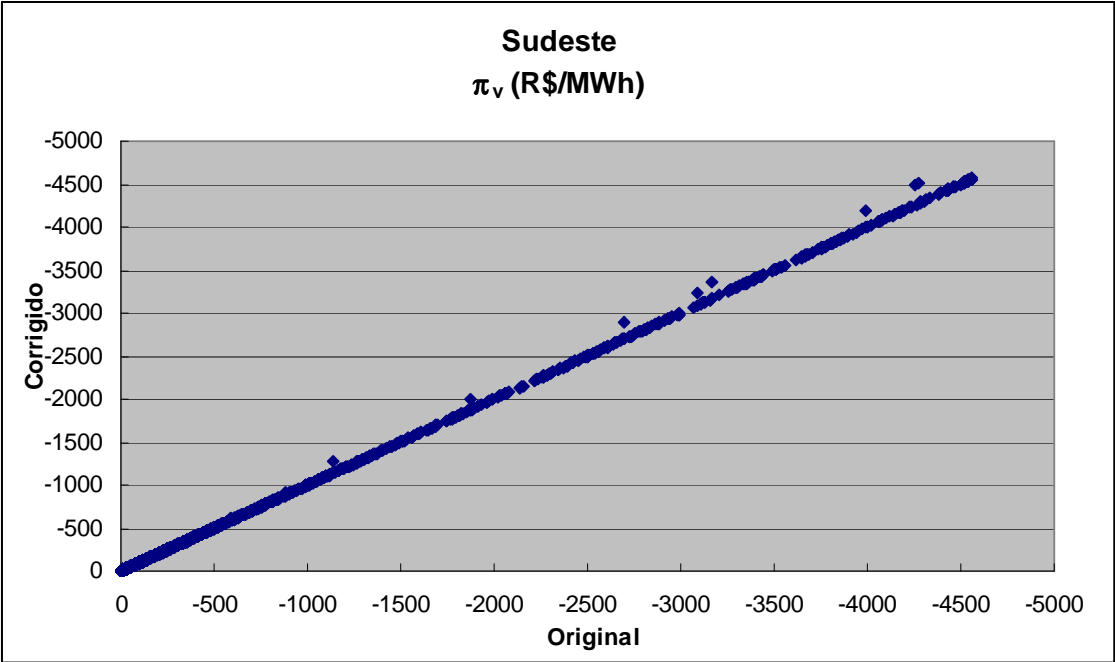
## **7 NOTA TÉCNICA nº 36 - Avaliação do impacto do ajuste na parcela de meta de desvio de água controlável da derivada do corte de Benders na versão 12b do NEWAVE (Setembro 2007)**

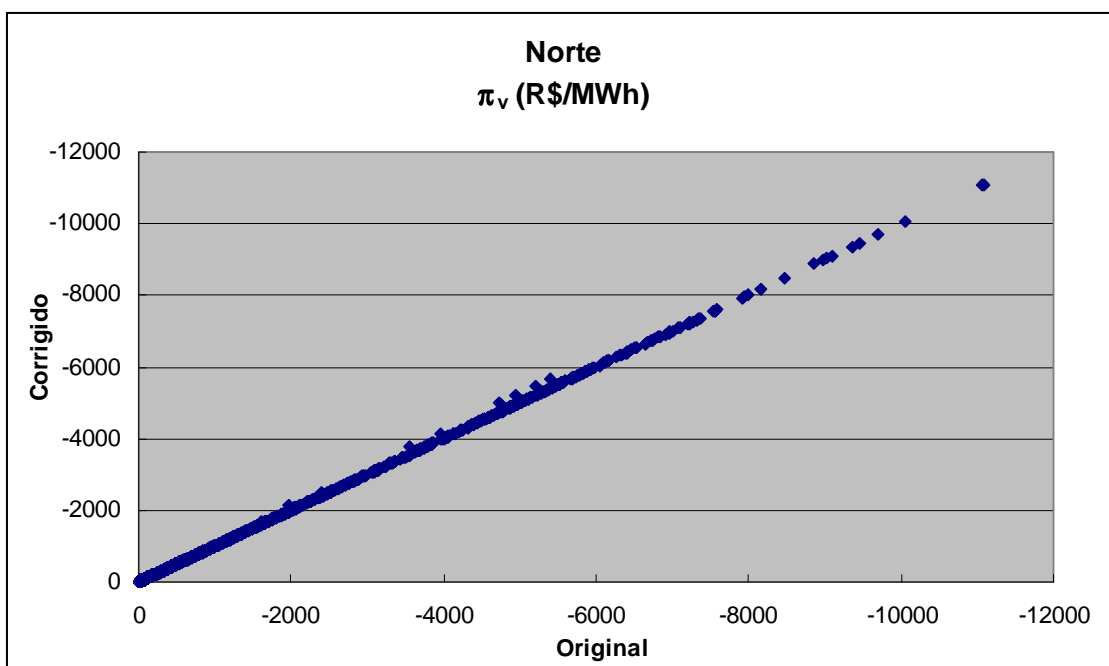
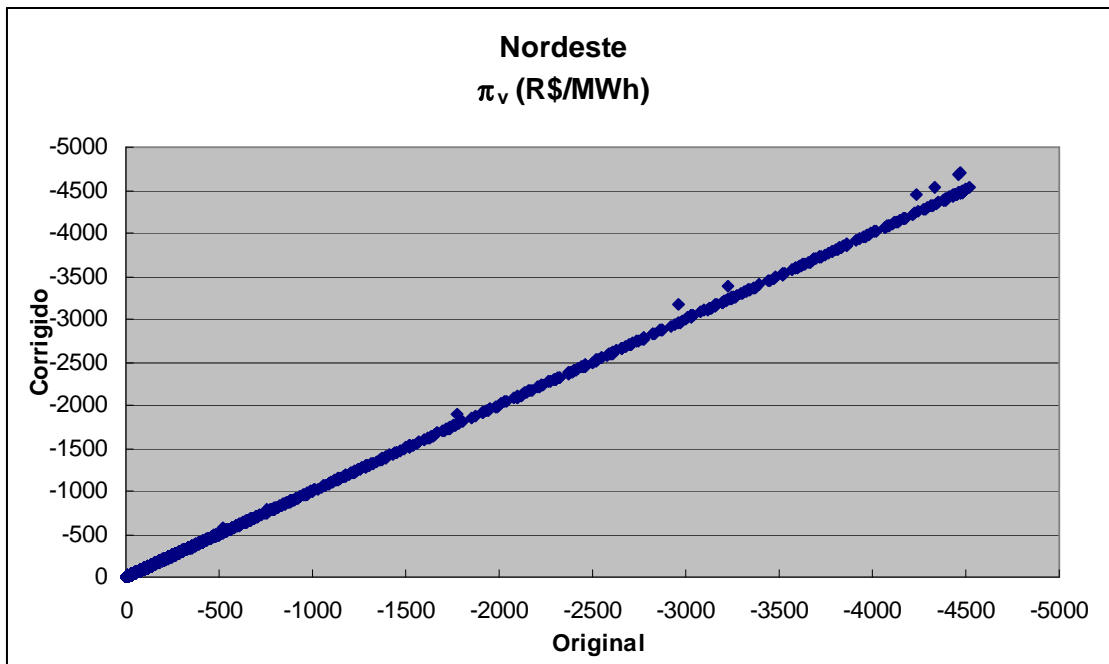
Na versão 12b, enviada em setembro de 2006 ao ONS, o cálculo da derivada que gera o corte de Benders referente à parcela de meta de desvio de água controlável foi ajustado. O impacto deste ajuste é apenas relacionado com o coeficiente relacionado à energia armazenada no início do mês ( $\pi_v$ ).

Como nos casos de PMOs estudados até o momento, o montante de energia de desvio de água por subsistema é sempre pequeno, espera-se que os impactos associados a este ajuste não tenham relevância nos resultados finais.

A título de ilustração, é analisado o impacto deste ajuste nos valores dos coeficientes dos cortes de Benders para o PMO de agosto de 2007. Para isso realizou-se um levantamento dessas diferenças considerando a versão 12f, por meio da comparação dos coeficientes  $\pi_v$  considerando o ajuste e aqueles que seriam obtidos caso o ajuste não fosse feito nesta mesma versão.

Nos gráficos abaixo são comparados 2400 pares de valores (com e sem ajuste), relativos a 12 meses e 200 cenários de afluências forward. Observa-se que, de fato, independentemente do subsistema, este ajuste tem impacto bastante reduzido tanto no número de cortes de Benders quanto no montante da diferença entre os coeficientes. Assim pode-se considerar que o impacto deste ajuste é desprezível nos resultados finais.





## **8 NOTA TÉCNICA nº 39 - Alteração na indexação do vetor de configuração da parábola de geração hidráulica máxima - Revisão 1 (Setembro 2007)**

No NEWAVE, as variáveis que compõem o sistema equivalente de energia são indexadas em função da configuração hidroelétrica. Uma configuração hidroelétrica é definida em função da topologia do sistema e dos estados das usinas hidrelétricas. Diversas alterações nos estados do sistema podem caracterizar uma mudança de configuração, como por exemplo, o término do enchimento de volume morto de um reservatório, a entrada em operação de uma máquina ou a variação na cota do canal de fuga de uma usina.

No NEWAVE versão 12, são definidos três tipos de configuração:

1. CONFIGURAÇÕES POR ENTRADA DE RESERVATÓRIO E/OU POR POTÊNCIA DE BASE: que é definida quando uma usina hidrelétrica termina o enchimento do volume morto de seu reservatório, quando entra em operação uma máquina de tal forma que o número de máquinas dessa usina seja igual número de máquinas de base ou quando há variação na cota média do canal de fuga da usina. As variáveis do sistema equivalente de energia indexadas por este tipo de configuração são: energia armazenável máxima, energia armazenável máxima por restrição de volume de espera, energia armazenável mínima por restrição de operação, energia de vazão mínima, energia evaporada, energia de desvio de água, energia natural afluyente bruta, energia controlável e energia fio d'água bruta.
2. CONFIGURAÇÕES POR ALTERAÇÃO DE POTÊNCIA: que é definida quando há a entrada em operação de uma máquina da usina. As variáveis do sistema equivalente de energia indexadas por este tipo de configuração são os coeficientes da parábola de geração hidráulica máxima.
3. CONFIGURAÇÕES POR QUALQUER MODIFICAÇÃO NA EXPANSÃO: definida quando ocorre qualquer evento que defina uma configuração por entrada de reservatório e/ou por potência de base ou por alteração de potência. As variáveis do sistema equivalente de energia indexadas por este tipo de configuração são: energia natural afluyente líquida e energia fio d'água líquida.

No relatório de convergência do programa NEWAVE são impressas três tabelas informando a ordem cronológica dessas configurações.

A capacidade de geração hidráulica leva em conta que durante a simulação da operação os níveis de armazenamento se modificam, e conseqüentemente, implicam em mudanças no valor da disponibilidade de geração hidráulica do sistema.

Uma vez que a agregação das usinas para a determinação do sistema equivalente de energia é feita *a priori* no NEWAVE, antes da simulação da operação, para cada configuração do tipo 2 (configurações por alteração de potência), são calculados três valores de geração hidráulica máxima, correspondentes à queda líquida considerando o reservatório no volume mínimo, no volume correspondente a 65% do volume útil e no volume máximo.

A partir destes três pontos, ajusta-se uma parábola, a partir da qual se obtém a geração hidráulica máxima em função da energia armazenada no mês, conforme ilustrado na Figura 1. A estimativa da parábola é feita somente no início de cada configuração do tipo 2, considerando as energias indexadas pela configuração tipo 1 corresponde.

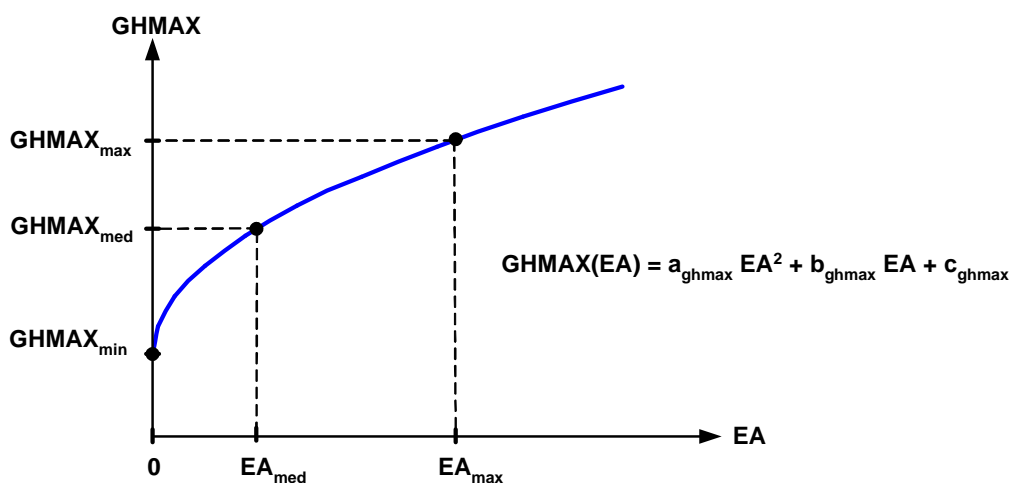


Figura 1 – Parábola de Geração Hidráulica Máxima

Durante a simulação forward ou recursão backward, caso um reservatório tenha concluído o enchimento do volume morto (o que ocasiona uma nova configuração do tipo 1) sem que tenha ocorrido qualquer entrada de máquina no mesmo subsistema desse reservatório (o que ocasiona uma nova configuração do tipo 2), pode haver uma consulta à parábola fora do domínio onde essa foi definida.

Para ilustrar o descrito acima, seja a evolução temporal das configurações apresentada na Tabela 1.

**Tabela 1 – Evolução das Configurações**

Configuração	jan	fev	mar	abr	mai	Jun	jul	ago	set	out	nov	dez
<b>Tipo 1</b>	3	3	3	4	4	4	4	4	4	4	4	5
<b>Tipo 2</b>	1	1	1	1	2	2	2	2	2	2	2	3
<b>Tipo 3</b>	3	3	3	4	5	5	5	5	5	5	5	6

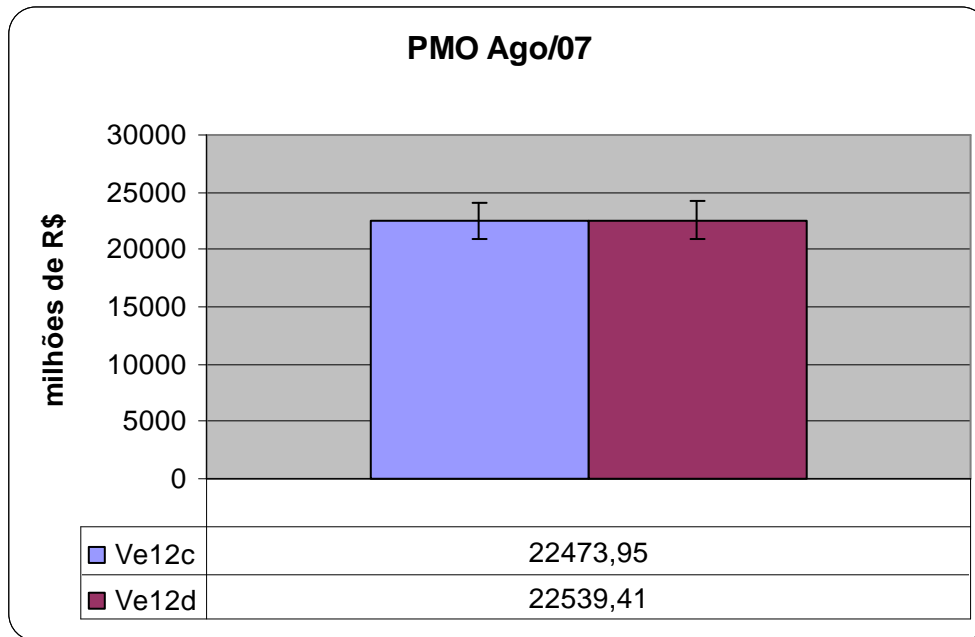
Para o exemplo da Tabela 1 houve uma entrada de reservatório do mês de abril, uma entrada de máquina no mês de maio e uma motorização no mês de dezembro.

Durante a simulação do mês de abril o valor da geração hidráulica máxima é calculado considerando a parábola da primeira configuração do tipo 2 (que durante os cálculos considerou a energia armazenável máxima da terceira configuração do tipo 1). Se em alguma série a energia armazenada for maior do que a energia armazenável máxima da terceira configuração do tipo 1, a consulta à parábola estará sendo para valores de armazenamento feita fora do domínio onde essa foi definida.

Para eliminar este efeito, a partir da versão 12d do NEWAVE a estimativa da parábola de geração hidráulica máxima é feita sempre que houver qualquer alteração de configuração no sistema, independente do tipo desta configuração. Isto corresponde a indexar os coeficientes desta parábola através da configuração do tipo 3 (configurações por qualquer modificação na expansão).

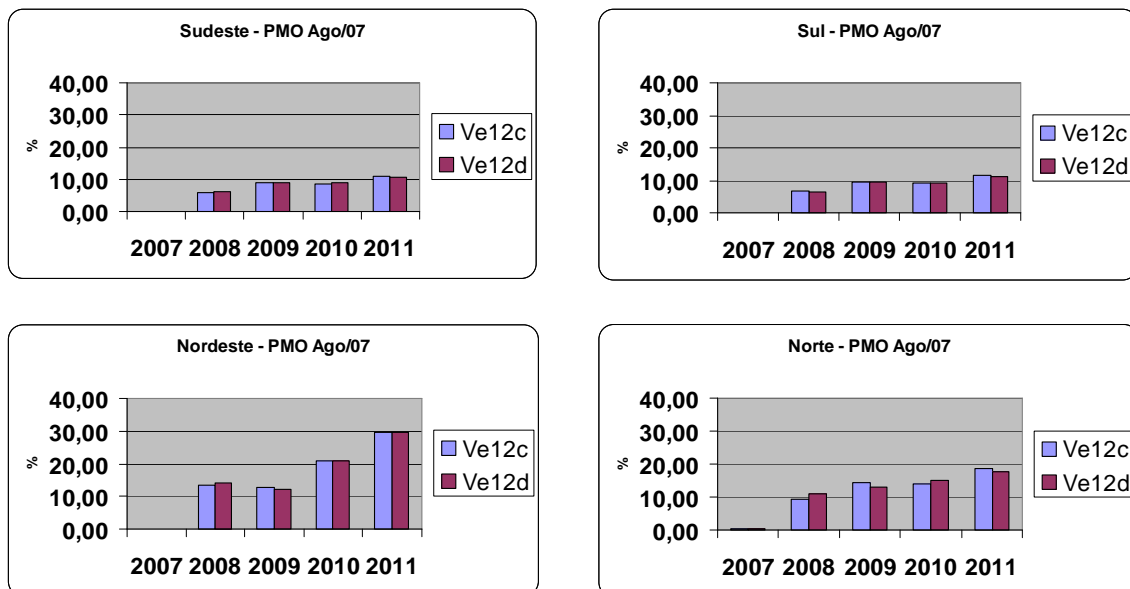
A título de ilustração, é analisado o impacto desta alteração nos valores do custo total de operação, custo marginal de operação, energia não suprida anual e risco anual de qualquer déficit para um caso baseado no PMO de agosto de 2007. Para tanto são comparados os resultados obtidos com as versões 12d e 12c.

Na Figura 2 é apresentado o valor esperado do custo total de operação. A diferença de aproximadamente 0,3% não é estatisticamente significativa, estando essa diferença dentro do intervalo de confiança de 95%.



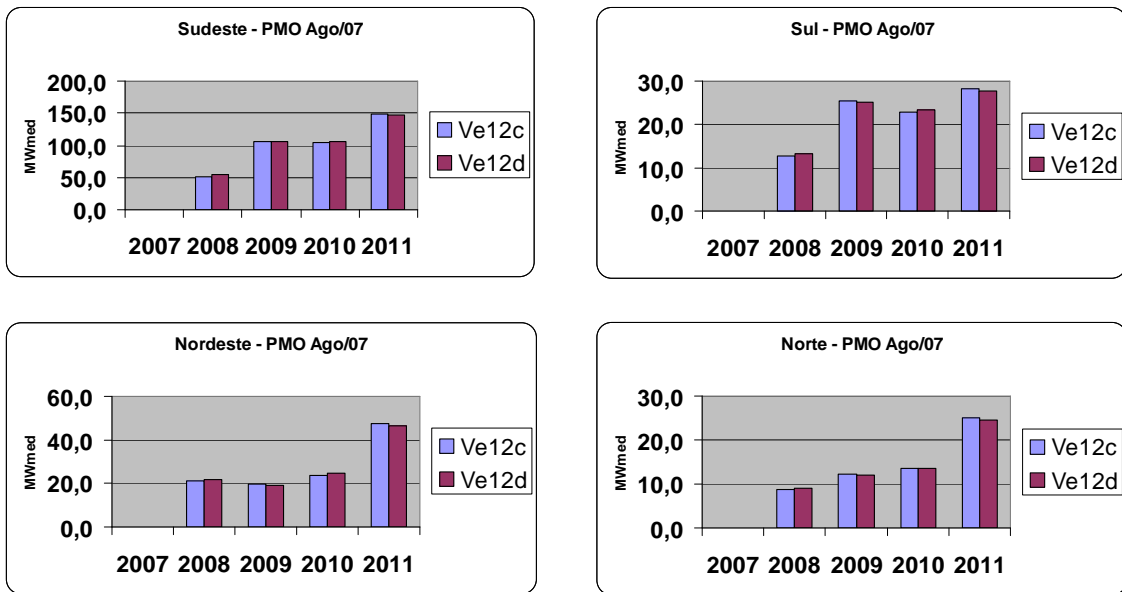
**Figura 2 – Valor Esperado do Custo Total de Operação**

Os valores de risco anual de déficit e o do valor esperado da energia não suprida, obtidos na simulação final, são apresentados nas Figuras 3 e 4, respectivamente.



**Figura 3 – Risco Anual de Déficit (qualquer déficit)**

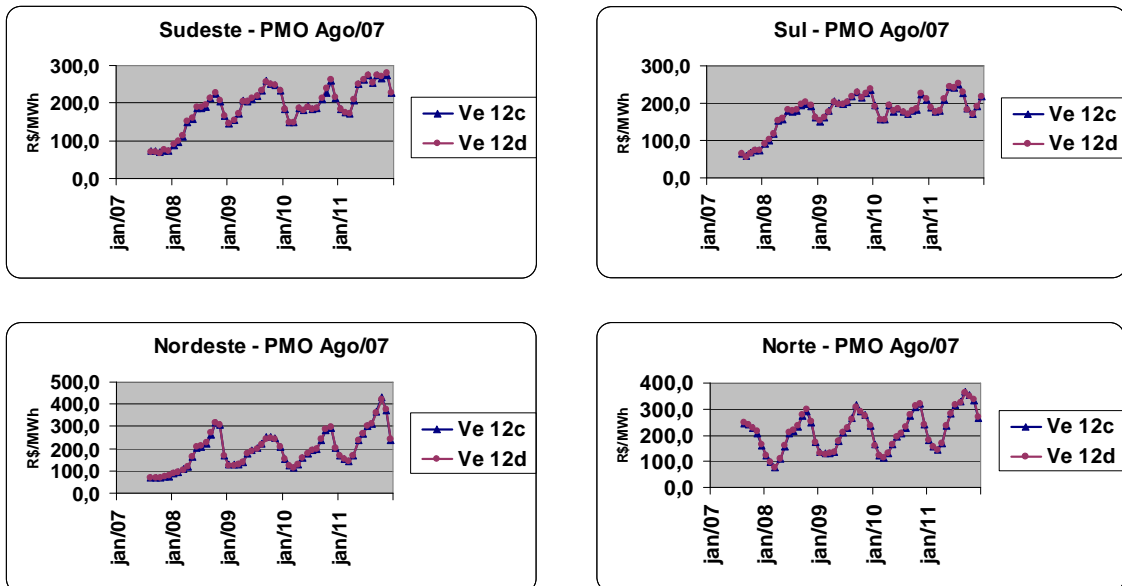




**Figura 4 – Valor Esperado da Energia Não Suprida**

Observa-se que, independentemente do subsistema, não há uma variação acentuada nos valores de energia não suprida e risco de déficit.

A média do custo marginal de operação, para o período do planejamento, é mostrada na Figura 5. Da mesma forma, a alteração efetuada na versão 12d tem um impacto bastante reduzido nos valores médios de CMO.



**Figura 5 – Custo Marginal de Operação**

Assim pode-se considerar que o impacto da alteração na indexação do vetor de configuração da parábola de geração hidráulica máxima é desprezível nos resultados finais.

## **9 NOTA TÉCNICA nº 40a - Análise da nota técnica ONS - Metodologia para consideração do intervalo de confiança para as auto-correlações temporais para uso no modelo PAR(p) (Janeiro 2009)**

A Nota Técnica “Metodologia para Consideração do Intervalo de Confiança para as Auto-Correlações Temporais para Uso no Modelo Par(p)” preparada pelo ONS preconiza que existe a possibilidade de que a propagação da tendência hidrológica no modelo PAR(p) influencie de forma exagerada os resultados do modelo de otimização usado no planejamento da operação do SIN. Conseqüentemente, o modelo de otimização pode ter uma disponibilidade distorcida de recursos no futuro, a partir de uma pequena variação na tendência hidrológica. Como forma de reduzir a propagação da tendência hidrológica a Nota Técnica propõe um método de estimação das auto-correlações de séries temporais trocando as auto-correlações amostrais do histórico por limites inferiores de seus intervalos de confiança. Esta Nota Técnica apresenta os comentários da equipe do CEPEL sobre a Nota Técnica do ONS.

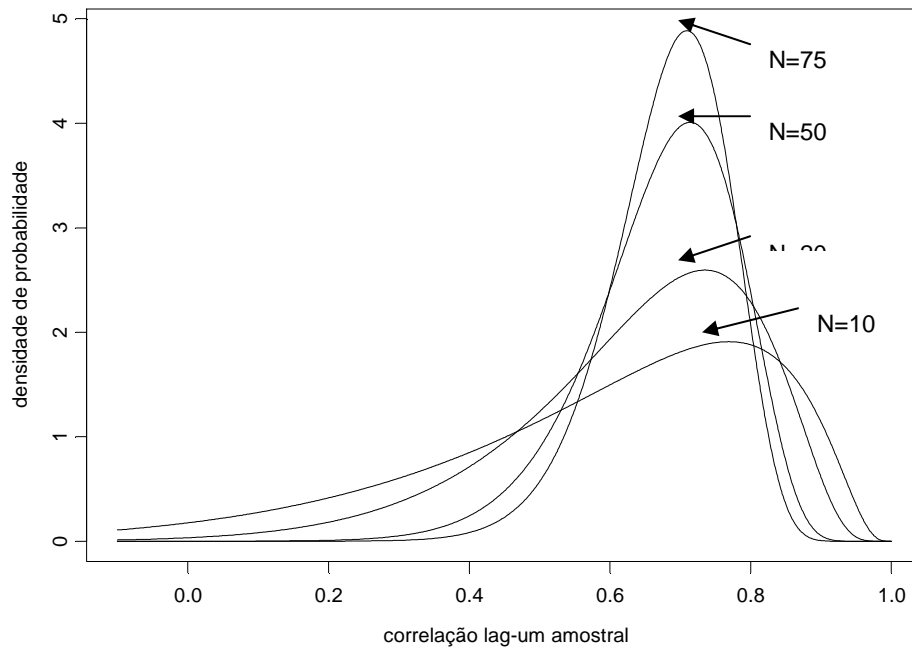
### **9.1 Considerações Iniciais**

Com efeito, a possibilidade preconizada pela Nota Técnica do ONS, de propagação exagerada da tendência hidrológica, existe já que a metodologia adotada no modelo GEVAZP procura estimar a função de auto-correlação teórica da população de séries temporais possíveis de ocorrer, a partir de uma única série temporal finita disponível (amostra), a saber: a série histórica com 76 anos de dados mensais. Portanto, esta estimação está submetida ao que se chama de variação amostral. Com outra amostra do mesmo comprimento (se possível) chegaríamos à outra estimativa, felizmente, com grande chance de se situar bem próxima da estimativa obtida com a primeira amostra, já que 76 anos compõem uma amostra de tamanho razoável. Ainda assim, existe a possibilidade da amostra particular disponível levar a uma superestimação das auto-correlações, levando à influência exagerada na propagação da tendência hidrológica citada na Nota Técnica do ONS. Por outro lado existe também a possibilidade da amostra particular disponível levar a uma subestimação da função de auto-correlação. Neste caso, a propagação da tendência hidrológica no modelo PAR(p) estaria influenciando de forma reduzida os resultados.

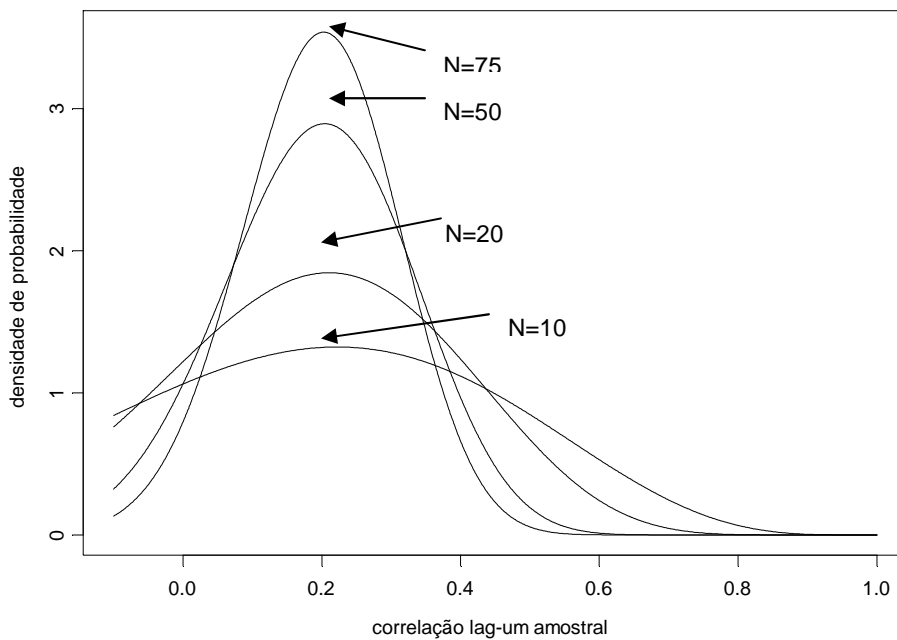
A metodologia de estimação adotada no modelo GEVAZP utiliza a função de auto-correlação amostral (correlograma). Em geral, o uso de correlogramas para estimação da função de auto-correlação populacional apresenta uma tendência a subestimar os valores de população, tendência esta que se reduz à medida que a amostra cresce. Por exemplo, Vujica Yevjevich (Stochastic Process in Hydrology, pp 50), apresenta uma distribuição para a auto-correlação amostral lag 1 de um processo AR(1), cujo valor esperado é igual a  $\rho N/(N+2)$ , onde  $\rho$  é o valor de população. No caso de uma amostra de 76 anos, isso indica uma subestimativa esperada de cerca de 2,6%. Nesta mesma distribuição, o desvio-padrão da distribuição amostral diminui com o tamanho da amostra e levemente com a correlação de população. Ou seja, quanto menor a correlação de população, maior o desvio-padrão, ou incerteza, da estimativa.

A Figura 1 mostra as funções densidades de probabilidades da auto-correlação amostral lag 1 para amostras de 75, 50, 20 e 10 observações para um valor populacional qualitativamente alto correspondente a  $\rho=0.70$ . Nesta figura, podem-se observar as chances relativas de sub e super estimação para os diferentes tamanhos de amostra, verificando-se como à medida que o tamanho da amostra aumenta diminui a chance de se obter valores amostrais afastados do valor populacional. Em particular, para uma amostra de 75 observações, a probabilidade de se obter correlações amostrais baixas é bastante reduzida, caso a correlação populacional for alta. Ainda neste caso, o valor esperado da distribuição dos valores amostrais é igual 0,6818, confirmando o “bias” na estimação desta estatística, mencionada no parágrafo anterior. Pode-se observar também que este valor esperado é um dos valores de maior frequência. Por sua vez, o desvio-padrão desta distribuição é igual a 0,0852.

Pode-se realizar a mesma análise considerando-se uma população de auto-correlação lag um baixa, por exemplo usando  $\rho=0,2$ . A Figura 2 mostra que a partir de uma amostra de 75 observações, a probabilidade de obter um valor amostral alto, por exemplo,  $r=0,7$  é aproximadamente nula. Ainda neste caso de  $N=75$ , a distribuição tem valor esperado igual 0,1948, confirmando novamente o “bias” na estimação deste estatística, mencionado anteriormente. Pode-se observar também que este valor esperado é um dos valores de maior frequência. Por sua vez, o desvio-padrão desta distribuição é igual a 0,1119, superior ao da correlação igual a 0,7.

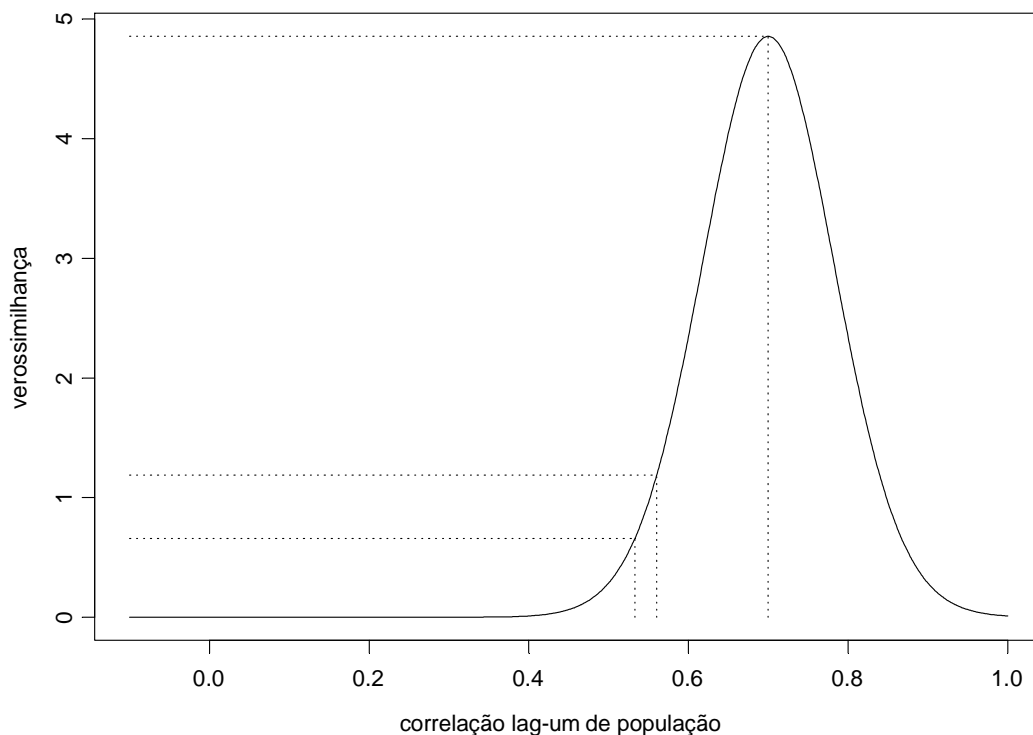


**Figura 1 - Distribuições da auto-correlação amostral lag 1 para diferentes tamanhos de amostra ( $\rho=0,7$ )**



**Figura 2 - Distribuições da auto-correlação amostral lag 1 para diferentes tamanhos de amostra ( $\rho=0,2$ )**

Uma análise orientadora consiste em fixar o valor amostral e calcular as densidades de probabilidades para diferentes valores de população, configurando uma função de verossimilhança. A Figura 3 mostra a função de verossimilhança correspondente a uma amostra de 75 observações com correlação amostral igual a 0,70. É importante assinalar que o valor mais provável (estimativa de máxima verossimilhança) é exatamente o valor amostral. Observa-se também, a perda de verossimilhança associada à troca da estimativa amostral (0,7) pelos limites inferiores de intervalos de confiança uni caudais de 95% (0,5603) ou 97,5% (0,533). No caso de 95% a verossimilhança cai de 4,854 para 1,186 e, no caso de 97,5% cai de 4,854 para 0,6592.



**Figura 3 - Verossimilhança da auto-correlação de população lag 1 para amostra de tamanho 75 e  $r=0,7$**

Podemos então, em relação à correlação lag um amostral de amostras finitas de um processo AR(1), concluir que:

- para qualquer tamanho de amostra pode fornecer estimativas acima ou abaixo do valor populacional;
- a variabilidade é decrescente com o tamanho da amostra;

- é uma estimativa tendenciosa para menos do correspondente valor populacional. À medida que o tamanho da amostra cresce a tendência desta estimativa se reduz, e assintoticamente a estimativa é não tendenciosa;
- é razoável propor uma correção desta tendência. Esta correção vai na direção contrária do proposto na Nota técnica do ONS.

Especificamente, para uma amostra de 75 observações de um processo AR(1) pode-se levantar que:

- Se a população apresentar uma correlação lag um alta, a correlação lag um amostral muito provavelmente será alta também;
- Se a população apresentar uma correlação lag um baixa, a correlação lag um amostral muito provavelmente será baixa também;
- Se a correlação lag um amostral obtida for elevada (em torno de 0,7) reduzi-la para o limite inferior do intervalo de 95% fornecerá uma estimativa da correlação de população 4,1 vezes menos provável que o valor amostral e, reduzi-la para o limite inferior do intervalo de 97,5% fornecerá uma estimativa da correlação de população 7,4 vezes menos provável que o valor amostral.

Em geral, em situações de modelagem de séries temporais estacionárias, onde o modelo AR(1) é uma caso particular, vale observar que:

As auto-correlações amostrais obtidas da série histórica fornecem informações relevantes sobre a estrutura de auto-correlação da população de séries possíveis de ocorrer. Uma postura clássica consiste em considerar estas informações como as únicas disponíveis.

