

**RELATÓRIO TÉCNICO**

Nº: DP/DEA – 46994/09      Nº DE PÁGINAS: 309      ANEXOS: 3

**TÍTULO:** Definição da árvore de cenários de afluências para o planejamento da operação energética de médio prazo.

**DEPARTAMENTO:** DEPARTAMENTO DE OTIMIZAÇÃO ENERGÉTICA E MEIO AMBIENTE – DEA  
**ÁREA DE RESPONSABILIDADE:** B200      **Nº DO PROJETO:** 1798


**DESTINATÁRIOS:**  
Centrais Elétricas Brasileiras S/A – ELETROBRÁS  
Av. Presidente Vargas, 409 – 12º andar  
20071-003 – Rio de Janeiro / RJ  
  
**ATENÇÃO:**  
Dr. Paulo Roberto de Holanda Sales

**EQUIPE DE ACOMPANHAMENTO:**  
Renata Falcão – ELETROBRÁS  
Lilian Sampaio – ELETROBRÁS  
Arlete Neves – ELETROBRÁS  
Carlos Alberto - ELETROBRÁS

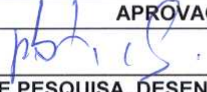
**RESUMO:** Este trabalho propõe um método para definir uma árvore de cenários completa, levando em consideração a representação da dependência temporal e espacial, tal que o espaço de estados relativo ao processo estocástico das afluências seja coberto da melhor forma possível. A partir da árvore completa de cenários seleciona-se uma sub-árvore adequada à solução do problema por algoritmos da PDDE, que não degrade excessivamente a qualidade de representação da árvore completa, que permita viabilidade computacional, que não traga impactos negativos na convergência do processo iterativo de cálculo da política ótima de operação e que seja robusta a variações no tamanho das amostras e ao uso de diferentes sementes. Desta maneira, pretende-se melhorar a qualidade da função de custo futuro que é construída pelo NEWAVE e repassada aos demais modelos da cadeia de planejamento energético e produzir resultados mais robustos de forma que pequenas variações na árvore de afluências não ocasionem grandes variações na solução do problema.

**AUTORES:**  
Débora Dias Jardim Penna - CEPEL  
Maria Elvira Piñeiro Maceira – CEPEL  
Jorge Machado Damazio – CEPEL

**PALAVRAS-CHAVE:**  
Planejamento da operação energética; geração de árvore de cenários; técnicas de agregação; técnicas de amostragem  
**CLASSIFICAÇÃO:** CONTROLADO

  
**GERENTE DO PROJETO**  
**NOME:** Débora Dias Jardim Penna  
**TEL:** (21) 2598- 6450      **FAX:** (21) 2598-6482  
**EMAIL:** [debora@cepel.br](mailto:debora@cepel.br)

  
**CHEFE DE DEPARTAMENTO**  
**NOME:** Maria Elvira Piñeiro Maceira  
**TEL:** (21) 2598-6454      **FAX:** (21) 2598-6482  
**EMAIL:** [elvira@cepel.br](mailto:elvira@cepel.br)

**APROVAÇÃO**  
  
**DIRETOR DE PESQUISA, DESENVOLVIMENTO E INOVAÇÃO**  
**ROBERTO PEREIRA CALDAS**  
*29/12/09*

# **PROJETO GEVAZP**

**DEFINIÇÃO DA ÁRVORE DE CENÁRIOS DE  
AFLUÊNCIAS PARA O PLANEJAMENTO DA  
OPERAÇÃO ENERGÉTICA DE MÉDIO PRAZO**

**[gevazp@cepel.br](mailto:gevazp@cepel.br)**

## Agradecimentos

Os autores agradecem aos engenheiros do Operador Nacional do Sistema Elétrico (ONS) Alberto Kligerman, Cecília Mércio, Maria Cândida Lima, Maria Helena de Azevedo, Joari Costa e Murilo Soares pelas enriquecedoras discussões.

## Sumário

1	Introdução	1
1.1.	Considerações Iniciais	1
1.2.	Contexto do Trabalho	3
1.3.	Objetivos do Trabalho	5
1.4.	Metodologia Proposta	5
1.5.	Relevância do Trabalho	6
1.6.	Organização do Trabalho	6
2	Planejamento da Operação Energética	8
2.1.	Considerações Iniciais	8
2.2.	Formulação do Problema	9
2.2.1.	Sistemas Térmicos	9
2.2.2.	Sistemas Hidrotérmicos	10
2.3.	Planejamento da Operação Energética no Sistema Brasileiro	13
2.3.1.	Cadeia de Modelos	15
2.3.2.	Planejamento de Médio Prazo	16
2.3.3.	Planejamento de Curto Prazo	19
2.3.4.	Programação da Operação	19
2.4.	Resumo	20
3	Planejamento da Operação de Médio Prazo	21
3.1.	Considerações Iniciais	21
3.2.	Representação da Incerteza Hidrológica na PDDE	22
3.3.	Construção da Árvore de Cenários	26
3.4.	Resumo	30
4	Modelo de Geração de Séries Sintéticas de Energia	31
4.1.	Considerações Iniciais	31
4.2.	Processos Estocásticos e Séries Temporais	32
4.3.	O Modelo Autorregressivo Periódico	35

4.3.1. Descrição do Modelo	35
4.3.2. Ajuste do Modelo	38
4.3.3. Geração de Séries Sintéticas com o Modelo PAR( $\rho$ )	40
4.3.4. Vazões Incrementais Negativas	44
4.3.5. Correlação Espacial	45
4.4. Avaliação do Desempenho do Modelo	46
4.4.1. Geração em Paralelo	46
4.4.2. Geração em Árvore	51
4.5. Resumo	54
5 Definição da Árvore de Cenários	55
5.1. Considerações Iniciais	55
5.2. Técnicas de Seleção de Cenários	55
5.2.1. Métodos de Agrupamento	57
5.2.2. Método K-MEANS	60
5.3. Técnicas de Amostragem	61
5.3.1. Amostragem por Hipercubo Latino	61
5.3.2. Amostragem Descritiva	64
5.3.3. Quase-Monte Carlo	65
5.4. Resumo	68
6 Aplicação do Modelo de Geração de Cenários	69
6.1. Considerações Iniciais	69
6.2. Alternativas para Construção da Árvore de Afluências	69
6.3. Construção dos Cortes de Benders	75
6.4. Reamostragem de Cenários (recombinação de ruídos)	77
6.5. Considerações Finais	80
6.6. Resumo	81
7 Avaliação dos Cenários Hidrológicos Gerados	83
7.1. Considerações Iniciais	83
7.2. Análise da amostra de ruídos	84
7.2.1. Amostra de ruídos utilizada no processo de agregação	84
7.2.2. Amostra de ruídos utilizada na geração dos cenários do passo forward	90
7.2.3. Amostra de ruídos utilizada na geração dos cenários do passo	

backward	94
7.2.4. Variando o tamanho da amostra de ruídos forward e backward	98
7.2.5. Amostra de ruídos utilizando LHC e AD	106
7.3. Análise dos Cenários Gerados para Passo Forward	109
7.3.1. Envoltória dos Cenários	110
7.3.2. Média dos Cenários	115
7.3.3. Desvio-padrão dos Cenários	117
7.3.4. Distribuição Univariada dos Cenários	119
7.3.5. Correlação Cruzada dos Cenários	122
7.3.6. Análise de Seqüências Negativas	129
7.4. Análise dos Cenários Gerados para Passo Backward	135
7.4.1. Envoltória dos Cenários	136
7.4.2. Testes não condicionados para média e desvio	141
7.4.3. Distribuição Univariada dos Cenários	143
7.4.4. Testes condicionados para média e desvio	148
7.4.5. Correlação Cruzada dos Cenários	151
7.5. Resumo	159
8 Avaliação dos Resultados do Problema de Planejamento da Operação	161
8.1. Considerações Iniciais	161
8.2. Avaliação da convergência	163
8.3. Alteração da amostra de cenários	166
8.4. Alteração do número de cenários backward	171
8.5. Alteração do número de cenários forward	174
8.6. Alteração do tamanho da amostra para o processo de agregação	179
8.7. Alteração do método de amostragem	182
8.8. Combinação de método de amostragem e agregação	186
8.9. Verificação da estabilidade dos resultados	192
8.10. Resumo	202
9 Conclusões e Sugestões para Trabalhos Futuros	203
9.1. Sugestões para trabalhos futuros	206
10 Referências Bibliográficas	208

Apêndice A Modelo Newave	215
A.1. Introdução	215
A.2. Modelo Newave	216
A.2.1. Módulo de Cálculo do Sistema Equivalente	218
A.2.2. Módulo de Cálculo da Política de Operação	221
A.2.2.1. Programação Dinâmica Dual Estocástica Aplicada ao Planejamento da Operação Hidrotérmica	222
A.2.2.2 Despacho de Operação Hidrotérmica em Sistemas Equivalentes de Energia	226
A.2.3. Módulo de Simulação da Operação	230
Apêndice B Análise dos Cenários Gerados de Energia Natural Afluente (Complementação)	231
B.1 Análise dos Cenários Gerados para Passo Forward	231
B.1.1 Média dos Cenários	231
B.1.2 Desvio-padrão dos Cenários	236
B.2. Análise dos Cenários Gerados para Passo Backward	240
B.2.2 Testes não Condicionados para Média e Desvio-padrão	240
B.2.2 Testes Condicionados para Média e Desvio-padrão	246
B.2.3 Testes Condicionados para Correlação Cruzada	251
Apêndice C Avaliação dos Resultados do Problema de Planejamento da Operação (Complementação)	262

## Lista de figuras

Figura 1: Processo de decisão para um sistema hidrotérmico	3
Figura 2: Custos Imediato e Futuro X Armazenamento	11
Figura 3: Uso ótimo da água	12
Figura 4: Horizontes do Planejamento da Operação e Principais Decisões em cada Estágio	14
Figura 5: Representação de Incertezas e do Detalhamento do Sistema com o Horizonte de Estudo	15
Figura 6: Cadeia de Modelos Computacionais para o Planejamento da Expansão e Operação Energética	18
Figura 7: Exemplo de árvore de cenários (Árvore Completa)	23
Figura 8: Exemplo de sub-árvore de cenários do NEWAVE	24
Figura 9: Exemplo CEPEL - árvore completa	25
Figura 10: Exemplo CEPEL – sub-árvores	26
Figura 11: Opção Atual	28
Figura 12: Sorteio Condicionado	28
Figura 13: Opção 1AAS	29
Figura 14: Relação Demanda x Risco em um Ano Qualquer	31
Figura 15: Vazão Afluente Anual x Anos	32
Figura 16: Tendência Hidrológica	41
Figura 17: Geração não condicionada – Primeira Etapa	42
Figura 18: Geração não condicionada – Segunda Etapa	42
Figura 19: Geração em Paralelo (Pente) para a simulação Forward	43
Figura 20: Seqüência de geração para a simulação Backward	44
Figura 21: Geração em Paralelo (Pente) para a simulação Backward	44
Figura 22: Seqüência Negativa	47
Figura 23: Volume de regularização (Déficit)	49
Figura 24: Aplicação técnicas de agregação	57
Figura 25: Exemplo ilustrativo do processo aglomerativo	58
Figura 26: Dendograma	59
Figura 27: Divisão em 5 intervalos do domínio de duas VA (a) distribuição normal (b) distribuição uniforme	62
Figura 28: Valores sorteados para cada variável	63



Figura 29: Representação bi-dimensional de uma possível amostragem por hipercubo latino	63
Figura 30: Representação da amostragem utilizando método Quase-Monte Carlo (Sobol) (a) 100 pontos (b) 1000 pontos	66
Figura 31: Amostra (a) Seqüência de Sobol (b) Amostragem Aleatória Simples	67
Figura 32: Amostra (a) Seqüência de Sobol [50] (b) Seqüência de Halton [20]	67
Figura 33: Aplicação do Procedimento de Agregação	70
Figura 34: Escolha do Objeto Representativo	70
Figura 35: Probabilidade dos Cenários Forward	72
Figura 36: Opção 0	72
Figura 37: Opção 1	73
Figura 38: Opção 2	73
Figura 39: Opção 3	74
Figura 40: Opção 4	75
Figura 41: Construção da FCF – modelo NEWAVE	76
Figura 42: Árvore Completa – Exemplo	78
Figura 43: Sub-árvore (1º. amostra) – Exemplo	78
Figura 44: – Sub-árvore – Exemplo (a) 2º. amostra e (b) 3º. Amostra	79
Figura 45: Sub-árvore após diversas reamostragens – Exemplo	79
Figura 46: Tendência Hidrológica Recente (%MLT)	84
Figura 47: Estimativa Média (amostra de ruídos)	85
Figura 48: Estimativa Desvio-Padrão (amostra de ruídos)	86
Figura 49: Número de Rejeições (amostra de ruídos)	87
Figura 50: Intervalos da Distribuição Univariada	87
Figura 51: Exemplificação Classe 34	88
Figura 52: Distribuição Multivariada – Amostra com Tamanho 2 mil a 100 mil	89
Figura 53: Estatística t – Média – Amostra Forward (ruído)	91
Figura 54: Estatística t – Desvio-padrão – Amostra Forward (ruído)	92
Figura 55: Teste de Aderência – Amostra Forward (ruído)	93
Figura 56: Distribuição Multivariada – Amostra Forward (ruído)	94
Figura 57: Estatística t – Desvio-padrão Média – Amostra Backward (ruído)	95
Figura 58: Estatística t – Desvio-padrão – Amostra Backward (ruído)	96
Figura 59: Teste de Aderência – Amostra Backward (ruído)	97
Figura 60: Distribuição Multivariada – Amostra Backward (ruído)	97

Figura 61: Estatística t – Média – Amostra Forward (ruído) – Variando Tamanho Amostra Forward	98
Figura 62: – Estatística t – Desvio-Padrão – Amostra Forward (ruído) – Variando Tamanho Amostra Forward	99
Figura 63: Distribuição – Amostra Forward (ruído) – 250x20 – Opção Atual	99
Figura 64: Distribuição – Amostra Forward (ruído) – 300x20 – Opção Atual	100
Figura 65: Distribuição – Amostra Forward (ruído) – 250x20 – Opção 0	100
Figura 66: Distribuição – Amostra Forward (ruído) – 300x20 – Opção 0	100
Figura 67: Distribuição – Amostra Forward (ruído) – 250x20 – Opção 4	101
Figura 68: Distribuição – Amostra Forward (ruído) – 300x20 – Opção 4	101
Figura 69: Estatística t – Média – Amostra Backward (ruído) – Variando Tamanho Amostra Backward	102
Figura 70: Estatística t – Desvio-Padrão – Amostra Backward (ruído) – Variando Tamanho Amostra Backward	102
Figura 71: Estatística t – Média – Amostra Forward (ruído) – Variando Tamanho Amostra Backward	103
Figura 72: Estatística t – Desvio-Padrão – Amostra Forward (ruído) – Variando Tamanho Amostra Backward	104
Figura 73: Distribuição – Amostra Backward (ruído) – 200x50 – Opção Atual	104
Figura 74: Distribuição – Amostra Backward (ruído) – 200x100 – Opção Atual	105
Figura 75: Distribuição – Amostra Backward (ruído) – 200x50 – Opção 0 e 4	105
Figura 76: Distribuição – Amostra Backward (ruído) – 200x100 – Opção 0 e 4	105
Figura 77: Estatística t – Média – Amostra Forward (ruído) – Variando Método Amostragem	106
Figura 78: Estatística t – DP – Amostra Forward (ruído) – Variando Método Amostragem	107
Figura 79: Estatística t – Média – Amostra Backward (ruído) – Variando Método Amostragem	107
Figura 80: Estatística t – DP – Amostra Backward (ruído) – Variando Método Amostragem	108
Figura 81: Distribuição – Amostra Forward (ruído) – LHC	108
Figura 82: Distribuição – Amostra Backward (ruído) – LHC	109
Figura 83: Distribuição – Amostra Forward (ruído) – AD	109
Figura 84: Distribuição – Amostra Backward (ruído) – AD	109
Figura 85: Envoltória Cenários Forward – Sudeste	111

Figura 86: Envoltória Cenários Forward – Sul	112
Figura 87: Envoltória Cenários Forward – Nordeste	113
Figura 88: Envoltória Cenários Forward – Norte	114
Figura 89: Média Cenários Forward – Sudeste	116
Figura 90: Desvio-Padrão Cenários Forward – Sudeste	118
Figura 91: Teste de Aderência Cenários Forward – Sudeste	120
Figura 92: Teste de Aderência Cenários Forward – Sul	120
Figura 93: Teste de Aderência Cenários Forward – Nordeste	121
Figura 94: Teste de Aderência Cenários Forward – Norte	122
Figura 95: Correlação Cruzada Cenários Forward – Sudeste x Sul	123
Figura 96: Correlação Cruzada Cenários Forward – Sudeste x Nordeste	124
Figura 97: Correlação Cruzada Cenários Forward – Sudeste x Norte	125
Figura 98: Correlação Cruzada Cenários Forward – Sul x Nordeste	126
Figura 99: Correlação Cruzada Cenários Forward – Sul x Norte	127
Figura 100: Correlação Cruzada Cenários Forward – Nordeste x Norte	128
Figura 101: Teste com Diferentes Níveis de Regularização – Opção Atual	131
Figura 102: Teste com Diferentes Níveis de Regularização – Opção 0	133
Figura 103: Teste com Diferentes Níveis de Regularização – Opções 1 e 1AAS	134
Figura 104: Teste com Diferentes Níveis de Regularização – Opções 2, 3 e 4	134
Figura 105: Amostra utilizada no teste não condicionado	135
Figura 106: Amostras utilizadas no teste condicionado	136
Figura 107: Envoltória Cenários Forward – Sudeste	137
Figura 108: Envoltória Cenários Backward – Sul	138
Figura 109: – Envoltória Cenários Backward – Nordeste	139
Figura 110: Envoltória Cenários Backward – Norte	140
Figura 111: Média Cenários Backward – Sudeste	143
Figura 112: Teste de Aderência Cenários Backward – Sudeste	145
Figura 113: Teste de Aderência Cenários Backward – Sul	146
Figura 114: Teste de Aderência Cenários Backward – Nordeste	147
Figura 115: Teste de Aderência Cenários Backward – Norte	148
Figura 116: Estatística t Média e Desvio-Padrão – Cenários Backward - Sudeste	150
Figura 117: Correlação Cruzada Cenários Backward – Sudeste x Sul	153

Figura 118: CC Backward x Forward – Sudeste x Sul – Opção Atual	154
Figura 119: CC Backward x Forward – Sudeste x Sul – Opção 0	154
Figura 120: CC Backward x Forward – Sudeste x Sul – Opção 1	155
Figura 121: CC Backward x Forward – Sudeste x Sul – Opção 4	155
Figura 122: CC Backward x Forward – Sudeste x Sul – Opção 1AAS	156
Figura 123: CC Backward x Forward – Sudeste x Sul – Opção LHC	156
Figura 124: CC Backward x Forward – Sudeste x Sul – Opção AD	157
Figura 125: Comparação entre CC Forward e Backward	159
Figura 126: Diagrama esquemático das usinas hidroelétricas do SIN 2008-2012	161
Figura 127: Representação dos subsistemas equivalentes	162
Figura 128: Convergência – Última Iteração	166
Figura 129: Convergência – Última Iteração – Variação Amostra – FEV/07	167
Figura 130: Trajetória Convergência – Variação Amostra – Opção Atual – FEV/07	168
Figura 131: Trajetória Convergência – Variação Amostra – Opção 4 – FEV/07	169
Figura 132: Convergência – Última Iteração – Variação Amostra – Com Reamostragem - FEV/07	170
Figura 133: CMO – Variação Amostra – Com Reamostragem - FEV/07	170
Figura 134: COPER – Variação Amostra – Com Reamostragem - FEV/07	171
Figura 135: Risco – Variação Amostra – Com Reamostragem - FEV/07	171
Figura 136: EENS – Variação Amostra – Com Reamostragem - FEV/07	171
Figura 137: Convergência – Última Iteração – Variação Backward – Com Reamostragem - FEV/07	172
Figura 138: COPER – Variação Backward – Com Reamostragem - FEV/07	173
Figura 139: CMO – Variação Backward – Com Reamostragem - FEV/07	173
Figura 140: Risco – Variação Backward – Com Reamostragem - FEV/07	174
Figura 141: EENS – Variação Backward – Com Reamostragem - FEV/07	174
Figura 142: Convergência – Última Iteração – Variação Forward – Com Reamostragem - FEV/07	175
Figura 143: Resultados – Variação Forward – Com Reamostragem - FEV/07	175
Figura 144: Convergência – Última Iteração – Variação Forward & Amostra – Opção Atual - FEV/07	176
Figura 145: Resultados – Variação Forward & Amostra –	

Opção Atual - FEV/07	176
Figura 146: Convergência – Última Iteração – Variação Forward & Amostra – Opção 0 - FEV/07	177
Figura 147: Resultados – Variação Forward & Amostra – Opção 0 - FEV/07	178
Figura 148: Convergência – Última Iteração – Variação Forward & Amostra – Opção 4 - FEV/07	178
Figura 149: Resultados – Variação Forward & Amostra – Opção 4 - FEV/07	179
Figura 150: Convergência – Última Iteração – Variação Forward & Amostra – Opção 4 - FEV/07	180
Figura 151: COPER – Variação Tamanho da Amostra - FEV/07	180
Figura 152: CMO – Variação Tamanho da Amostra - FEV/07	181
Figura 153: Risco – Variação Tamanho da Amostra - FEV/07	181
Figura 154: EENS – Variação Tamanho da Amostra - FEV/07	182
Figura 155: Convergência – Última Iteração – Variação Método de Amostragem - FEV/07	183
Figura 156: Resultados- Variação Método de Amostragem - FEV/07	184
Figura 157: Convergência – Última Iteração – Variação Método de Amostragem e Backward - FEV/07	185
Figura 158: Resultados- Variação Método de Amostragem e Backward - FEV/07	185
Figura 159: Convergência – Última Iteração – Variação Método de Amostragem e Forward - FEV/07	186
Figura 160: Resultados- Variação Método de Amostragem e Forward - FEV/07	186
Figura 161: Opções com combinação de LHC com agregação	187
Figura 162: Convergência – Última Iteração – Variação Amostra – Combinação LHC e Agregação - FEV/07	188
Figura 163: CMO - Variação Amostra – Combinação LHC e Agregação - FEV/07	188
Figura 164: COPER - Variação Amostra – Combinação LHC e Agregação - FEV/07	189
Figura 165: Risco - Variação Amostra – Combinação LHC e Agregação - FEV/07	189
Figura 166: EENS - Variação Amostra – Combinação LHC e Agregação - FEV/07	189

Figura 167: Convergência – Última Iteração – Variação Forward – Combinação LHC e Agregação - FEV/07	190
Figura 168: Resultados - Variação Forward – Combinação LHC e Agregação - FEV/07	191
Figura 169: Convergência – Última Iteração – Variação Backward – Combinação LHC e Agregação - FEV/07	191
Figura 170: Resultados - Variação Backward – Combinação LHC e Agregação - FEV/07	192
Figura 171: Resultados 45ª iteração – Opção Atual	195
Figura 172: Resultados 45ª iteração – Opção 0	196
Figura 173: Resultados 45ª iteração – Opção 4	198
Figura 174: Resultados 45ª iteração – Opção 0 com reamostragem	199
Figura 175: Resultados 45ª iteração – Opção 4 com reamostragem	201
Figura 176: Esquema de relacionamento entre os módulos do programa NEWAVE	218
Figura 177: Cálculo da Função de Custo Futuro através da Programação Dinâmica	224
Figura 178: Média Cenários Forward – Sul	232
Figura 179: Média Cenários Forward – Nordeste	234
Figura 180: Média Cenários Forward – Norte	235
Figura 181: Desvio-Padrão Cenários Forward – Sul	237
Figura 182: Desvio-Padrão Cenários Forward – Nordeste	239
Figura 183: Desvio-Padrão Cenários Forward – Norte	240
Figura 184: Média Cenários Backward – Sul	242
Figura 185: Média Cenários Backward – Nordeste	244
Figura 186: Média Cenários Backward – Norte	246
Figura 187: Estatística t Média e Desvio-Padrão – Cenários Backward – Sul	248
Figura 188: Estatística t Média e Desvio-Padrão – Cenários Backward - Nordeste	249
Figura 189: Estatística t Média e Desvio-Padrão – Cenários Backward - Norte	251
Figura 190: Correlação Cruzada Cenários Backward – Sudeste x Nordeste	253
Figura 191: Correlação Cruzada Cenários Backward – Sudeste x Norte	255
Figura 192: Correlação Cruzada Cenários Backward – Sul x Nordeste	257

Figura 193: Correlação Cruzada Cenários Backward –Sul x Norte	259
Figura 194: Correlação Cruzada Cenários Backward – Nordeste x Norte	261
Figura 195: Resultados– Variação Backward - PMO Mai/04	262
Figura 196: Resultados– Variação Backward - PMO Set/04	263
Figura 197: Resultados– Variação Backward - PMO Set/09	264
Figura 198: Resultados– Variação Backward - PMO Jun/06	265
Figura 199: Resultados– Variação Backward - PMO Jul/06	266
Figura 200: Resultados– Variação :Backward - PMO Jan/07	267
Figura 201: Resultados– Variação Backward - PMO Fev/07	267
Figura 202: Resultados– Variação Backward - PMO Mai/07	268
Figura 203: Resultados– Variação Backward - PMO Jul/07	269
Figura 204: Resultados– Variação Backward - PMO Ago/07	269
Figura 205: Resultados– Variação Backward - PMO Jan/08	270
Figura 206: Resultados– Variação Forward - PMO Mai/04	271
Figura 207: Resultados– Variação Forward - PMO Set/04	272
Figura 208: Resultados– Variação Forward - PMO Set/09	272
Figura 209: Resultados– Variação Forward - PMO Jun/06	273
Figura 210: Resultados– Variação Forward - PMO Jul/06	274
Figura 211: Resultados– Variação Forward - PMO Jan/07	275
Figura 212: Resultados– Variação Forward - PMO Fev/07	275
Figura 213: Resultados– Variação Forward - PMO Mai/07	276
Figura 214: Resultados– Variação Forward - PMO Jul/07	277
Figura 215: Resultados– Variação Forward - PMO Ago/07	278
Figura 216: Resultados– Variação Forward - PMO Jan/08	278
Figura 217: Resultados– Variação Semente – Opção 4 - PMO Set/05	279
Figura 218: Resultados– Variação Semente – PMO Set/04	280
Figura 219: Resultados– Variação Semente – PMO Jul/07	281
Figura 220: Resultados– Variação Semente – PMO Jan/08	281
Figura 221: Resultados– Variação Amostra & Semente – Opção 4 - PMO Jan/08	282
Figura 222: Resultados– Variação Forward & Semente – Opção 4 - PMO Set/04	283
Figura 223: Resultados– Variação Forward & Semente – Opção 4 - PMO Jul/07	284
Figura 224: Resultados– Variação Forward & Semente –	

Opção 4 - PMO Jan/08	284
Figura 225:Resultados– Variação Forward & Amostra – Opção 4 - PMO Jul/07	285
Figura 226:Resultados– Variação Forward & Amostra – Opção 4 - PMO Jan/08	286
Figura 227:Resultados– Variação Semente – Opção 4 & Variantes - PMO Set/04	287
Figura 228:Resultados– Variação Amostra – Opção 4 & Variantes - PMO Jan/08	287
Figura 229:Resultados– Variação Backward – Opção 4 & Variantes - PMO Set/04	288
Figura 230:Resultados– Variação Backward – Opção 4 & Variantes - PMO Jul/06	289
Figura 231:Resultados– Variação Backward – Opção 4 & Variantes - PMO Jan/08	290
Figura 232:Resultados– Variação Forward – Opção 4 & Variantes - PMO Set/04	290
Figura 233:Resultados– Variação Forward – Opção 4 & Variantes - PMO Jul/06	291
Figura 234:Resultados– Variação Forward – Opção 4 & Variantes - PMO Jan/08	292



## Lista de tabelas

Tabela 1: Valores Padrão para IC 95%	53
Tabela 2: Correlação Cruzada Média Cenários Forward	128
Tabela 3: Análise das Seqüências Negativas – Opção Atual	129
Tabela 4: Teste de Máximos – Opção Atual	130
Tabela 5: Análise das Seqüências Negativas – Opção 0	131
Tabela 6: Análise das Seqüências Negativas – Opções 1 e 1AAS	131
Tabela 7: Análise das Seqüências Negativas – Opções 2, 3 e 4	132
Tabela 8: Teste de Máximos – Opção 0	132
Tabela 9: Teste de Máximos – Opções 1 e 1AAS	132
Tabela 10: Teste de Máximos – Opções 2, 3 e 4	133
Tabela 11: Correlação Cruzada Média Cenários Backward - Não Condicionado	157
Tabela 12: Número de iterações para convergência	164
Tabela 13: Número de iterações para convergência (cont.)	164

## Lista de siglas e abreviaturas

AD – amostragem descritiva  
AAS – amostragem aleatória simples  
CCEE – Câmara de Comercialização de Energia Elétrica  
Cepel – Centro de Pesquisas de Energia Elétrica  
CMO – custo marginal de operação  
COPER – valor esperado do custo total de operação para a simulação final  
EENS - valor esperado da energia não suprida  
ENA – energia natural afluyente  
EPE – Empresa de Pesquisa Energética  
FCF – função de custo futuro  
FCI – função de custo imediato  
LHC – hipercubo latino (sigla em inglês)  
MME – Ministério de Minas e Energia  
NLEQ – número de cenários backward (aberturas)  
NSIM – número de cenários forward  
ONS – Operador Nacional do Sistema Elétrico  
PAR(p) – modelo autorregressivo periódico de ordem p (sigla em inglês)  
PDD – programação dinâmica determinística  
PDDE – programação dinâmica dual estocástica ( em inglês SDDP)  
PDE – programação dinâmica estocástica  
PMO – programa mensal de operação  
SIN – sistema interligado nacional  
ZINF - limite inferior do valor esperado do custo total de operação  
ZSUP – limite superior do valor esperado do custo total de operação

# 1 Introdução

## 1.1. Considerações Iniciais

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

O planejamento da operação energética de um sistema hidrotérmico é uma tarefa difícil e deve ser feita de forma coordenada por se tratar de um problema:

- Acoplado no tempo: os reservatórios das usinas hidrelétricas possuem capacidade limitada de armazenamento e a decisão de geração em um estágio tem como consequência o nível de armazenamento no final do estágio, que será o nível de armazenamento inicial do próximo estágio. Logo, a decisão de geração em um estágio altera o montante disponível de geração para os estágios futuros. Desta forma, surge uma dependência entre a decisão operativa de um determinado período e as decisões futuras.
- Acoplado no espaço: normalmente o parque hidrelétrico está disposto em diversas bacias hidrográficas, com várias usinas hidrelétricas dispostas em uma mesma cascata, onde a operação de usinas a montante afeta a operação das demais usinas a jusante.
- Estocástico: há grande incerteza em relação às afluências futuras, que variam sazonal e regionalmente e podem possuir períodos secos de longa duração. Além disso, existem ainda as incertezas com relação à demanda de energia elétrica.

O problema de planejamento da operação energética deve ainda considerar as restrições operativas associadas às usinas hidrelétricas e termelétricas e as restrições elétricas do sistema.

A princípio pode-se imaginar que a água armazenada nos reservatórios não possui valor associado, logo seria mais vantajoso atender a demanda utilizando primeiramente a água estocada nos reservatórios evitando o acionar de usinas termelétricas, que possuem um custo direto associado ao seu combustível. Porém, considerando que a capacidade de armazenamento em um sistema é limitada pela capacidade de seus reservatórios e que a afluência futura aos aproveitamentos é desconhecida, pode-se dizer que há uma dependência entre a decisão operativa de hoje e os custos operativos do futuro. Na Figura 1 é ilustrada essa dependência.

Na Figura 1 é ilustrada a relação que existe entre uma decisão de despacho tomada no presente e sua consequência no futuro. Se, por exemplo, for tomada uma decisão de utilizar a energia hidrelétrica para atender a demanda presente e, no futuro, ocorrer uma seca (baixas afluências), poderá ser necessário o uso de geração térmica de custo elevado ou, até mesmo, interromper o fornecimento de energia. Logo o custo de operação será elevado. Se, por outro lado, a opção for de atender a demanda presente com geração térmica e no futuro ocorrer uma cheia (afluências altas), poderá ocorrer um vertimento no subsistema, caracterizando um desperdício de energia e consequentemente um aumento desnecessário do custo de operação.

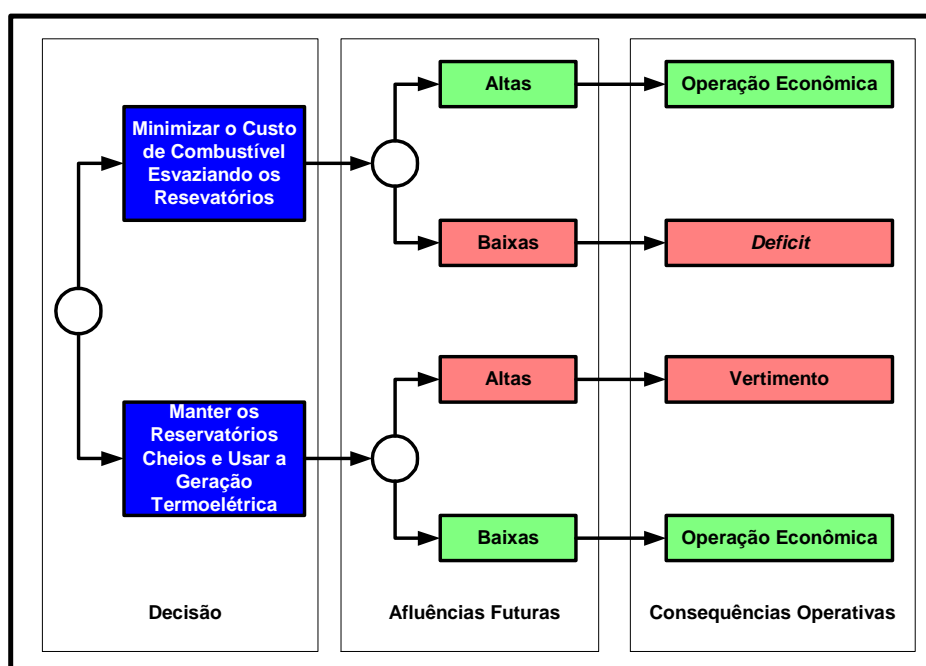


Figura 1: Processo de decisão para um sistema hidrotérmico

Desta forma, em sistemas hidrotérmicos o benefício da utilização da água armazenada nos reservatórios em um determinado estágio é medido em função da economia de combustível das térmicas e déficits futuros.

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

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

## **1.2. Contexto do Trabalho**

Com as mudanças do marco regulatório em 1999, o Operador Nacional do Sistema Elétrico (ONS) se tornou o responsável pelo planejamento da operação energética do Sistema Interligado Nacional (SIN). Nessa tarefa o ONS utiliza uma cadeia de modelos matemáticos desenvolvida pelo Centro de Pesquisas de Energia Elétrica (CEPEL). No eixo principal dessa cadeia estão os modelos NEWAVE, DECOMP e DESSEM<sup>1</sup> que têm como objetivo definir a política ótima de operação para os problemas de médio prazo, curto prazo e programação diária, respectivamente, considerando o critério de minimização dos custos de operação ao longo do horizonte de planejamento.

Para o planejamento da expansão de mais longo prazo (20 a 30 anos) é utilizado o modelo MELP com o objetivo de selecionar dentre uma lista de projetos aqueles que serão candidatos à construção. Com a retomada dos estudos de planejamento da expansão pelo Ministério de Minas e Energia (MME) e pela Empresa de Pesquisa Energética (EPE), o modelo NEWAVE vem sendo utilizado para auxiliar a elaboração dos Planos Decenais de Expansão (PDE).

---

<sup>1</sup> Vale ressaltar que o modelo DESSEM ainda se encontra em fase de validação pelo ONS e pelos agentes do setor elétrico brasileiro

O modelo de planejamento de médio prazo, NEWAVE, pode considerar um horizonte temporal de 5 a 10 anos com discretização mensal. Para reduzir a dimensionalidade do problema é utilizada uma representação simplificada do parque gerador, através de sistemas equivalentes de energia. A estocasticidade das afluições é representada por um número muito grande de cenários hidrológicos gerados sinteticamente por um modelo estocástico periódico autorregressivo de ordem  $p$  (PAR( $p$ )). O cálculo da política de operação é baseado na programação dinâmica dual estocástica (PDDE).

No NEWAVE, a política de operação, ou estratégia ótima de operação, é representada pela função de custo futuro e calculada por um processo iterativo para um conjunto de estados (energia armazenada no início do estágio e tendência hidrológica<sup>2</sup>). Para cada estado, o corte da função de custo futuro corresponde a uma média estimada para um conjunto de aberturas<sup>3</sup>.

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

Como a árvore de cenários do problema de planejamento de médio prazo possui uma cardinalidade bastante elevada, igual ao número de aberturas elevado ao número de estágios do horizonte de planejamento (normalmente igual  $20^{120-1}$ ), torna-se impossível do ponto de vista computacional percorrer completamente a árvore. Portanto, apenas uma porção da árvore (sub-árvore) é percorrida.

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

---

<sup>2</sup> O conjunto dos últimos valores observados em uma série hidrológica

<sup>3</sup> Conjunto de afluições utilizada durante a recursão backward

### **1.3. Objetivos do Trabalho**

Este trabalho propõe um método para definir uma árvore de cenários completa, levando em consideração a representação da dependência temporal e espacial, tal que o espaço de estados relativo ao processo estocástico das afluências seja coberto da melhor forma possível. A partir da árvore completa de cenários seleciona-se uma sub-árvore adequada à solução do problema por algoritmos da PDDE, que não degrade excessivamente a qualidade de representação da árvore completa, que permita viabilidade computacional, que não traga impactos negativos na convergência do processo iterativo de cálculo da política ótima de operação e que seja robusta a variações no tamanho das amostras e ao uso de diferentes sementes.

Desta maneira, pretende-se melhorar a qualidade da função de custo futuro que é construída pelo NEWAVE e repassada aos demais modelos da cadeia de planejamento energético e produzir resultados mais robustos de forma que pequenas variações na árvore de afluências não ocasionem grandes variações na solução do problema.

### **1.4. Metodologia Proposta**

Para a definição da árvore completa e da sub-árvore a ser visitada durante o processo do cálculo da estratégia ótima de operação serão empregadas no modelo de geração de cenários hidrológico:

- Técnicas de seleção de cenários: as técnicas de agregação de cenários, quando aplicadas a um grande número de cenários hidrológicos gerados, proporcionam a escolha de um conjunto representativo de cenários.
- Técnicas de amostragem: são propostos métodos que permitam uma forma de amostragem mais eficiente utilizando, por exemplo, amostragem estratificada e amostragem descritiva.

Tais técnicas podem ser aplicadas separadamente ou em conjunto com o intuito de representar de forma mais acurada o processo de estocástico de afluências.

Adicionalmente, um procedimento de reamostragem de cenários é aplicado durante o cálculo da política ótima de operação com o objetivo de percorrer uma porção maior da árvore completa de cenários.

## **1.5. Relevância do Trabalho**

No contexto do planejamento da operação energética para o setor elétrico brasileiro, o trabalho proposto pode ser utilizado por diversos agentes como: o ONS, na determinação de metas mais robustas para o planejamento energético de médio prazo; a Câmara de Comercialização de Energia Elétrica (CCEE), na definição de preços para a energia negociada no mercado a vista; o MME e a EPE, na elaboração dos planos de planejamento da expansão e no cálculo da garantia física para os empreendimentos de geração de energia elétrica; e pelas empresas do setor elétrico, para realização de estudos que auxiliem o processo de tomada de decisão e a elaboração estratégias de ação para o médio prazo.

## **1.6. Organização do Trabalho**

No presente capítulo foi feita uma contextualização do trabalho dentro do setor elétrico, foram apresentados a motivação que levou ao desenvolvimento deste trabalho e, em linhas gerais, a metodologia proposta para definir a árvore de cenários de afluência para o planejamento da operação energética de médio prazo.

No capítulo 2 discute-se o problema de planejamento da operação, em especial do sistema elétrico brasileiro, suas características básicas, as etapas do planejamento da operação e os modelos desenvolvidos para cada uma delas.

No capítulo 3 é apresentado um detalhamento da consideração hidrológica no modelo NEWAVE. Também são discutidas as atuais alternativas para construção da sub-árvore de cenários e os possíveis problemas que podem ocorrer caso a escolha da sub-árvore não seja apropriada.

Em seguida, no capítulo 4 é realizada uma revisão do modelo autorregressivo periódico e da metodologia de identificação, estimação e verificação para o modelo. Este modelo é atualmente utilizado para a geração de cenários hidrológicos para os modelos de planejamento da operação energética do sistema hidrelétrico brasileiro.

No capítulo 5 são descritas as técnicas de seleção de cenários e as técnicas de amostragem propostas para serem aplicadas ao modelo de geração de cenários hidrológicos utilizado no planejamento de médio prazo.



O capítulo 6 constitui a principal contribuição deste trabalho. Nele são apresentadas alternativas para definição da sub-árvore de cenários de afluência utilizando-se os subsídios dados nos capítulos anteriores.

No capítulo 7 é apresentada uma avaliação do ponto de vista hidrológico dos cenários de afluências gerados pelas alternativas propostas no capítulo anterior. Também é realizada uma avaliação das amostras que são utilizadas nas técnicas de seleção de cenários.

O capítulo 8 são feitos estudos de casos com o SIN, no qual são analisados os resultados do modelo NEWAVE considerando as alternativas propostas no capítulo 6.

No capítulo 9 constam as principais conclusões do trabalho e sugestões para desenvolvimentos futuros.

No Apêndice A é apresentado um detalhamento da formulação do modelo NEWAVE, utilizado no planejamento da operação de médio prazo. Adicionalmente, é feita uma breve descrição da programação dinâmica dual estocástica.

Finalmente, no Apêndice B é apresentada uma complementação da análise dos cenários gerados de energia natural afluente, e no Apêndice C é realizada uma complementação da avaliação dos resultados do problema de planejamento da operação.

Este trabalho baseia-se no desenvolvimento de doutorado da Engenheira Débora Dias Jardim Penna, sob supervisão de Maria Elvira P. Maceira e Reinaldo Castro Souza, tendo sido apresentado à PUC-Rio (Penna, 2009).

## 2 Planejamento da Operação Energética

### 2.1. Considerações Iniciais

O planejamento da operação energética tem por objetivo determinar uma política de operação para o sistema de geração de energia elétrica, composto por usinas hidrelétricas, usinas termelétricas e intercâmbios de energia entre os subsistemas, que minimize o custo esperado de operação ao longo do horizonte de planejamento. O custo de operação inclui custo de combustível das usinas térmicas e penalidades pelo não atendimento da demanda de energia.

Para o cálculo da política ótima de operação são usados modelos energéticos, que também podem ser utilizados em diversos estudos, tais como:

- Informações sobre consumo de combustível: através da simulação da operação do sistema para diversos cenários hidrológicos e de demanda, podem ser estimadas média e variância do consumo de combustível e valores de pico deste consumo;
- Estudo de políticas tarifárias: um dos resultados dos modelos de planejamento é o custo marginal de operação, que mede a variação do custo esperado de operação com relação a incrementos de demanda. Este índice é fundamental para a determinação de preços de intercâmbio e de tarifas de pedágio pelo uso do sistema de transmissão;
- Estudos de política de racionamento: os modelos de planejamento da operação representam o efeito de diferentes políticas de racionamento de energia, em termos de duração, profundidade e impactos econômicos e financeiros;
- Estudos de expansão da geração e transmissão do sistema.

Neste capítulo é apresentado de forma sucinta o problema de planejamento da operação energética com suas principais características e sua aplicação no sistema brasileiro. Ao final é descrita a cadeia de modelos desenvolvida pelo CEPEL e utilizada pelo ONS para auxiliar a operação do SIN.

## **2.2. Formulação do Problema**

### **2.2.1. Sistemas Térmicos**

Sistemas termelétricos são aqueles formados unicamente por usinas térmicas (Fortunato et al, 1990), como as usinas nucleares, usinas com combustíveis fósseis, usinas a gás, além de usinas de cogeração. De forma simplificada, pode-se dizer que essas usinas usam a queima de combustíveis para a obtenção de vapor, que aciona as turbinas acopladas ao gerador. As usinas nucleares trabalham de forma semelhante, porém obtêm calor a partir de reações nucleares.

Em estudos de planejamento, as usinas termelétricas são representadas através de características físicas e operativas, tais como potência máxima, tipo de combustível, taxa de tomada de carga, geração mínima operativa, dentre outros índices.

Em sistemas de geração puramente térmicos, o custo variável do combustível é o responsável pela definição do custo de cada usina. De forma simplificada, o problema de planejamento da operação resume-se em classificar as usinas em ordem crescente de custo de combustível e despachar cada MWh incremental para a usina com o menor custo, até que a demanda seja suprida. Este processo utilizado para a resolução deste problema de operação é chamado de despacho por ordem de mérito. A última usina a ser despachada será o gerador marginal e o custo marginal de operação será igual ao seu custo de combustível.

Na prática, este problema de otimização é um pouco mais complexo, já que devem ser levados em conta: a taxa de tomada de carga, tempos mínimos e máximos em operação, além de tempos de resfriamento (Silva, 2001).

As principais características dos sistemas puramente térmicos são:

- Desacoplados no tempo: uma decisão tomada em um período não tem impactos sobre o custo de operação de períodos posteriores, já que depende apenas da disponibilidade de combustível;

- Custo direto de operação: o custo de operação de cada usina térmica independe do nível de geração ou da disponibilidade de outras usinas;
- Confiabilidade de fornecimento: não apresentam restrições de armazenamento, dependendo apenas da capacidade total de atendimento das usinas.

### **2.2.2. Sistemas Hidrotérmicos**

Os sistemas hidrotérmicos caracterizam-se pela presença de usinas hidrelétricas, adicionalmente às térmicas. Diferentemente do sistema visto no item anterior, esses sistemas contam com a energia armazenada nos reservatórios de água, de custo teoricamente zero, em substituição às dispendiosas usinas térmicas.

Entretanto, o volume limitado dos reservatórios aliado à variabilidade das futuras afluições produz uma relação entre a decisão operativa em um determinado estágio e as conseqüências futuras desta decisão, conforme ilustrado na Figura 1. Desta forma, por exemplo, se for tomada uma decisão de utilizar a energia hidrelétrica para atender a carga própria presente e baixas afluições ocorrerem no futuro, as usinas hidrelétricas podem não ter água suficiente para suprir a demanda no futuro. Como conseqüência, poderá ser necessária a utilização de recursos térmicos extremamente caros ou, até mesmo, a interrupção do fornecimento de energia. Se, por outro lado, a opção for de atender a carga própria presente com geração térmica, e altos valores de afluições ocorrerem no futuro, já estando elevados os níveis dos reservatórios, poderá haver vertimento no sistema, caracterizando um desperdício de energia e aumento desnecessário do custo de operação.

O problema do planejamento da operação para sistemas hidrotérmicos possui as seguintes características:

- Acoplamento temporal: como visto, as decisões tomadas no presente têm conseqüências no futuro. A solução ótima é obtida minimizando-se o benefício presente do uso da água mais o benefício futuro de seu armazenamento;
- Natureza estocástica: há incerteza a respeito das afluições futuras no momento em que a decisão operativa é tomada;

- Acoplamento espacial: como as usinas podem estar dispostas em cascata, a decisão de deplecionamento de uma usina a montante afeta a afluência total à usina a jusante;
- Custos não-diretos associados à geração de uma hidrelétrica: o custo do uso da água armazenada nos reservatório pode ser medido em termos da economia resultante do custo de combustível das térmicas não despachadas ou déficits evitados devido ao seu uso no futuro;
- Competição entre os objetivos do problema: a máxima utilização da energia hidrelétrica disponível a cada estágio minimiza os custos de combustível no estágio. Entretanto, esta política resulta em maiores riscos de déficits futuros. Por sua vez, a máxima confiabilidade de fornecimento é obtida conservando o nível dos reservatórios o mais elevado possível. Entretanto, isto leva a uma maior utilização de geração térmica e, portanto, aumento nos custos de operação do estágio.

Como visto no item anterior, na tomada de decisão da operação de um sistema hidrotérmico deve-se comparar o benefício imediato do uso da água e o benefício futuro de seu armazenamento. O benefício do uso imediato da água pode ser representado por uma função chamada Função de Custo Imediato (FCI), enquanto que o benefício de armazená-la no presente para o seu uso futuro pode ser representado através de uma Função de Custo Futuro (FCF). Estas duas funções estão ilustradas no gráfico da Figura 2. O eixo x, das abscissas, representa o volume final armazenado nos reservatórios das usinas hidrelétricas do sistema, e o eixo y, das ordenadas, representa os valores da FCF ou FCI expressos em unidades monetárias.

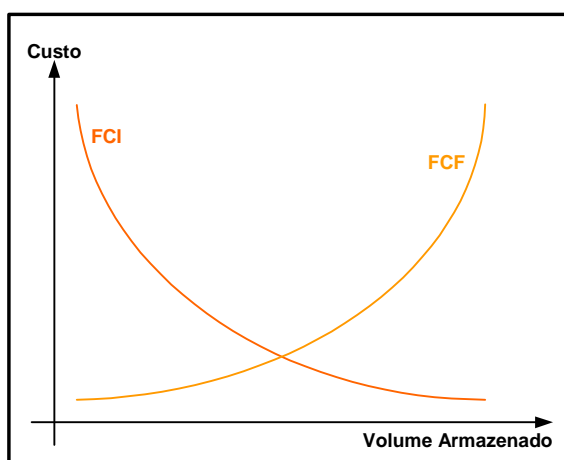


Figura 2: Custos Imediato e Futuro X Armazenamento

Como é de se esperar, a FCI aumenta com o volume final armazenado nos reservatórios. Isto ocorre porque a decisão de economizar água no presente está relacionada a um maior gasto com geração térmica no estágio atual para o atendimento à carga. Por outro lado, a FCF diminui com o volume final armazenado nos reservatórios, porque a decisão de economizar água no presente está relacionada a um menor uso de geração térmica no futuro.

O uso ótimo da água armazenada corresponde ao ponto que minimiza a soma dos custos imediato e futuro. Como é mostrado na Figura 3, o ponto de mínimo custo global também corresponde ao ponto onde as derivadas da FCI e da FCF com relação ao armazenamento (V) se igualam em módulo (equação 2.1). Estas derivadas são conhecidas como valores da água.

$$\frac{\partial(FCI + FCF)}{\partial V} = \frac{\partial FCI}{\partial V} + \frac{\partial FCF}{\partial V} = 0 \Rightarrow \frac{\partial FCI}{\partial V} = -\frac{\partial FCF}{\partial V} \quad (2.1)$$

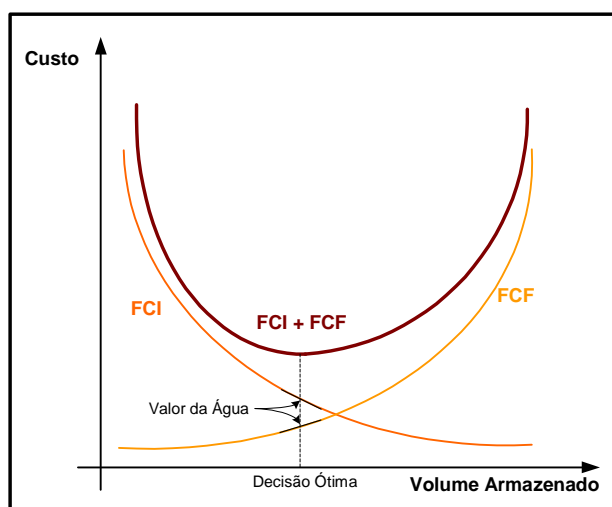


Figura 3: Uso ótimo da água

Vale ressaltar ainda, que a existência de interligações com sistemas vizinhos permite uma redução dos custos de operação, através do intercâmbio de energia e um aumento da confiabilidade de fornecimento, através da repartição das reservas. Por exemplo, quando sistemas interligados possuem regimes hidrológicos complementares, a transferência de energia em um determinado período se dará do sistema em regime de cheias para o sistema que está em período seco.

O valor da energia hidráulica (valor da água) é o valor da geração térmica mais déficit que se poderia substituir hoje ou no futuro. Este valor é resultante do processo de determinação da política ótima de operação. Com este conceito, pode-se representar uma hidrelétrica como sendo uma térmica cujo custo de operação é o valor da água. Entretanto, é importante observar que este valor não se mede de maneira isolada em cada usina, pois depende da operação conjunta do sistema. Em outras palavras, se a política ótima de operação hidrotérmica de cada usina hidrelétrica for calculada de maneira isolada, os intercâmbios de energia, mesmo baseados nos valores da água de cada usina hidrelétrica, não resultam na operação mais econômica possível.

Dessa forma, para obter os ganhos operativos máximos de um sistema hidrotérmico interligado, é necessário calcular a política de operação e operar o sistema de maneira integrada, isto é, otimizando conjuntamente a operação de todos os subsistemas, com o objetivo de minimizar o custo total de operação.

### **2.3. Planejamento da Operação Energética no Sistema Brasileiro**

O sistema de geração de energia elétrica brasileiro apresenta características únicas quando comparado a outros: é um sistema de grande porte, com predominância de geração hidráulica, e constituído por grandes reservatórios com capacidade de regularização plurianual, distribuídos em diversas bacias hidrográficas, com a existência de várias usinas hidrelétricas em um mesmo rio. Adicionalmente, o registro histórico de vazões afluentes às usinas hidroelétricas que compõem o parque gerador apresenta períodos hidrológicamente desfavoráveis de longa duração, tais como o observado no período de 1951 a 1956.

Em um sistema com estas características, é indubitável o ganho obtido através da coordenação e otimização do despacho do parque gerador. No entanto, esta coordenação é bastante complexa do ponto de vista técnico, tendo sido necessário o desenvolvimento, com tecnologia nacional, de modelos matemáticos e programas computacionais para dar suporte às decisões, capaz de considerar os ganhos energéticos com a operação integrada dos diversos subsistemas e aderente aos horizontes de tomada de decisão e níveis de incertezas associadas. Estes desenvolvimentos vêm sendo realizados, preservando todo o conhecimento e experiência já adquiridos, há cerca de 20 anos, pelos técnicos das empresas do setor,

no âmbito dos extintos GCOI<sup>4</sup> e GCPS<sup>5</sup>, capitaneados pela ELETROBRÁS e pelo Centro de Pesquisas de Energia Elétrica, o CEPEL. A partir de 1999, com as mudanças do marco regulatório, o Operador Nacional do Sistema Elétrico, o Ministério de Minas e Energia e mais recentemente a Empresa de Pesquisa Energética passaram a atuar e contribuir decisivamente para o desenvolvimento.

Tipicamente, o problema de planejamento da operação é desagregado em horizontes temporais, de acordo com as principais decisões a serem tomadas e o nível de representação do sistema, conforme apresentado na Figura 4. Por exemplo, esta figura indica que os estudos de longo prazo possuem horizonte de 5 a 10 anos com desagregação mensal, o sistema hidrelétrico é representado de forma agregada, por subsistemas e reservatórios equivalentes, e as variáveis de interesse são: a geração hidrelétrica por subsistema, a geração termelétrica e os intercâmbios energéticos entre os subsistemas.

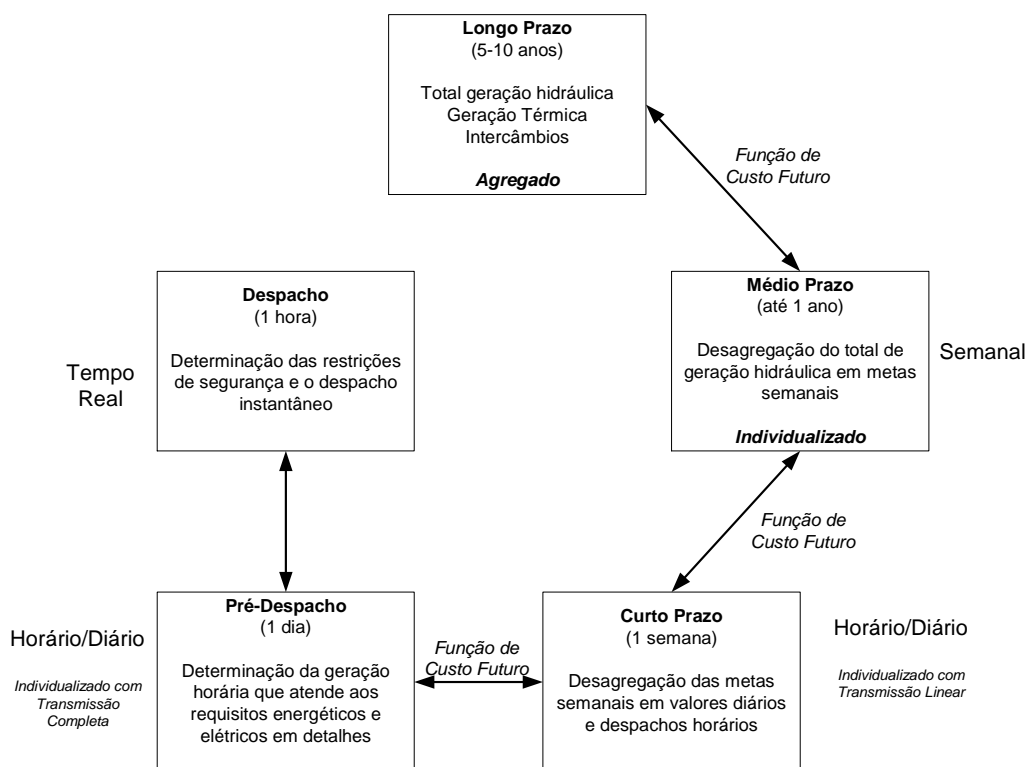


Figura 4: Horizontes do Planejamento da Operação e Principais Decisões em cada Estágio

Observa-se que quanto mais distante é o horizonte de tomada de decisão, maior é a necessidade de consideração das incertezas associadas ao problema e menor é o

<sup>4</sup> Grupo Coordenador da Operação Interligada



grau de detalhamento na representação do sistema (usinas hidrelétricas e rede de transmissão). A Figura 5 ilustra este fato.

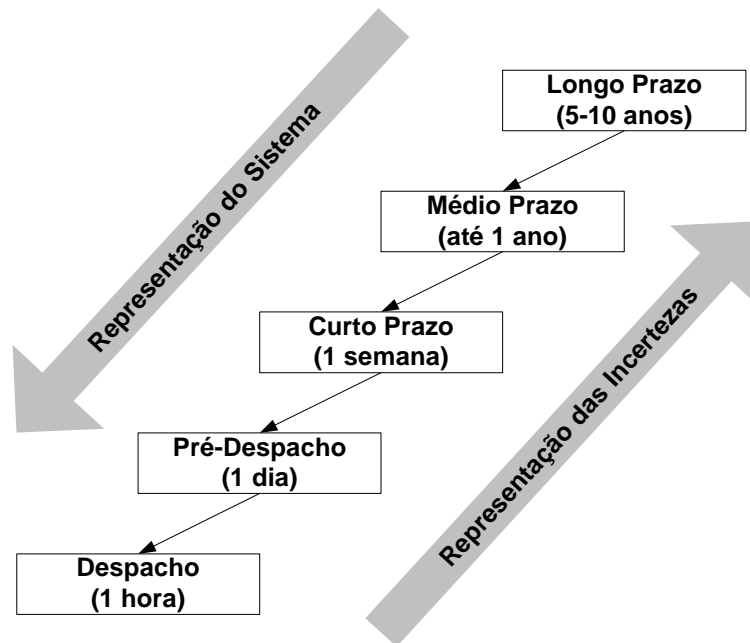


Figura 5: Representação de Incertezas e do Detalhamento do Sistema com o Horizonte de Estudo

A nomenclatura associada ao horizonte de planejamento, apresentada na Figura 4 e na Figura 5, segue a utilizada internacionalmente. A partir de 1999 convencionou-se chamar, em nível nacional, os estudos com horizonte de 5 anos, de estudos de médio prazo; os estudos com horizonte de até 1 ano, de estudos de curto prazo, e os estudos com horizonte de 1 semana, de programação da operação.