Por outro lado, o uso da postura clássica não implica que os valores das auto-correlações amostrais devam ser adotados na íntegra no modelo. Com efeito, as técnicas clássicas de inferência estatística sugerem que na construção do modelo aplique-se o princípio da parcimônia sem necessariamente reproduzir todo o correlograma amostral.

Por exemplo, de posse de uma amostra com 10 observações e auto-correlação lag-um amostral baixa, igual a 0.20, pode-se optar por considerar o modelo mais simples de aleatoriedade pura (AR(0)). A formalização de um intervalo de confiança para a auto-correlação amostral indica que este intervalo engloba o zero. Diz-se que a auto-correlação amostral é não significativa e decide-se adotar o modelo mais simples, sem reprodução exata do correlograma amostral. Uma análise de

verossimilhança usando a distribuição referida acima indicaria uma perda de verossimilhança muito pequena (ver a curva para  $N=10$  da Figura 2 para uma análise aproximada).

O modelo GEVAZP emprega uma metodologia visando obter uma representação parcimoniosa do correlograma. Na fase de escolha da ordem  $p$  do modelo, são utilizados intervalos de confiança referentes aos coeficientes amostrais de auto-correlação parcial para identificar os não significativos. Com isto, escolhe-se quantos coeficientes de auto-correlação amostrais serão utilizados para o ajuste do modelo (ordem do modelo). A idéia é levar para o modelo a parte mais importante da experiência empírica, representada pelos primeiros coeficientes de auto-correlação, da maneira como foi observada na série histórica. Os coeficientes de auto-correlação amostral mais defasados, que pouco acrescentam à explicação da variabilidade futura, são deixados de lado, não sendo incorporados ao modelo.

Em adição ao uso do conceito de representação parcimoniosa do correlograma, o modelo GEVAZP adota o pressuposto que a tendência hidrológica tem efeito positivo ou nulo nas afluições futuras, não sendo possível que um aumento da tendência atual se reflita em redução de valores esperados para afluições futuras. Se as auto-correlações amostrais calculadas no histórico levarem a um modelo com efeito negativo da tendência hidrológica em afluição futura, isto se deve a variação amostral, e a parcela específica da experiência empírica não deve ser levada para a modelagem. O procedimento implementado revê a auto-correlação parcial amostral de maior lag anteriormente considerada significativa, identifica ela como não significativa, e simplifica o modelo, adotando-se menos um termo auto-regressivo. Com isto provoca-se o menor afastamento possível do pico de verossimilhança consistente com a informação adicional sobre o sinal do efeito da tendência hidrológica nas afluições futuras. Os testes mostraram que com a aplicação desta correção a reprodução das estatísticas referentes às secas ocorridas no histórico não são prejudicadas.

## **9.2 Análise da Metodologia Proposta**

Em relação à proposta descrita na Nota Técnica preparada pelo ONS, trata-se também de um desvio em relação à informação empírica representada pelo histórico de afluições e, portanto corresponde também a uma perda de verossimilhança. A justificativa deste afastamento seria a existência de uma informação adicional além da série histórica, no caso informação referente à “curva de



propagação do delta”. A análise desta informação será feita depois da análise do procedimento de redução das auto-correlações proposto.

No procedimento sugerido, ocorre redução simultânea de todas as auto-correlações do correlograma. Sendo assim, a perda de verossimilhança tende a ser da ordem ou maior do que as perdas observadas no exercício do item anterior para os mesmos níveis de confiança. O importante aqui é observar que sem o apoio de uma técnica formal de inferência estatística que minimize a perda de verossimilhança, as séries sintéticas geradas a partir da nova metodologia ou serão muito parecidas com as atualmente geradas (caso do quantil de 50%) e não atingem o objetivo de reduzir a propagação da tendência hidrológica atual, ou serão muito diferentes atingindo o objetivo (caso de quantis menores de 50%), mas com perdas substanciais de verossimilhança, dificilmente justificáveis numa análise formal de inferência estatística. A consequência imediata é que as estatísticas de reprodução de características dos períodos críticos serão afetadas já que a estrutura de auto-correlação é responsável pela reprodução das características das secas históricas, conforme os desenvolvimentos realizados pelo CEPEL desde a década de 80. Como observação adicional referente ao procedimento proposto, ao se reduzir as auto-correlações amostrais antes da fase de escolha das ordens dos termos auto-regressivo do modelo, os testes de significância dos coeficientes de auto-correlação parcial resultante não são mais exatos.

Em relação à informação adicional a ser utilizada para justificar a redução das auto-correlações amostrais (“curva de propagação do delta”), há de se adiantar na definição sobre o que existe de informação além da série histórica. No caso do sinal do efeito da tendência hidrológica, existe a informação de especialistas em hidrologia que no regime hidrológico das bacias do sistema hidrotérmico brasileiro, este sinal deve ser não-negativo. Que informação existe acerca do limite superior deste efeito? É necessário fazer estudos sistemáticos acerca do limite superior deste efeito a fim de criar sensibilidade para responder as perguntas: Existe indicio empírico para admitir que os valores de propagação do delta das curvas da Figura 1 da nota técnica do ONS estão superestimando os correspondentes valores populacionais? Que limites superiores podem ser impostos a estes valores?

Como observação adicional, as definições de delta do mês  $m$  como simples diferença da afluência em relação à média do mês  $m$  e de propagação como a reflexão no delta do mês  $m$  causada por uma variação unitária ocorrida na afluência do mês  $m-1$  não levam em conta a variação sazonal dos desvios padrões. Por outro lado,

de forma incoerente, o relatório do ONS omite na fórmula de cálculo da propagação o efeito da variação sazonal dos desvios-padrões. Por exemplo, o último parágrafo da

página 4, ficaria coerente se o coeficiente  $\phi_1^{(m)}$  fosse substituído por  $\frac{\sigma_m \phi_1^{(m)}}{\sigma_{m-1}}$  e o

coeficiente  $\phi_2^{(m+1)}$  fosse substituído por  $\frac{\sigma_{m+1} \phi_2^{(m+1)}}{\sigma_{m-1}}$ . De forma similar ao início da

página 5, o coeficiente  $\phi_1^{(m+1)}$  deveria ser substituído por  $\frac{\sigma_{m+1} \phi_1^{(m+1)}}{\sigma_m}$  e correções

devem ser adicionadas às duas equações seguintes desta página. A Figura 1 parece ter sido obtida usando as correções acima. Isto explica porque a curva de propagação do delta do subsistema Norte se eleva fortemente para os meses da estação seca, meses de desvio padrão muito menor do que os outros meses do ano. Note que nestes meses uma variação unitária de afluência é muito menos provável do que a mesma variação, por exemplo, em fevereiro e março. Unir numa curva estas medidas mensais de propagação por unidade se está misturando propagações devidas a variações não igualmente prováveis.

Uma outra maneira de ver o problema é a mistura dos efeitos da sazonalidade da dispersão com os efeitos da estrutura de dependência temporal, motivo da análise. Uma medida de propagação que isola o efeito desejado e une efeitos de variações mensais aproximadamente equiprováveis é trabalhar com diferenças da média mensal divididas pelos desvios padrões mensais, o que corresponde a usar como fórmula de cálculo da curva de propagação do delta as fórmulas apresentadas na Nota Técnica do ONS.

### 9.3 Conclusão

O procedimento proposto na Nota Técnica elaborada pelo ONS de redução de todas as auto-correlações amostrais afasta a modelagem da estrutura de correlação temporal da experiência empírica presente na série histórica, causando perda de verossimilhança e prejuízos à representação das características das secas históricas.

Os testes estatísticos usados para escolher as ordens dos termos auto-regressivos do modelo passam a não serem exatos.

Uma investigação mais elaborada deveria ser feita para verificar a plausibilidade de estar ocorrendo propagação exagerada da tendência hidrológica.

#### **9.4 Avaliação da nova proposta do ONS**

Após leitura desta Nota Técnica, o ONS apresentou uma proposta alternativa onde a política de operação passa a ser construída pelo algoritmo de programação dinâmica dual estocástica com base em um conjunto de cenários de energias afluentes gerados por um modelo PAR(p) estimado trocando as autocorrelações amostrais do histórico por limites inferiores de seus intervalos de confiança conforme proposta anterior. Já os cenários de energias afluentes utilizados para avaliar a política de operação obtida e as condições de atendimento do sistema seriam gerados utilizando o modelo PAR(p) atual, com base em técnicas de estimação consagradas na literatura especializada. Isto significa dizer que se acredita que o processo de energias afluentes segue o modelo PAR(p) em uso, porém admite-se construir uma política de operação com base em outro processo estocástico de energias afluentes, na esperança de se obter uma menor volatilidade dos custos marginais de operação. Entende-se aqui, volatilidade como uma flutuação temporal do custo marginal de operação, base do preço da energia de curto prazo.

Duas conseqüências negativas da proposta do ONS são imediatas:

- i) Não otimalidade – as políticas de operação construídas com a aplicação desta proposta são em geral não-ótimas, com custo total esperado de operação maior ou igual que a política construída da forma tradicional. Se o aumento de custo for significativo, abre-se a discussão sobre a oportunidade de se adotar esta nova política de operação.
- ii) Não-unicidade da política - A cada nível de significância adotado no cálculo dos limites inferiores das autocorrelações corresponde uma política diferente. Este aspecto pode gerar uma grande discussão em torno de qual intervalo de confiança adotar, criando desconfiança junto aos agentes do sistema elétrico, já que o procedimento proposto de estimação de função de autocorrelação não tem respaldo técnico na literatura.

Não obstante essas conseqüências negativas, o CEPEL realizou algumas simulações para se obter sensibilidade quanto aos impactos desta proposta nos custos esperados totais de operação e nas evoluções de custo marginal de operação.

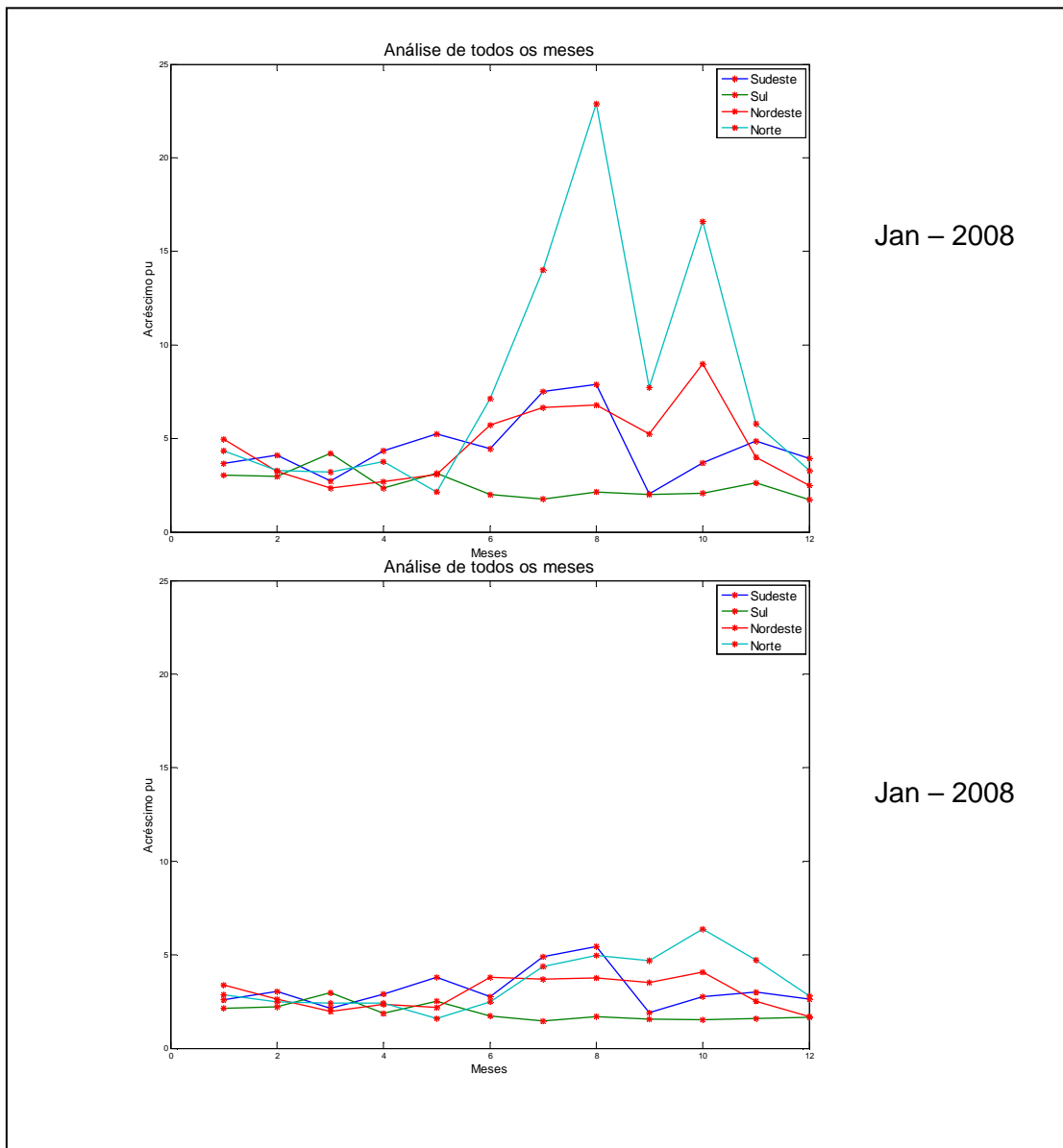
Foram analisados dois procedimentos:

Procedimento 1: política de operação e simulação da operação utilizando 2000 cenários de energias afluentes calculadas pelo modelo PAR(p) atualmente em uso.

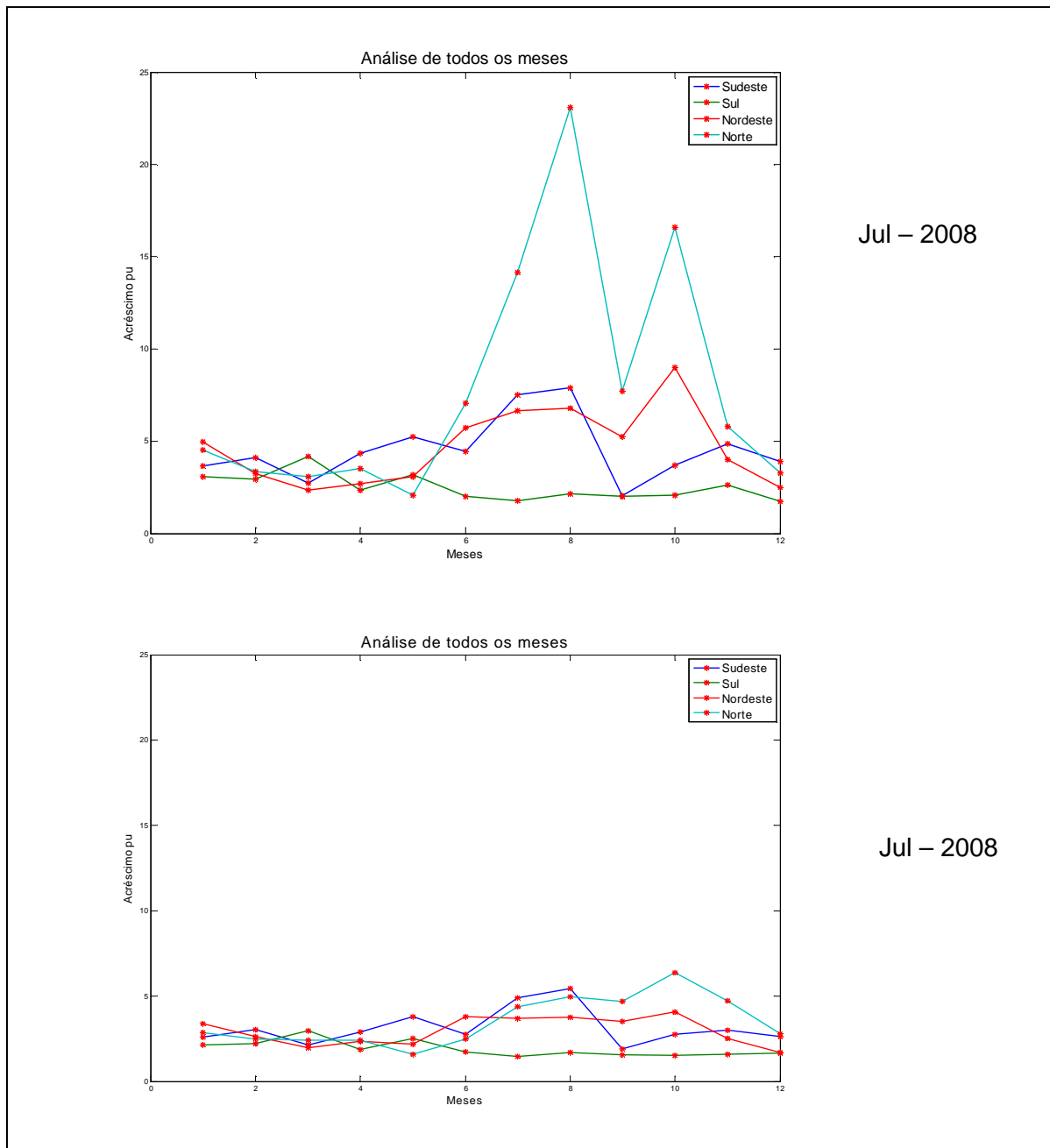
Procedimento 2: política de operação utilizando cenários de energias afluentes calculadas pelo modelo PAR(p) estimado utilizando os limites inferiores dos intervalos de confiança das funções de auto-correlação amostrais a um desvio-padrão. Já a simulação final empregou os mesmos cenários do procedimento 1.

Cada procedimento foi avaliado utilizando os conjuntos de dados correspondentes aos PMOs de janeiro e julho de 2008.

As Figuras 4 e 5, respectivamente correspondentes aos PMOs de janeiro e julho de 2008, comparam as “curvas de propagação do delta do mês” para os modelos PAR(p) ajustados para este exercício, observando-se que o ajuste trocando as auto-correlações pelo limite inferior reduziu estas curvas para todos os sistemas.



**Figura 4- Curvas de propagação do delta - PMO jan/2008**



**Figura 5 – Curva de Propagação do delta - PMO jul/2008**

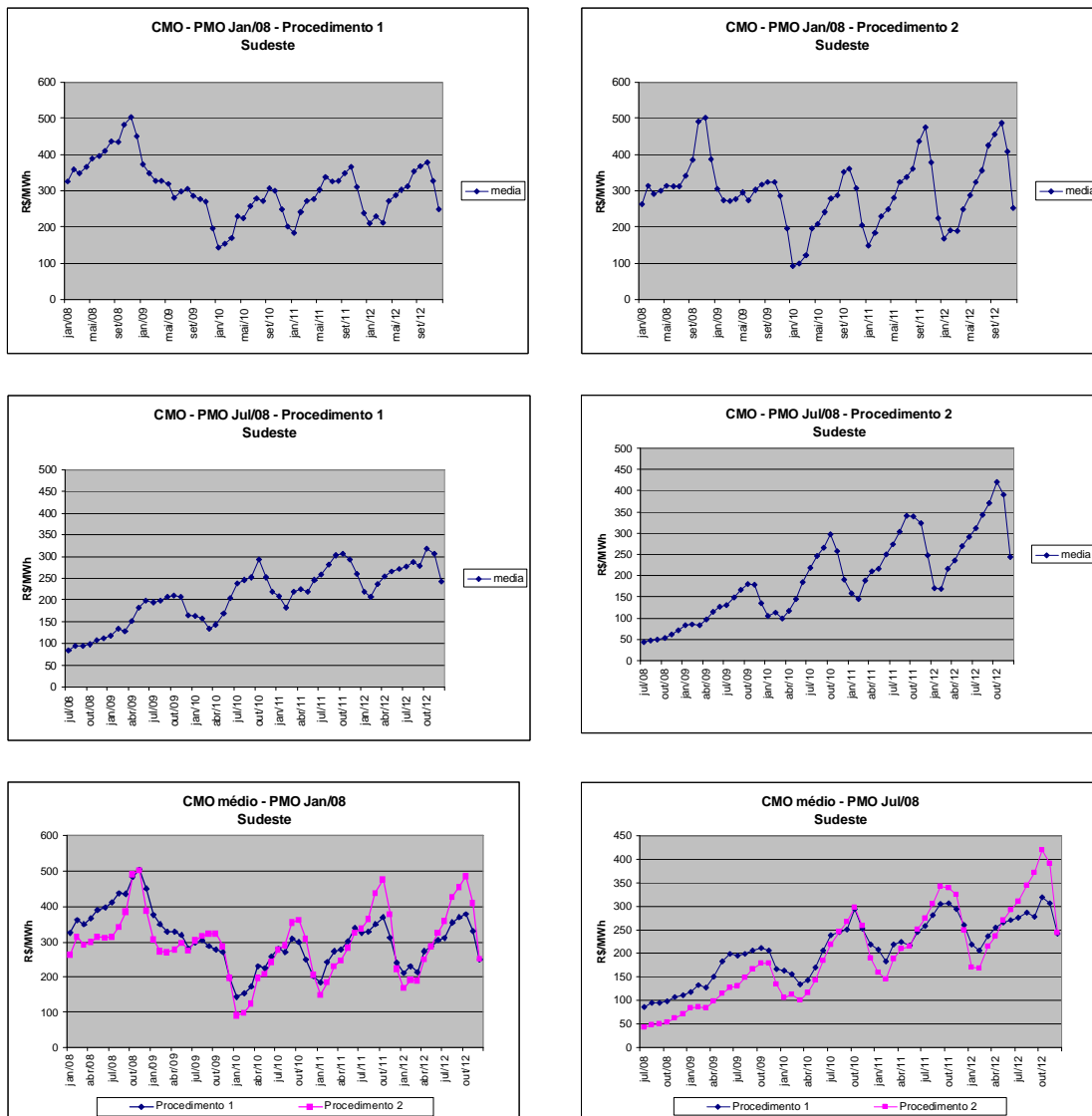
A Tabela 1 fornece para cada PMO analisado o custo total esperado de operação das simulações de cada procedimento, junto com o desvio padrão e nas duas últimas colunas são apresentadas as variações em termos percentuais e em número de desvios padrão.

**Tabela 1 - Custo Total Esperado de Operação**

PMO	Procedimento 1	Procedimento 2	Variação (%)	Variação (nº d.pad.)
Jan/2008	50.677,84 ± 1.486,65	60.406,84 ± 1856,23	19,20	6,54
Jul/2008	31.190,67 ± 992,7	35.422,04 ± 11.303,41	13,57	4,26

Pode-se observar da Tabela 1 que conforme esperado o procedimento 2 leva a custos esperados totais de operação superiores aos custos do procedimento 1. A variação é estatisticamente significativa, superando largamente os desvios padrões de cada estimativa. Numericamente, são diferenças nada desprezíveis, 20% no PMO de janeiro e 14% no PMO de julho.

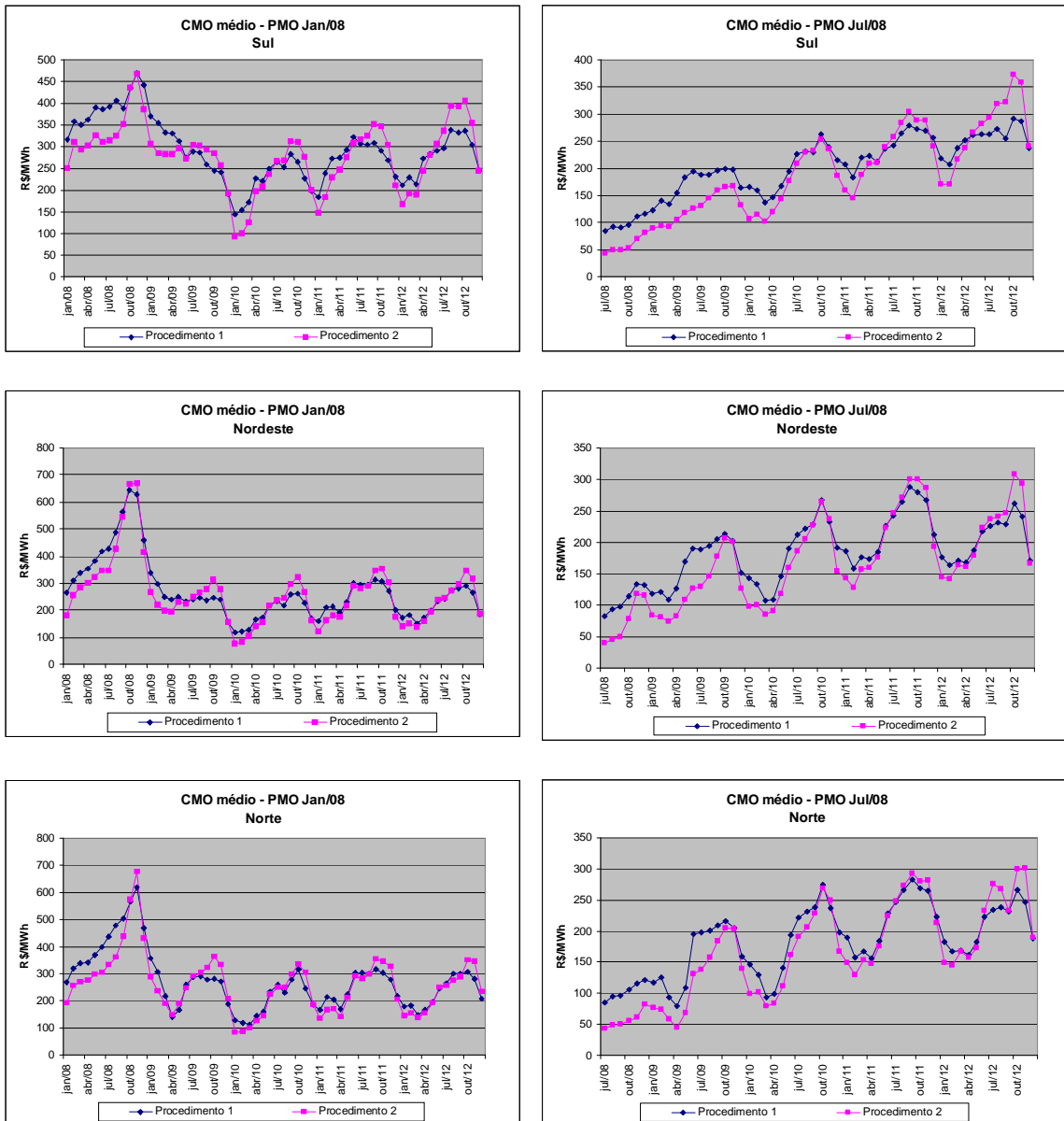
A Figura 6 apresenta na primeira linha, para os procedimento 1 e 2, as evoluções ao longo dos meses dos valores esperados dos CMOs do Sudeste das 2000 séries da simulação para o PMO de Janeiro de 2008. Na segunda linha, apresentam-se os mesmos resultados para o PMO de Julho de 2008. A última linha compara as evoluções dos valores esperados dos CMOs dos dois procedimentos obtidos com os dois procedimentos para um cada um dos PMOs.



**Figura 6 – Análise do CMO médio - Sudeste**

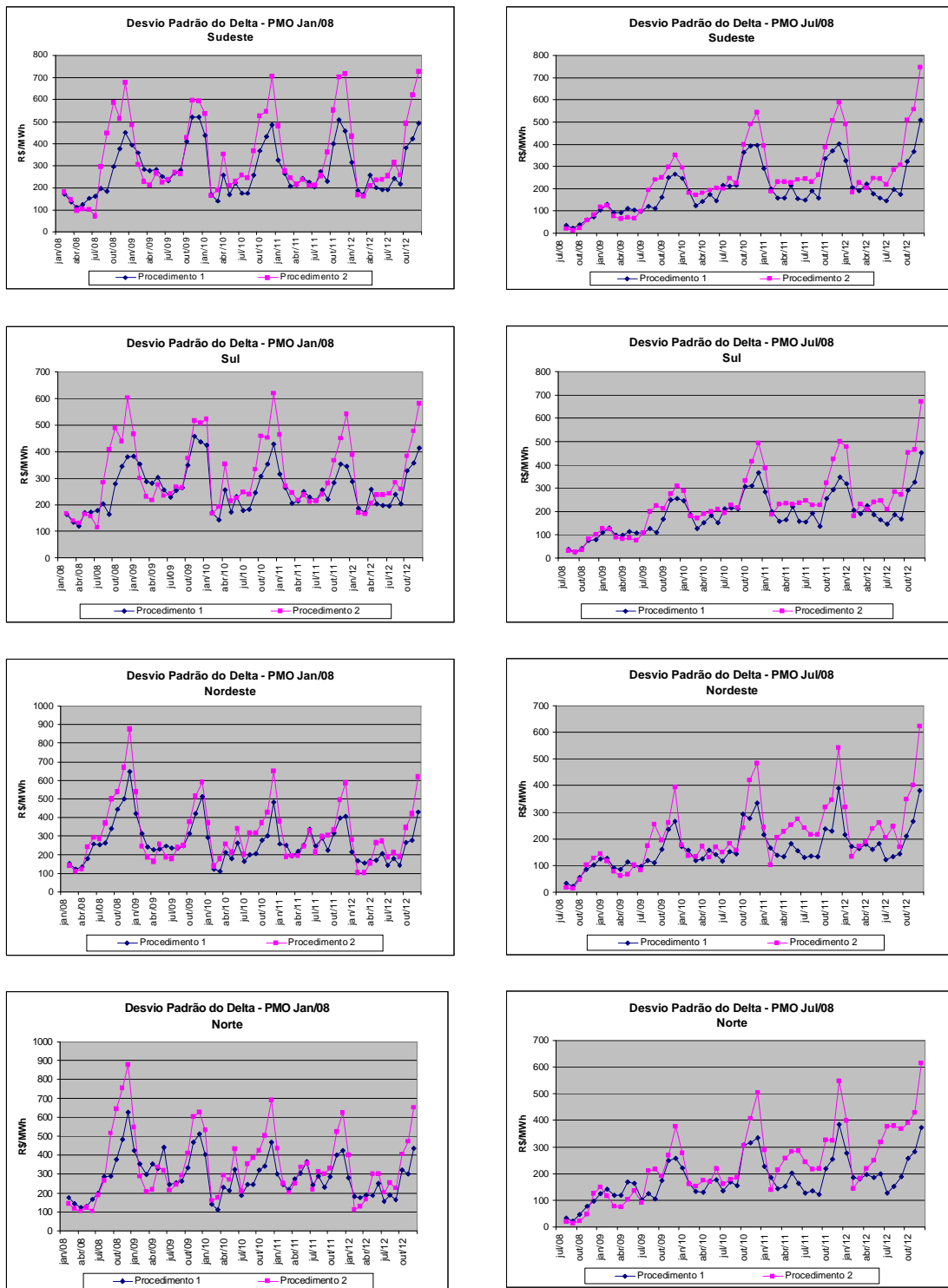
Comparando-se os resultados, pode-se observar que o procedimento 2 leva a uma flutuação mensal maior da evolução mensal dos custos marginais de operação esperados que o procedimento 2, exatamente ao contrário do desejado. Resultados similares foram obtidos para os outros sistemas. O gráfico com a evolução do CMO médio é apresentado na Figura 7.





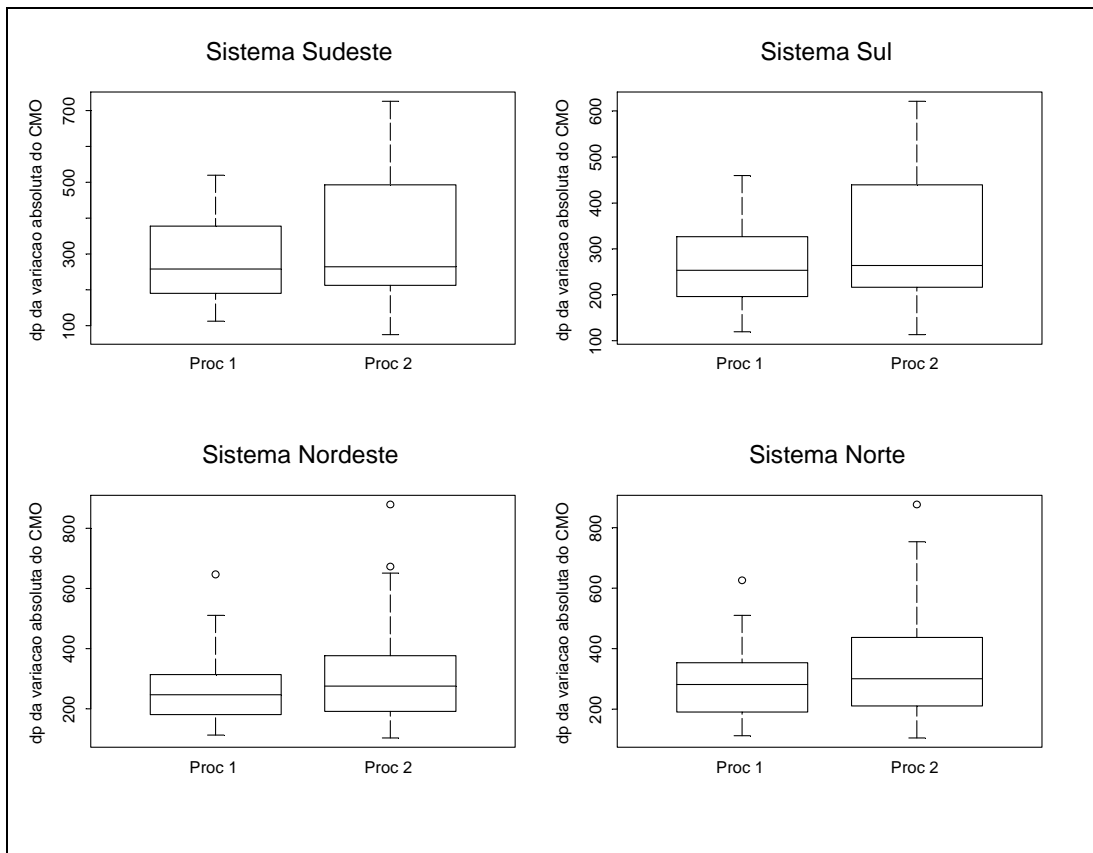
**Figura 7 – CMO médio para os sistemas Sul, Nordeste e Norte**

Para avaliar numericamente a volatilidade do CMO considerou-se a diferença absoluta entre CMOs de meses consecutivos. Como primeira métrica, foi calculado o desvio padrão desta diferença absoluta ao longo das 2000 simulações. A Figura 8 compara esta métrica para os procedimentos 1 e 2 para os PMOs de Jan/2008 e Jul/2008 para cada sistema. Pode-se notar que na maioria dos meses o desvio padrão da diferença absoluta do procedimento 2 é superior ao do procedimento 1 para todos os sistemas.

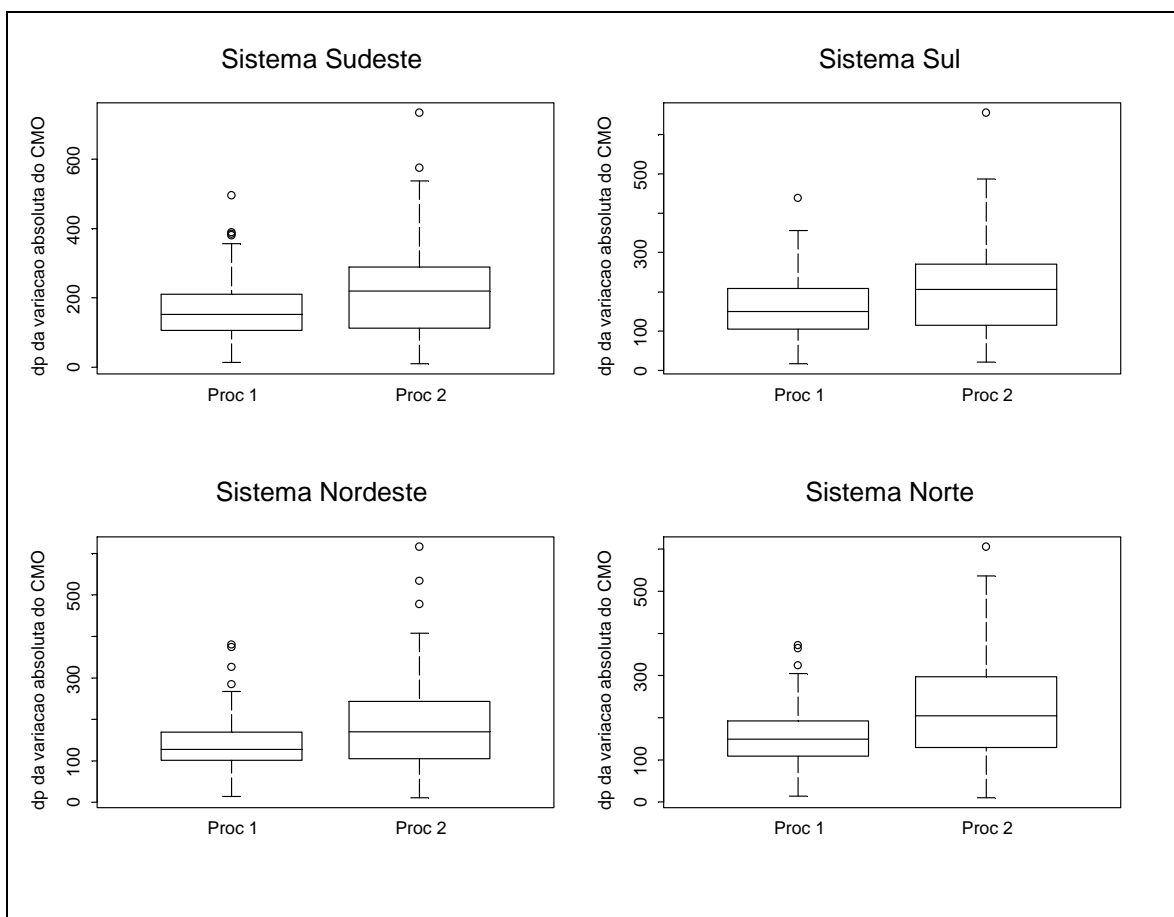


**Figura 8 – Desvio Padrão das diferenças absolutas mensais de CMO**

As figuras 9 e 10 comparam os boxplots dos desvios padrões das diferenças absolutas mensais de CMO para os PMOs de janeiro de 2008 e as Tabelas 2 e 3 apresentam os valores médios. Observa-se a maior volatilidade do procedimento 2 em todos os sistemas nos dois PMOs.