### 2.3.1. Cadeia de Modelos

Para fazer face a esta desagregação temporal, foi necessário o desenvolvimento, pelo CEPEL, de uma cadeia de modelos matemáticos e computacionais para o planejamento da expansão da geração e para o planejamento e programação da operação energética (Maceira et al, 2002), ilustrada na Figura 6.

O princípio básico dos modelos dessa cadeia é coordenar a operação das usinas hidroelétricas do sistema, visando diminuir o uso de geração térmica e operar o sistema de maneira mais eficiente, de forma a contribuir para a redução dos custos operativos e dos riscos de déficit.

<sup>5</sup> Grupo Coordenador do Planejamento do Sistema

É feita, a seguir, uma breve descrição dos modelos computacionais utilizados pelo Operador Nacional do Sistema Elétrico para o planejamento da operação energética de médio prazo, curto prazo e para a programação da operação. Estes modelos computacionais são o NEWAVE (Maceira et al, 1999 e 2008), o DECOMP (Costa et al, 1999 e Xavier et al, 2005) e o DESSEM (Belloni et al, 2003), todos desenvolvidos pelo Centro de Pesquisas de Energia Elétrica – CEPEL.

### **2.3.2. Planejamento de Médio Prazo**

Para o planejamento da operação de médio prazo foi desenvolvido o modelo NEWAVE (Maceira et al., 1998 e 2008). Ele define para cada mês do período de planejamento, que pode variar de 5 a 10 anos, a alocação ótima dos recursos hídricos e térmicos de forma a minimizar o valor esperado do custo de operação ao longo de todo o período de planejamento. O parque hidrelétrico é representado de forma agregada e o cálculo da política de operação é baseado na programação dinâmica dual estocástica (Pereira & Pinto, 1991).

O modelo SUIHI-O (Maceira & Cunha, 1995) é um modelo de simulação a usinas individualizadas da operação energética de sistemas hidrotérmicos interligados. É capaz de simular subsistemas hidrotérmicos eletricamente interligados em malha, mas hidraulicamente independentes. Pode ser acoplado ao modelo de decisão estratégica NEWAVE através da função de custo futuro para cada estágio e pode considerar restrições locais de operação decorrente, como por exemplo, o uso múltiplo da água.

Os modelos NEWAVE e SUIHI-O descritos acima simulam um grande número de séries hidrológicas em paralelo, calculando assim índices probabilísticos de desempenho do sistema para cada estágio da simulação.

O modelo GEVAZP (Maceira & Bezerra, 1997) é o responsável pela geração de cenários sintéticos de afluências mensais multivariados para os modelos NEWAVE e SUIHI-O. Ele ajusta modelos autorregressivos periódicos de tal forma a garantir a semelhança estatística entre as seqüências histórica e sintética e a geração de períodos secos e úmidos mais críticos do que aqueles encontrados no registro histórico. Neste modelo, a afluência do período  $t$  é função das afluências passadas ( $t-1, t-2, \dots$ ) e a estrutura de dependência temporal é sazonal.

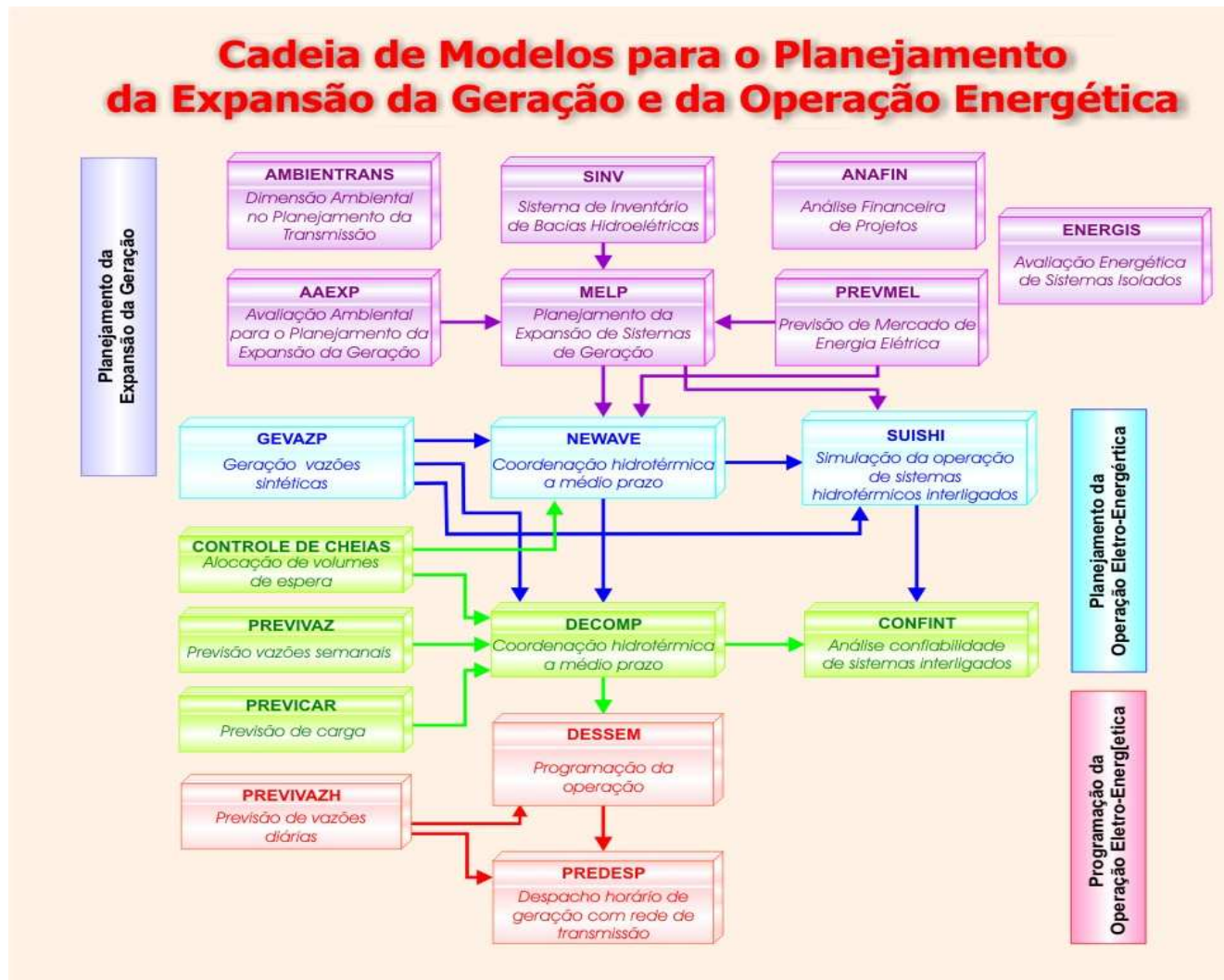


Figura 6: Cadeia de Modelos Computacionais para o Planejamento da Expansão e Operação Energética

### **2.3.3. Planejamento de Curto Prazo**

O modelo DECOMP foi desenvolvido para aplicação no horizonte de curto prazo. Nesta etapa, há um acoplamento com a etapa anterior através da função de custo futuro (que representa a política ótima do médio prazo) gerada num estágio que seja coincidente com o final do horizonte do modelo de curto prazo. Desta forma, é gerada uma função que retrata o valor econômico da água armazenada nos reservatórios em função dos níveis de armazenamento dos mesmos.

O objetivo do planejamento a curto prazo é minimizar o valor esperado do custo de operação ao longo do período de planejamento de um ano, com discretização semanal para o primeiro mês e mensal para os demais. Para isto, é determinada a trajetória ótima de evolução dos armazenamentos dos reservatórios e, para cada patamar de carga, as metas de geração de cada usina de um sistema hidrotérmico, os intercâmbios entre os subsistemas e os custos marginais de operação. Também são consideradas nesta etapa as diversas restrições de geração e algumas restrições elétricas.

A incerteza das afluências aos diversos aproveitamentos do sistema é consideravelmente menor que a presente no planejamento a médio prazo, e é representada através de cenários hidrológicos. Estes cenários são representados através de uma árvore de afluências com probabilidades de ocorrência associadas a cada ramo e são gerados pelo modelo GEVAZP. As previsões de vazões médias semanais são geradas utilizando o modelo PREVIVAZ (Costa et al., 2003).

### **2.3.4. Programação da Operação**

Para o planejamento e programação diária da operação do SIN foi desenvolvido o modelo DESSEM (Lima, 2007). O DESSEM trabalha com horizonte de uma a duas semanas, sendo os dois primeiros dias discretizados de trinta em trinta minutos, os demais dias têm discretização horária. A fim de que o despacho fornecido pelo DESSEM esteja bastante próximo do despacho que efetivamente será colocado em prática, o modelo considera em sua formulação a dinâmica das unidades termelétricas, bem como uma representação detalhada das funções de produção das unidades hidrelétricas. A rede de transmissão é representada por um fluxo de potência DC (corrente contínua). A função de custo futuro gerada pelo modelo de curto prazo

no estágio que coincide com o último estágio do modelo de programação diária é utilizada para se definir a meta de geração de cada unidade geradora.

## **2.4. Resumo**

Neste capítulo foi introduzido o problema de planejamento da operação e suas principais características, tanto para sistemas térmicos quanto sistemas hidrotérmicos. Foi dado um enfoque especial ao sistema elétrico brasileiro, mostrando suas principais características e como o problema de planejamento da operação energética é abordado. Além disso, foram apresentadas as etapas nas quais o problema de planejamento é desagregado e a cadeia de modelos desenvolvidos para cada uma das etapas.

No próximo capítulo será abordado mais especificamente o problema de planejamento da operação de médio prazo, a estratégia de solução adotada para solução do problema e como a incerteza hidrológica é representada no problema. Também serão apresentados métodos para construção da árvore de cenários e o método atualmente utilizado pelo modelo NEWAVE.

## **3** **Planejamento da Operação de Médio Prazo**

### **3.1.** **Considerações Iniciais**

Diversos métodos têm sido propostos para a solução do problema de planejamento energético de médio prazo de sistemas hidrotérmicos interligados, tais como os métodos baseados em algoritmos de fluxos em rede (Rosenthal, 1981 e Oliveira, 1992), os métodos que combinam programação linear com programação dinâmica (Yeh et al, 1992), e aqueles que utilizam decomposição de Benders multi-estágio (Pereira & Pinto, 1991; Rotting & Gjesviska, 1992; Maceira et al, 2002).

No modelo NEWAVE, o problema de planejamento da operação energética de médio prazo é representado por um problema de programação estocástica linear multi-estágio. A técnica utilizada para encontrar a solução ótima é a PDDE, proposta por Pereira & Pinto (1991), baseada na decomposição de Benders (Benders, 1962). A formulação da PDDE atualmente empregada no modelo leva em conta a correlação temporal das afluições aos reservatórios (Maceira, 1993). Assim, consideram-se como variáveis de estado do problema o armazenamento no início do período e as afluições passadas (tendência hidrológica).

Os primeiros métodos de decomposição desenvolvidos para resolver problemas de programação linear estocástico percorriam a árvore de cenários em sua totalidade, como o método L-shaped (Van Slyke & Wets, 1969) para 2 estágios e o método de decomposição aninhada, proposto por Birge (1985) para casos multi-estágio. Para problemas onde a árvore de cenários apresenta uma cardinalidade elevada, percorrer a árvore em sua totalidade é impossível do ponto de vista computacional ou prático. Para contornar este problema foram desenvolvidos diversos métodos que utilizam técnicas de amostragens para selecionar uma sub-árvore de cenários com tamanho reduzido. O primeiro método a fazer uso da amostragem em programação estocástica foi o PDDE. Outras abordagens para resolver problemas de programação estocástica multi-estágio foram sugeridas, dentre as quais se destacam os métodos de decomposição aninhada reduzida – AND - (Donohue & Birge, 2006), de amostragem reduzida – ReSa (Hindsgerber & Philpott, 2001) e o algoritmo de planos cortantes e

amostragem parcial – CUPPS – (Chen & Powell, 1999). Segundo Hindsgerber & Philpott (2001), o método AND tem bom desempenho para árvores com poucos estágios e muitas aberturas, enquanto a PDDE é melhor para árvores com muitos estágios e poucas aberturas, já o método ReSa incorpora idéias do método AND para árvores com muitos estágios e poucas aberturas.

Em (Philpott & Guan, 2007) é apresentada a prova de convergência para estes métodos em um número infinito de iterações. Neste artigo, os autores não analisam a robustez dos algoritmos com relação à árvore completa de cenários, restringindo-se a analisar a convergência. Na formulação apresentada no artigo, pressupõe-se que o processo estocástico é um processo discreto com um número finito de possíveis realizações, e que a sub-árvore de cenários é re-sorteada a cada iteração do algoritmo. A concepção do problema adotada pelo NEWAVE, utiliza um processo estocástico contínuo, onde os ruídos são modelados por uma distribuição multivariada lognormal. Além disso, no algoritmo de solução implementado, uma única amostra é utilizada em todas as iterações para definição da sub-árvore de cenários. Portanto, pode-se notar que o algoritmo do NEWAVE não se enquadra nos pressupostos analisados em Philpott & Guan (2007).

Neste capítulo é discutida a representação da incerteza hidrológica no algoritmo da PDDE, os procedimentos atuais para a construção da sub-árvore a ser utilizada pelo algoritmo e os problemas que podem ocorrer quando a escolha da sub-árvore não for feita de forma adequada. No apêndice A é apresentada de forma sucinta o algoritmo da PDDE. Para um exemplo detalhada da aplicação da PDDE no planejamento da operação de médio prazo consultar Marcato (2002).

### **3.2. Representação da Incerteza Hidrológica na PDDE**

A representação do NEWAVE é essencialmente estocástica, considerando-se que as aflúncias hidrológicas que ocorrerão em cada período do horizonte de planejamento não são previamente conhecidas. A incerteza das aflúncias é representada por uma árvore de cenários, conforme exemplo apresentado na Figura 7, onde cada caminho na árvore é chamado de cenário de aflúncia ou cenário hidrológico, e cada nó representa uma possível realização da variável aleatória (aflúncia) correspondente. Essas realizações seguem um processo estocástico com propriedades estatísticas inferidas do histórico de aflúncias e que devem ser

preservadas na construção da árvore, ou de suas aproximações, tais como média, desvio padrão e correlação temporal e espacial.

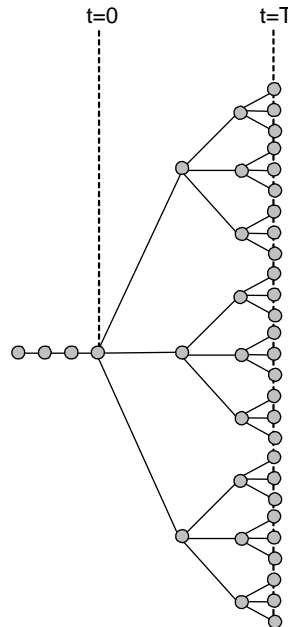


Figura 7: Exemplo de árvore de cenários (Árvore Completa)

Da formulação do processo estocástico considerado e de seus parâmetros, por meio de sucessivas operações, incluindo desconto do valor médio, divisão pelo desvio padrão, aplicação de filtros autorregressivos e transformação logarítmica, é possível transformar uma árvore de cenários de afluência em uma árvore de cenários de variáveis aleatórias independentes entre si com média zero e variância constante (ruído). Esse processo é chamado branqueamento.

A solução do problema de planejamento associado à árvore de cenários completa, em princípio, exigiria a solução de um problema para cada nó ( $NLEQ^{T-1}$  problemas, onde NLEQ é o número de aberturas). Supondo a árvore completa da Figura 7, onde cada nó está ligado a três outros nós e que o total de períodos ao longo do horizonte de planejamento seja igual ao valor utilizado atualmente no programa mensal de operação (PMO), igual a 120, a cardinalidade da árvore de cenários completa é igual a  $3^{120}$  e o número de problemas a ser resolvido desta árvore é da ordem de  $10^{57}$ . Ou seja, mesmo uma representação simples da incerteza hidrológica (3 ligações a cada nó) torna o percurso de toda a árvore de cenários inviável do ponto de vista computacional. Portanto, para solução deste problema deve-se procurar aplicar técnicas que contornem o problema da dimensionalidade da árvore de cenários, evitando-se percorrê-la em sua totalidade.



Para a solução do problema, atualmente é utilizada uma sub-árvore selecionada por sorteio aleatório a partir de uma árvore de cenários considerada completa, que por sua vez é obtida por sorteio aleatório da população contínua teórica de ruídos correspondente ao processo estocástico utilizado (PAR(p)). A Figura 8a ilustra a seleção de uma sub-árvore a partir da árvore completa da Figura 7. A partir dos cenários da sub-árvore selecionada obtém-se por simulação uma estimativa (e um intervalo de confiança) para o limite superior do valor esperado do custo total de operação (simulação forward – árvore em negro). Todas as ramificações da sub-árvore selecionada são consideradas na construção de uma aproximação da função de custo futuro que fornece no estágio inicial o limite inferior do valor esperado do custo total de operação (simulação backward – árvore tracejada em negro), Figura 8b. Maiores detalhes sobre a estimativa dos limites superior e inferior do valor esperado do custo total de operação podem ser obtidos em Pereira & Pinto (1991).

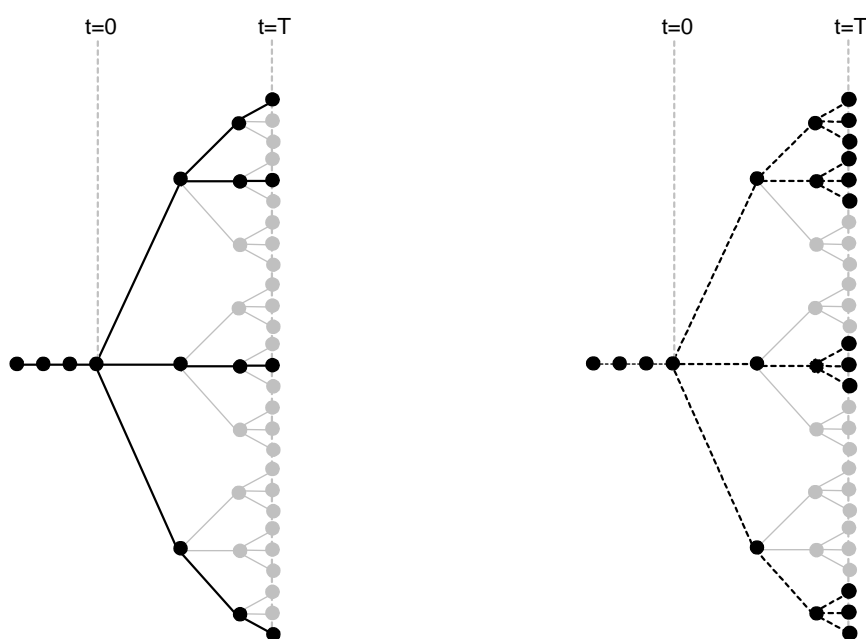


Figura 8: Exemplo de sub-árvore de cenários do NEWAVE  
 (a) Simulação Forward; (b) Simulação Backward

Além do problema de representação do processo estocástico pela árvore completa e pela sub-árvore, um outro problema potencial do processo randômico de escolha da sub-árvore é o impacto no processo de convergência do problema de otimização. O algoritmo de solução é dito convergido quando o limite inferior do valor esperado do custo total de operação, intitulado ZINF no NEWAVE, estiver no intervalo

de confiança do limite superior do valor esperado do custo total de operação, chamado de ZSUP no NEWAVE. Com efeito, pode-se mostrar que o critério de convergência não é alcançado, isto é, ZINF não se aproxima de ZSUP, se algumas das possíveis sub-árvores forem selecionadas.

Seja o exemplo apresentado pelo CEPEL<sup>6</sup>: Suponha que o processo estocástico das afluições seja representado adequadamente pela árvore binária apresentada na Figura 9. Duas possíveis sub-árvores com quatro cenários selecionadas a partir da árvore completa estão ilustradas na Figura 10. Se fosse utilizada a árvore completa tanto na simulação backward ( $T-1$  até  $1$ ), quanto na simulação forward ( $1$  até  $T$ ), os valores de ZINF e ZSUP iriam convergir para o valor ótimo, aqui chamado de  $Z^*$ .

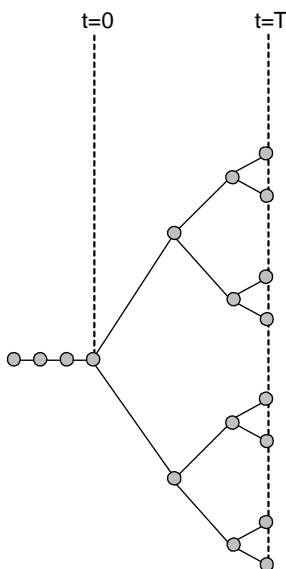


Figura 9: Exemplo CEPEL - árvore completa

Considerando a primeira sub-árvore (Figura 10a), pode-se dizer que a simulação backward percorre todo o espaço de estados, logo ZINF tende ao valor  $Z^*$ . Já a simulação forward não percorre todo o espaço de estados, desta forma ZSUP pode estabilizar acima ou abaixo de  $Z^*$ , dependendo se a sub-árvore escolhida for pessimista ou otimista. No caso pessimista, dependendo de quão acima de  $Z^*$  o valor de ZSUP estabilizar, o processo de convergência pode não ser satisfeito, considerando o intervalo de confiança adotado.

<sup>6</sup> Apresentação realizada pelo CEPEL na 23ª reunião da Força Tarefa NEWAVE em 13 de agosto de 2007 no ONS

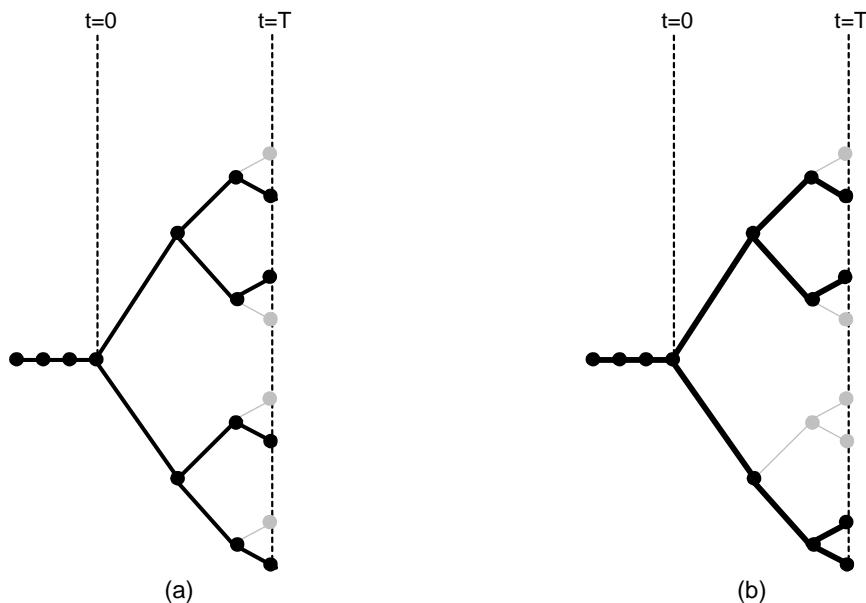


Figura 10: Exemplo CEPEL – sub-árvores  
 (a) ZINF  $\rightarrow$   $Z^*$ ; (b) ZINF  $\rightarrow$   $Z^{**} < Z^*$

Suponha agora que a segunda sub-árvore (Figura 10b) seja selecionada. Neste caso, a simulação backward não percorre todo o espaço de estados, logo ZINF tende a um valor  $Z^{**}$  que é menor do que o valor ótimo  $Z^*$ . A simulação forward também não percorre todo o espaço de estados, e ZSUP pode estabilizar ou abaixo de  $Z^*$ , dependendo se a sub-árvore escolhida for pessimista ou otimista. Outra vez, não é garantida a convergência do processo de otimização, tanto no caso pessimista quanto no caso otimista.

### 3.3. Construção da Árvore de Cenários

Como já mencionado anteriormente, o método da PDDE utiliza uma árvore de cenários para representar o processo estocástico. Essa árvore de cenários é escolhida de forma aleatória e representa uma pequena porção da árvore completa de cenários.

Existem diversos métodos que propõem construir uma árvore de cenários de tamanho controlado. Latorre et al (2007) apresentam métodos para construção da árvore de cenários a ser utilizada em problemas de otimização estocástica. Após a construção da árvore são aplicadas técnicas para agrupar cenários com o intuito de reduzir cardinalidade desta. Em todos os métodos sugeridos há problemas para preservação da correlação temporal do processo.

Em Heitsh & Romish (2001) é sugerida a utilização de métricas probabilísticas, tal como a métrica de Fournier-Mourier, para a determinação da redução ótima de cenários a ser aplicada em algoritmos de programação estocástica. Outros trabalhos que também seguem esta linha são Dupacová et al. (2000) e Growe-Kuska et al. (2003). O processo modelado nestes trabalhos não apresenta correlação temporal, portanto o método proposto não garante que essas correlações sejam preservadas na árvore reduzida.

Técnicas de agregação de cenários são utilizadas em Jardim (2002) com o intuito de gerar uma árvore de cenários hidrológicos multivariados com boa representatividade do processo estocástico das afluições e baixa cardinalidade. A estrutura da árvore de cenários é definida a priori, e é levada em consideração a preservação da estrutura de correlação temporal e espacial do processo estocástico de afluições. A árvore de cenários obtida com o método proposto é aplicada ao problema de planejamento de médio prazo.

Em Oliveira (2007) são empregadas técnicas de redução ótima de cenários na construção da árvore de cenários hidrológicos multivariados utilizada no problema de planejamento de médio prazo. No método proposto há a necessidade de armazenar um grande número de informações para calcular a distância entre os cenários. Para contornar esse problema, em Oliveira et al. (2008) é proposto o uso de um algoritmo de redução local. Além disso, também é proposto um ajuste na medida de distância com o intuito de reduzir a degradação da variância dos cenários preservados. Nos métodos mencionados acima, a estrutura de correlação temporal e espacial são preservadas.

Costa (2007) apresenta uma proposta para tratamento das incertezas em modelos de programação estocástica baseada em árvore de cenários de tamanho limitado, porém representativa. Para reduzir a dimensionalidade do espaço de estados é utilizado o método dos componentes principais antes de empregar técnicas para definir uma amostra representativa. São avaliados diversos métodos para a geração da árvore de cenários, entre eles o método do reticulado. O trabalho concentra-se na geração de árvores para o planejamento de curto-prazo, que tem um horizonte bem menor do que o planejamento de médio prazo.

Na versão atualmente em uso do modelo NEWAVE, considera-se a utilização de um único conjunto de ruídos para a geração dos cenários da simulação backward e forward (opção Atual). Nesta alternativa, é sorteada de forma aleatória a amostra de vetores de ruídos que são utilizados para a construção dos cenários da backward. A

partir desta amostra são sorteados, por meio de um sorteio condicionado, os vetores de ruídos que são utilizados para a geração dos cenários da simulação forward, conforme Figura 11.

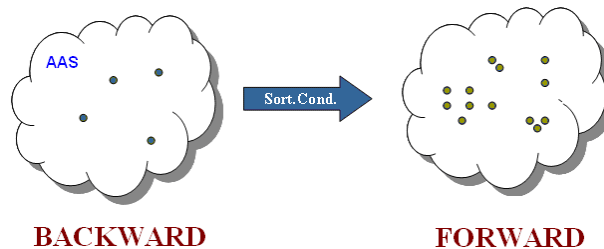


Figura 11: Opção Atual

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 àquele valor sorteado. No exemplo da Figura 12, o número aleatório uniforme sorteado foi 0,79. O ruído cuja função acumulada corresponde a 0,79 é o ruído número 191.

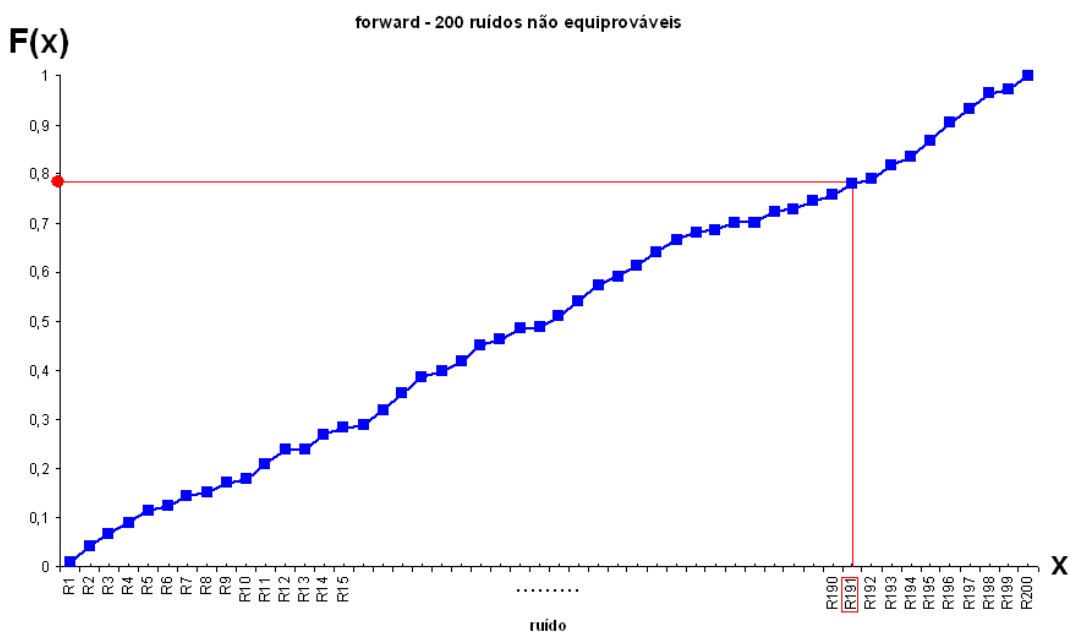


Figura 12: Sorteio Condicionado

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.

Existe disponível na versão atualmente em uso do modelo NEWAVE uma segunda alternativa para geração da árvore de cenários. Nesta alternativa são utilizados diferentes ruídos na backward e forward, todos sorteados da mesma população contínua. A árvore de cenários do passo forward e backward são obtidas através de amostragem aleatória simples (opção 1AAS), Figura 13.

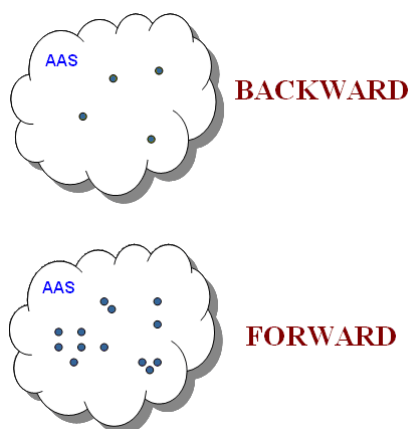


Figura 13: Opção 1AAS

Vale ressaltar que em ambas opções os cenários backward e forward são equiprováveis.

A opção Atual é a opção utilizada na elaboração dos PMOs pelo ONS atualmente. Apesar de esta opção representar os ruídos da amostra forward de uma forma muito simplificada, sorteados a partir de uma população reduzida (normalmente a amostra backward possui apenas 20 representantes), ela foi escolhida por se aproximar melhor dos pressupostos indicados em Philpott & Guan (2007) para garantia assintótica da convergência do algoritmo de solução. Com efeito, pode-se interpretar a amostra backward como o conjunto de todas as possíveis realizações, e a amostragem da forward passa a ser realizada a partir deste conjunto.

Por outro lado, na concepção de um conjunto contínuo de possíveis realizações, para efeito de convergência é de se supor que os cenários visitados nos passos forward e backward não precisem ser necessariamente os mesmos, mas sim que devam visitar as mesmas regiões da distribuição de probabilidades do processo

estocástico. Vale ressaltar que, sob o ponto de vista da robustez dos resultados, a opção Atual tende a ser inferior à segunda alternativa, uma vez que a amostragem da forward está restrita à amostra backward.

Na época em que a escolha pela opção Atual foi realizada, para os casos estudados foi verificado que aparentemente as características do processo estocástico de afluência estavam sendo igualmente preservadas nas opções Atual e 1AAS.

### **3.4. Resumo**

Neste capítulo foi apresentada a PDDE, método utilizado pelo modelo NEWAVE. A importância e os problemas de uma representação inadequada da incerteza hidrológica neste método também foram discutidos.

Na PDDE a incerteza hidrológica é representada através de uma árvore de cenários. Desta forma, foi apresentado como a árvore de cenários é atualmente construída e foram discutidos alguns outros métodos para construção de árvores de cenários encontrados na literatura.

No próximo capítulo será apresentado o modelo estatístico utilizado para gerar a árvore de cenários hidrológicos utilizada na solução do problema de planejamento da operação energética.

## 4 Modelo de Geração de Séries Sintéticas de Energia

### 4.1. Considerações Iniciais

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

Por exemplo, um grande número de cenários pode ser utilizado para estimar a relação entre demanda energética e risco de não atendimento em um ano qualquer (energia garantida). Na Figura 14 é mostrado que um sistema hidroelétrico pode suprir a demanda e que está associada ao risco  $p$ , previamente selecionado.

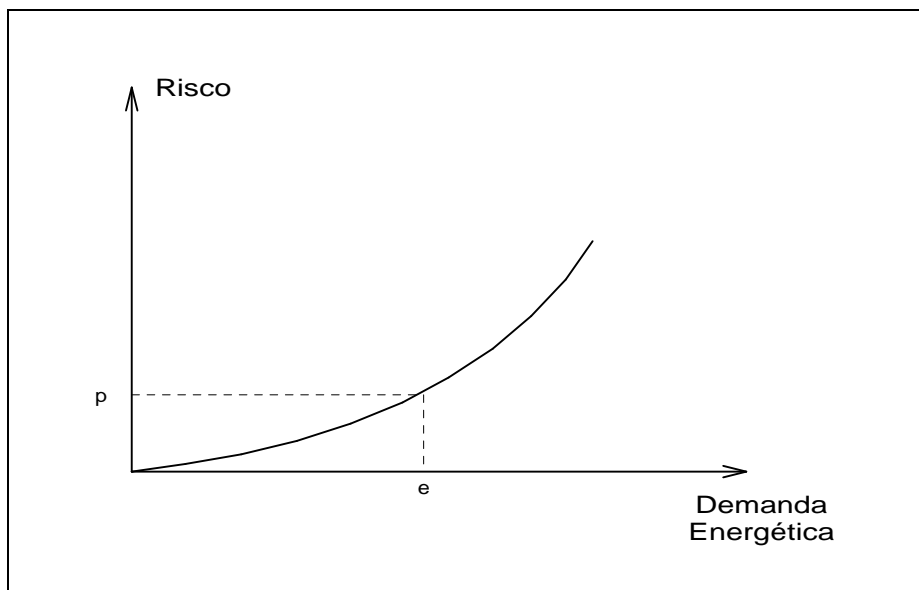


Figura 14: Relação Demanda x Risco em um Ano Qualquer



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

Nesta seção será descrito o modelo autorregressivo periódico, PAR(p), proposto para ser utilizado no Modelo Estratégico de Geração Hidrotérmica a Subsistemas Equivalentes Interligados – NEWAVE, no Modelo de Simulação a Usinas Individualizadas para Subsistemas Hidrotérmicos Interligados – SUIISHI-O e no Modelo de Determinação da Coordenação da Operação a Médio Prazo – DECOMP.

#### 4.2. Processos Estocásticos e Séries Temporais

Medindo-se, por exemplo, a vazão afluyente mensal a um dado local por um período de cinco anos, pode-se obter a curva 1 da Figura 15. Realizando-se as mesmas medições para outro segmento de cinco anos, obtém-se outra curva, que é em geral diferente da primeira.

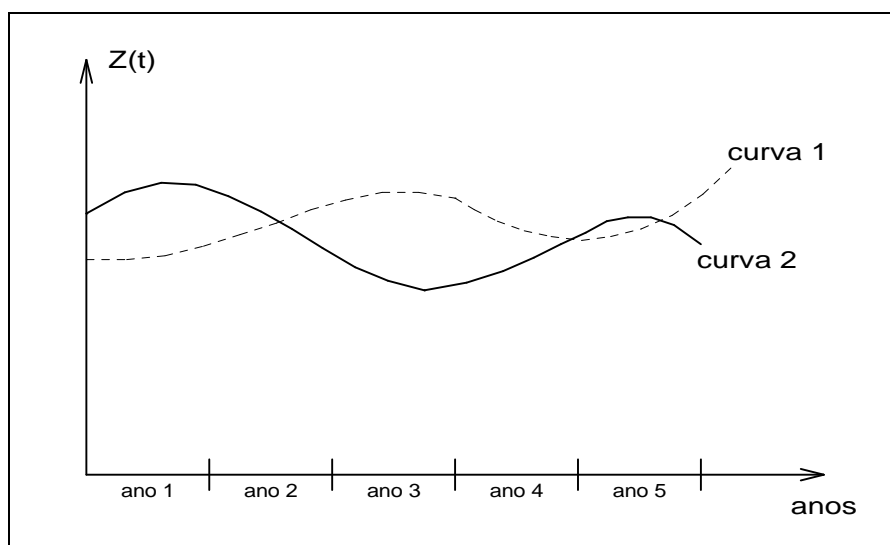


Figura 15: Vazão Afluyente Anual x Anos

Estas curvas são chamadas de trajetórias ou realizações do processo físico que está sendo observado. Este pode ser modelado por um processo estocástico, que nada mais é do que o conjunto de todas as possíveis trajetórias que podem ser observadas com as respectivas probabilidades. Cada trajetória é também chamada de série temporal.

Na prática, só está disponível uma realização do processo estocástico, a série histórica. Assim, para cada instante do tempo (dia, mês, ano, etc), o processo estocástico é uma variável aleatória. O valor observado em um instante  $t$  qualquer (valor da série histórica no instante  $t$ ) nada mais é do que o valor 'amostrado' da distribuição de probabilidade associado à variável aleatória do processo estocástico no instante  $t$ .

Um processo estocástico é totalmente descrito pelo conjunto de todas as séries temporais que o compõem, com suas respectivas probabilidades, ou pela distribuição de probabilidade conjunta de todas as variáveis aleatórias envolvidas. Como na realidade não estão disponíveis o conjunto de todas as séries temporais e a distribuição de probabilidade conjunta de todas as variáveis aleatórias, a modelagem de séries sintéticas tem por objetivo ajustar um modelo pelo qual acreditamos que a série histórica tenha sido produzida e a partir dele gerar séries sintéticas que representam as séries temporais que podem ser 'amostradas' pelo processo físico que se está observando, o processo estocástico.

É comum considerar algumas hipóteses simplificadoras para o problema, por exemplo, estacionaridade. Um processo estocástico é estacionário se ao longo do tempo as suas propriedades estocásticas não se alteram. Isto significa que a média, o desvio padrão etc, não sofrem modificações ao longo do tempo, ou de uma forma mais abrangente, significa que a distribuição de probabilidade em um instante  $t$  qualquer é válida para qualquer outro instante. A não estacionaridade de um processo estocástico pode ser causada pela intervenção direta do homem, ou da natureza, no processo físico, ou ainda pela presença de ciclos, sazonalidades (características que se repetem dentro de um ano). Quanto ao último fator, o processo de afluições anuais é considerado estacionário e o processo de afluições mensais é considerado não estacionário.

Dada uma série temporal anual observada  $[z_1, z_2, \dots, z_N]$  podemos estimar os seguintes índices estatísticos:

- média amostral:

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N z_i \quad (4.1)$$

- desvio padrão amostral, que mede o grau de dispersão da amostra em torno da média:

$$\hat{\sigma} = \sqrt{\frac{1}{N} \sum_{i=1}^N (z_i - \hat{\mu})^2} \quad (4.2)$$

- a auto-covariância de ordem um, que mede o grau de dependência linear entre duas variáveis aleatórias espaçadas de um intervalo de tempo. Neste caso, a auto-covariância de ordem um indica a dependência linear entre a vazão de um ano com a vazão do ano imediatamente anterior:

$$\hat{\gamma}(1) = \frac{1}{N} \sum_{i=2}^N (z_i - \hat{\mu})(z_{i-1} - \hat{\mu}) \quad (4.3)$$

A auto-covariância possui a dimensão da variável Z ao quadrado. Para se obter um índice de análise mais direto divide-se a covariância de ordem um pelo desvio-padrão das variáveis envolvidas, resultando na correlação de ordem um, também conhecida por dependência serial anual.

$$\hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\sigma}^2} \quad (4.4)$$

Se este valor é igual a 1, a vazão de um ano pode ser perfeitamente descrita pela vazão do ano anterior. Caso seja igual a -1 as variáveis são perfeitamente correlacionadas, mas o crescimento de uma implica no decréscimo da outra. Se este valor é zero, as variáveis não têm dependência linear.

Pode-se medir também a dependência espacial entre as vazões afluentes a duas usinas hidroelétricas, isto é, se a série histórica de uma usina é linearmente dependente da série histórica de outra, por meio da correlação cruzada ou correlação espacial entre elas.

Por exemplo, a correlação espacial anual entre as vazões afluentes às usinas hidroelétricas a e b pode ser obtida por:

$$\hat{\rho}_{a,b} = \frac{\frac{1}{N} \sum_{i=1}^N (z_i^{(a)} - \hat{\mu}^{(a)})(z_i^{(b)} - \hat{\mu}^{(b)})}{\hat{\sigma}^{(a)}\hat{\sigma}^{(b)}} \quad (4.5)$$

Em geral, processos naturais em escala mensal apresentam um comportamento periódico descrito pelos pela componente sazonal. Cada período apresenta um

conjunto de características estatísticas próprias descritas pela média, desvio-padrão e estrutura de correlações sazonais.

A média amostral de cada mês é dada por:

$$\hat{\mu}_m = \frac{1}{N} \sum_{i=1}^N z_{12*(i-1)+m} \quad m = 1, \dots, 12 \quad (4.6)$$

De forma análoga, o desvio padrão amostral de cada mês é dado por:

$$\hat{\sigma}_m = \sqrt{\frac{1}{N} \sum_{i=1}^N (z_{12*(i-1)+m} - \hat{\mu}_m)^2} \quad m = 1, \dots, 12 \quad (4.7)$$

É comum em processos mensais se definir um conjunto de valores que descreva a estrutura de correlação linear de um dado mês com os meses anteriores. Esses valores podem ser definidos pela correlação de ordem 1, que descreve a dependência linear da variável aleatória de um mês qualquer  $m$  com a variável aleatória do mês imediatamente anterior ( $m-1$ ); correlação de ordem 2, que descreve a dependência linear da variável aleatória do mês  $m$  com a variável aleatória do mês ( $m-2$ ); ...; correlação de ordem  $k$ , que descreve a dependência linear da variável aleatória do mês  $m$  com a variável aleatória do mês ( $m-k$ ), onde  $k$  é um valor qualquer. A esse conjunto de valores dá-se o nome de função de autocorrelação do mês  $m$ , também conhecida por dependência serial mensal.

Os valores amostrais desses índices estatísticos podem ser obtidos da seguinte forma:

$$\hat{\gamma}_m(k) = \frac{1}{N} \sum_{i=1}^N (z_{12*(i-1)+m} - \hat{\mu}_m)(z_{12*(i-1)+m-k} - \hat{\mu}_m) \quad m = 1, \dots, 12 \quad (4.8)$$

$$\hat{\rho}_m(k) = \frac{\hat{\gamma}_m(k)}{\hat{\sigma}_m \hat{\sigma}_{m-k}} \quad m = 1, \dots, 12 \quad (4.9)$$

### 4.3. O Modelo Autorregressivo Periódico

#### 4.3.1. Descrição do Modelo

Séries hidrológicas de intervalo de tempo menor que o ano, tais como séries mensais, têm como característica o comportamento periódico das suas propriedades probabilísticas, como, por exemplo, a média, a variância, a assimetria e a estrutura de autocorrelação. A análise deste tipo de série pode ser feita pelo uso de formulações

autorregressivas cujos parâmetros apresentam um comportamento periódico. Esta classe de modelos costuma ser denominada modelos autorregressivos periódicos (Maceira, 1989). Estes são referenciados por modelos PAR(p), onde p é a ordem do modelo, ou seja, o número de termos autorregressivos do modelo. Em geral, p é um vetor,  $p = (p_1, p_2, \dots, p_{12})$ , onde cada elemento fornece a ordem de cada período.

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

$$\left( \frac{Z_t - \mu_m}{\sigma_m} \right) = \phi_1^m \left( \frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \left( \frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \quad (4.10a)$$

ou

$$\phi^m(B) \left( \frac{Z_t - \mu_m}{\sigma_m} \right) = a_t \quad (4.10b)$$

onde:

$Z_t$  é uma série sazonal de período s

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

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

N é o número de anos

$\mu_m$  é a média sazonal de período s

$\sigma_m$  é desvio-padrão sazonal de período s

$\phi^m(B)$  é o operador autorregressivo de ordem  $p_m$

$$\phi^m(B) = (1 - \phi_1^m B - \phi_2^m B^2 - \dots - \phi_{p_m}^m B^{p_m})$$

$B^i$  aplicado a  $Z_t$  resulta em  $Z_{t-i}$  ( $B^i Z_t = Z_{t-i}$ )

$p_m$  é a ordem do operador autorregressivo do período m

$a_t$  é a série de ruídos independentes com média zero e variância  $\sigma_a^{2(m)}$

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

$$\rho^m(k) = E \left[ \left( \frac{Z_t - \mu_m}{\sigma_m} \right) \left( \frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}} \right) \right] \quad (4.11)$$

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

multiplicando-se ambos os lados da equação (4.10a) por  $\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}}\right)$  e tomando o valor esperado obtemos para cada período:

$$E\left[\left(\frac{Z_t - \mu_m}{\sigma_m}\right)\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}}\right)\right] = \phi_1^m E\left[\left(\frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}}\right)\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}}\right)\right] + \dots + \phi_{p_m}^m E\left[\left(\frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}}\right)\left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}}\right)\right] + E\left[a_t \left(\frac{Z_{t-k} - \mu_{m-k}}{\sigma_{m-k}}\right)\right] \quad (4.12)$$

Por exemplo, para  $k = 1$  a expressão (4.12) resulta em:

$$\rho^m(1) = \phi_1^m + \phi_2^m \rho^{m-1}(1) + \dots + \phi_{p_m}^m \rho^{m-1}(p_m - 1) \quad (4.13)$$

Conhecidos os parâmetros de um modelo PAR(p), as funções  $\rho^m(k)$  são dadas pela solução de (4.12) e podem ser expressas por uma combinação de decaimentos exponenciais ou ondas senoidais, o que faz com que cada  $\rho^m(k)$  tenda a zero à medida que  $k$  cresce.

Fixando-se  $m$  e variando  $k$  de 1 a  $p_m$  em (4.12) obtém-se, para cada período, um conjunto de equações comumente denominado de equações de Yule-Walker. Para um período  $m$  qualquer:

$$\begin{bmatrix} 1 & \rho^{m-1}(1) & \rho^{m-1}(2) & \dots & \rho^{m-1}(p_m-1) \\ \rho^{m-1}(1) & 1 & \rho^{m-2}(1) & \dots & \rho^{m-2}(p_m-2) \\ \rho^{m-1}(2) & \rho^{m-2}(1) & 1 & \dots & \rho^{m-3}(p_m-3) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho^{m-1}(p_m-1) & \rho^{m-2}(p_m-2) & \rho^{m-3}(p_m-3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_1^m \\ \phi_2^m \\ \phi_3^m \\ \vdots \\ \phi_{p_m}^m \end{bmatrix} = \begin{bmatrix} \rho^m(1) \\ \rho^m(2) \\ \rho^m(3) \\ \vdots \\ \rho^m(p_m) \end{bmatrix} \quad (4.14)$$

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

$$\begin{bmatrix} 1 & \rho^{m-1}(1) & \rho^{m-1}(2) & \dots & \rho^{m-1}(k-1) \\ \rho^{m-1}(1) & 1 & \rho^{m-2}(1) & \dots & \rho^{m-2}(k-2) \\ \rho^{m-1}(2) & \rho^{m-2}(1) & 1 & \dots & \rho^{m-3}(k-3) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho^{m-1}(k-1) & \rho^{m-2}(k-2) & \rho^{m-3}(k-3) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1}^m \\ \phi_{k2}^m \\ \phi_{k3}^m \\ \vdots \\ \phi_{kk}^m \end{bmatrix} = \begin{bmatrix} \rho^m(1) \\ \rho^m(2) \\ \rho^m(3) \\ \vdots \\ \rho^m(k) \end{bmatrix} \quad (4.15)$$

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

Para  $k = 0$ , a expressão (4.12) fica:

$$1 = \phi_1^m \rho^m(1) + \phi_2^m \rho^m(2) + \dots + \phi_{p_m}^m \rho^m(p_m) + E \left[ a_t \left( \frac{Z_t - \mu_m}{\sigma_m} \right) \right] \quad (4.16)$$

Multiplicando ambos os lados da equação (4.10a) por  $a_t$  e tomando o valor esperado:

$$E \left[ a_t \left( \frac{Z_t - \mu_m}{\sigma_m} \right) \right] = \sigma_a^{2(m)} \quad (4.17)$$

Substituindo (4.17) em (4.16):

$$\sigma_a^{2(m)} = 1 - \phi_1^m \rho^m(1) - \phi_2^m \rho^m(2) - \dots - \phi_{p_m}^m \rho^m(p_m) \quad (4.18)$$

Esta expressão é válida para qualquer período  $m$ .

### 4.3.2. Ajuste do Modelo

Box e Jenkins (Box & Jenkins, 1976) sugeriram uma metodologia bastante elaborada para ajuste de modelos estocásticos da família ARIMA a séries temporais, que pode ser estendida para modelos da família PAR( $p$ ). Nesta metodologia a estratégia de seleção do modelo é dividida em três etapas. A primeira etapa, denominada por Box e Jenkins de **identificação do modelo**, consiste em escolher, tentativamente, a ordem do modelo, com base em estimativas das funções  $\rho^m(k)$  e  $\phi_{kk}^m$  obtidas a partir da série amostral. Na modelagem autorregressiva periódica isto consiste em escolher o vetor  $p$ . A segunda etapa refere-se à **estimação do modelo**, ou seja, estimação dos seus parâmetros, sendo em geral recomendado o uso de estimadores de máxima verossimilhança ou suas aproximações. A terceira etapa diz respeito à **verificação do modelo**, isto é, verificar através de testes estatísticos se as hipóteses consideradas durante as etapas anteriores são atendidas. Se as hipóteses não são verificadas, deve-se retornar à primeira etapa até que os resultados sejam satisfatórios.

Ressalta-se que esta estratégia pode muitas vezes resultar em mais de um modelo capaz de descrever o processo estocástico em estudo.

#### 4.3.2.1. Identificação

A identificação do modelo consiste em determinar as ordens mais apropriadas dos operadores autorregressivos de cada período  $p_m$ ,  $m = 1, \dots, s$ . Isto pode ser feito obtendo-se estimativas  $\hat{\phi}_{kk}^m$ ,  $k = 1, N/4$  e substituindo em (4.15) as autocorrelações pelos respectivos valores amostrais (4.8) e (4.9). Segundo Maceira (1989), se a ordem do operador autorregressivo de um período qualquer  $m$  é  $p_m$ , então  $\hat{\phi}_{kk}^m$  para  $k > p_m$  tem distribuição aproximadamente Normal com média zero e variância  $N^{-1}$ . Para cada período  $m$  procura-se a maior ordem  $i$  tal que todas as estimativas  $\hat{\phi}_{kk}^m$  para  $k > i$  não sejam mais significativas

#### 4.3.2.2. Estimação

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

As médias e os desvios-padrão sazonais são estimados pelas equações (4.6) e (4.7) respectivamente.

Os parâmetros  $\phi_i^m$ ,  $i = 1, \dots, p_m$ , são estimados substituindo-se em (4.14)  $\rho^{m-j}(k)$ ,  $j = 0, \dots, p_m$ ,  $k = 1, \dots, p_m$ , por suas estimativas dadas pelas expressões (4.8) e (4.9).

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

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

Finalmente, as estimativas de  $\sigma_a^{2(m)}$  podem ser obtidas usando-se a expressão (4.18).



### 4.3.2.3. Verificação

Na etapa de verificação são utilizados procedimentos estatísticos que permitem verificar se o modelo estimado é adequado e portanto, capaz de gerar as séries sintéticas. A grande maioria dos testes se baseia numa análise detalhada dos resíduos ajustados  $a_t$ . Se o modelo estimado for adequado, a média dos resíduos deve ser igual zero, a variância deve ser pequena e constante e a função de correlação dos resíduos,  $\rho_a^m(k)$ , deve ser não significativa para  $k > 0$ .

### 4.3.3. Geração de Séries Sintéticas com o Modelo PAR(p)

Quando se dispõe de uma série temporal, por exemplo, de medições de vazões naturais, comumente a denominamos de série histórica. A série histórica é apenas uma das possíveis realizações de um processo estocástico, ou seja, pode-se imaginar que a natureza ‘sorteou’ a série histórica segundo algum conjunto de leis probabilísticas. A um novo sorteio corresponde uma outra série diferente da série histórica, mas igualmente possível (Kelman, 1992).

Utiliza-se o modelo PAR para aproximar este comportamento estocástico, ou seja, o modelo deve permitir que, artificialmente, se façam tantos sorteios quantos forem necessários para o estudo em foco.

Cada sorteio está associado a uma série temporal, chamada de série sintética.

Seja a equação do modelo (4.10a):

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

De posse da expressão (4.18) pode-se escrever (4.10a) na forma:

$$Z_t = \mu_m + \phi_1^m \sigma_m \left( \frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \sigma_m \left( \frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + \sigma_m \sigma_a^{(m)} E_t \quad (4.19a)$$

ou

$$Z_t = \mu_m + \phi_1^m \sigma_m \left( \frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \sigma_m \left( \frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + \sigma_m \sqrt{1 - \phi_1^m \rho^m(1) - \phi_2^m \rho^m(2) - \dots - \phi_{p_m}^m \rho^m(p_m)} E_t \tag{4.19b}$$

onde a série de ruídos  $\{E_t\}$  é independente com média zero e variância unitária.

A expressão (4.19b) é particularmente conveniente na geração de séries sintéticas, pois permite um esquema de geração seqüencial no qual, a cada instante  $t$ ,  $Z_t$  é obtido como uma função de valores  $Z_{t-\tau}$ ,  $\tau = 1, \dots, p_m$ , e de um sorteio  $E_t$ .

Para se iniciar o processo de geração é necessário arbitrar valores iniciais  $Z_{t-\tau}$ , onde  $\tau$  pode ir de 1 a  $p_m$ . Estimativas para esses valores iniciais podem ser, por exemplo, as médias mensais ou os últimos valores amostrais.

O conjunto dos últimos valores observados ( $Z_{t-\tau}$ ) em uma série hidrológica é denominado de tendência hidrológica e a média mensal da série histórica é chamada de MLT mensal.

A estrutura de geração das séries sintéticas mostrada na Figura 16 é usualmente conhecida como geração em paralelo ou pente. Esta é a estrutura de geração utilizada pelo modelo NEWAVE.

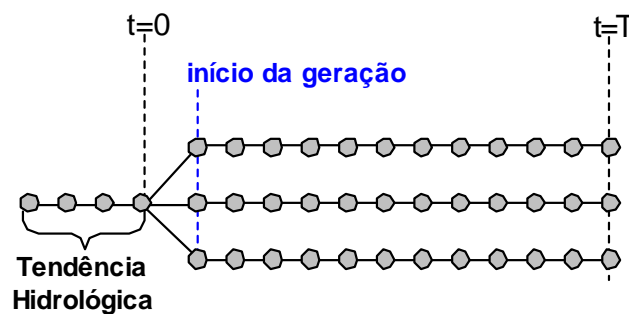


Figura 16: Tendência Hidrológica

A geração das séries sintéticas pode ser realizada de forma condicionada ou não aos valores iniciais. Para não condicionar a série sintética aos valores iniciais devem-se desprezar, por exemplo, os primeiros valores gerados. Para tanto, o processo de geração da série sintética deve ser realizado em duas etapas. Na primeira etapa, a geração das séries sintéticas utiliza como valores iniciais as médias correspondentes a cada período (MLT mensal), conforme ilustrado na Figura 17. Os valores gerados, resultantes da primeira etapa, para os últimos períodos são utilizados como valores iniciais na segunda etapa do processo de geração, conforme Figura 18. Desta forma, as séries sintéticas geradas durante a segunda etapa são ditas ‘não condicionadas’.

Para gerar séries sintéticas ‘condicionadas’ deve-se efetuar apenas a primeira etapa do processo descrito acima, utilizando como valores iniciais uma tendência hidrológica apropriada.

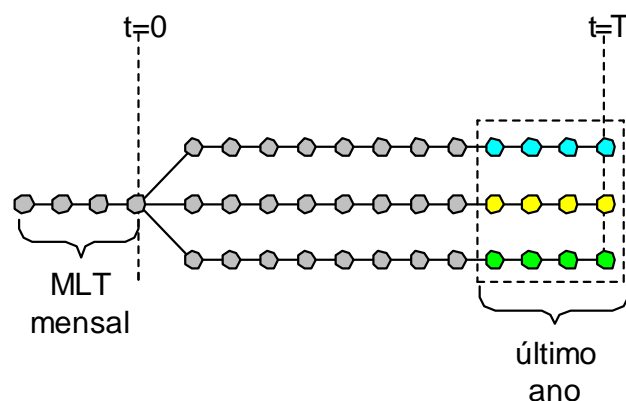


Figura 17: Geração não condicionada – Primeira Etapa

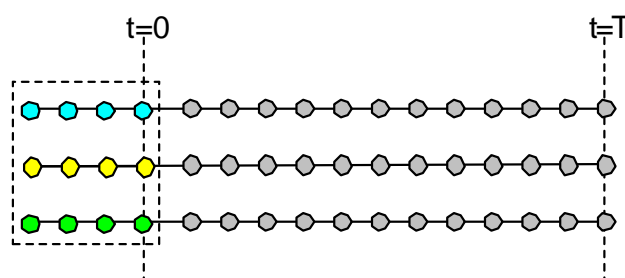


Figura 18: Geração não condicionada – Segunda Etapa

No modelo NEWAVE as séries sintéticas de energias naturais afluentes são necessárias no módulo de cálculo da política de operação hidrotérmica e no módulo de simulação da operação. A política de operação é calculada através de um processo iterativo onde a cada iteração são realizadas uma simulação *backward* e uma simulação *forward*. De maneira simplificada, a simulação da operação pode ser considerada como uma simulação *forward*.

As séries sintéticas utilizadas nas simulações *forward* e *backward* são geradas a partir do mesmo modelo matemático, porém o procedimento para o cálculo destas séries é ligeiramente diferente.

#### 4.3.3.1. Simulação Forward

A geração das séries sintéticas para a simulação *forward* é feita de forma seqüencial, isto é, a cada instante  $t$ ,  $Z_t$  é obtido como uma função dos  $p_m$  valores

passados e de um sorteio  $E_t$ . Assim, para cada instante  $t$  é sorteado um conjunto diferente de ruídos. Na Figura 19 é apresentado o esquema de geração para a simulação *forward*, supondo três séries sintéticas.

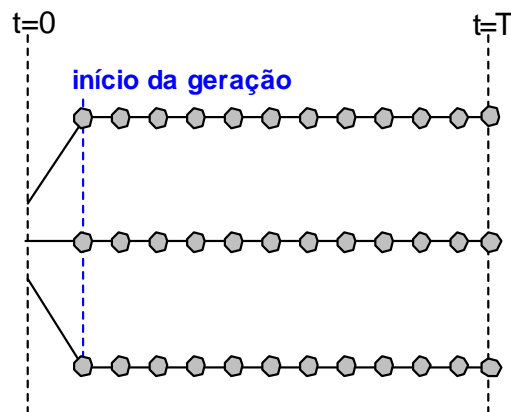


Figura 19: Geração em Paralelo (Pente) para a simulação Forward

Durante a simulação *forward* apenas um valor de afluência é utilizado a cada período e série. Já na simulação *backward*, um número maior de valores de afluência é utilizado a cada período e série. Cada valor de afluência que é utilizado em um determinado período e série é denominado de 'abertura'.

A geração das afluências para as aberturas é feita de forma seqüencial, isto é, a cada instante  $t$  e série  $i$ ,  $Z_{t,i}$  é obtido como uma função dos  $p_m$  valores passados da série  $i$  e de um sorteio  $E_t$ . Os valores passados da série  $i$  são aqueles gerados para a simulação *forward*, Figura 19.

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

Na Figura 20 é mostrada a seqüência de geração para as aberturas da simulação *backward*, supondo três séries sintéticas com três aberturas.

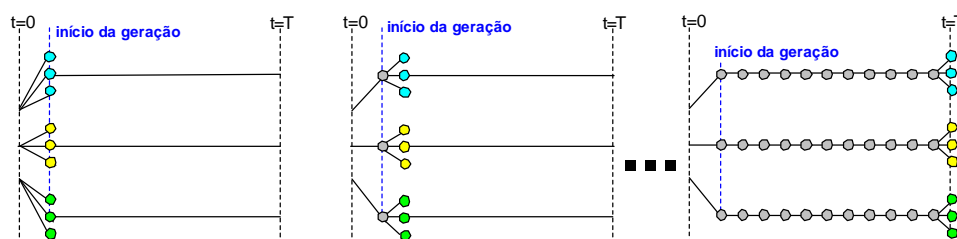


Figura 20: Seqüência de geração para a simulação Backward

Na Figura 21 é apresentado o esquema de geração para a simulação backward.

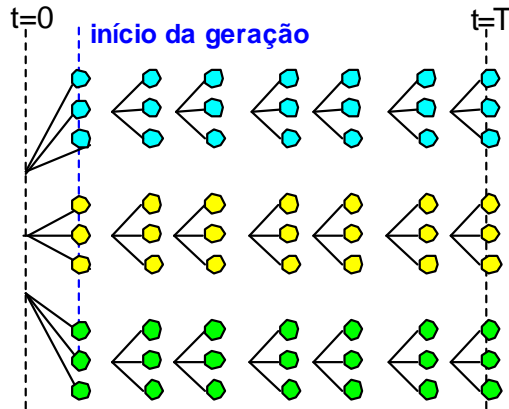


Figura 21: Geração em Paralelo (Pente) para a simulação Backward

#### 4.3.4. Vazões Incrementais Negativas

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

$$Z_t = \mu_m + \phi_1^m \sigma_m \left( \frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \sigma_m \left( \frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + \sigma_m a_t > 0 \quad (4.20a)$$

ou

$$a_t > - \left( \frac{\mu_m}{\sigma_m} \right) - \phi_1^m \left( \frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) - \dots - \phi_{p_m}^m \left( \frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \quad (4.20b)$$

Muitos pesquisadores consideram que os resíduos  $a_t$  apresentam distribuição Normal e uma possível não-normalidade pode ser corrigida pela transformação Box-Cox (Maceira, 1989). Como as séries sintéticas produzidas serão utilizadas em modelos que calculam as estratégias ótimas de operação de um sistema multi-reservatórios, baseados em programação dinâmica dual estocástica, o modelo de geração de séries sintéticas deve ser aplicado diretamente à série temporal original e deve ser capaz de lidar com resíduos que apresentam um forte coeficiente de assimetria.

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

$$\xi_t = \ln(a_t - \Delta) \quad (4.21)$$

Os parâmetros  $\Delta$  e  $\sigma_\xi^{2(m)}$  são estimados de tal forma a preservar os momentos dos resíduos, através das seguintes relações (Charbeneau, 1978):

$$\sigma_\xi^{2(m)} = \ln(\theta) \quad (4.22)$$

$$\Delta = \frac{-\sigma_a^{(m)}}{\sqrt{(\theta - 1)}} \quad (4.23)$$

Fazendo  $\Delta$  igual à expressão (4.20b) pode-se escrever:

$$\theta = 1 + \frac{\sigma_a^{2(m)}}{(-\Delta)^2} \quad (4.24)$$

$\sigma_\xi^{2(m)}$  pode ser diretamente obtido a partir das expressões (4.22) e (4.24).

#### 4.3.5. Correlação Espacial

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

$$W_t = D\xi_t \quad (4.25)$$

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

A matriz  $D$  pode ser estimada por:

$$DD^T = \hat{U} \quad (4.26)$$

onde  $\hat{U}$  pode ser, em princípio, a estimativa da matriz de covariâncias  $[E(\hat{\xi}_t, \hat{\xi}_t^T)]$ . Por exemplo, o elemento  $u_{ab}$  é a covariância espacial lag zero entre os resíduos das usinas  $a$  e  $b$ . Na prática, o comportamento dos resíduos não segue o comportamento das vazões: os resíduos não são espacialmente correlacionados. No entanto, a fim de se preservar as dependências espaciais entre as usinas, utilizou-se a correlação

espacial entre as vazões em substituição à correlação espacial entre os resíduos (Maceira & Jardim, 2000). A matriz D pode ser estimada adotando-se D como uma matriz triangular inferior ou por decomposição espectral.

#### **4.4. Avaliação do Desempenho do Modelo**

##### **4.4.1. Geração em Paralelo**

Um modelo de geração de séries sintéticas deve preservar as principais características da série observada. Em termos práticos, a utilidade de um modelo pode ser aferida por sua capacidade de reproduzir distribuições de probabilidade de variáveis aleatórias relevantes ao processo. Estas variáveis devem naturalmente refletir os requisitos da aplicação proposta. No caso do planejamento da expansão e operação de um sistema de geração de energia elétrica, a representação de períodos críticos, as correlações serial e espacial entre as vazões afluentes aos diversos reservatórios, curvas de regularização para diversos reservatórios, etc., devem de alguma forma ser levados em conta nesta avaliação (Kelman & Pereira, 1977).

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

Na Figura 22 é ilustrado o conceito de seqüência e das variáveis associadas. A linha contínua representa afluências e a linha pontilhada um limite pré-estabelecido. Os intervalos  $(t_1 - t_2)$  e  $(t_3 - t_4)$  correspondem às seqüências negativas, isto é, as vazões estão abaixo dos limites.

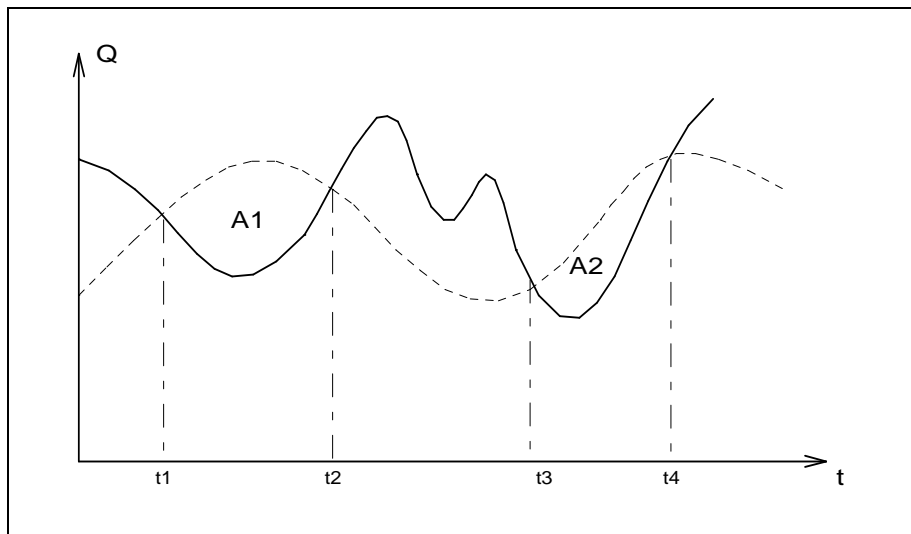


Figura 22: Seqüência Negativa

As variáveis associadas são:

**Comprimento de seqüência:** corresponde ao comprimento dos intervalos  $(t_2 - t_1)$ ,  $(t_4 - t_3)$ , na Figura 22 e é calculado com em (4.27).

$$C = t_2 - t_1 \quad (4.27)$$

**Soma de seqüência:** corresponde à área abaixo do limite durante a seqüência, equação (4.28). Na Figura 22, as somas de seqüências são as áreas A1 e A2

$$S = \sum_{i=t_1}^{t_2} (Z_i - \hat{\mu}_i) \quad (4.28)$$

**Intensidade de seqüência:** corresponde ao valor médio abaixo do limite, isto é, a soma de seqüência dividida pelo respectivo comprimento de seqüência.

$$I = \frac{S}{C} \quad (4.29)$$

Repetindo estes cálculos para cada seqüência negativa é possível obter da série histórica amostras de três variáveis aleatórias:

- Soma de seqüência negativa
- Comprimento de seqüência negativa
- Intensidade de seqüência negativa

O mesmo procedimento é realizado para a série sintética, onde o limite pré-determinado é o mesmo usado na série histórica, e obtemos outras três amostras.



Podemos, então, testar a hipótese de que as duas amostras são provenientes de distribuições idênticas.

Seja  $n_1$  o número de observações da variável aleatória  $X$ , obtida a partir da série histórica e  $n_2$  o número de observações da variável aleatória  $Y$ , obtida a partir da série gerada. Caso essas variáveis aleatórias sejam discretas (ex.: comprimento de seqüência) utilizamos o teste de igualdade de duas distribuições multinomiais. Uma descrição detalhada deste teste é fornecida por Mood et al (1974).

O teste se baseia na divisão do espaço amostral em  $(k+1)$  subconjuntos e a hipótese nula estabelece que:

$$H_0: p_{1,j} = p_{2,j} \quad j = 1, \dots, k+1 \quad (4.30)$$

onde  $p_{1,j}$  é a probabilidade de que uma observação da primeira população pertença ao  $j$ -ésimo subconjunto e  $p_{2,j}$  é a correspondente probabilidade para a segunda população.

Demonstra-se que:

$$\chi_k^2 = \sum_{i=1}^2 \sum_{j=1}^{k+1} \frac{\left( N_{ij} - n_i \frac{(N_{1j} + N_{2j})}{(n_1 + n_2)} \right)^2}{n_i \frac{(N_{1j} + N_{2j})}{(n_1 + n_2)}} \quad (4.31)$$

tem assintoticamente distribuição Qui-quadrado com  $k$  graus de liberdade, onde  $N_{ij}$  é o número de observações do  $j$ -ésimo subconjunto da  $i$ -ésima população e  $n_i$  é o número de observações da  $i$ -ésima variável aleatória.

Caso as variáveis aleatórias em estudo sejam contínuas, por exemplo, a soma e intensidade de seqüência, o teste de Kolmogorov-Smirnov (Stephens, 1974) para duas amostras é bastante adequado.

Sob a hipótese nula da igualdade entre duas distribuições, a variável aleatória  $W = \max_x |S_1(x) - S_2(x)|$  tem distribuição conhecida.  $S_1(x)$  é a distribuição de freqüências acumulada relativa à série histórica e  $S_2(x)$  é a correspondente distribuição relativa à série gerada. Para grandes valores de  $n_1$  e  $n_2$ , o percentual de 95% de  $W$  é aproximadamente dado por:

$$W_{\text{crítico}} = 1,358 \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad (4.32)$$

Quando  $W > W_{\text{crítico}}$  a hipótese de igualdade entre as distribuições é rejeitada. Descrição detalhada é fornecida por Bradley (1968).

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

$$S_t = \sum_{j=1}^t (Z_j - \beta \bar{\mu}) \quad (4.33)$$

onde  $\beta$  é chamado de nível de regularização.

Neste caso a seqüência de somas parciais é dada por:

$$S_t = \sum_{j=1}^t Z_j - \beta t \bar{\mu} \quad (4.34)$$

e está ilustrada na Figura 23.

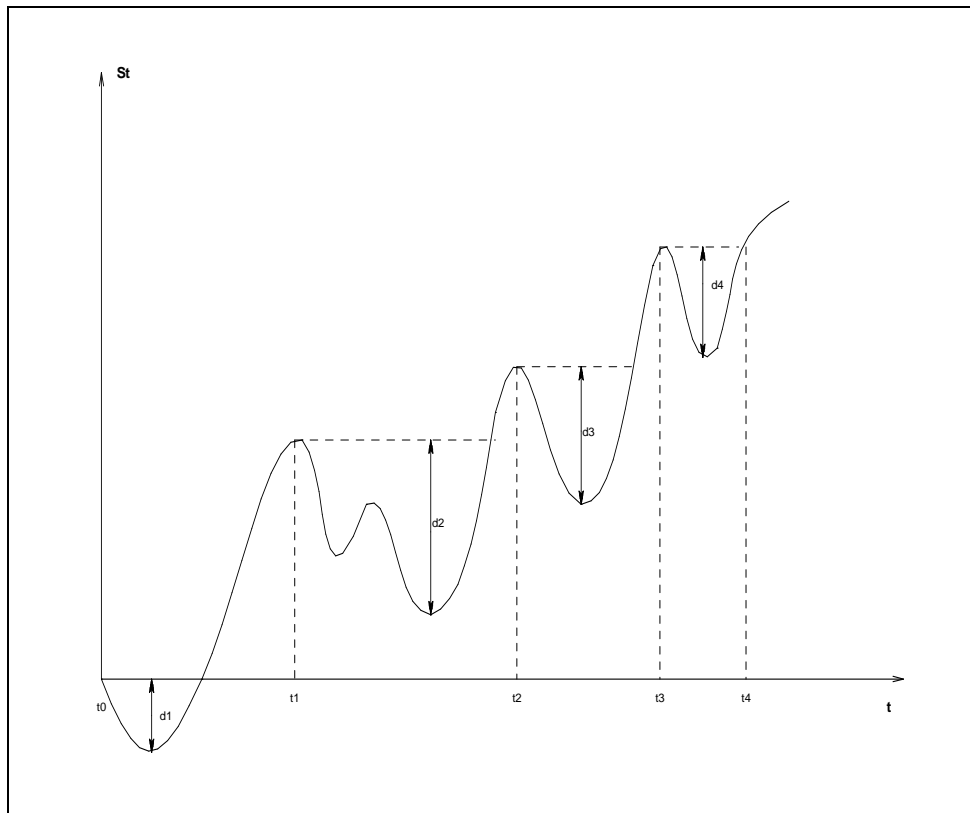


Figura 23: Volume de regularização (Déficit)

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

Um modelo orientado para aplicação no setor energético deve estar apto a produzir séries sintéticas que revelem uma distribuição amostral para a variável aleatória **déficit** estatisticamente indistinguível da distribuição amostral obtida a partir da série histórica.

Verificou-se, entretanto, que os índices mais relevantes para o planejamento estavam em geral associados a valores extremos das distribuições. O período crítico, por exemplo, corresponde à pior situação hidrológica em todo o histórico. Esta constatação é consequência da existência de reservatórios plurianuais, pouco sensíveis a flutuações de curto prazo. O primeiro índice do tipo 'máximo' a ser definido foi o chamado máximo déficit, que corresponde ao volume do reservatório capaz de regularizar uma vazão especificada.

Rippl (1883) propôs utilizar o **máximo déficit**, obtido a partir da série histórica, como valor de projeto para a capacidade de um reservatório.

Para estudo das variáveis do tipo 'máximo', o procedimento é o seguinte:

- Divide-se a série gerada em M segmentos de comprimento igual ao da série histórica;
- Calcula-se o valor do índice para cada segmento, obtendo-se assim uma distribuição de frequências.

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

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

Esta análise pode ser realizada para as seguintes estatísticas:

- Máximo comprimento de seqüência
- Máxima soma de seqüência
- Máxima intensidade de seqüência
- Máximo déficit para um determinado nível de regularização
- Comprimento do período crítico
- Vazão média de período crítico

#### 4.4.2. Geração em Árvore

Alguns testes foram implementados para verificar se os cenários gerados estão coerentes com o modelo ajustado. Para tanto foram avaliados os momentos mensais condicionados de primeira e segunda ordem para os valores gerados.

##### 4.4.2.1. Teste de Momentos Condicionados de Ordem Um e Dois

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

Seja o modelo PAR(p) para uma usina hidrelétrica expresso como:

$$\left( \frac{Z_t - \mu_m}{\sigma_m} \right) = \phi_1^m \left( \frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \left( \frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t \quad (4.35)$$

onde  $a_t$  é uma série de ruídos independentes com média zero e variância  $\sigma_a^{2(m)}$ ,  $Z_t$  é a série de energias naturais afluentes a um subsistema, e os parâmetros  $\mu_m$  e  $\sigma_m$  são a média e o desvio padrão históricos, respectivamente, do subsistema considerada.

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

$$E(Z_t | Z_{t-}) = \left[ \phi_1^m \left( \frac{Z_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \phi_{p_m}^m \left( \frac{Z_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) \right] \sigma_m + \mu_m \quad (4.36)$$

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

$$\text{VAR}(Z_t | Z_{t-}) = \sigma_m^2 \sigma_a^{2(m)} = \sigma_m^2 \left[ 1 - \phi_1^m \rho^m(1) - \phi_2^m \rho^m(2) - \dots - \phi_{p_m}^m \rho^m(p_m) \right] \quad (4.37)$$

E o desvio padrão condicionado é dado por:

$$\sigma_{Z_t|Z_{t-}} = \sigma_m \sigma_a^{(m)} = \sigma_m \sqrt{1 - \phi_1^m \rho^m(1) - \phi_2^m \rho^m(2) - \dots - \phi_{p_m}^m \rho^m(p_m)} \quad (4.38)$$

#### 4.4.2.2. Parâmetros das distribuições amostrais

Em Yevjevich (1972) são apresentados os parâmetros da distribuição amostral para a média e o desvio-padrão. Estes parâmetros são necessários para verificar se os momentos condicionados de primeira e segunda ordem obtidos a partir dos N cenários gerados são condizentes com os valores teóricos calculados para estes momentos.

##### a) Média

O valor esperado da média amostral,  $E(\bar{x})$ , é igual à média da população (4.39) e a variância da média amostral,  $VAR(\bar{x})$ , é igual à variância da população dividida pelo número de elementos da amostra (4.40).

$$E(\bar{x}) = \mu = E(Z_t | Z_{t-}) \quad (4.39)$$

$$VAR(\bar{x}) = \frac{\mu_2}{N} = \frac{VAR(Z_t | Z_{t-})}{N} \quad (4.40)$$

onde  $\mu$  e  $\mu_2$  são a média e a variância da população.

Pelo teorema do limite central pode-se dizer que a média tem distribuição assintoticamente Normal. Então o intervalo de confiança de 95% calculado para a média amostral estimada a partir dos N cenários hidrológicos gerados, para um dado subsistema, é dado por:

$$IC_\mu = E(Z_t | Z_{t-}) \pm 1,96 \frac{\sigma_{Z_t|Z_{t-}}}{\sqrt{N}} \quad (4.41)$$

##### b) Desvio Padrão

O valor esperado do desvio padrão amostral é igual ao desvio padrão da população (4.42).

$$E(s) = \sigma = \sigma_{Z_t|Z_{t-}} \quad (4.42)$$

A variância do desvio padrão amostral, VAR(s), é definida em (4.43a). Substituindo o momento de quarta ordem em (4.43a) obtém-se (4.43b). A definição do momento central de quarta ordem de população correspondente à distribuição Lognormal pode ser encontrada em Yevjevich (1972).

$$\text{VAR}(s) = \frac{\mu_4 - \mu_2^2}{4\mu_2 N} \quad (4.43a)$$

$$\text{VAR}(s) = \frac{\mu_4 - \mu_2^2}{4\mu_2 N} = \frac{\mu_2(\eta^8 + 6\eta^6 + 15\eta^4 + 16\eta^2 + 2)}{4N} \quad (4.43b)$$

onde:

$\mu_2$  e  $\mu_4$  são os momentos centrados de segunda e quarta ordem da população

$\eta = \sqrt{e^{\sigma_n^2} - 1}$  é o coeficiente de variação da população e  $\sigma_n^2$  é o parâmetro de forma da Distribuição Lognormal 3 parâmetros ajustada para a geração dos cenários de afluências.

Com o intuito de estabelecer um intervalo de confiança de 95% (4.44) para a estimativa do desvio padrão, foi calculada uma tabela com valores-padrão, VP (Jardim, 2002). A Tabela 1 foi construída através da simulação de 100.000 amostras com 1000 cenários, geradas por distribuições Lognormal com  $\sigma_n^2$  variável e  $\mu_n$  igual a 1.

$$\text{IC}_\sigma = \sigma_{Z_t|Z_{t-}} + \text{VP} \frac{\sigma_{Z_t|Z_{t-}} \sqrt{\eta^8 + 6\eta^6 + 15\eta^4 + 16\eta^2 + 2}}{\sqrt{4N}} \quad (4.44)$$

Tabela 1: Valores Padrão para IC 95%

<b>Valores-Padrão (VP)</b>		
$\sigma_n$	<b>Limite Inferior</b>	<b>Limite Superior</b>
0.05	-1.9661	1.9570
0.07	-1.9650	1.9648
0.10	-1.9439	1.9850
0.30	-1.9311	2.0045
0.50	-1.8305	2.0673
0.70	-1.6777	2.1220

---

**Valores-Padrão (VP)**

---

$\sigma_n$	Limite Inferior	Limite Superior
0.80	-1.5493	2.0849
0.90	-1.3965	2.0005
1.00	-1.1970	1.8652
1.10	-0.9889	1.6254

---

#### **4.5. Resumo**

Neste capítulo foi apresentada a formulação do modelo autorregressivo periódico e todas as etapas que envolvem o seu ajuste. Também foi mostrada a estratégia adotada para evitar a geração de cenários de energia negativos e para geração de cenários multivariados. Finalmente foram apresentados os testes utilizados para aferir o modelo ajustado.

No próximo capítulo serão apresentadas técnicas de agregação de cenários e técnicas de amostragem que serão aplicadas ao modelo PAR(p) para definição da árvore de cenários a ser utilizada no problema de planejamento da operação energética.

## **5** **Definição da Árvore de Cenários**

### **5.1.** **Considerações Iniciais**

Com o intuito de propor um método para definir a sub-árvore a ser visitada durante o processo do cálculo da estratégia ótima de operação, de forma a tornar os resultados obtidos mais robustos com relação ao uso de diferentes amostras e à variações da cardinalidade destas amostras, neste capítulo serão apresentadas técnicas de seleção de cenários e técnicas de amostragem. As técnicas propostas poderão ser aplicadas separadamente ou em conjunto com o objetivo de representar de forma mais acurada o processo estocástico de afluências a ser utilizado durante o cálculo da política ótima de operação.

As técnicas de seleção de cenários, quando aplicadas a um grande número de cenários hidrológicos gerados, proporcionam a escolha de um conjunto representativo de cenários. O conjunto resultante de cenários deve representar de forma adequada o processo estocástico que o deu origem, no caso o processo estocástico de afluências.

Para reduzir o erro da estimativa de uma variável obtida por simulação de Monte Carlo deve-se aumentar o tamanho da amostra utilizada, reduzindo desta forma o desvio padrão da estimativa, até que se atinja a precisão desejada. Porém, nem sempre é viável computacionalmente aumentar o tamanho da amostra. Então, outra maneira de reduzir a variância é espalhar a amostra o mais uniformemente possível sobre o espaço amostral. As técnicas de amostragem abordadas neste item têm esse objetivo.

### **5.2.** **Técnicas de Seleção de Cenários**

O principal objetivo da aplicação das técnicas de seleção de cenários é a redução do número de cenários hidrológicos através da escolha de um conjunto representativo de cenários hidrológicos. Os cenários que fazem parte deste conjunto



representativo devem conter toda a informação necessária para representar o processo estocástico de afluições.

Em Jardim (2002) foi tratado o problema de reduzir a árvore de afluições utilizada pelos modelos de planejamento de tal forma que a árvore reduzida preserve as características do processo estocástico do qual as vazões são geradas. Foram empregadas técnicas estatísticas multivariadas capazes de elaborar critérios que possibilitam agrupar cenários similares em determinados grupos. Estas técnicas podem ser reunidas sob o nome genérico de Análise de Conglomerados.

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 podem-se citar trabalhos que empregam as técnicas de agregação para a caracterização de curvas de carga (Velásquez 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 utiliza para a realização de diversos estudos como os citados em Aldenderfer & 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.

Aplicadas a um grande número de cenários hidrológicos gerados, obtidos através de modelos autorregressivos periódicos, as técnicas de agregação proporcionam a escolha de um conjunto representativo de cenários. No exemplo ilustrado na Figura 24, são gerados  $n$  cenários para o quarto período, com base nos valores observados passados ( $Z_3$ ,  $Z_2$  e  $Z_1$ ), e após a aplicação das técnicas de

agregação são obtidos 3 cenários representativos. Os cenários que fazem parte deste conjunto são obtidos através do agrupamento de cenários semelhantes e possuem características similares aos demais componentes do grupo em que estão localizados.

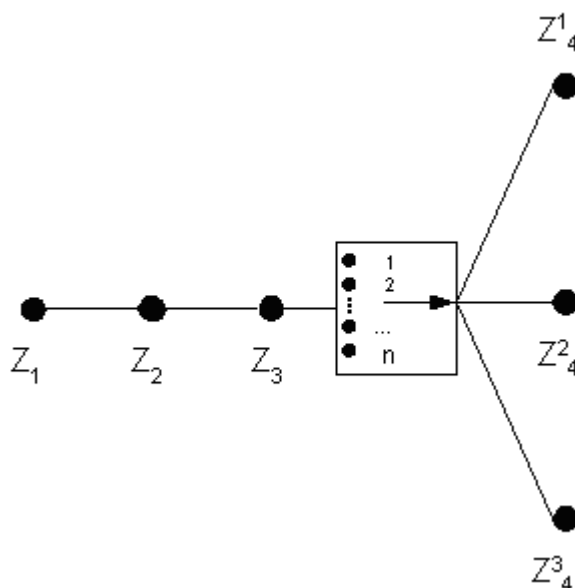


Figura 24: Aplicação técnicas de agregação

Desta forma, usando as técnicas de agregação pode-se escolher um conjunto representativo de cenários hidrológicos a partir de um grande número de séries de vazões geradas, diminuindo assim o esforço computacional sem contudo deixar de representar de forma adequada o processo estocástico das afluições.

### 5.2.1. 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 de 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.

Apesar da aparente simplicidade dos algoritmos utilizados, as técnicas de agregação permitem sugerir hipóteses sobre as relações multivariadas existentes nos dados, muito úteis na elaboração de modelos estatísticos mais sofisticados.

As técnicas de agregação também 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.

### 5.2.1.1. 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. Segundo Aldenderfer & Blashfield (1984), os métodos aglomerativos são mais difundidos e utilizados na literatura.

Os resultados de ambos podem ser apresentados graficamente na forma de um diagrama bidimensional denominado dendograma, que ilustra as fusões ou divisões realizadas em níveis sucessivos. Na Figura 25: Exemplo ilustrativo do processo aglomerativo

é mostrado 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.

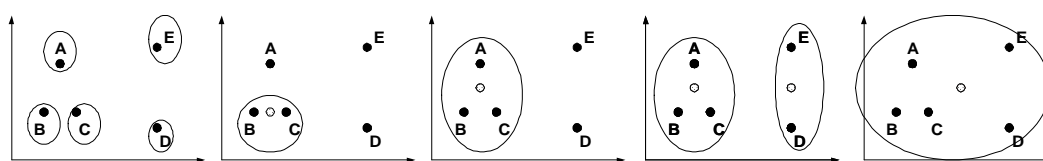


Figura 25: Exemplo ilustrativo do processo aglomerativo

O dendograma resultante desta seqüência de fusões é mostrado na Figura 26.

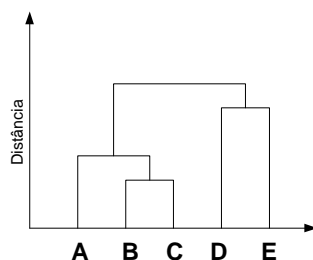


Figura 26: Dendograma

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 realocação 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 & Odell (1970) e Johnson & Wichern (1998).

#### **5.2.1.2. 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 verificar todas as possíveis partições do conjunto de dados, porém o número de possibilidades é muito grande, assintoticamente de ordem de  $K^{N-1}$  (Bussab et al., 1990), 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 (Hartigan & Wong, 1979). 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 & Blashfield (1984), Hair Jr. et al. (1998), Johnson & Wichern (1998) e Bouroche & Saporta (1980).

### **5.2.2. 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). Vale a pena lembrar que toda vez que um objeto for realocado os centróides devem ser recalculados. Este último passo é repetido até que não haja mais realocações de objetos. O algoritmo K-Means pode ser resumido nos seguintes passos:

- Passo 1: Divida os N objetos em K agrupamentos através de uma partição inicial ou especificação de K centróides iniciais;
- Passo 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;
- Passo 3: Repita o passo 2 até que não haja mais realocações de objetos de um grupo para outro.

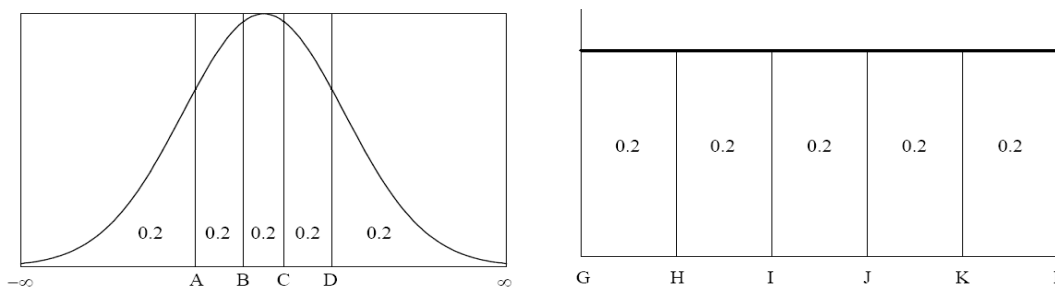
Com o intuito de aperfeiçoar, tornar mais rápido e mais eficiente o algoritmo apresentado, alguns procedimentos podem ser modificados, gerando assim variações deste método. A inicialização dos grupos pode ser feita de forma aleatória através do sorteio de pontos (objetos) para serem usados como semente inicial dos grupos ou pela partição aleatória do conjunto de dados. Os pontos sorteados podem ser sorteados do próprio conjunto de dados ou não. Estes pontos também podem ser escolhidos um a um pelo especialista ou retirados de forma programada 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.

### 5.3. Técnicas de Amostragem

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

#### 5.3.1. Amostragem por Hipercubo Latino

A amostragem por hipercubo latino foi sugerida por McKay et al (1979). Nela, o domínio de cada variável aleatória (VA)  $X_k$  ( $k = 1, \dots, M$ ) é dividido em  $N$  intervalos,  $\Delta X_k^i$  ( $i=1, 2, \dots, N$ ), de igual probabilidade  $1/N$ , como mostrado nas Figura 27a e b.



(a) (b)  
 Figura 27: Divisão em 5 intervalos do domínio de duas VA  
 (a) distribuição normal (b) distribuição uniforme

Na Figura 27 são apresentadas duas variáveis aleatórias, uma tendo distribuição normal e outra, distribuição uniforme. Neste exemplo, o domínio de cada variável foi dividido em cinco intervalos.

O número de intervalos  $N$  na amostragem por hipercubo latino deve ser igual ao tamanho da amostra desejada, ou seja, igual ao número total de simulações. Para cada intervalo é amostrado apenas um valor ( $x_k^i$ ), isto é, este valor será usado em uma e apenas uma simulação.

Os valores amostrados  $x_k^i$ , para um valor  $i$  qualquer, são obtidos pela resolução da equação (5.1):

$$F_k(x_k^i) = \frac{i-1 + R_i}{N} \quad i = 1, 2, \dots, N \quad (5.1)$$

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

A amostragem é realizada utilizando a transformada inversa da função de distribuição de probabilidade em questão, como em (5.2):

$$x_k^i = F_k^{-1}\left(\frac{i-1 + R_i}{N}\right) \quad (5.2)$$

Para exemplificar a amostragem por hipercubo latino, suponha uma distribuição bivariada onde uma variável tem distribuição normal e a outra, uniforme. Na Figura 28 são apresentados os cinco valores sorteados para cada uma das duas variáveis. Esses valores estão marcados nos respectivos eixos. Note que foi sorteado apenas um valor para cada intervalo.

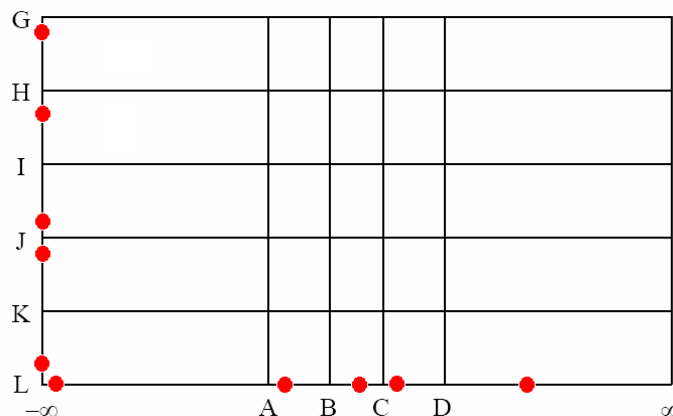


Figura 28: Valores sorteados para cada variável

Após obtidos os  $N$  valores para cada variável  $X_k$ , esses devem ser emparelhados de forma aleatória com os valores das demais variáveis. Dessa forma, são formados  $N$  vetores de dimensão  $M$ . A seleção aleatória do  $i$ -ésimo valor para cada variável é realizada mediante a permutação aleatória dos inteiros  $1, 2, \dots, N$ . Na Figura 29 é apresentada uma possível amostra de cinco pontos gerados utilizando a amostragem por hipercubo latino do exemplo anterior.

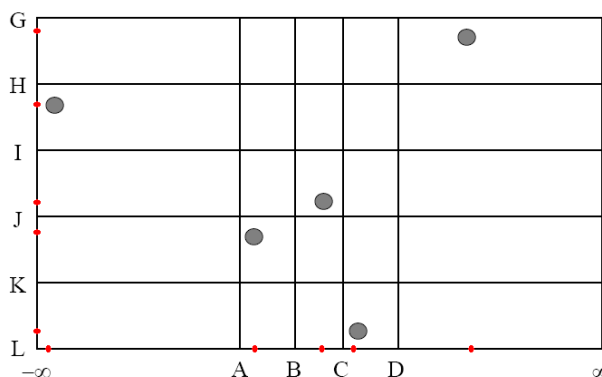


Figura 29: Representação bi-dimensional de uma possível amostragem por hipercubo latino

A amostra dos  $N$  pontos da amostragem por hipercubo latino no espaço Euclidiano de  $M$ -dimensões contém apenas um ponto em cada intervalo de cada uma das  $M$  variáveis.

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



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

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

Uma desvantagem deste método é não poder aumentar o tamanho da amostra aproveitando a amostra já disponível. Amostras pequenas podem não fornecer resultados estatísticos aceitáveis enquanto que amostras grandes demais podem tornar o problema inviável computacionalmente. Em Tong (2006) e Sallaberry et al (2008) são apresentados procedimentos para estender ou refinar amostras geradas por amostragem por hipercubo latino, porém a nova cardinalidade da amostra deve ser um múltiplo da cardinalidade original.

### 5.3.2. Amostragem Descritiva

Na amostragem descritiva (Saliby, 1997), o procedimento se assemelha bastante ao procedimento da amostragem por hipercubo latino, a diferença está na escolha dos valores amostrados. O domínio de cada variável aleatória  $X_k$  também é dividido em  $N$  intervalos de igual probabilidade ( $p=1/N$ ). No caso da amostragem descritiva o ponto  $x_k^i$  é escolhido como sendo o centróide do intervalo  $i$ .

A fórmula usada para a geração do conjunto de valores descritivos, a serem depois permutados aleatoriamente é dada em (5.3).

$$x_k^i = F_k^{-1}\left(\frac{i-1+0.5}{N}\right) = F_k^{-1}\left(\frac{i-0.5}{N}\right) \quad (5.3)$$

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

Em Saliby & Pacheco (2002) é feita uma comparação entre as técnicas de amostragem descritiva, amostragem por hipercubo latino, métodos quase-Monte Carlo e a amostragem aleatória simples. Os resultados obtidos com amostragem descritiva e por hipercubo latino tendem a ser equivalentes à medida que o tamanho da amostra cresce.

Em Jirutitijaroen & Singh (2008) é proposta a utilização da amostragem descritiva nas simulações de Monte Carlo utilizadas para avaliar a eficácia dos métodos analíticos de fluxos de carga probabilístico. Para este problema os resultados obtidos são melhores do que aqueles encontrados utilizando-se a amostragem aleatória simples. Com o intuito de reduzir uma correlação indesejada introduzida pelo método de emparelhamento aleatório das amostras das variáveis de entrada, Yu et al. (2009) propõem o uso da amostragem descritiva combinada com o método de decomposição de Cholesky.

Assim como ocorre com o método de hipercubo latino, no método de amostragem descritiva também não é possível aumentar o tamanho da amostra aproveitando os valores já amostrados. Os trabalhos de Tong (2006) e Sallaberry et al (2008) também podem ser aplicados neste método de amostragem.

### **5.3.3. Quase-Monte Carlo**

Os métodos de quase-Monte Carlo são métodos determinísticos baseados em seqüências de baixa discrepância. Essas seqüências são construídas tal que cada ponto deve ser adicionado de forma que fique o mais distante possível dos demais pontos da seqüência. Desta forma, são preenchidos seqüencialmente os maiores espaços entre os pontos da seqüência.

As seqüências de baixa discrepância são geradas sem qualquer aleatoriedade e existem diversas formas de obtê-las, como por exemplo, seqüências de Halton, Sobol e Faure (Morokoff e Caflisch, 1995).

A seqüência de Halton de uma dimensão é baseada na escolha de um número primo  $p$  e na expansão de uma seqüência de inteiros na base  $p$ . Uma vez definido o número primo  $p$ , o  $k$ -ésimo elemento da seqüência pode ser derivado a partir do seguinte procedimento:

- i) decomponha o número  $k-1$  na base  $p$ , conforme (5.4):

$$k-1 = [a_n \ a_{n-1} \ \dots \ a_2 \ a_1 \ a_0]_p \quad (5.4)$$

ii) some os termos resultantes da decomposição, divididos por potências crescentes do número primo  $p$ , conforme (5.5):

$$x_k = \frac{a_n}{p^{n+1}} + \frac{a_{n-1}}{p^n} + \dots + \frac{a_2}{p^3} + \frac{a_1}{p^2} + \frac{a_0}{p} \quad (5.5)$$

No caso multivariado, a seqüência de Halton para cada dimensão utiliza um número primo  $p$  diferente. A primeira dimensão usa como base o número 2, a segunda dimensão usa o número 3, e assim por diante.

A seqüência de Sobol segue o mesmo princípio da seqüência de Halton, porém para o caso multivariado o número primo utilizado em cada uma das dimensões é o mesmo e é igual a 2 ( $p=2$ ). Os elementos da seqüência são reordenados através de técnicas de permutação aleatória para cada dimensão.

Na Figura 30 é apresentado um exemplo de duas amostras obtidas com seqüências de Sobol.

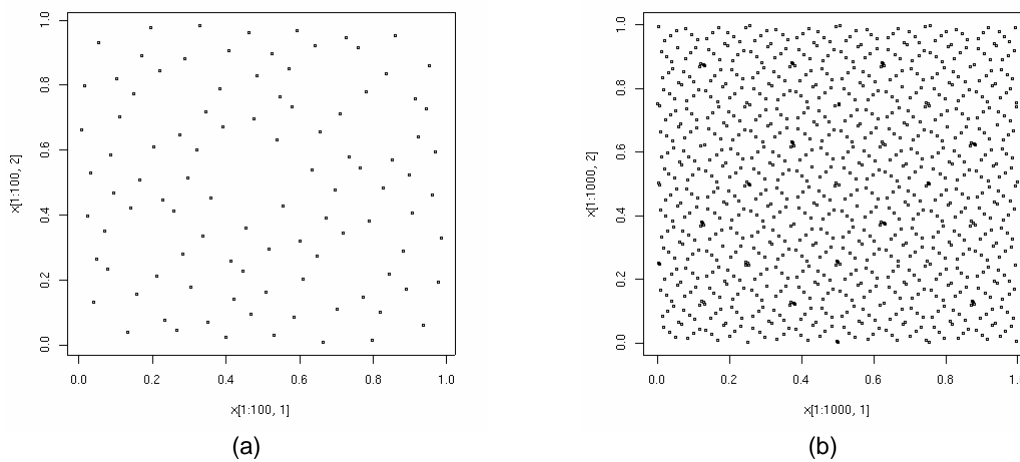


Figura 30: Representação da amostragem utilizando método Quase-Monte Carlo (Sobol)  
 (a) 100 pontos (b) 1000 pontos

Fonte: Frota (2003)

A seqüência de Faure segue o mesmo princípio da seqüência de Halton. Porém, apesar de usar o mesmo número primo como base para todas as dimensões, este não é fixo e depende do tamanho da amostra. Esse número é definido como sendo o menor número primo maior ou igual ao tamanho da amostra. Assim como a seqüência de Sobol, no caso multivariado os valores da seqüência são reordenados utilizando-se técnicas de permutação aleatória para cada dimensão.

Na Figura 31 é apresentada uma comparação de uma amostra de 10000 pontos gerados com amostragem aleatória simples e utilizando um método de quase-Monte Carlo (seqüência de Sobol). Note que a amostra gerada pela seqüência Sobol está mais uniformemente distribuída do que a amostra aleatória simples.

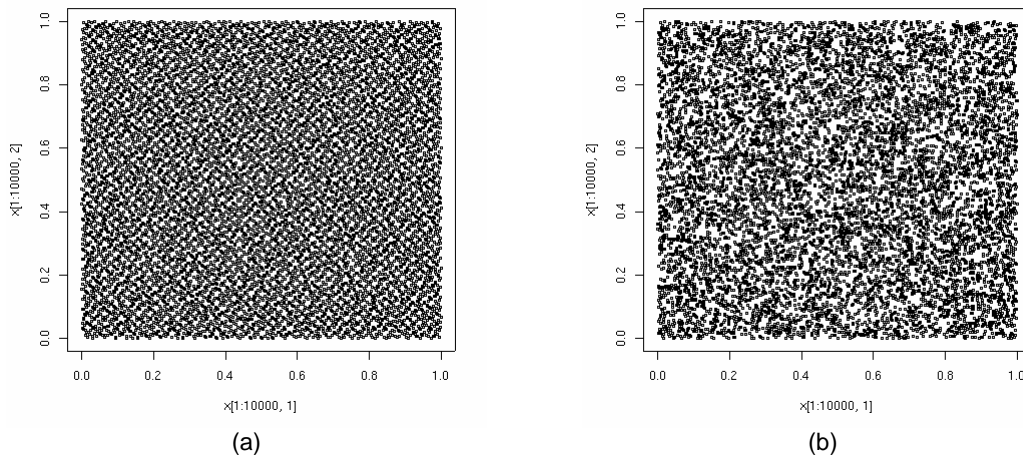


Figura 31: Amostra (a) Seqüência de Sobol (b) Amostragem Aleatória Simples

Fonte: Frota (2003)

Morokoff e Caflisch (1995) comparam o uso das seqüências de Halton, Sobol e Faure com a amostragem aleatória simples. Os métodos quase-Monte Carlo tiveram um desempenho melhor, mas essa vantagem se reduz à medida que o número de variáveis considerado aumenta (problema da dimensionalidade). Na Figura 32 são apresentados pontos gerados pela seqüência de Sobol com dimensão 50 e pela seqüência de Halton com dimensão 20, para a mesma distribuição.

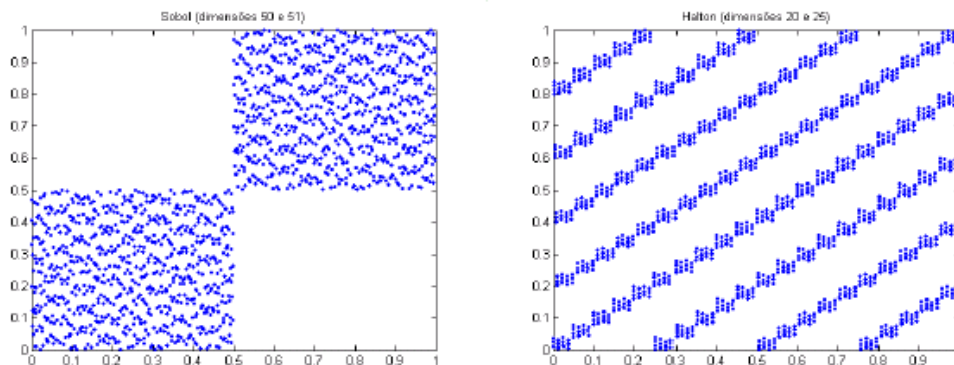


Figura 32: Amostra (a) Seqüência de Sobol [50] (b) Seqüência de Halton [20]

Fonte: Frota (2003)

A técnica de geração adotada pelo gerador de cenários hidrológicos deve ser capaz de gerar cenários multivariados tanto para a abordagem que utiliza a representação da configuração hidráulica por subsistemas equivalentes, atualmente adotada pelo modelo NEWAVE, quanto para a representação híbrida de subsistemas equivalentes e individualizados proposta por Marcato (2002). Neste último caso, a dimensionalidade dos cenários hidrológicos multivariados passa a ser elevada, pois o SIN é um sistema de grande porte. Logo, a utilização dos métodos de quase-Monte Carlo não é recomendada.

#### **5.4. Resumo**

Neste capítulo foram descritas técnicas de seleção de cenários que têm como objetivo achar objetos representativos dentro de uma grande amostra de dados, de forma a reduzir a cardinalidade do problema sem afetar a representatividade da amostra. Além disso, foram apresentadas técnicas de amostragem, alternativas à amostragem aleatória simples.

No próximo capítulo serão apresentadas as propostas para construção da árvore de cenários, aplicando-se as técnicas discutidas neste capítulo.

## 6

### Aplicação do Modelo de Geração de Cenários

#### 6.1.

##### Considerações Iniciais

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 autorregressivo periódico de ordem  $p$ ,  $PAR(p)$ , que modela a afluência de um mês como sendo função das afluências dos  $p$  meses anteriores. A amostra de ruídos aleatórios utilizada pelo modelo  $PAR(p)$  é obtida atualmente através de amostragem aleatória simples, opção Atual. Neste capítulo são propostas alternativas para construção da sub-árvore a ser visitada durante a definição da estratégia ótima de operação.

#### 6.2.

##### Alternativas para Construção da Árvore de Afluências

O método proposto 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  $N(0,1)$  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 tratados como equiprováveis. Logo após é realizada a agregação desses objetos de forma a reduzir a dimensionalidade da amostra original. Na Figura 33 é ilustrado o procedimento proposto.

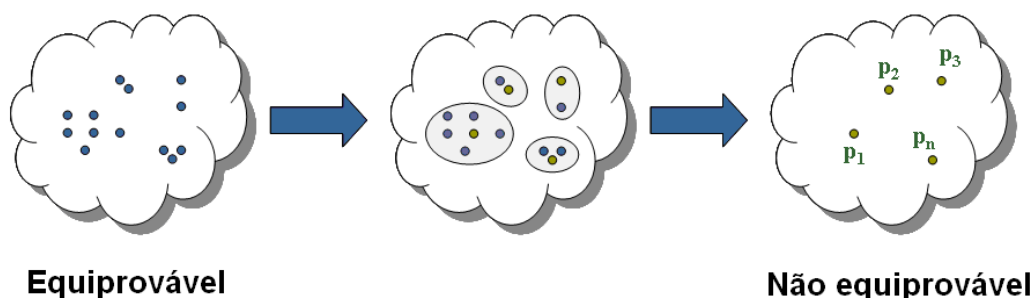


Figura 33: Aplicação do Procedimento de Agregação

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 por um 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. Na Figura 34 é ilustrado como é escolhido o representante de cada grupo formado, o ponto médio dos grupos está assinalado com um x.

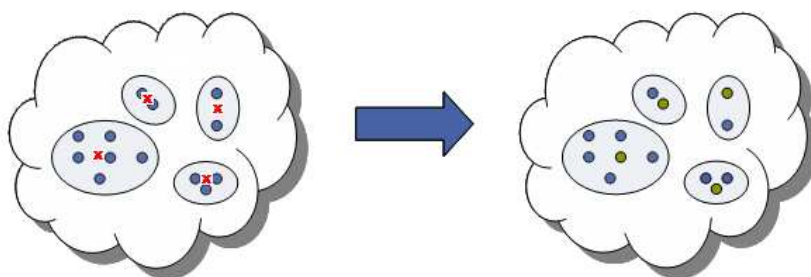


Figura 34: Escolha do Objeto Representativo

Os vetores de ruídos multivariados resultantes do processo de agregação não são mais tratados como 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 (6.1):

$$P_k = \frac{NO_k}{N} \quad (6.1)$$

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. A probabilidade de um cenário no passo forward é calculada como o produtório das probabilidades do ruído em cada um dos estágios. Essa probabilidade é padronizada de forma que a soma da probabilidade de todos os cenários somem 1. Na Figura 35 é apresentada a evolução das probabilidades dos cenários do passo forward ao longo dos estágios. A probabilidade no último estágio é a probabilidade de ocorrência do cenário. A ocorrência de uma extensa seqüência de ruídos com elevada probabilidade, ao longo de períodos consecutivos, pode gera cenários com probabilidade muito maior do que os demais cenários. Observe na Figura 35 que um dos cenários apresenta uma probabilidade de ocorrência de aproximadamente 70%. Cenários no passo forward com probabilidades muito diferentes ocasionam problemas na convergência do processo de otimização. Como o limite superior do valor ótimo da função objetivo (ZSUP) é calculado como a média ponderada dos custos totais de operação de cada um dos cenários, se um dos cenários for muito mais provável que os demais cenários, o valor de ZSUP refletirá basicamente a operação do cenário com elevada probabilidade.

Para contornar este problema, uma nova amostra de ruídos equiprováveis será construída para o passo forward a partir da amostra resultante do processo de agregação por meio de um sorteio condicionado.



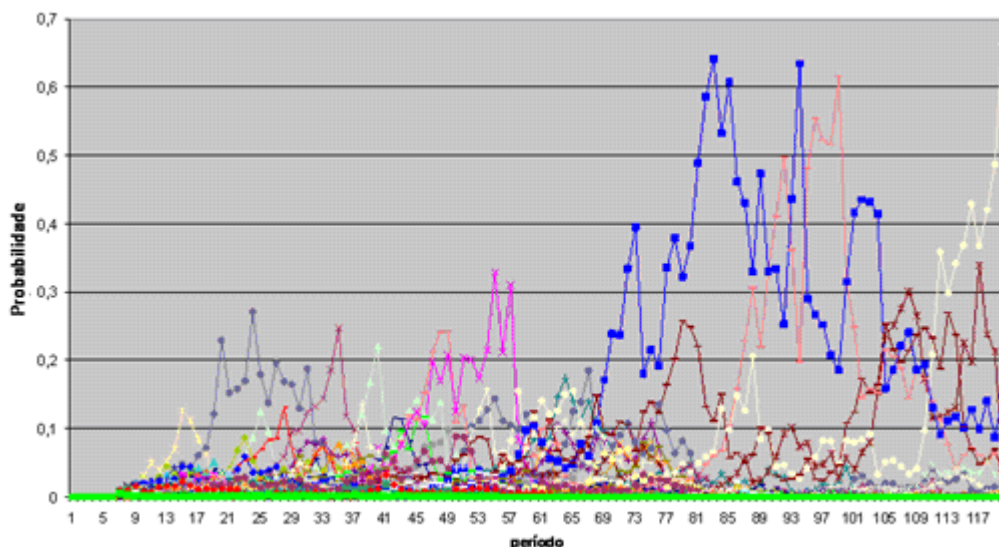


Figura 35: Probabilidade dos Cenários Forward

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 36.

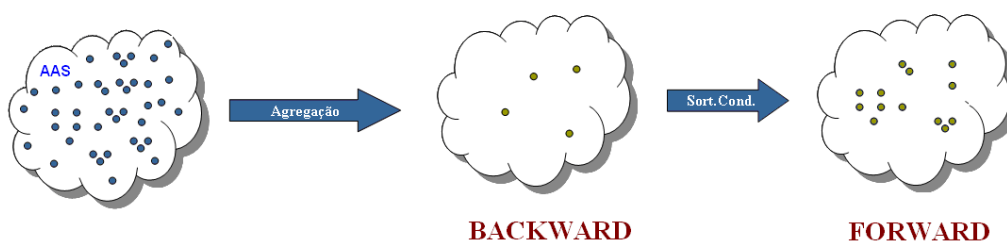


Figura 36: Opção 0

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 37.

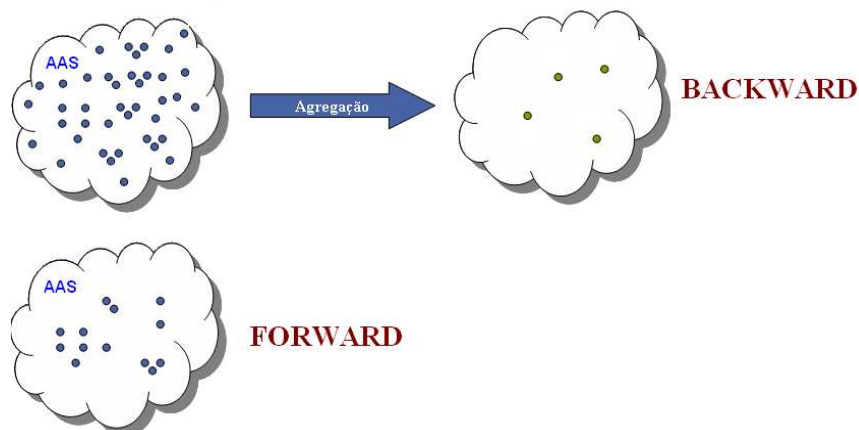


Figura 37: Opção 1

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 tratados como equiprováveis.

Na terceira alternativa, Figura 38, 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**.

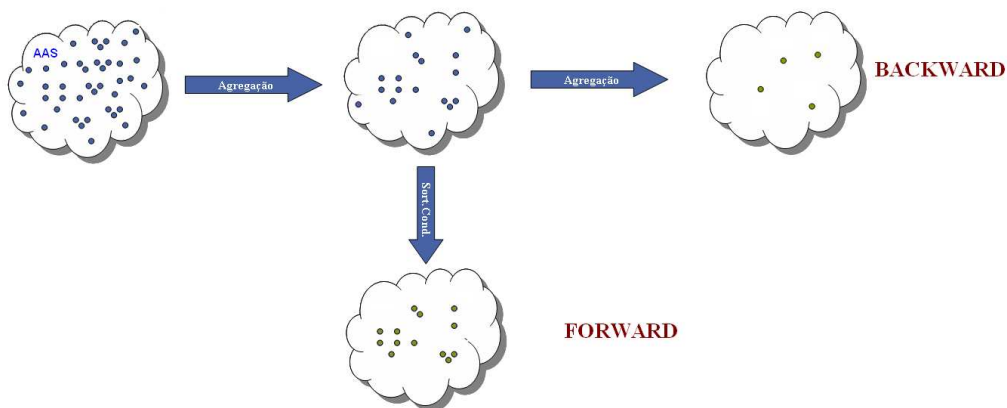


Figura 38: Opção 2

Na quarta alternativa (**opção 3**), Figura 39, 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.

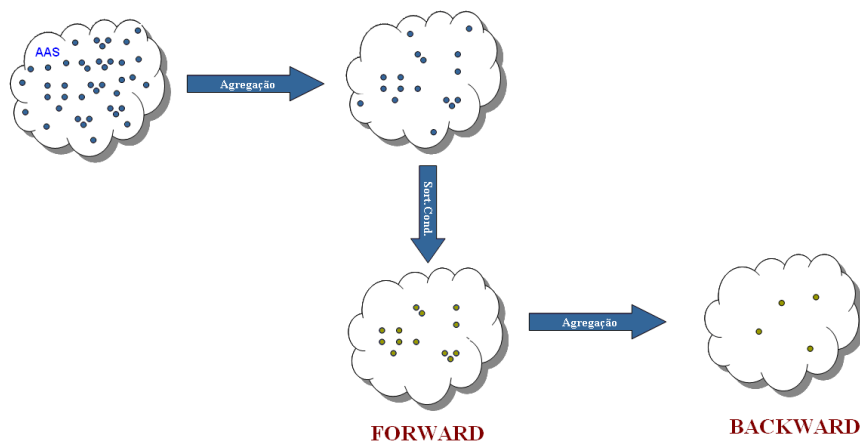


Figura 39: Opção 3

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 40.

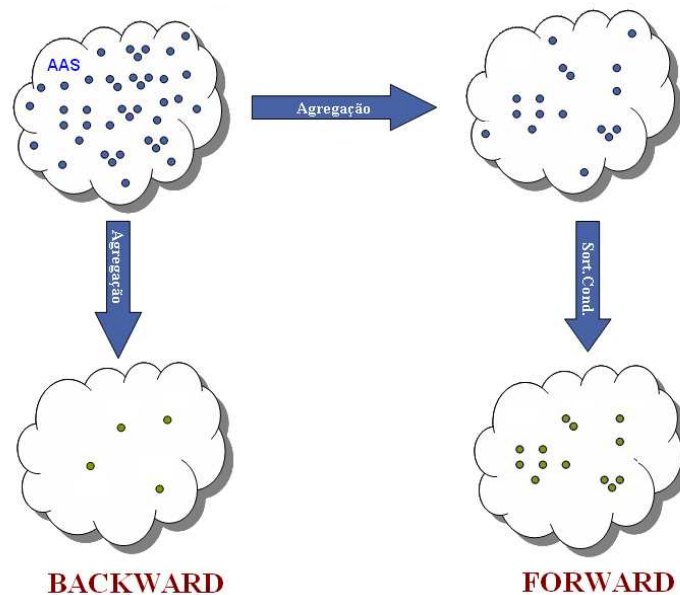


Figura 40: Opção 4

As técnicas da amostragem por hipercubo latino e amostragem descritiva podem ser aplicadas nas cinco alternativas apresentadas neste capítulo e nas duas opções atualmente disponíveis no modelo NEWAVE, descrita no item 3.3. No caso das opções que utilizam técnicas de agregação, os métodos de amostragem alternativos são usados para gerar a amostra de ruídos original que será agregada.