**Figura 9 – Box-plot do D. P. das Dif. Absolutas Mensais de CMO PMO Jan 08**



**Figura 10 – Box-plot do D. P. das Dif. Absolutas Mensais de CMO PMO Jul 08**

**Tabela 2 – Média do D.P. das diferenças absolutas de CMO PMO Jan 2008**

Sistema	Procedimento 1	Procedimento 2	Varição (%)
Sudeste	280,19	343,24	22,50
Sul	261,83	311,16	18,83
Nordeste	264,96	315,48	19,06
Norte	287,0	344,82	20,14

**Tabela 3 – Média do D.P. das diferenças absolutas de CMO PMO Jul 2008**

Sistema	Procedimento 1	Procedimento 2	Variação (%)
Sudeste	184,46	245,18	32,92
Sul	175,68	230,72	31,33
Nordeste	152,19	204,74	34,52
Norte	166,76	229,31	37,51

As Tabelas 4 e 5 apresentam o desvio padrão ao longo das 2000 simulações de todas as diferenças absolutas entre CMO's de meses consecutivos para os PMOs de janeiro de 2008 e julho de 2008. Observa-se que a maior volatilidade do procedimento 2 em todos os sistemas nos dois PMOs.

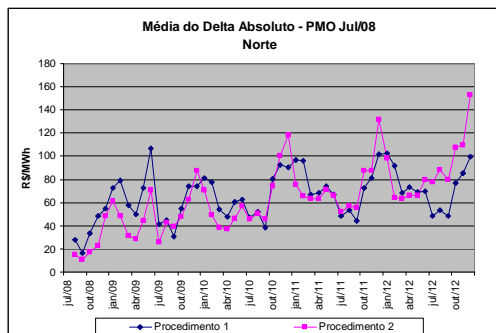
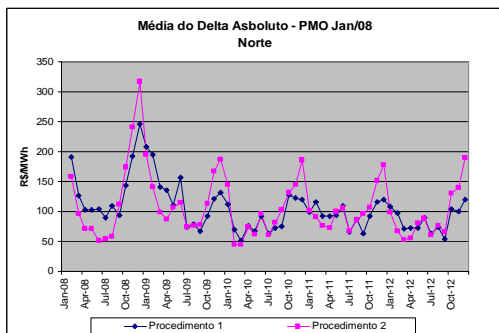
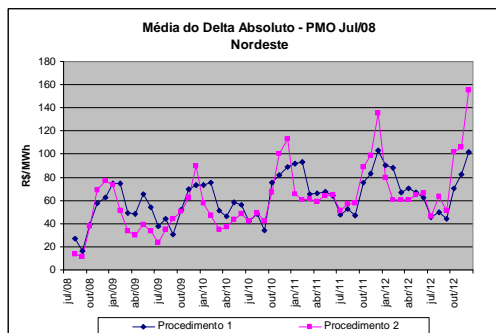
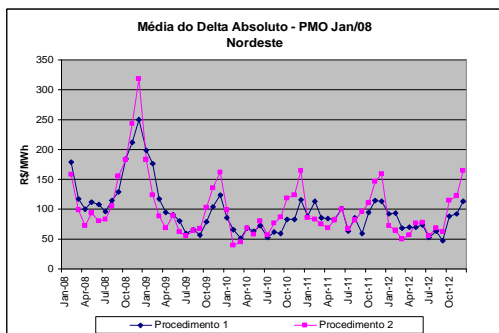
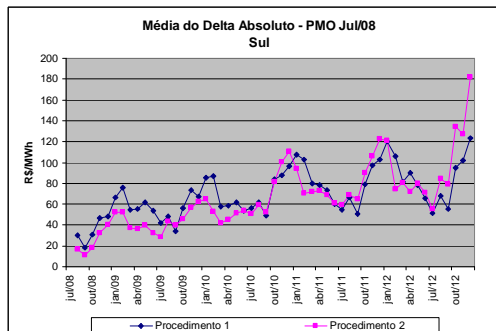
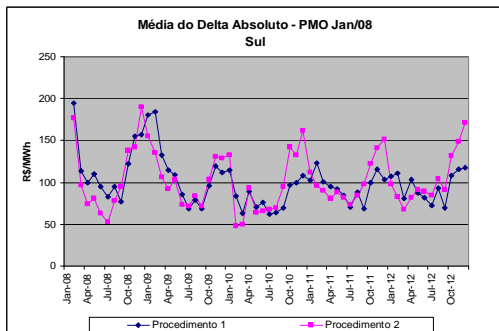
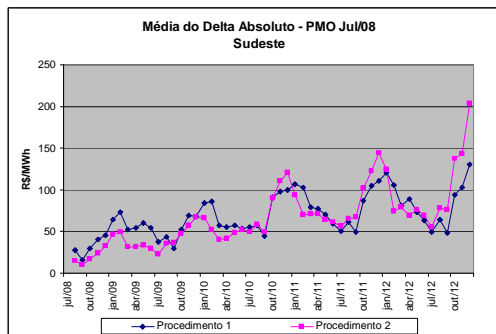
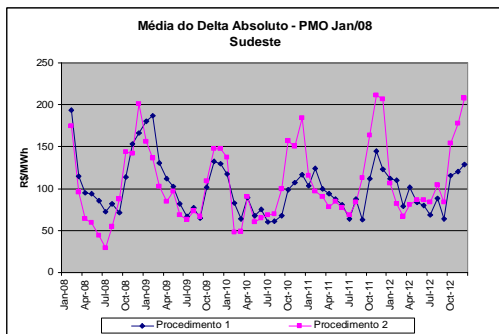
**Tabela 4 – Desvio padrão da variação mensal de CMO PMO Jan 2008**

Sistema	Procedimento 1	Procedimento 2	Variação (%)
Sudeste	303,33	389,79	28,50
Sul	276,41	339,58	22,86
Nordeste	289,43	358,21	23,76
Norte	309,28	391,55	26,60

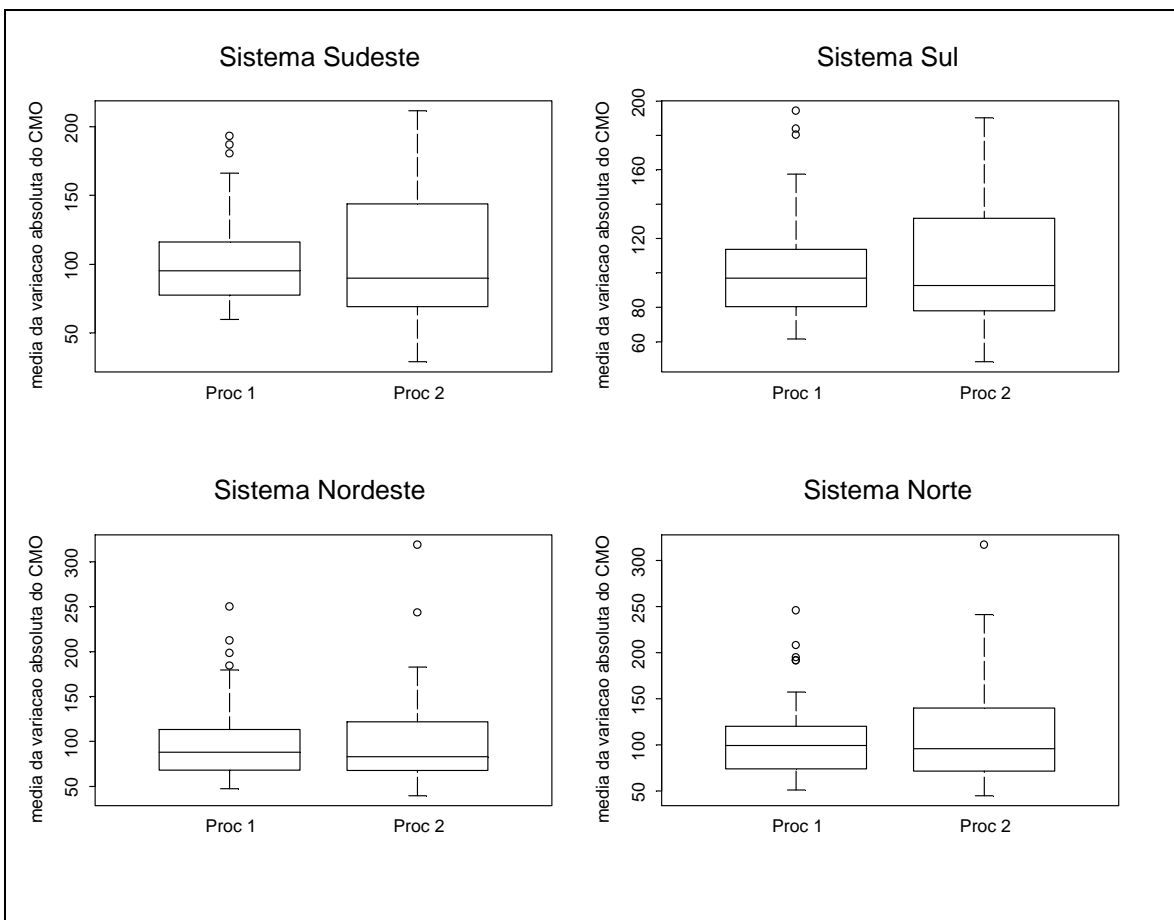
**Tabela 5 – Desvio padrão da variação mensal de CMO PMO Jul 2008**

Sistema	Procedimento 1	Procedimento 2	Variação (%)
Sudeste	214,02	292,99	36,90
Sul	197,62	266,79	35,00
Nordeste	171,46	248,80	45,11
Norte	184,97	265,41	43,49

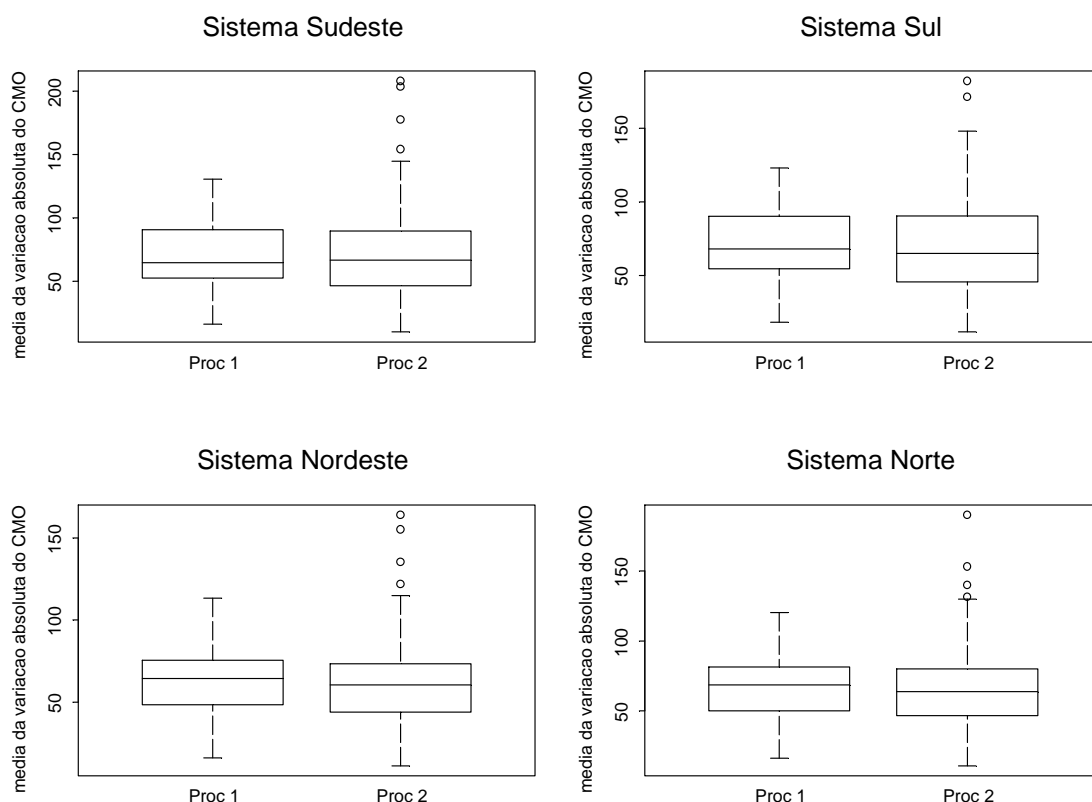
Como segunda métrica, foi calculada a média da diferença absoluta dos CMOs ao longo das 2000 simulações. A Figura 11 e os boxplots das Figuras 12 e 13 comparam esta métrica para os procedimentos 1 e 2 para os PMOs de Jan/2008 e Jul/2008 para cada sistema. Observa-se que a flutuação desta métrica ao longo dos meses no procedimento 2 é superior à flutuação ao longo dos meses para o procedimento 1 em todos os sistemas.



**Figura 11 – Média das diferenças absolutas mensais de CMO**



**Figura 12 – Box-plot Média das Dif. Abs. Mensais CMO PMO Jan 08**



**Figura 13 – Box-plot Média das Dif. Abs. Mensais CMO PMO Jul 08**

Os valores médios ao longo dos meses das médias das diferenças absolutas mensais de CMO para os PMOs de janeiro de 2008 e julho de 2008 são apresentados nas Tabelas 6 e 7. Nesta avaliação, observa-se pouca diferença entre os dois procedimentos em todos os sistemas nos dois PMOs.

**Tabela 6 – Média da média das diferenças absolutas de CMO PMO Jan 2008**

Sistema	Procedimento 1	Procedimento 2	Variação (%)
Sudeste	100,83	105,66	4,79
Sul	100,66	102,08	1,41
Nordeste	96,88	100,77	4,02
Norte	105,53	107,54	1,91

**Tabela 7 – Média da média das diferenças absolutas de CMO PMO Jul 2008**



Sistema	Procedimento 1	Procedimento 2	Varição (%)
Sudeste	68,70	66,76	-2,81
Sul	69,64	66,47	-4,55
Nordeste	62,34	61,07	-2,02
Norte	65,62	63,46	-3,33

## 9.5 Conclusão

Da primeira fase do estudo, realizado em maio de 2008, as seguintes conclusões foram obtidas:

“O procedimento proposto na Nota Técnica elaborada pelo ONS de redução de todas as auto-correlações amostrais afasta a modelagem da estrutura de correlação temporal da experiência empírica presente na série histórica, causando perda de verossimilhança e prejuízos à representação das características das secas históricas.

Os testes estatísticos usados para escolher as ordens dos termos auto-regressivos do modelo passam a não serem exatos.

Uma investigação mais elaborada deveria ser feita para verificar a plausibilidade de estar ocorrendo propagação exagerada da tendência hidrológica.”

Para a análise da nova proposta do ONS, foram realizados experimentos com os PMOs de janeiro e julho de 2008. Estes experimentos conduziram às seguintes conclusões:

O modelo PAR(p) ajustado trocando as auto-correlações amostrais do histórico por limites inferiores de seus intervalos de confiança tende a fornecer modelagens com redução na denominada “propagação da tendência hidrológica”, conforme medida pela “curva de propagação do delta”. Para o intervalo de confiança de um desvio padrão, esta redução é sensível.

O procedimento sugerido na nova proposta do ONS produz políticas de operação 10 a 20% mais caras do que as políticas obtidas pelo procedimento atual, conforme avaliado pelo custo total de operação avaliado nas simulações com 2000 cenários gerados pelo modelo PAR(p) tradicional.

Os experimentos realizados com os PMOs de janeiro e julho de 2008 indicaram ainda que as reduções na “curva de propagação do delta” obtidos com o modelo PAR(p) ajustado trocando as auto-correlações amostrais do histórico por

limites inferiores de seus intervalos de confiança de um desvio-padrão não se refletem em reduções nas flutuações mensais de CMO quando o modelo ajustado desta forma é utilizado para obter a política de operação. Pelo contrário, na avaliação com simulações com 2000 cenários gerados pelo modelo PAR(p) tradicional, observou-se um aumento da flutuação ao longo do horizonte dos CMOs médios mensais. Considerando como métrica de volatilidade o desvio padrão das variações absolutas de CMO de meses consecutivos, a política de operação obtida com o modelo PAR(p) ajustado conforme sugerido pelo ONS apresentou uma volatilidade 20% a 40% maior do que a política de operação tradicional. Considerando como métrica de volatilidade a média das variações absolutas de CMO de meses consecutivos, as duas políticas não se diferenciam.

Em resumo, os estudos realizados forneceram resultados não promissores para este caminho de obtenção de política de operação com redução da volatilidade dos CMOs.

## **10 Nota Técnica nº 42 - Aplicação de Técnicas de Agregação na Geração de Cenários Hidrológicos para o Planejamento de Médio Prazo (Novembro 2008)**

### **10.1 Introdução**

O modelo NEWAVE, utilizado no planejamento da operação de médio prazo, 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. Essa estratégia é definida por uma política de operação ótima representada por uma função de custo futuro estimada utilizando a programação dinâmica dual estocástica. Além disso, o parque hidrelétrico é representado de forma agregada e a estocasticidade das afluições é representada por um número muito grande de cenários hidrológicos gerados sinteticamente por um modelo estocástico periódico auto-regressivo de ordem  $p$ .

A incerteza referente às afluições é considerada explicitamente no cálculo da função de custo futuro bem como 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  $20^{120}$ ), 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 escolhida utilizando o método de Monte-Carlo clássico que usa a amostragem aleatória simples.

Nesta Nota Técnica é descrito um método para a definição da sub-árvore a ser visitada durante o processo do cálculo da estratégia ótima de operação com o intuito de tornar mais robusto os resultados obtidos por esta 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. Para tanto se propõe que sejam aplicados ao modelo de geração de cenários hidrológicos multivariados técnicas estatísticas multivariadas capazes de elaborar critérios que possibilitam agrupar objetos similares em determinados grupos (técnicas de agregação). Estas técnicas podem ser reunidas sob o nome genérico de Análise de Conglomerados.

Usando as técnicas de agregação pretende-se escolher um conjunto representativo de cenários hidrológicos a partir de um grande número de cenários, reduzindo a variação nos resultados do modelo de médio prazo do planejamento da operação, sem deixar de representar de forma adequada o processo estocástico das afluições.

## **10.2 Análise de Conglomerados**

### **10.2.1 Considerações Gerais**

A Análise de Conglomerados é usada para reduzir uma grande massa de dados, na medida em que possibilita a partição/classificação dos dados em um número menor de grupos. Também é utilizada para desenvolver hipóteses a respeito da natureza dos dados ou para examinar hipóteses previamente estabelecidas. Representa uma poderosa ferramenta com aplicações em diversos problemas de formação de grupos. Elas podem ser empregadas, por exemplo, para identificar padrões similares de demanda de energia elétrica, na construção de segmentos de mercados, para agrupar programas de TV em tipos similares de acordo com tendências registradas de audiência, etc.

A Análise de Conglomerados tem grande aplicação na pesquisa científica em diversas áreas do conhecimento. Na literatura existem vários trabalhos que utilizam técnicas de agregação. Na linha de estudos elétricos pode-se citar trabalhos que empregam as técnicas de agregação para na construção da árvore de cenários hidrológicos para o planejamento da operação de curto-prazo (JARDIM et al, 2002) e na caracterização de curvas de carga (VELASQUEZ et al., 2001). Na área das Ciências da Computação, a Análise de Conglomerados está sendo amplamente utilizada para a classificação e comparação de documentos na Internet (STEINBACH et al., 2000). As Ciências Sociais também a utilizam para a realização de diversos estudos como os citados em ALDENDERFER e BLASHFIELD (1984). Existem ainda outras aplicações nas áreas de Ecologia (VALENTIN, 2000), Marketing (ZIKMUND,

1999) e Finanças (FARREL, 1997). Em HARTIGAN (1975) são mostrados diversos trabalhos em áreas distintas que empregam as técnicas de agregação.

Segundo JOHNSON e WICHERN (1998) a utilização de técnicas de agregação como procedimento exploratório é importante para o estudo da natureza complexa das inter-relações multivariáveis, indicando que a formulação de uma estrutura de agrupamentos naturais a partir de dados observados pode produzir meios informais para avaliar dimensionalidade, identificar casos marginais e sugestões de hipóteses sobre correlações. A Análise de Conglomerados difere dos métodos tradicionais de classificação porque, enquanto estes procuram associar novos itens a classes pré-determinadas, no caso de agrupamento não há nenhum conhecimento prévio acerca da estrutura de agrupamentos, e a divisão em grupos baseia-se unicamente nas similaridades ou diferenças entre os objetos, observados através dos dados coletados.

A utilização das técnicas de agregação na pesquisa científica vem aumentando devido à disponibilidade de computadores mais capazes e poderosos. Em geral, para um conjunto de objetos que se deseja agrupar, o número de estrutura de grupos possíveis é extremamente alto, de modo que se devem procurar algoritmos que levem em consideração algumas restrições prévias como forma de se reduzir o esforço matemático e determinar uma estrutura que seja satisfatória segundo um critério escolhido.

Tendo pesquisado alguns exemplos de aplicação de técnicas de agregação ALDEANDERFER e BLASHFIELD (1984) afirmam que, apesar das diferenças quanto a objetivos, tipos de dados e métodos usados em cada experiência, os cinco procedimentos básicos necessários para caracterizar todos os estudos de Análise de Conglomerados são:

- 1 Seleção de uma amostra de objetos que deverão ser agrupados
- 2 Definição de um conjunto de variáveis que serão medidas para todos os objetos da amostra
- 3 Cálculo das similaridades entre os objetos
- 4 Utilização de um método de Análise de Conglomerados para gerar grupos de objetos similares
- 5 Validação da estrutura dos grupos resultantes

### 10.3 Medidas de Similaridade

A escolha de um critério que quantifique o grau de associação entre os objetos ou variáveis tem um papel crucial nos estudos que utilizam a Análise de Conglomerados. Esta medida é chamada de coeficiente de similaridade e pode ser classificada como medida de similaridade ou de dissimilaridade. Na primeira, quanto maior for o valor da medida, mais similares são os objetos, enquanto que na segunda quanto menor o valor observado, mais parecidos são os objetos. De modo geral, medidas de dissimilaridade podem ser convertidas em medidas de similaridade através de uma relação inversa.

Apesar da aparente simplicidade, o conceito de similaridade e, especialmente, os procedimentos usados para medir a similaridade estão longe de serem simples. A similaridade entre os objetos pode ser medida de várias maneiras, dentre as quais as três que mais se destacam são as medidas de correlação, as medidas de distância, e as medidas de associação. As duas primeiras medidas requerem dados quantitativos enquanto que a última trata de dados qualitativos.

As medidas de distância representam a similaridade como a proximidade de um objeto a outro através de suas variáveis. Como é a medida mais intuitiva, as medidas de distâncias se tornaram as mais difundidas e utilizadas. Na verdade, medidas de distância representam uma medida de não similaridade, pois quanto maior a distância entre dois objetos maior a diferença entre eles. A distância é então convertida em uma medida de similaridade através do uso de uma relação inversa.

A estimação quantitativa da similaridade tem sido dominada pelo conceito de métrica. Objetos são representados como pontos no espaço e a similaridade entre eles é medida através da distância entre os pontos. A dimensionalidade do espaço é determinada pelo número de atributos (variáveis) usados para descrever os objetos, por exemplo,  $R_p$  caso os objetos sejam descritos por  $P$  variáveis.

Entre uma das mais populares representações de distância está a distância Euclidiana. Sejam  $X_i = (x_{i1}, \dots, x_{ip})$  e  $X_j = (x_{j1}, \dots, x_{jp})$  dois objetos caracterizados por  $P$  atributos, então a distância Euclidiana entre os dois objetos é definida como em (1a e 1b).

$$d_{i,j} = \sqrt{\sum_{r=1}^p (x_{ir} - x_{jr})^2} \quad 1a)$$

ou de forma matricial:

$$d_{i,j} = \sqrt{(X_i - X_j)'(X_i - X_j)} \quad 1b)$$

onde  $d_{i,j}$  é a distância entre os objetos  $i$  e  $j$ ,  $X_i$  é  $i$ -ésimo objeto e  $x_{ir}$  é o valor da  $r$ -ésima variável do  $i$ -ésimo objeto.

Uma outra importante métrica é a distância de Mahalanobis, também conhecida como distância generalizada. Esta métrica é definida como em (2).

$$d_{i,j} = \sqrt{(X_i - X_j)' \Sigma^{-1} (X_i - X_j)} \quad 2)$$

onde  $\Sigma^{-1} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & & \sigma_{2p} \\ \vdots & & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}^{-1}$  é a inversa da matriz de covariância.

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.

Muitas medidas de distância são sensíveis a variações na escala ou na magnitude entre as variáveis. A forma mais comum de padronização é a conversão de cada variável em valores padrão, subtraindo-se pelo valor médio (3) e dividindo-se pelo seu respectivo desvio padrão (4). A transformação (5) resulta em uma variável com média zero e desvio igual a 1. Também elimina a influência introduzida pelo uso de diferentes escalas nas variáveis usadas na análise. Não existe diferença nos valores padrão quando a escala é alterada.

$$\mu_r = \frac{1}{N} \sum_{i=1}^N x_{ir} \quad 3)$$

onde  $N$  é o número de objetos.

$$\sigma_r = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{ir} - \mu_r)^2} \quad 4)$$

$$y_{ir} = \frac{x_{ir} - \mu_r}{\sigma_r}$$

5)

As distâncias entre dois objetos podem ser organizadas na forma de uma matriz, conhecida como matriz de distâncias ou matriz de similaridade, ilustrada na Figura 1. A matriz de similaridade é uma matriz simétrica, de ordem igual ao número de objetos N, onde o elemento  $d_{i,j}$  é a medida de distância entre os objetos i e j.

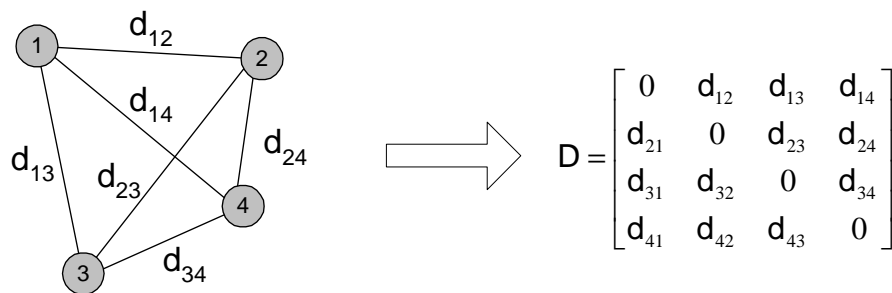


Figura 1 – Matriz de Similaridades

## 10.4 Métodos de Agrupamento

O principal objetivo quando se usa a Análise de Conglomerados é encontrar grupos de objetos similares em um conjunto de dados de tal forma que as variâncias entre os grupos seja máxima, e dentro deles, mínima. Considerando-se a enorme dificuldade em examinar todas as formas de agrupamentos possíveis, foram propostos vários algoritmos que promovem a divisão de objetos em grupos sem a necessidade de testar todas as configurações.

As técnicas de agregação constituem um meio para a redução da dimensionalidade de um conjunto de dados, pois se as classes obtidas forem internamente homogêneas, pode-se associar a cada classe um objeto típico, em geral a média dos objetos da classe, e assim, ao invés de analisar todo conjunto de dados, pode-se analisar apenas um pequeno número de objetos típicos, que capturam a maior parte da diversidade, ou melhor, da variância de todo conjunto.

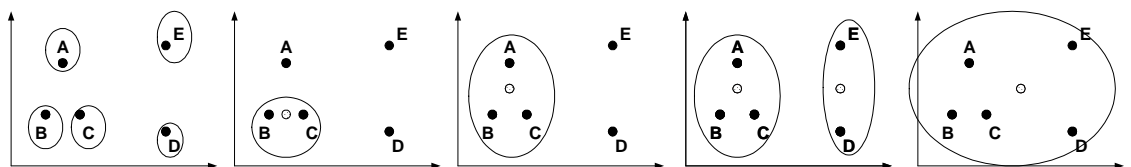
Os algoritmos mais comumente utilizados para problemas de agregação podem ser classificados em duas categorias: (1) métodos hierárquicos e (2) métodos não hierárquicos.



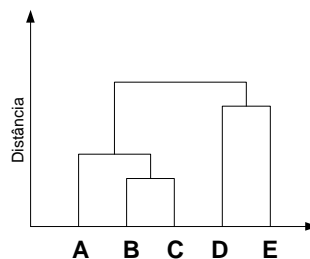
### 10.4.1 Métodos Hierárquicos

As técnicas hierárquicas podem ser aglomerativas ou divisivas. Nos métodos aglomerativos, os objetos individuais são agrupados de acordo com suas similaridades, enquanto que os métodos divisivos partem de um único grupo de objetos que é sucessivamente dividido até que cada subgrupo contenha somente um objeto.

Os resultados de ambos podem ser apresentados graficamente na forma de um diagrama bidimensional denominado dendograma, que ilustra as fusões ou divisões realizados em níveis sucessivos. A Figura 2 mostra o processo aglomerativo sendo aplicado a 5 objetos (A,B,C,D e E). A cada etapa é mostrado o centróide dos grupos que vão se formando. Na etapa inicial todos os objetos estão sós em um grupo e na etapa final todos os objetos estão reunidos no mesmo grupo. O dendograma resultante desta seqüência de fusões é mostrado na Figura 3.



**Figura 2** – Exemplo ilustrativo do processo aglomerativo



**Figura 3** – Dendograma

O método é denominado hierárquico porque uma vez que dois objetos ou grupos são agrupados/separados, estes permanecem juntos/separados até o final da agregação, isto é, não há realocação dos objetos. Isto é uma desvantagem do método, pois se algum objeto for incorretamente agrupado em um estágio anterior não há possibilidade de realocá-lo em um estágio posterior. Uma outra desvantagem é a necessidade da construção e armazenamento da matriz de similaridade. A construção desta matriz pode representar uma limitação para a maioria das aplicações em

microcomputadores, por este motivo os métodos hierárquicos não são indicados para conjuntos grandes de dados.

Mais detalhes sobre métodos hierárquicos, seus algoritmos e características podem ser encontrados em HARTIGAN (1975), ANDERBERGER (1973), HAIR JR. et al. (1998), DURAN e ODELL (1970) e JOHNSON e WICHERN (1998).

#### **10.4.2 Métodos Não Hierárquicos**

Nos métodos não hierárquicos os objetos são divididos em um número de grupos previamente fixado. Estes grupos são formados de modo que duas premissas básicas sejam atendidas: coesão interna e isolamento dos grupos.

Diferentemente dos métodos hierárquicos, as técnicas não hierárquicas não exigem a determinação e o armazenamento da matriz de similaridade, cuja ordem depende do número de objetos a ser analisados. Por este motivo, os métodos não hierárquicos são computacionalmente mais eficientes quando se trabalha com um grande conjunto de dados.

O caminho mais intuitivo para encontrar a melhor partição é checar todas as possíveis partições do conjunto de dados, porém o número de possibilidades é muito grande, assintoticamente de ordem de  $KN-1$ , onde  $K$  é número de grupos e  $N$  o número de objetos que se deseja agrupar. Para resolver um problema de pequeno porte com 20 objetos e 3 grupos, é preciso investigar cerca de um bilhão de possíveis partições únicas. Dado a inviabilidade da análise de todas as partições possíveis, pesquisadores desenvolveram vários procedimentos heurísticos que investigam algumas partições com o intuito de encontrar a melhor partição, ou uma alternativa que seja quase ótima.

Dentre os procedimentos heurísticos desenvolvidos, o mais conhecido é o método K-Means. Este método, com pequenas variações, é um dos mais usados na Análise de Conglomerados quando se tem muitos objetos.

Mais informações sobre métodos não hierárquicos, suas características e sua utilização são encontradas em HARTIGAN (1975), ANDERBERGER (1973), ALDENDERFER e BLASHFIELD (1984), HAIR JR. et al. (1998), JOHNSON e WICHERN (1998) e BOUROCHE e SAPORTA (1980).

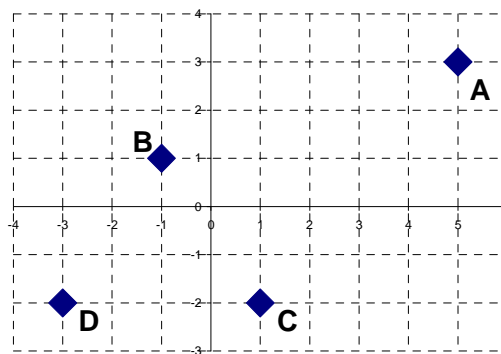
## 10.5 Método K-Means

O primeiro passo deste método é formar uma partição inicial aleatória no conjunto de dados. O número de grupos deve ser estabelecido previamente. O próximo passo é o cálculo dos centróides destes grupos. Então, a distância entre cada objeto e cada centróide é calculada. Os objetos são realocados para o grupo que tiver o centróide mais próximo (menor distância). Este último passo é repetido até que não haja mais realocações de objetos. Vale a pena lembrar que toda vez que um objeto for realocado os centróides devem ser recalculados. O algoritmo K-Means pode ser resumido nos seguintes passos:

- 1 Divida os  $N$  objetos em  $K$  agrupamentos através de uma partição inicial ou especificação de  $K$  centróides iniciais;
- 2 Realoque um objeto para o grupo cujo centróide é o mais próximo deste objeto e recalcule o centróide do grupo que recebeu e que perdeu o objeto;
- 3 Repita o passo 2 até que não haja mais realocações de objetos de um grupo para outro.

Com o intuito aperfeiçoar, tornar mais rápido e mais eficiente o algoritmo apresentado, alguns procedimentos podem ser modificados, gerando assim variações deste método. A inicialização dos grupos pode ser feita de forma aleatória através do sorteio de pontos (objetos) para serem usados como semente inicial dos grupos ou pela partição aleatória do conjunto de dados. Os pontos sorteados podem ser sorteados de dentro do conjunto de dados ou não. Estes pontos também podem ser escolhidos um a um pelo especialista ou retirados de forma programada de dentro do conjunto de dados. Outra modificação que pode ser realizada é quanto à atualização dos centróides durante processo de realocação dos objetos. Esta atualização pode ser feita a cada vez que um objeto for realocado ou somente quando todos os objetos forem realocados. A primeira alternativa é a mais utilizada.

Para ilustrar como funciona o algoritmo do método não hierárquico descrito anteriormente, é utilizado um exemplo extraído de JOHNSON e WICHERN (1998). Considere um conjunto com 4 objetos ( $N=4$ ) descritos por 2 variáveis ( $P=2$ ),  $x_1$  e  $x_2$ , onde se procura formar dois grupos ( $K=2$ ). Os objetos são apresentados na Figura 4 e as suas coordenadas, na Tabela 1.

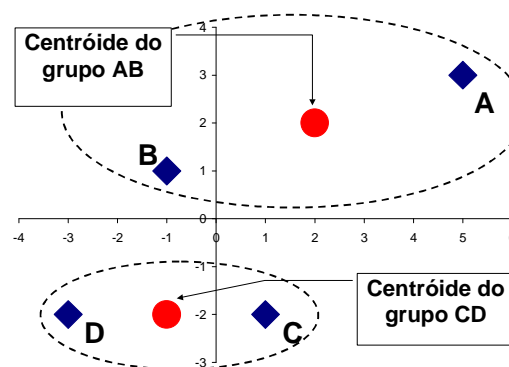


**Figura 4** – Configuração dos 4 objetos

**Tabela 1** – Coordenadas dos 4 objetos

Objetos	Variáveis	
	$x_1$	$x_2$
A	5	3
B	-1	1
C	1	-2
D	-3	-2

Para inicializar o processo de agregação, os objetos foram particionados em dois grupos AB e CD, Figura 5, e a partir deles são calculadas as coordenadas dos centróides, Tabela 2.



**Figura 5** – Grupos AB e CD com os respectivos centróides

**Tabela 2** – Coordenadas dos centróides (1ª etapa)

Grupos	Variáveis	
	$x_1$	$x_2$
AB	$\frac{5 + (-1)}{2} = 2$	$\frac{3 + 1}{2} = 2$
CD	$\frac{1 + (-3)}{2} = -1$	$\frac{-2 + (-2)}{2} = -2$

Determinadas as coordenadas dos centróides, são calculadas as distâncias de cada objeto com relação aos centróides para verificar a necessidade de realocação. Neste exemplo é utilizada a distância Euclidiana.

Iniciando com o objeto A, se tem a seguinte distância:

$$D(A, (AB)) = \sqrt{(5 - 2)^2 + (3 - 2)^2} = 3.16$$

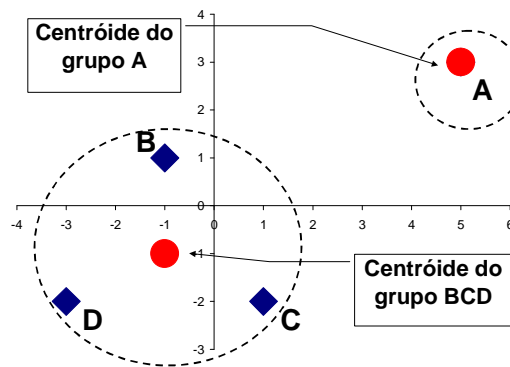
$$D(A, (CD)) = \sqrt{(5 - (-1))^2 + (3 - (-2))^2} = 7.81$$

Como a distância entre A e o centróide do grupo AB é menor que a distância entre A e o centróide do grupo CD, não há realocação, isto é, o objeto A permanece no grupo AB. Seguindo agora com o objeto B tem-se as seguintes distâncias ao quadrado:

$$D(B, (AB)) = \sqrt{(-1 - 2)^2 + (1 - 2)^2} = 3.16$$

$$D(B, (CD)) = \sqrt{(-1 - (-1))^2 + (1 - (-2))^2} = 3$$

Neste caso a distância entre B e o centróide do grupo CD é a menor distância, então o objeto B deve ser realocado para o grupo CD, Figura 6, e as novas coordenadas dos centróides são apresentadas na Tabela 3:



**Figura 6** – Grupos A e BCD com os respectivos centróides

**Tabela 3** – Coordenadas dos centróides (2ª etapa)

Grupos	Variáveis	
	$x_1$	$x_2$
A	5	3
BCD	$\frac{-1+1+(-3)}{3} = -1$	$\frac{1+(-2)+(-2)}{3} = -1$

Continuando com os objetos C e D verifica-se que não há realocação.

A Tabela 4 mostra as distâncias de cada objeto com relação ao centróides. Em todos os casos a distância do objeto ao centróide do grupo onde está alocado é sempre a menor distância, assim não há realocação de objetos e a execução do algoritmo pode ser finalizada.

**Tabela 4** – Distância entre objetos e centróides

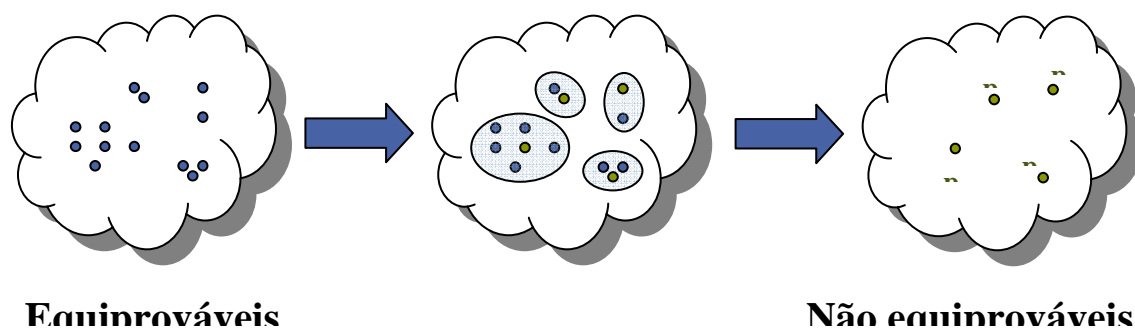
Grupos	Distância Euclidiana			
	Objetos			
	A	B	C	D
A	0	6.32	6.40	9.43
BCD	7.21	2	2.24	2.24

## 10.6 Aplicação no Modelo de Geração de Cenários

Os cenários de energia natural afluyente, que são utilizados durante as simulações forward e backward do processo de definição da política ótima de operação, são obtidos através de um modelo auto-regressivo periódico de ordem  $p$ , PAR( $p$ ), que modela a aflluência de um mês como sendo função das afluências dos  $p$  meses anteriores (MACEIRA e MERCIO, 1997). A amostra de ruídos aleatórios utilizada pelo modelo PAR( $p$ ), é obtida atualmente através de amostragem aleatória simples.

O método proposto nesta Nota Técnica consiste em aplicar as técnicas de agregação no procedimento de geração dos cenários de energia natural afluyente das simulações forward e backward. Neste caso, as técnicas de agregação são empregadas para a geração da amostra de ruídos aleatórios multivariados que é utilizada pelo modelo PAR( $p$ ).

Inicialmente será gerada uma amostra muito grande utilizando a amostragem aleatória simples, onde cada objeto é um vetor de ruídos aleatórios (um ruído para cada subsistema considerado na configuração). Os vetores de ruídos que compõem essa amostra são equiprováveis. Logo após é realizada a agregação desses objetos de forma a reduzir a dimensionalidade da amostra original. A Figura 7 ilustra o procedimento proposto.



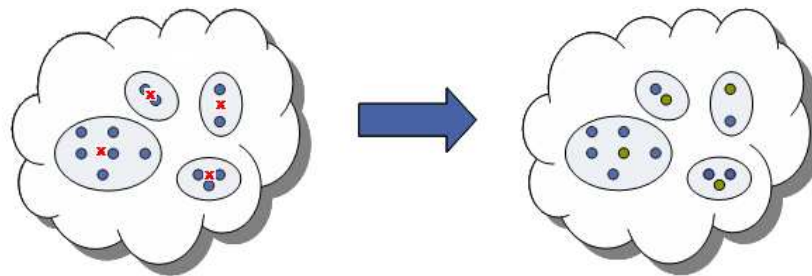
**Figura 7**– Aplicação do Procedimento de Agregação

O método de agregação escolhido é o método não hierárquico K-Means, pois o tamanho da amostra que é fornecida para o processo de agregação é grande. Os métodos não hierárquicos são ideais para trabalhar com grandes conjuntos de dados, pois não requerem o cálculo da matriz de similaridade.

O processo de agregação é inicializado através do sorteio aleatório de pontos iniciais para representar os centróides dos grupos. Estes pontos iniciais são

objetos do conjunto de entrada, logo são vetores de ruídos pertencentes à amostra original. Desta maneira, pode-se garantir que nenhum grupo ficará vazio.

Nos passos seguintes até a convergência do processo de agregação, o centróide dos grupos será o ponto médio destes grupos. Após a convergência do processo de agregação, o centróide dos grupos será o objeto mais próximo do ponto médio deste grupo. A Figura 8 ilustra como é escolhido o representante de cada grupo formado, o ponto médio dos grupos está assinalado com um **x**.



**Figura 8**– Escolha do Objeto Representativo

Os vetores de ruídos multivariados resultantes do processo de agregação não são mais equiprováveis. A probabilidade dos objetos representantes irá refletir a representatividade do grupo em que ele se encontra. A probabilidade  $P_k$  associada ao grupo  $k$  é calculada como em (6):

$$P_k = \frac{NO_k}{N} \quad 6)$$

onde  $N$  é tamanho da amostra antes do processo de agregação e  $NO_k$  é número de objetos alocados no grupo  $k$ .

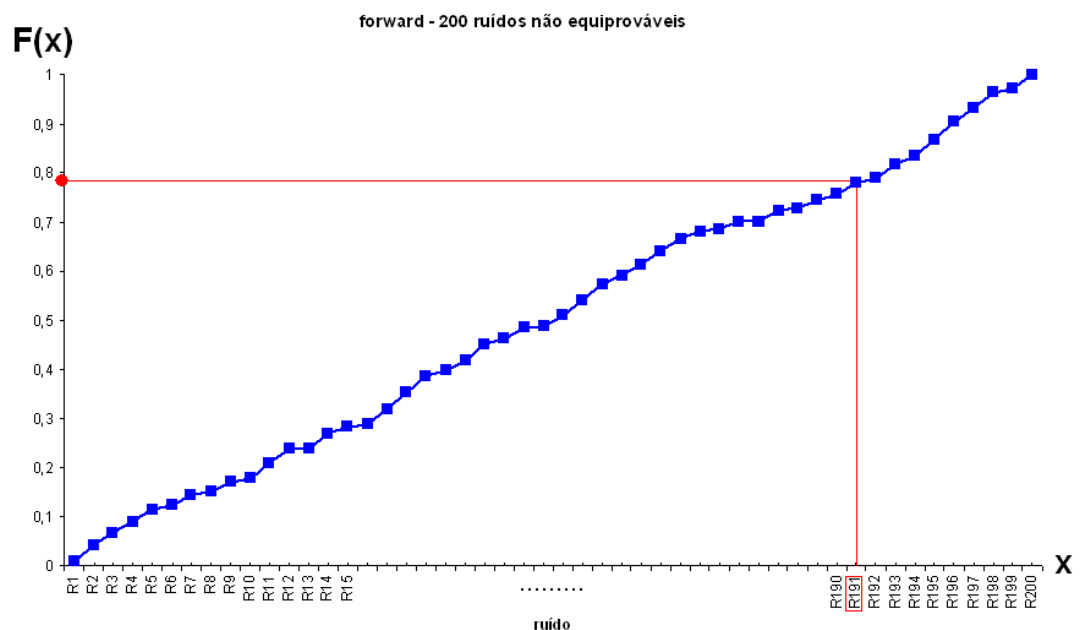
A probabilidade do cenário de energia natural afluyente é a igual a probabilidade do vetor de ruídos resultantes a partir do qual ele foi gerado.

O processo de agregação pode ser incorporado tanto no processo de construção da árvore do passo forward quanto do passo backward. Porém, estudos exploratórios com cenários hidrológicos não equiprováveis no passo forward mostraram resultados muito instáveis. Neste sentido, uma nova amostra de ruídos equiprováveis será construída para o passo forward a partir da amostra resultante do processo de agregação através de um sorteio condicionado.



O sorteio condicionado é baseado no teorema da transformação inversa: “Se  $X$  é uma variável aleatória de distribuição acumulada  $F(x)$ , então a variável aleatória  $Y=F(x)$  tem distribuição uniforme  $(0,1)$ ”.

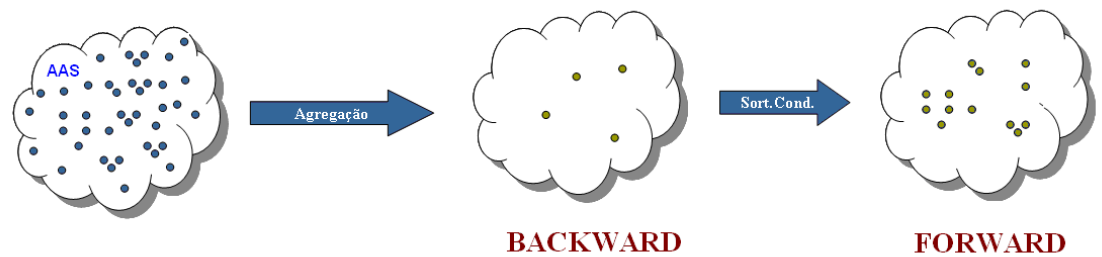
Inicialmente é calculada a distribuição acumulada empírica da amostra de ruídos não equiprováveis resultante do processo de agregação para o passo forward. Logo em seguida é sorteado um número aleatório uniforme  $[0,1]$ , e a partir de uma consulta à função acumulada é identificado o ruído associado aquele valor sorteado. No exemplo da Figura 9, o número aleatório uniforme sorteado foi 0,89. O ruído cuja função acumulada corresponde a 0,89 é o ruído número 191.



**Figura 9** – Sorteio Condicionado

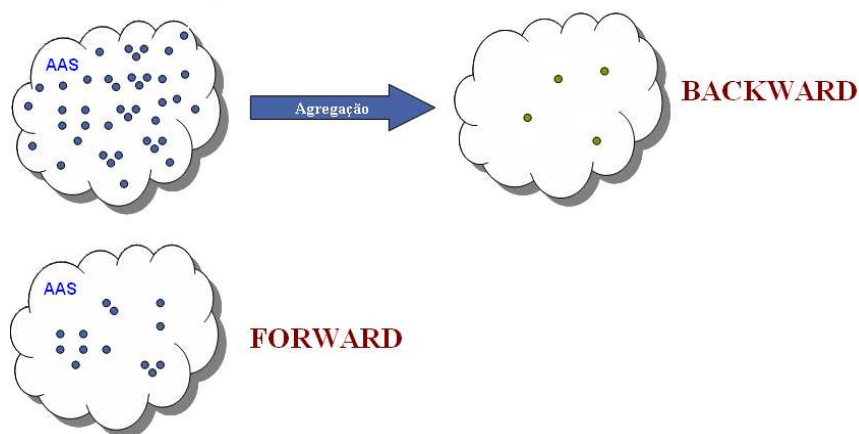
O sorteio condicionado é realizado tantas vezes quanto for o tamanho da amostra de ruídos do passo forward. Os ruídos que compõem essa nova amostra de ruídos são equiprováveis. Logo, os cenários hidrológicos do passo forward construídos a partir dessa amostras de ruídos também são equiprováveis.

Foram estudadas cinco alternativas de aplicação do processo de agregação na construção das árvores de cenários hidrológicos. Na primeira alternativa, chamada de opção 0, o processo de agregação é aplicado para definir a amostra de ruídos do passo backward. A partir dessa amostra é realizado um sorteio condicionado para definir a amostra de ruídos a ser utilizada na construção dos cenários hidrológicos a serem utilizados pelo passo forward. Esse procedimento é ilustrado na Figura 10.



**Figura 10** – Opção 0

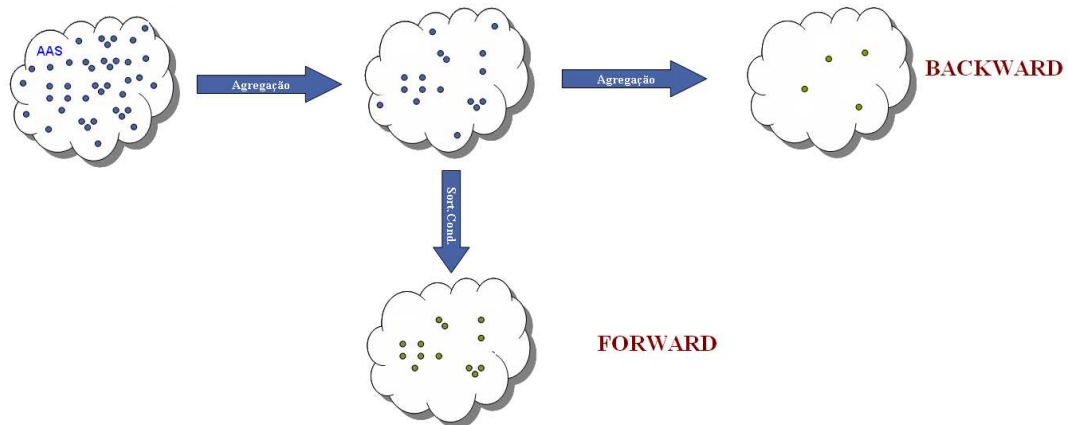
Na segunda alternativa, chamada de opção 1, o processo de agregação é aplicado apenas na construção da árvore de cenários do passo backward, de acordo com o descrito anteriormente. A árvore de cenários do passo forward é obtida através de amostragem aleatória simples (AAS), Figura 11.



**Figura 11** – Opção 1

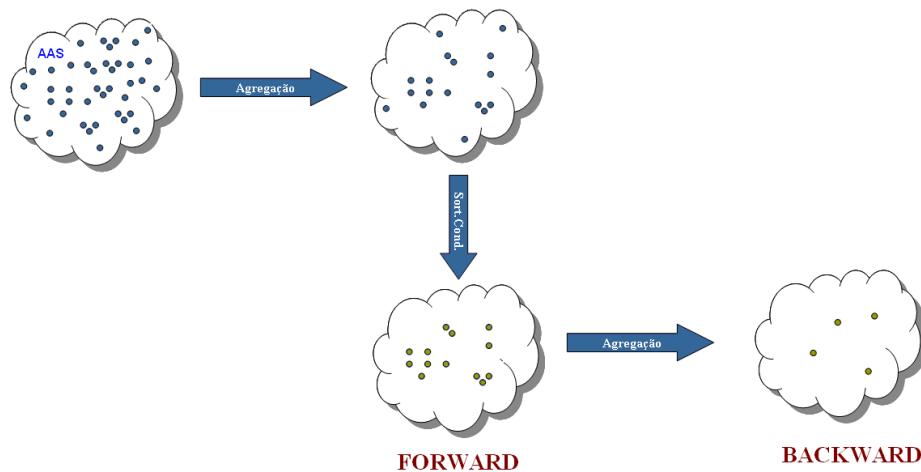
Os ruídos aleatórios multivariados pertencentes à amostra de ruídos que será utilizada para a construção dos cenários da backward são não equiprováveis. A probabilidade do cenário de energia natural afluente da simulação backward é igual à probabilidade do ruído multivariado a partir do qual ele foi gerado. Já os cenários hidrológicos da simulação forward são equiprováveis.

Na terceira alternativa, Figura 12, o processo de agregação é aplicado para obter a amostra de ruídos do passo forward. A árvore de cenários do passo backward é obtida aplicando-se o processo de agregação na amostra de ruídos construída para o passo forward. Neste caso, o algoritmo de agregação deve levar em conta que os objetos da amostra a ser agregada são não equiprováveis. Uma amostra com objetos equiprováveis é construída para o passo forward através de sorteio condicionado. Essa alternativa é chamada **opção 2**.



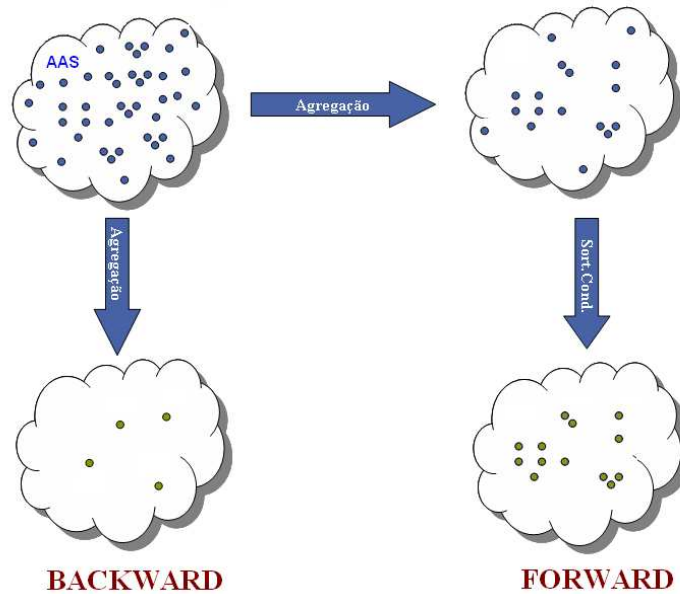
**Figura 12 – Opção 2**

Na quarta alternativa (**opção 3**), Figura 13, o processo de agregação é aplicado para obter a amostra de ruídos do passo forward. Uma amostra com objetos equiprováveis é construída para o passo forward através de sorteio condicionado. A árvore de cenários do passo backward é obtida aplicando-se o processo de agregação na amostra de ruídos construída para o passo forward, após o sorteio condicionado.



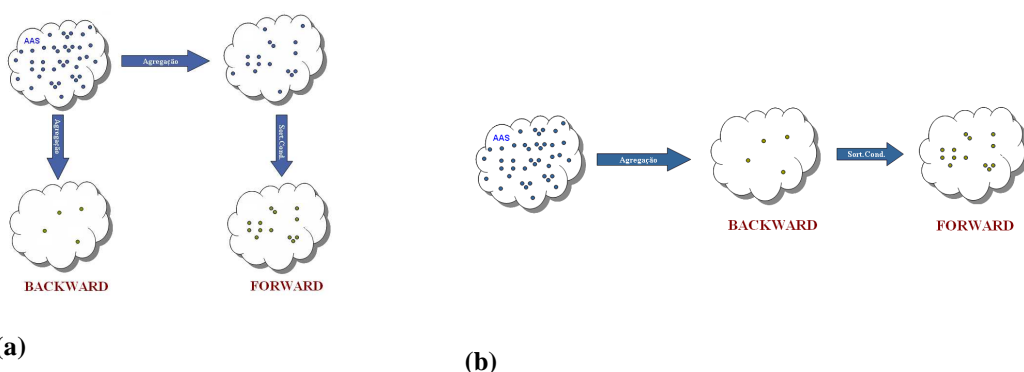
**Figura 13 – Opção 3**

Na quinta alternativa, o processo de agregação é aplicado na construção da árvore de cenários do passo forward. A árvore de cenários do passo backward é obtida aplicando-se novamente o processo de agregação na amostra originalmente gerada. Novamente, uma amostra com objetos equiprováveis é construída para o passo forward através de sorteio condicionado. A opção 4 é ilustrada na Figura 14.



**Figura 14 – Opção 4**

A amostra de cenários hidrológicos do passo forward gerada para o primeiro período do horizonte de estudo utilizando a opção 0 possui no máximo  $nleq$  valores distintos (onde  $nleq$  é o número de cenários da amostra backward). Para aprimorar a representatividade dos cenários forward utilizados no primeiro período de estudo é proposto um aperfeiçoamento na opção 0. Para o primeiro período, a amostra forward passa a ser construída conforme descrito na opção 4 e para os demais períodos a amostra é construída de acordo com o descrito na opção 0 original, Figura 15.



**Figura 15 – Evolução da Opção 0. (a) 1º período (b) demais períodos**

## 10.7 Cortes de Benders

No NEWAVE, a estratégia 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 hidrológica). Para cada estado, o corte da função de custo futuro corresponde a uma média calculada para um conjunto de afluências utilizadas durante o cálculo da política de operação no passo backward, Figura 16.

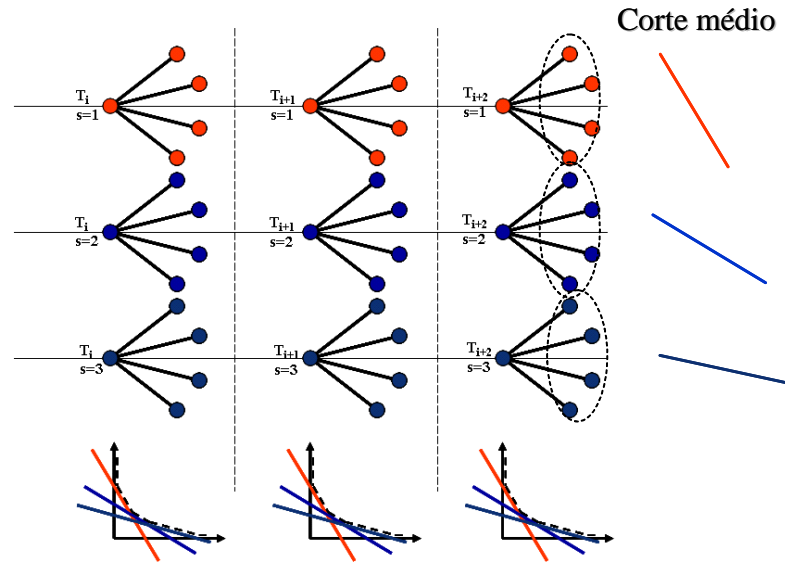


Figura 16 – Construção da FCF

A aplicação das técnicas na agregação no processo de geração dos cenários hidrológicos utilizados durante a recursão backward, resulta em um conjunto de cenários não equiprováveis. Logo, o cálculo do corte médio a ser adicionado à função de custo futuro deve ser modificado de forma a levar em conta a probabilidade de cada cenário hidrológico do conjunto backward.

$$\overline{\pi}_V^{isim} = \sum_{i=1}^{NLEQ} \pi_V^{i,isim} * P_i \quad (7)$$

$$\overline{\pi}_{\Lambda_j}^{isim} = \sum_{i=1}^{NLEQ} \pi_{\Lambda_j}^{i,isim} * P_i \quad (8)$$

$j = 1, \dots,$

NARP

$$\overline{RHS}^{isim} = \sum_{i=1}^{NLEQ} RHS^{i,isim} * P_i \quad (9)$$

onde:

NLEQ é o número de aberturas (tamanho do conjunto de afluências utilizado na recursão backward);

NARP é a ordem do modelo PAR(p);

Pi é a probabilidade do i-ésimo cenário hidrológico do conjunto backward;

$\pi_V^{i, isim}$  é o coeficiente do corte de Benders associado ao estado armazenamento inicial do estágio, calculado no cenário forward isim e na i-ésima abertura;

$\overline{\pi_V^{isim}}$  é o coeficiente do corte de Benders médio associado ao estado armazenamento inicial do estágio, calculado no cenário forward isim;

$\pi_{A_j}^{i, isim}$  é o coeficiente do corte de Benders associado ao estado energia afluente passada do estágio t-j, calculado no cenário forward isim e na i-ésima abertura;

$\overline{\pi_{A_j}^{isim}}$  é o coeficiente do corte de Benders médio associado ao estado energia afluente passada do estágio t-j, calculado no cenário forward isim;

$RHS^{i, isim}$  é o termo independente do corte de Benders calculado no cenário forward isim e na i-ésima abertura;

$\overline{RHS^{isim}}$  é o termo independente do corte de Benders médio calculado no cenário forward isim.

## 10.8 Conclusões

Nesta Nota Técnica foram descritos procedimentos, baseados na análise de conglomerados, para a definição da sub-árvore a ser visitada durante o processo de cálculo da política ótima de operação com o intuito de tornar os resultados mais robustos, com relação a variações no número de cenários de simulação forward e backward, e com relação à amostra de cenários hidrológicos utilizada.

Foram propostas cinco alternativas para a geração dos cenários hidrológicos para os passos forward e backward. Todas as alternativas apresentadas utilizam técnicas estatísticas multivariadas capazes de elaborar critérios que

possibilitam agrupar objetos similares em determinados grupos (técnicas de agregação).

# 11 NOTA TÉCNICA nº 43 - Representação da Correlação Cruzada na Geração de Cenários de Afluências para o Planejamento da Operação Energética de Médio Prazo (Setembro 2009)

## 11.1 Introdução

No modelo GEVAZP [1], atualmente utilizado para representar as incertezas hidrológicas no planejamento da operação energética de médio prazo, as seqüências de energias naturais mensais afluentes aos subsistemas são modeladas como um processo autoregressivo periódico de ordem  $p$ . Nesta modelagem, a cada mês a energia natural afluyente (ENA) de cada subsistema é composta pela soma de uma parcela determinística, dada por uma combinação linear das  $p$  afluências passadas mais recentes mais uma constante (igual a média do mês), com uma parcela estocástica, dada por um ruído aleatório de valor esperado nulo. O ruído é concebido como um sorteio de uma distribuição lognormal três parâmetros multivariada (cada dimensão corresponde a um subsistema).

De acordo com [1], o terceiro parâmetro da distribuição lognormal tem o objetivo de se garantir que no processo de geração das séries sintéticas não ocorram valores negativos de afluências aos subsistemas. Este terceiro parâmetro, denominado deslocamento, explicita um limite inferior para os ruídos, oferecendo assim um instrumento ajustável a cada situação da parcela determinística do modelo e utilizada no módulo de geração para garantir energias naturais afluentes sintéticas estritamente positivas.

Uma análise detalhada das estatísticas de correlação entre as séries sintéticas de energias naturais afluentes de diferentes subsistemas (correlação cruzada) geradas pelo modelo GEVAZP interno ao programa NEWAVE [2], [3], revelou a existência de uma interação entre os limites inferiores da distribuição lognormal multivariada e a matriz de covariância amostral dos ruídos gerados, não atendendo a condição de independência entre a matriz de covariância dos ruídos e as afluências passadas. Como será observado na seção 3, quanto maior o deslocamento, maior o impacto na assimetria da distribuição lognormal 3 parâmetros e conseqüentemente, maior a dificuldade na reprodução da correlação cruzada.



No modelo GEVAZP a estimação dos parâmetros do modelo PARP(p) tem como objetivo a reprodução das estatísticas das séries observadas (série histórica) nas séries sintéticas. Assim, a constante (média do mês) e os coeficientes das combinações lineares da parcela determinística de cada subsistema a cada mês são estimados a partir do histórico pelo método dos momentos visando reproduzir médias, desvios-padrões e coeficientes de auto-correlação temporal até o lag p das ENA históricas de cada mês. Os valores esperados e variâncias de cada dimensão (subsistema) da distribuição multivariada dos ruídos são também obtidos pelo método dos momentos, novamente visando reproduzir médias e desvios padrões das ENA de cada mês.

No caso das covariâncias da distribuição dos ruídos, o objetivo é representar as correlações cruzadas observadas das séries de ENA históricas anuais nas séries de ruídos gerados. O efeito imediato da violação da condição de independência entre a matriz de covariância dos ruídos e as influências passadas é prejudicar esta representação. Esta Nota Técnica descreve uma metodologia desenvolvida para aprimorar o processo de sorteio de ruídos multivariados lognormais três parâmetros com o propósito de reduzir o impacto dos limites inferiores na matriz de covariância. Apresenta-se, uma revisão da modelagem matemática do modelo GEVAZP, seguida da análise da qualidade da geração onde se evidencia a dificuldade de representação das correlações cruzadas. Apresenta-se então o aprimoramento proposto e seus resultados.

## **11.2 Modelagem Matemática no GEVAZP**

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 [4] 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:

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

onde,  $Z_t$  é a série sazonal,  $m$  é o índice do período sazonal,  $\mu_m$  é a média sazonal,  $\sigma_m$  é o desvio-padrão sazonal,  $\phi_i^m$  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 [5]. 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.

Para garantir o atendimento à equação (2), foi ajustada uma distribuição Lognormal com três parâmetros aos resíduos mensais  $a_t$ , cujo terceiro parâmetro  $\Delta_t$  é conhecido como deslocamento. Conseqüentemente, a variável  $\xi_t$  tem distribuição Normal com média  $\mu_{\xi_t}$  e desvio-padrão igual a  $\sigma_{\xi_t}^2$ :

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

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

$$\mu_{\xi_t} = \frac{1}{2} \ln \left( \frac{\sigma_a^{2(m)}}{\theta^2 - \theta} \right)$$

(4)

$$\sigma_{\xi_t}^2 = \ln(\theta)$$

(5)

$$\Delta_t = \mu_a^{(m)} - \frac{\sigma_a^{(m)}}{\sqrt{(\theta - 1)}}$$

(6)

Fazendo  $\Delta_t$  igual ao lado direito da expressão (2), podemos escrever:

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

(7)

$\sigma_{\xi_t}^2$  pode ser diretamente obtido a partir das expressões (7) e (5).

Com o intuito de gerar ENA mensais multivariadas, assume-se que os resíduos espacialmente não correlacionados,  $\xi_t$ , podem ser transformados em resíduos espacialmente correlacionados,  $W_t$ , da seguinte forma:

$$W_t = D\xi_t$$

(8)

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

A matriz D é estimada utilizando:

$$DD^T = \hat{U}$$

(9)

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.

### 11.3 Estatísticas Médias e Desvio-Padrão

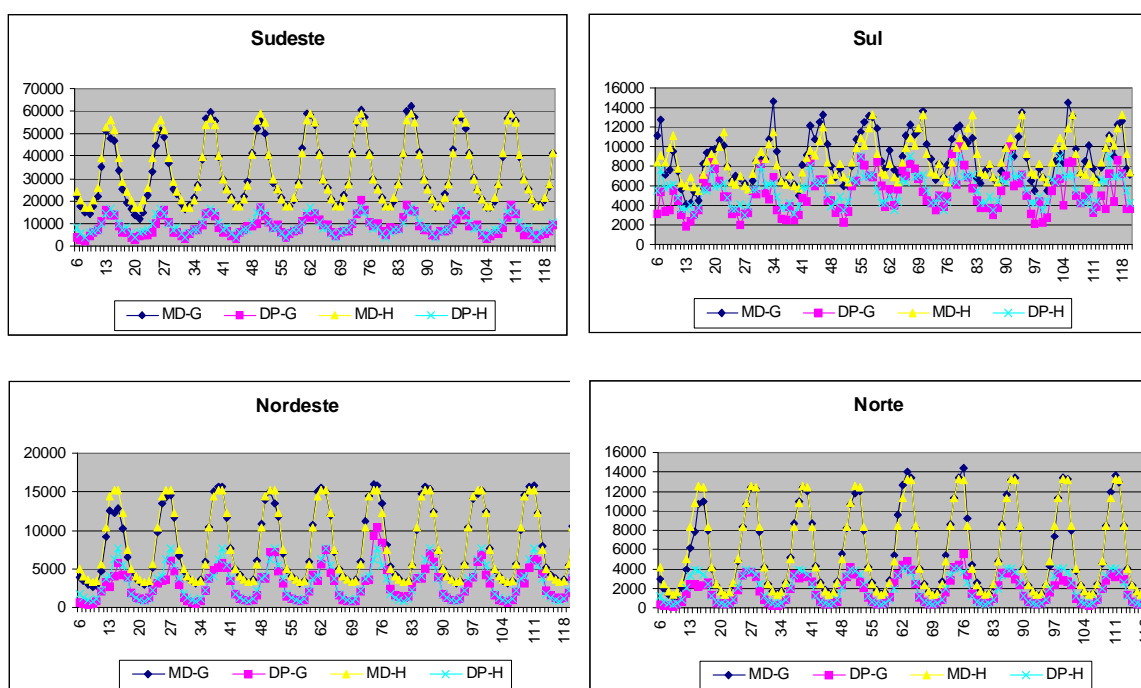
Na definição da política ótima de operação, o modelo NEWAVE utiliza dois tipos de cenários sintéticos de ENA gerados pelo modelo GEVAZP. Para a simulação

forward é gerada uma amostra de seqüências sintéticas de ENA para todo o horizonte do estudo, sendo normalmente utilizadas 200 seqüências. Já para a simulação backward, para cada mês e para cada seqüência sintética da simulação forward é gerada uma amostra de ENA condicionadas ao correspondente passado recente. Normalmente, utilizam-se amostras de 20 ENA condicionadas, chamadas de aberturas.

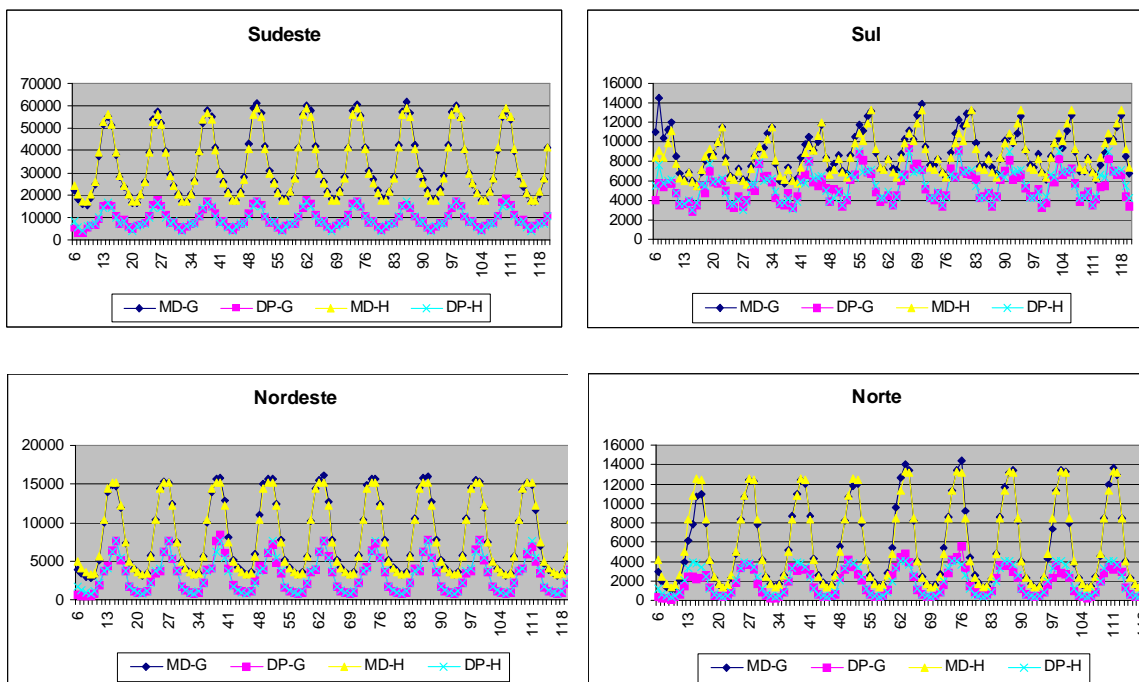
Para as análises a seguir é utilizado um caso baseado no PMO Junho de 2007. A versão utilizada do programa NEWAVE foi a versão 12e.

Para a análise da simulação forward foram comparadas as evoluções temporais da média e desvio-padrão das ENA dos cenários gerados com as correspondentes estatísticas da série histórica. Foram utilizadas amostras forward com 20, 200 e 2000 séries sintéticas. A Figura 1 ilustra as comparações para os subsistemas Sudeste, Sul, Nordeste e Norte.

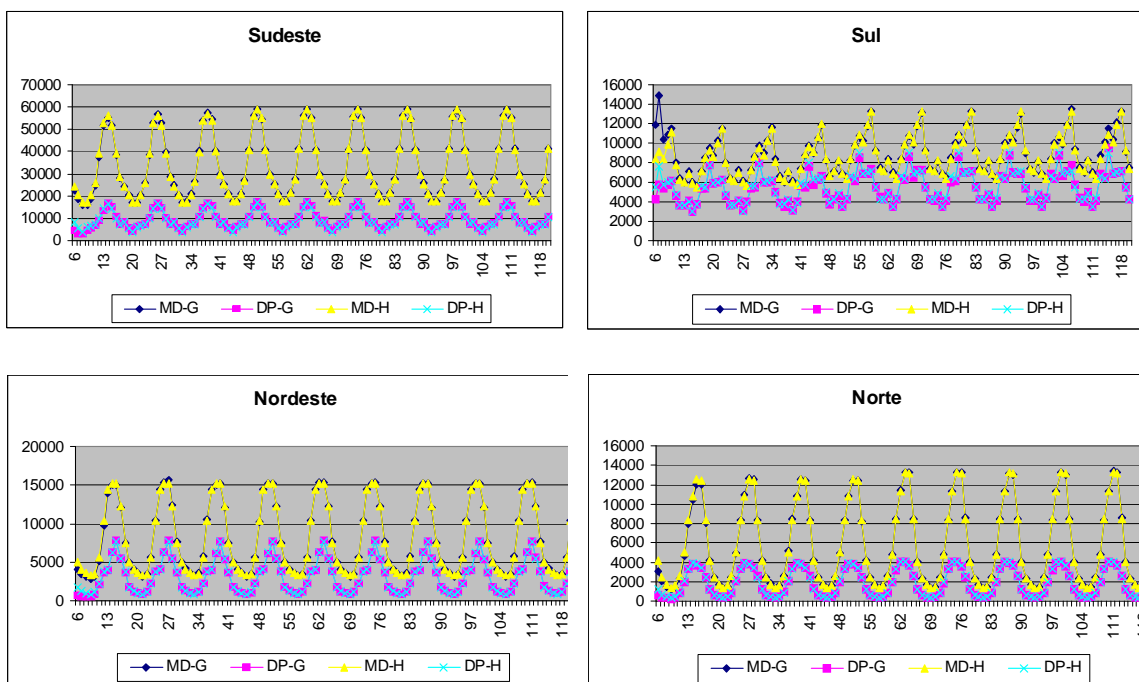
Pode-se observar que à medida que cresce o tamanho da amostra fica maior a aderência aos valores do histórico. Observa-se ainda o efeito do condicionamento ao passado recente nos meses iniciais do estudo.