### 6.3. Construção dos 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 de Benders da função de custo futuro corresponde a uma média calculada para um conjunto de aflúências utilizadas durante o cálculo da política de operação no passo backward, Figura 41. No Apêndice A é apresentado um detalhamento do cálculo dos cortes de Benders realizado no modelo NEWAVE.

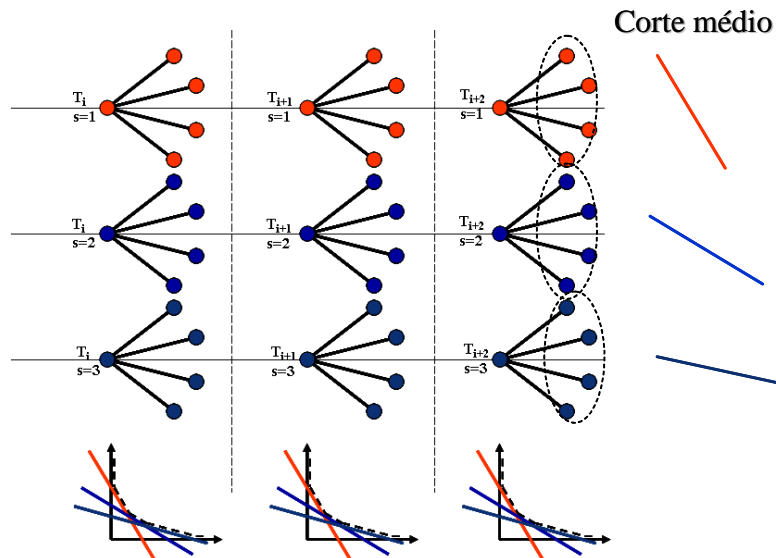


Figura 41: Construção da FCF – modelo NEWAVE

A aplicação das técnicas de 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.

O novo cálculo dos coeficientes dos cortes de Benders é mostrado nas equações 6.2 a 6.4.

$$\overline{\pi}_V^{isim} = \sum_{i=1}^{NLEQ} \pi_V^{i,isim} * P_i \quad isim = 1, \dots, NSIM \quad (6.2)$$

$$\overline{\pi}_{A_j}^{isim} = \sum_{i=1}^{NLEQ} \pi_{A_j}^{i,isim} * P_i \quad isim = 1, \dots, NSIM \quad (6.3)$$

j = 1, ..., NARP

$$\overline{RHS}^{isim} = \sum_{i=1}^{NLEQ} RHS^{i,isim} * P_i \quad isim = 1, \dots, NSIM \quad (6.4)$$

onde:

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

NARP: ordem do modelo PAR(p);

NSIM: número de cenários utilizados no passo forward;

P<sub>i</sub>: probabilidade do i-ésimo cenário hidrológico do conjunto backward;

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

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

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

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

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

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

#### 6.4. Reamostragem de Cenários (recombinação de ruídos)

Após a construção da sub-árvore de cenários com as alternativas descritas neste capítulo, pode-se aplicar o procedimento de reamostragem das sub-árvores a cada iteração do processo de convergência do cálculo da política ótima de operação. Com isto, pode-se percorrer uma porção maior da árvore completa de cenários. Atualmente, a sub-árvore utilizada pelo modelo NEWAVE é fixa durante todo o processo de convergência.

Neste trabalho, somente os ruídos da amostra forward são recombinados a cada iteração. No processo de reamostragem empregado neste trabalho, a cada iteração do processo de convergência, os ruídos da amostra forward, obtidos após o processo de agregação e do sorteio condicionado, são combinados de forma diferente ao longo dos estágios de maneira a gerar outros cenários de afluência. Esse procedimento é ilustrado no exemplo da Figura 42 até a Figura 45.

Neste exemplo tem-se uma árvore com cinco estágios e com duas aberturas por estágio, perfazendo um total de 16 cenários. Na Figura 42 é apresentada a árvore completa de cenários. Cada nó da árvore corresponde a um vetor de ruídos.

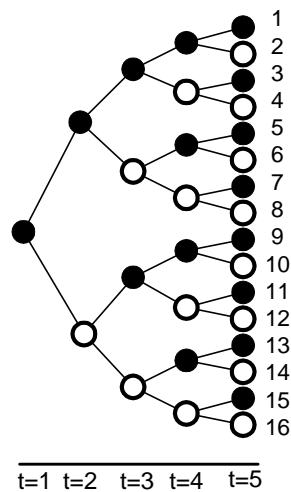


Figura 42: Árvore Completa – Exemplo

A sub-árvore a ser percorrida durante a primeira iteração do processo é apresentada na Figura 43. Essa sub-árvore pode ser obtida por amostragem aleatória simples ou pelas alternativas descritas neste capítulo. Note que a sub-árvore representa apenas uma pequena porção da árvore completa.

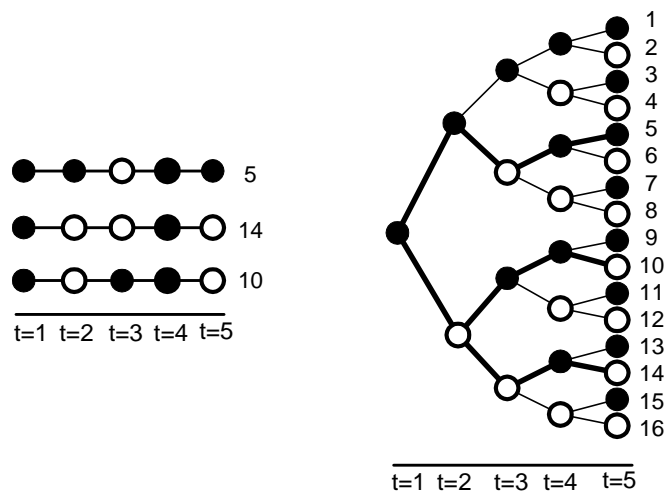


Figura 43: Sub-árvore (1ª amostra) – Exemplo

Para obter a sub-árvore a ser utilizada nas demais iterações do processo, é realizada uma combinação dos ruídos ao longo dos estágios. O conjunto de ruídos utilizado em um determinado estágio é o mesmo utilizado para a construção da sub-árvore da primeira iteração. A diferença é que a combinação dos ruídos entre os

estágios é diferente, Figura 44 (a) e (b), gerando desta forma outra sub-árvore de cenários.

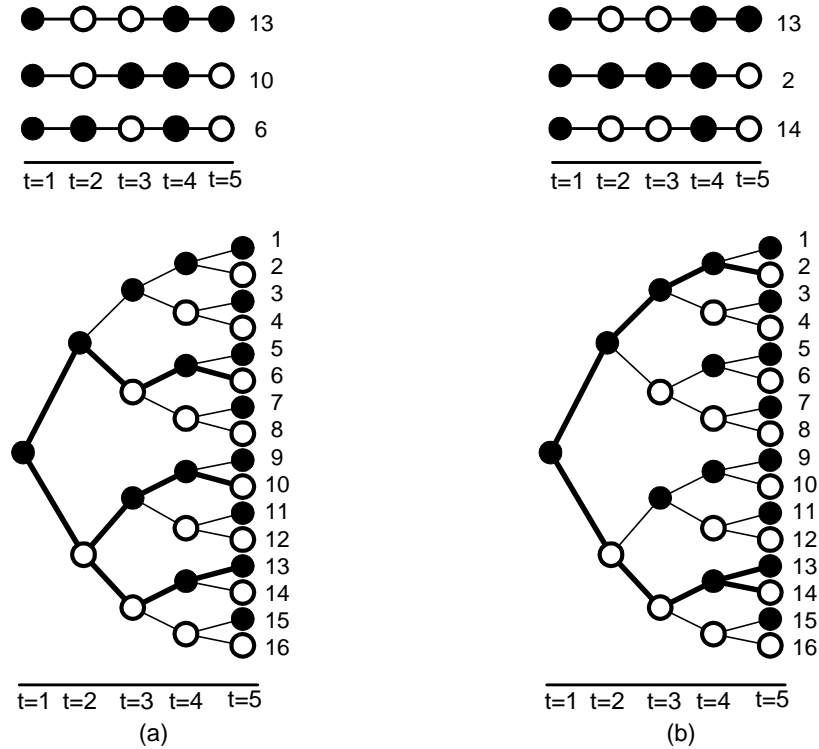


Figura 44: – Sub-árvore – Exemplo (a) 2º. amostra e (b) 3º. Amostra

Após a aplicação do processo de reamostragem, todos os cenários em destaque na Figura 45 terão sido percorridos. Ainda existe uma porção da árvore completa que não foi visitada durante o processo de otimização (cenários marcados de cinza), porém o número de cenários visitados foi bem maior do que aquele mostrado na Figura 42.

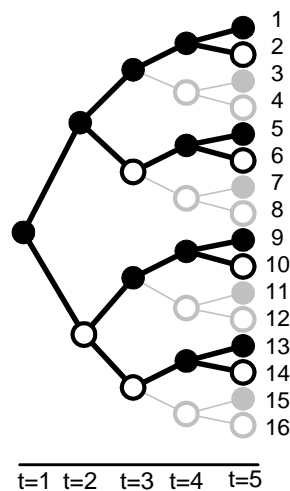


Figura 45: Sub-árvore após diversas reamostragens – Exemplo



Com a aplicação da técnica de reamostragem é esperado uma melhora na convergência do processo de solução do problema de otimização, uma vez que uma maior parcela da árvore completa será percorrida.

Normalmente, o critério de parada utilizado nos algoritmos que aplicam decomposição de Benders é a proximidade entre os valores de ZINF e ZSUP, limite inferior e superior do valor ótimo da função objetivo, respectivamente. O limite inferior (ZINF) cresce monotonicamente a cada iteração, o que não necessariamente ocorre com o limite superior (ZSUP).

Em (Geoffrion, 1972) é dito que o valor de ZSUP sempre é estabelecido como o melhor valor ótimo já obtido em qualquer iteração passada. Isto é, no caso de problemas de minimização deve-se então considerar o menor valor obtido para ZSUP. Em (Noonan & Giglio, 1977) é apresentado um problema de planejamento da expansão e neste trabalho o processo é dito convergido quando o limite inferior para o custo mínimo está dentro de uma tolerância pré-especificada em relação ao custo do melhor planejamento obtido até o momento. No trabalho de (Rotting & Gjelsvik, 1992) é utilizado o valor de ZUP obtido em cada iteração (sem atualização) e o critério de parada é quando ZINF e ZSUP estão suficientemente próximos, usando como tolerância para ZSUP o desvio padrão amostral dos ZSUP das iterações passadas.

No modelo NEWAVE é calculado um intervalo de confiança para a estimativa de ZSUP e quando o valor de ZINF estiver neste intervalo, o processo é dito concluído. O valor considerado para ZSUP é sempre o menor valor obtido ao longo das iterações.

Quando for aplicado o processo de reamostragem de cenários, o valor de ZSUP obtido a cada iteração será mantido, e o processo será dito convergido quando o valor de ZINF estiver completamente contido no intervalo de confiança calculado a partir de ZSUP. Outros critérios de parada podem ser propostos tais como a estabilidade no valor de ZINF e estabilidade em alguns resultados de operação (geração hidráulica, intercâmbios, etc).

## **6.5. Considerações Finais**

A aplicação das técnicas de agregação no modelo de geração de cenários sintéticos para o planejamento de médio prazo aumentou bastante o esforço computacional gasto neste processo, principalmente quando o tamanho da amostra

crece. No caso de uma amostra de vetores de ruído com tamanho de 100 mil objetos, o tempo gasto verificado na etapa de geração dos cenários hidrológicos foi de 6 horas em um microcomputador pentium com dois núcleos e 1GHz. O aumento no tempo computacional pode tornar impraticável o emprego desta metodologia no setor elétrico. Então foram implementadas técnicas de paralelização no processo de geração das amostras de ruídos. Como as amostras de ruídos não possuem dependência temporal, o processo de paralelização pode ser aplicado. O tempo gasto, considerando uma amostra de 100 mil vetores de ruídos, reduziu de 6 horas para aproximadamente 13 minutos, considerando 16 processadores em paralelo.

## **6.6. Resumo**

Neste capítulo foram apresentadas diversas propostas para a construção da sub-árvore a ser considerada durante a resolução do problema de planejamento da operação energética utilizando PDDE. Para definição da sub-árvore foram propostas aplicação de técnicas de agregação, reamostragem de cenários e adoção dos métodos de amostragem hipercubo latino e amostragem descritiva.

No próximo capítulo será feita uma avaliação dos cenários hidrológicos gerados considerando as propostas deste capítulo.



## 7

# Avaliação dos Cenários Hidrológicos Gerados

### 7.1.

#### Considerações Iniciais

O objetivo neste capítulo é avaliar a qualidade dos cenários hidrológicos gerados aplicando-se os métodos propostos neste trabalho. Primeiramente é avaliado o tamanho da amostra de ruídos utilizada no processo de agregação. Após a definição do tamanho desta amostra é analisada as amostras de ruídos geradas para construção dos cenários hidrológicos utilizados nos passos forward e backward. São realizados testes de aderência, análise das estimativas de média, desvio padrão e correlação espacial. Finalmente, são avaliados os cenários hidrológicos. São realizados testes estatísticos condicionados, para o caso dos cenários do passo backward, testes estatísticos não condicionados e testes de seqüência negativa para os cenários do passo forward.

Para verificar a acurácia da geração da árvore de cenários é utilizado o Programa Mensal de Operação de janeiro de 2008. A configuração hidroelétrica utilizada neste estudo é composta por 145 usinas hidroelétricas. Estas usinas são agregadas em quatro subsistemas equivalentes de energia (Sudeste/Centro-Oeste, Sul, Nordeste e Norte). Além disso, fazem parte do parque gerador do SIN 112 usinas termoelétricas. O horizonte de planejamento é de cinco anos divididos em estágios mensais e mais um período de pós-estudo com cinco anos, totalizando 120 períodos.

Como o modelo estatístico utilizado para geração dos cenários sintéticos de energia é o PAR(p), a energia natural afluyente de um período depende das p afluências passadas. Atualmente a ordem máxima considerada para o modelo é igual a seis ( $p \leq 6$ ). A tendência hidrológica recente adotada neste estudo é mostrada na Figura 46 em percentual da MLT (média de longo termo mensal). Nota-se que a situação hidrológica recente não é favorável para os subsistemas Nordeste e Norte (abaixo de 50% MLT), está um pouco abaixo da média para o subsistema Sudeste e é favorável para o subsistema Sul. Desta forma, a média dos cenários gerados para os primeiros meses do horizonte de estudo devem estar abaixo da MLT para os

subsistemas Nordeste e Norte, acima da média histórica para o subsistema Sul e um pouco abaixo para o subsistema Sudeste. O efeito condicionante ao passado recente vai se atenuando ao longo dos períodos e praticamente não é mais verificado após os primeiros meses.

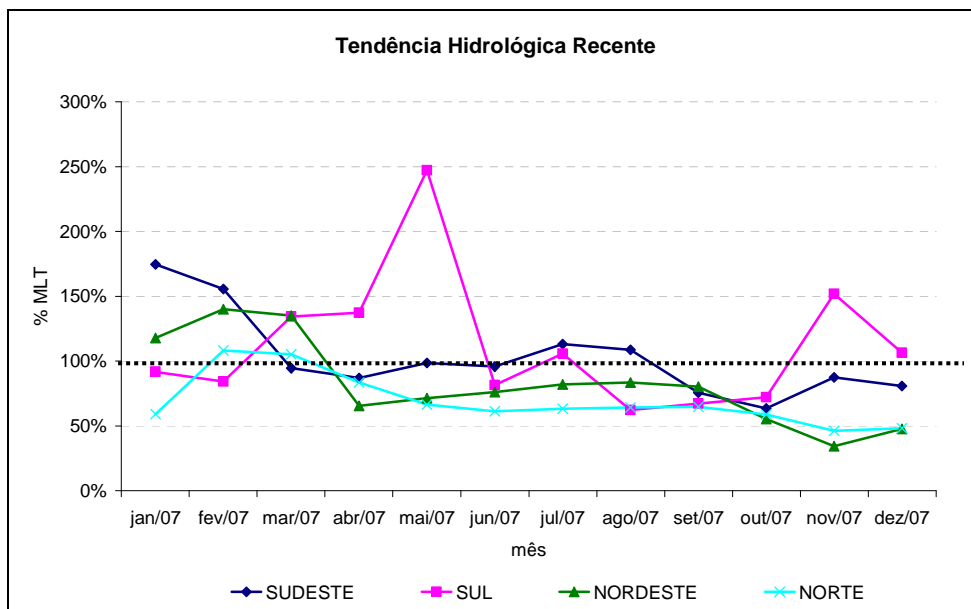


Figura 46: Tendência Hidrológica Recente (%MLT)

## 7.2. Análise da amostra de ruídos

### 7.2.1. Amostra de ruídos utilizada no processo de agregação

A amostra considerada no processo de agregação é formada por um número pré-definido de vetores de ruídos, onde cada posição do vetor representa um ruído sorteado da distribuição normal padrão para um dos subsistemas considerados na representação do SIN. Vale lembrar que neste ponto os vetores de ruídos não apresentam correlação espacial. Neste item são avaliadas amostras com tamanho variando de 2 mil a 100 mil vetores de ruídos, totalizando um total de 9 casos.

Para cada período do horizonte de estudo é gerada uma amostra de ruídos e calculados a média, o desvio-padrão para cada subsistema e a distribuição de freqüência multivariada. No final, são calculadas a média e o intervalo de confiança para as estimativas da média e do desvio-padrão. A distribuição multivariada estimada é comparada com a distribuição multivariada teórica.

A distribuição da estimativa da média é dada por (7.1), onde os parâmetros  $\mu$  e  $\sigma^2$  são iguais a zero e um, respectivamente. O valor de  $n$  é igual ao número de períodos vezes o tamanho da amostra. Na Figura 47 são apresentados a estimativa para média (gráfico de barras) e o intervalo de confiança (gráfico de linha) considerando um nível de significância de 5%. O resultado obtido é satisfatório para cada caso analisado (exceto para os casos com amostra de 5 mil no subsistema Sul e com amostra 75 mil para o subsistema Norte).

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad (7.1)$$

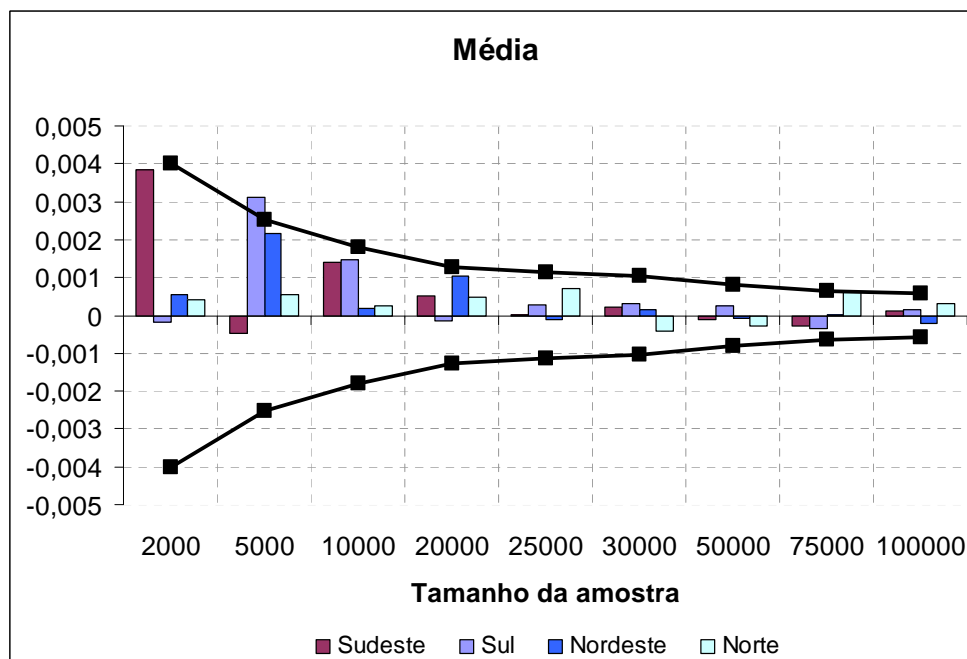


Figura 47: Estimativa Média (amostra de ruídos)

Na Figura 48 é mostrado o desvio-padrão médio para cada subsistema (gráfico de barras) e um intervalo de confiança (gráfico de linha), considerando um nível de significância de 5%, para esta estimativa. A distribuição do desvio-padrão, no caso de variáveis com distribuição normal padrão, é dada por (7.2). Novamente, pode-se observar que os resultados são satisfatórios para todos os casos analisados, com exceção da amostra de tamanho 5 mil, para o subsistema Nordeste.

$$(n - 1) \frac{S^2}{\sigma^2} \sim \chi_{n-1}^2 \quad (7.2)$$

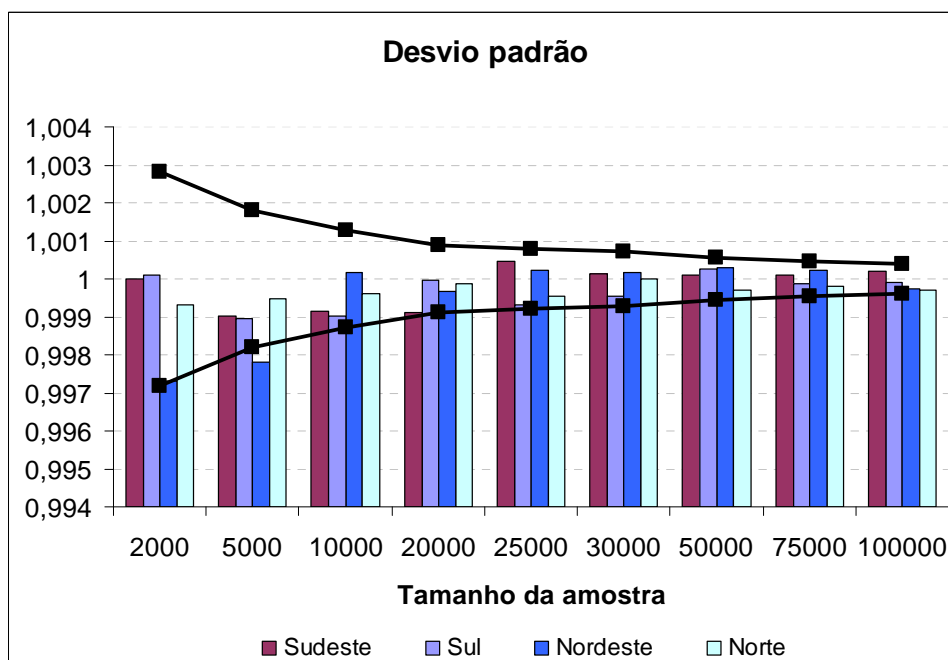


Figura 48: Estimativa Desvio-Padrão (amostra de ruídos)

Na Figura 49 é apresentado o número de rejeições para os testes estatísticos da média e desvio-padrão. Considerando para cada caso analisado um conjunto de 480 testes (120 períodos x 4 subsistemas) e supondo um nível de significância de 5%, é admissível que 24 testes sejam reprovados. Este limite é ultrapassado nos casos com amostras de tamanho 5 mil, 10mil e 75 mil. O número de rejeições do caso 5 mil é compatível com um nível de significância de aproximadamente 7%.

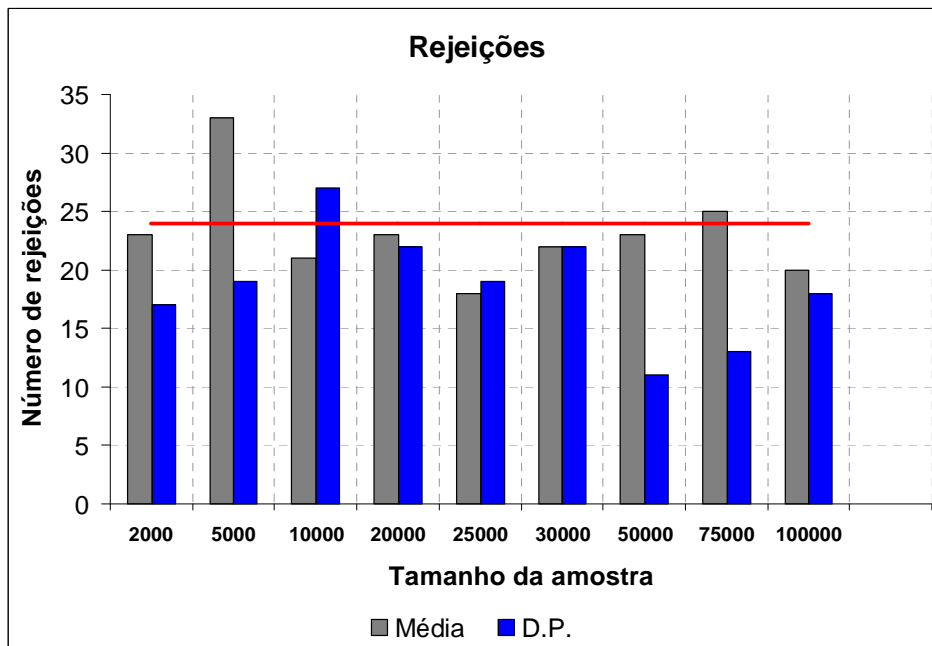


Figura 49: Número de Rejeições (amostra de ruídos)

A distribuição multivariada é calculada dividindo-se a distribuição univariada normal padrão de cada subsistema em três intervalos e então é feita a combinação dos intervalos de cada um dos quatro subsistemas. Desta forma, são definidas 81 classes ( $3^4 = 81$  classes). Para este estudo, a distribuição univariada é dividida em 3 intervalos com probabilidades iguais a 0,25; 0,50 e 0,25 (quantis da distribuição), respectivamente, de acordo com a Figura 50.

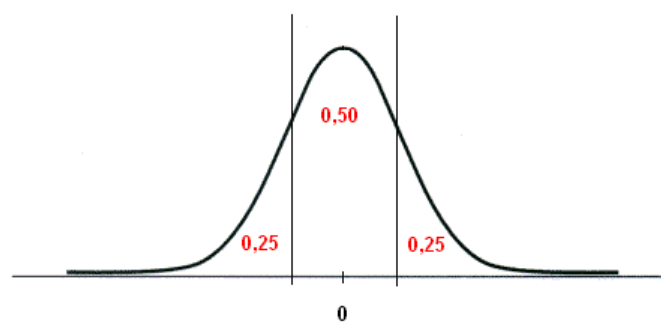


Figura 50: Intervalos da Distribuição Univariada

Na Figura 51 é ilustrada a classe 34, isto é, o ruído dos subsistemas Sudeste e Norte devem estar contidos no segundo intervalo enquanto os ruídos do Sul e do Nordeste devem pertencer ao terceiro e primeiro intervalos, respectivamente. A probabilidade teórica da classe 34 é igual a



$$P(-0,6745 < SE < 0,6745) \quad * \quad P(0,6745 < SE < \infty) \quad * \quad P(-\infty < NE < -0,6745) \quad *$$

$$* P(-0,6745 < N < 0,6745) = 0,5 * 0,25 * 0,25 * 0,5 = 0,015625$$

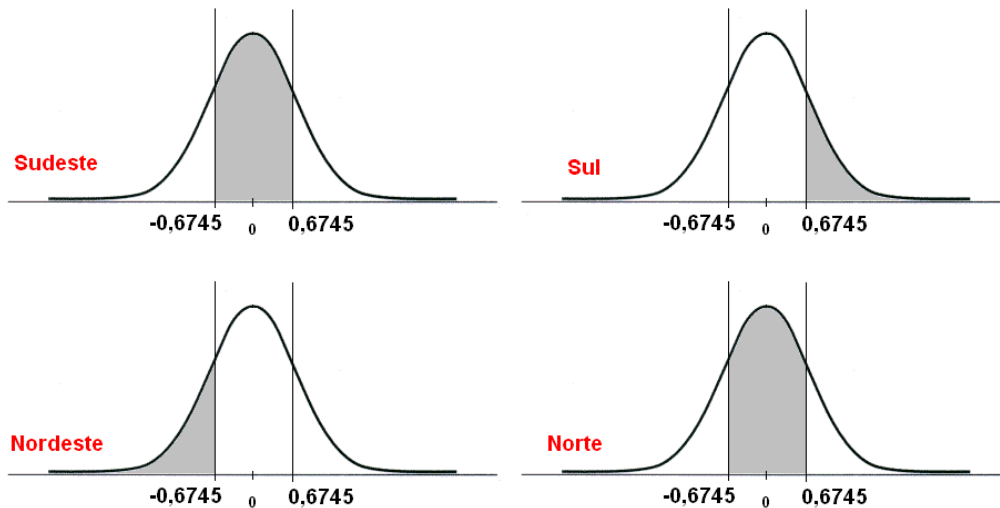
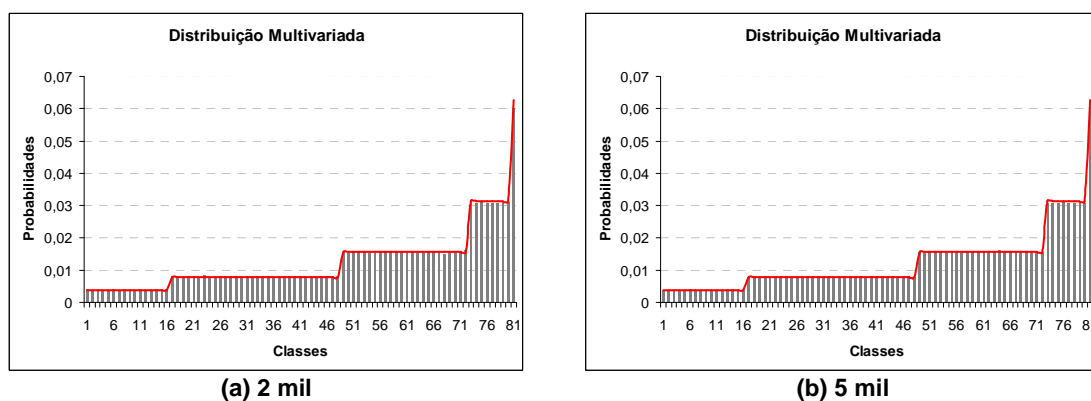
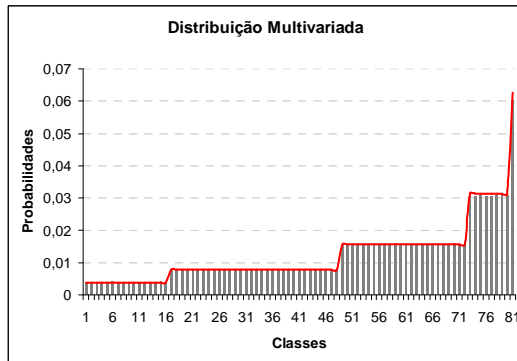


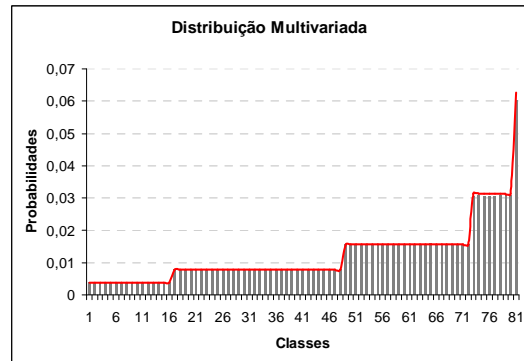
Figura 51: Exemplificação Classe 34

Para facilitar a visualização da distribuição multivariada, as classes são ordenadas de acordo com sua probabilidade teórica. Da Figura 52a até a Figura 52i são ilustradas as distribuições multivariadas médias para os casos considerando amostras de tamanho 2 mil a 100 mil. A linha contínua indica a probabilidade teórica da classe.

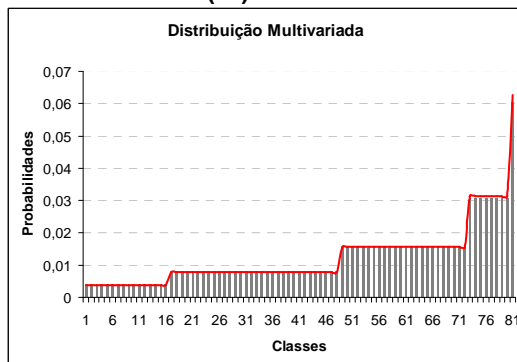




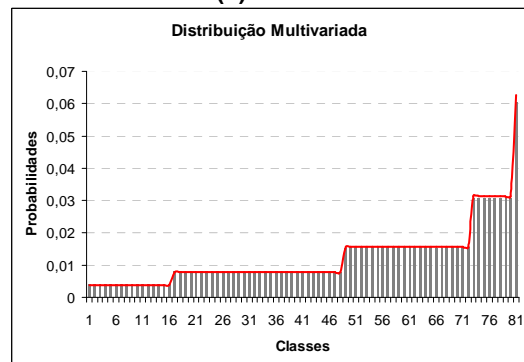
**(ac) 10 mil**



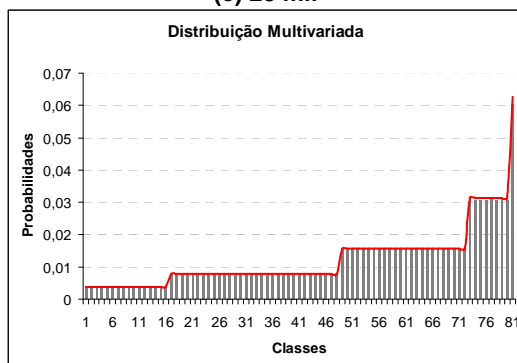
**(d) 20 mil**



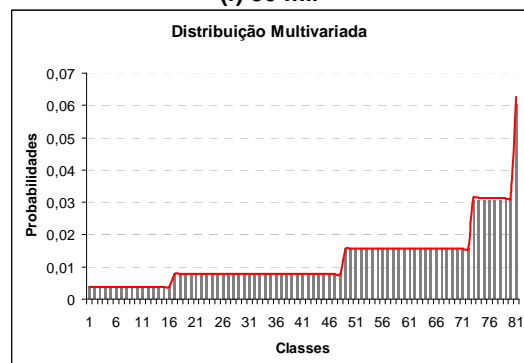
**(e) 25 mil**



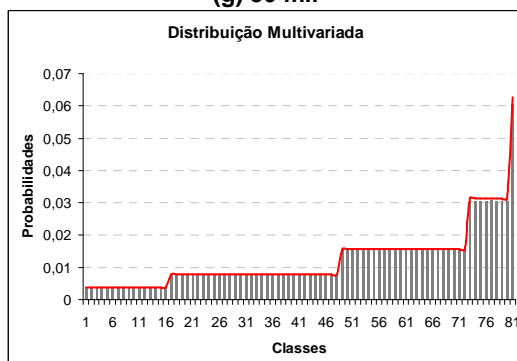
**(f) 30 mil**



**(g) 50 mil**



**(h) 75 mil**



**(i) 100 mil**

Figura 52: Distribuição Multivariada – Amostra com Tamanho 2 mil a 100 mil

Pode-se observar que a distribuição multivariada teórica é preservada satisfatoriamente independente do caso analisado.

O tempo computacional gasto durante o processo de agregação aumenta consideravelmente à medida que o número de elementos da amostra de ruídos cresce. Por exemplo, para gerar e agregar 120 amostras com tamanho 2 mil são necessários, em um micro computador Pentium duo-core 1GHz, 2 minutos. Para gerar e agregar 120 amostras com tamanho 100 mil são necessárias aproximadamente 4 horas.

De acordo com o exposto neste item, doravante é considerada a amostra de tamanho 2 mil para as demais análises realizadas neste capítulo.

### 7.2.2.

#### **Amostra de ruídos utilizada na geração dos cenários do passo forward**

Neste item são avaliadas as amostras de ruídos que são utilizadas para gerar os cenários hidrológicos do passo forward. São apresentados testes estatísticos para média e desvio-padrão, bem como testes de aderência e a distribuição multivariada de freqüências. São testadas todas as opções de construção de árvore propostas neste trabalho, além das duas opções existentes atualmente no modelo NEWAVE. Também é testada a amostra forward resultante do processo de agregação e que dá origem à amostra forward utilizada nas opções 2 a 4.

Nesta análise é considerada uma amostra de ruídos para passo forward com tamanho igual a 200 e para passo backward igual a 20. Lembre-se que nas opções Atual e 0 a amostra de ruídos forward é proveniente de uma amostra de ruídos backward. A amostra inicial utilizada no processo de agregação é formada por 2000 vetores de ruídos.

Neste trabalho, é definido como estatística t o desvio relativo de uma estatística  $\hat{\theta}$  (média, desvio padrão etc), isto é, a diferença entre a estimativa e o seu valor esperado dividido pelo desvio padrão da estimativa, conforme (7.3):

$$\frac{\hat{\theta} - \mu_{\theta}}{\sigma_{\theta}} \sim t_n \quad (7.3)$$

As estatísticas t para as estimativas da média e do desvio-padrão são apresentadas na Figura 53 e na Figura 54, respectivamente. As linhas tracejadas são os limites relativos a um nível de significância de 5%. Para cada opção analisada são considerados 480 valores (120 períodos x 4 subsistemas).

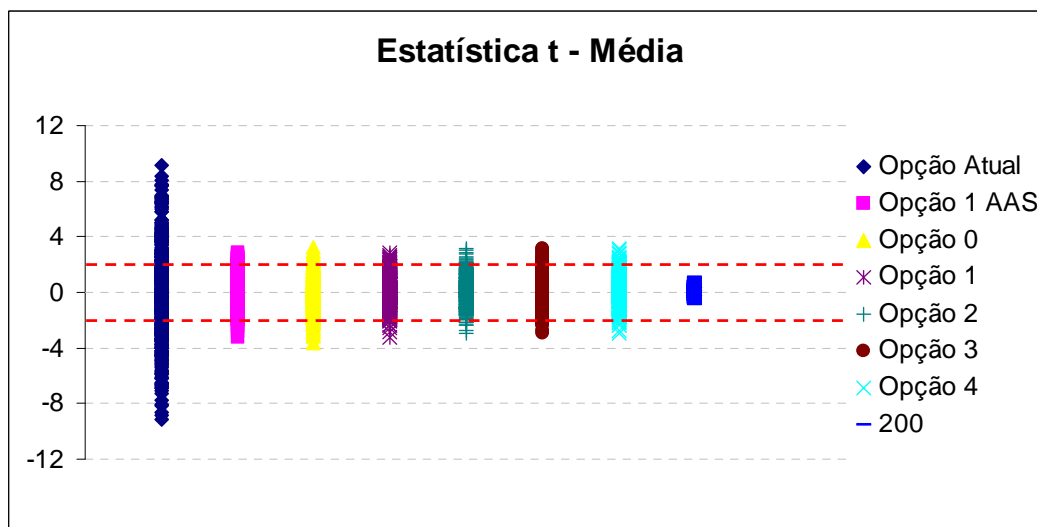


Figura 53: Estatística t – Média – Amostra Forward (ruído)

Note que a opção atual é aquela que apresenta maior dispersão com relação aos valores de população (zero no caso da média e um no caso do desvio-padrão). Isto ocorre, pois a amostra do passo forward é sorteada de uma amostra backward de tamanho reduzido, neste caso igual a 20, obtida por amostragem aleatória simples. Note que a adoção do processo de agregação para obtenção da amostra backward (opção 0) já é suficiente para reduzir essa dispersão. Todavia, pode-se notar uma tendência em gerar desvios-padrão menores do que o valor de população.

A degradação na representação do desvio-padrão ocorre, porém em menor escala, com todas as opções que utilizam o processo de agregação para obtenção, direta ou indireta, da amostra forward, conforme pode ser observado na Figura 54.

As amostras forward das opções 2 a 4 são idênticas e são obtidas a partir de um sorteio condicionado da amostra 200 (última opção apresentada na Figura 53 e na Figura 54). Como esperado, há um aumento na dispersão das estimativas, uma vez que a amostra resultante do sorteio condicionado possui menos informação.

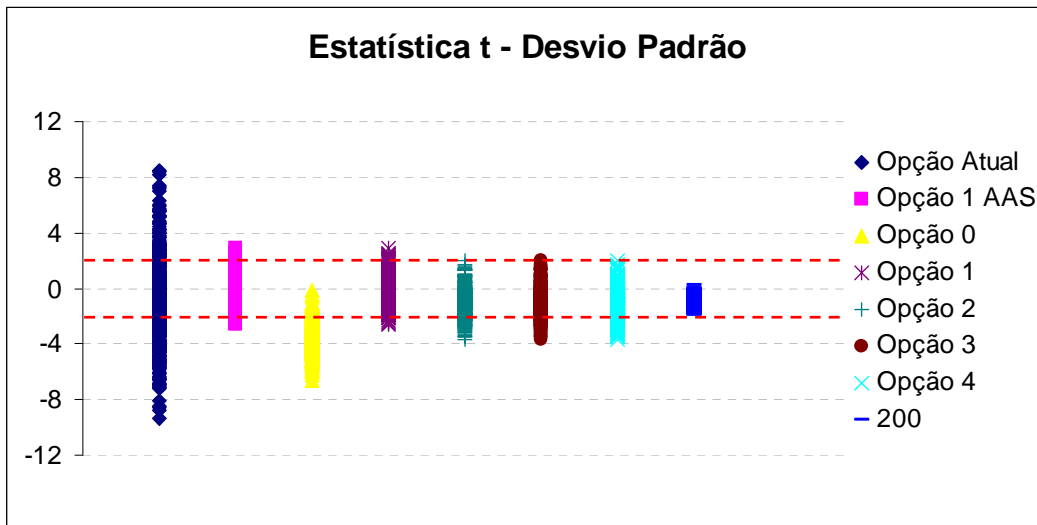
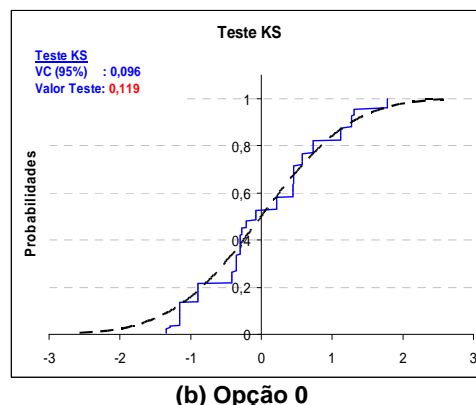
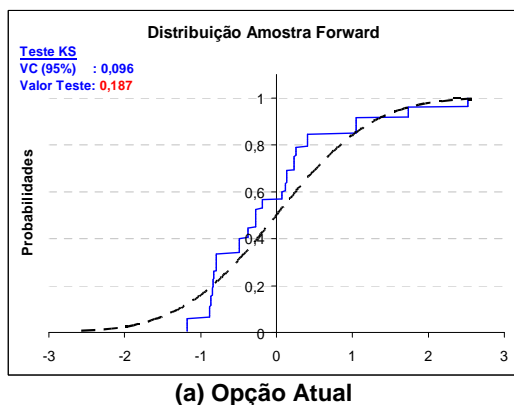
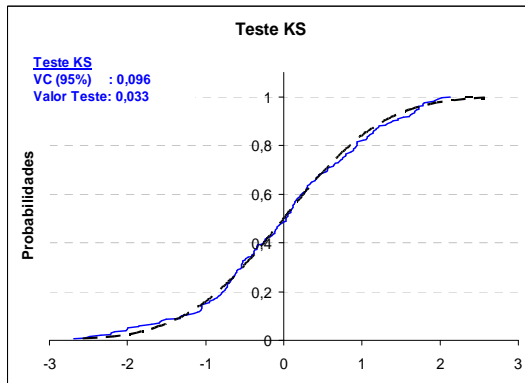


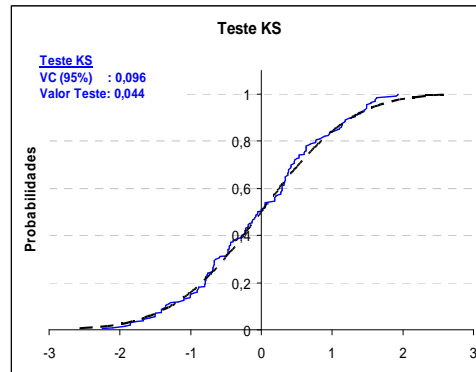
Figura 54: Estatística t – Desvio-padrão – Amostra Forward (ruído)

O teste de aderência Kolmogorov-Smirnov (Stephens, 1974) é apresentado para a distribuição univariada somente do subsistema Sudeste. Os resultados obtidos para os demais subsistemas são semelhantes. Da Figura 55a até a Figura 55e estão ilustradas as distribuições acumuladas da normal padrão (linha pontilhada) e da amostra forward (linha contínua) obtida considerando as opções em análise. O nível de significância adotado para o teste é de 5% e nos gráficos são apresentados o valor crítico e o valor obtido do teste.

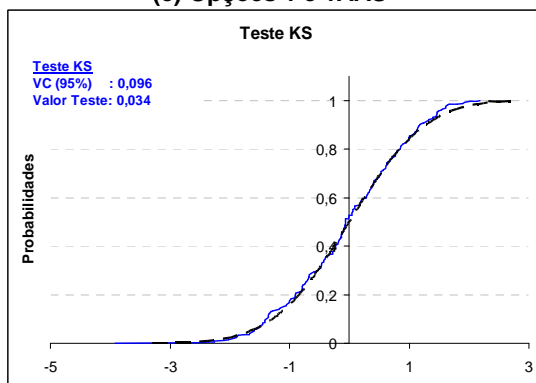




(c) Opções 1 e 1AAS



(d) Opções 2 a 4

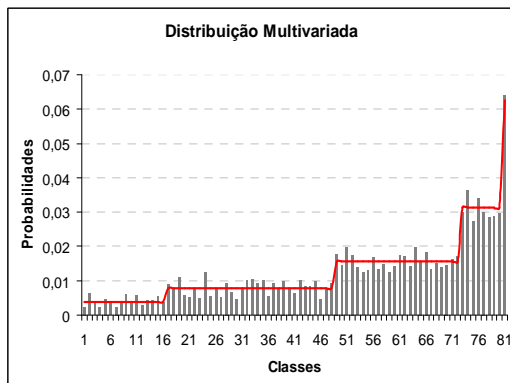


(e) 200

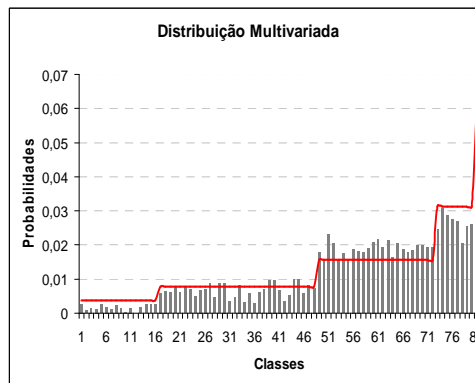
Figura 55: Teste de Aderência – Amostra Forward (ruído)

Pode-se observar que as opções Atual e 0 degradam excessivamente a distribuição da amostra forward. Nas demais opções, a distribuição normal padrão univariada é bem preservada.

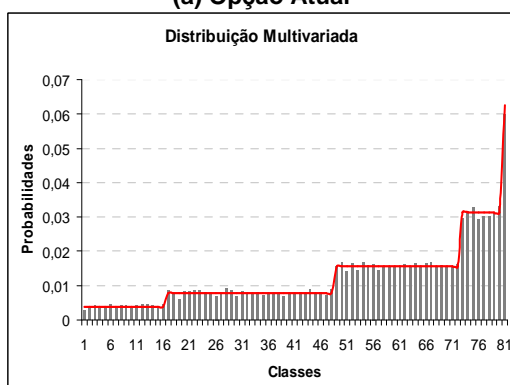
Complementando, é feita uma análise da distribuição multivariada seguindo o mesmo procedimento utilizado anteriormente. Da Figura 56a até a Figura 56e são apresentadas as distribuições multivariadas para todas as opções. Novamente, nota-se uma pior representação desta distribuição no caso das opções Atual e 0, com relação às demais opções.



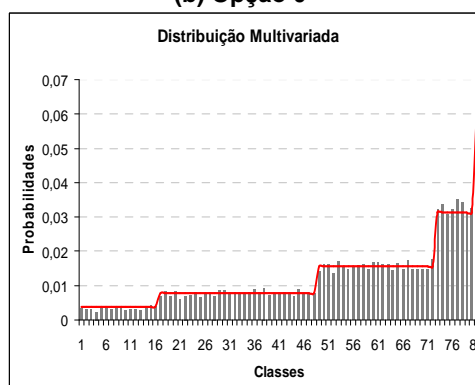
(a) Opção Atual



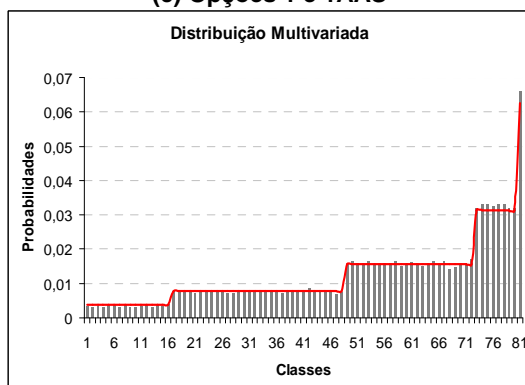
(b) Opção 0



(c) Opções 1 e 1AAS



(d) Opções 2 a 4



(e) 200

Figura 56: Distribuição Multivariada – Amostra Forward (ruído)

### 7.2.3.

#### Amostra de ruídos utilizada na geração dos cenários do passo backward

Para avaliar as amostras de ruídos que são utilizadas para gerar os cenários hidrológicos do passo backward são apresentados testes estatísticos para média e desvio-padrão, bem como testes de aderência para a distribuição univariada. A distribuição multivariada de freqüências também é avaliada. São testadas todas as opções de construção de árvore propostas neste trabalho, além das duas opções existentes atualmente no modelo NEWAVE. Vale ressaltar que o procedimento para

obtenção da amostra backward é o mesmo para as opções Atual e 0, logo as amostras são iguais. Nas opções 1, 0 e 4 as amostras obtidas são idênticas. Para todas as opções é analisada uma amostra de ruído backward com tamanho igual a 20.

Na Figura 57 e na Figura 58 são apresentadas as estatísticas t para a estimativa da média e do desvio-padrão. O intervalo de confiança representado pelas linhas tracejadas é calculado considerando-se um nível de significância de 5%. As opções que utilizam o processo de agregação apresentam uma dispersão menor com relação aos valores da população.

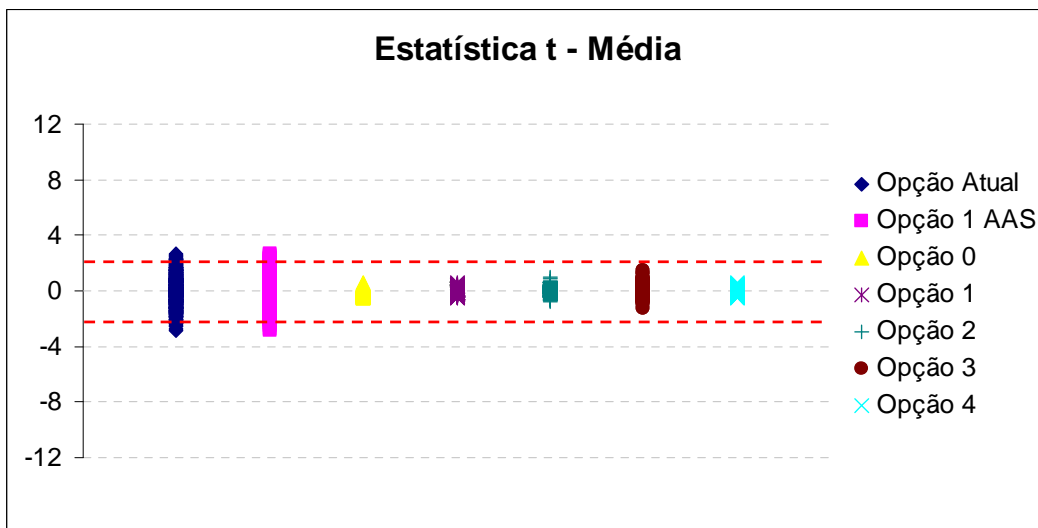


Figura 57: Estatística t – Desvio-padrãoMédia – Amostra Backward (ruído)

Como já observado para as amostras de ruídos forward, aqui também há uma pequena degradação no desvio-padrão quando é aplicado o processo de agregação. Há uma tendência em gerar desvios menores do que o valor desejado (igual a 1).



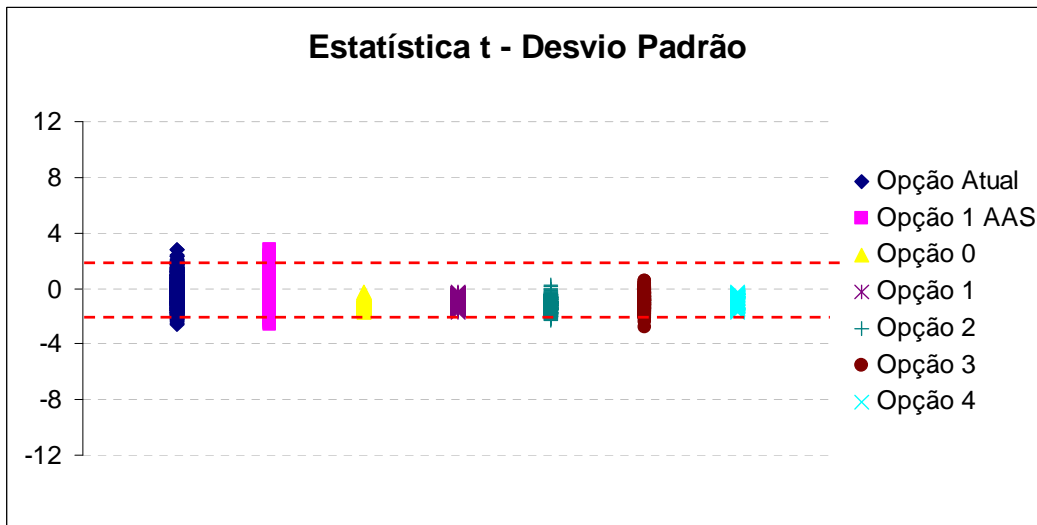
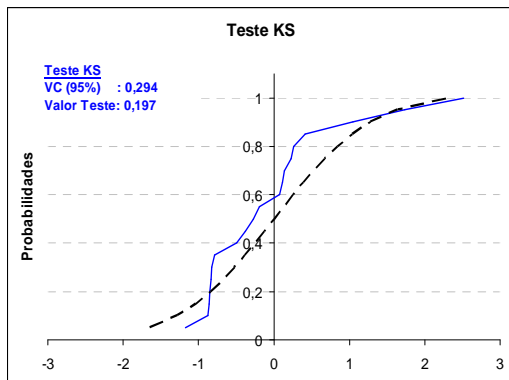
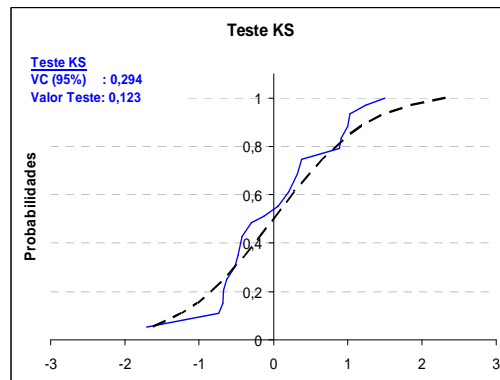


Figura 58: Estatística t – Desvio-padrão – Amostra Backward (ruído)

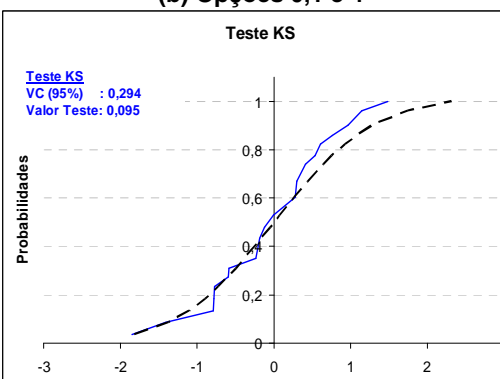
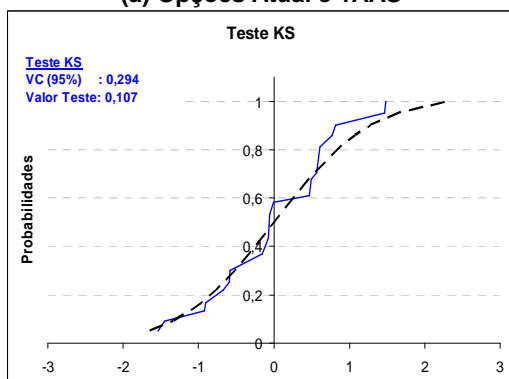
A seguir é utilizado o teste de Kolmogorov-Smirnov para medir a aderência entre a distribuição univariada da amostra backward e a distribuição normal padrão. Da Figura 59a até a Figura 59d são apresentadas as distribuições acumuladas de freqüência para o subsistema Sudeste e os resultados do teste de aderência. Resultados semelhantes foram obtidos para os demais subsistemas.



(a) Opções Atual e 1AAS



(b) Opções 0,1 e 4



**(c) Opção 2**

**(d) Opção 3**

Figura 59: Teste de Aderência – Amostra Backward (ruído)

Conforme pode ser observado na Figura 59, todas as opções analisadas passam no teste de aderência, porém pode-se verificar visualmente uma melhor aderência entre as distribuições nas opções que utilizam processo de agregação direta ou indiretamente para gerar a amostra backward.

Da Figura 60a até a Figura 60d são apresentadas as distribuições multivariadas, calculadas considerando o procedimento de cálculo já descrito nos itens anteriores.

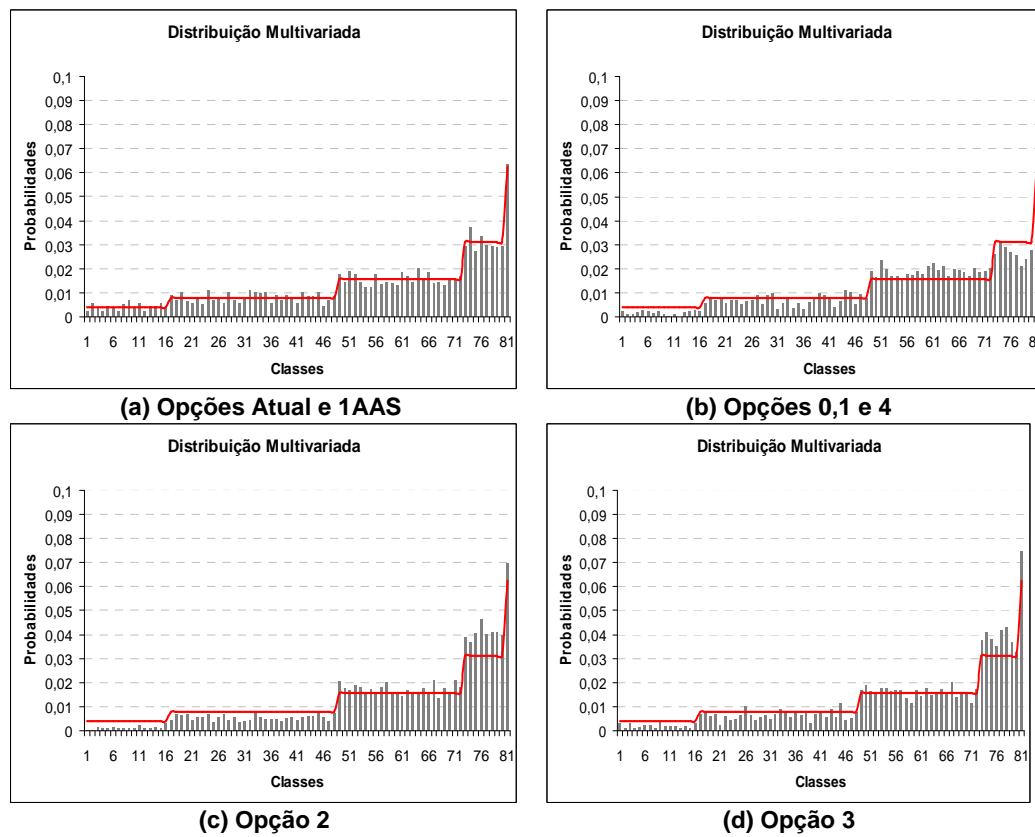


Figura 60: Distribuição Multivariada – Amostra Backward (ruído)

Neste caso, as opções que preservam melhor a distribuição multivariada são as opções Atual e 1 AAS, que utilizam a amostragem aleatória simples para gerar a amostra de ruídos da backward. Contudo, à medida que o tamanho da amostra backward cresce, as distribuições multivariadas de todas as opções tendem ao valor teórico. Essas avaliações são apresentadas no próximo item.

## 7.2.4.

### Variando o tamanho da amostra de ruídos forward e backward

Nos itens anteriores são analisadas as amostras de ruídos forward e backward com tamanho fixo, iguais a 200 e 20, respectivamente. Neste item, são analisadas amostras com diferentes tamanhos a fim de verificar a robustez das opções propostas.

#### 7.2.4.1.

### Variando o tamanho da amostra forward

São testadas variações no tamanho da amostra forward para as opções Atual, 0 e 4. A amostra forward para opção 1 AAS é idêntica à amostra da opção Atual e as amostras forward das opções 2 e 3 são idênticas à amostra da opção 4. São apresentadas avaliações somente para a amostra forward, uma vez que a geração da amostra backward não é influenciada pela amostra forward nas opções analisadas neste item.

Na Figura 61 e na Figura 62 são apresentadas as estatísticas t para as estimativas da média e desvio padrão das amostras de ruídos forward com tamanho igual a 200, 250 e 300. Mesmo aumentando o tamanho da amostra forward, os resultados obtidos para a opção Atual são bastante dispersos e a tendência observada na estimativa do desvio-padrão para opção 0 não se altera. Já a pequena degradação no desvio-padrão verificada na opção 4 se reduz à medida que o tamanho da amostra aumenta.

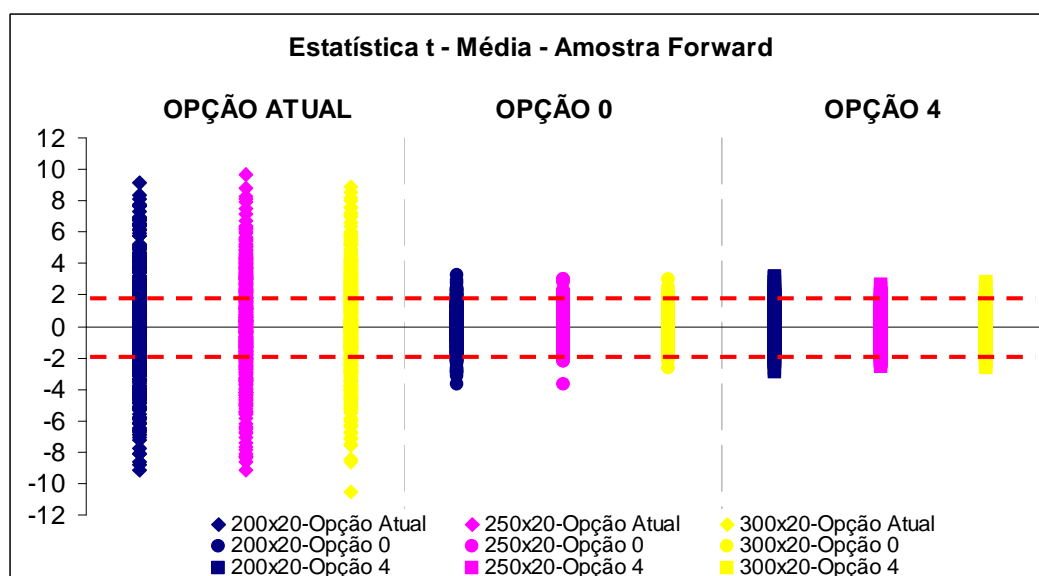


Figura 61: Estatística t – Média – Amostra Forward (ruído) – Variando Tamanho Amostra Forward

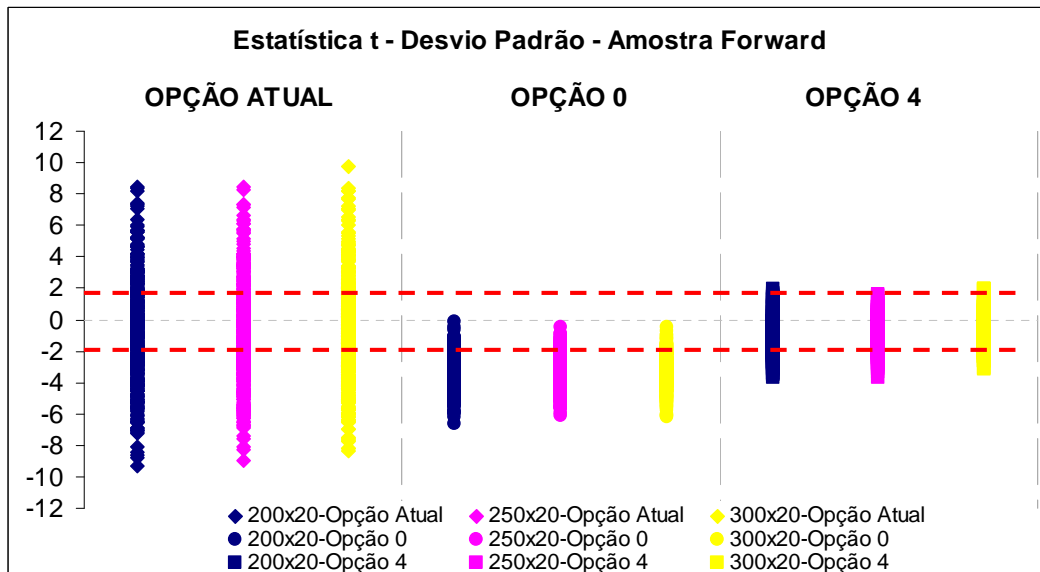


Figura 62: – Estatística t – Desvio-Padrão – Amostra Forward (ruído) – Variando Tamanho Amostra Forward

Na Figura 63 e na Figura 64 são apresentadas as distribuição univariada para o subsistema Sudeste, o resultado do teste de aderência e a distribuição multivariada para opção Atual, considerando amostras forward com tamanho 250 e 300, respectivamente. Conforme o esperado, pode-se observar que o aumento do tamanho da amostra forward não traz melhoria na representatividade desta amostra, isto ocorre, pois a amostra backward, da qual é sorteada a amostra forward, permanece inalterada.

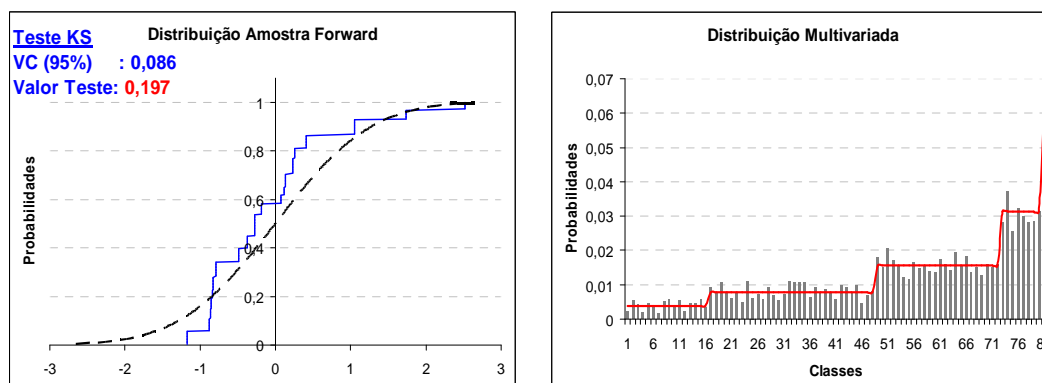


Figura 63: Distribuição – Amostra Forward (ruído) – 250x20 – Opção Atual

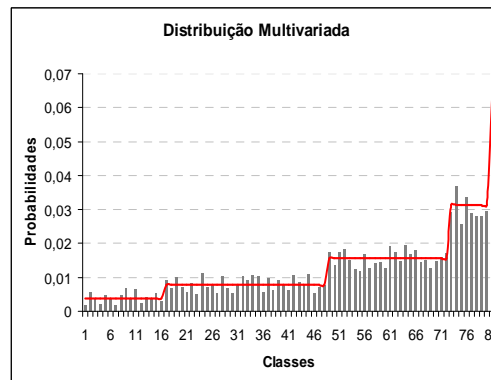
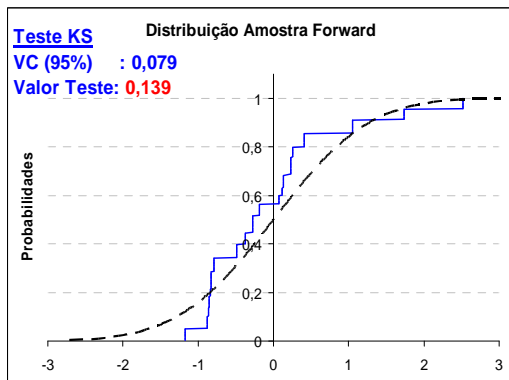


Figura 64: Distribuição – Amostra Forward (ruído) – 300x20 – Opção Atual

O mesmo pode ser observado na Figura 65 e na Figura 66, onde são apresentados os mesmos resultados, porém considerando-se a opção 0.

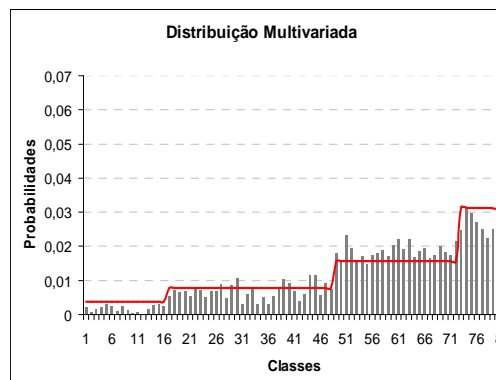
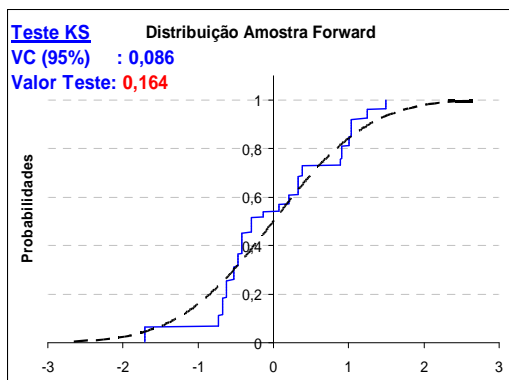


Figura 65: Distribuição – Amostra Forward (ruído) – 250x20 – Opção 0

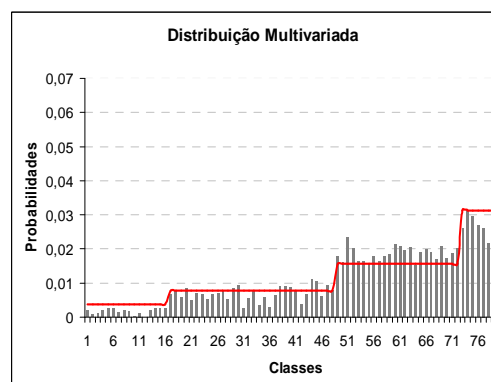
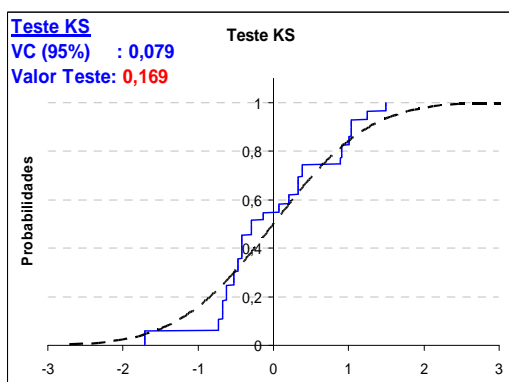


Figura 66: Distribuição – Amostra Forward (ruído) – 300x20 – Opção 0

Na Figura 67 e na Figura 70 são apresentadas as distribuição univariada para o subsistema Sudeste, o resultado do teste de aderência e a distribuição multivariada para opção 4, considerando amostras forward com tamanho 250 e 300,

respectivamente. Observa-se o aumento do tamanho da amostra forward traz melhoria na representatividade desta amostra.

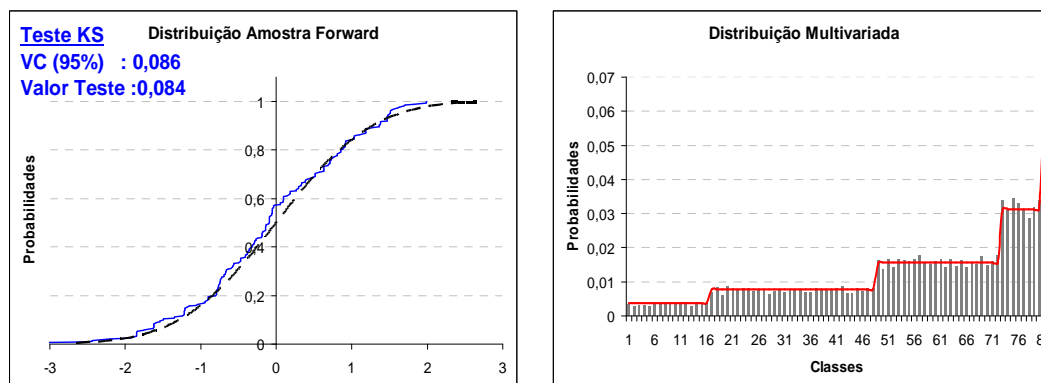


Figura 67: Distribuição – Amostra Forward (ruído) – 250x20 – Opção 4

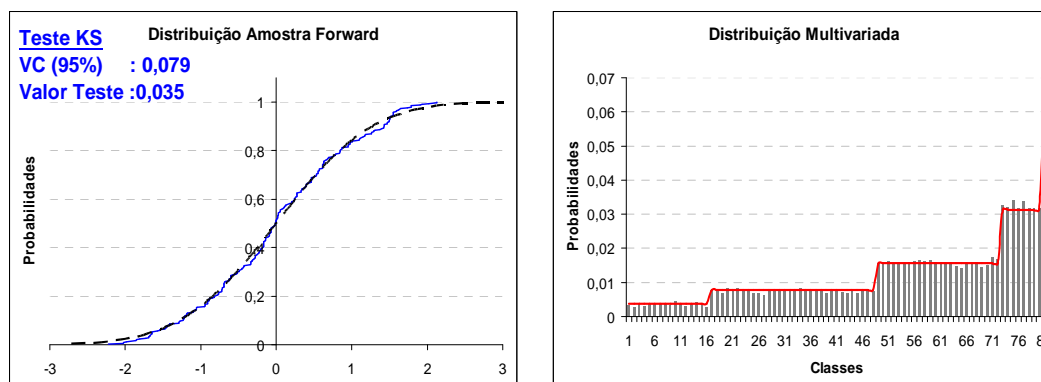


Figura 68: Distribuição – Amostra Forward (ruído) – 300x20 – Opção 4

#### 7.2.4.2.

#### Variando o tamanho da amostra backward

Agora são avaliadas variações no tamanho da amostra backward para as opções Atual, 0 e 4. Como a geração da amostra backward influencia na geração da amostra forward, no caso das opções Atual e 0, as análises são realizadas tanto para as amostras backward quanto forward. Vale lembrar que para esta análise o tamanho da amostra forward permanece fixo e igual a 200.

Na Figura 69 e na Figura 70 são apresentadas as estatísticas t para as estimativas da média e desvio-padrão para amostras backward com tamanho iguais a 20, 50 100. A amostra backward das opções 0 e 4 são idênticas. Observe que para

todas as opções há uma redução na dispersão das estimativas à medida que o tamanho da amostra bakward aumenta.

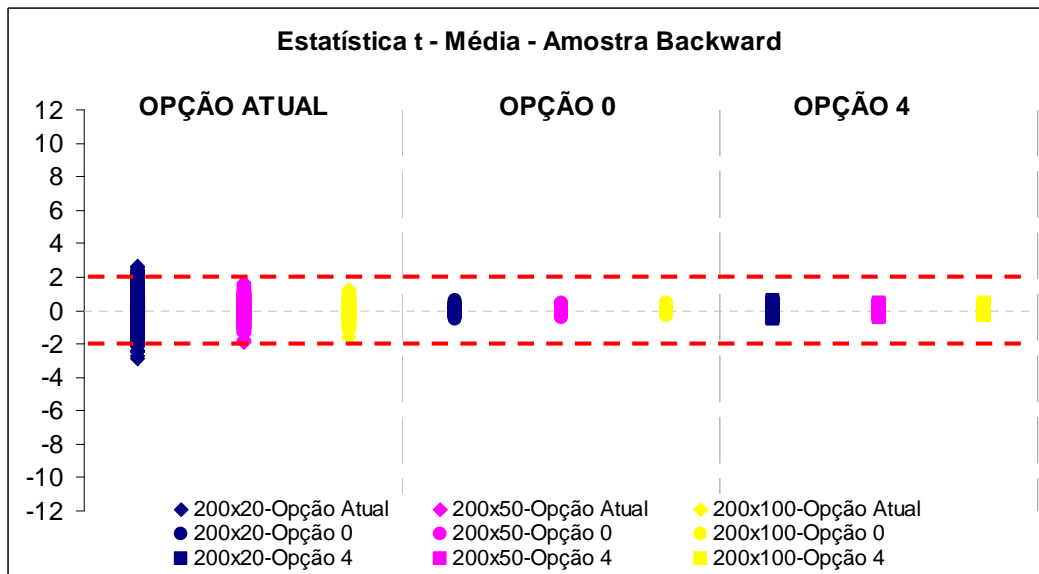


Figura 69: Estatística t –Média – Amostra Backward (ruído) – Variando Tamanho Amostra Backward

Esse ganho é bastante significativo com relação ao desvio-padrão, note que a degradação com relação ao valor de população se reduz para as opções que utilizam o processo de agregação.

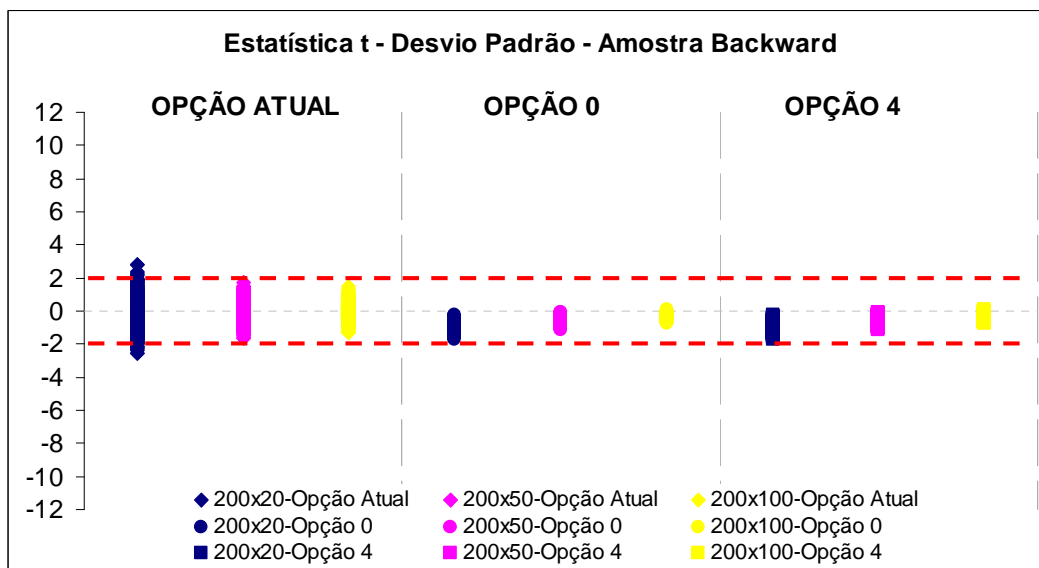


Figura 70: Estatística t –Desvio-Padrão – Amostra Backward (ruído) – Variando Tamanho Amostra Backward

Na Figura 71 e na Figura 72 são apresentados resultados das mesmas análises, porém considerando as amostras de ruído forward. Para a opção Atual há uma redução na dispersão das estimativas uma vez que a amostra backward que dá origem à amostra forward fica mais robusta. Com relação à opção 4, conforme o esperado, não se verifica nenhum ganho ao aumentar o tamanho da amostra backward, pois a amostra forward é obtida de forma independente.

A estimativa média praticamente não se altera para opção 0, indicando que uma amostra com 20 valores obtidos com agregação fornece a mesma representatividade que uma amostra com 100 valores também obtida pelo processo de agregação. Porém, há ganho na representação da estimativa do desvio-padrão, o que era esperado pois quanto maior a amostra resultante de um mesmo processo de agregação, menor é a degradação do desvio-padrão.

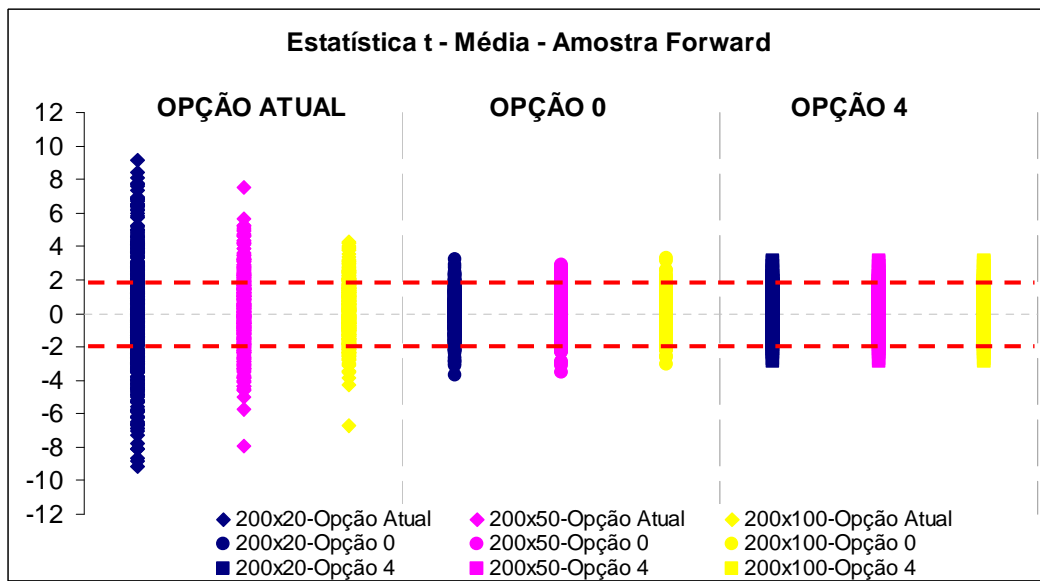


Figura 71: Estatística t –Média – Amostra Forward (ruído) – Variando Tamanho Amostra Backward



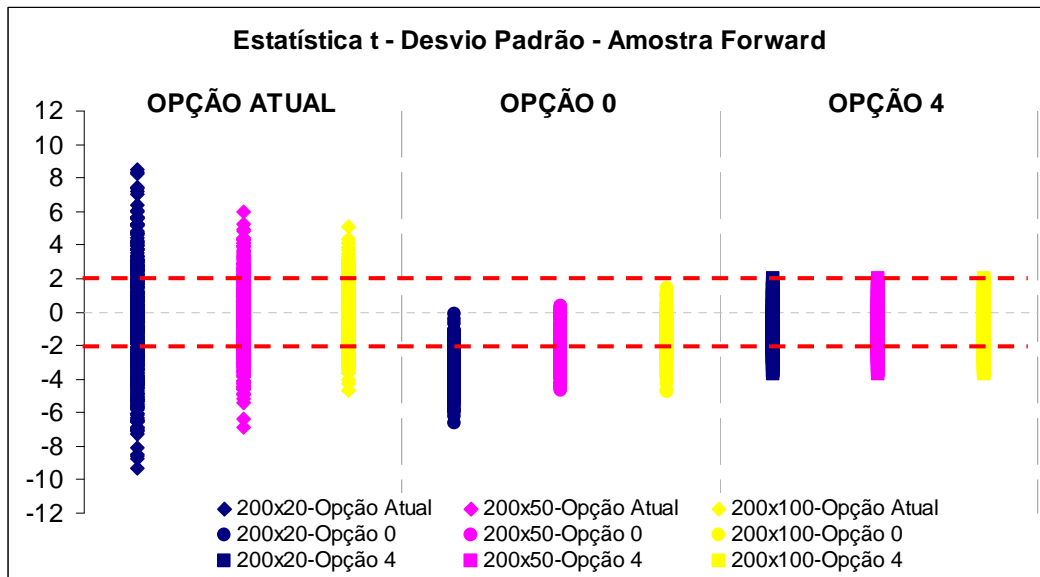


Figura 72: Estatística t –Desvio-Padrão – Amostra Forward (ruído) – Variando Tamanho Amostra Backward

Na Figura 73 e na Figura 74 são apresentadas as distribuição univariada para o subsistema Sudeste, o resultado do teste de aderência e a distribuição multivariada para opção Atual, considerando amostras backward com tamanho 50 e 100, respectivamente. Pode-se observar uma melhor aderência entre as distribuições teórica e calculada à medida que o tamanho da amostra backward aumenta.

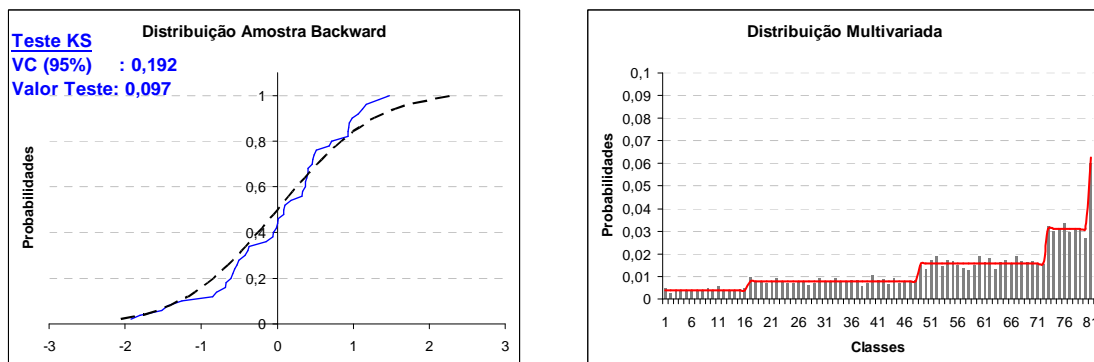


Figura 73: Distribuição – Amostra Backward (ruído) – 200x50 – Opção Atual

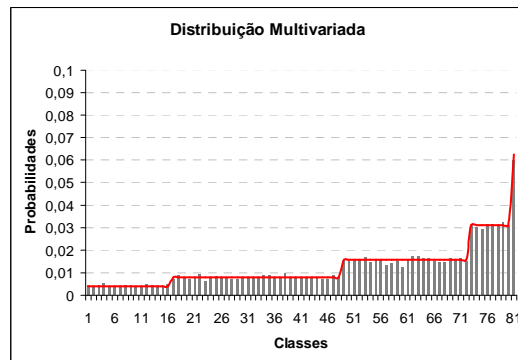
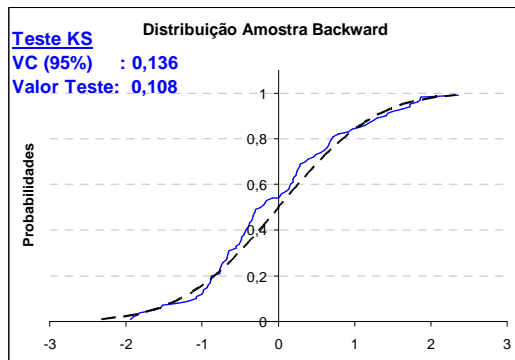


Figura 74: Distribuição – Amostra Backward (ruído) – 200x100 – Opção Atual

O mesmo pode ser verificado na Figura 75 e na Figura 76, onde são apresentados os mesmos resultados, porém considerando-se as opções 0 e 4. Pode-se observar que a distribuição univariada da amostra obtida utilizando processo de agregação continua sendo mais aderente à distribuição teórica.

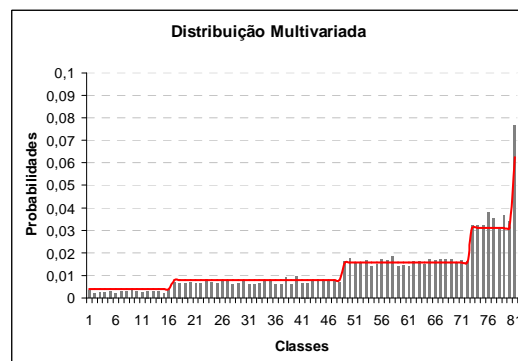
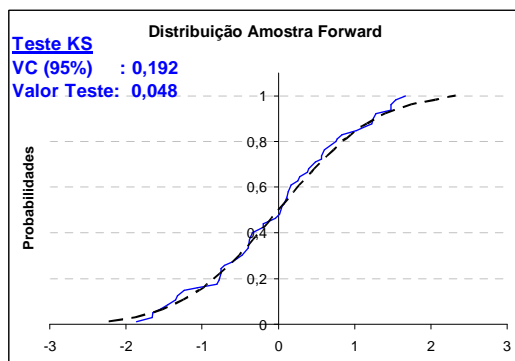


Figura 75: Distribuição – Amostra Backward (ruído) – 200x50 – Opção 0 e 4

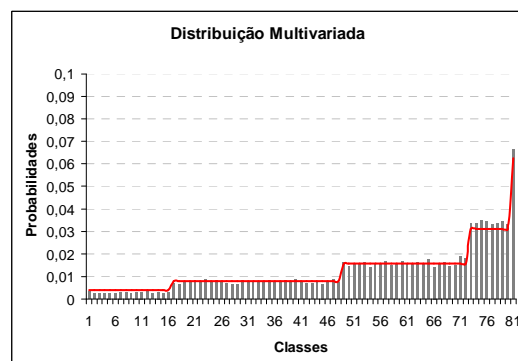
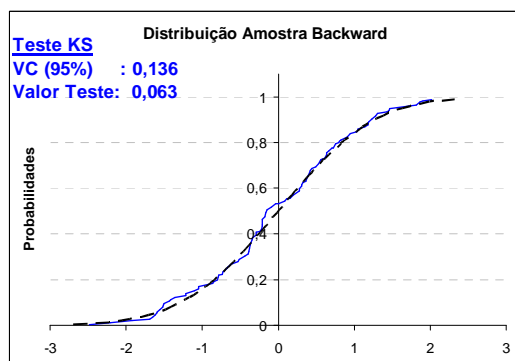


Figura 76: Distribuição – Amostra Backward (ruído) – 200x100 – Opção 0 e 4

### 7.2.5. Amostra de ruídos utilizando LHC e AD

Neste item são comparados os resultados obtidos empregando os métodos de amostragem AAS, LHC e AD no processo de geração das amostras de ruídos backward e forward. O tamanho das amostras são fixos e iguais a 200 para forward e 20 para backward.

Na Figura 77 e na Figura 78 são apresentadas as estimativas para média e desvio-padrão, assim como um intervalo de confiança construído considerando um nível de significância de 5%, para a amostra forward. Como era esperado, os métodos LHC e AD apresentam uma menor dispersão com relação aos valores de população.

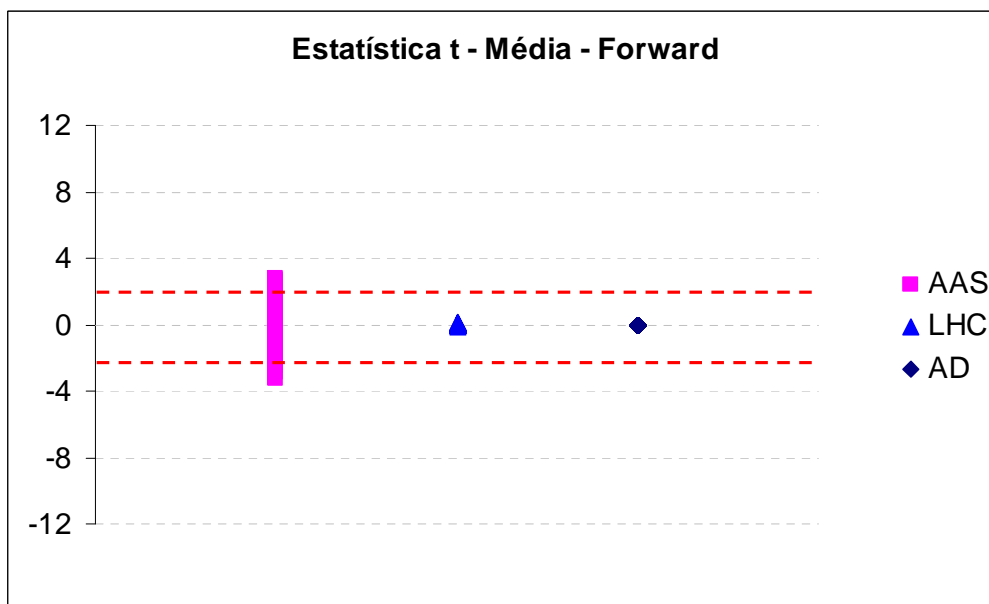


Figura 77: Estatística t – Média – Amostra Forward (ruído) – Variando Método Amostragem

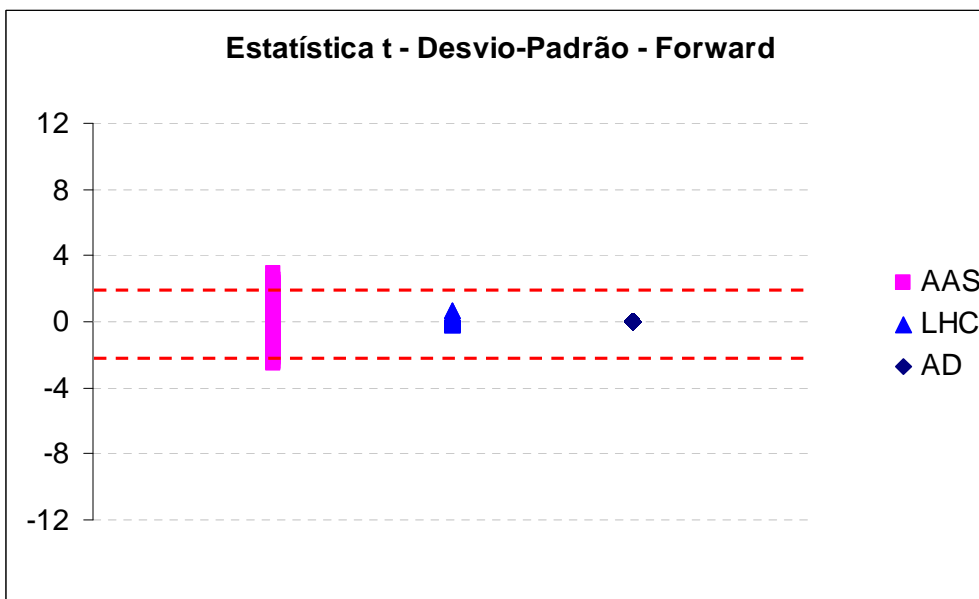


Figura 78: Estatística t – DP – Amostra Forward (ruído) – Variando Método Amostragem

Análise semelhante é realizada para amostra backward e mostrada na Figura 79 e na Figura 80. Novamente a dispersão das estimativas com relação aos valores de população são maiores no caso AAS.

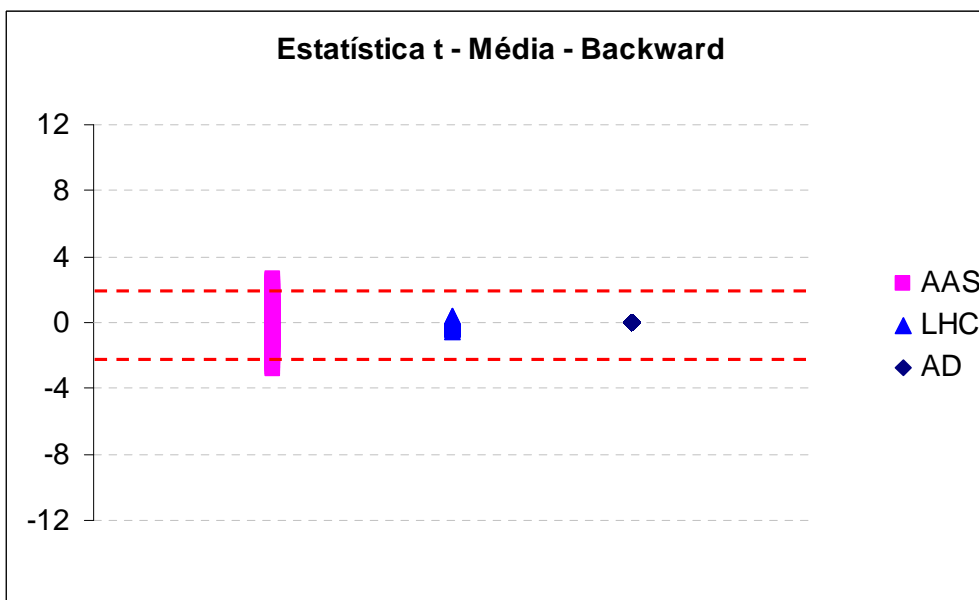


Figura 79: Estatística t – Média – Amostra Backward (ruído) – Variando Método Amostragem

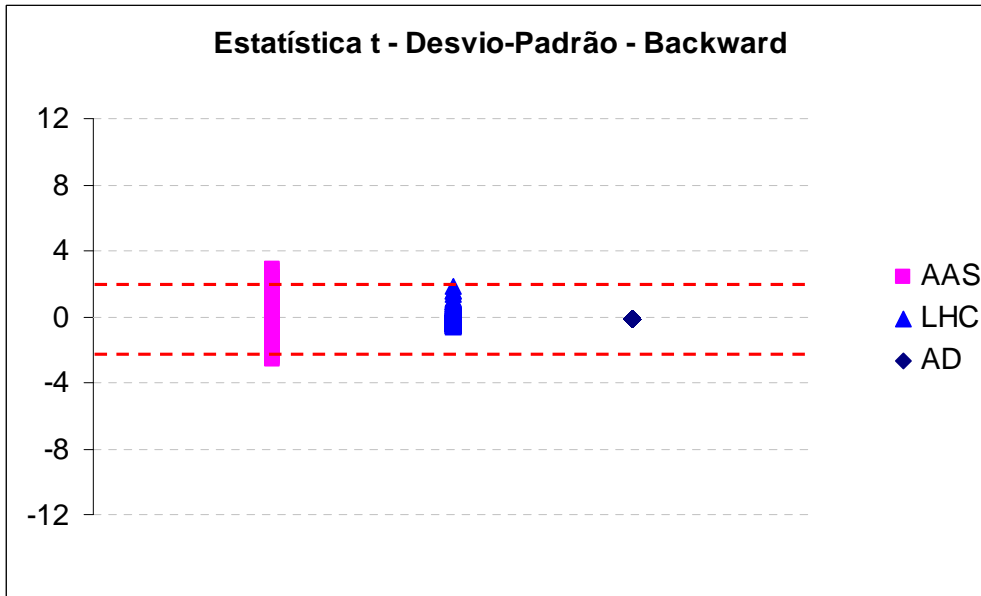


Figura 80: Estatística t – DP – Amostra Backward (ruído) – Variando Método Amostragem

As distribuições univariadas e multivariadas e o teste de aderência para as amostras forward e backward, considerando LHC e AD, são apresentadas da Figura 81 até a Figura 84. As distribuições univariadas das amostras forward geradas com as opções LHC e AD são tão aderentes quanto a amostra gerada com AAS (Figuras 55c e 56c), porém para as amostras backward, os métodos LHC e AD apresentam uma aderência bem superior àquela obtida com AAS (Figuras 59a e 60a). As distribuições multivariadas são preservadas de forma satisfatória nos três métodos de amostragem analisados.

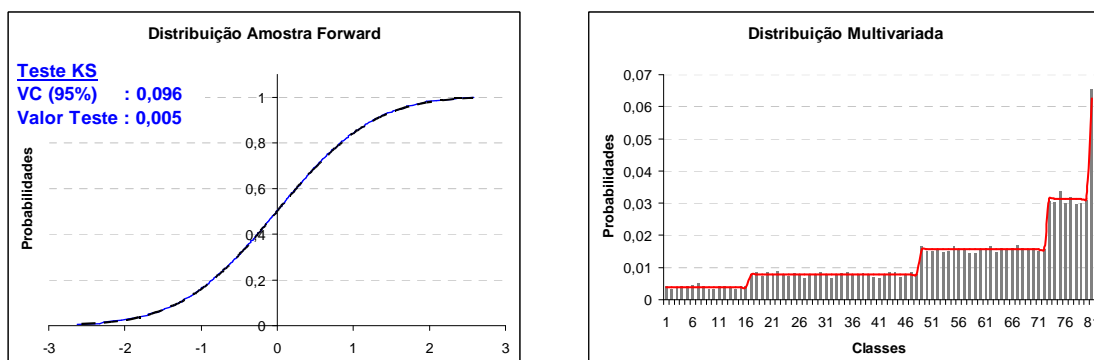


Figura 81: Distribuição – Amostra Forward (ruído) – LHC

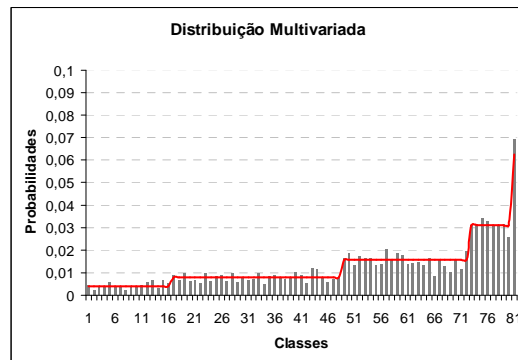
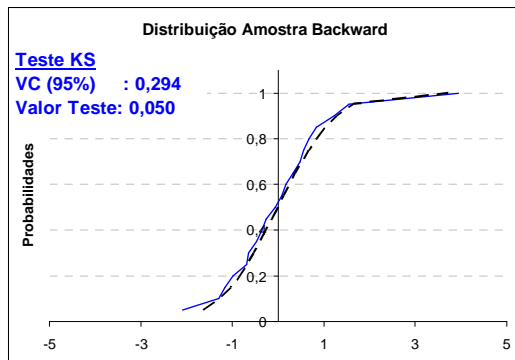


Figura 82: Distribuição – Amostra Backward (ruído) – LHC

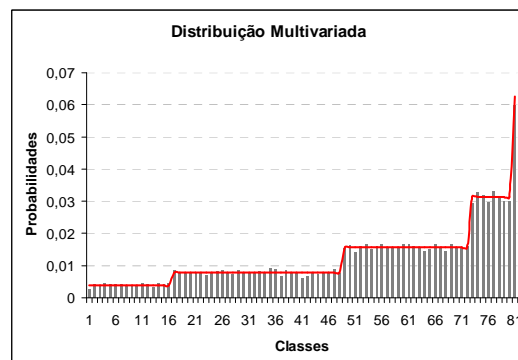
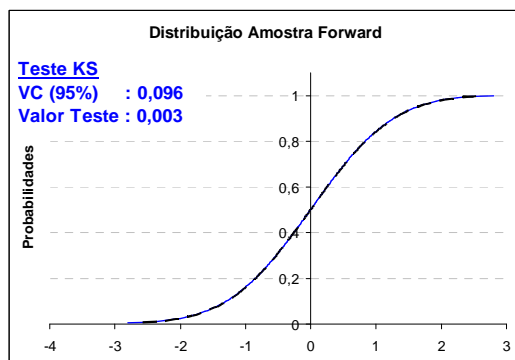


Figura 83: Distribuição – Amostra Forward (ruído) – AD

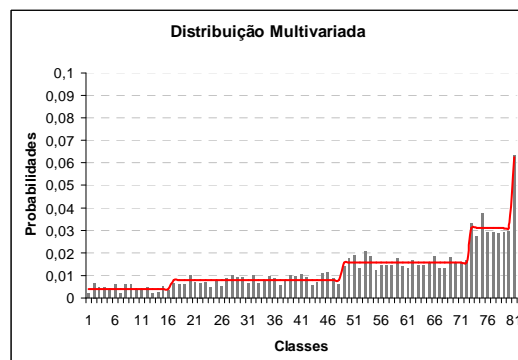
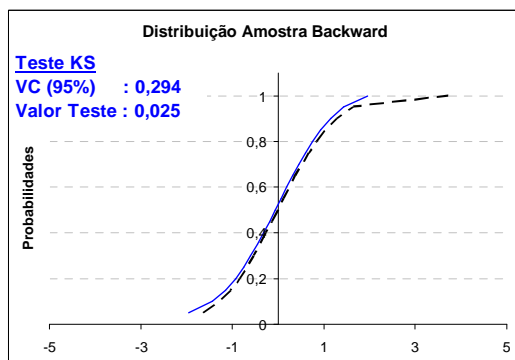


Figura 84: Distribuição – Amostra Backward (ruído) – AD

### 7.3. Análise dos Cenários Gerados para Passo Forward

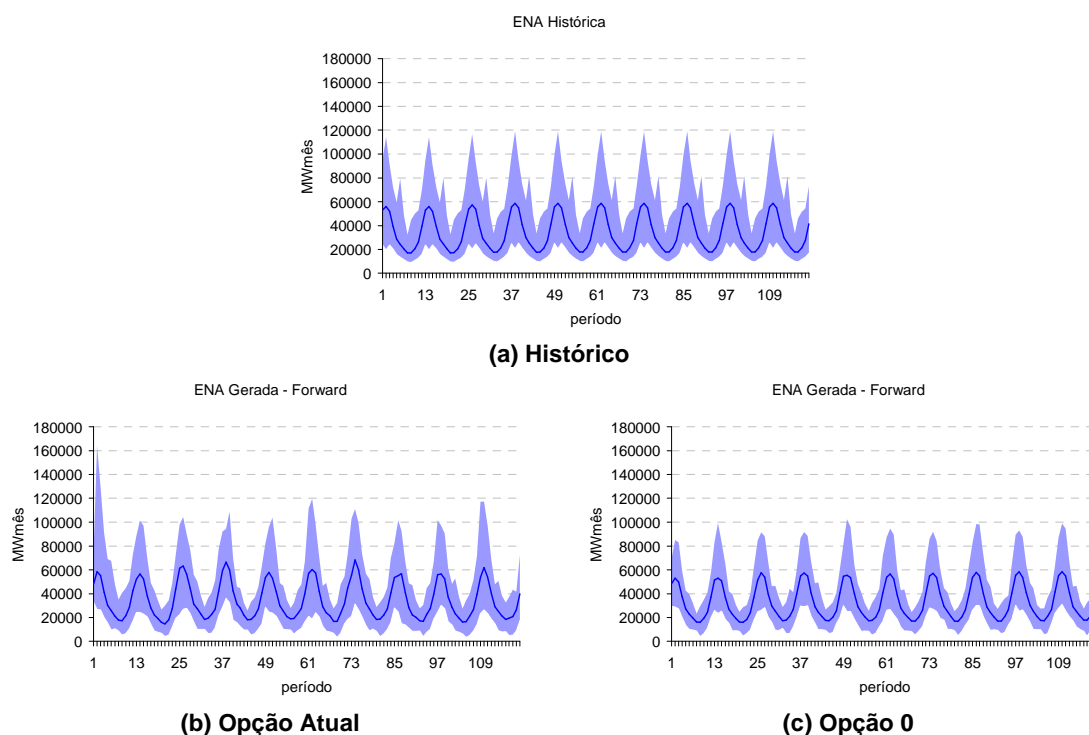
Neste item é apresentada uma avaliação dos cenários hidrológicos obtidos para o passo forward a partir da utilização das opções descritas no trabalho, além dos métodos de amostragem LHC e AD. Nesta análise são considerados 200 cenários forward e 20 cenários backward, e para o processo de agregação é fornecida uma amostra de 2000 vetores de ruídos.

São utilizados testes não condicionados para a avaliação da capacidade de preservação das estatísticas média, desvio-padrão e correlação cruzada, e testes para verificação da aderência entre as distribuições gerada e histórica.

Inicialmente é realizada uma comparação entre as envoltórias dos cenários hidrológicos históricos e gerados com todas as opções descritas.

### 7.3.1. Envoltória dos Cenários

Na Figura 85b até a Figura 85g são mostradas as envoltórias e os valores médios dos 200 cenários gerados de energia natural afluyente para o subsistema Sudeste. A envoltória da série histórica é apresentada na Figura 85a. Observa-se que todas as opções de geração avaliadas são capazes de reproduzir bem o histórico. As opções Atual e LHC apresentam um valor máximo bem maior que o histórico para o mês de fevereiro do primeiro e sexto anos, respectivamente. Por se tratar de apenas um cenário atípico, o valor da média não é muito afetado.



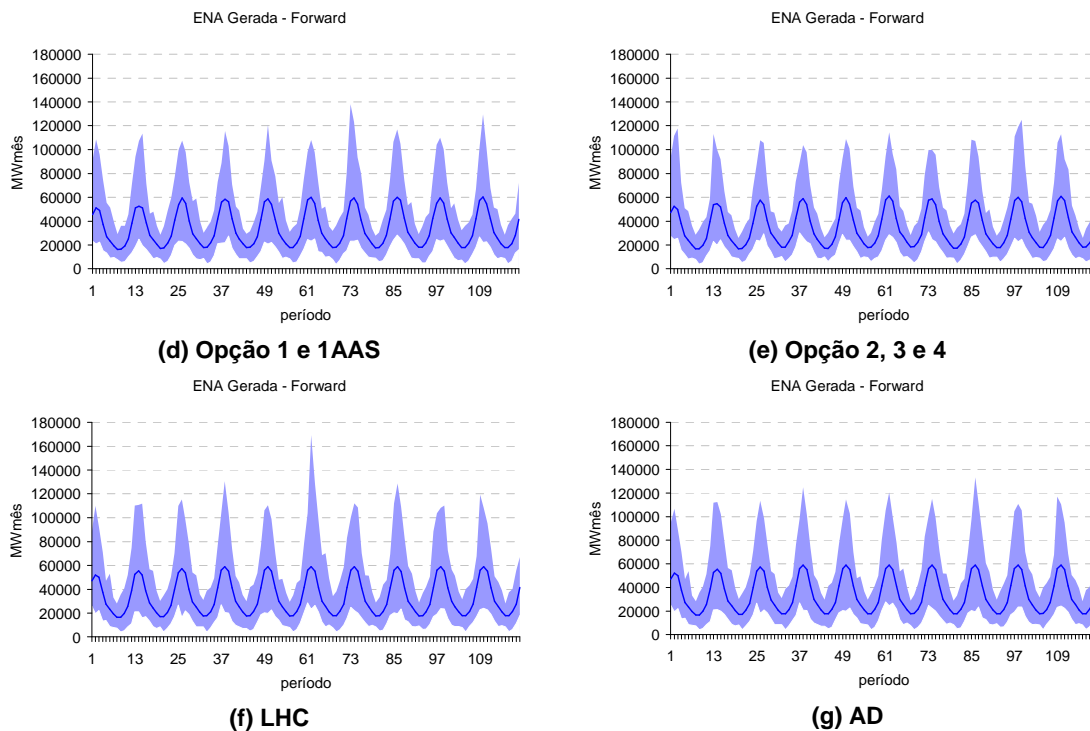
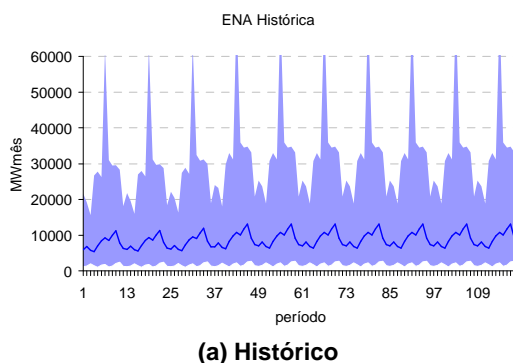


Figura 85: Envoltória Cenários Forward – Sudeste

Na Figura 86b até a Figura 86g são apresentadas as envoltórias e os valores médios para o subsistema Sul, assim como seu valor histórico, Figura 86a. O Sul é o subsistema que apresenta maior variabilidade no seu comportamento hidrológico, seu coeficiente de variação mensal é bem próximo de um. O valor máximo de ENA observado no Sul foi no mês de julho de 1983, quando ocorreu uma grande cheia na região. Excluindo o valor máximo de julho, pode-se observar que as opções 2, 3, 4, LHC e AD são as que mais se aproximam da envoltória histórica.

Note que o valor médio dos cenários gerados não é preservado de maneira satisfatória pela Opção Atual.



(a) Histórico



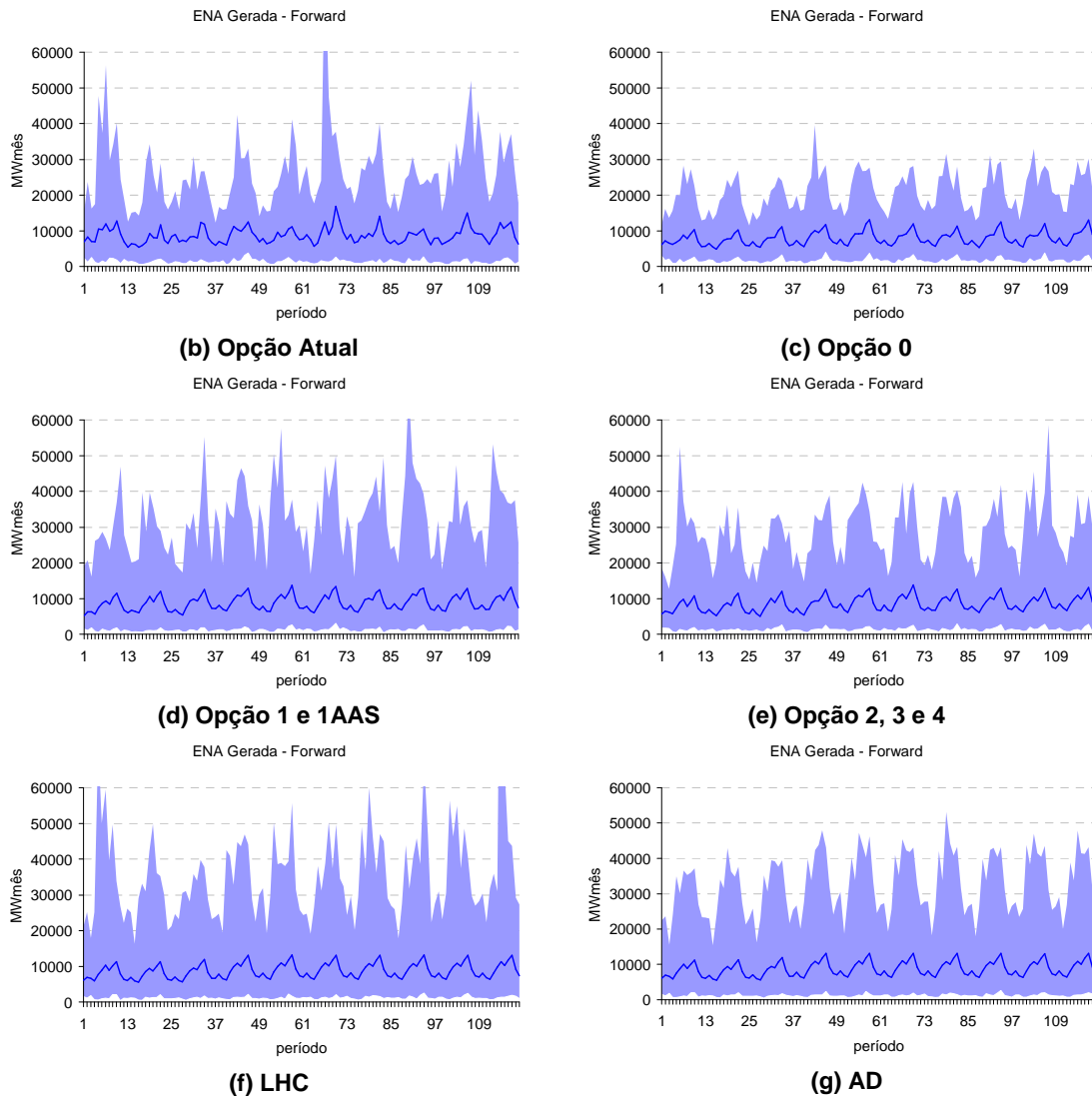
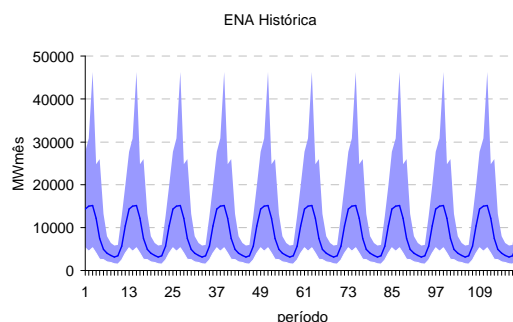
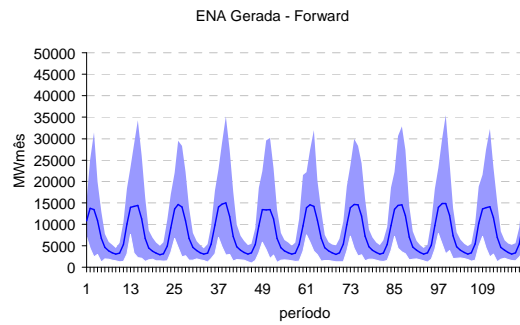
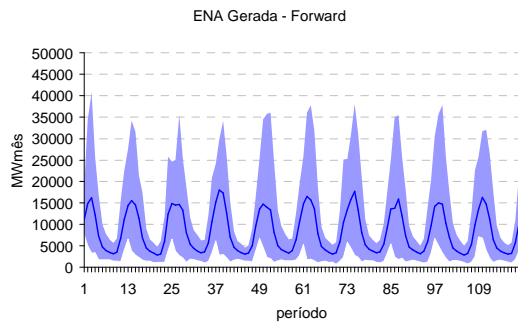


Figura 86: Envoltória Cenários Forward – Sul

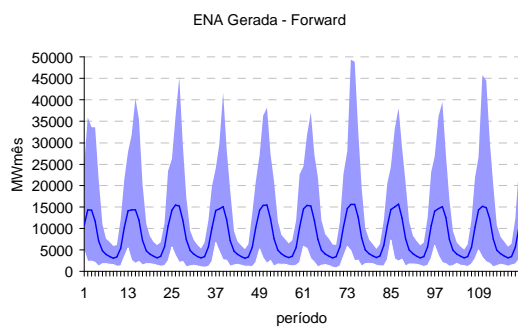
A mesma análise é realizada para o subsistema Nordeste, Figura 87a até a Figura 87g. Todas as opções analisadas preservam de forma satisfatória a envoltória histórica. A opção 0 apresenta dificuldade em preservar valores mais elevados.