Amostra com 20 Séries Sintéticas



Amostra com 200 Séries Sintéticas

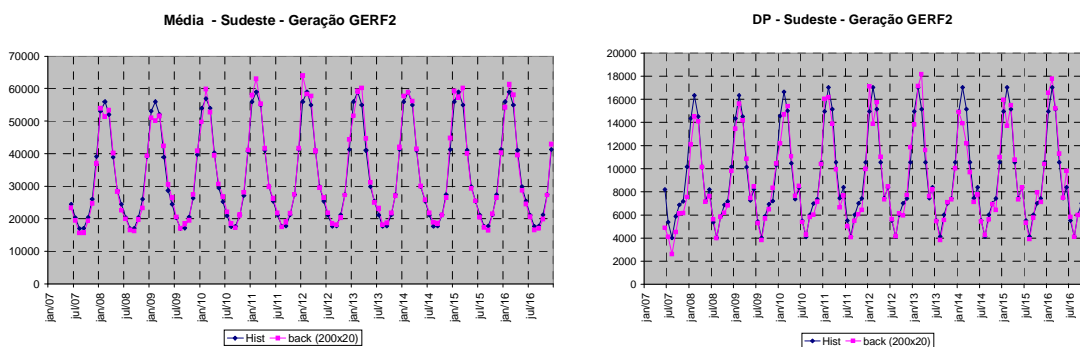


Amostra com 2000 Séries Sintéticas

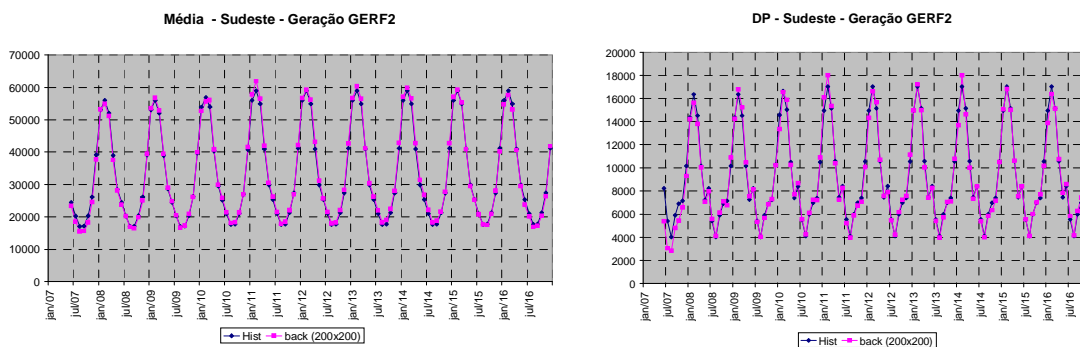
Figura 1 - Evolução temporal da Média e Desvio-Padrão das ENA

A seguir foram comparadas as evoluções temporais não condicionadas<sup>6</sup> de média e desvio-padrão das ENA dos cenários gerados para a simulação backward com as mesmas estatísticas da série histórica. Para configurar a análise, as médias e desvios-padrão das ENA geradas foram obtidos agrupando-se as aberturas de todas as seqüências forward. Foram consideradas 200 seqüências forward, tendo sido analisados os casos de 20 e 200 aberturas. A Figura 2 ilustra as comparações para os subsistemas Sudeste, Sul, Nordeste e Norte.

Pode-se observar que à medida que aumenta o número de aberturas, cresce a aderência aos valores do histórico. Novamente, observa-se o efeito do condicionamento ao passado recente nos meses iniciais do estudo.



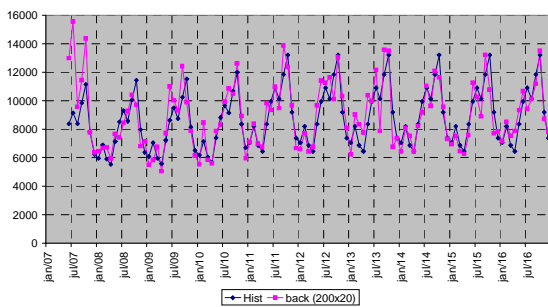
Sudeste - Caso de 20 Aberturas



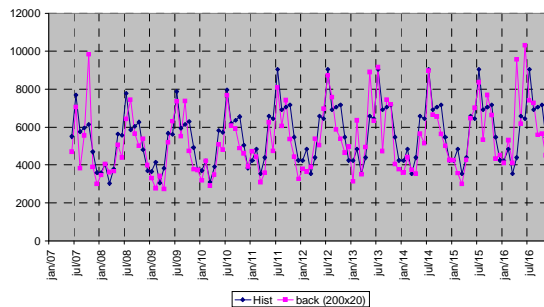
Sudeste - Caso de 200 Aberturas

<sup>6</sup> Para realização dos testes não condicionados, a cada período são considerados todos os cenários backward gerados, independentemente do estado de afluência passada

Média - Sul - Geração GERF2

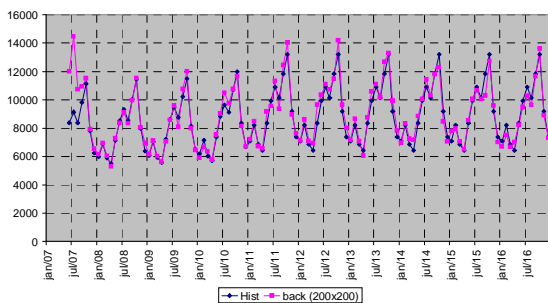


DP - Sul - Geração GERF2

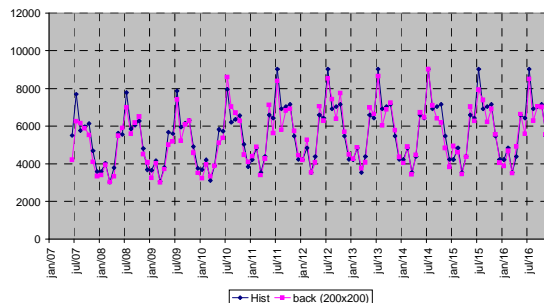


Sul - Caso de 20 Aberturas

Média - Sul - Geração GERF2

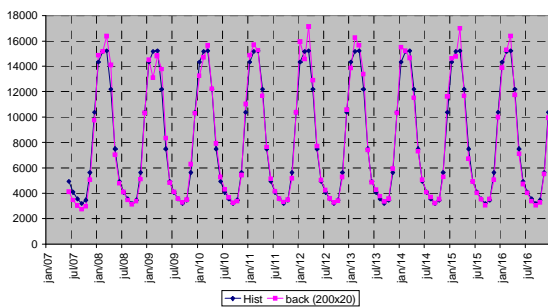


DP - Sul - Geração GERF2

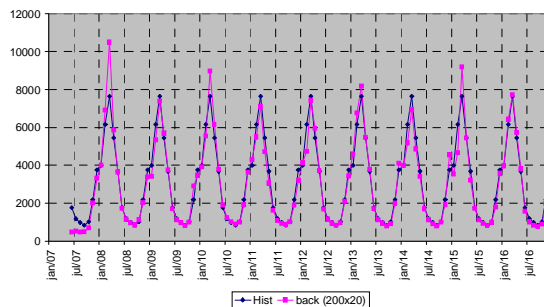


Sul - Caso de 200 Aberturas

Média - Nordeste - Geração GERF2

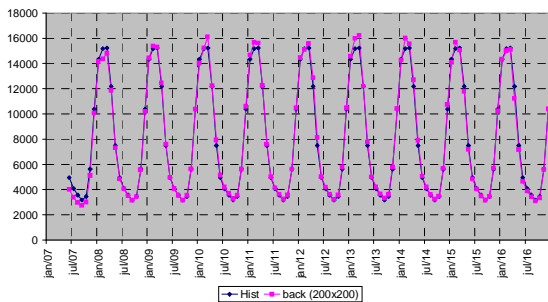


DP - Nordeste - Geração GERF2

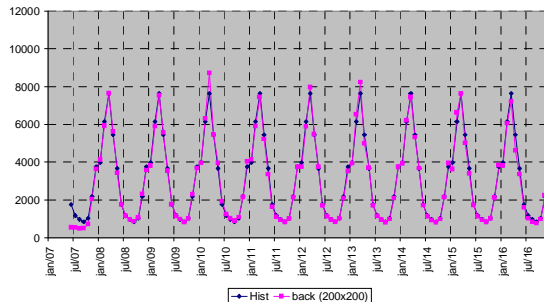


Nordeste - Caso de 20 Aberturas

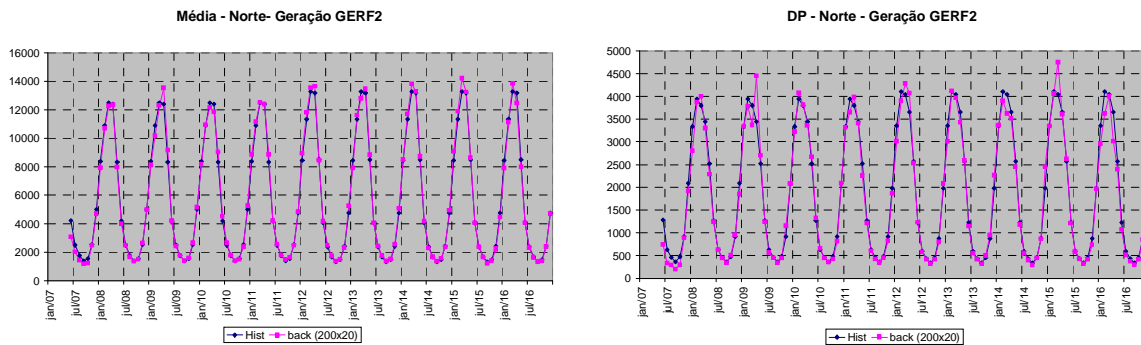
Média - Nordeste - Geração GERF2



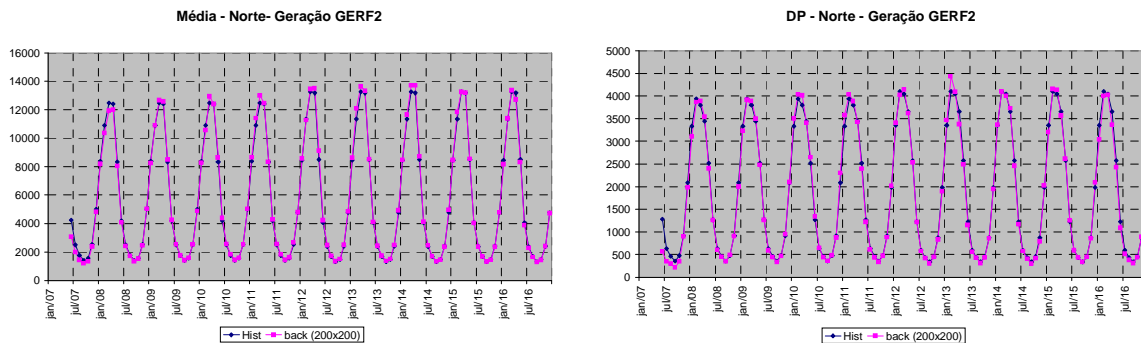
DP - Nordeste - Geração GERF2



### Nordeste- Caso de 200 Aberturas



### Norte - Caso de 20 Aberturas



### Norte- Caso de 200 Aberturas

Figura 2 - Evolução Temporal Não Condicionadas da Média e Desvio-Padrão

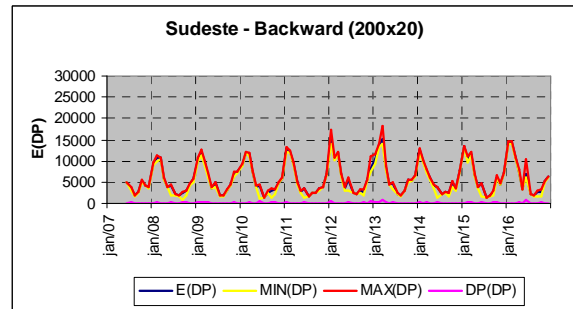
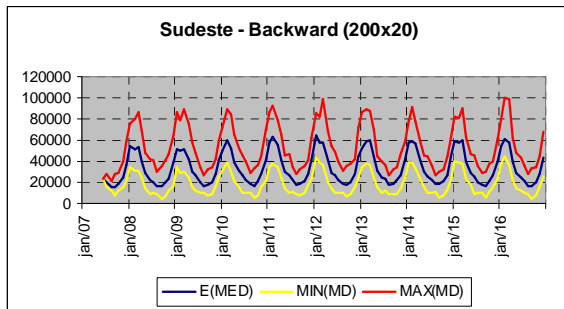
Para a análise condicionada ao passado recente, obtiveram-se as médias e os desvios-padrão de cada conjunto de aberturas, resultando em uma amostra de 200 valores para cada estatística. No caso das médias condicionadas, o que se espera é uma variação em parte explicada pela variação da parcela determinística das séries forwards, que é função das aflúências passadas (tendência hidrológica), e em parte devido à flutuação amostral. Já com relação ao desvio-padrão condicionado, pela hipótese de homoscedácia<sup>7</sup> dos resíduos, o desvio-padrão condicionado de população das ENAS não varia conforme o passado recente, e, portanto, idealmente, o que se espera são variações de desvios-padrão explicáveis por flutuação amostral.

A Figura 3 apresenta a evolução temporal dos valores médio, mínimo e máximo destas médias e desvios-padrão considerando-se 20 e 200 aberturas, para

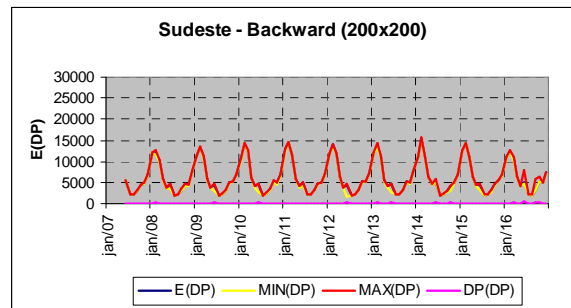
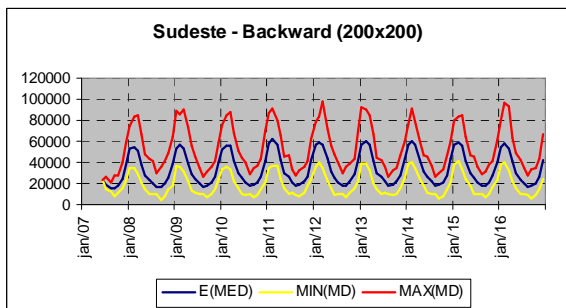
<sup>7</sup> Variância constante



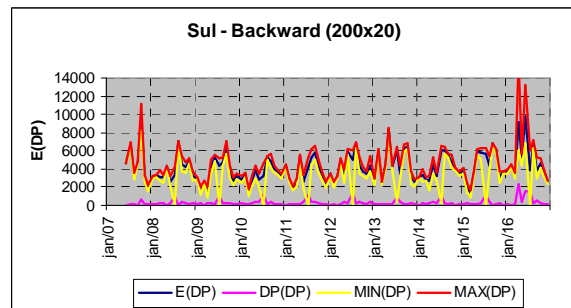
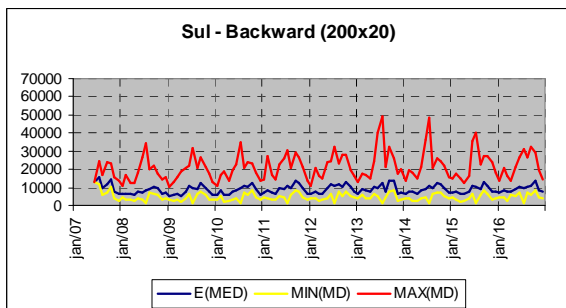
todos os subsistemas. Para o desvio-padrão, apresenta-se também a evolução do desvio-padrão de seus valores ao longo das 200 séries forwards.



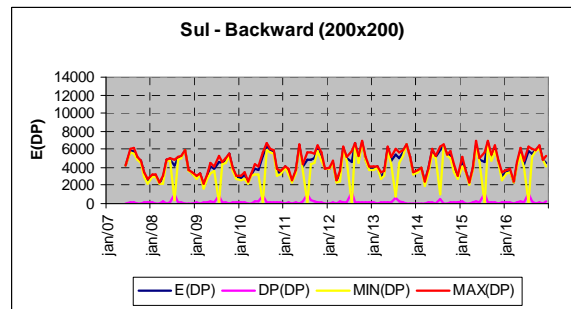
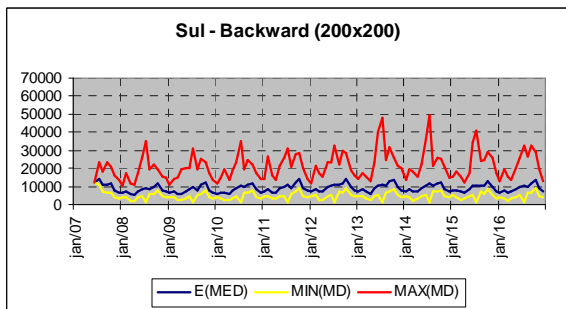
### Sudeste - Caso de 20 Aberturas



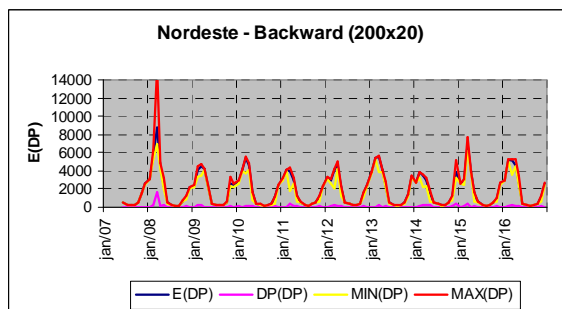
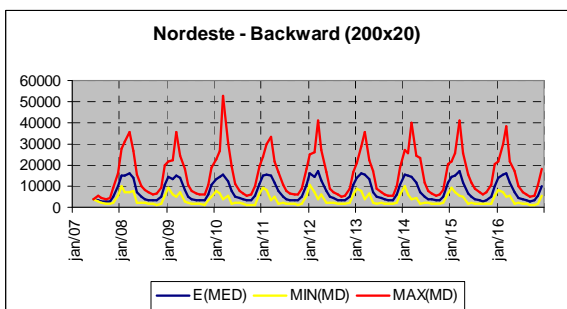
### Sudeste - Caso de 200 Aberturas



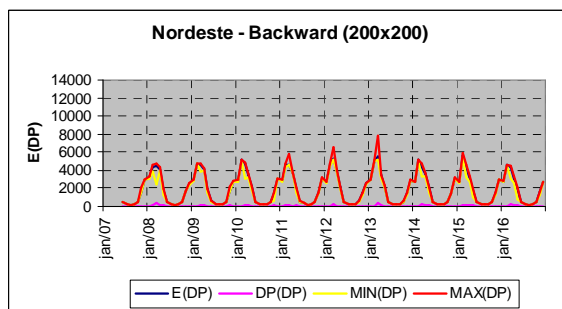
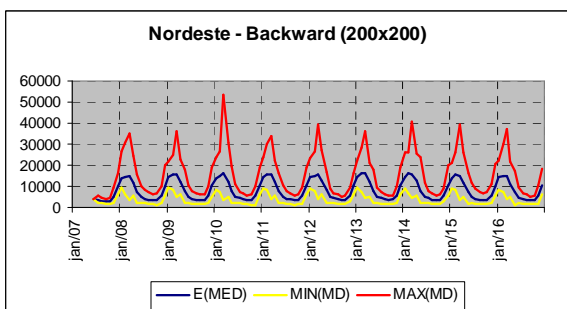
### Sul - Caso de 20 Aberturas



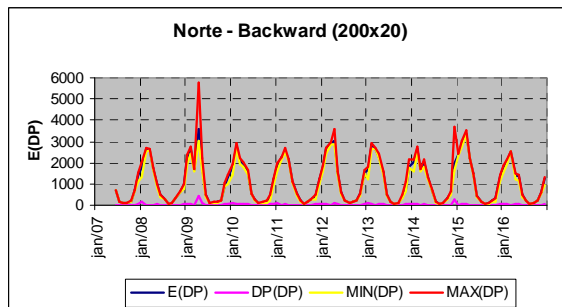
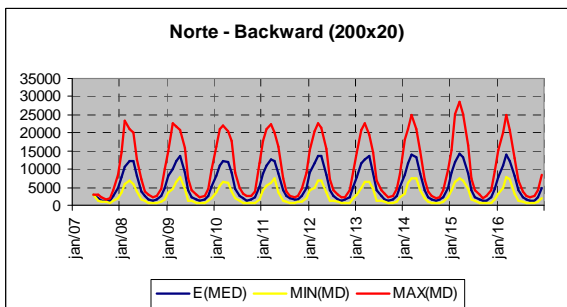
### Sul - Caso de 200 Aberturas



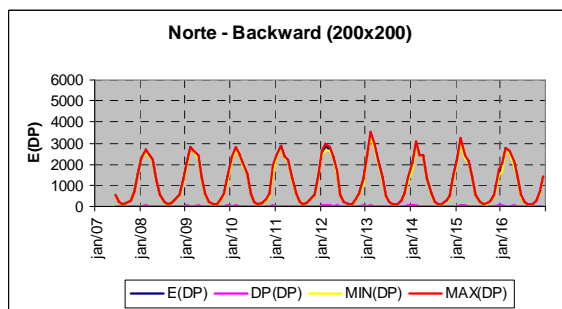
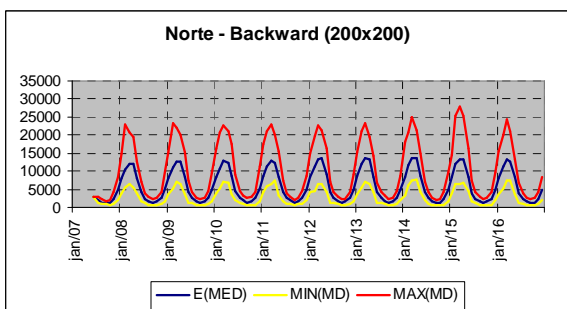
Nordeste - Caso de 20 Aberturas



Nordeste - Caso de 200 Aberturas



Norte - Caso de 20 Aberturas



Norte - Caso de 200 Aberturas

Figura 3 - Evolução Temporal dos Valores Médio, Mínimo e Máximo das Médias e Desvios-Padrão

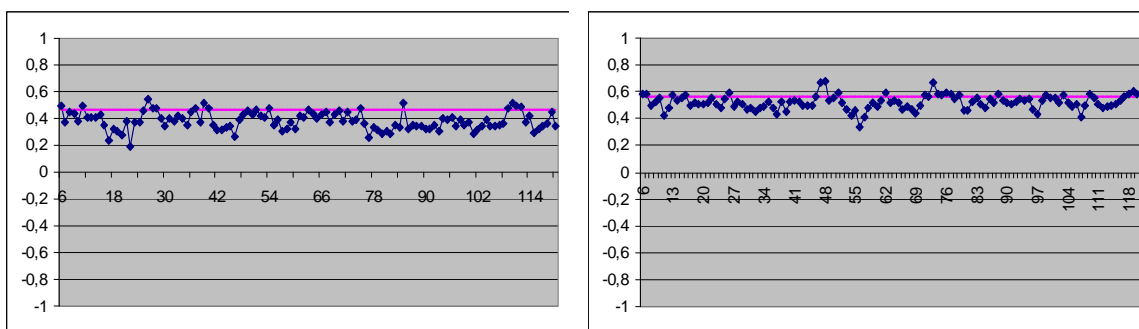
Em termos de médias condicionadas observa-se uma diferença entre os valores máximos e mínimos atribuíveis à variação da parcela determinística e à flutuação amostral, onde se percebe que o aumento no número de aberturas traz pouco impacto.

Com relação ao desvio-padrão condicionado, observa-se em alguns meses, principalmente para o subsistema Sul, a ocorrência de valores atípicos de desvio-padrão condicionado, com aumento também atípico do desvio-padrão do desvio-padrão. Este fenômeno pode se explicado por uma interação entre os limites inferiores das distribuições lognormal três parâmetros e a matriz de covariância dos ruídos.

### 11.4 Estatística Correlações Cruzadas

Para a análise das correlações cruzadas foram calculados três conjuntos de correlações cruzadas entre afluências geradas concomitantes: correlações cruzadas dos cenários forward, correlações cruzadas não condicionadas dos cenários backward, e correlações cruzadas condicionadas de cada conjunto de aberturas de cenários da simulação backward. Estas últimas correspondem às correlações cruzadas entre ruídos, pois cada conjunto de aberturas é condicionado à mesma tendência hidrológica.

As correlações cruzadas dos cenários forward foram comparadas com os valores de correlações cruzadas das séries históricas anuais de energia afluyente. A Figura 4 mostra esta comparação para todos os pares de subsistemas, onde se considerou 200 cenários. Para todos os pares de subsistemas obteve-se boa aderência.



SE-S

SE-NE

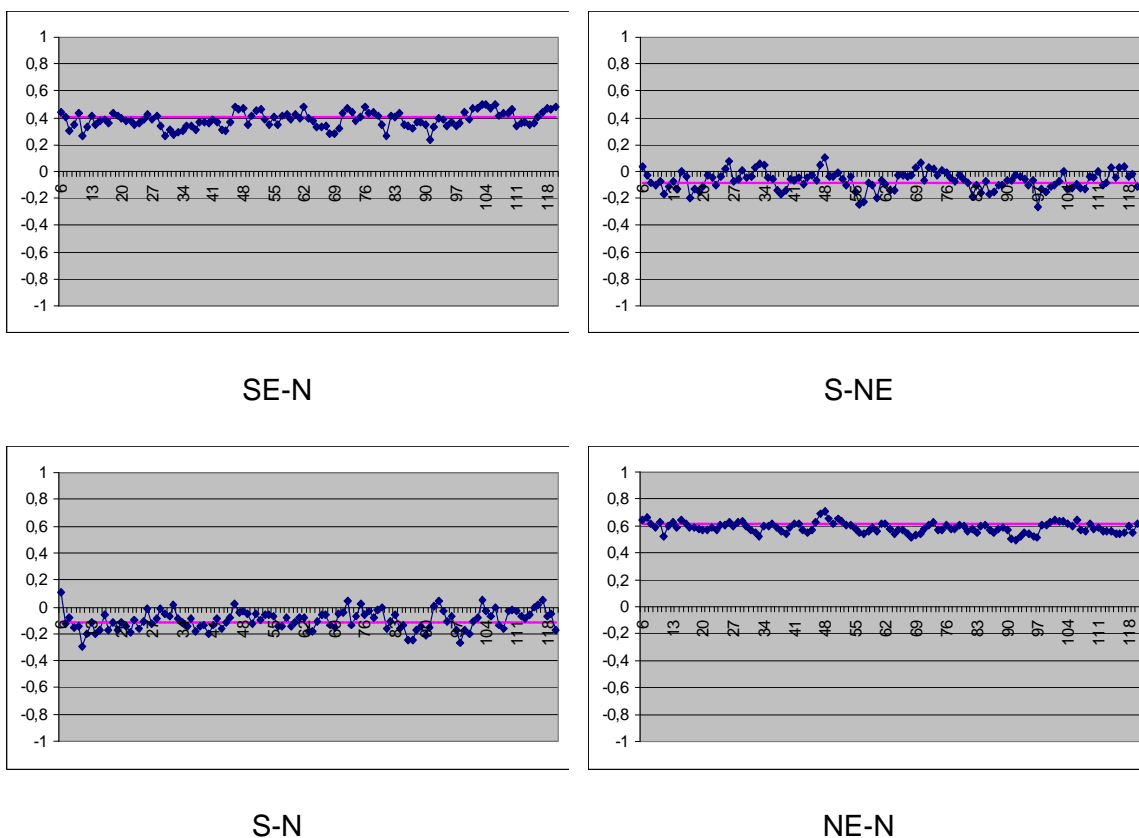
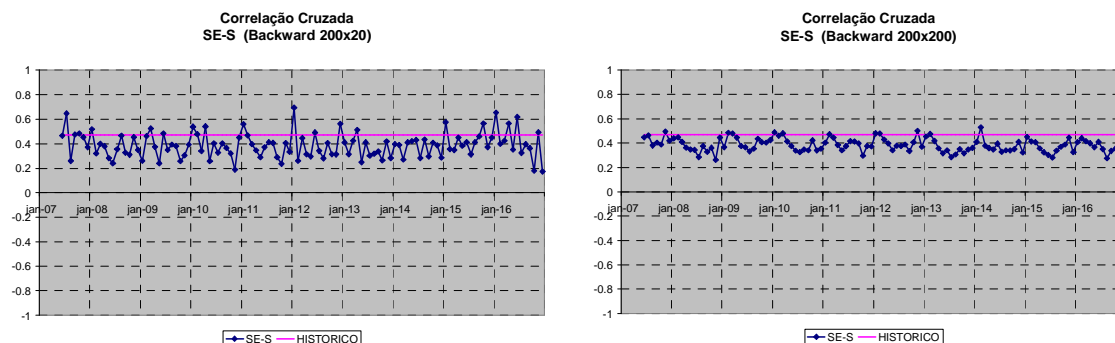
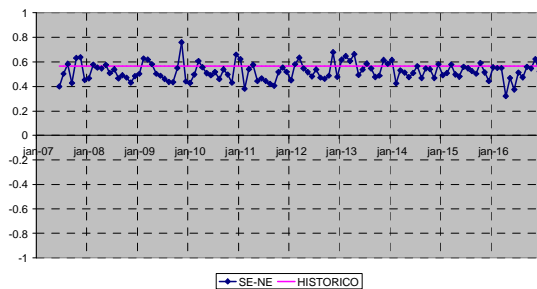


Figura 4 – Correlação Cruzada Simulação Forward

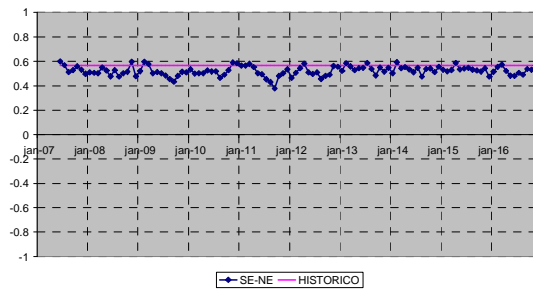
No cálculo das correlações cruzadas não condicionadas das séries backward, da mesma forma que no cálculo das médias e desvios-padrão não condicionados, as aberturas de cada série forward foram agrupadas. A Figura 5 mostra os valores obtidos para todos os pares de subsistemas considerando-se 20 e 200 aberturas. Novamente, para todos os pares de subsistemas obteve-se boa aderência.



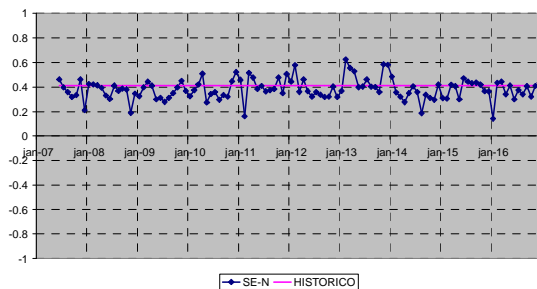
**Correlação Cruzada  
SE-NE (Backward 200x20)**



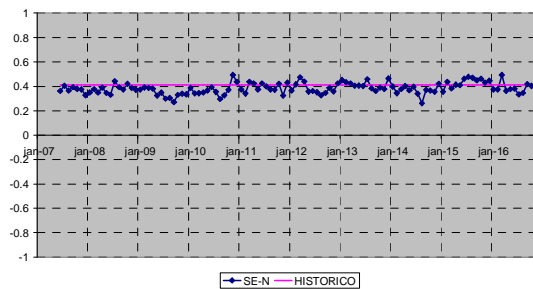
**Correlação Cruzada  
SE-NE (Backward 200x200)**



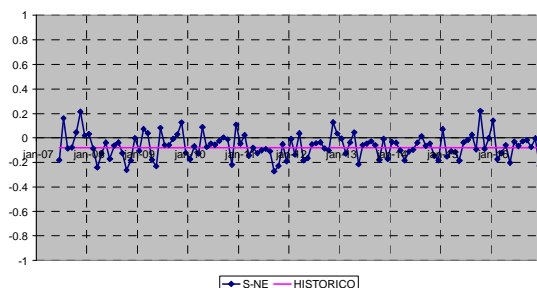
**Correlação Cruzada  
SE-N (Backward 200x20)**



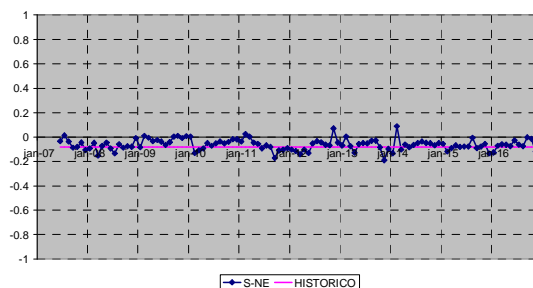
**Correlação Cruzada  
SE-N (Backward 200x200)**



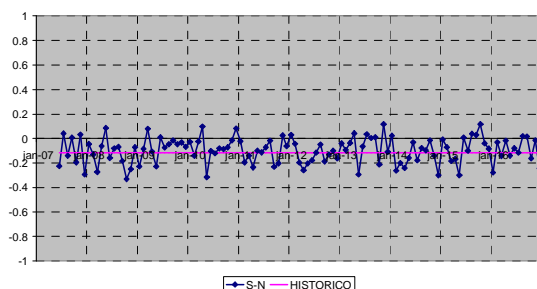
**Correlação Cruzada  
S-NE (Backward 200x20)**



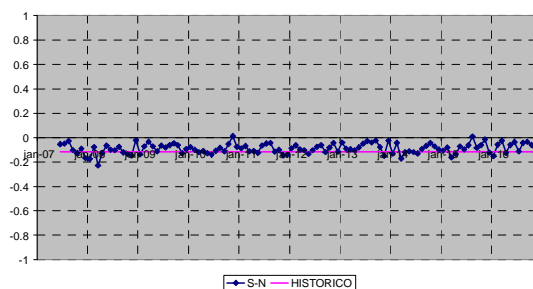
**Correlação Cruzada  
S-NE (Backward 200x200)**



**Correlação Cruzada  
S-N (Backward 200x20)**



**Correlação Cruzada  
S-N (Backward 200x200)**



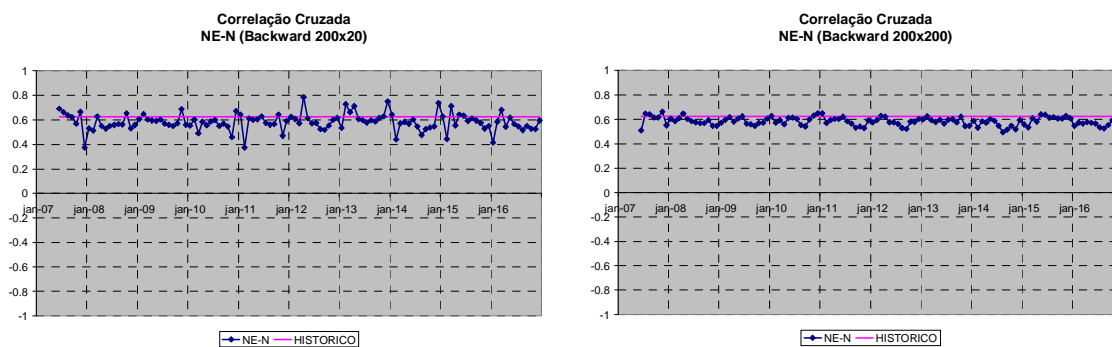
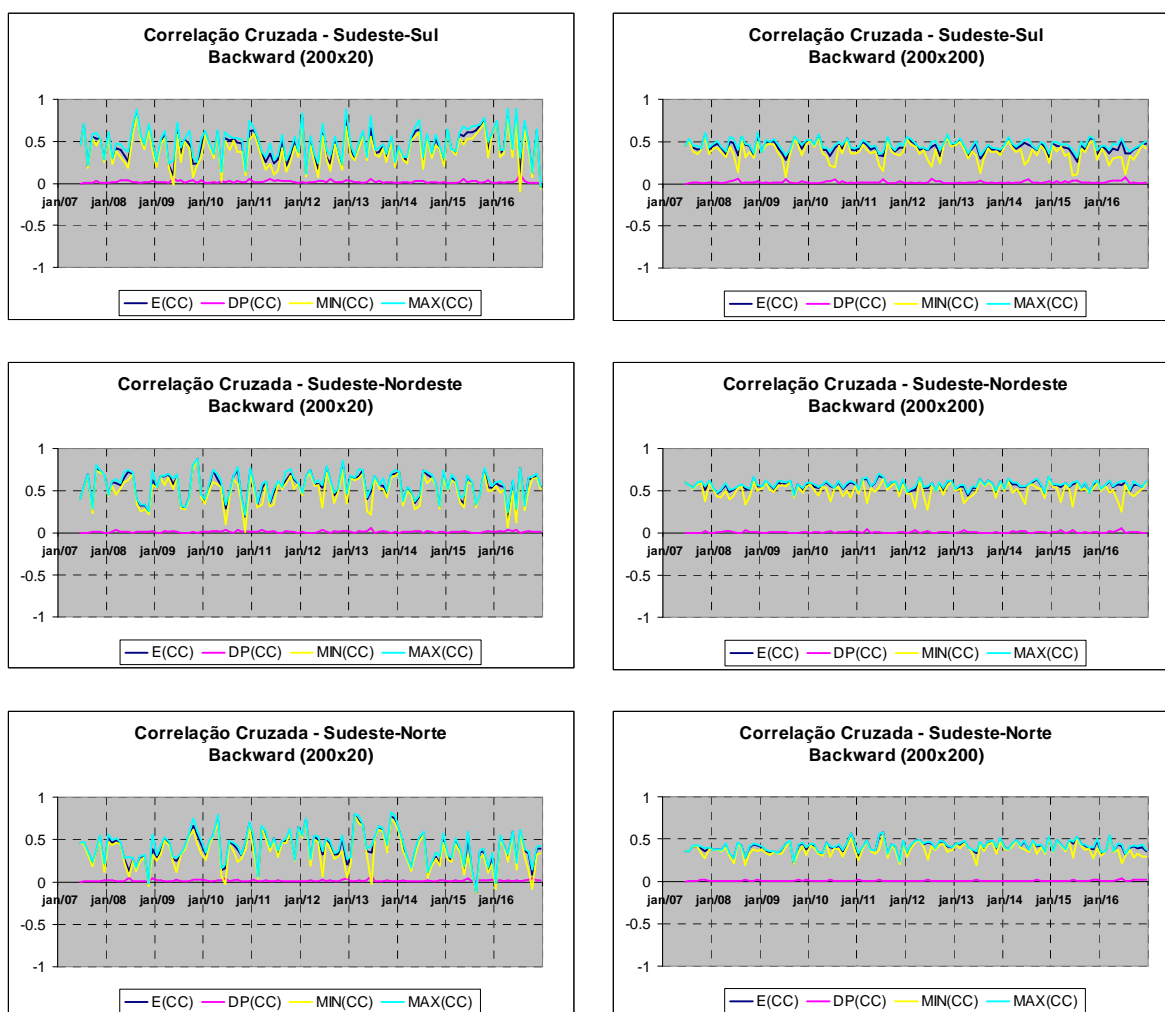


Figura 5 - Correlação Cruzada Simulação Backward

A Figura 6 mostra as estatísticas das correlações cruzadas condicionada de cada conjunto de aberturas de cenários backward obtidos para todos os pares de subsistemas considerando-se 20 e 200 aberturas. Pode-se observar, principalmente para os meses de Julho, a ocorrência de valores muito baixos no par Sudeste-Sul.