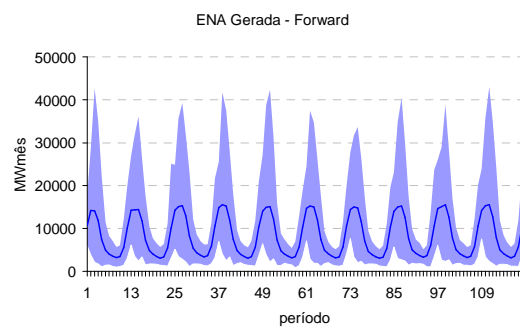
**(a) Histórico**



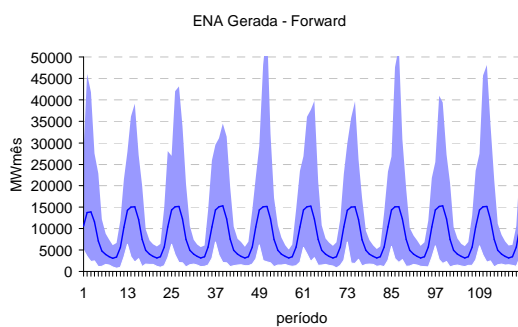
**(b) Opção Atual**



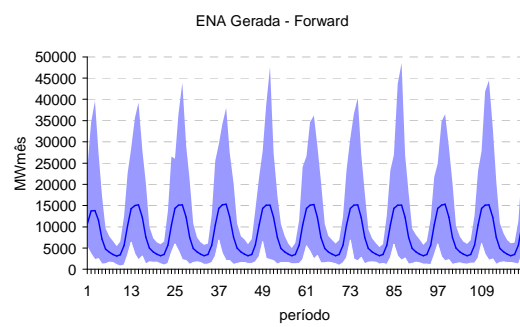
**(c) Opção 0**



**(d) Opção 1 e 1AAS**



**(e) Opção 2, 3 e 4**



**(f) LHC**



**(g) AD**

Figura 87: Envoltória Cenários Forward – Nordeste

Novamente, todas as opções analisadas são capazes de reproduzir satisfatoriamente a envoltória histórica do subsistema Norte, Figura 88a até Figura 88g, exceto a opção 0 com relação a valores superiores a 25000 MWh/mês.

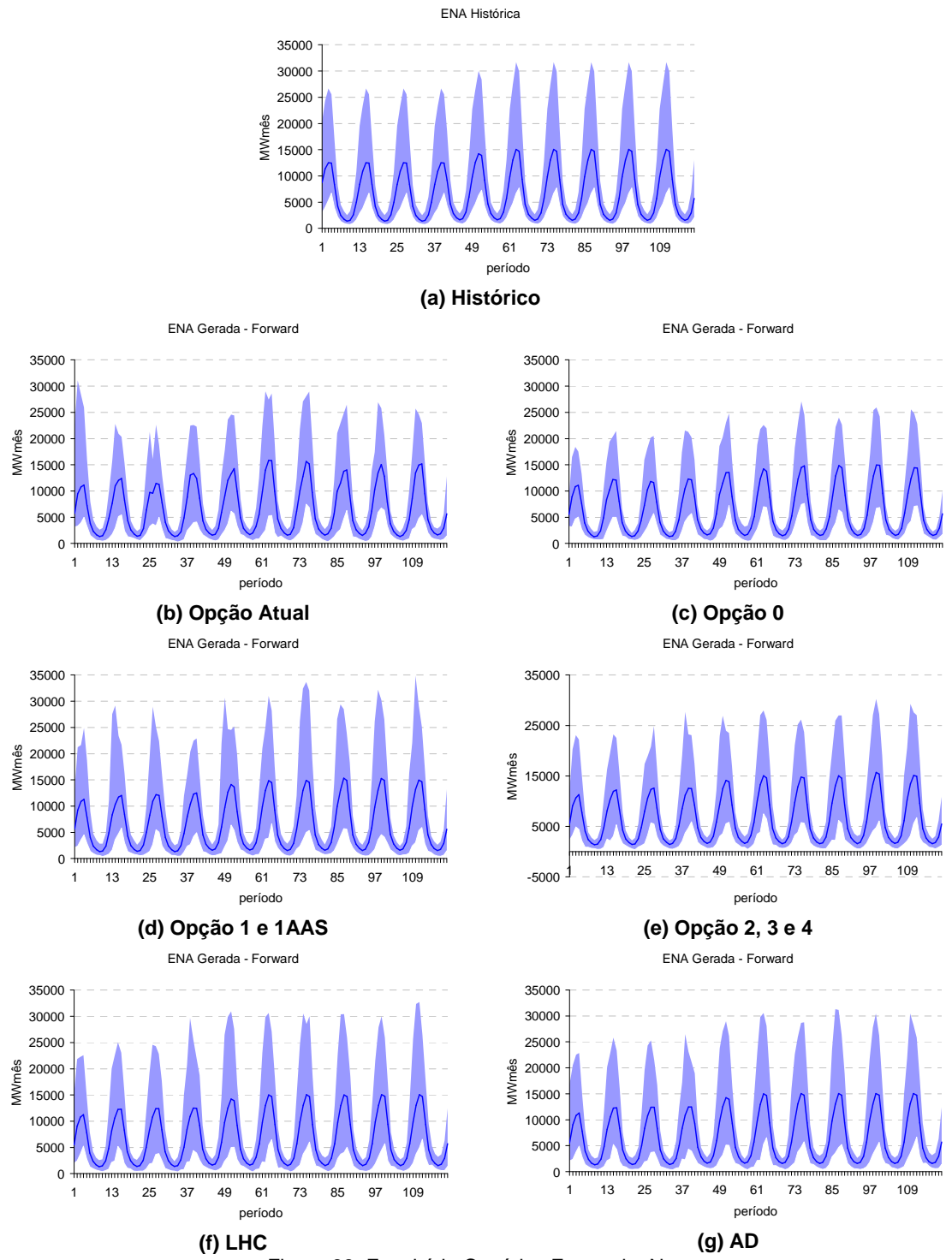


Figura 88: Envolvória Cenários Forward – Norte

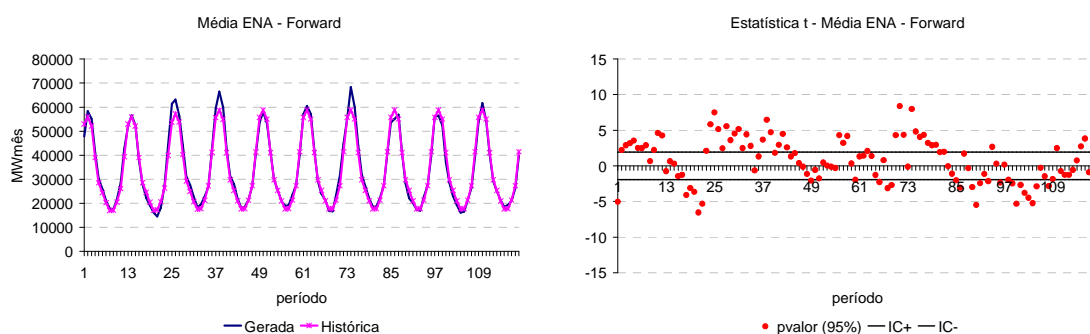
### 7.3.2. Média dos Cenários

Neste item é realizada uma análise mais detalhada da média dos cenários hidrológicos. O nível de significância do teste é de 5% e como parâmetros de população são considerados os valores históricos (testes não condicionados). Na Figura 89 são apresentadas a evolução temporal da média histórica e dos cenários gerados e a estatística t desta estimativa, para o subsistema Sudeste.

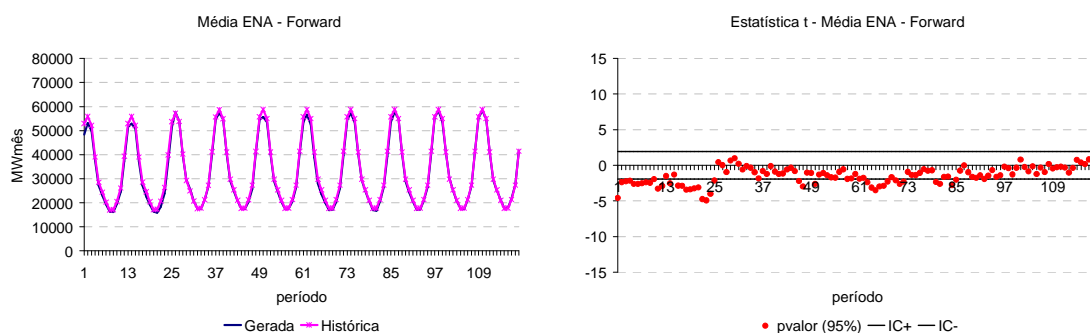
Os cenários gerados para os primeiros períodos são fortemente condicionados ao passado recente (tendência hidrológica recente). Desta forma, as médias dos primeiros períodos diferem do valor histórico, uma vez que o passado recente é diferente da MLT, conforme mostrado na Figura 46.

Analisando a Figura 89a, observa-se que a Opção Atual é aquela que apresenta maior variabilidade com relação à média. Essa variabilidade é reduzida quando se aplica técnicas de agregação tanto na obtenção da amostra backward (opção 0), quanto na obtenção da amostra forward (opções 2,3 e 4). Figura 89b e Figura 89d.

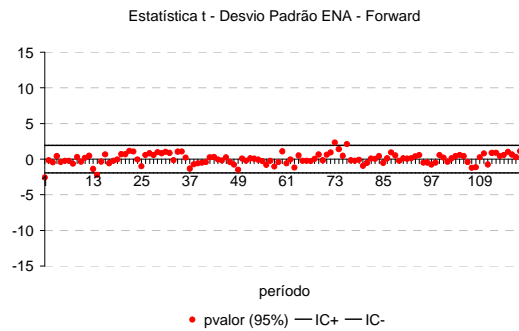
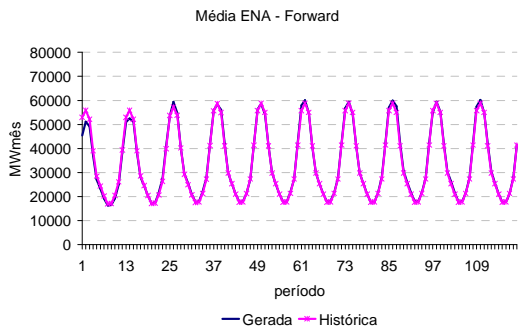
Comparando as opções 1/1AAS, LHC e AD, verifica-se que as técnicas de amostragem estratificadas reduzem de sobremaneira a variabilidade da estimativa da média.



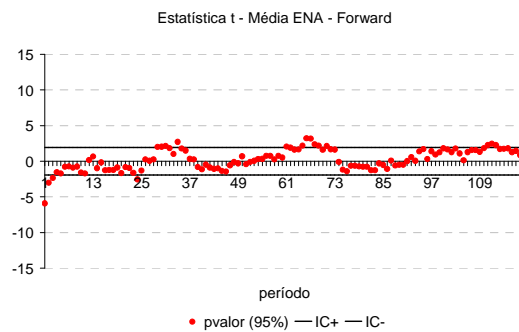
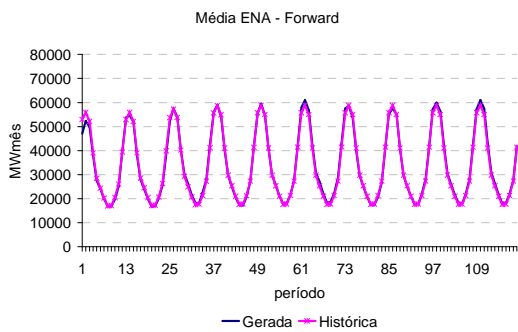
(a) Opção Atual



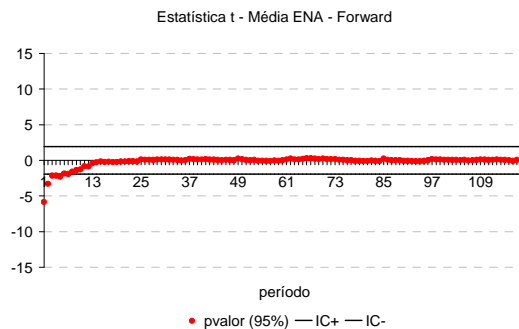
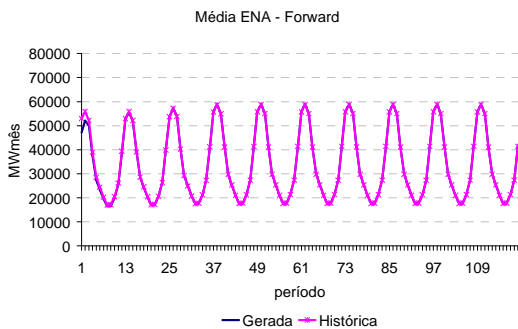
(b) Opção 0



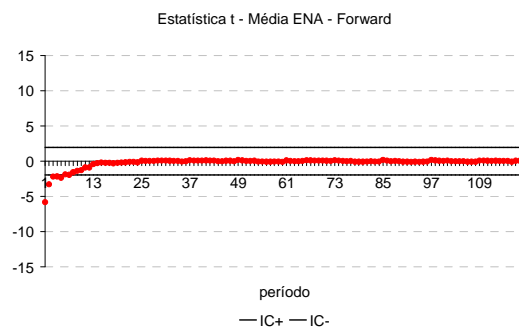
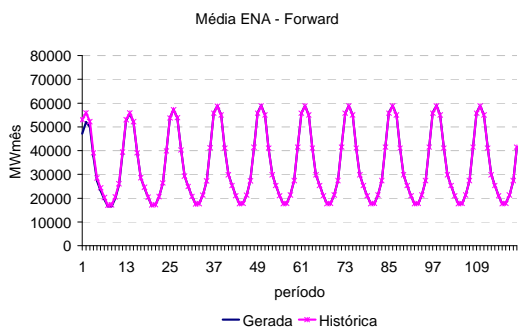
**(c) Opção 1 e 1AAS**



**(d) Opção 2, 3 e 4**



**(e) Opção LHC**



**(f) Opção AD**

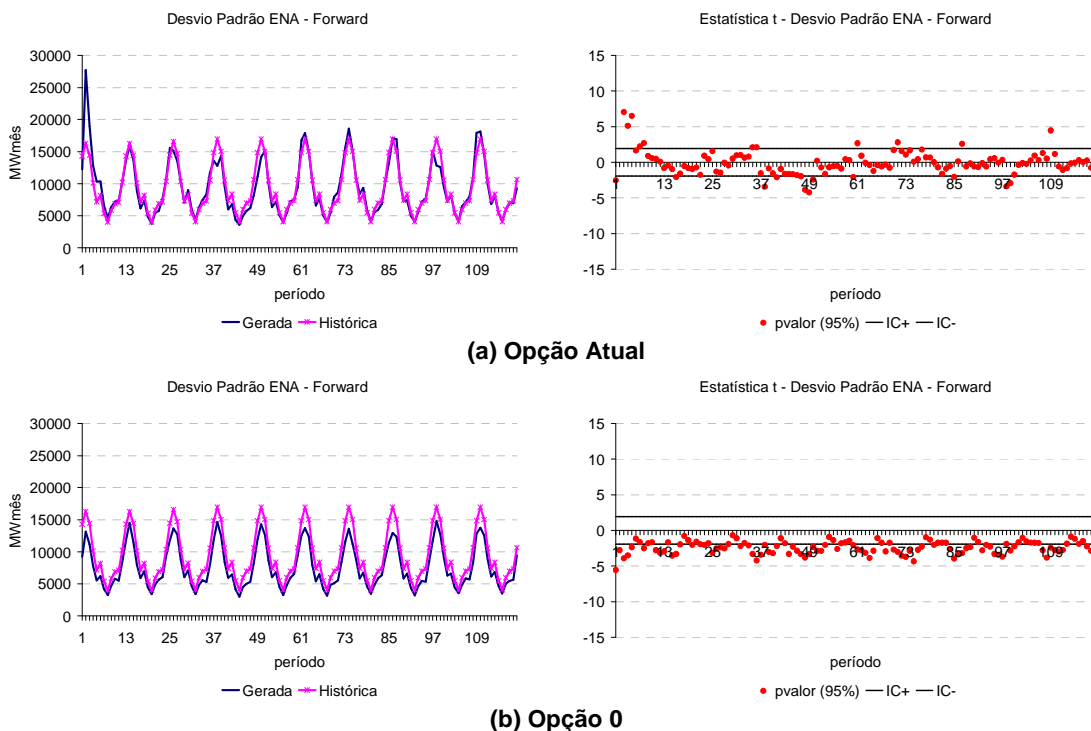
Figura 89: Média Cenários Forward – Sudeste

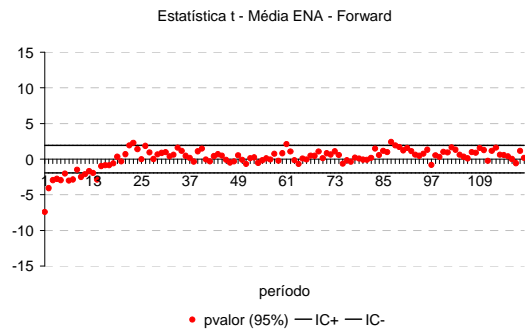
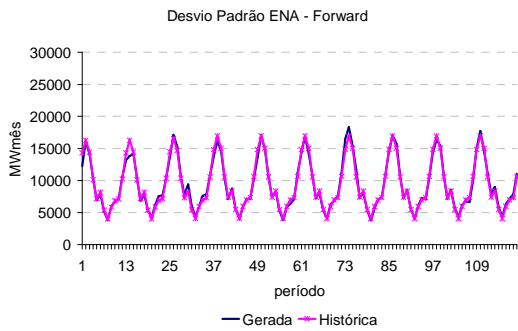
O comportamento observado para os demais subsistemas é bastante semelhante. Os resultados obtidos para os demais subsistemas são apresentados no Anexo B.

### 7.3.3. Desvio-padrão dos Cenários

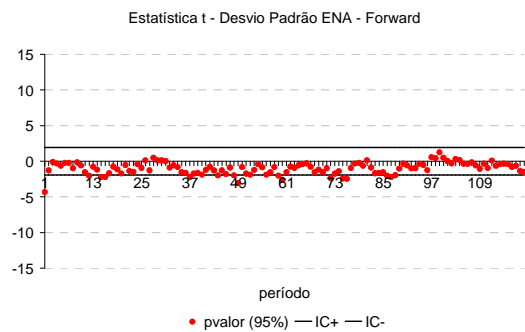
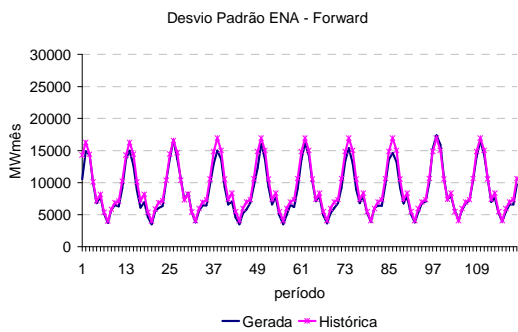
Na Figura 90 é avaliada a estimativa do desvio-padrão para o subsistema Sudeste, com relação ao desvio-padrão histórico. A estimativa do desvio-padrão nos primeiros meses do estudo está muito condicionada ao passado recente, assim como ocorre com a estimativa da média.

Pode-se observar que as opções Atual e 0 apresentam uma maior variabilidade para a estimativa do desvio-padrão, isto ocorre pois a amostra forward é gerada a partir de uma amostra de tamanho bem reduzido, no caso analisado a amostra backward tem tamanho igual a 20. Além disso, observa-se que a opção 0 degrada bastante o desvio-padrão com relação ao valor histórico. As opções 2, 3 e 4 que utilizam o processo de agregação para obtenção da amostra forward representam de forma satisfatória o desvio padrão, porém pode-se notar uma pequena tendência de redução deste valor. A preservação do desvio histórico é alcançada pela opção 1/1AAS, com exceção dos primeiros meses. As opções LHC e AD apresentam uma variabilidade menor para a estimativa do desvio-padrão do que a observada na opção 1/1AAS.

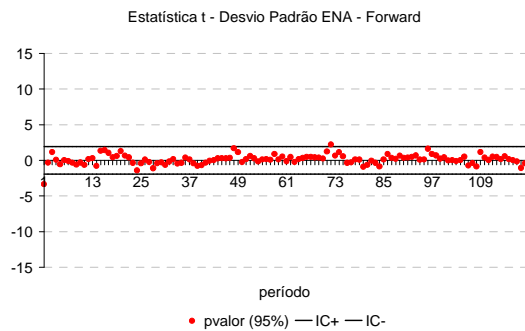
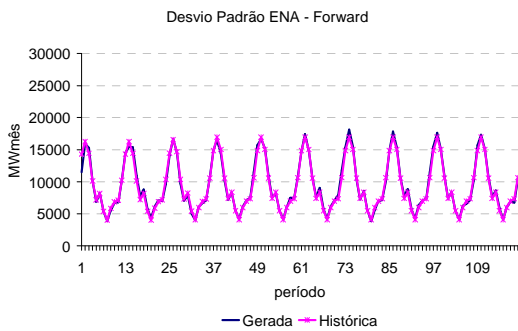




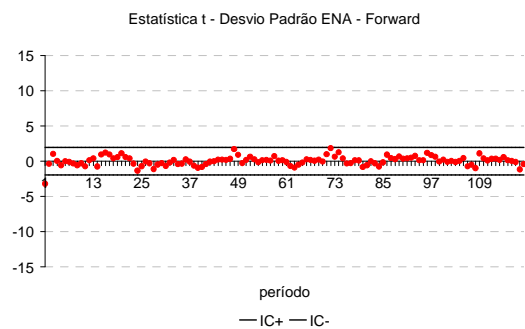
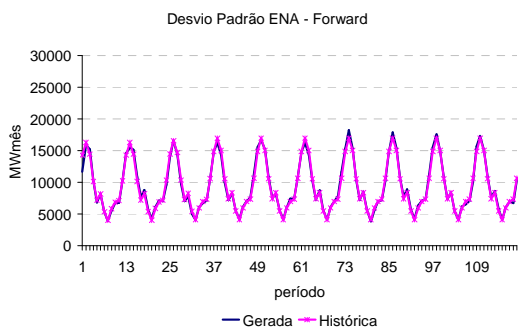
**(c) Opção 1 e 1AAS**



**(d) Opção 2, 3 e 4**



**(e) Opção LHC**



**(f) Opção AD**

Figura 90: Desvio-Padrão Cenários Forward – Sudeste

Novamente, o comportamento observado para os demais subsistema é similar ao comportamento observado para o subsistema Sudeste. Esses resultados podem ser conferidos no Anexo B.

### 7.3.4. Distribuição Univariada dos Cenários

Na próxima análise é aplicado o teste de Kolmogorov-Smirnov para verificar a aderência entre as distribuições histórica e gerada a cada período do horizonte de estudo. O nível de significância adotado é de 5% e o valor crítico do teste é representado por uma linha tracejada nas figuras a seguir.

Da Figura 91a até a Figura 91f são apresentados os resultados deste teste para os cenários hidrológicos do subsistema Sudeste. Pode-se destacar o grande número de testes rejeitados quando se utiliza a opção Atual. As melhores representações da distribuição univariada do subsistema Sudeste são obtidas com as opções LHC e AD. As demais opções apresentam um comportamento satisfatório.

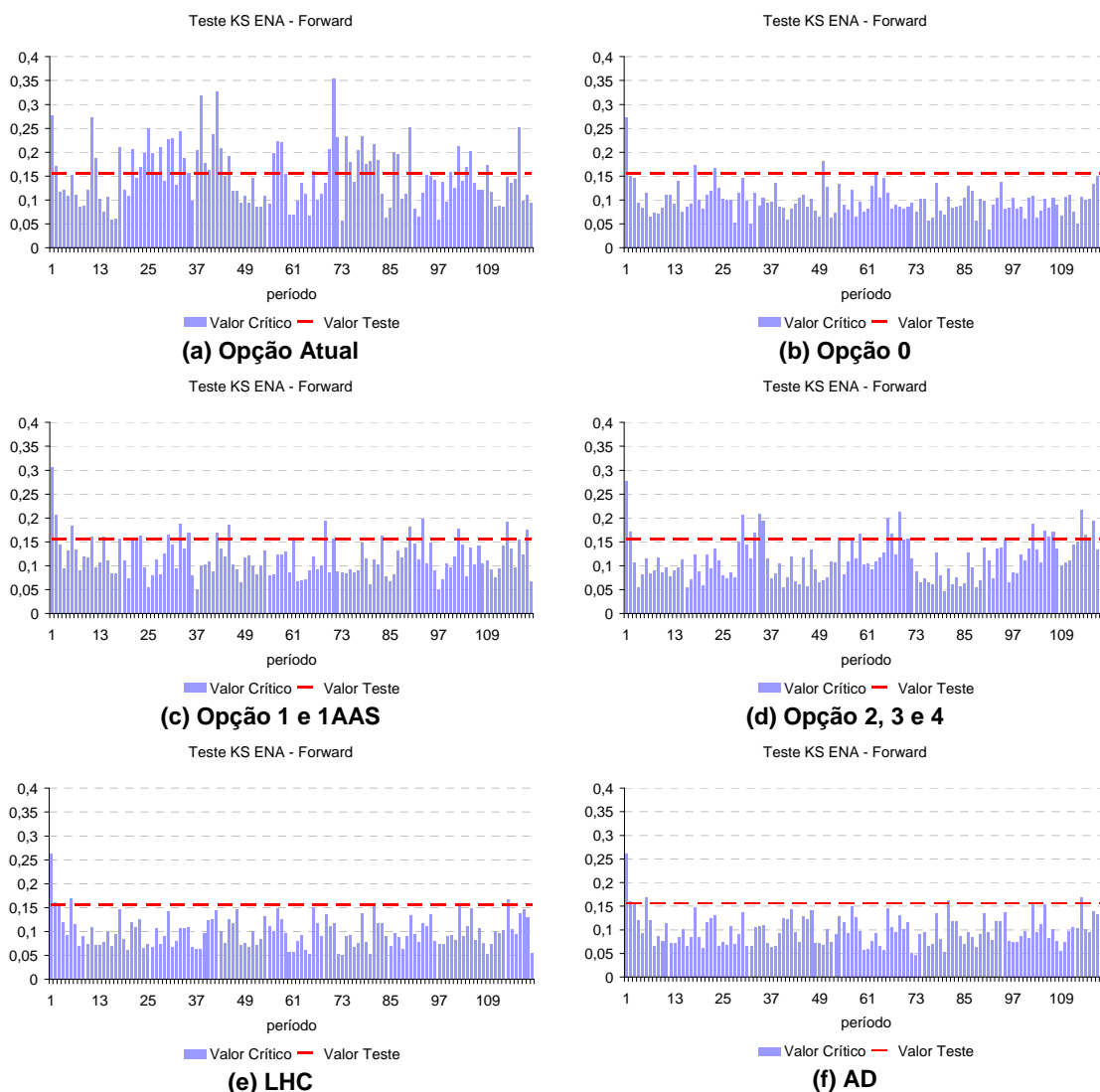




Figura 91: Teste de Aderência Cenários Forward – Sudeste

Os resultados obtidos com a aplicação do teste KS para a distribuição univariada de energia natural afluyente do Sul são apresentados da Figura 92a até a Figura 92f. Conforme observado anteriormente, a opção Atual tem um grande número de rejeições, porém ao aplicar as técnicas de agregação para obter a amostra backward, a partir da qual é gerada a amostra forward (opção 0), a aderência entre as distribuições histórica e gerada é alcançada. As demais opções apresentam um bom comportamento com relação ao teste de aderência.

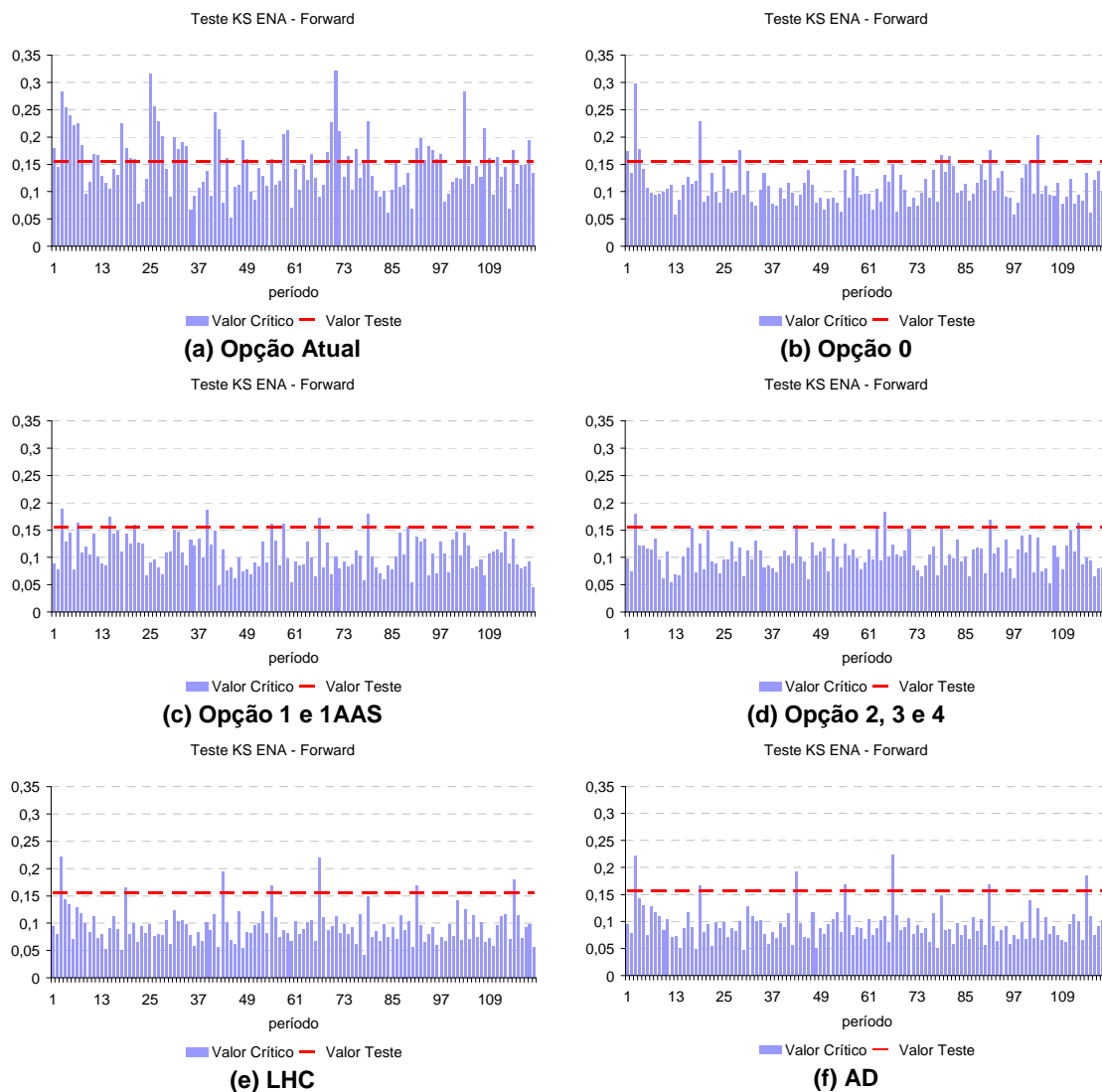


Figura 92: Teste de Aderência Cenários Forward – Sul

Analisando a Figura 93a até Figura 93f pode-se verificar que, em geral, todas as opções apresentam bom nível de aceitação dos testes de aderência, para o

subsistema Nordeste. Novamente a opção que apresenta o maior índice de rejeição é a opção Atual, porém em nível menor do que o observado nos subsistemas Sul e Sudeste.

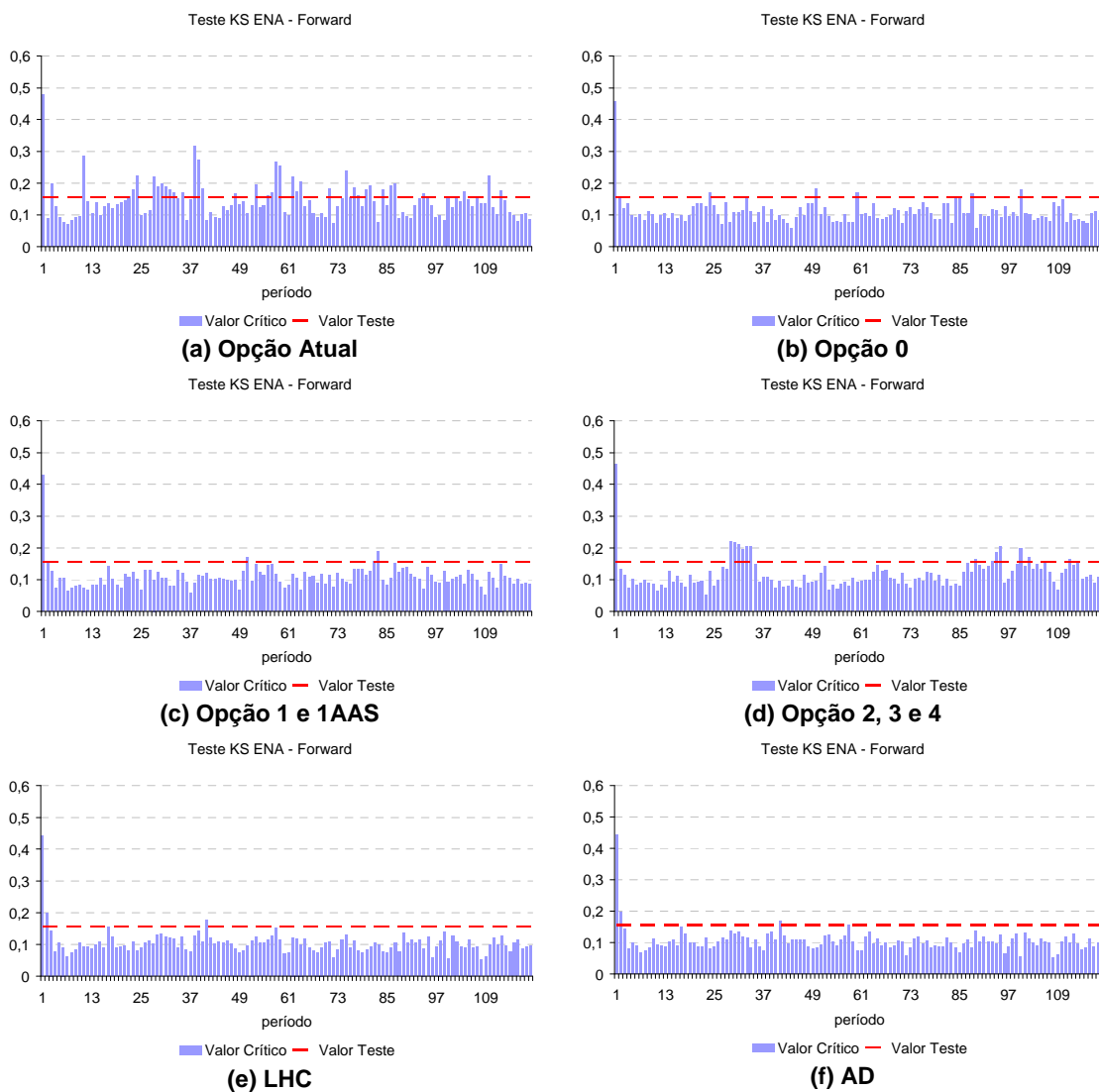


Figura 93: Teste de Aderência Cenários Forward – Nordeste

As mesmas considerações feitas para o subsistema Nordeste se aplicam ao subsistema Norte, de acordo com a Figura 94a até Figura 94f.

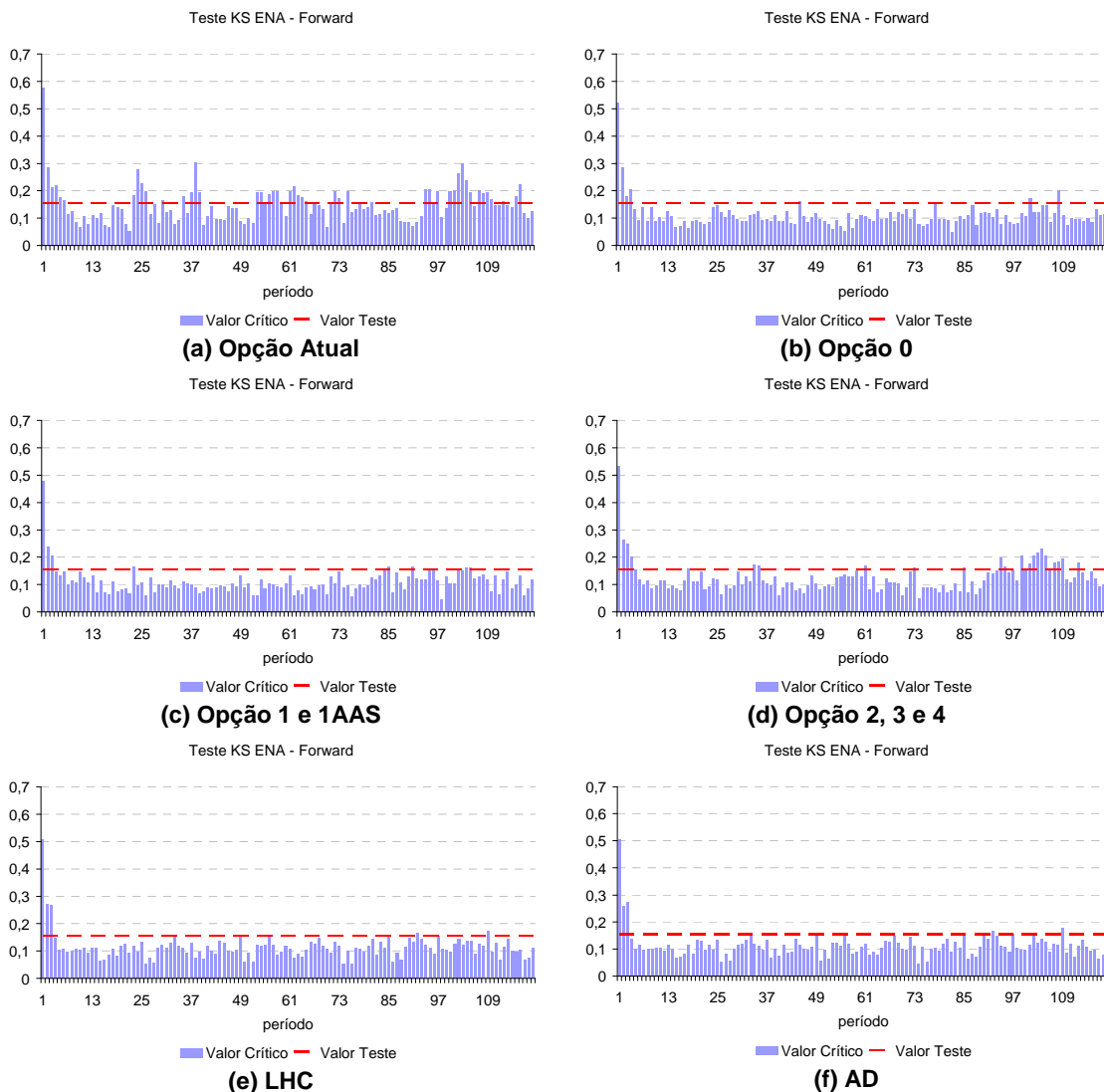


Figura 94: Teste de Aderência Cenários Forward – Norte

### 7.3.5. Correlação Cruzada dos Cenários

A preservação da correlação cruzada é um objetivo a ser perseguido pelo processo de geração de cenários sintéticos multivariados. A seguir, é feita uma análise desta grandeza através de inspeção visual dos resultados obtidos. Os valores calculados são comparados com o valor histórico.

Da Figura 95a até a Figura 95f é apresentada a evolução temporal da correlação cruzada entre os subsistemas Sudeste e Sul. A linha mais clara indica a correlação cruzada histórica. Observe que há uma variação em torno do valor histórico para as opções Atual e 0 a 4. A variabilidade é bastante acentuada para a opção Atual,

reforçando a indicação de que a maneira utilizada para obter a amostra forward não é a mais apropriada. Nas opções LHC e AD pode-se observar uma tendência de redução no valor da correlação cruzada estimada.

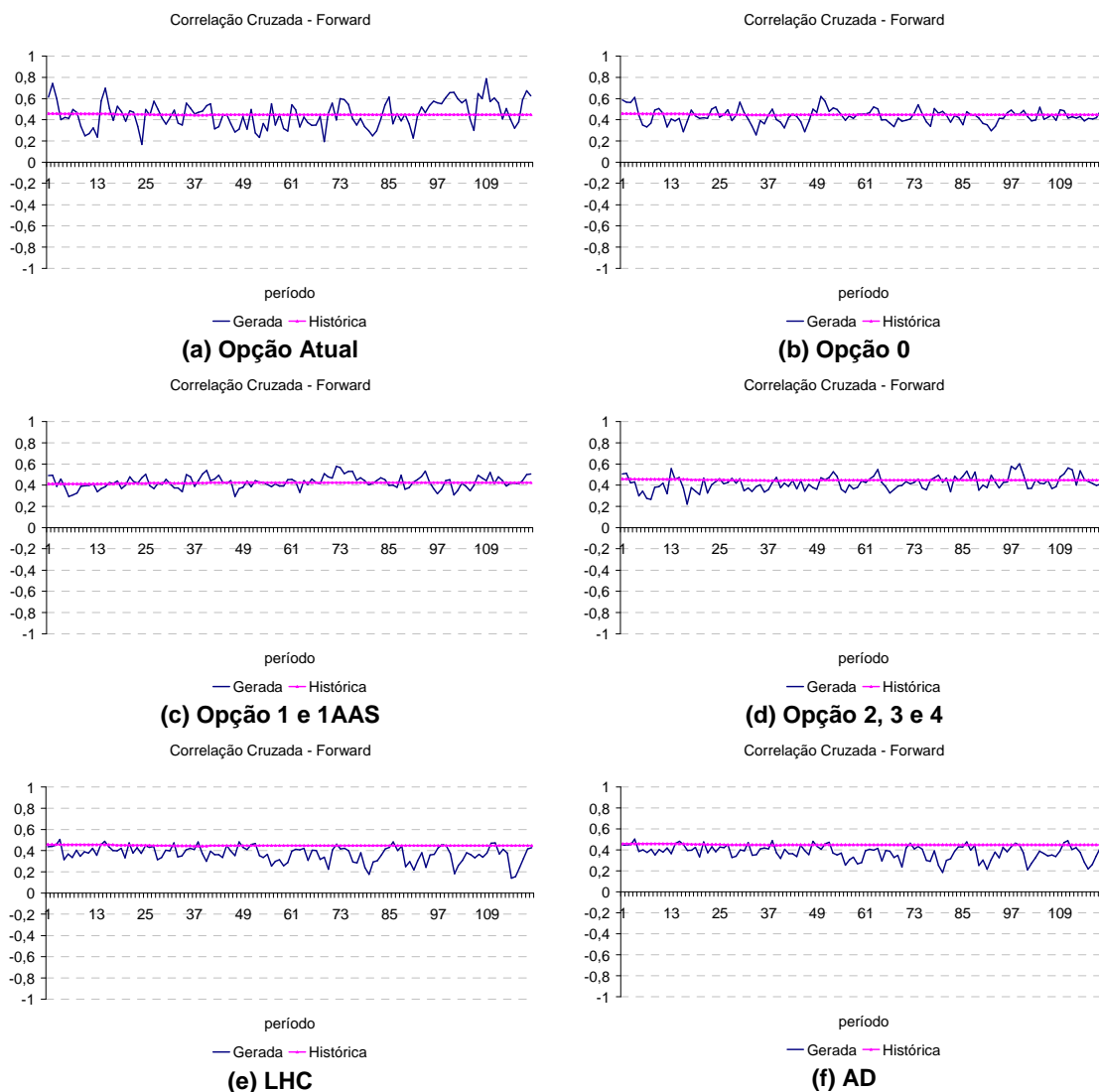


Figura 95: Correlação Cruzada Cenários Forward – Sudeste x Sul

Da Figura 96a até a Figura 96f pode-se analisar o comportamento da correlação cruzada entre os subsistemas Sudeste e Nordeste. As correlações cruzadas obtidas com as opções LHC e AD continuam, em geral, abaixo do valor histórico, porém as diferenças relação ao histórico são bem menores. Conforme observado anteriormente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico.

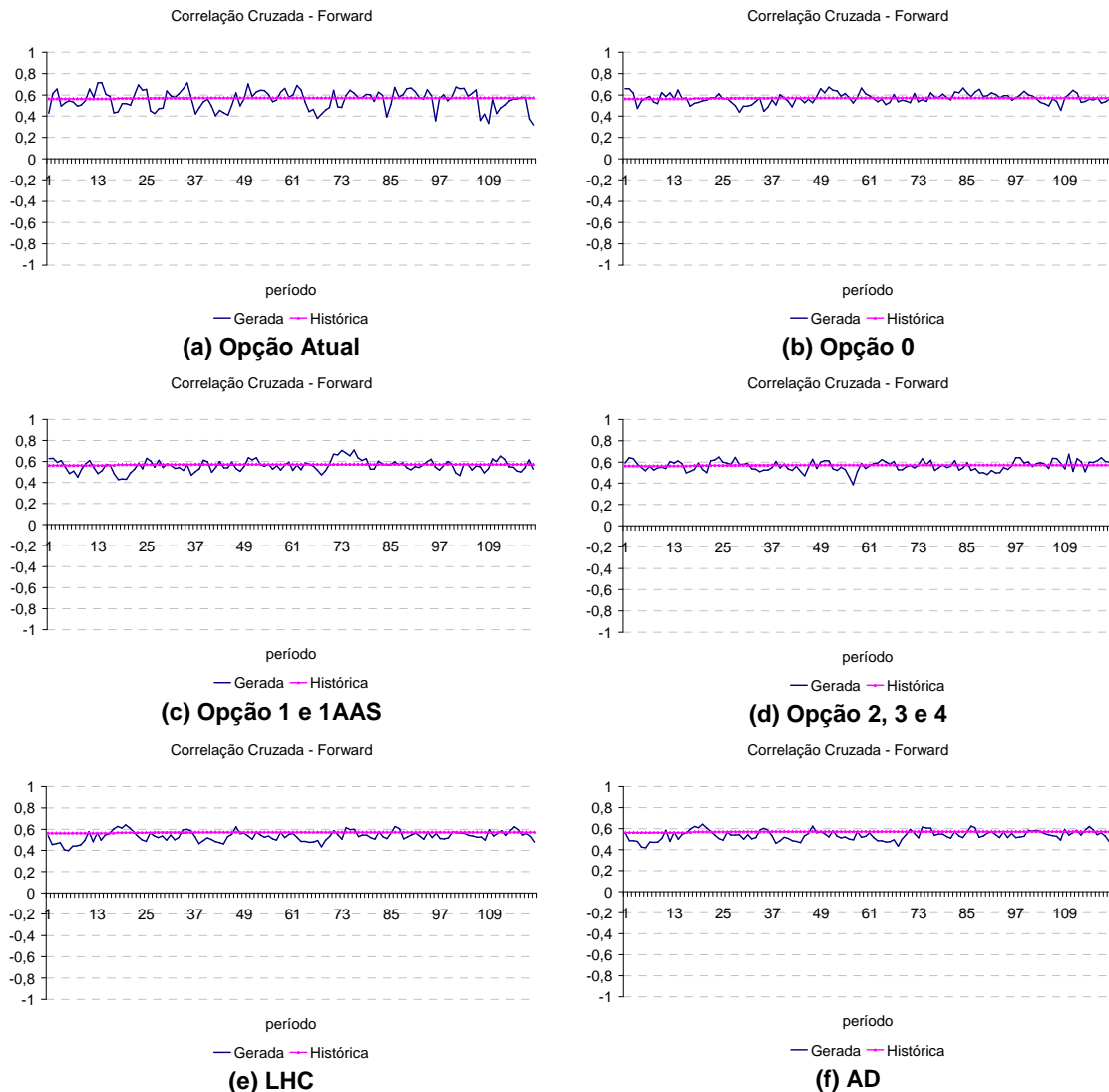


Figura 96: Correlação Cruzada Cenários Forward – Sudeste x Nordeste

Da Figura 97a até a Figura 97f é apresentada a evolução temporal da correlação cruzada entre os subsistemas Sudeste e Norte. Conforme observado anteriormente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico. A correlação cruzada dos cenários gerados com a opção 0 fica um pouco mais dispersa em relação ao valor histórico. As correlações cruzadas obtidas com as demais opções se aproximam razoavelmente do valor observado no histórico.

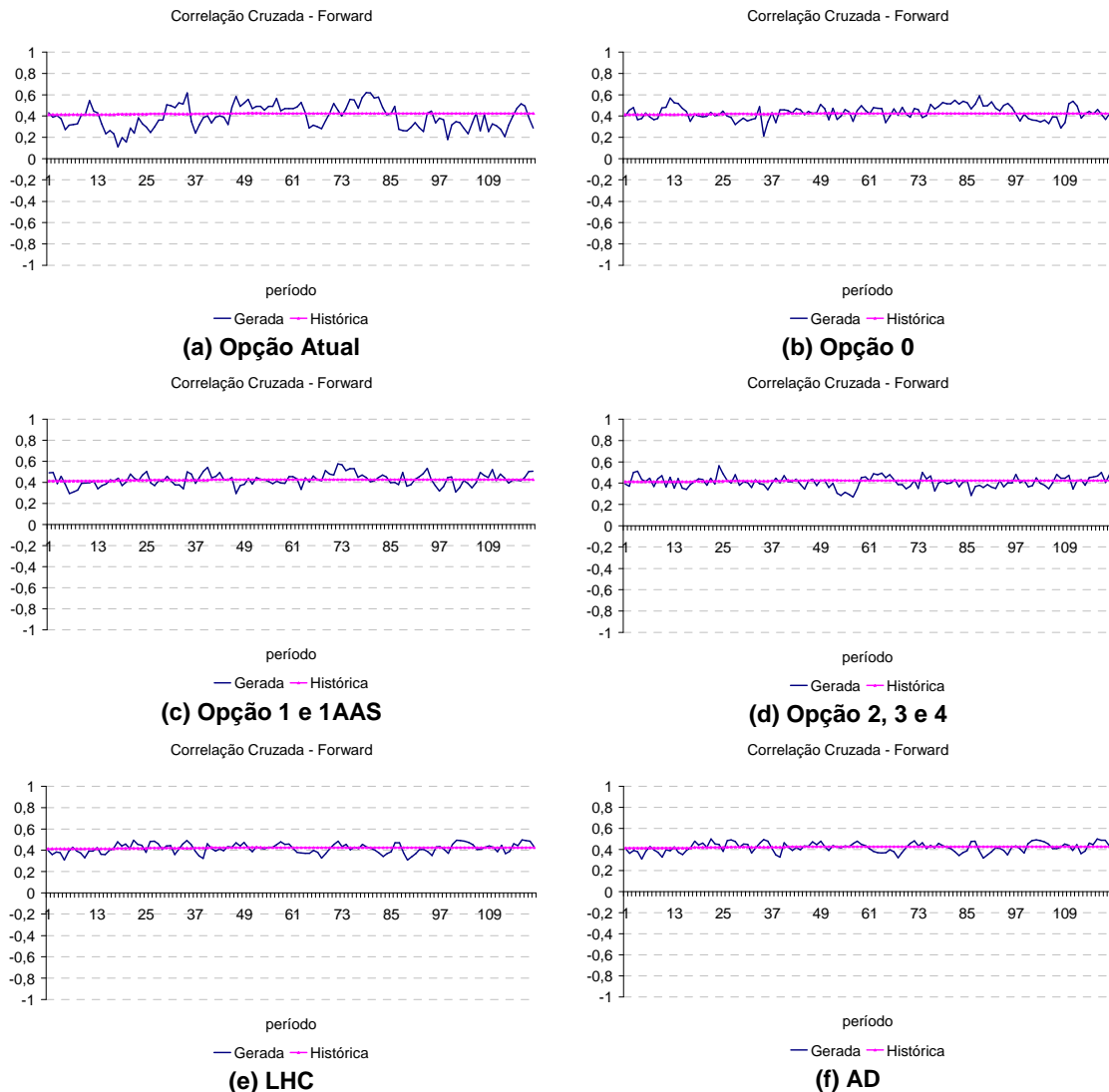


Figura 97: Correlação Cruzada Cenários Forward – Sudeste x Norte

Da Figura 98a até a Figura 98f pode-se analisar o comportamento da correlação cruzada entre os subsistemas Sul e Nordeste. As correlações cruzadas obtidas com as opções LHC e AD continuam, em geral, abaixo do valor histórico em valores absolutos. Conforme observado anteriormente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico.

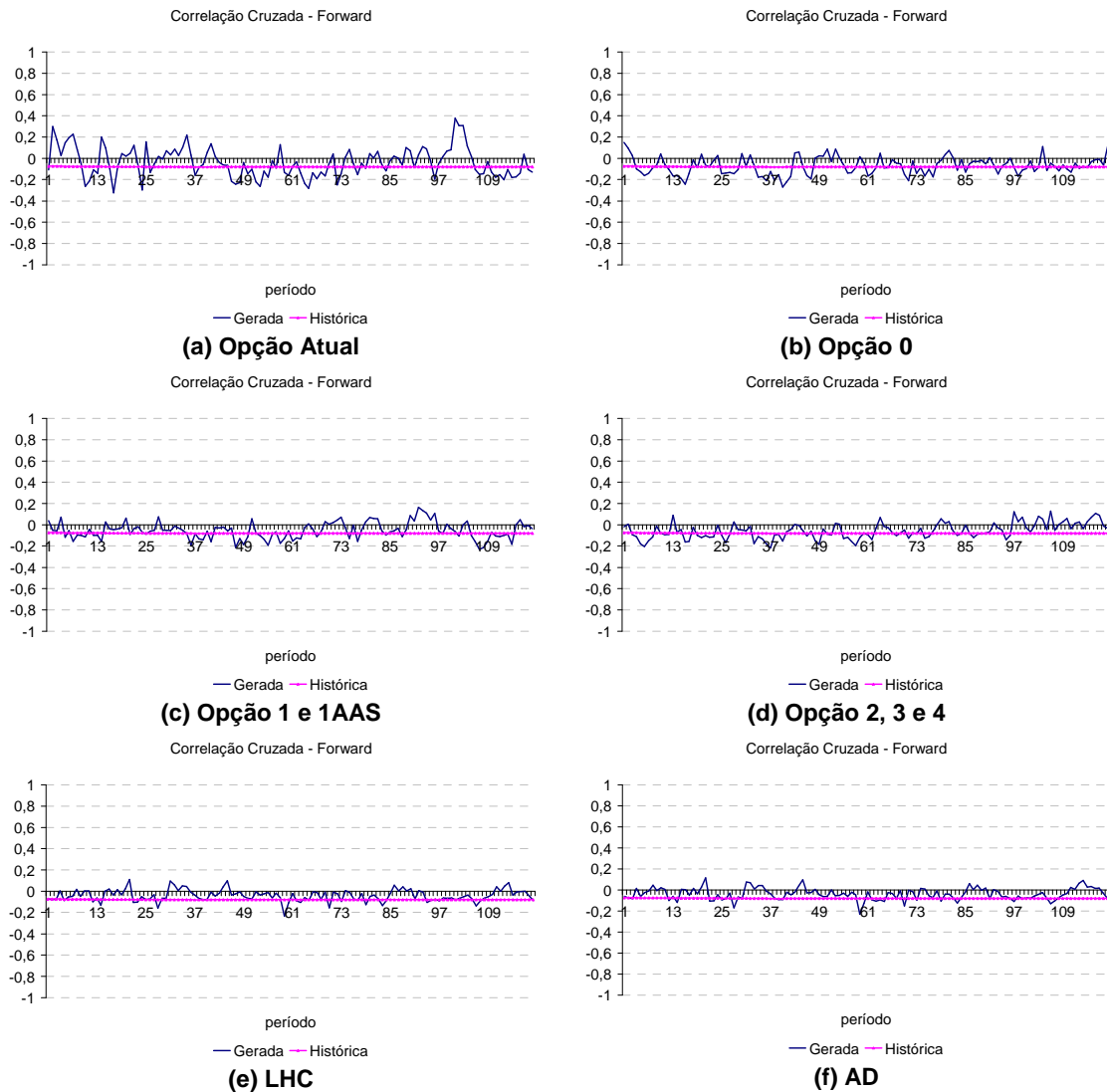
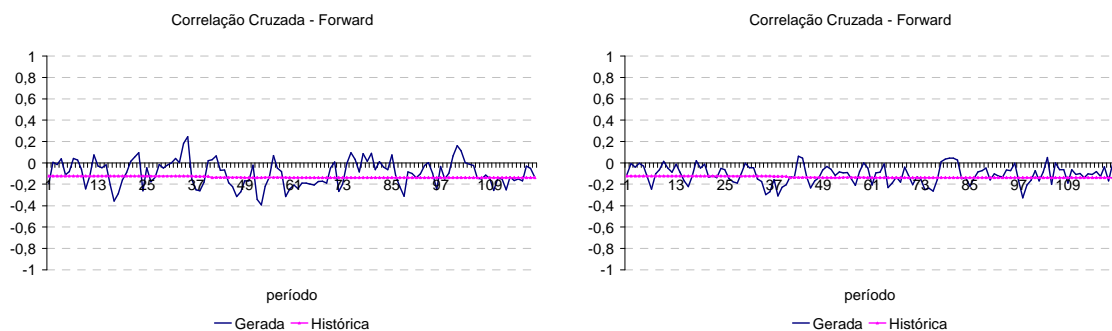


Figura 98: Correlação Cruzada Cenários Forward – Sul x Nordeste

As mesmas observações feitas na análise da correlação cruzada entre Sul e Nordeste valem para a correlação cruzada entre Sul e Norte, Figura 99a até a Figura 99f.



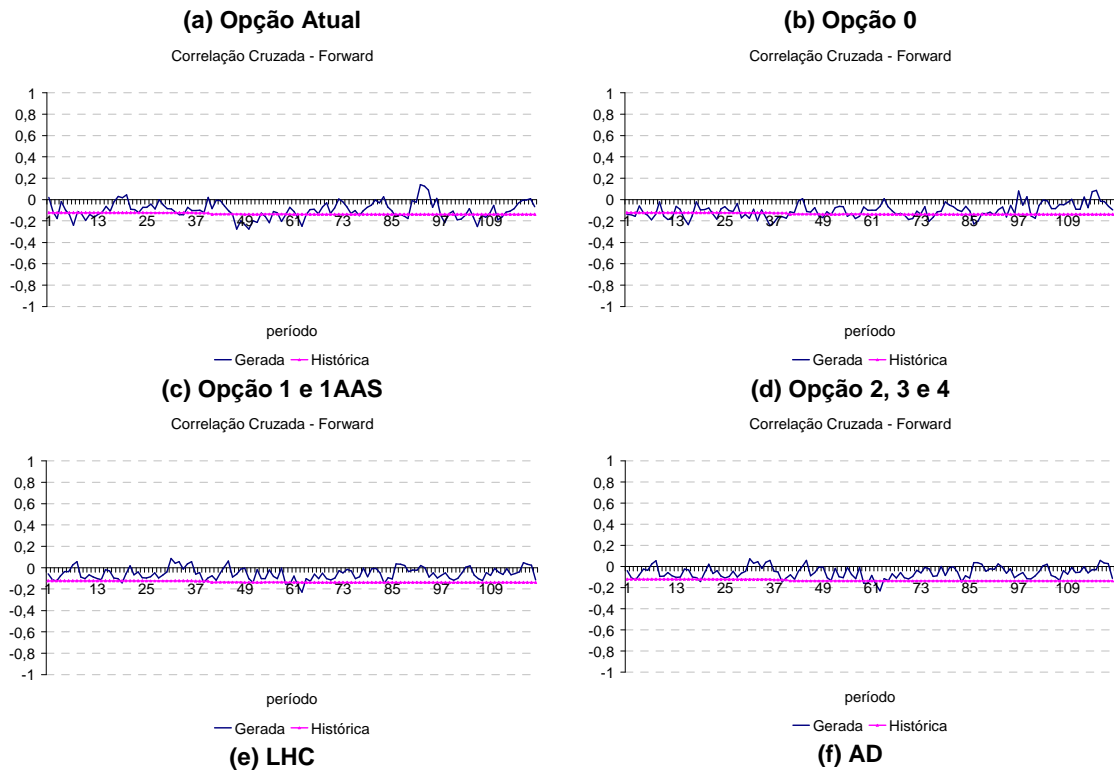
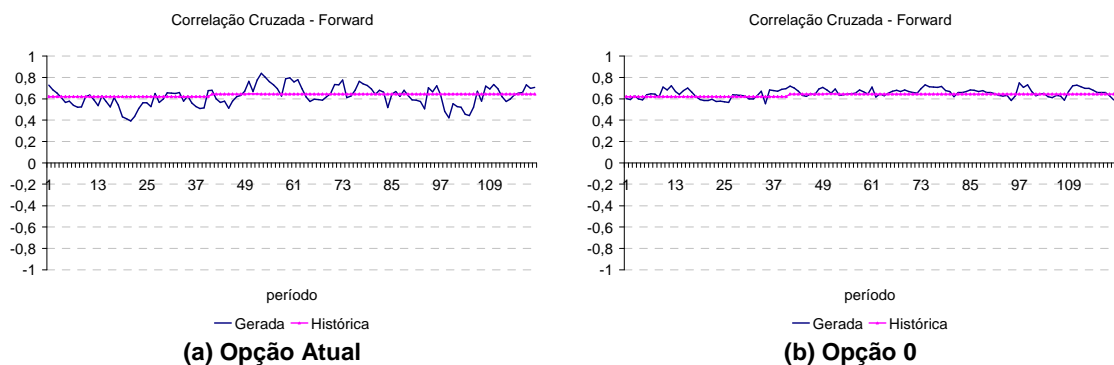


Figura 99: Correlação Cruzada Cenários Forward –Sul x Norte

Da Figura 100a até a Figura 100f é comparada evolução da correlação cruzada entre os subsistemas Nordeste e Norte com o valor observado no histórico. As correlações cruzadas obtidas com as opções LHC e AD permanecem, em geral, abaixo do valor histórico, porém as diferenças relação ao histórico são bem menores. Novamente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico.





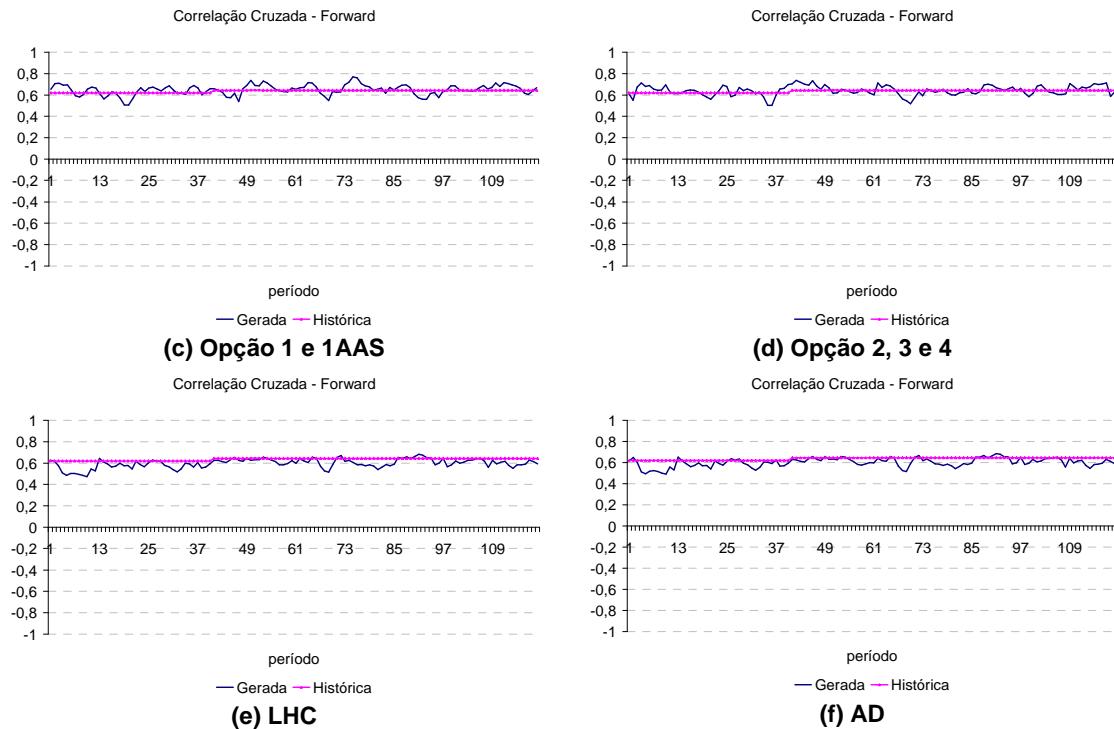


Figura 100: Correlação Cruzada Cenários Forward – Nordeste x Norte

Na Tabela 2 é mostrada a correlação cruzada média ao longo de todo horizonte de estudo para opções analisadas neste item. Conforme o observado nas figuras anteriores, em geral as opções LHC e AD têm dificuldade em preservar o valor da correlação cruzada histórica.

Tabela 2: Correlação Cruzada Média Cenários Forward

OPÇÃO	Sudeste-Sul	Sudeste-Nordeste	Sudeste-Norte	Sul-Nordeste	Sul-Norte	Nordeste-Norte
Histórica	0,4502	0,5709	0,4265	-0,0792	-0,1313	0,6372
Atual	0,4495	0,5534	0,3911	-0,0375	-0,0961	0,6236
0	0,4344	0,5691	0,4332	-0,0711	-0,1089	0,6538
1 e 1AAS	0,4208	0,5594	0,4269	-0,0543	-0,0989	0,6486
2, 3 e 4	0,4207	0,5649	0,4096	-0,0572	-0,1035	0,6422
LHC	0,3710	0,5342	0,4173	-0,0396	-0,0563	0,5972
AD	0,3779	0,5367	0,4188	-0,0371	-0,0569	0,6004

### 7.3.6. Análise de Seqüências Negativas

Neste item é realizada uma avaliação dos cenários sintéticos de energia natural afluyente para o passo forward com o intuito de verificar a reprodução dos longos períodos de seca verificados no registro histórico. Para tanto, são realizados testes de seqüência negativa, testes de máximo déficit e testes relacionado ao período crítico<sup>7</sup>.

Este estudo de caso tem um horizonte de planejamento de cinco anos com um período pós de mais cinco anos e um período pré de dez anos, totalizando um horizonte de estudo de 20 anos. É considerada a configuração estática de janeiro de 2008, ou seja, as usinas hidroelétricas e térmicas em operação nesta data são consideradas estáticas ao longo dos vinte anos e neste período não ocorre a entrada em operação de nenhuma outra usina. O mercado e o intercâmbio de janeiro de 2008 são mantidos constantes para todo o horizonte de estudo. São utilizados 200 cenários hidrológicos para o passo forward.

Vale ressaltar que os cenários forward gerados com as opções 2, 3 e 4 são idênticos. O mesmo ocorre com os cenários gerados utilizando as opções 1 e 1AAS.

Na Tabela 3 são apresentados os resultados obtidos para o teste de aderência (coluna VT) das distribuições da soma, comprimento e intensidade de seqüência negativa para os cenários hidrológicos gerados com a opção Atual. Na coluna VC 95% é indicado o valor crítico do teste com nível de aceitação de 5%. Para as variáveis contínuas soma e intensidade de seqüência negativa é utilizado o teste Kolmogorov-Smirnov, e para a variável discreta comprimento da seqüência negativa é utilizado o teste Qui-Quadrado, conforme descrito no item 4.4.1. Quando o valor do teste (coluna VT) for menor que o valor crítico (coluna VC), as duas distribuições são consideradas aderentes.

Tabela 3: Análise das Seqüências Negativas – Opção Atual

---

<sup>7</sup> O período crítico de um sistema é definido como o período compreendido entre o instante em que o sistema está cheio (armazenamento igual a 100%) e o instante em que o sistema está vazio (reservatório vazio), sem que haja reenchimentos intermediários do reservatório.

Soma Seq. Neg. Teste KS			Comprimento Seq. Neg. Teste Qui-quadrado			Intensidade Seq. Neg. Teste KS		
	VC 95%	VT		VC 95%	VT		VC 95%	VT
<b>SE</b>	0,139	0,069	<b>SE</b>	12,600	2,736	<b>SE</b>	0,139	0,113
<b>S</b>	0,125	0,044	<b>S</b>	14,100	3,246	<b>S</b>	0,125	0,065
<b>NE</b>	0,149	0,058	<b>NE</b>	11,100	4,703	<b>NE</b>	0,149	0,122
<b>N</b>	0,161	0,080	<b>N</b>	12,600	1,926	<b>N</b>	0,161	0,141

Analisando a Tabela 4 é possível verificar que a opção Atual é capaz de gerar cenários hidrológicos com seqüências negativas mais críticas do que a pior seqüência negativa do histórico. No subsistema Norte nenhum cenário gerado apresentou uma seqüência negativa com soma e comprimento mais críticos do que aquela observada no histórico, porém foram gerados cenários com períodos secos de intensidade superior ao pior caso já observado.

Tabela 4: Teste de Máximos – Opção Atual

Teste de Máximos P(G<H)			
	Soma	Comp.	Intens.
<b>SE</b>	73%	62%	39%
<b>S</b>	67%	48%	72%
<b>NE</b>	72%	92%	32%
<b>N</b>	100%	100%	56%

Na Figura 101a-c são apresentados os resultados dos testes de máximo déficit, máximo comprimento de período crítico e a vazão média durante este período. Nestes testes é verificado se os cenários gerados são capazes de reproduzir valores superiores ao máximo valor observado no histórico para estas variáveis. As barras indicam a probabilidade do valor da variável analisada (máximo déficit, máximo comprimento de período crítico e a vazão média) ser menor do que o valor observado no histórico.

Os testes são realizados para diversos níveis de regularização (alfa). Pode-se observar que, para os valores de alfa entre 0,7 e 0,85 (valores mais utilizados na prática), há sempre uma probabilidade de gerar valores maiores do que o máximo observado no histórico.

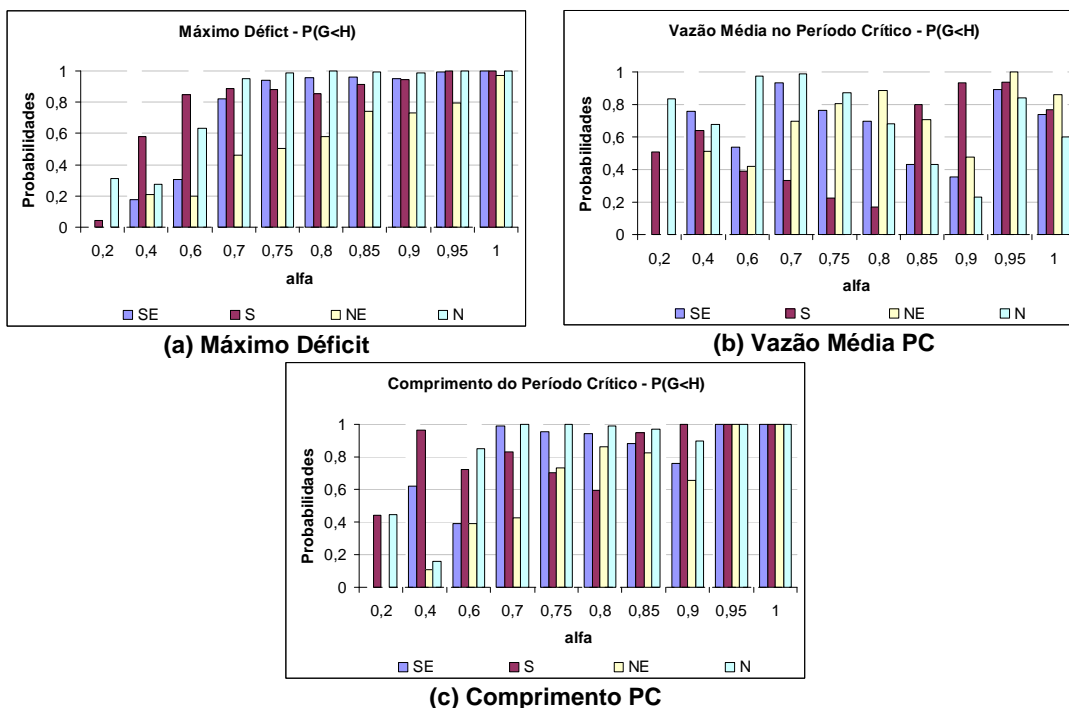


Figura 101: Teste com Diferentes Níveis de Regularização – Opção Atual

Na Tabela 5, Tabela 6 e Tabela 7 são apresentados os resultados dos testes de aderência para as variáveis relacionadas à seqüência negativa, para as opções 0, 1 e 4, respectivamente. Assim como verificado para a opção Atual, as distribuições obtidas pelos cenários hidrológicos gerados por todas as opções são aderentes à distribuição histórica.

Tabela 5: Análise das Seqüências Negativas – Opção 0

	Soma Seq. Neg. Teste KS		Comprimento Seq. Neg. Teste Qui-quadrado		Intensidade Seq. Neg. Teste KS			
	VC 95%	VT		VC 95%	VT	VC 95%	VT	
<b>SE</b>	0,139	0,057	<b>SE</b>	12,600	1,901	<b>SE</b>	0,139	0,055
<b>S</b>	0,125	0,058	<b>S</b>	14,100	3,891	<b>S</b>	0,125	0,078
<b>NE</b>	0,149	0,079	<b>NE</b>	11,100	4,350	<b>NE</b>	0,149	0,085
<b>N</b>	0,161	0,092	<b>N</b>	12,600	1,612	<b>N</b>	0,161	0,113

Tabela 6: Análise das Seqüências Negativas – Opções 1 e 1AAS

Soma Seq. Neg. Teste KS			Comprimento Seq. Neg. Teste Qui-quadrado			Intensidade Seq. Neg. Teste KS		
	VC 95%	VT		VC 95%	VT		VC 95%	VT
<b>SE</b>	0,139	0,062	<b>SE</b>	12,600	2,368	<b>SE</b>	0,139	0,102
<b>S</b>	0,125	0,049	<b>S</b>	14,100	1,605	<b>S</b>	0,125	0,063
<b>NE</b>	0,149	0,063	<b>NE</b>	11,100	4,311	<b>NE</b>	0,149	0,127
<b>N</b>	0,161	0,085	<b>N</b>	12,600	2,547	<b>N</b>	0,161	0,147

Tabela 7: Análise das Seqüências Negativas – Opções 2, 3 e 4

Soma Seq. Neg. Teste KS			Comprimento Seq. Neg. Teste Qui-quadrado			Intensidade Seq. Neg. Teste KS		
	VC 95%	VT		VC 95%	VT		VC 95%	VT
<b>SE</b>	0,139	0,057	<b>SE</b>	12,600	2,281	<b>SE</b>	0,139	0,076
<b>S</b>	0,125	0,046	<b>S</b>	14,100	3,094	<b>S</b>	0,125	0,057
<b>NE</b>	0,149	0,063	<b>NE</b>	11,100	3,580	<b>NE</b>	0,149	0,101
<b>N</b>	0,161	0,084	<b>N</b>	12,600	2,330	<b>N</b>	0,161	0,130

O complemento dos valores apresentados na Tabela 8, Tabela 9 e Tabela 10 indicam a probabilidade de ocorrência de eventos mais severos do que os observados no histórico. Novamente, para todos os subsistemas, exceto o Norte nas variáveis soma e comprimento, todas as opções geram cenários mais críticos do que o observado no histórico.

Tabela 8: Teste de Máximos – Opção 0

Teste de Máximos P(G<H)			
	Soma	Comp.	Intens.
<b>SE</b>	77%	57%	70%
<b>S</b>	61%	29%	86%
<b>NE</b>	71%	79%	38%
<b>N</b>	100%	100%	79%

Tabela 9: Teste de Máximos – Opções 1 e 1AAS

Teste de Máximos P(G<H)			
	Soma	Comp.	Intens.
<b>SE</b>	78%	66%	45%
<b>S</b>	61%	43%	72%
<b>NE</b>	70%	59%	24%
<b>N</b>	100%	100%	59%

Tabela 10: Teste de Máximos – Opções 2, 3 e 4

**Teste de Máximos  
P(G<H)**

	Soma	Comp.	Intens.
<b>SE</b>	75%	61%	48%
<b>S</b>	59%	43%	79%
<b>NE</b>	66%	85%	33%
<b>N</b>	100%	100%	63%

Analisando os testes com relação às variáveis máximo déficit, máximo período crítico e vazão média durante o período crítico, mostrados na Figura 102a-c, Figura 103a-c e Figura 104a-c respectivamente, pode-se dizer os resultados obtidos são similares e, portanto, todas as opções analisadas são capazes de gerar eventos mais críticos que o verificado no histórico.

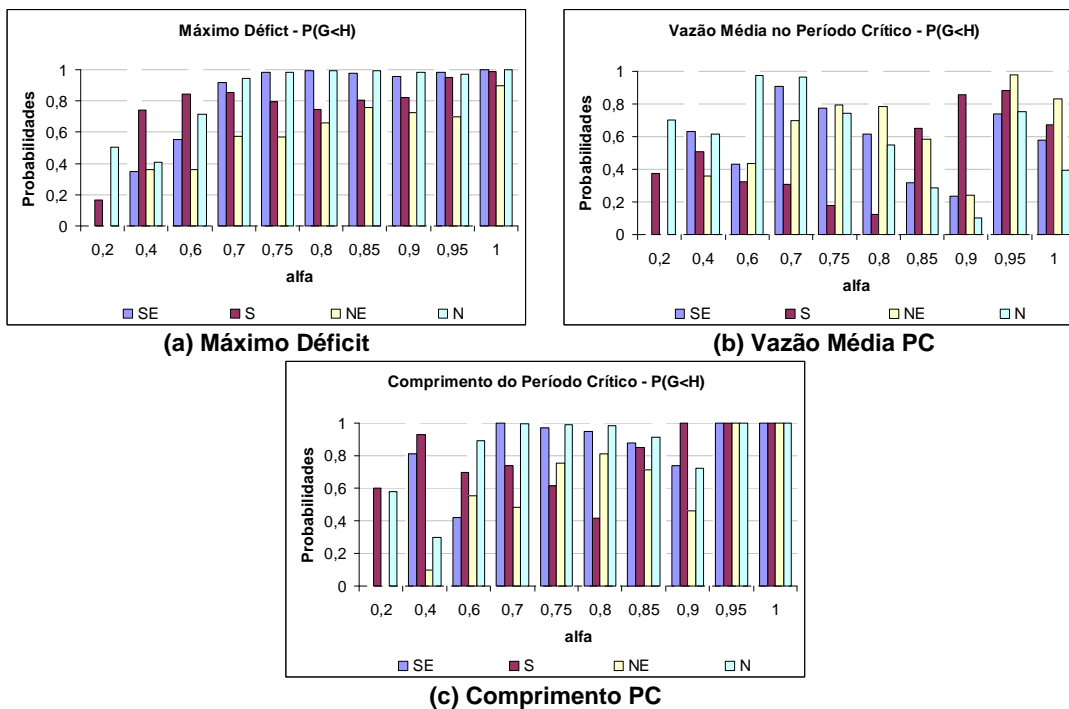


Figura 102: Teste com Diferentes Níveis de Regularização – Opção 0

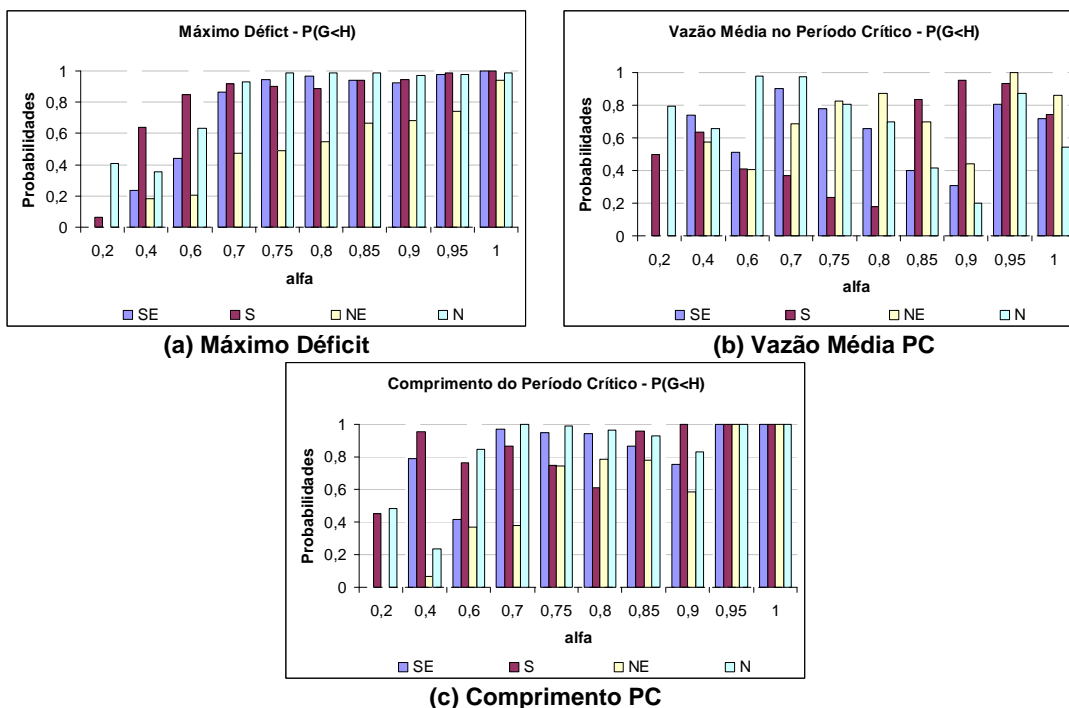


Figura 103: Teste com Diferentes Níveis de Regularização – Opções 1 e 1AAS

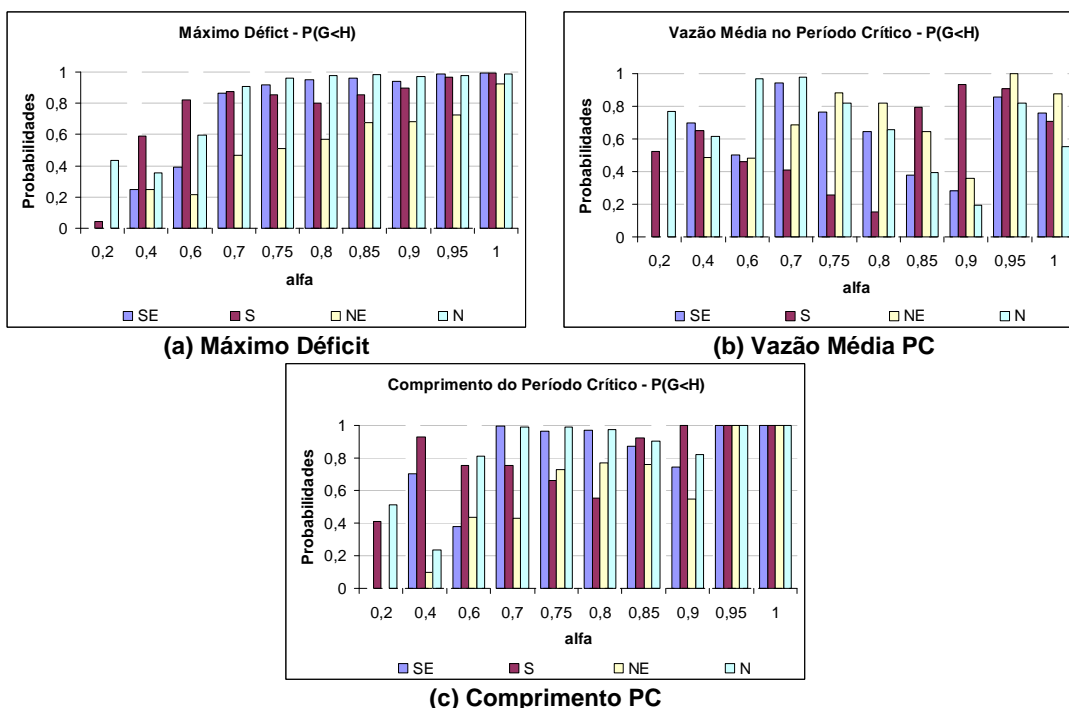


Figura 104: Teste com Diferentes Níveis de Regularização – Opções 2, 3 e 4

É sempre desejável que o modelo escolhido para representar um processo estocástico seja capaz de gerar eventos mais favoráveis e mais críticos do que os observados no registro histórico. Analisando os resultados dos testes aplicados neste

item, pode-se concluir que todas as opções analisadas para gerar cenários hidrológicos são capazes de reproduzir os eventos críticos ocorridos no histórico.

#### 7.4. Análise dos Cenários Gerados para Passo Backward

Neste item é apresentada uma avaliação dos cenários hidrológicos obtidos para o passo backward a partir da utilização das opções descritas no trabalho, além dos métodos de amostragem LHC e AD. Nesta análise são considerados 200 cenários forward e 20 cenários backward, e para o processo de agregação é fornecida uma amostra de 2000 vetores de ruídos.

São utilizados testes não condicionados para a avaliação da capacidade de preservação das estatísticas média, desvio-padrão e correlação cruzada, e testes para verificação da aderência entre as distribuições gerada e histórica. Para a realização dos testes não condicionados, a cada período são considerados todos os cenários backward gerados, independentemente do estado de afluência passada, conforme ilustrado na Figura 105. Desta forma, neste estudo é considerada uma amostra com 4000 vetores multivariados de ENA (200x20).

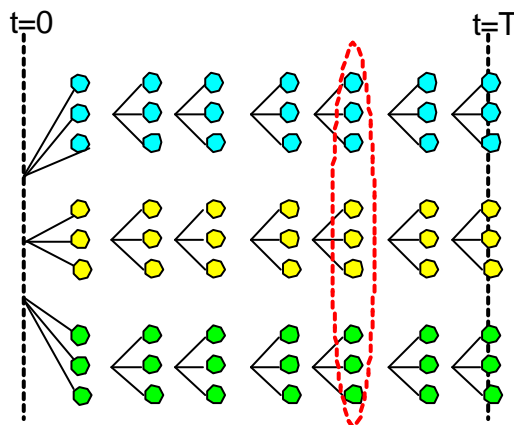


Figura 105: Amostra utilizada no teste não condicionado

Para a avaliação da qualidade dos cenários gerados para cada conjunto de aberturas são realizados testes condicionados. Neste caso, o teste é aplicado a cada estado de afluência passada, conforme mostrado na Figura 106. Neste estudo são realizados a cada período 200 testes considerando uma amostra de 20 vetores multivariados de ENA.



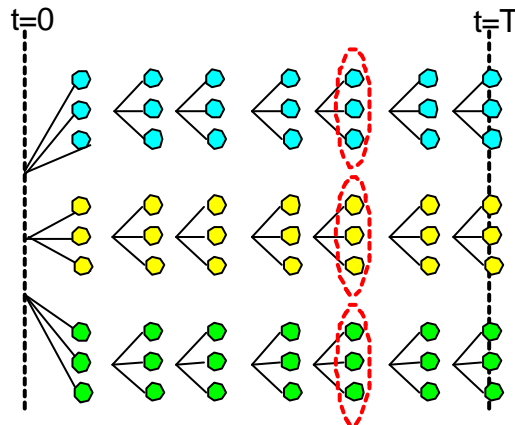
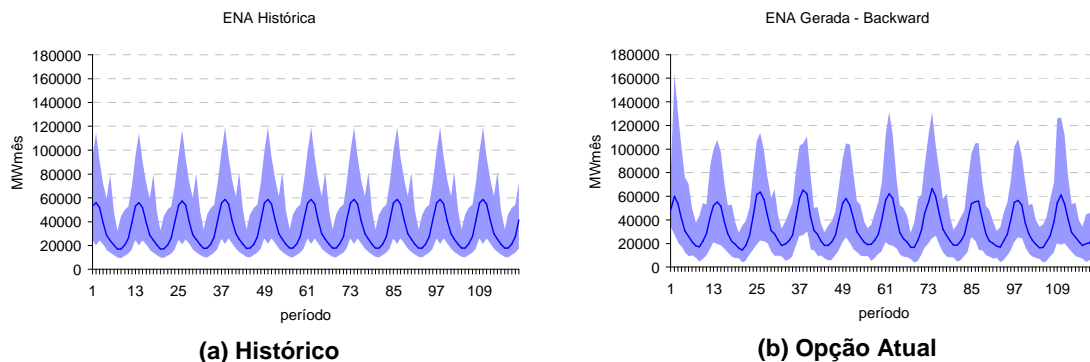


Figura 106: Amostras utilizadas no teste condicionado

Para iniciar a avaliação dos cenários de ENA é realizada uma comparação entre as envoltórias dos cenários hidrológicos históricos e gerados com todas as opções e métodos de amostragem descritos.

#### 7.4.1. Envoltória dos Cenários

Da Figura 107b até a Figura 107g são mostradas as envoltórias e os valores médios dos 4000 cenários gerados de energia natural afluyente para o subsistema Sudeste, e na Figura 107a é apresentado o histórico. Observe que todas as opções de geração avaliadas são capazes de reproduzir bem o histórico, exceto um valor máximo atípico ocorrido no mês de junho. As opções Atual e LHC apresentam um valor máximo bem maior que o histórico para o mês de fevereiro do primeiro e sexto anos, respectivamente. Por se tratar de apenas um cenário atípico, o valor da média não é muito afetado.



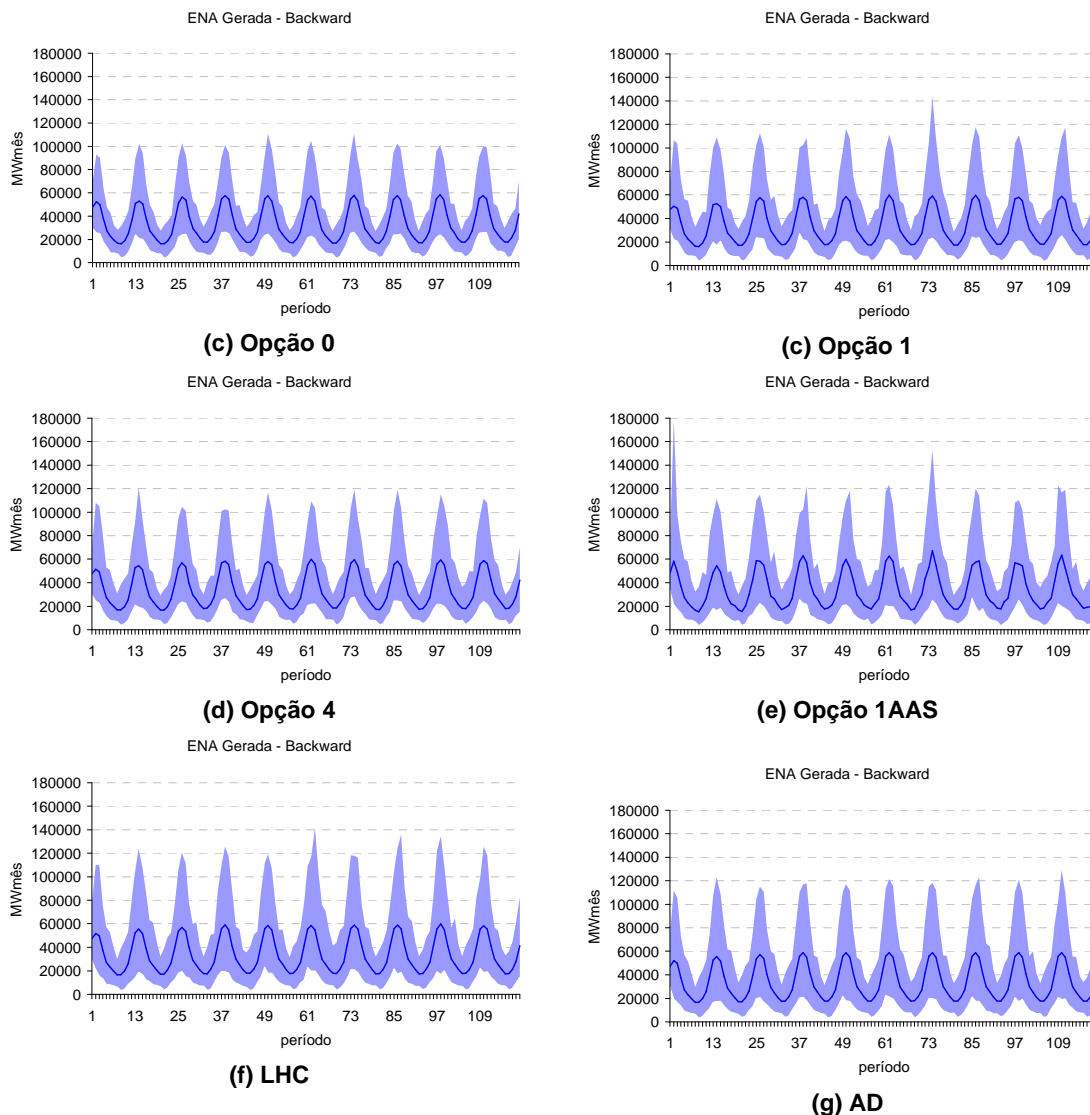


Figura 107: Envoltória Cenários Forward – Sudeste

Da Figura 108b até a Figura 108g são apresentadas as envoltórias e os valores médios para o subsistema Sul, assim como seu valor histórico, Figura 108a. O Sul é o subsistema que apresenta maior variabilidade no seu comportamento hidrológico, seu coeficiente de variação mensal é bem próximo de um. O valor máximo de ENA observado no Sul foi no mês de julho de 1983, quando ocorreu uma grande cheia na região. Excluindo o valor máximo de julho, pode-se observar que as opções 2, 3, 4, LHC e AD são as que mais se aproximam da envoltória histórica.

Nota-se que o valor médio dos cenários gerados não é preservado de maneira satisfatória pela Opção Atual.

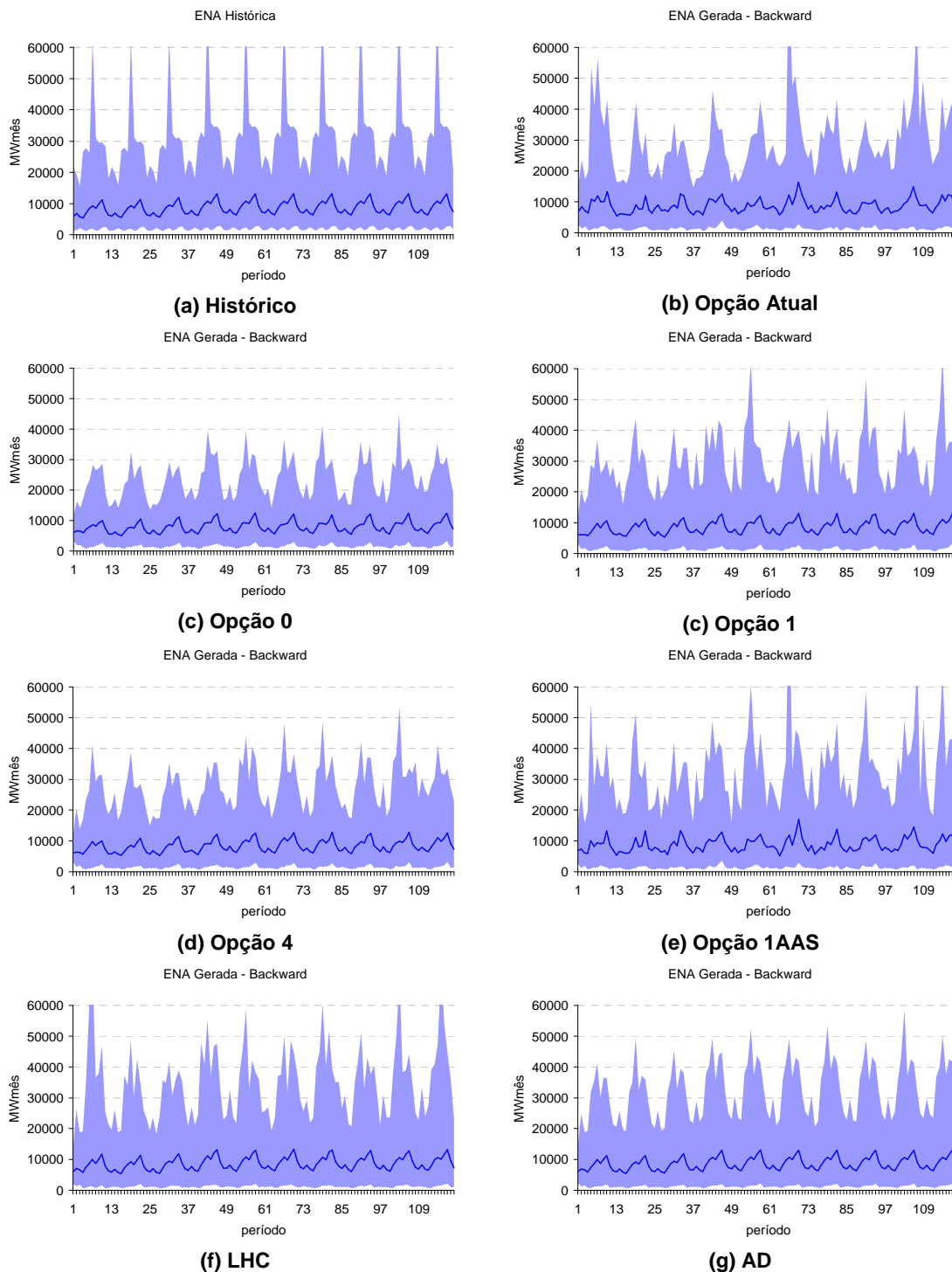


Figura 108: Envoltória Cenários Backward – Sul

A mesma análise é realizada para o subsistema Nordeste, Figura 109a até a Figura 109g. Todas as opções analisadas preservam de forma satisfatória a envoltória histórica. Observa-se na opção 0 uma dificuldade em preservar valores mais elevados.

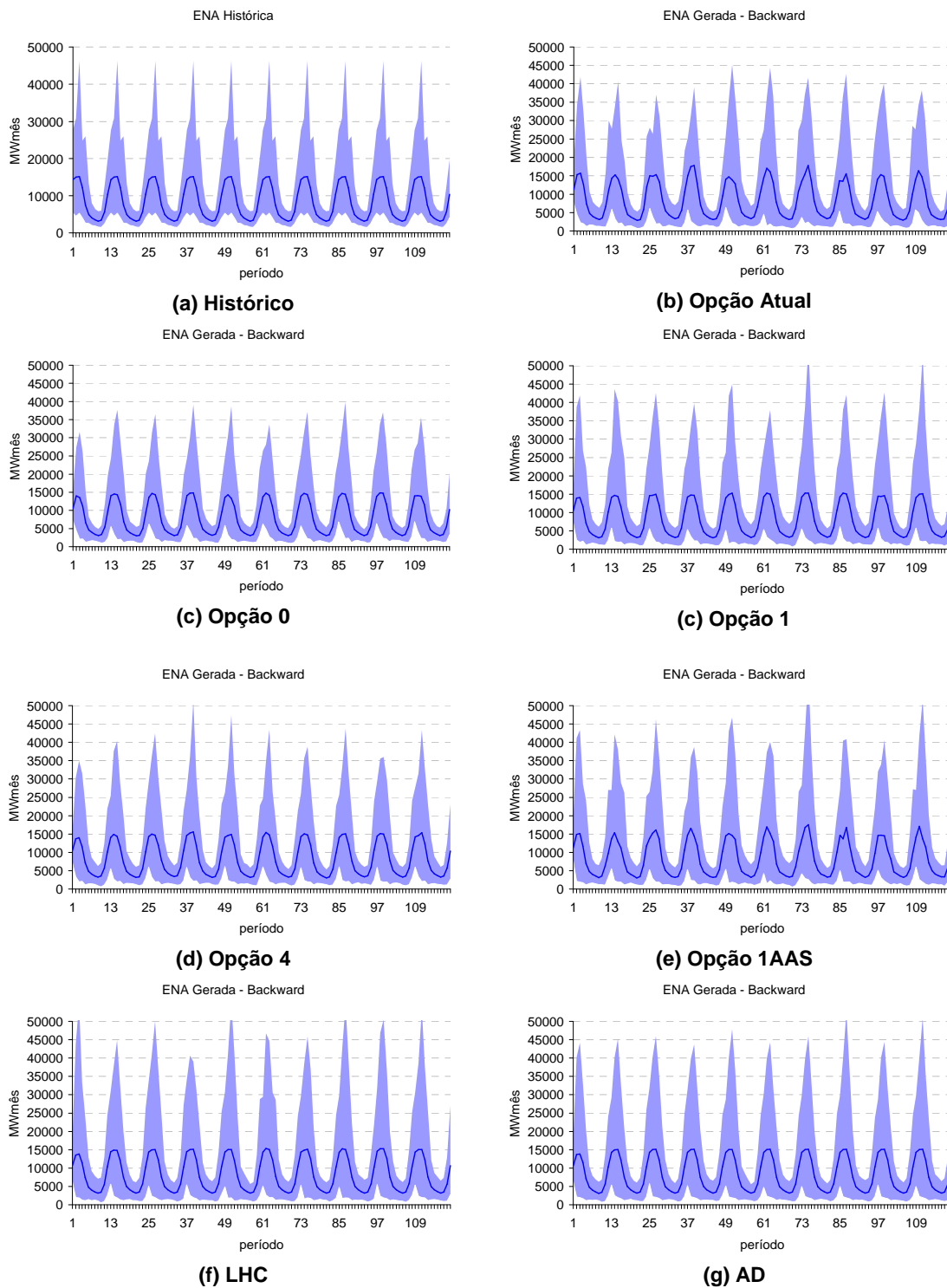


Figura 109: – Envolvória Cenários Backward – Nordeste

Novamente, todas as opções analisadas são capazes de reproduzir satisfatoriamente a envoltória histórica do subsistema Norte, Figura 110a até Figura

110g, exceto pelas opções Atual e 0 que apresentam um pouco de dificuldade para representar valores elevados.

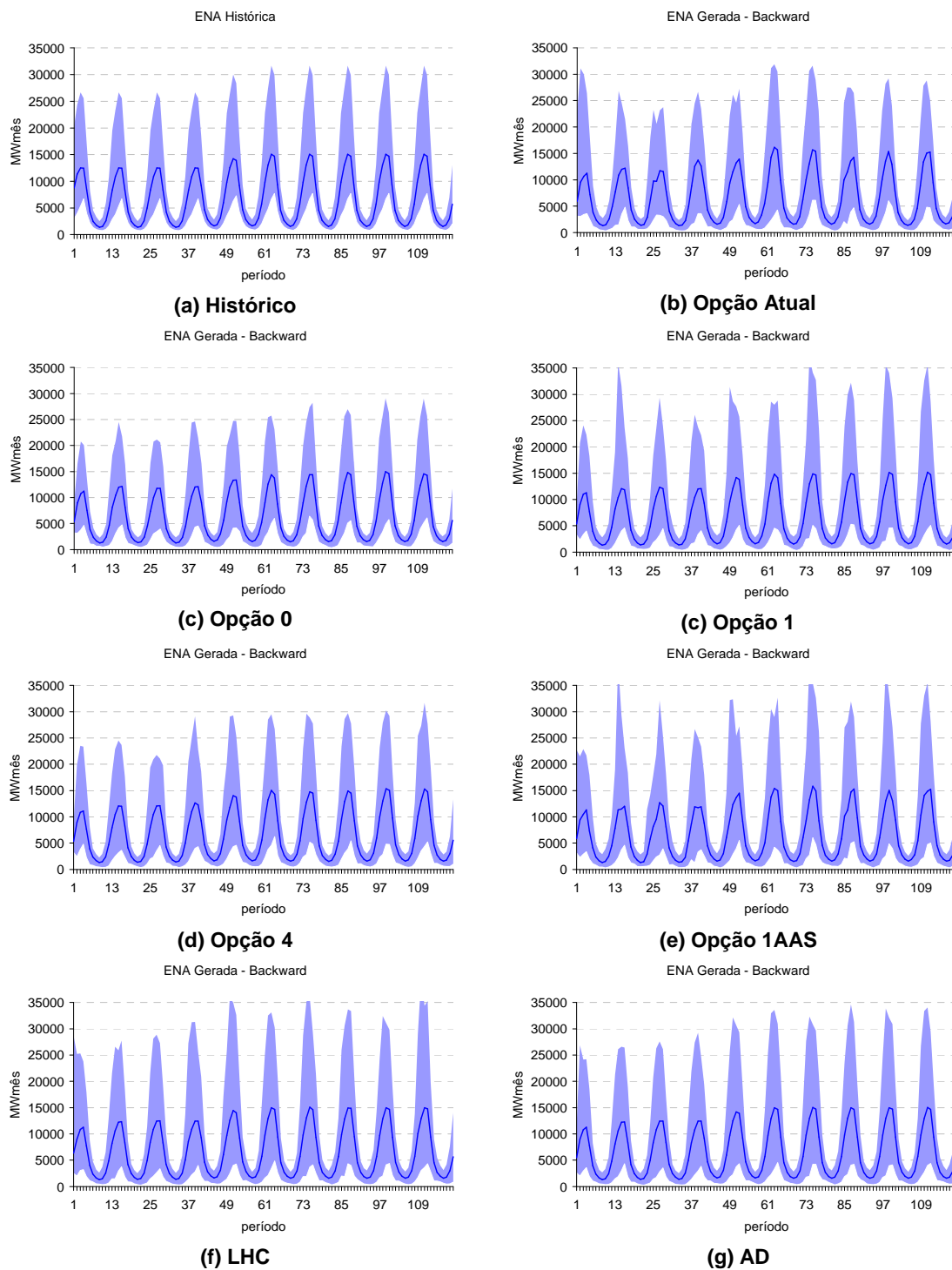


Figura 110: Envoltória Cenários Backward – Norte

#### **7.4.2. Testes não condicionados para média e desvio**

Neste item é realizada uma análise mais detalhada da média e do desvio-padrão dos cenários hidrológicos. O nível de significância do teste é de 5% e como parâmetros de população são considerados os valores históricos. Na primeira coluna da Figura 111a até a Figura 111g são apresentadas as evoluções temporais da média e do desvio padrão para a série histórica e para os cenários gerados, para o subsistema Sudeste.

Os cenários gerados para os primeiros períodos são fortemente condicionados ao passado recente (tendência hidrológica recente). Desta forma, as médias e os desvios-padrão dos primeiros períodos diferem do valor histórico, uma vez que o passado recente é diferente da MLT, conforme mostrado na Figura 46.

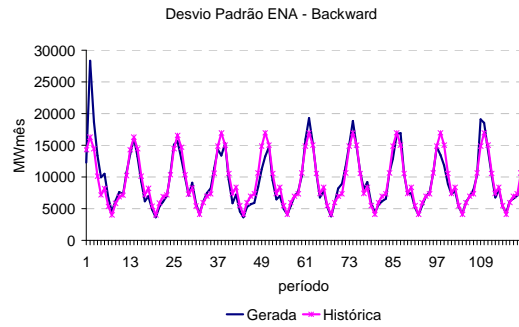
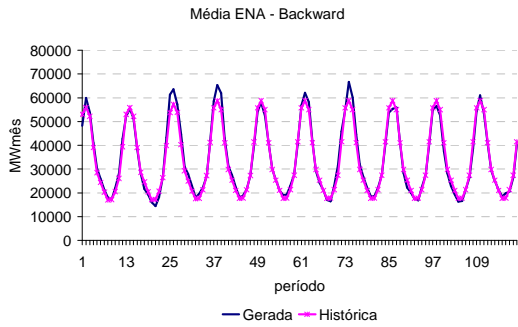
Vale ressaltar que as estimativas não condicionadas da média e desvio-padrão dos cenários backward são influenciadas tanto pela qualidade da geração dos cenários backward, propriamente ditos, quanto pela geração dos cenários forward.

Analisando a Figura 111a até a Figura 111g, observa-se que a Opção Atual é aquela que apresenta maior diferença com relação à média histórica. As estimativas de média se aproximam mais da média histórica quando se aplica técnicas de agregação tanto na obtenção da amostra backward (opções 0 e 1), quanto na obtenção da amostra forward (opções 2,3 e 4). Comparando as opções 1AAS, LHC e AD, verifica-se que as técnicas de amostragem estratificadas reduzem bastante a variabilidade da estimativa da média.

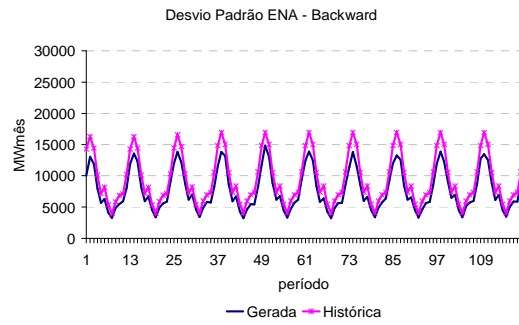
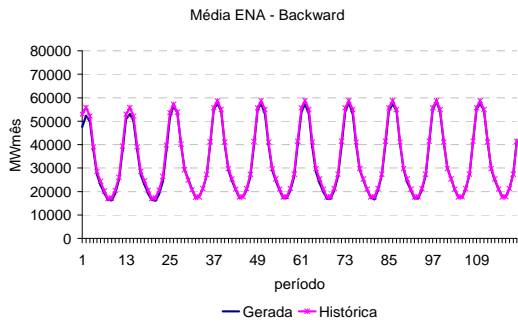
Como pode ser observado, as opções que empregam técnicas de agregação na obtenção dos cenários da backward apresentam uma ligeira degradação com relação ao desvio-padrão. O maior nível de degradação na estimativa do desvio ocorre na opção 0 e parte desta degradação deve-se à degradação observada nos cenários forward.

Também pode-se observado nas opções Atual e 1AAS uma dificuldade em preservar o desvio-padrão histórico, porém as diferenças com relação ao histórico não apresentam uma tendência de decréscimo.

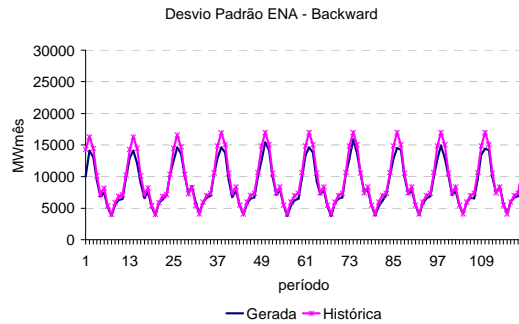
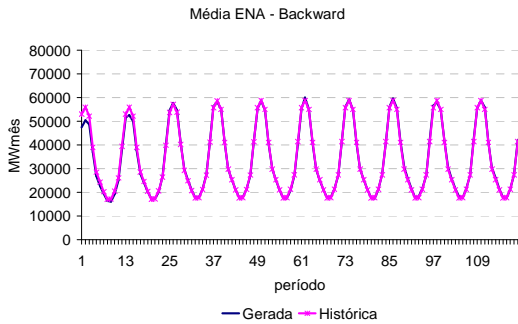
A preservação do desvio histórico é alcançada quase com perfeição nas opções LHC e AD, com exceção dos primeiros meses.



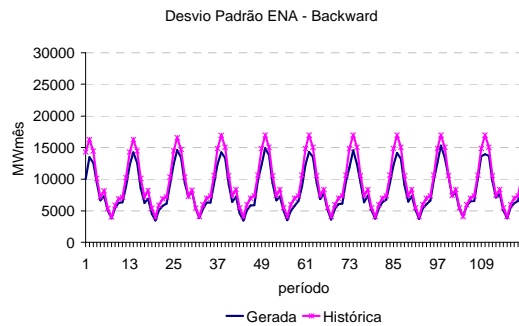
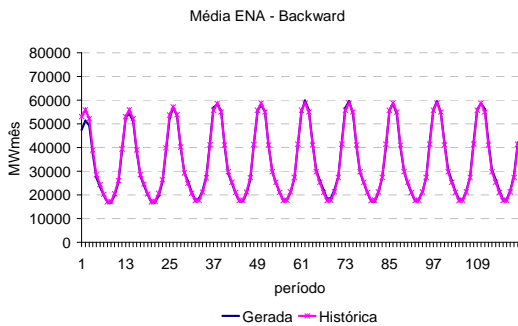
**(a) Opção Atual**



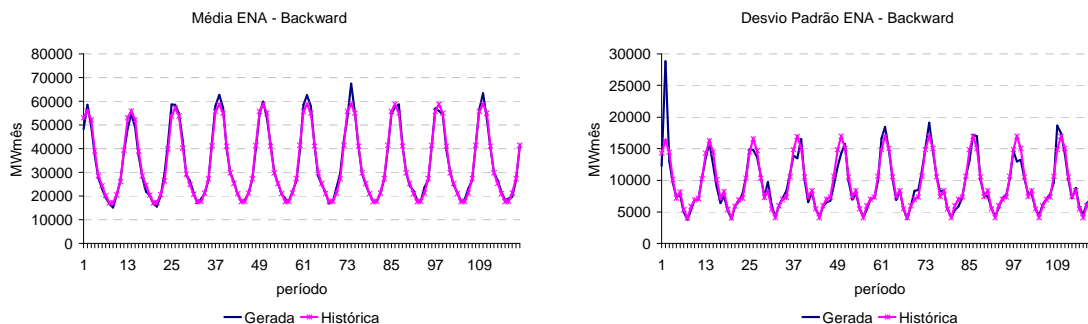
**(b) Opção 0**



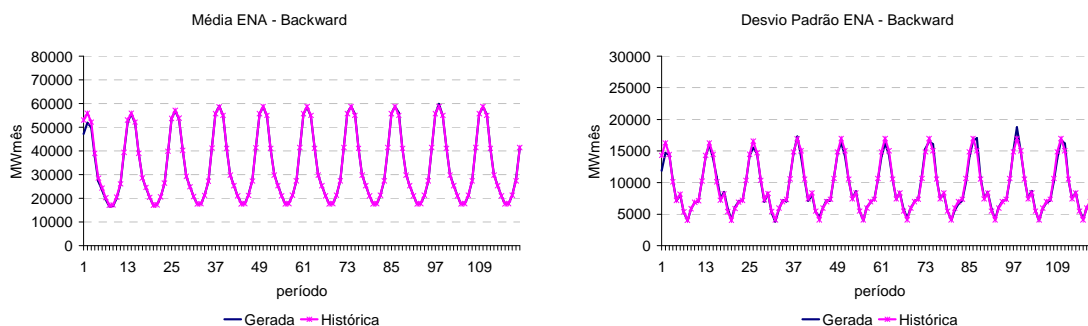
**(c) Opção 1**



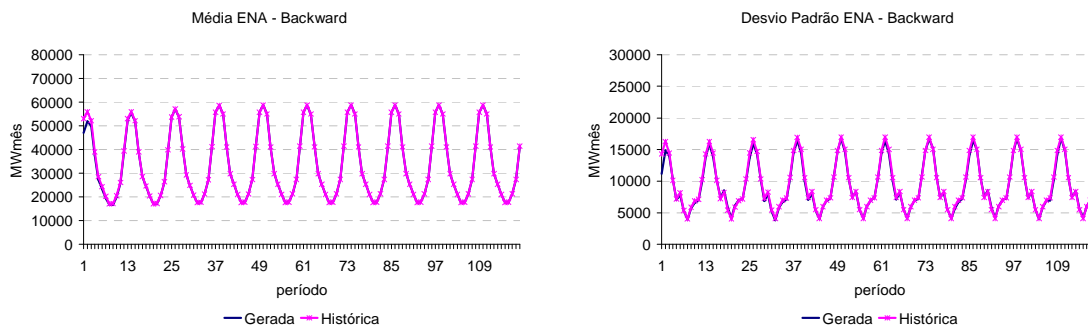
**(d) Opção 4**



**(e) Opção 1AAS**



**(f) Opção LHC**



**(g) Opção AD**

Figura 111: Média Cenários Backward – Sudeste

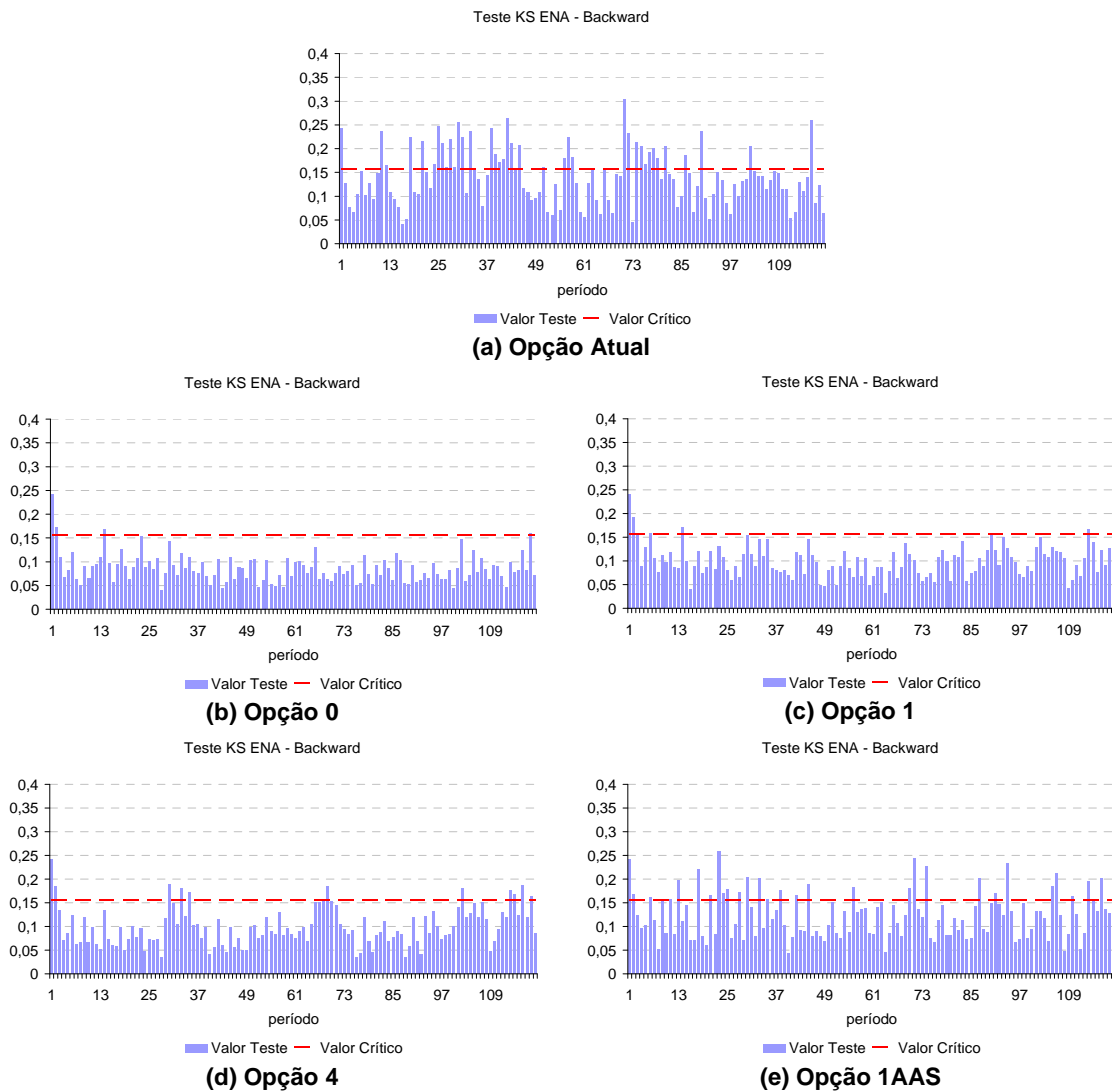
As observações resultantes desta análise também são válidas para os demais subsistemas. As estimativas de média e desvio-padrão para os demais subsistemas podem ser avaliadas no Anexo B.