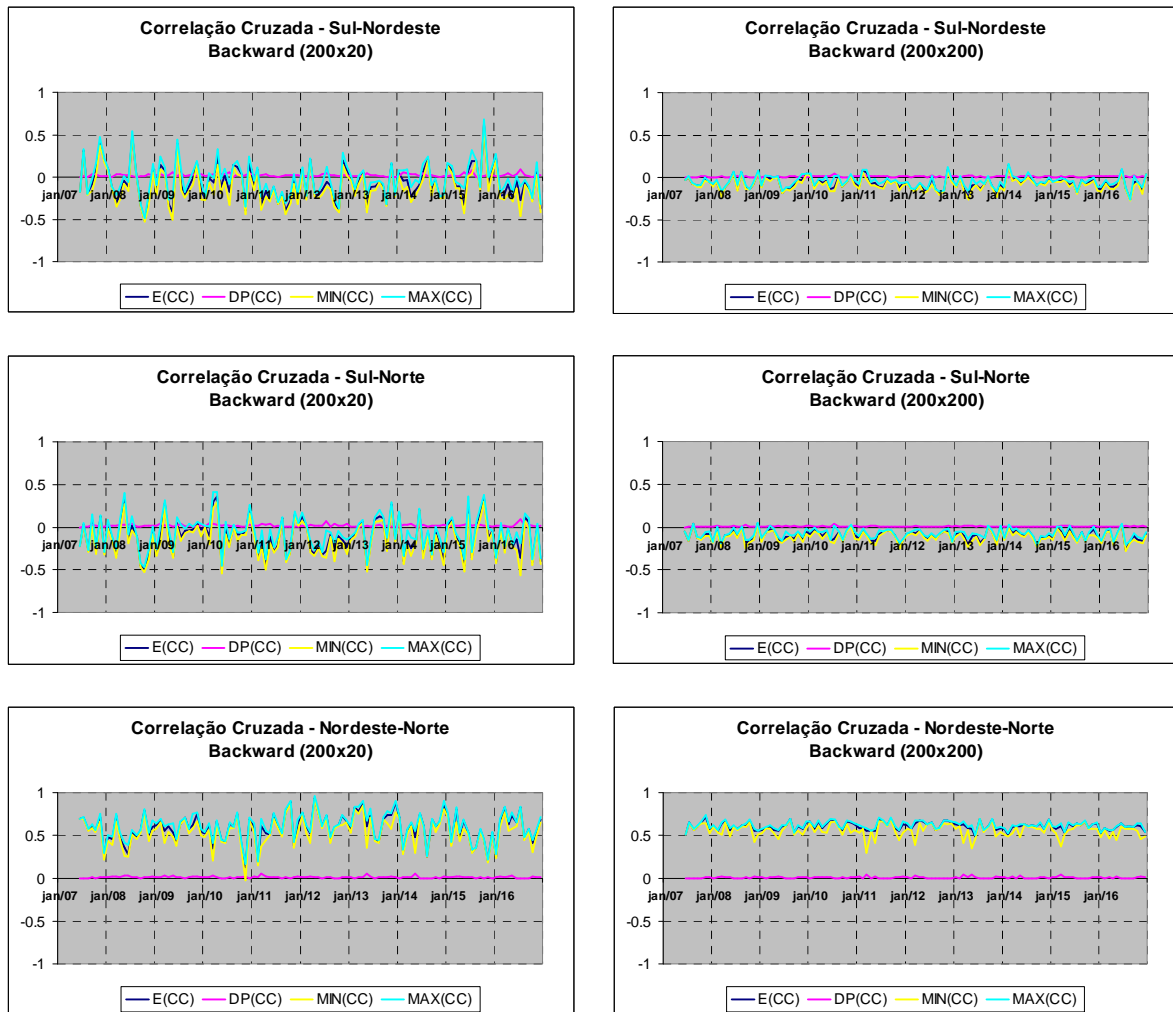


Figura 6 - Correlações Cruzadas Condicionadas – Simulação Backward

Nas versões anteriores à versão 13 do programa NEWAVE, durante o processo de geração dos cenários de aflúncias na simulação backward, era utilizado o mesmo ruído normal para todos os subsistemas. Desta forma, a correlação cruzada entre os ruídos normais de todos os pares de subsistemas era igual a um. Na Figura 7 são apresentadas as estatísticas das correlações cruzadas condicionada de cada conjunto de aberturas de cenários backward obtidos para todos os pares de subsistemas, considerando-se 20 aberturas. Note que há a ocorrência de valores de correlação cruzada das ENAs muito menores do que um, principalmente nos meses de Julho para os pares de subsistemas onde o Sul é analisado.



Figura 7 - Correlações Cruzadas Condicionadas – Sim. Back. – Ruídos Totalmente Correlacionados

A Figura 8 detalha a variabilidade das correlações cruzadas condicionadas Sudeste-Sul para os diversos meses de julho do horizonte de planejamento através de gráficos Box-plot. Pode-se observar a ocorrência de diversos valores atípicos bem inferiores ao valor histórico utilizado na geração dos ruídos, especialmente nos meses de julho.



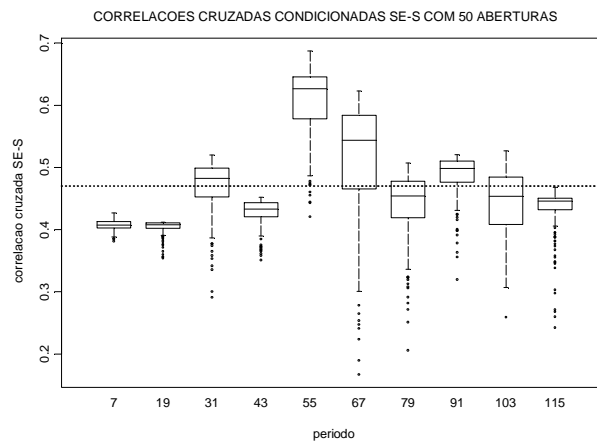


Figura 8 - Variação das correlações cruzadas Sudeste-Sul

Com o objetivo de identificar os motivos desta variabilidade, plotou-se os gráficos Box-plot dos coeficientes de variação<sup>8</sup> dos resíduos do subsistema Sul de cada julho do horizonte (Figuras 9a e 9b). A Figura 9b é um corte da Figura 9a em relação ao eixo do coeficiente de variação, considerando o valor máximo de coeficiente de variação igual a 10. Pode-se notar a ocorrência de valores atípicos superiores. Vale observar que quanto menor o valor da parte determinística da equação do modelo PAR(p), maior será o coeficiente de variação do ruído de um determinado subsistema.

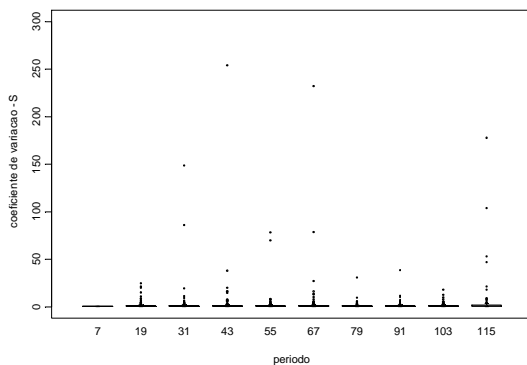


Figura 9a – C. de Variação dos Ruídos do Sul

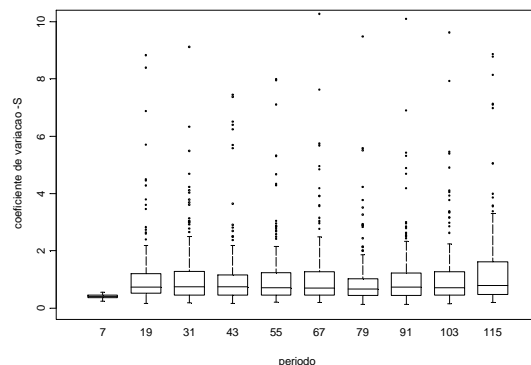


Figura 9b – C. de Variação dos Ruídos do Sul

(detalhamento da Figura 9a)

<sup>8</sup> o coeficiente de variação é uma medida de dispersão calculada como [desvio-padrão](#) dividido pela [média](#).

A Figura 10 comprova que as variações de correlações cruzadas entre Sudeste e Sul em julho de 2012 estão diretamente relacionadas às variações do coeficiente de variação dos ruídos do sistema Sul e, portanto, às condições anteriores de afluências refletidas na parte determinística do modelo.

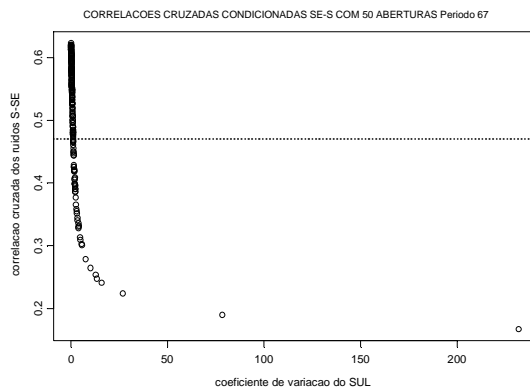
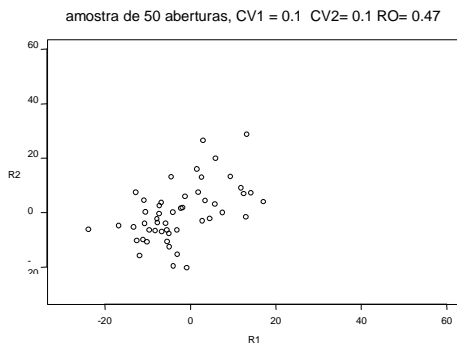


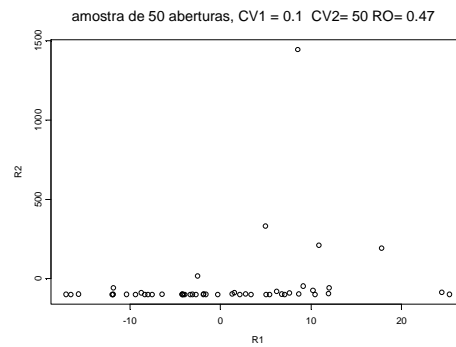
Figura 10 – Coeficiente de variação do ruído do Sul e Correlação Cruzada Sudeste-Sul

Na modelagem Lognormal multivariada, a correlação cruzada entre os ruídos de dois subsistemas quaisquer é influenciada pelo coeficiente de variação de um deles, conforme mostra o exemplo das Figuras 11a e 11b. Nas duas Figuras 11a e 11b foram gerados 50 vetores bivariados de ruídos lognormais com correlação cruzada de população igual a 0.47. Na Figura 11a os coeficientes de variação dos dois ruídos são iguais a 0.1 e a correlação cruzada gerada (0,51) é semelhante à de população. Na Figura 11b, o coeficiente de variação de um dos ruídos foi aumentado para 50. Observa-se, neste caso, que a correlação cruzada gerada (0,21) foi bem aquém do valor populacional.

Muito embora, as correlações cruzadas anuais do histórico estarem sendo bem representadas nos cenários forward e nos cenários backward, quando analisados de forma não condicionada, observa-se que as correlações cruzadas condicionadas entre as afluências aos diversos subsistemas nos cenários gerados pelo modelo para a simulação backward dependem das afluências anteriores. Uma das possibilidades de se evitar esta dependência é abandonar a modelagem lognormal, o que prejudicaria a condição de garantir apenas afluências positivas.



(a)  $r = 0,51$

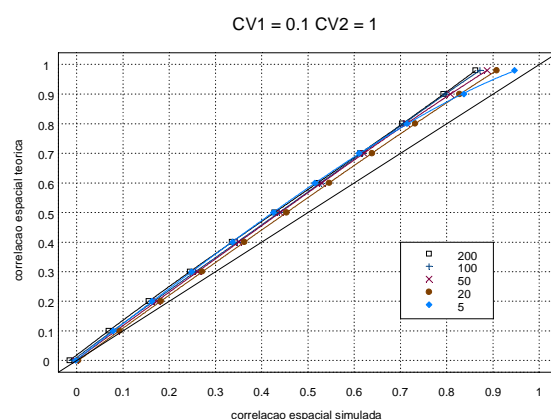
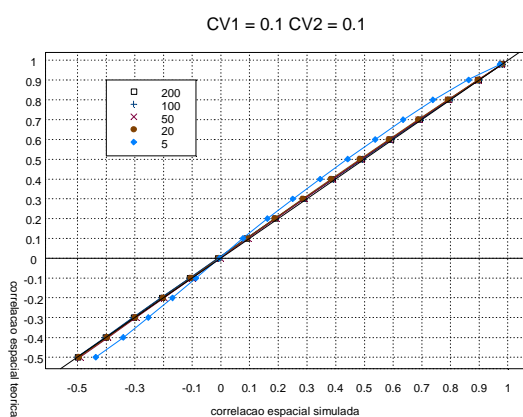


(b)  $r = 0,21$

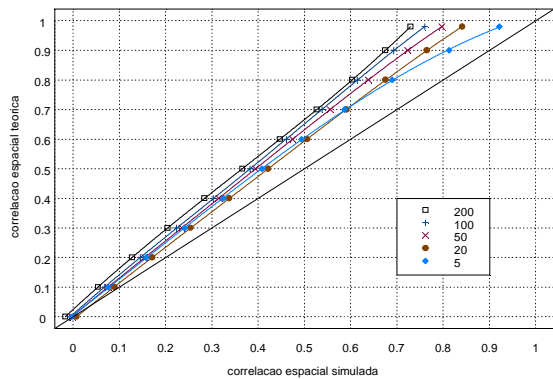
Figura 11 – Exemplo de sorteio bivariado de variáveis Lognormal

### 11.5 Compensação da Correlação Cruzada

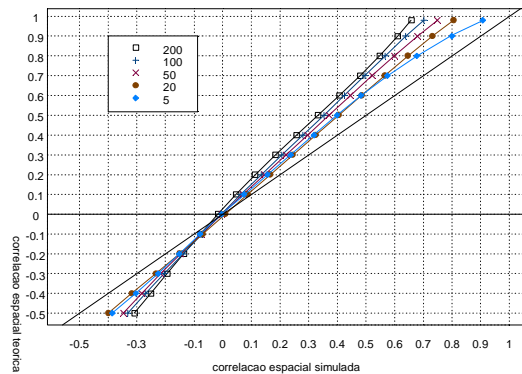
Como solução para evitar a dependência das correlações cruzadas condicionadas às  $p$  afluições recentes, optou-se por variar a matriz  $\hat{U}$  utilizada na geração dos ruídos em (9) conforme os coeficientes de variação dos ruídos, de forma a compensar a distorção causada pela transformação logarítmica. Esta compensação será realizada tanto na simulação forward quanto na backward. Para a aplicação desta compensação foram construídos polinômios do quarto grau ajustados por mínimos quadrados a valores obtidos por simulação de Monte-Carlo. A Figura 12 mostra o ajuste obtido para diferentes valores de coeficientes de variação dos ruídos e número de aberturas, para qualquer par de subsistemas.



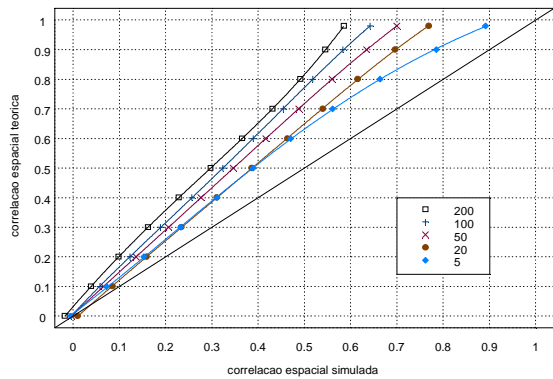
CV1 = 0.1 CV2 = 2



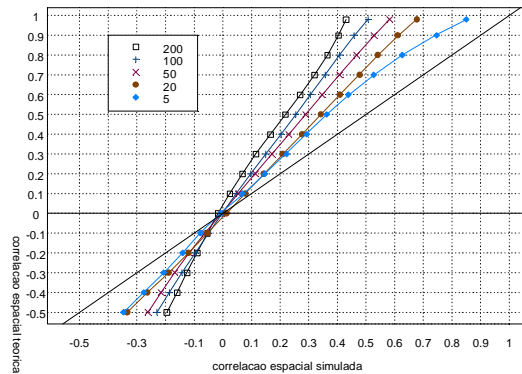
CV1 = 0.1 CV2 = 3



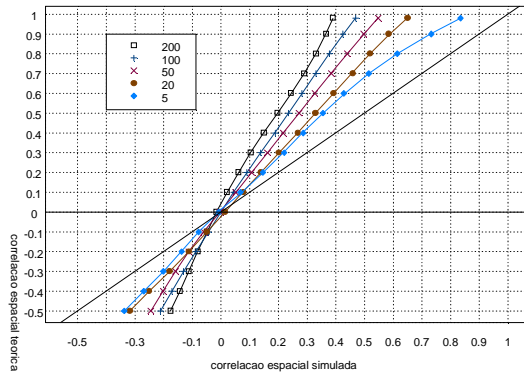
CV1 = 0.1 CV2 = 5



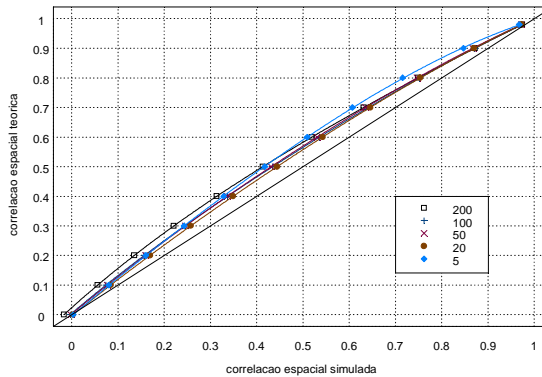
CV1 = 0.1 CV2 = 50



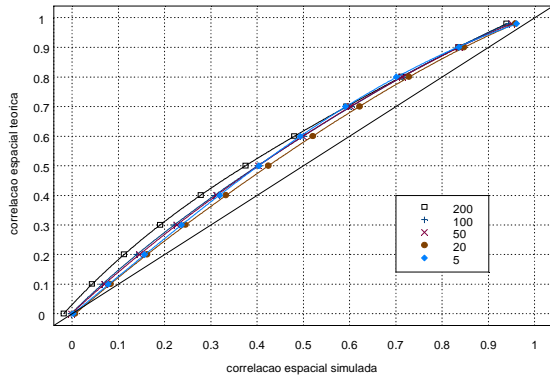
CV1 = 0.1 CV2 = 200



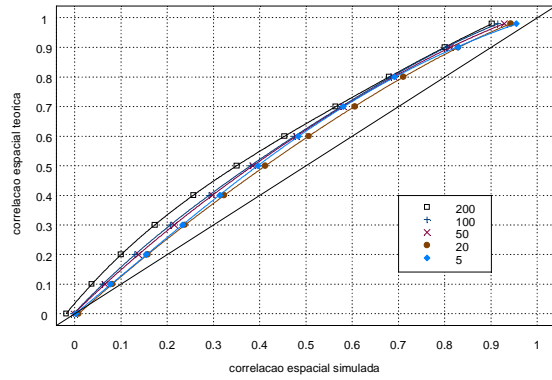
CV1 = 1 CV2 = 1



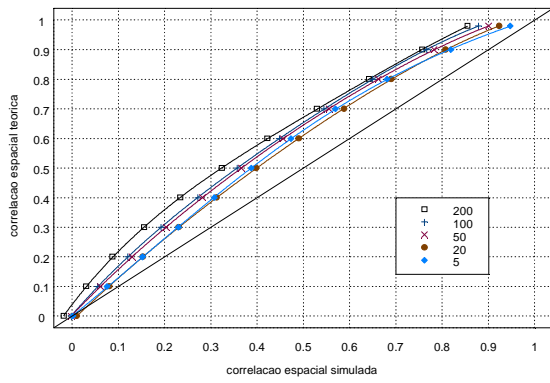
CV1 = 1 CV2 = 2



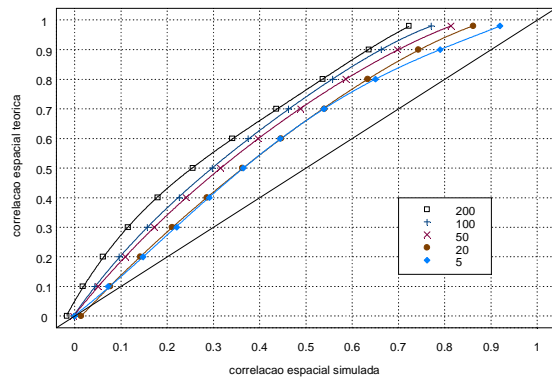
CV1 = 1 CV2 = 3



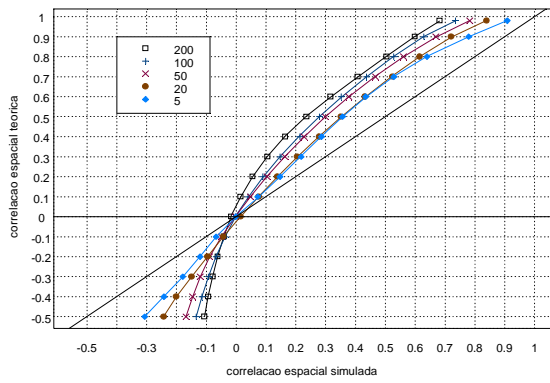
CV1 = 1 CV2 = 5



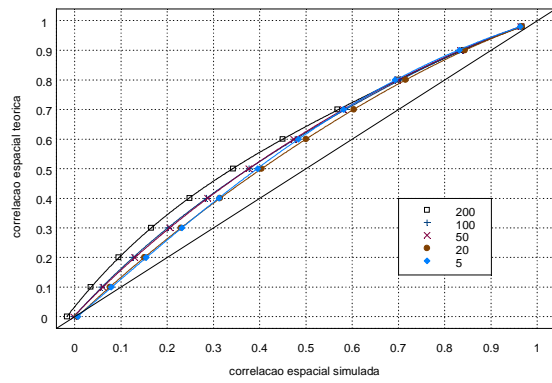
CV1 = 1 CV2 = 50



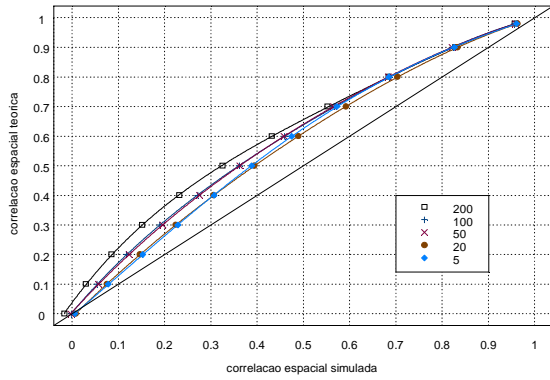
CV1 = 1 CV2 = 200



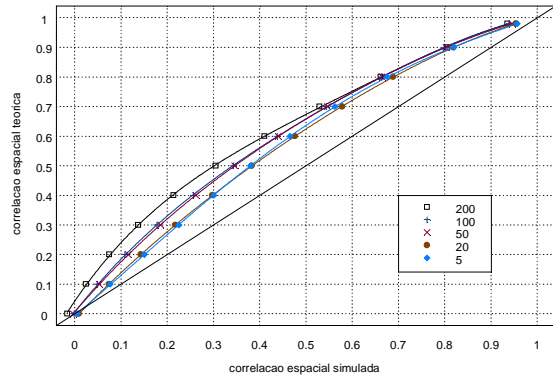
CV1 = 2 CV2 = 2



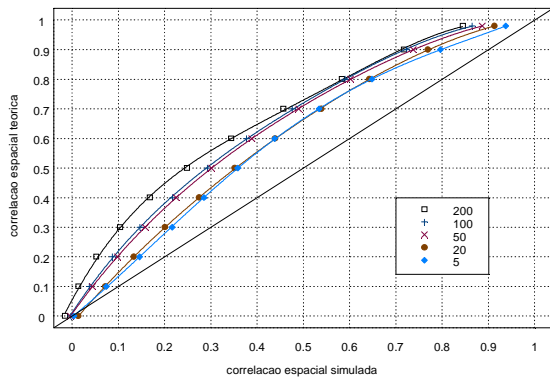
CV1 = 2 CV2 = 3



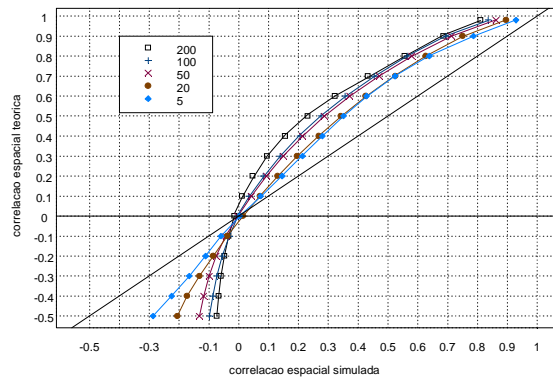
CV1 = 2 CV2 = 5



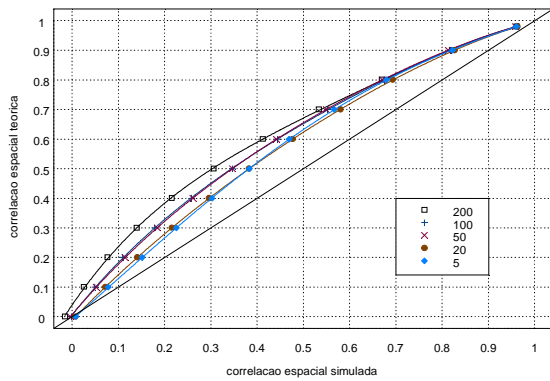
CV1 = 2 CV2 = 50



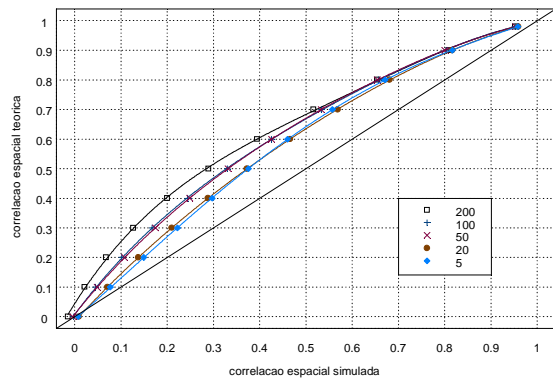
CV1 = 2 CV2 = 200



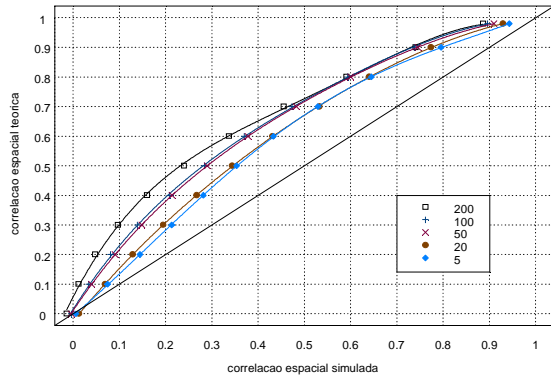
CV1 = 3 CV2 = 3



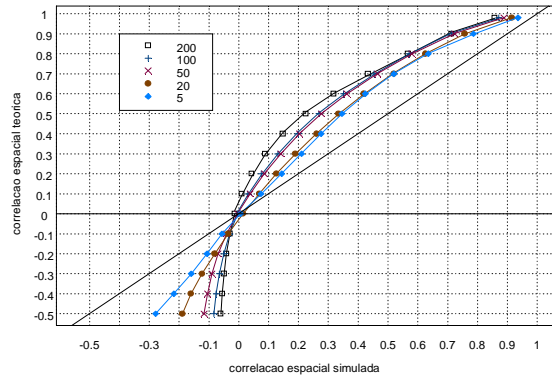
CV1 = 3 CV2 = 5



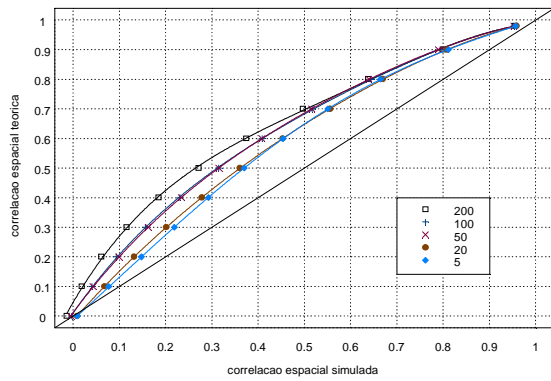
CV1 = 3 CV2 = 50



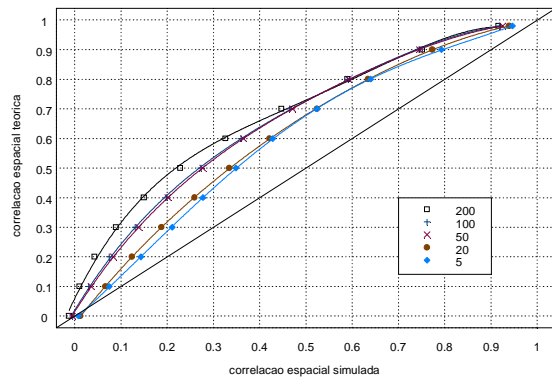
CV1 = 3 CV2 = 200



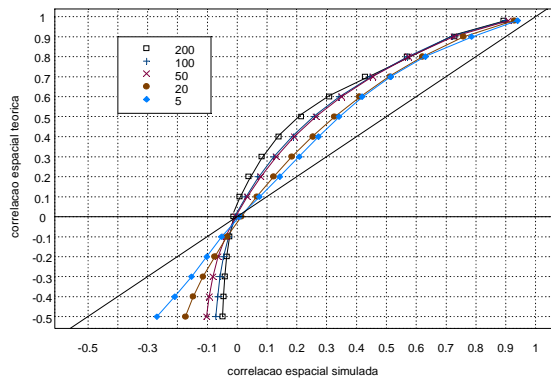
CV1 = 5 CV2 = 5



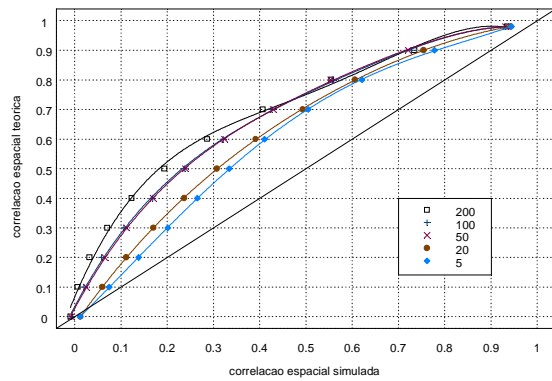
CV1 = 5 CV2 = 50



CV1 = 5 CV2 = 200



CV1 = 50 CV2 = 50



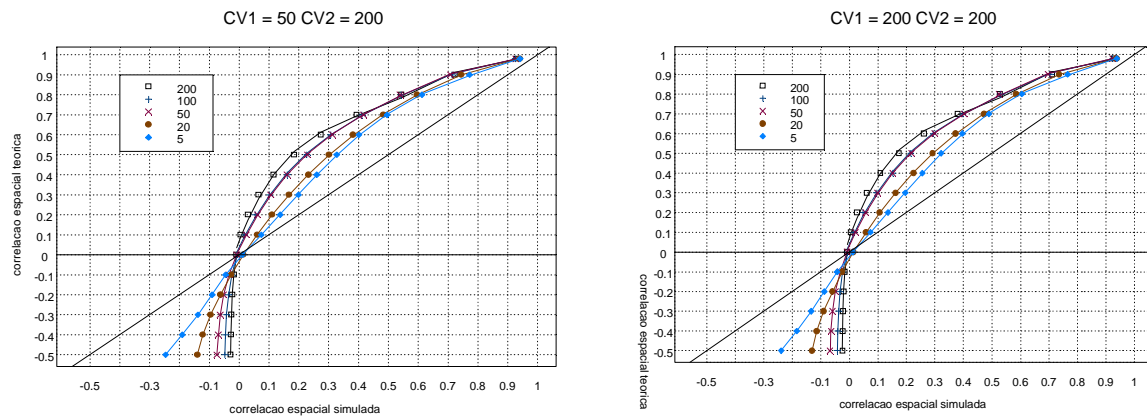


Figura 12 – Ajustes dos Polinômios para Compensação das Correlações Cruzadas

Nas Tabelas 1a e 1b são apresentados os polinômios utilizados para a compensação da correlação cruzada considerando-se valores positivos e negativos, respectivamente. Os polinômios apresentados nas tabelas abaixo são referentes ao ajuste obtido para o número de aberturas igual a 200.

Tabela 1a- Coeficientes do Polinômio para Compensação da Correlação Cruzada (valores positivos)

		Correlação Cruzada Positiva				
CV (ISIS)	CV (JSIS)	A (X <sup>0</sup> )	B (X <sup>1</sup> )	C (X <sup>2</sup> )	D (X <sup>3</sup> )	E (X <sup>4</sup> )
0,10	0,10	9,64E-03	1,01E+00	-2,33E-02	-2,91E-03	4,02E-03
0,10	1,00	1,81E-02	1,20E+00	-2,14E-01	7,90E-02	6,78E-02
0,10	2,00	2,43E-02	1,44E+00	-5,17E-01	2,78E-01	2,57E-01
0,10	3,00	2,78E-02	1,60E+00	-7,36E-01	3,76E-01	5,86E-01
0,10	5,00	3,15E-02	1,80E+00	-9,96E-01	3,68E-01	1,37E+00
0,10	50,00	4,12E-02	2,45E+00	-1,74E+00	-1,21E+00	8,91E+00
0,10	200,00	4,44E-02	2,70E+00	-2,00E+00	-2,64E+00	1,48E+01
1,00	0,10	1,81E-02	1,20E+00	-2,14E-01	7,90E-02	6,78E-02
1,00	1,00	2,35E-02	1,40E+00	-7,94E-01	5,28E-01	-1,60E-01
1,00	2,00	3,07E-02	1,66E+00	-1,53E+00	1,36E+00	-4,98E-01
1,00	3,00	3,46E-02	1,83E+00	-2,05E+00	2,04E+00	-8,10E-01
1,00	5,00	3,89E-02	2,04E+00	-2,72E+00	3,00E+00	-1,29E+00
1,00	50,00	5,03E-02	2,68E+00	-5,14E+00	7,15E+00	-3,74E+00
1,00	200,00	5,41E-02	2,92E+00	-6,19E+00	9,20E+00	-5,11E+00
2,00	0,10	2,43E-02	1,44E+00	-5,17E-01	2,78E-01	2,57E-01
2,00	1,00	3,07E-02	1,66E+00	-1,53E+00	1,36E+00	-4,98E-01
2,00	2,00	3,35E-02	1,94E+00	-2,32E+00	2,19E+00	-8,39E-01
2,00	3,00	3,77E-02	2,11E+00	-2,92E+00	2,96E+00	-1,18E+00
2,00	5,00	4,22E-02	2,30E+00	-3,64E+00	3,95E+00	-1,65E+00
2,00	50,00	5,42E-02	2,90E+00	-6,05E+00	7,76E+00	-3,71E+00
2,00	200,00	5,81E-02	3,12E+00	-7,04E+00	9,53E+00	-4,76E+00
3,00	0,10	2,78E-02	1,60E+00	-7,36E-01	3,76E-01	5,86E-01
3,00	1,00	3,46E-02	1,83E+00	-2,05E+00	2,04E+00	-8,10E-01
3,00	2,00	3,77E-02	2,11E+00	-2,92E+00	2,96E+00	-1,18E+00



		<b>Correlação Cruzada Positiva</b>				
<b>CV (ISIS)</b>	<b>CV (JSIS)</b>	<b>A (X^0)</b>	<b>B (X^1)</b>	<b>C (X^2)</b>	<b>D (X^3)</b>	<b>E (X^4)</b>
3,00	3,00	3,92E-02	2,29E+00	-3,49E+00	3,64E+00	-1,47E+00
3,00	5,00	4,39E-02	2,47E+00	-4,22E+00	4,62E+00	-1,93E+00
3,00	50,00	5,63E-02	3,04E+00	-6,55E+00	8,22E+00	-3,80E+00
3,00	200,00	6,03E-02	3,24E+00	-7,49E+00	9,83E+00	-4,73E+00
5,00	0,10	3,15E-02	1,80E+00	-9,96E-01	3,68E-01	1,37E+00
5,00	1,00	3,89E-02	2,04E+00	-2,72E+00	3,00E+00	-1,29E+00
5,00	2,00	4,22E-02	2,30E+00	-3,64E+00	3,95E+00	-1,65E+00
5,00	3,00	4,39E-02	2,47E+00	-4,22E+00	4,62E+00	-1,93E+00
5,00	5,00	4,58E-02	2,67E+00	-4,91E+00	5,51E+00	-2,33E+00
5,00	50,00	5,87E-02	3,20E+00	-7,14E+00	8,86E+00	-4,02E+00
5,00	200,00	6,28E-02	3,39E+00	-8,02E+00	1,03E+01	-4,84E+00
50,00	0,10	4,12E-02	2,45E+00	-1,74E+00	-1,21E+00	8,91E+00
50,00	1,00	5,03E-02	2,68E+00	-5,14E+00	7,15E+00	-3,74E+00
50,00	2,00	5,42E-02	2,90E+00	-6,05E+00	7,76E+00	-3,71E+00
50,00	3,00	5,63E-02	3,04E+00	-6,55E+00	8,22E+00	-3,80E+00
50,00	5,00	5,87E-02	3,20E+00	-7,14E+00	8,86E+00	-4,02E+00
50,00	50,00	6,58E-02	3,68E+00	-9,08E+00	1,15E+01	-5,26E+00
50,00	200,00	7,00E-02	3,84E+00	-9,82E+00	1,27E+01	-5,87E+00
200,00	0,10	4,44E-02	2,70E+00	-2,00E+00	-2,64E+00	1,48E+01
200,00	1,00	5,41E-02	2,92E+00	-6,19E+00	9,20E+00	-5,11E+00
200,00	2,00	5,81E-02	3,12E+00	-7,04E+00	9,53E+00	-4,76E+00
200,00	3,00	6,03E-02	3,24E+00	-7,49E+00	9,83E+00	-4,73E+00
200,00	5,00	6,28E-02	3,39E+00	-8,02E+00	1,03E+01	-4,84E+00
200,00	50,00	7,00E-02	3,84E+00	-9,82E+00	1,27E+01	-5,87E+00
200,00	200,00	7,23E-02	4,01E+00	-1,06E+01	1,39E+01	-6,47E+00

Tabela 1b - Coeficientes do Polinômio para Compensação da Correlação Cruzada (valores negativos)

		<b>Correlação Cruzada Negativa</b>				
<b>CV (ISIS)</b>	<b>CV (JSIS)</b>	<b>A (X^0)</b>	<b>B (X^1)</b>	<b>C (X^2)</b>	<b>D (X^3)</b>	<b>E (X^4)</b>
0,10	0,10	9,64E-03	1,01E+00	-2,34E-02	-3,76E-03	1,30E-02
0,10	1,00	1,81E-02	1,19E+00	-1,85E-01	-1,39E-01	4,00E-02
0,10	2,00	2,42E-02	1,43E+00	-4,64E-01	-5,90E-01	-5,44E-02
0,10	3,00	2,76E-02	1,59E+00	-7,12E-01	-1,08E+00	-1,48E-01
0,10	5,00	3,13E-02	1,79E+00	-1,05E+00	-1,70E+00	1,81E-01
0,10	50,00	4,23E-02	2,56E+00	-1,27E+00	5,33E+00	3,21E+01
0,10	200,00	4,67E-02	2,92E+00	-1,51E-01	2,04E+01	8,73E+01
1,00	0,10	1,81E-02	1,19E+00	-1,85E-01	-1,39E-01	4,00E-02
1,00	1,00	2,35E-02	1,40E+00	-7,82E-01	5,74E-01	4,36E-01
1,00	2,00	3,07E-02	1,68E+00	-1,62E+00	2,55E+00	1,63E+00
1,00	3,00	3,48E-02	1,88E+00	-2,19E+00	5,83E+00	4,19E+00
1,00	5,00	3,95E-02	2,14E+00	-2,55E+00	1,54E+01	1,62E+01
1,00	50,00	5,79E-02	3,53E+00	1,00E+01	2,31E+02	6,05E+02
1,00	200,00	6,81E-02	4,50E+00	2,92E+01	5,25E+02	1,66E+03
2,00	0,10	2,42E-02	1,43E+00	-4,64E-01	-5,90E-01	-5,44E-02
2,00	1,00	3,07E-02	1,68E+00	-1,62E+00	2,55E+00	1,63E+00
2,00	2,00	3,34E-02	2,01E+00	-3,51E+00	4,55E+00	-2,65E+01

Correlação Cruzada Negativa						
CV (ISIS)	CV (JSIS)	A (X^0)	B (X^1)	C (X^2)	D (X^3)	E (X^4)
2,00	3,00	3,75E-02	2,23E+00	-5,32E+00	4,22E+00	-9,07E+01
2,00	5,00	4,22E-02	2,51E+00	-7,91E+00	-8,31E-01	-2,88E+02
2,00	50,00	6,15E-02	4,08E+00	-2,65E+00	6,87E+01	-3,12E+03
2,00	200,00	7,37E-02	5,35E+00	2,20E+01	3,47E+02	-5,34E+03
3,00	0,10	2,76E-02	1,59E+00	-7,12E-01	-1,08E+00	-1,48E-01
3,00	1,00	3,48E-02	1,88E+00	-2,19E+00	5,83E+00	4,19E+00
3,00	2,00	3,75E-02	2,23E+00	-5,32E+00	4,22E+00	-9,07E+01
3,00	3,00	3,78E-02	2,39E+00	-1,10E+01	-3,74E+01	-4,40E+02
3,00	5,00	4,12E-02	2,53E+00	-2,17E+01	-1,57E+02	-1,53E+03
3,00	50,00	4,67E-02	2,06E+00	-1,33E+02	-2,59E+03	-2,73E+04
3,00	200,00	4,70E-02	1,34E+00	-2,33E+02	-5,52E+03	-6,28E+04
5,00	0,10	3,13E-02	1,79E+00	-1,05E+00	-1,70E+00	1,81E-01
5,00	1,00	3,95E-02	2,14E+00	-2,55E+00	1,54E+01	1,62E+01
5,00	2,00	4,22E-02	2,51E+00	-7,91E+00	-8,31E-01	-2,88E+02
5,00	3,00	4,12E-02	2,53E+00	-2,17E+01	-1,57E+02	-1,53E+03
5,00	5,00	3,62E-02	2,00E+00	-6,60E+01	-8,59E+02	-6,90E+03
5,00	50,00	-8,55E-03	-7,29E+00	-7,23E+02	-1,69E+04	-1,62E+05
5,00	200,00	-5,75E-02	-1,78E+01	-1,51E+03	-3,97E+04	-4,12E+05
50,00	0,10	4,23E-02	2,56E+00	-1,27E+00	5,33E+00	3,21E+01
50,00	1,00	5,79E-02	3,53E+00	1,00E+01	2,31E+02	6,05E+02
50,00	2,00	6,15E-02	4,08E+00	-2,65E+00	6,87E+01	-3,12E+03
50,00	3,00	4,67E-02	2,06E+00	-1,33E+02	-2,59E+03	-2,73E+04
50,00	5,00	-8,55E-03	-7,29E+00	-7,23E+02	-1,69E+04	-1,62E+05
50,00	50,00	-9,38E-01	-2,30E+02	-1,93E+04	-6,56E+05	-8,22E+06
50,00	200,00	-2,07E+00	-5,40E+02	-4,88E+04	-1,84E+06	-2,55E+07
200,00	0,10	4,67E-02	2,92E+00	-1,51E-01	2,04E+01	8,73E+01
200,00	1,00	6,81E-02	4,50E+00	2,92E+01	5,25E+02	1,66E+03
200,00	2,00	7,37E-02	5,35E+00	2,20E+01	3,47E+02	-5,34E+03
200,00	3,00	4,70E-02	1,34E+00	-2,33E+02	-5,52E+03	-6,28E+04
200,00	5,00	-5,75E-02	-1,78E+01	-1,51E+03	-3,97E+04	-4,12E+05
200,00	50,00	-2,07E+00	-5,40E+02	-4,88E+04	-1,84E+06	-2,55E+07
200,00	200,00	-4,47E+00	-1,28E+03	-1,28E+05	-5,41E+06	-8,35E+07

Logo, para cada período e estado analisado a compensação da correlação cruzada é função da correlação cruzada histórica e do polinômio selecionado. A seleção do polinômio é realizada em função do coeficiente de variação de cada subsistema do par de subsistemas e do número de séries considerado.

Para calcular a compensação da correlação cruzada é realizada uma ponderação entre os valores dos coeficientes de correlação cruzada compensada de quatro polinômios selecionados (um maior e outro menor do que o valor dos coeficientes de variação de cada subsistema do par analisado).

O fator de ponderação é função do inverso da distância euclidiana entre os coeficientes de variação observados para o par de subsistemas analisado e aqueles tabelados no procedimento de compensação da correlação cruzada (Tabelas 1a e 1b).

Para ilustrar o cálculo da compensação da correlação cruzada, considere os seguintes dados:

- correlação cruzada histórica entre os subsistemas SE-S: 0,459
- coeficiente de variação dos subsistemas SE e S para um determinado período e estado: 0,3689 e 0,4816

São analisados os polinômios da Tabela 1a cujos pares de coeficientes de variação estão na vizinhança do par analisado (0,3689;0,4816). Neste caso, os pares que estão na vizinhança são (0,10;0,10), (0,10;1,00), (1,00;0,10) e (1,00;1,00). Os quatro polinômios selecionados são reproduzidos na Tabela 2 e na última coluna são apresentados os valores da correlação cruzada compensada para cada polinômio.

Tabela 2 – Caso Exemplo: Cálculo dos coeficientes de correlação cruzada compensada

CV (ISIS)	CV (JSIS)	A (X^0)	B (X^1)	C (X^2)	D (X^3)	E (X^4)	X (CorrSE-s)	F(X)
0,10	0,10	9,64E-03	1,01E+00	-2,33E-02	-2,91E-03	4,02E-03	0,459	0,46822
0,10	1,00	1,81E-02	1,20E+00	-2,14E-01	7,90E-02	6,78E-02	0,459	0,53446
1,00	0,10	1,81E-02	1,20E+00	-2,14E-01	7,90E-02	6,78E-02	0,459	0,53446
1,00	1,00	2,35E-02	1,40E+00	-7,94E-01	5,28E-01	-1,60E-01	0,459	0,54278

A correlação cruzada compensada é calculada como a soma ponderada dos valores calculados para cada polinômio da Tabela 2. O cálculo dos fatores de ponderação (FP) é apresentado na Tabela 3.

Tabela 3 – Caso Exemplo: Cálculo dos fatores de ponderação

CV (ISIS)	CV (JSIS)	distância $[(CV_{isis}-0,3689)^2+(CV_{jsis}-0,4816)^2]^{1/2}$	ID (1/distância)	FP (FP=ID/SD)
0,10	0,10	0,4668	2,1421	0,3329
0,10	1,00	0,5840	1,7124	0,2661
1,00	0,10	0,7375	1,3559	0,2107
1,00	1,00	0,8167	1,2244	0,1903

Soma de ID (SD) **6,4348**

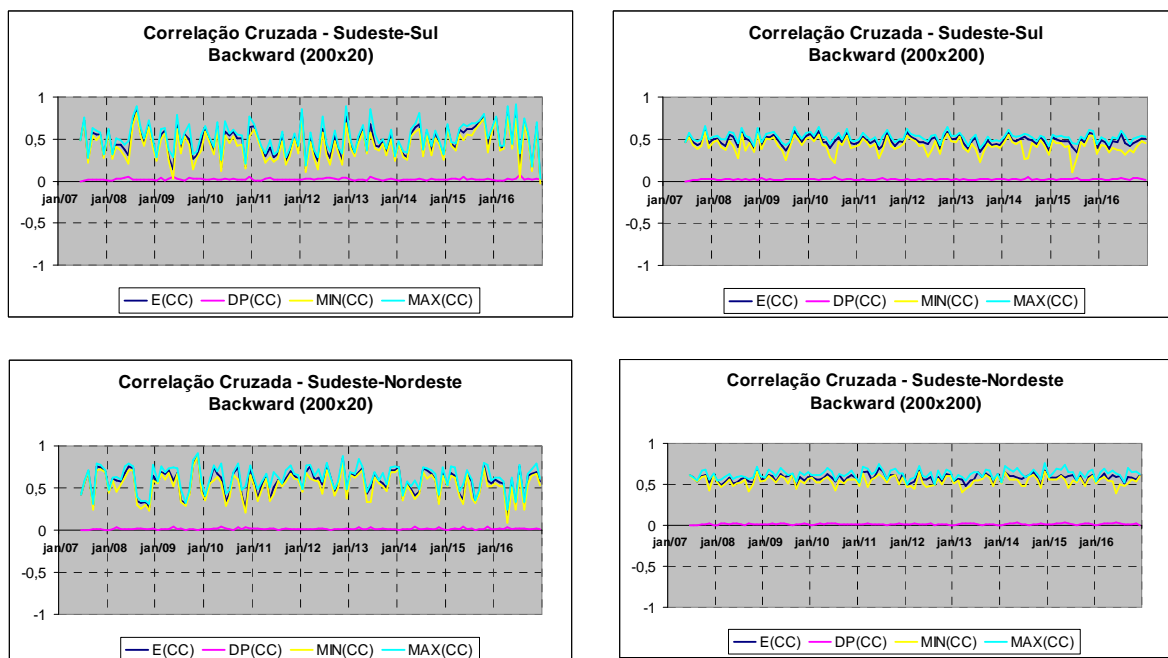
Finalmente, o cálculo da correlação cruzada compensada entre os subsistemas SE-S é mostrado na Tabela 4.

Tabela 4 – Caso Exemplo: Cálculo da correlação cruzada compensada SE-S

CV (ISIS)	CV (JSIS)	F(X)	FP	F(X)*FP
0,10	0,10	0,46822	0,3329	0,1559
0,10	1,00	0,53446	0,2661	0,1422
1,00	0,10	0,53446	0,2107	0,1126
1,00	1,00	0,54278	0,1903	0,1033
<b>Correlação cruzada compensada</b>				<b>0,5140</b>

## 11.6 Estudo de Caso

A Figura 13 mostra as correlações cruzadas condicionadas das séries geradas para simulação backward para todos os pares de subsistemas levando-se em consideração as funções de compensação apresentadas no item anterior. Comparando-se com a Figura 6, observa-se uma significativa melhoria.



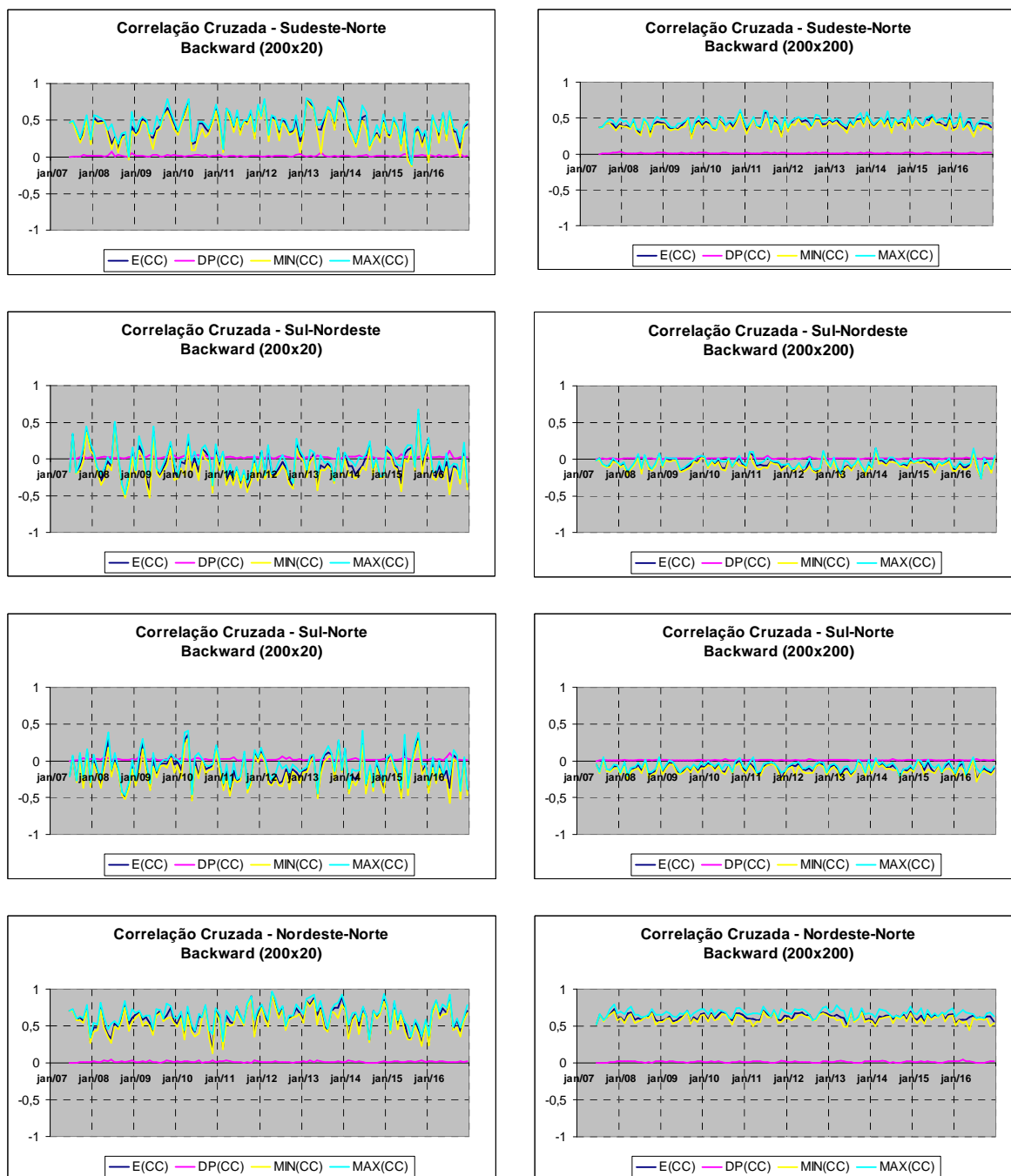
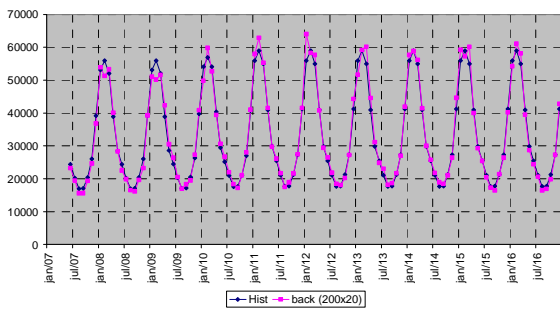


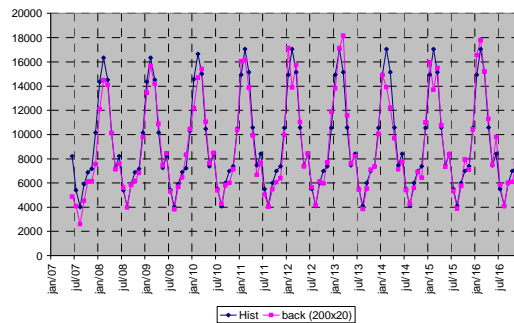
Figura 13 – Correlações Cruzadas Condicionadas – Simulação Backward – Uso de Compensação

A mesma comparação mostrada na Figura 2 é realizada para os cenários gerados considerando o uso da compensação da correlação cruzada. Na Figura 14 são comparadas as evoluções temporais não condicionadas de média e desvio-padrão das ENA dos cenários gerados para a simulação backward com as mesmas estatísticas da série histórica. Pode-se observar que a aderência entre os valores gerados e históricos não foi afetada pelo uso da compensação da correlação cruzada.

Média - Sudeste - G2\_CorrigeCCruza

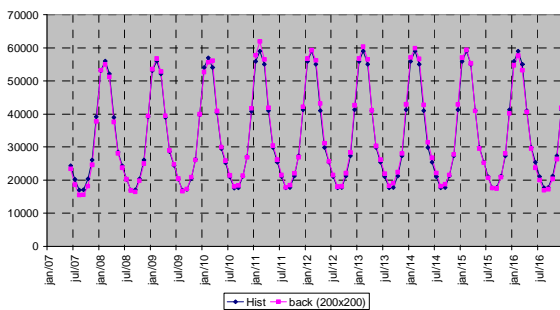


DP - Sudeste - G2\_CorrigeCCruza

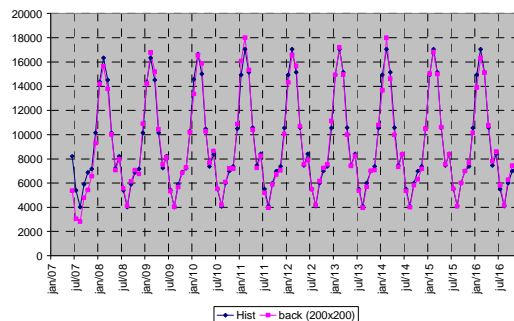


Sudeste - Caso de 20 Aberturas

Média - Sudeste - G2\_CorrigeCCruza

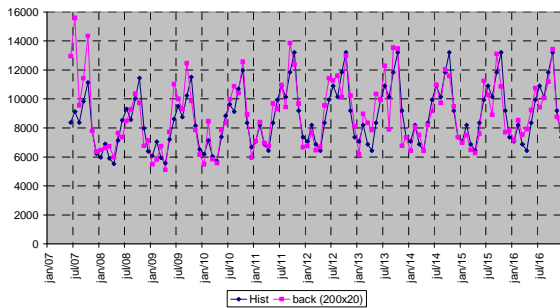


DP - Sudeste - G2\_CorrigeCCruza

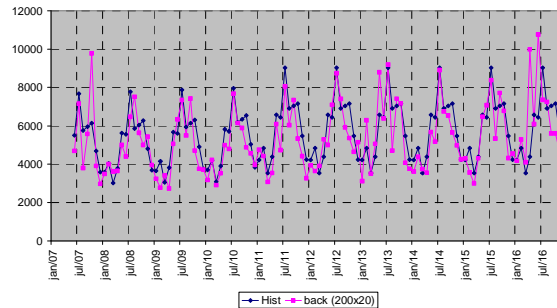


Sudeste - Caso de 200 Aberturas

Média - Sul - G2\_CorrigeCCruza

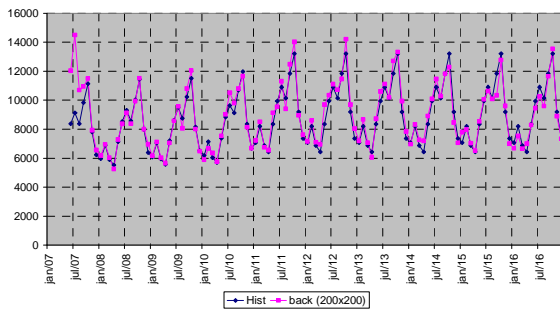


DP - Sul - G2\_CorrigeCCruza

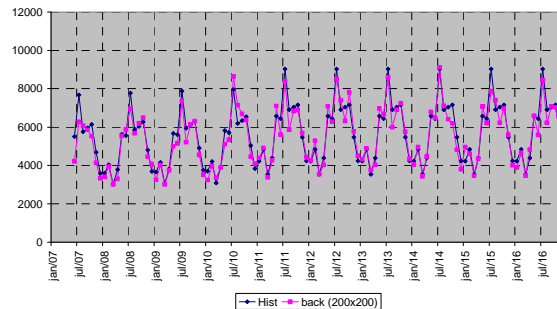


Sul - Caso de 20 Aberturas

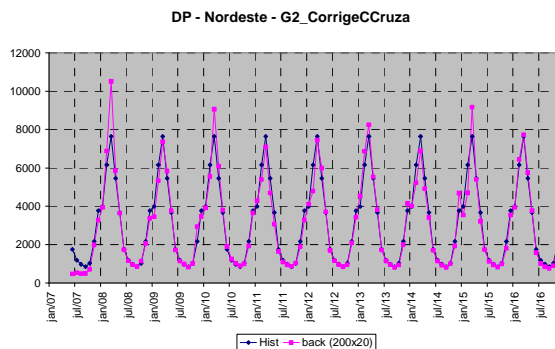
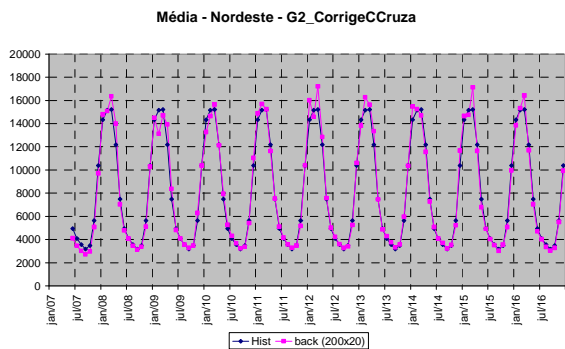
Média - Sul - G2\_CorrigeCCruza



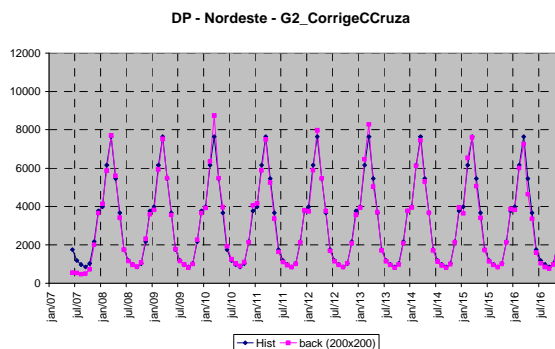
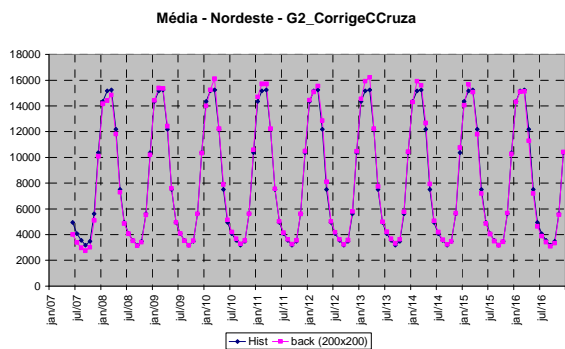
DP - Sul - G2\_CorrigeCCruza



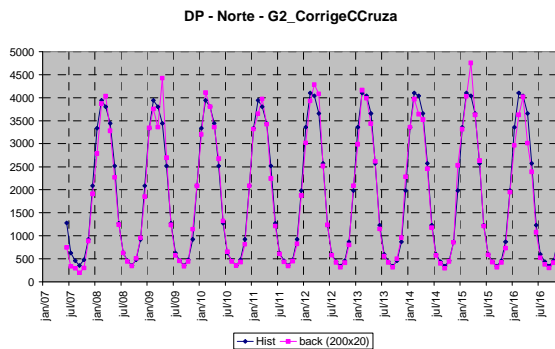
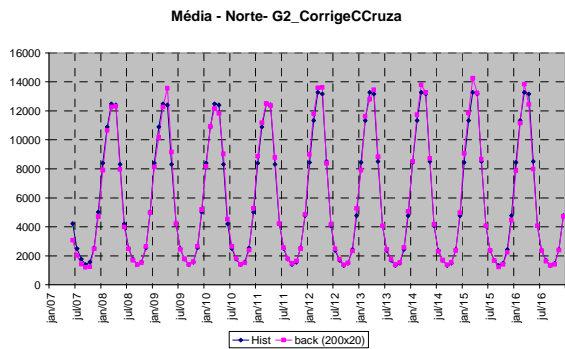
### Sul - Caso de 200 Aberturas



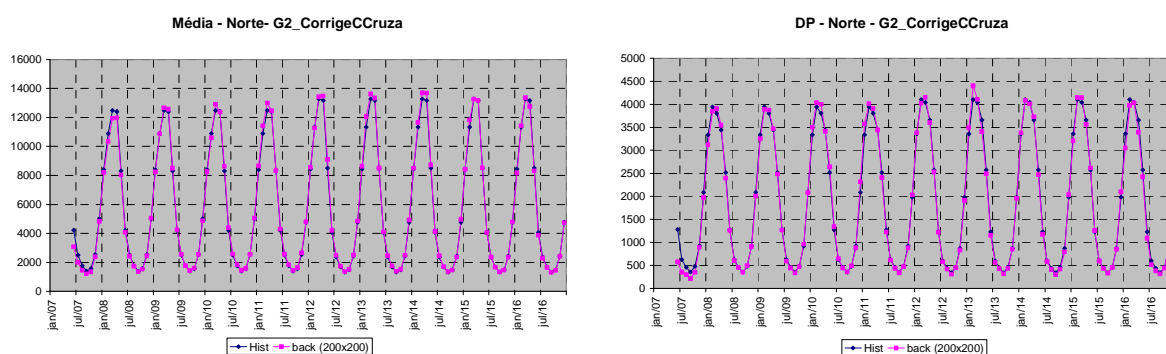
### Nordeste- Caso de 20 Aberturas



### Nordeste- Caso de 200 Aberturas



### Norte - Caso de 20 Aberturas



Norte- Caso de 200 Aberturas

Figura 14 - Evolução Temporal Não Condicionadas da Média e Desvio-Padrão – Uso de Compensação

## 11.7 Conclusões

Esta Nota Técnica apresentou uma modificação da metodologia empregada no modelo GEVAZP para gerar cenários sintéticos de energias afluentes aos subsistemas de geração de energia elétrica, a fim de aprimorar o desempenho do modelo em relação às correlações cruzadas condicionadas entre os subsistemas. O aprimoramento proposto utiliza polinômios de quarto grau, função dos coeficientes de variação dos ruídos de cada subsistema e do número de aberturas utilizado, para alterar a matriz de correlação cruzada utilizada na geração. Os resultados obtidos mostraram o bom desempenho do aprimoramento proposto.

## 11.8 BIBLIOGRAFIA

- [1] M.E.P. Maceira, C.V. Bezerra, "Stochastic Streamflow model for Hydroelectric Systems", Proceedings of 5<sup>th</sup> International Conference on Probabilistic Methods Applied to Power Systems - PMAPS, pp. 305-310, Vancouver, Canada, 1997.
- [2] M.E.P. Maceira, L.A. Terry, F.S. Costa, J.M. Damázio, A.C.G. Melo, "Chain of Optimization Models for Setting the Energy Dispatch and Spot Price in the Brazilian System", Proceedings of 14<sup>th</sup> Power Systems Computation Conference, Seville, Spain, 2002.
- [3] 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", Proceedings of 16<sup>th</sup> Power Systems Computation Conference.



- [4] J.D.Salas, J.W.Delleur, V.Yevjevich, W.L.Lane, Applied Modeling of Hydroelectric Series, Water Resources Publications, 1980.
- [5] G.E.P.Box, D.R.Cox, "An Analysis of Transformations", *Journal of The Royal Statistical Society*, A127, pp. 211-252, 1964.

## 12 NOTA TÉCNICA nº 44 - INTERCÂMBIO MÍNIMO COM VARIÁVEL DE FOLGA (Junho 2009)

### 12.1 Introdução

O modelo NEWAVE, em sua versão 14.1, permite a representação de intercâmbios mínimos obrigatórios. Para essas restrições é fornecida pelo usuário a capacidade mínima da interligação, que deve ser inferior à capacidade máxima de intercâmbio da linha. Adicionalmente, para cada mês do período de estudo que tenha a restrição de intercâmbio mínimo obrigatório, a capacidade de intercâmbio no sentido oposto da linha deve ser zero. O fator de ponderação utilizado para representar um dado patamar de carga na restrição de intercâmbio mínimo obrigatório é o mesmo utilizado para a restrição de intercâmbio de energia. A motivação da restrição de intercâmbio mínimo obrigatório é a representação da geração de Itaipu em cada um dos setores (50 e 60 Hz)

Na versão 14.1, a restrição de intercâmbio mínimo obrigatório não pode ser violada. Desta forma, quando não há recurso suficiente para atender essa restrição, a solução ótima do problema é o acionamento da térmica deficit. Esta é uma solução matematicamente viável, porém não é uma solução possível na operação do sistema.

A funcionalidade descrita nesta Nota Técnica consiste na inclusão de uma variável de folga na restrição de intercâmbio mínimo obrigatório, penalizada na função objetivo. Dessa forma, a térmica deficit não será mais acionada para o atendimento deste requisito, a não ser que a penalidade declarada para o não-atendimento dessa restrição seja maior do que o custo de déficit.

Essa nova funcionalidade foi implementada a partir da versão 14.1 do programa NEWAVE gerando, dessa forma, uma nova versão intitulada versão 14.2. As restrições de mínimo intercâmbio descritas nesta Nota Técnica, quando representadas, o são no cálculo da estratégia e na simulação final.

### 12.2 Modelagem

Na modelagem do problema de planejamento da operação energética da versão 14.1, as restrições de intercâmbio mínimo obrigatório são dadas conforme (1).

$$\underline{\text{int}}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}} \leq \text{int}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}} \quad (1)$$

onde:

$\text{int}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}}$  : é o intercâmbio de energia (MWh) entre os subsistemas *isis* e *jsis*, no patamar *ipat* e período *iper*;

$\underline{\text{int}}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}}$  : é o intercâmbio mínimo de energia (MWh) entre os subsistemas *isis* e *jsis*, no patamar *ipat* e período *iper*, dado por (2)

$$\underline{\text{int}}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}} = \text{INTMIN}_{\text{isis}, \text{jsis}}^{\text{iper}} * \text{PINTER}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}} * \text{FPENG}_{\text{ipat}}^{\text{iper}} \quad (2)$$

$\text{INTMIN}_{\text{isis}, \text{jsis}}^{\text{iper}}$  : é limite de intercâmbio mínimo obrigatório (MW) entre os subsistemas *isis* e *jsis*, no período *iper*;

$\text{FPENG}_{\text{ipat}}^{\text{iper}}$  : é duração do patamar *ipat* no período *iper* (pu);

$\text{PINTER}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}}$  : é o fator que deve ser aplicado ao intercâmbio para compor o intercâmbio do subsistema *isis* para o subsistema *jsis* do patamar *ipat* no período *iper* (pu). Esse fator é o mesmo utilizado para compor a capacidade máxima de intercâmbio por patamar.

Na modelagem utilizada na versão 14.2, uma variável de decisão  $\delta_{\text{intmin}}$  é inserida no problema como uma variável de folga para a meta de intercâmbio mínimo obrigatório. A restrição (1) passa, então, a ser escrita como em (3):

$$\underline{\text{int}}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}} \leq \text{int}_{\text{isis}, \text{jsis}, \text{ipat}}^{\text{iper}} + \delta_{\text{intmin}}^{\text{iper}} \quad (3)$$

$$\delta_{\text{intmin}}^{\text{iper}} \geq 0$$

onde:

$\delta_{\text{intmin}}^{\text{iper}}$  : é a variável de folga associada à restrição de intercâmbio mínimo de energia entre os subsistemas *isis* e *jsis*, no patamar *ipat* e período *iper*.

Existe uma penalidade associada ao não atendimento da meta de intercâmbio mínimo obrigatório, informada pelo usuário. Desta forma, a função objetivo do problema passa a considerar o custo associado à violação da meta de intercâmbio mínimo obrigatório. A parcela referente a esse custo é dada pela equação (4):

$$\sum_{isis=1}^{nsis} \sum_{jsis=1}^{nsis} \sum_{ipat=1}^{npsc} \delta_{intmin}^{iper_{isis,jsis,ipat}} \cdot C_{int\ min} \quad (4)$$

onde:

$C_{int\ min}$  : é a penalidade (\$/MWh) associada ao não atendimento da meta de intercâmbio mínimo obrigatório.

### **12.3 Entrada de dados**

Para implementação dessa nova funcionalidade foi alterado o formato do arquivo de entrada Penalidades (PENALID.DAT). No arquivo PENALID.DAT foi incluída uma nova palavra-chave, INTMIN, que corresponde à penalidade pela violação das restrições de intercâmbio mínimo. A declaração da capacidade mínima de intercâmbio é realizada no arquivo de subsistemas (SISTEMA.DAT), no mesmo bloco onde é declarada a capacidade máxima de intercâmbio. Há um campo específico para indicar se as informações fornecidas são relativas à capacidade mínima de intercâmbio.

A descrição completa dos formatos dos arquivos citados acima pode ser obtida no Manual do Usuário.

### **12.4 Rebatimento nos demais modelos da cadeia**

A consideração explícita do intercâmbio mínimo obrigatório com variável de folga não está implementada nos demais modelos da cadeia energética.

## **13 NOTA TÉCNICA nº 45 - REPRESENTAÇÃO DE RESTRIÇÕES DE INTERCÂMBIOS COM AGRUPAMENTO LIVRE - Atualização da Nota Técnica nº 21 (Março 2009)**

### **13.1 Introdução**

O modelo NEWAVE, em sua versão 14.5, permite a representação de restrições de intercâmbio máximo de energia entre os subsistemas e intercâmbio mínimo obrigatório entre os subsistemas. Para as restrições de intercâmbio de energia entre os subsistemas, é informada pelo usuário a capacidade da interligação, um fator de ponderação que irá representar a restrição em um dado patamar de carga, além do fator de perda associado a essa interligação.

Por questões de natureza elétrica, pode ser necessária a representação das restrições de intercâmbio não apenas pela capacidade máxima individual da linha, mas através de agrupamentos de interligações.

A funcionalidade descrita nesta Nota Técnica consiste em permitir a representação de restrições de intercâmbio que relacionem as interligações representadas, com a possibilidade de múltiplos agrupamentos destas interligações, considerando subsistemas reais e fictícios. Esta funcionalidade deve ser flexível a ponto de considerar, além dos limites individuais das interligações representadas, o limite do agrupamento como uma combinação linear de interligações.

Essa nova funcionalidade foi implementada a partir da versão 14.5 do programa NEWAVE gerando, dessa forma, uma nova versão intitulada versão 14.6.

### **13.2 Alterações**

Para representar os agrupamentos de intercâmbios foram incluídas novas restrições nos problemas de despacho hidrotérmico. Essas restrições são consideradas durante o cálculo da estratégia ótima de operação e durante a simulação final da política de operação. As restrições de agrupamento de intercâmbio não são relaxadas, de modo que devem ser sempre atendidas.

Cada restrição de agrupamento de intercâmbio, representada por uma combinação linear de limites de interligação existentes, está sujeita a um limite

máximo. As restrições são fornecidas por patamar de carga e o limite máximo para cada agrupamento é informado pelo usuário em arquivo específico.

De maneira geral as equações de agrupamento têm a seguinte forma, dada pela equação (1).

$$\sum_{j=1}^{N_{INT}^a} k_j^a F_{t,i,j,k_j}^m \leq LIMITE_a^m, 0 \leq k_j \leq 1, \forall j = 1, \dots, N_{INT}^a \quad (1)$$

onde:

$N_{INT}^a$ : número de interligações no agrupamento  $a$ ;

$k_j^a$ : coeficiente de ponderação da interligação  $j$  no agrupamento  $a$ ;

$F_{t,i,j,k_j}^m$ : intercâmbio de energia dos subsistemas  $i$  para  $k$ , no estágio  $t$ , patamar de carga  $m$  (MWmês);

$LIMITE_a^m$ : limite superior por patamar  $m$ , do agrupamento de interligações  $a$  (MWmês).

O valor  $LIMITE_a^m$  é calculado considerando-se a duração do mês correspondente e o limite do agrupamento  $a$  para o patamar  $m$  (MWmédio), informado em arquivo específico.

### 13.3 Entrada de dados

Para implementação dessa nova funcionalidade foi criado um novo arquivo de dados (AGRINT.DAT), onde são fornecidas as informações relativas aos agrupamentos de intercâmbio. Esse arquivo é composto de dois blocos de dados, onde no primeiro bloco são definidos os agrupamentos, as interligações que compõem a restrição e os coeficientes que relacionam os valores destas interligações na restrição. No segundo bloco são definidos os limites por patamar de carga para todos os agrupamentos declarados no primeiro bloco. Estes limites podem ser definidos ao longo do horizonte de planejamento. Para os períodos estáticos inicial e final serão considerados os limites do primeiro e do último período de planejamento, respectivamente.

Além disso, dois arquivos de entrada tiveram seu formato alterado: Dados Gerais (DGER.DAT) e Lista de arquivos (ARQUIVOS.DAT). No DGER.DAT pode-se optar pela consideração ou não da nova funcionalidade. No arquivo ARQUIVOS.DAT,

foi incluído um novo registro referente ao nome do arquivo com os parâmetros a serem utilizados nessa nova funcionalidade. Esse registro é de existência obrigatória.

#### ***13.4 Rebatimento nos demais modelos da cadeia***

A consideração de agrupamentos de intercâmbios já está implementada no modelo DECOMP, porém esta funcionalidade não está implementada nos demais modelos da cadeia energética.

## 14 NOTA TÉCNICA nº 46 - Equalização das penalidades de intercâmbio de energia entre os subsistemas (Março 2009)

Nessa nota técnica é descrita a funcionalidade de equalização das penalidades de intercâmbio de energia entre subsistemas, implementada na versão 14.4 do modelo NEWAVE.

Para que a solução ótima do problema de planejamento da operação seja aderente à operação energética do sistema elétrico brasileiro, são adotadas pequenas penalidades para as variáveis de decisão intercâmbio de energia, vertimento e excesso de energia. Sem a adoção dessas penalidades, soluções com exportação de deficit ou exportação de vertimento poderiam estar dentre as soluções ótimas do problema e impressas pelo programa como solução. Os valores adotados para as penalidades são de R\$ 0,005/MWh para o intercâmbio, R\$ 0,001 para o vertimento e R\$ 0,002 para o excesso de energia. Esses valores foram obtidos através de estudos, levando-se em conta que as penalidades deveriam ter os menores valores possíveis de forma que sua influência nos custos marginais de operação seja desprezível e que não causem problemas numéricos durante a execução do programa.

Na Figura 7 é apresentada a configuração adotada nos estudos de planejamento da operação energética. Seja a penalidade associada ao intercâmbio de energia representada por  $p$ . Para essa configuração, pode-se observar que intercâmbios entre os subsistemas Sudeste e Norte têm o custo unitário de  $2p$ , entre os subsistemas Sudeste e Sul o custo unitário de intercâmbio é  $p$  e entre os subsistemas Sudeste e Nordeste, o custo pode ser igual a  $p$  ou  $2p$ , dependendo da linha.

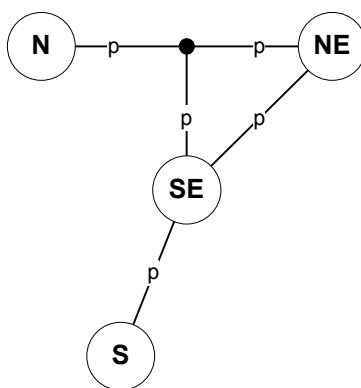
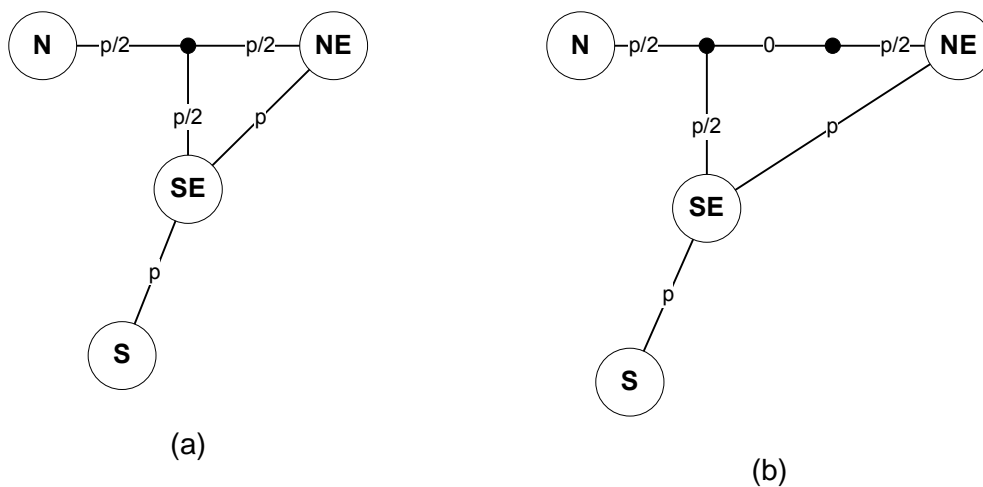


Figura 7 – Configuração com 4 subsistemas, com interligação SE-NE



Pela configuração da Figura 7, intercâmbios de energia entre os subsistemas Sudeste e Nordeste ocorrem preferencialmente pela linha que interliga esses dois sistemas diretamente, em detrimento do caminho Sudeste – Imperatriz – Nordeste.

Com a funcionalidade descrita nessa nota técnica, qualquer intercâmbio de energia entre dois subsistemas terá o custo unitário  $p$ . Isso é garantido adotando-se como penalidade de intercâmbio entre um sistema real e um fictício o valor de  $p/2$  e entre dois subsistemas fictícios o valor zero, conforme Figura 8.



**Figura 8 – Penalidades adotadas para os intercâmbios de energia.**

A opção de adotar a funcionalidade de equalização de penalidades de intercâmbio de energia é feita através de um registro no arquivo de dados gerais (e.g. *dger.dat*).

## 15 NOTA TÉCNICA nº 47 - Consideração de penalidades ao não atendimento da restrição de vazão mínima obrigatória (Junho 2009)

### 15.1 Introdução

A vazão mínima de um aproveitamento é uma vazão defluente obrigatória, turbinada ou vertida, que tem por finalidade manter um nível mínimo de um rio para fins de navegação, abastecimento, controle de poluição, irrigação e lazer, entre outros. No NEWAVE, onde o sistema hidrotérmico tem uma representação agregada, é calculada uma energia de vazão mínima correspondente a essa defluência obrigatória, para o conjunto de usinas com capacidade de regularização mensal. Essa energia é considerada no problema de despacho hidrotérmico como uma geração hidráulica controlável mínima, ou seja, obrigatória.

Como é possível que um subsistema não disponha de recurso suficiente para atender à restrição, na versão 14.7 do NEWAVE, descrita por essa nota técnica, foi implementada a possibilidade de consideração de uma penalidade associada ao não atendimento dos requisitos de energia de vazão mínima obrigatória. Essa penalidade é fornecida pelo usuário através de um arquivo específico para tal finalidade (e.g. *PENALID.DAT*). Foram incluídas restrições que representam o atendimento à meta de energia de vazão mínima obrigatória e variáveis de folga, representando o não atendimento desse requisito. Considerando que a meta em energia de vazão mínima é uma função da energia armazenada inicial do subsistema, que é um estado do problema, a consideração desta restrição provoca alterações no cálculo dos cortes de Benders. Adicionalmente foi implementada a opção de flexibilização das restrições de meta de vazão mínima obrigatória, permitindo que uma parte da violação seja penalizada a um valor menor.

Nas versões onde a violação ao requisito de defluência mínima obrigatória não era considerada explicitamente como uma variável de decisão do problema de otimização, a energia de vazão mínima era considerada como sendo controlável durante o cálculo da política de operação. Essa consideração se fez necessária para evitar problemas relacionados à convexidade da função de custo futuro. Com a implementação descrita nesse documento, a consideração dessa restrição no cálculo da política ótima de operação tornou-se possível.

## 15.2 Modelagem

A implementação da alteração descrita nessa nota técnica caracteriza-se pela criação de uma penalização à violação dos requisitos de atendimento à vazão mínima obrigatória, através da inclusão de uma variável de folga nas restrições de atendimento à demanda e balanço hídrico, penalizada na função objetivo, de forma a evitar que diferentes condições iniciais possam levar a um mesmo custo de operação, conforme colocado anteriormente. Além disso, a penalização na função objetivo permite que a função de custo futuro possa vislumbrar um possível não atendimento à vazão mínima obrigatória no futuro, de forma que a violação à descarga mínima obrigatória possa ser minimizada.

Para um dado estado em um instante  $t$ , a meta de atendimento à energia de vazão mínima obrigatória,  $M_{EVMIN}(EA_t)$ , que é função da energia armazenada, pode ser expressa como a soma da energia efetivamente atendida ( $evmin$ ) mais uma folga ( $\delta_{EVMIN}$ ), representando a parcela da meta que não foi atendida. Isso é expresso pela equação (1), denominada Equação de Atendimento à Meta de Energia de Vazão Mínima - EMAVM.

$$M_{EVMIN}(EA_t) = evmin + \delta_{EVMIN} \quad (1)$$

A seguir serão apresentadas as restrições do problema de despacho hidrotérmico do NEWAVE afetadas por essa implementação, além da função objetivo do problema. Entretanto, esse problema é apresentado aqui de forma simplificada, omitindo-se diversas restrições e variáveis, visando facilitar o entendimento da nova implementação.

### Função objetivo:

A função objetivo do problema passa a considerar a parcela de vazão mínima eventualmente não atendida,  $\delta_{EVMIN}$ , penalizada por  $\psi_{EVMIN}$ . O valor de  $\psi_{EVMIN}$  é definido pelo usuário. Logo, a função objetivo do problema passa a ser:

$$z_t = \min \left[ \sum_{NSIS} \left( \sum_{NPMC} \left[ \sum_{TCLISIS} \psi_T \cdot g_T + \sum_{NPDF} \psi_D \cdot def \right] + \psi_{EVMIN} \cdot \delta_{EVMIN} \right) + \frac{1}{1+\beta} \alpha_{t+1} \right] \quad (2)$$

onde  $NSIS$  é o número de subsistemas,  $NPMC$  é o número de patamares de mercado,  $TCLISIS$  é o número de classes térmicas de um determinado subsistema,

$NPDF$  é o número de patamares de déficit,  $g_T$  é a geração térmica de uma determinada classe,  $def$  é o déficit em um subsistema,  $\beta$  é a taxa que traz o custo futuro a valor presente e  $\alpha_{t+1}$  é o valor do custo futuro.

### Equações de balanço hídrico – EBH:

Conforme dito acima, uma folga para o atendimento de vazão mínima obrigatória,  $\delta_{EVMIN}$ , deve ser incluída na equação de balanço hídrico. Considerando essa folga, tem-se, de uma forma simplificada:

$$ea_{t+1} + \sum^{NPMC} g_H + evert - \delta_{EVMIN} = EA_t + EC_t - \\ - M_{EVMIN}(EA_t) - EVAP(EA_t) - EMORTO \quad (3)$$

onde  $ea_{t+1}$  é a energia armazenada no final do mês  $t$ ,  $g_H$  é a geração hidráulica controlável,  $evert$  é o vertimento,  $EA_t$  é a energia armazenada no início do mês  $t$ ,  $EC_t$  é a energia controlável,  $EVAP(EA_t)$  é a evaporação e  $EMORTO$  é a energia de enchimento de volume morto.

Substituindo (1) em (3):

$$ea_{t+1} + \sum^{NPMC} g_H + evert + evmin = EA_t + EC_t - EVAP(EA_t) - EMORTO \quad (4)$$

### Equações de atendimento à demanda – EAD:

Assim como nas equações de balanço hídrico, nas de atendimento à demanda também será incluída a folga para energia de vazão mínima. Deve-se observar que, uma vez que o estudo pode considerar o mercado discretizado em mais de um patamar de carga, haverá, para cada subsistema, uma restrição de atendimento à demanda para cada patamar. Nesse caso, para cada restrição, tanto a meta de vazão mínima quanto a folga devem ser multiplicadas pelo número de horas do patamar, expresso na variável  $FPENG$ . Considerando a variável de folga, tem-se:

$$g_H + \sum^{TCLIS} g_T + \sum^{NPDF} def + \sum i_{recebido} - \sum i_{exportado} - exc - FPENG \cdot \delta_{EVMIN} = \\ [DEMLIQ - EFIO - M_{EVMIN}(EA_t)] \cdot FPENG \quad (5)$$

Substituindo (1) em (5) chega-se a:

$$g_H + \sum^{TCL\SIS} g_T + \sum^{NPDF} def + \sum i_{recebido} - \sum i_{exportado} - exc + FPENG \cdot evmin = \quad (6)$$

$$= [DEMLIQ - EFIO] \cdot FPENG$$

### Restrições de geração hidráulica máxima controlável – EEH:

A restrição de geração hidráulica máxima tem por objetivo limitar a geração hidráulica controlável ( $g_H$ ). A energia máxima que pode ser gerada pelos reservatórios é dada pela folga entre a geração hidráulica máxima e a geração de usinas a fio d'água mais energia de vazão mínima. Pela proposta do trabalho, a variável de folga para a energia de vazão mínima também deve ser incluída nesta restrição.

Considerando  $\delta_{EVMIN}$ , tem-se:

$$g_H - \delta_{EVMIN} \leq [\overline{GH}(EA_t) - EFIO - M_{EVMIN}(EA_t)] \cdot FPENG \quad (7)$$

Substituindo (1) em (7):

$$g_H + FPENG \cdot evmin \leq [\overline{GH}(EA_t) - EFIO] \cdot FPENG \quad (8)$$

#### 15.2.1 Construção dos Cortes

A introdução de uma nova restrição e novas variáveis ao problema de planejamento de operação causa alterações no cálculo dos cortes de Benders. A restrição (2), que representa a meta de geração de vazão mínima, tem a variável  $M_{EVMIN}$ , que é função do armazenamento inicial. Por outro lado, tal variável não figura mais nas restrições de balanço hídrico, atendimento a demanda e geração hidráulica máxima.

O coeficiente do corte de Benders associado ao estado armazenamento no início do período  $t$  é dado pela equação (9).

$$\pi_V = \sum^{NSIS} \eta_{BH} \frac{\partial}{\partial EA_t} EBH + \sum^{NPMC} \sum^{NSIS} \eta_{AD} \frac{\partial}{\partial EA_t} EAD + \sum^{NCOR} \eta_{COR} \frac{\partial}{\partial EA_t} ECOR +$$

$$+ \sum^{NPMC} \sum^{NSIS} \eta_{EH} \frac{\partial}{\partial EA_t} EEH + \sum^{NSIS} \eta_{EV} \frac{\partial}{\partial EA_t} EMAVM \quad (9)$$

Sejam as equações a seguir que representam as parábolas do fator de correção da energia controlável, da meta de energia de vazão mínima, energia evaporada e geração hidráulica máxima:

$$FC(EA_t) = a_{EC} EA_t^2 + b_{EC} EA_t + c_{EC} \quad (10)$$

$$M_{EVMIN}(EA_t) = a_{EZ} EA_t^2 + b_{EZ} EA_t + c_{EZ} \quad (11)$$

$$EVAP(EA_t) = a_{EV} EA_t^2 + b_{EV} EA_t + c_{EV} \quad (12)$$

$$\overline{GH}(EA_t) = a_{EH} EA_t^2 + b_{EH} EA_t + c_{EH} \quad (13)$$

Além disso:

$$EC_t = FC(EA_t) \cdot \gamma \cdot EAF_t \quad (14)$$

onde  $FC(EA_t)$  é o fator de correção de energia controlável,  $\gamma$  é a parcela da energia natural afluente que pode ser controlada,  $EAF_t$  é a energia natural afluente.

Substituindo (10) e (12) em (4):

$$ea_{t+1} + \sum^{NPMC} g_H + evert + evmin = EA_t + (a_{EC} EA_t^2 + b_{EC} EA_t + c_{EC}) \cdot \gamma \cdot EAF_t - (a_{EV} EA_t^2 + b_{EV} EA_t + c_{EV}) - EMORTO \quad (15)$$

A derivada de (15) em relação à  $EA_t$  é dada por:

$$\frac{\partial}{\partial EA_t} EBH = 1 + \gamma \cdot EAF_t (2a_{EC} EA_t + b_{EC}) - (2a_{EV} EA_t + b_{EV}) \quad (16)$$

A equação de atendimento à demanda, não possui nenhum termo em função de  $EA_t$ . Logo,

$$\frac{\partial}{\partial EA_t} EAD = 0 \quad (17)$$

Substituindo (13) em (8), a restrição de geração hidráulica máxima pode ser expressa como:

$$g_H + FPENG \cdot evmin \leq \left[ (a_{EH} EA_t^2 + b_{EH} EA_t + c_{EH}) - (1 - \gamma) \cdot EAF_t \right] \cdot FPENG \quad (18)$$

Derivando (18) em relação à  $EA_t$ :

$$\frac{\partial}{\partial EA_t} EEH = (2a_{EH} EA_t + b_{EH}) \cdot FPENG \quad (19)$$

Finalmente, a nova restrição de atendimento aos requisitos de energia de vazão mínima obrigatória, considerando (11), pode ser escrita como:

$$evmin + \delta_{evmin} = a_{EZ} EA_t^2 + b_{EZ} EA_t + c_{EZ} \quad (20)$$

A derivada de (20) em relação ao armazenamento inicial é dada por:

$$\frac{\partial}{\partial EA_t} EMAVM = (2a_{EV} EA_t + b_{EV}) \quad (21)$$

Finalmente, substituindo-se (16), (17), (19) e (21) em (9), é obtida para um subsistema  $i$  o seguinte coeficiente da equação de cortes de Benders:

$$\begin{aligned} \pi_V = & \eta_{BH} \left[ 1 + \gamma \cdot EAF_t (2a_{EC} EA_t + b_{EC}) - (2a_{EV} EA_t + b_{EV}) \right] + \\ & + \eta_{EH} (2a_{EH} EA_t + b_{EH}) + \\ & + \eta_{EV} (2a_{EV} EA_t + b_{EV}) \end{aligned} \quad (22)$$

Quanto à derivada do problema em relação às energias afluentes passadas, não há qualquer modificação, uma vez que a meta de geração de vazão não é função desses estados.

### 15.3 Flexibilização da restrição de vazão mínima obrigatória

Outra funcionalidade disponível é a possibilidade de flexibilização da restrição de vazão mínima obrigatória. A meta de vazão mínima será dividida em duas partes. Nesse caso, no arquivo de penalidades devem ser informados dois valores, um para cada violação. A partição aplicada à meta de vazão mínima,  $\varphi \in [0;1]$ , será calculada em função dos valores informados no arquivo de alteração cadastral (e.g. *modif.dat*). Nesse arquivo, podem ser informados até dois valores de vazão mínima, sendo que o segundo valor deve ser obrigatoriamente menor que o primeiro. O valor de  $\varphi$  é dado por:

$$\varphi = \frac{M'_{EVMIN}}{M_{EVMIN}}$$

onde  $M_{EVMIN}$  é a energia de vazão mínima das usinas calculada com os valores de cadastro ou o primeiro valor do arquivo de modificação cadastral, caso informado, e  $M'_{EVMIN}$  é a energia de vazão mínima das usinas calculada com os valores de cadastro ou o segundo valor do arquivo de modificação cadastral, caso informado.

Matematicamente, a restrição da equação (1) passa a ser descrita por duas restrições:

$$\varphi M_{EVMIN} (EA_t) = evmin1 + \delta1_{EVMIN} \quad (23a)$$

$$(1 - \varphi) M_{EVMIN} (EA_t) = evmin2 + \delta2_{EVMIN} \quad (23b)$$

onde  $\varphi \in [0;1]$ .

O problema de otimização é escrito como:

$$z_t = \min \sum^{NSIS} \left( \sum^{NPMC} \left[ \sum^{TCLISIS} \psi_T \cdot g_T + \sum^{NPDF} \psi_D \cdot def \right] + \psi1_{EVMIN} \cdot \delta1_{EVMIN} + \psi2_{EVMIN} \cdot \delta2_{EVMIN} \right) + \frac{1}{1 + \beta} \alpha_{t+1} \quad (24)$$

Sujeito a:

$$ea_{t+1} + \sum^{NPMC} g_H + evert + evmin1 + evmin2 = EA_t + EC_t - EVAP(EA_t) - EMORTO \quad (25)$$

$$g_H + \sum^{TCLISIS} gt + \sum^{NPDF} def + \sum i_{recebido} - \sum i_{exportado} - exc + FPENG \cdot (evmin1 + evmin2) = [DEMLIQ - EFIO] \cdot FPENG \quad (26)$$

$$g_H + FPENG \cdot (evmin1 + evmin2) \leq [\overline{GH}(EA_t) - EFIO] \cdot FPENG \quad (27)$$



$$\varphi M_{EVMIN}(EA_t) = evmin1 + \delta 1_{EVMIN} \quad (23a)$$

$$(1 - \varphi) M_{EVMIN}(EA_t) = evmin2 + \delta 2_{EVMIN} \quad (23b)$$

A consideração da inclusão destas restrições no cálculo do corte de Benders é feita de forma análoga, sendo que a equação (22) passa a ser escrita como:

$$\begin{aligned} \pi_v = & \eta_{BH} [1 + \gamma \cdot EAF_t (2a_{EC} EA_t + b_{EC}) - (2a_{EV} EA_t + b_{EV})] + \\ & + \eta_{EH} (2a_{EH} EA_t + b_{EH}) + \\ & + \eta_{EV1} (2a_{EV} EA_t + b_{EV}) \cdot \varphi + \\ & + \eta_{EV2} (2a_{EV} EA_t + b_{EV}) \cdot (1 - \varphi) \end{aligned} \quad (28)$$

## 16 NOTA TÉCNICA nº 55 - PROCEDIMENTO DE SOLUÇÃO DA BIBLIOTECA DE OTIMIZAÇÃO (Maio 2009)

O algoritmo de solução de problemas de programação linear adotado nos modelos de planejamento energético desenvolvidos no CEPEL é o Simplex. Esse algoritmo foi proposto por Dantzig em 1947 [1] e é amplamente utilizado em todo o mundo até os dias atuais. Esse é um método iterativo para encontrar uma solução que minimize (ou maximize) a função objetivo.

Pode ser mostrado que se existe ao menos um ótimo para um problema de programação linear, uma solução ótima estará em um vértice do poliedro que define a região viável do problema. Cada um dos vértices desse poliedro é chamado de *solução básica viável*.

A proposta do algoritmo Simplex é buscar entre as soluções básicas viáveis, qual aquela que otimiza o problema. A uma *base* corresponde um subconjunto das variáveis do problema. O processo iterativo do Simplex consiste em escolher, a cada iteração, qual a variável que não pertence a esse conjunto passará a pertencer e qual deixará de pertencer. A variável a ser escolhida para entrar na base é tal que a função objetivo do problema reduza (para o caso de minimização), ou ao menos não se altere. Para isso, calcula-se para cada variável o *custo reduzido*, correspondente ao impacto na função objetivo da variação do valor de cada variável. Nesse caso, a variável deve possuir o custo reduzido negativo ou nulo. A variável a sair da base é tal que se consiga a máxima redução possível na função objetivo do problema ao se introduzir a nova variável escolhida para a base, mantendo o problema viável.

Existem na literatura diversas heurísticas para a escolha da variável do problema a entrar na base (também chamadas *pricing*). Na ferramenta de solução de problemas de otimização adotado pelos modelos de planejamento energético desenvolvidos no CEPEL, o desenvolvedor disponibiliza duas dessas heurísticas: o *random pricing* e o *Devex pricing*[2].

Independente da heurística adotada (e da ferramenta de otimização), sucessivas resoluções de problemas de otimização pelo método Simplex, sem que a ferramenta seja reinicializada, podem levar a uma acumulação de erros numéricos, dificultando que o ótimo seja encontrado. Nesses casos, podem ser adotadas diversas ações. Normalmente, nos programas do CEPEL que envolvem programação linear, as

variáveis da biblioteca de otimização são reiniciadas e o problema remontado quando o ótimo não é encontrado. São tentadas então variantes do algoritmo Simplex (dual simplex, primal simplex). Finalmente, se o ótimo não for encontrado, a execução do programa é interrompida. Garante-se dessa forma que os resultados processados sejam sempre os ótimos. Por outro lado, conforme sugestão do próprio desenvolvedor, pode-se, antes do procedimento da reinicialização, mandar que a ferramenta resolva novamente o problema, partindo do ponto em que ela se encontrava, adotando a heurística que já estava sendo adotada. Essa ação foi implementada nas versões 14.0.1, 14.1.1, 14.6.1 e 14.8.2 do NEWAVE e no DECOMP a partir da versão 14.9.3, mantendo-se a heurística em uso naquele momento, ou seja, o *Devex Pricing*.

## REFERÊNCIAS

- [1] Dantzig, G. B. *Linear Programming and Extensions*. Princeton, New Jersey: Princeton University Press, 1963.
- [2] Harris, P. M. J. "Pivot Selection Methods of the Devex LP Code." *Mathematical Programming* 5,1–28 (1973).

## 17 NOTA TÉCNICA nº 56 - Execução do Módulo de Geração de Cenários Sintéticos de Energia Natural Afluente num Ambiente de Processamento Distribuído (Setembro 2009)

### 17.1 Introdução

O modelo NEWAVE, utilizado no planejamento da operação energética de médio prazo é composto basicamente pelos seguintes módulos:

- *Módulo de cálculo dos sistemas equivalentes*, onde são calculados os subsistemas equivalentes de energia;
- *Módulo de geração de cenários de energia natural afluente*, onde são estimados os parâmetros do modelo estocástico e geradas séries sintéticas de energias afluentes;
- *Módulo de cálculo da política de operação hidrotérmica*, onde é determinada a política de operação mais econômica para os subsistemas equivalentes, tendo em conta as incertezas nas afluências futuras, os patamares de demanda e a indisponibilidade dos equipamentos;
- *Módulo de simulação da operação*, onde é realizada a simulação da operação do sistema ao longo do período de planejamento, para distintos cenários de seqüências hidrológicas.

No módulo de geração de cenários de energia natural afluente são gerados diversos cenários hidrológicos que são utilizados durante as simulações forward e backward do processo de cálculo da política ótima de operação hidrotérmica. Esses cenários são obtidos através de um modelo autorregressivo periódico de ordem  $p$ , PAR( $p$ ), que modela a afluência de um mês como sendo função das afluências dos  $p$  meses anteriores.

Com o intuito de aprimorar a árvore de cenários a ser visitada durante o processo do cálculo da política ótima de operação é utilizada Amostragem Seletiva de Cenários. A Amostragem Seletiva consiste em aplicar técnicas de agregação no procedimento de geração dos cenários de energia natural afluente para as simulações

forward e backward. Neste caso, as técnicas de agregação são empregadas para a geração da amostra de ruídos  $N(0,1)$  multivariados que é utilizada pelo modelo  $PAR(p)$ .

A aplicação das técnicas de agregação no processo de geração de cenários de energia natural afluyente aumenta bastante o esforço computacional gasto neste processo, principalmente quando o tamanho da amostra cresce. De forma a reduzir o tempo computacional gasto no processo de geração das séries sintéticas, pode-se aplicar técnicas de processamento paralelo no módulo de geração de cenários de energia natural afluyente. O objetivo desta Nota Técnica é descrever como as técnicas de paralelização foram implementadas e quais os benefícios obtidos com sua utilização.

Esta implementação deu origem à versão 13.3.6-3 do programa NEWAVE.

## 17.2 Módulo de Geração de Cenários de ENA

No modelo  $PAR(p)$ , a energia natural afluyente de cada subsistema e a cada mês é composta pela soma de uma parcela determinística, dada por uma combinação linear das  $p$  afluências passadas mais recentes mais uma constante, com uma parcela estocástica, dada por um ruído aleatório de valor esperado nulo, de acordo com (1). O ruído é concebido como um sorteio de uma distribuição lognormal três parâmetros multivariada (cada dimensão corresponde a um subsistema). Para um maior detalhamento do modelo  $PAR(p)$ , consulte [1].

$$Z_t = \mu_m + \sum_{j=1}^{p_m} \phi_j^m \sigma_m \left( \frac{Z_{t-j} - \mu_{m-j}}{\sigma_{m-j}} \right) + \sigma_m a_t = PD_t + PE_t \quad (1)$$

O processo de geração de cenários de energia natural afluyente pode ser dividido de maneira bem sucinta em três etapas:

Etapa 1: Identificação e ajuste do modelo estatístico a ser utilizado, no caso do NEWAVE é utilizado o modelo autorregressivo periódico.

Etapa 2: Geração da amostra de ruídos (parcela estocástica) a ser utilizada para construção dos cenários hidrológicos utilizados nas simulações backward e forward.

Etapa 3: Geração dos cenários hidrológicos (parcela determinística + parcela estocástica) utilizados nas simulações backward e forward.

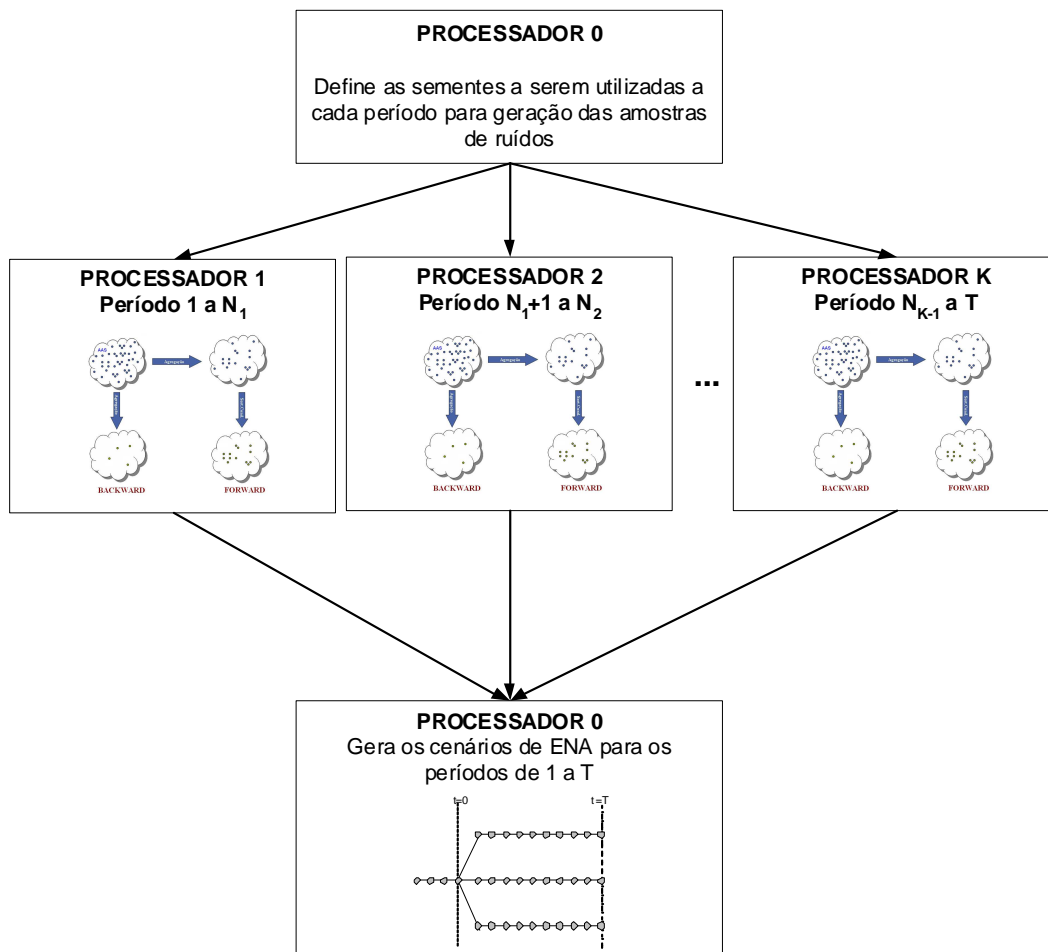
As Etapas 2 e 3 são repetidas a cada período do horizonte de estudo.

Na Etapa 2, inicialmente é gerada uma amostra de ruídos normais padrão independentes, em seguida é aplicada uma transformação linear de forma a tornar os ruídos correlacionados e, finalmente, é realizado o ajuste de uma distribuição lognormal 3 parâmetros. As técnicas de agregação são aplicadas nesta etapa, no passo referente à geração da amostras de ruídos normais padrão independentes. Para maiores detalhes sobre Amostragem Seletiva, consulte [2].

Em uma análise do processo de geração de cenários hidrológicos considerando a utilização de técnicas de agregação, foi observado que a segunda etapa é aquela que absorve a maior parcela do tempo computacional gasto neste processo. O tempo gasto com esta etapa aumenta consideravelmente à medida que o número de elementos da amostra original do processo de agregação cresce.

### ***17.3 Aplicação da Técnica de Paralelização***

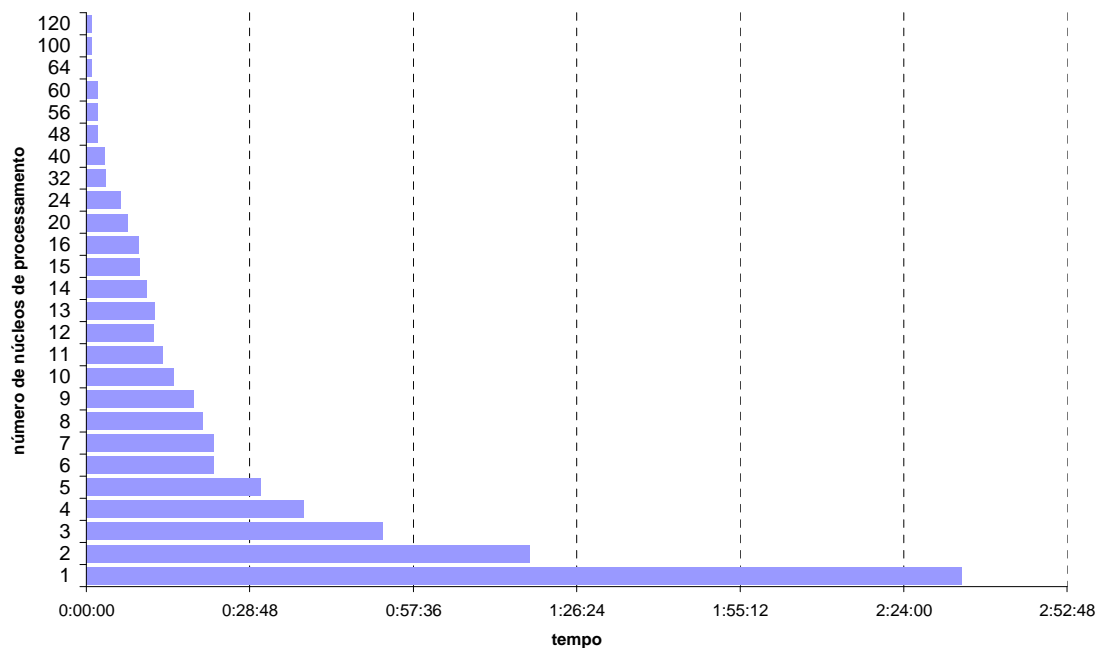
Como as amostras de ruídos normais padrão independentes geradas no primeiro passo da Etapa 2 não apresentam dependência temporal, é possível dividir/distribuir a tarefa de gerá-las a cada estágio entre diferentes processadores. Após a obtenção das amostras de ruídos para todos os estágios, é dada continuidade ao processo de geração de cenários hidrológicos. Na Figura 1 é ilustrada a distribuição dos processos entre os diversos núcleos de processamento.



**Figura 1 – Esquema de paralelização do processamento**

No caso de uma amostra original de ruído com tamanho de 100 mil objetos e dimensão igual a 4 (subsistemas), o tempo gasto verificado na etapa de geração dos cenários hidrológicos foi de 2 horas e meia com um núcleo de processamento de um processador quad-core Xeon da Intel modelo X5355 com frequência de processamento de 266 MHz. Após a aplicação de técnicas de paralelização no processo de geração das amostras de ruídos, o tempo gasto reduziu para aproximadamente 9 minutos, considerando 16 processadores em paralelo e, 3 minutos considerando 32 processadores. Na Figura 2 é apresentado o tempo de execução do módulo de geração de cenários de energia natural afluyente, considerando diversos números de núcleos de processamento.

**Tempo Gasto no Processo de Geração  
200x20 - Amostra original = 100 mil - 4 subsistemas**



**Figura 2 – Tempo de processamento**

## 17.4 Referências

[1] 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.

[2] CEPEL (2009). “Aplicação de Técnicas de Agregação na Geração de Cenários Hidrológicos para o Planejamento de Médio Prazo”, Nota Técnica 42 – Projeto NEWAVE.