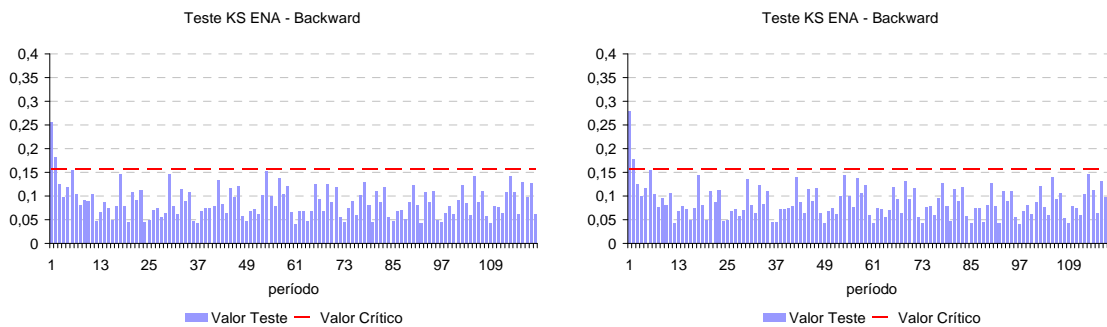
### 7.4.3. Distribuição Univariada dos Cenários

O objetivo desta análise é verificar a aderência entre as distribuições univariadas histórica e gerada a cada período do horizonte de estudo (teste não condicionado). Para tanto é aplicado o teste Kolmogorov-Smirnov para verificar com nível de significância de 5%. O valor crítico do teste é representado por uma linha tracejada nas figuras a seguir.



Da Figura 112a até a Figura 112g são apresentados os resultados do teste KS para os cenários hidrológicos do subsistema Sudeste. Pode-se destacar o grande número de testes rejeitados quando se utiliza a opção Atual. As melhores representações da distribuição univariada do subsistema Sudeste são obtidas com as opções LHC e AD. Um comportamento satisfatório é verificado para as demais opções.



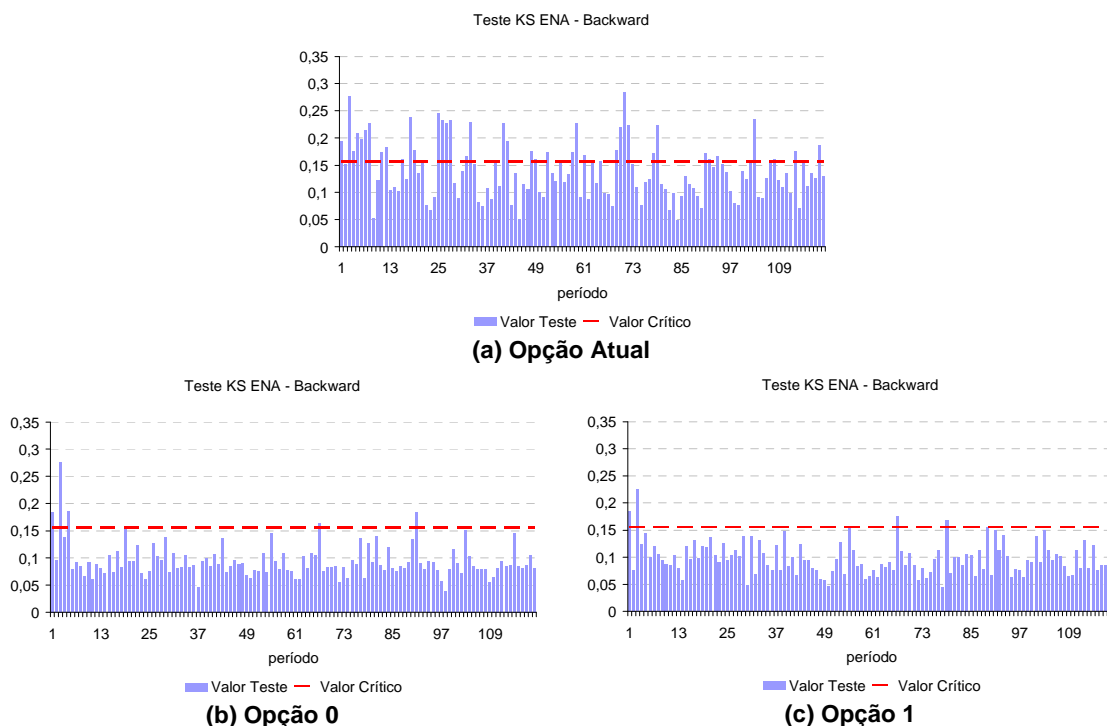


**(f) LHC**

**(g) AD**

Figura 112: Teste de Aderência Cenários Backward – Sudeste

Os resultados obtidos com a aplicação do teste KS para a distribuição univariada de energia natural afluyente do Sul são apresentados da Figura 113a até a Figura 113g. Conforme observado anteriormente, na opção Atual há um grande número de rejeições, porém ao aplicar as técnicas de agregação para obter a amostra backward, a partir da qual é gerada a amostra forward (opção 0), a aderência entre as distribuições histórica e gerada é alcançada. Nas demais opções é observado um bom comportamento com relação ao teste de aderência.



**(a) Opção Atual**

**(b) Opção 0**

**(c) Opção 1**

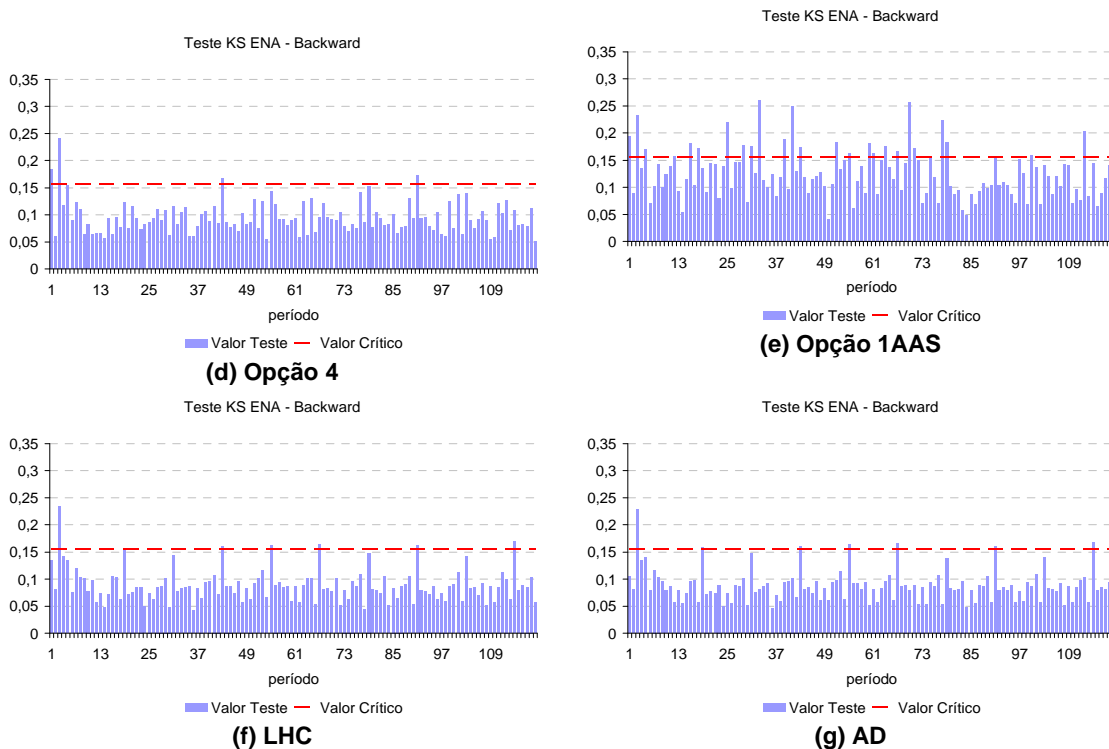
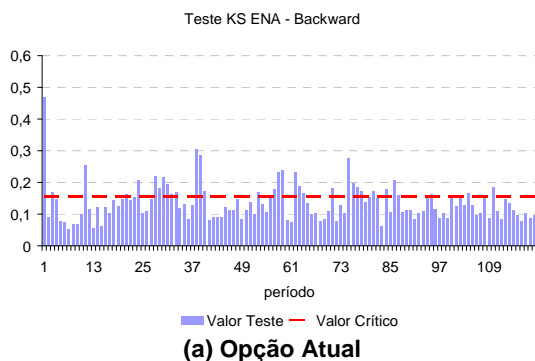


Figura 113: Teste de Aderência Cenários Backward – Sul

Analisando a Figura 114a até a Figura 114g pode-se verificar que, em geral, todas as opções apresentam bom nível de aceitação dos testes de aderência para o subsistema Nordeste. Novamente o maior índice de rejeição é observado na opção Atual, porém em nível menor do que o observado nos subsistemas Sul e Sudeste.



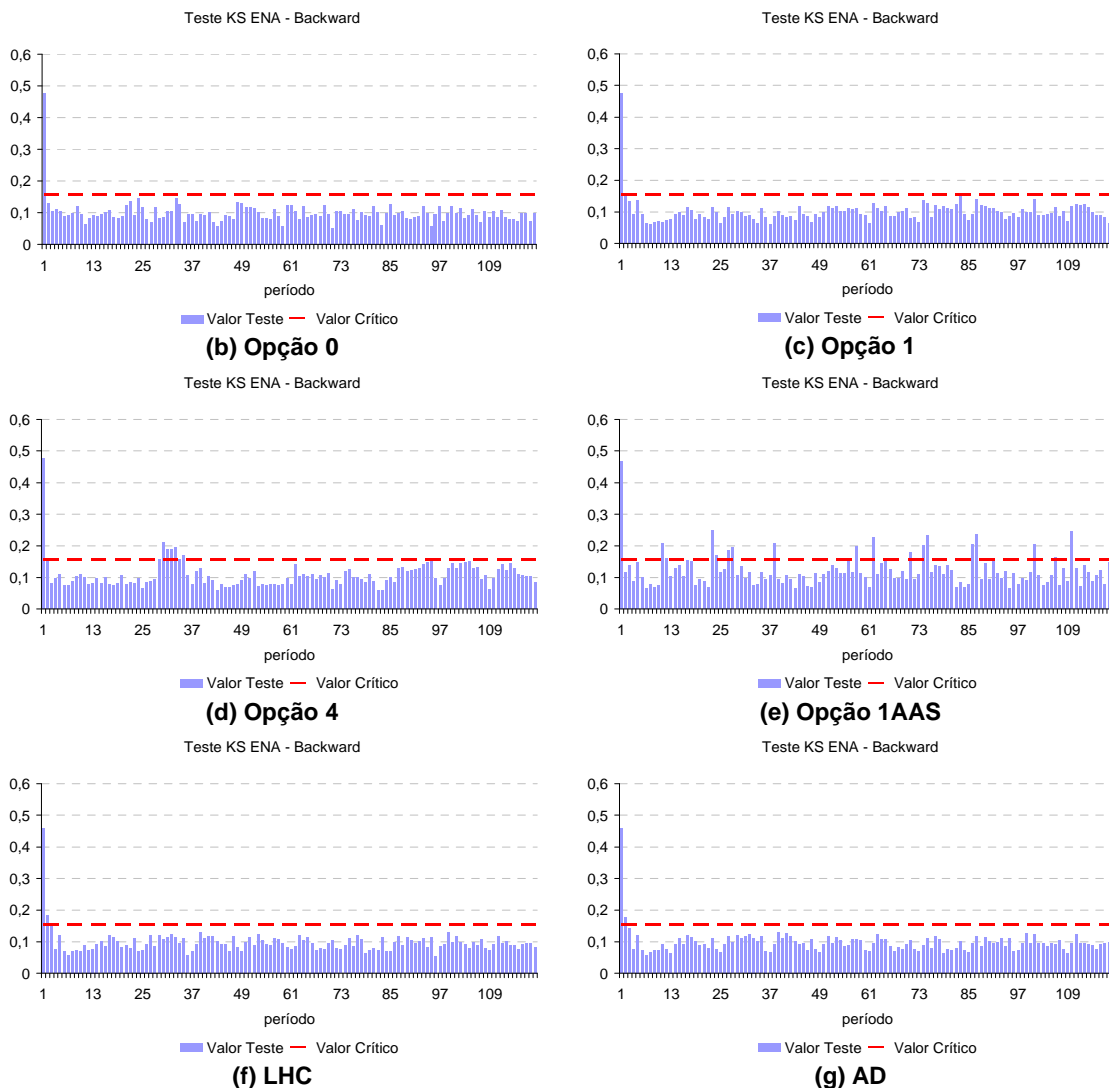
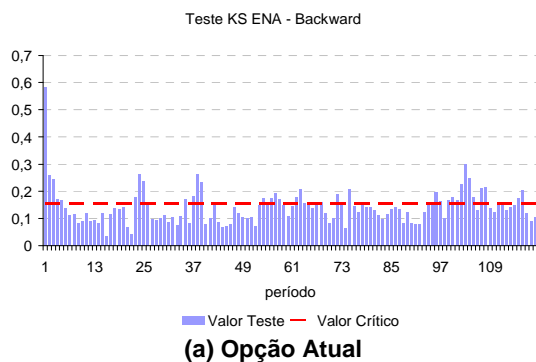


Figura 114: Teste de Aderência Cenários Backward – Nordeste

As mesmas considerações feitas para o subsistema Nordeste se aplicam ao subsistema Norte, como pode ser verificado da Figura 115a até a Figura 115g.



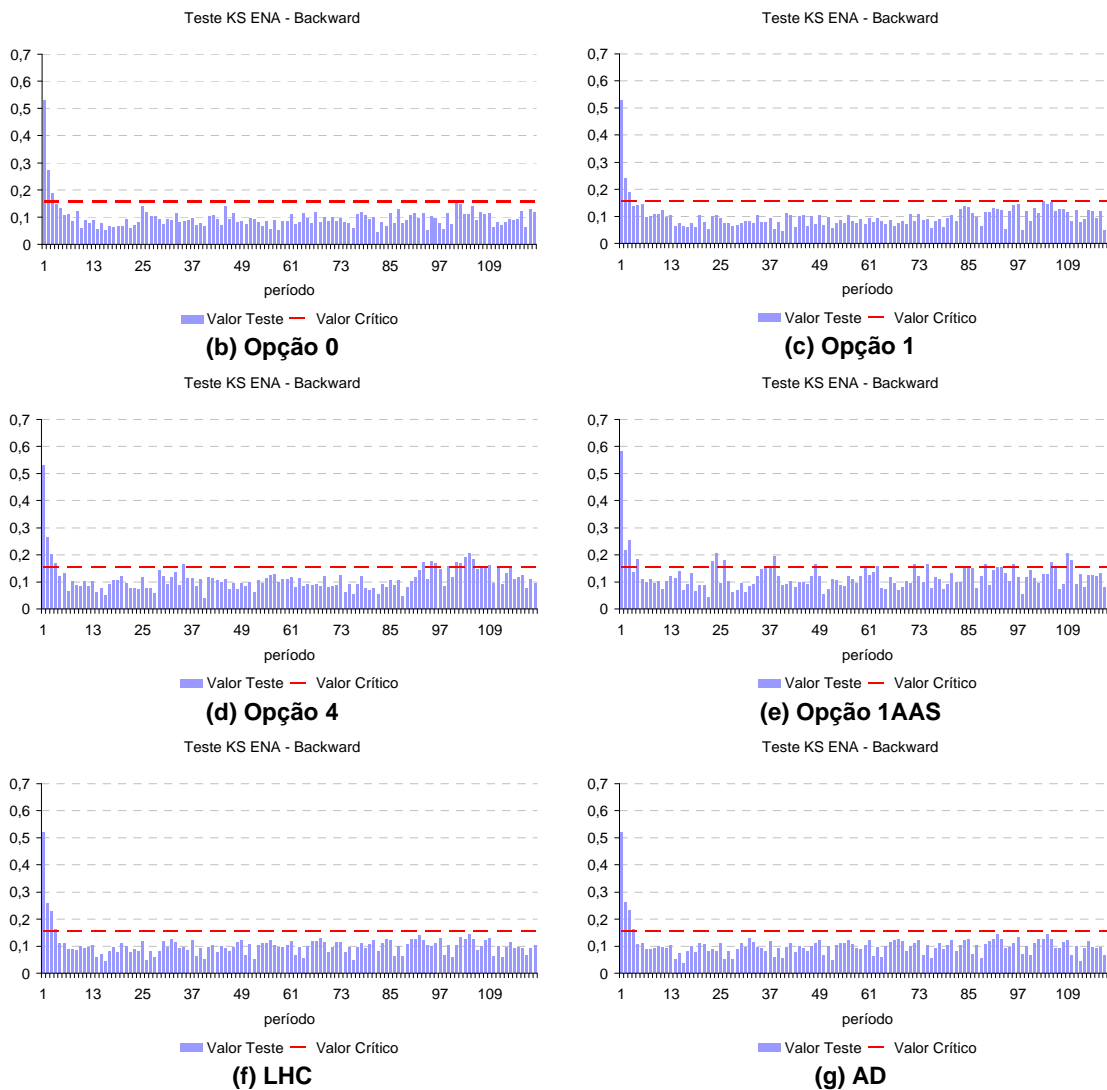


Figura 115: Teste de Aderência Cenários Backward – Norte

#### 7.4.4. Testes condicionados para média e desvio

A seguir são realizados os testes condicionados para média e desvio padrão dos cenários gerados para o passo backward. Os parâmetros de população considerados nos testes são os valores teóricos calculados de acordo como o item 4.5.2. O nível de significância adotado é de 5%, e o intervalo de confiança do teste está destacado nos gráficos por linhas pretas.

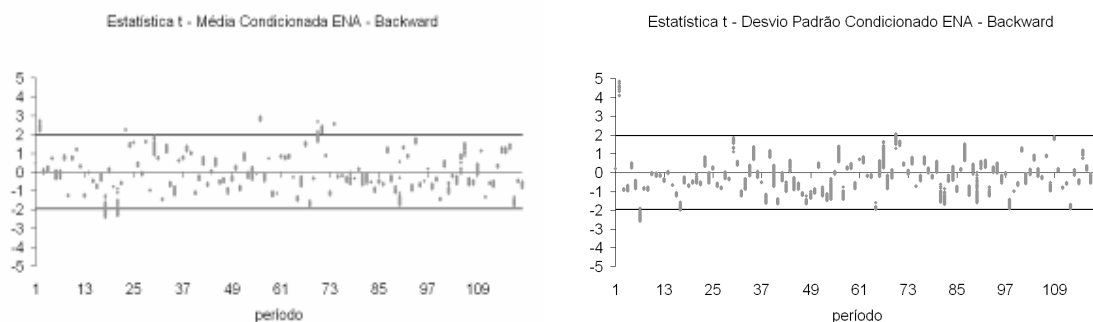
Para cada período são calculadas as estatísticas t para todos os conjuntos de aberturas. Neste caso, são considerados 200 conjuntos de aberturas por período.

Como pode ser constatado da Figura 116a até a Figura 116g, as opções Atual e 1AAS apresentam uma maior variabilidade com relação às estimativas da média e do

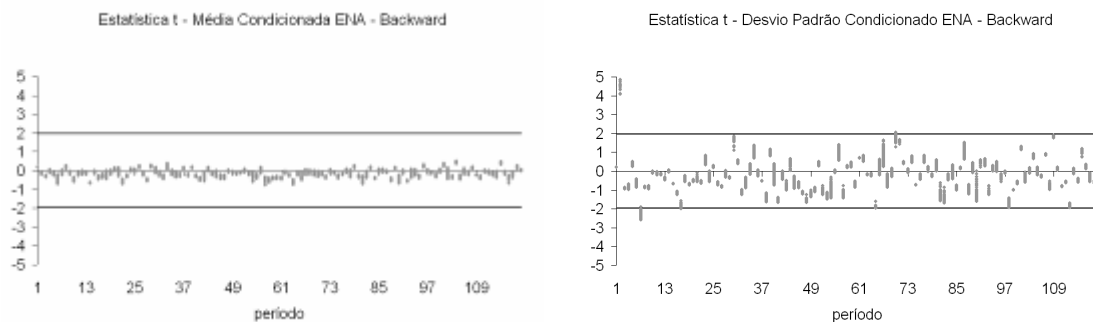
desvio-padrão. Isto ocorre porque o tamanho da amostra é relativamente pequeno (igual a 20 neste caso). Esta variabilidade tende a diminuir quando o tamanho da amostra aumenta, mas se o número de cenários backward for grande pode-se tornar inviável do ponto de vista computacional a resolução do problema de planejamento da operação energética.

Nota-se que mesmo utilizando uma amostra de tamanho reduzido, ao aplicar técnicas de agregação ou métodos de amostragem estratificada a variabilidade das estimativas de média e desvio se reduz bastante, exceto para o desvio-padrão da opção 0. A variabilidade do desvio-padrão obtida com a opção 0 está relacionada à degradação já existente nos cenários forward gerados por esta opção.

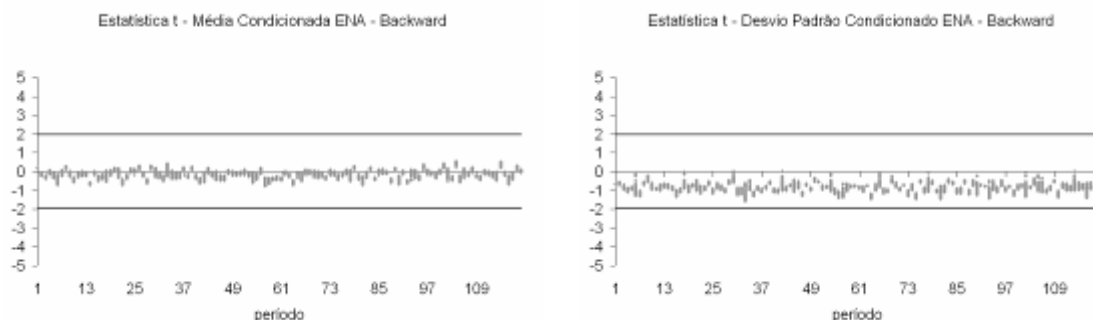
Conforme já mencionado anteriormente, é observado um pequeno viés para baixo na reprodução do desvio-padrão nas opções 1 e 4.



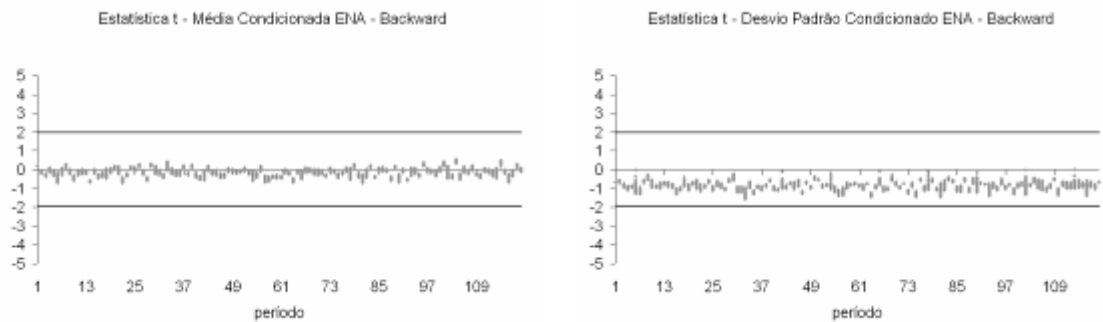
**(a) Opção Atual**



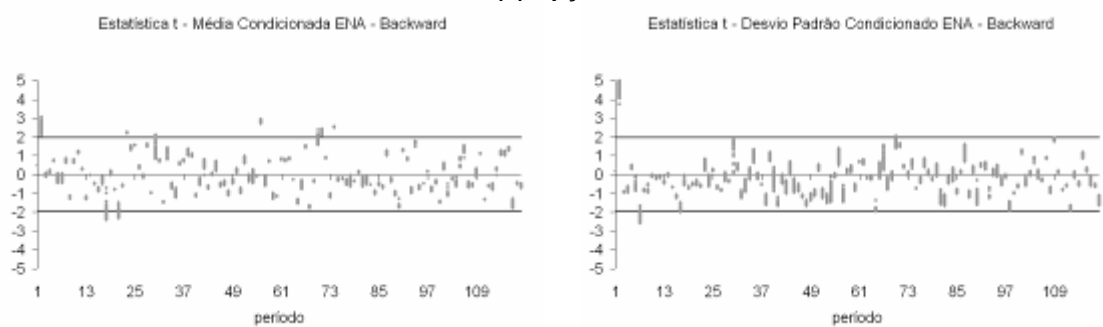
**(b) Opção 0**



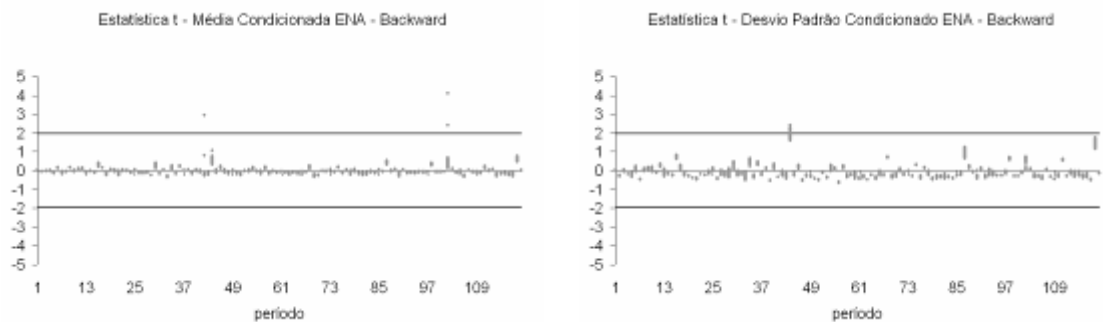
**(c) Opção 1**



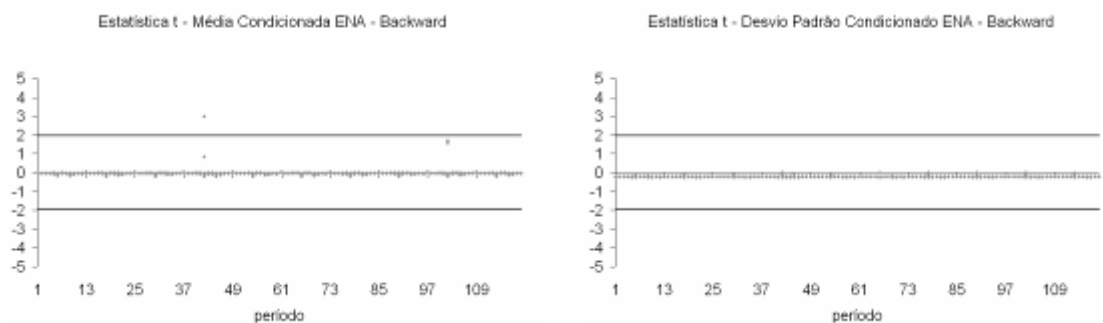
**(d) Opção 4**



**(e) Opção 1AAS**



**(f) Opção LHC**



**(g) Opção AD**

Figura 116: Estatística t Média e Desvio-Padrão – Cenários Backward - Sudeste

A análise e as conclusões obtidas para os demais subsistemas são análogas àquelas realizadas para o subsistema Sudeste. Os resultados podem ser verificados no Anexo B.

### 7.4.5. Correlação Cruzada dos Cenários

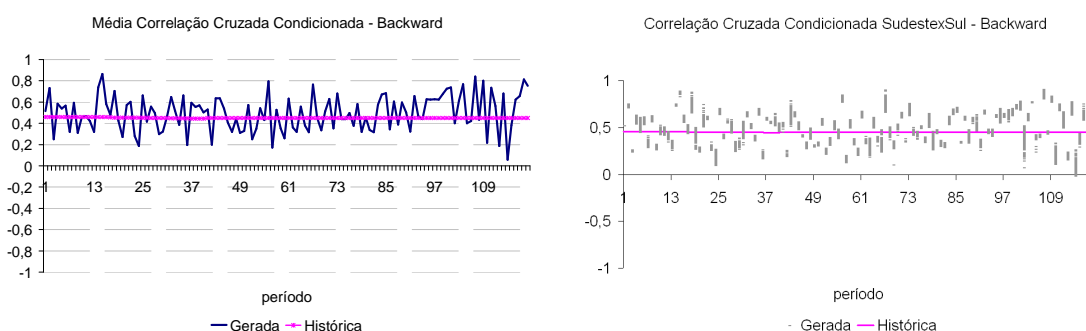
Neste item é avaliada a preservação da correlação cruzada para cada conjunto de cenários backward. A análise desta grandeza é realizada através de uma comparação dos resultados obtidos e os valores históricos.

Na primeira coluna da Figura 117a até a Figura 117g é apresentada a evolução temporal da média da correlação cruzada entre os subsistemas Sudeste e Sul. Esta média é calculada a partir dos 200 valores obtidos de correlação cruzada para cada período, que são ilustrados na segunda coluna desta figura. Em ambos gráficos, a linha mais clara indica a correlação cruzada histórica.

Observe que há uma grande variação em torno do valor histórico para as opções Atual e 1AAS. A variabilidade bastante acentuada verificada para estas opções está relacionada ao tamanho da amostra utilizada.

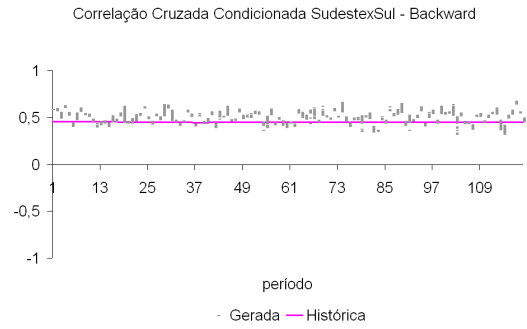
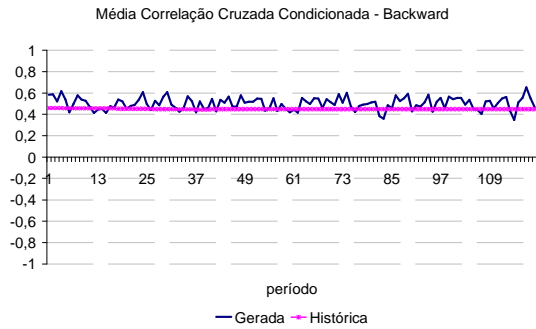
Como pode ser observado na Figura 117, letras (f) e (g), as opções LHC e AD apresentam uma variabilidade com relação ao valor histórico. Além disso, também apresentam uma tendência de redução no valor da correlação cruzada estimada. Isto ocorre por causa do tamanho limitado da amostra.

Mesmo considerando uma amostra de tamanho limitado, as opções que utilizam o processo de agregação conseguem preservar de maneira satisfatória a correlação cruzada histórica, conforme pode ser verificado nos resultados das opções 0, 1 e 4. Note também que ocorre uma menor variabilidade nos resultados.

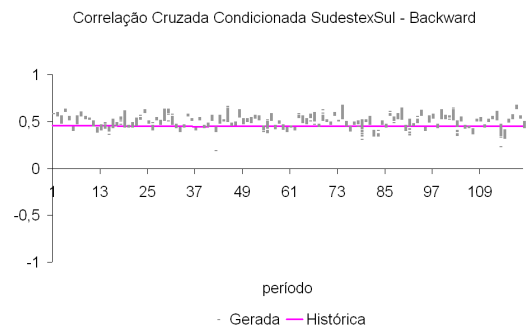
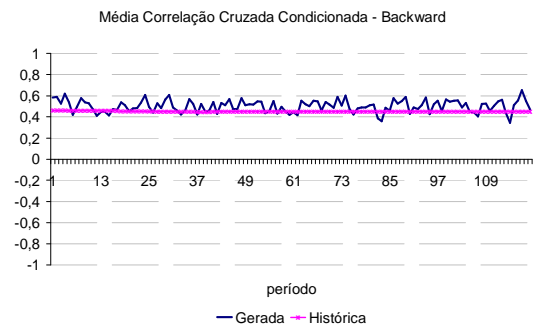


(a) Opção Atual

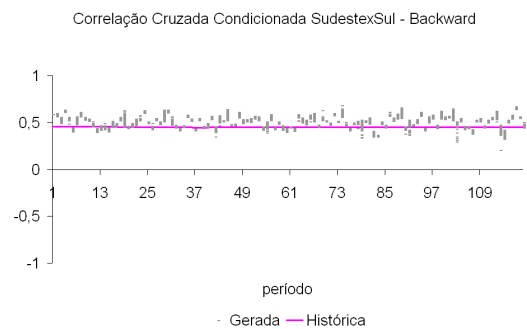
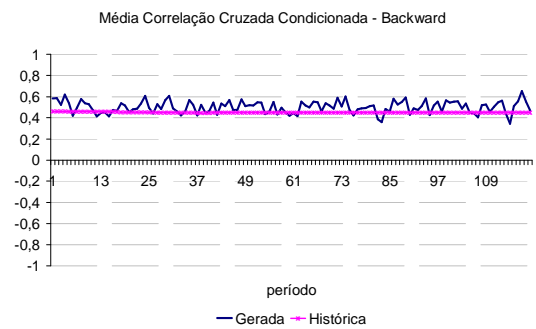




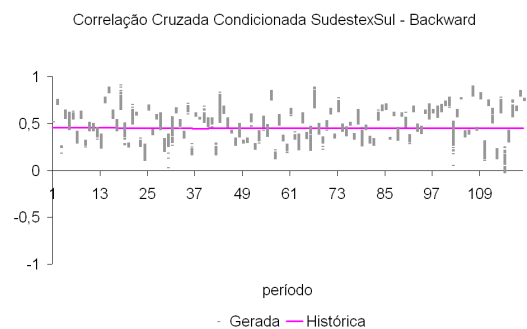
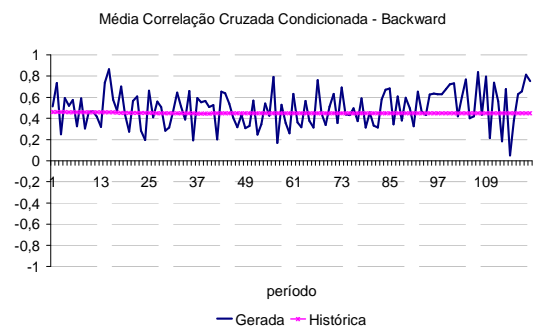
**(b) Opção 0**



**(c) Opção 1**



**(d) Opção 4**



**(e) Opção 1AAS**

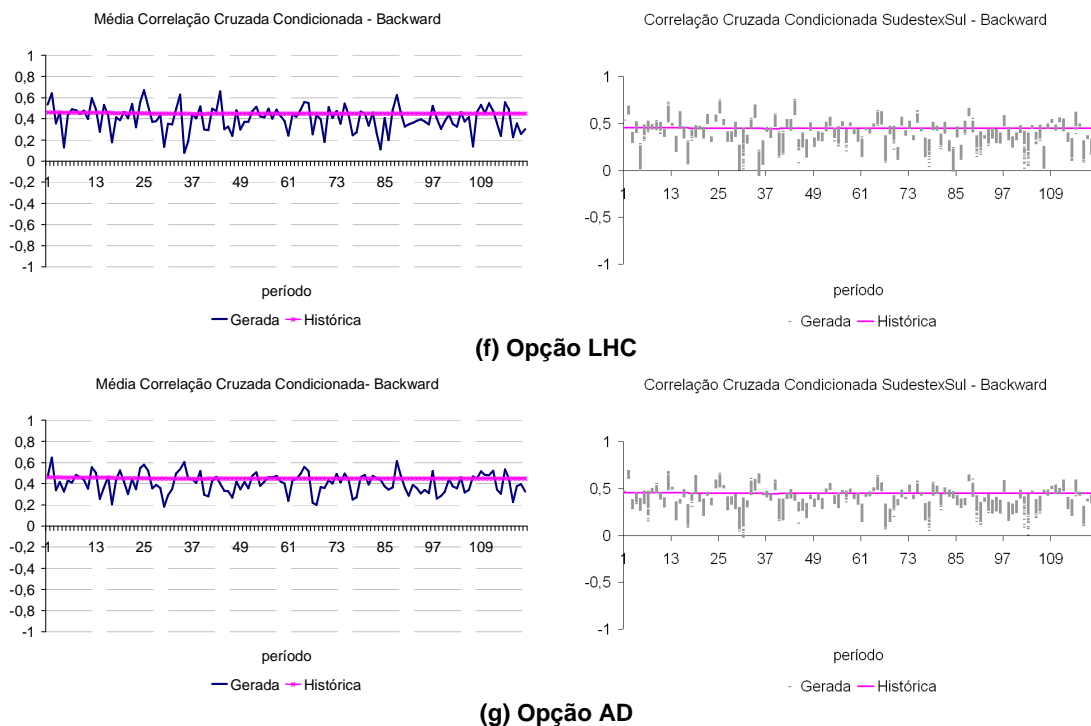


Figura 117: Correlação Cruzada Cenários Backward – Sudeste x Sul

Os resultados obtidos para as correlações entre os demais subsistemas podem ser conferidos no Anexo B. Em geral, as mesmas observações realizadas neste item podem ser feitas às demais correlações.

É desejável que os cenários dos passos forward e backward representem a correlação cruzada de forma compatível e que ambas se aproximem do valor observado no histórico. Nas figuras a seguir são apresentadas as correlações cruzadas de cada conjunto cenário backward e a correlação cruzada dos cenários forward entre os subsistemas Sudeste e Sul. Pode-se verificar que para a opção Atual, Figura 118, apresenta uma grande variabilidade da correlação cruzada tanto para o passo backward quanto para o passo forward. Também pode-se observar que a correlação do passo forward está bem próxima das correlações do passo backward. Nas opções que utilizam o processo de agregação, Figura 119 a Figura 121, a variabilidade das correlações é bem reduzida e a correlação do passo forward continua próxima das correlações do passo backward. Para as opções 1AAS, LHC e AD, observa-se uma grande variabilidade na correlação da backward, porém uma menor variação na correlação do passo forward, Figura 122 a Figura 124.

Correlação Cruzada Condicionada SudestexSul - Backward

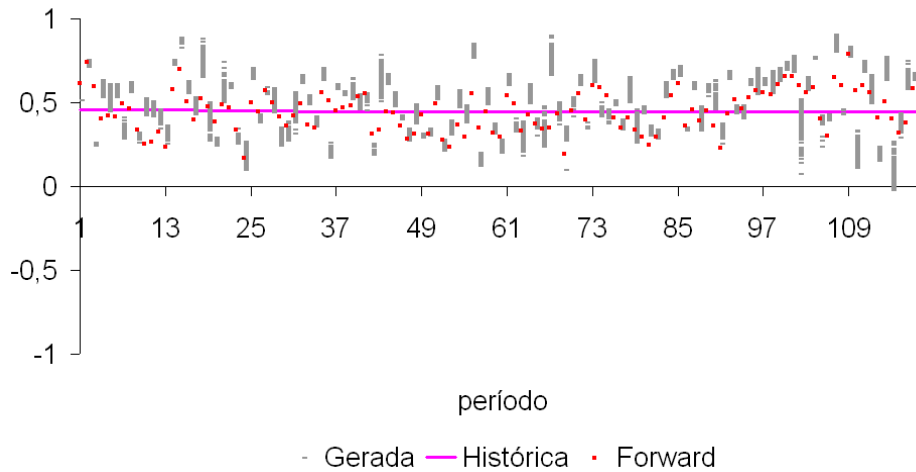


Figura 118: CC Backward x Forward – Sudeste x Sul – Opção Atual

Correlação Cruzada Condicionada SudestexSul - Backward

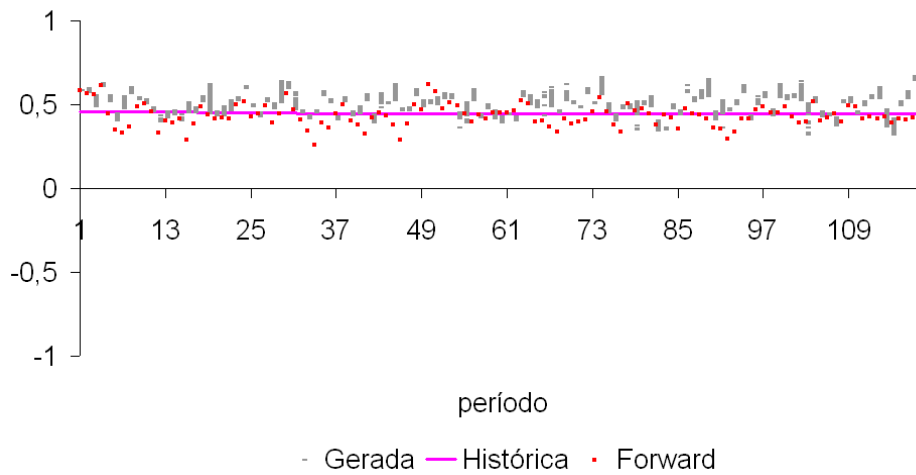


Figura 119: CC Backward x Forward – Sudeste x Sul – Opção 0

Correlação Cruzada Condicionada SudestexSul - Backward

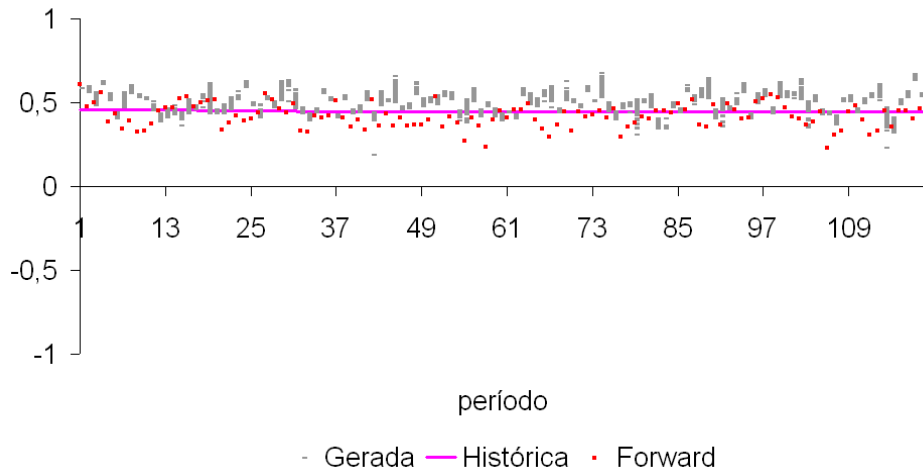


Figura 120: CC Backward x Forward – Sudeste x Sul – Opção 1

Correlação Cruzada Condicionada SudestexSul - Backward

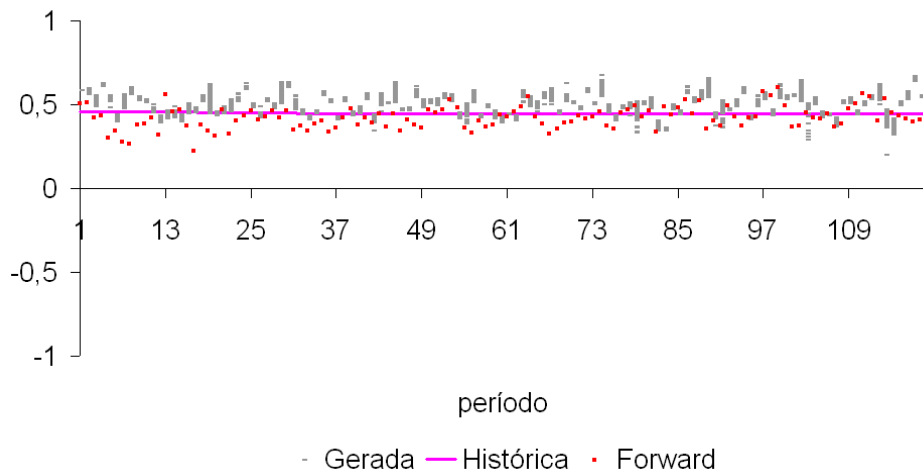


Figura 121: CC Backward x Forward – Sudeste x Sul – Opção 4

Correlação Cruzada Condicionada SudestexSul - Backward

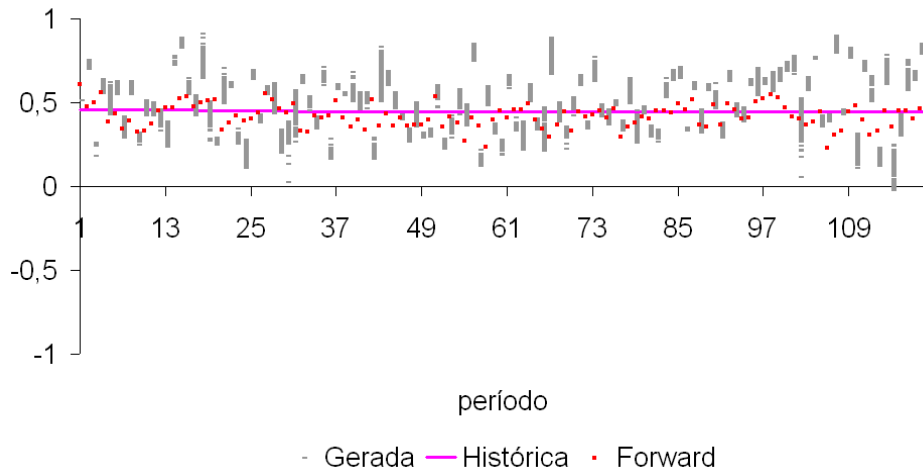


Figura 122: CC Backward x Forward – Sudeste x Sul – Opção 1AAS

Correlação Cruzada Condicionada SudestexSul - Backward

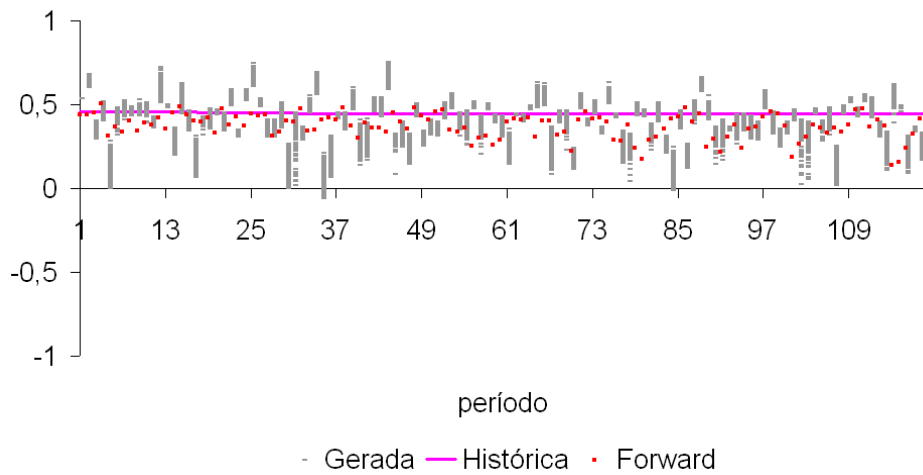


Figura 123: CC Backward x Forward – Sudeste x Sul – Opção LHC

Correlação Cruzada Condicionada SudestexSul - Backward

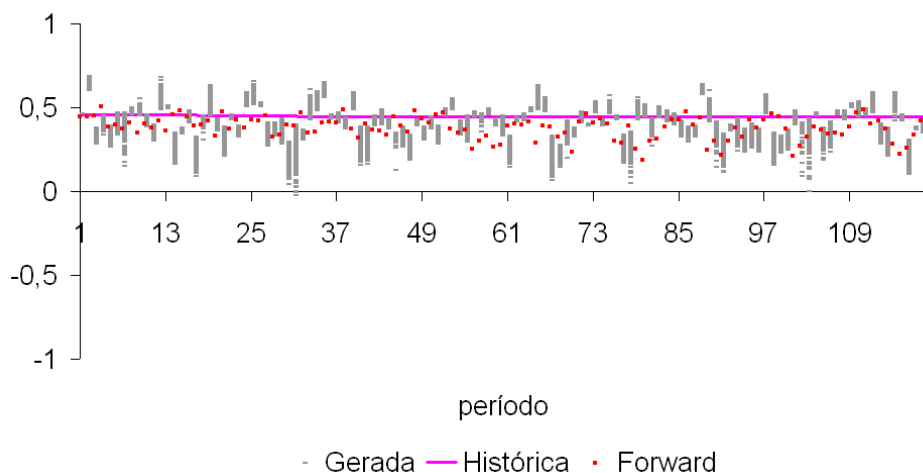


Figura 124: CC Backward x Forward – Sudeste x Sul – Opção AD

Na Tabela 11 é mostrada a correlação cruzada média não condicionada, ao longo de todo horizonte de estudo para opções. Da mesma forma como observado para a correlação cruzada condicionada, há dificuldade em preservar o valor da correlação cruzada histórica nas opções LHC e AD.

Tabela 11: Correlação Cruzada Média Cenários Backward – Não Condicionado

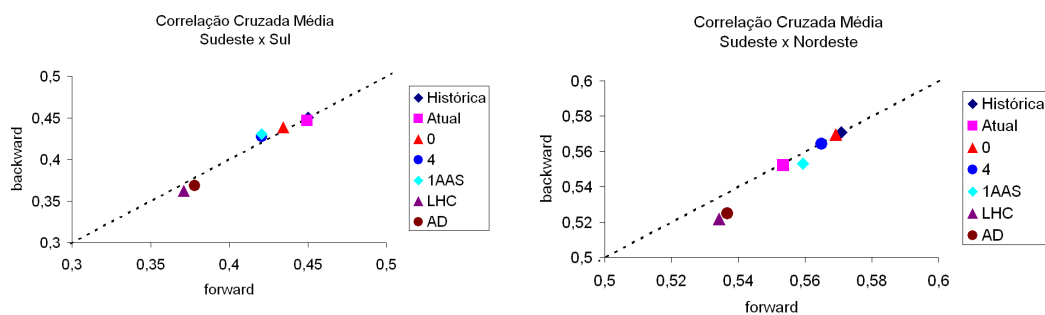
OPÇÃO	Sudeste-Sul	Sudeste-Nordeste	Sudeste-Norte	Sul-Nordeste	Sul-Norte	Nordeste-Norte
<b>Histórico</b>	0,4502	0,5709	0,4265	-0,0792	-0,1313	0,6372
<b>Atual</b>	0,4467	0,5521	0,3936	-0,0453	-0,1049	0,6253
<b>0</b>	0,4386	0,5696	0,4329	-0,0702	-0,1078	0,6528
<b>1</b>	0,4215	0,5616	0,4264	-0,0609	-0,1011	0,6453
<b>4</b>	0,4271	0,5644	0,4145	-0,061	-0,1036	0,6441
<b>1AAS</b>	0,4306	0,553	0,4151	-0,0586	-0,107	0,6416
<b>LHC</b>	0,362	0,5216	0,4009	-0,0525	-0,0753	0,5794
<b>AD</b>	0,3684	0,5249	0,4074	-0,0534	-0,0765	0,592

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

extremo onde a correlação cruzada representada nos cenários backward é próxima de um, isto é todos os subsistemas são quase perfeitamente correlacionados, e a correlação cruzada dos cenários forward é igual ao valor observado no histórico (diferente de 1). Neste caso, durante a simulação backward há uma grande probabilidade de visitar cenários muito críticos (todos os subsistemas secos) e cenários muito folgados (todos os subsistemas com bastante água). Durante a simulação forward a probabilidade de ocorrência destes cenários extremos é bem menor. Os cenários folgados não têm muito impacto na construção da FCF, pois a operação destes cenários tem custo praticamente igual a zero. Porém, cenários muito críticos podem elevar de sobremaneira o valor da FCF. Desta forma, o valor de ZINF, calculado considerando a função de custo futuro, pode assumir um valor bem alto, podendo inclusive ultrapassar o valor de ZSUP.

Caso ocorra o contrário, isto é, correlação cruzada dos cenários forward for maior do que a correlação dos cenários backward, o valor de ZSUP pode ser bem maior que o valor de ZINF, de forma que o processo não consiga convergir.

Na Figura 125 é ilustrada uma comparação entre as correlações cruzada dos cenários forward e a média das correlações cruzadas condicionadas dos cenários backward. A linha pontilhada indica uma reta com ângulo igual 45° e passando pela origem, isto é, quanto mais próximos os pontos estiverem desta linha maior a proximidade da representação da correlação cruzada nos cenários backward e forward. Note que em geral todas as opções, exceto as opções LHC e AD, representam da mesma forma a correlação cruzada nos cenários da backward e forward. As conseqüências desta representação para as opções LHC e AD serão discutidas no capítulo seguinte.



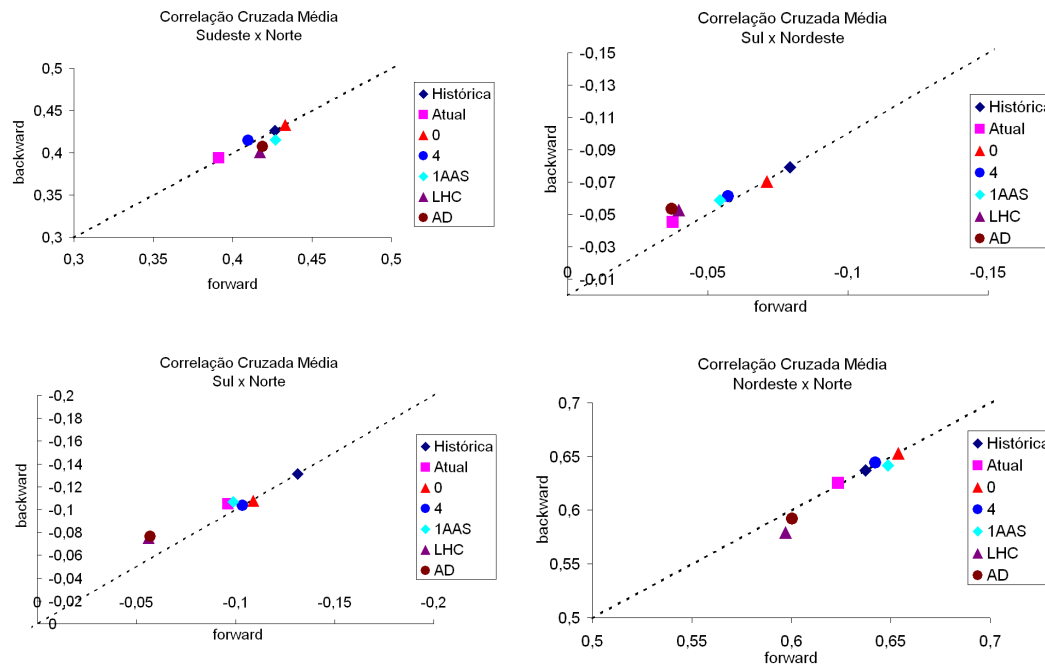


Figura 125: Comparação entre CC Forward e Backward

## 7.5. Resumo

Neste capítulo foi apresentada uma extensa avaliação dos cenários forward e backward gerados segundo as opções propostas neste trabalho. Além disso, foram testadas as amostras de ruídos que são usadas na geração dos cenários.

Durante as análises foi verificado que uma amostra de ruídos com tamanho igual a dois mil é suficiente para representar bem uma distribuição multivariada para quatro subsistemas. Para a geração da amostra de ruídos do passo forward, foi verificado uma dificuldade na preservação tanto da média quanto do desvio-padrão na opção Atual. O mesmo foi observado na opção 0 também. Para as demais opções, com destaque para as opções LHC e AD, foi observado que em geral as estatísticas analisadas são preservadas. Normalmente, as opções onde são utilizadas técnicas de agregação foi observada uma pequena degradação com relação ao desvio-padrão, principalmente com amostras de tamanho reduzido. Com relação à amostra utilizada no passo backward, pode-se verificar que os resultados das opções Atual e 1AAS apresentaram um pior desempenho se comparados com os resultados das demais opções.



Na opção Atual verificou-se um desempenho ruim para as estatísticas analisadas dos cenários de ENA para os passos forward e backward. Nas opções onde são utilizadas técnicas de agregação de cenários, foi observado um comportamento bem satisfatório para média, desvio-padrão, teste de aderência e correlação cruzada. Foi verificada novamente uma pequena degradação no desvio-padrão. Com relação à correlação cruzada foi observada uma grande variação para as opções 1AAS, LHC e AD, sendo que nas duas últimas há uma tendência a gerar valores de correlação menores do que o observado no histórico. Nos testes de seqüência negativa todas as opções foram aprovadas.

No próximo capítulo será apresentada uma avaliação dos resultados obtidos para o planejamento da operação energética de médio prazo utilizando os cenários gerados com as todas as opções propostas.

## 8 Avaliação dos Resultados do Problema de Planejamento da Operação

### 8.1. Considerações Iniciais

Este capítulo é destinado a aplicação das opções para construção da árvore de cenários de afluências propostas neste trabalho em casos reais do programa mensal de operação. É analisado o impacto da adoção dos métodos propostos em relação a representação atual.

Os estudos de caso têm um horizonte de planejamento de cinco anos com um período estático de mais cinco anos. A topologia dos casos estudo varia, porém um esquema com a topologia das usinas hidroelétricas do parque gerador brasileiro para o horizonte de 2008 a 2012 é apresentado na Figura 126.

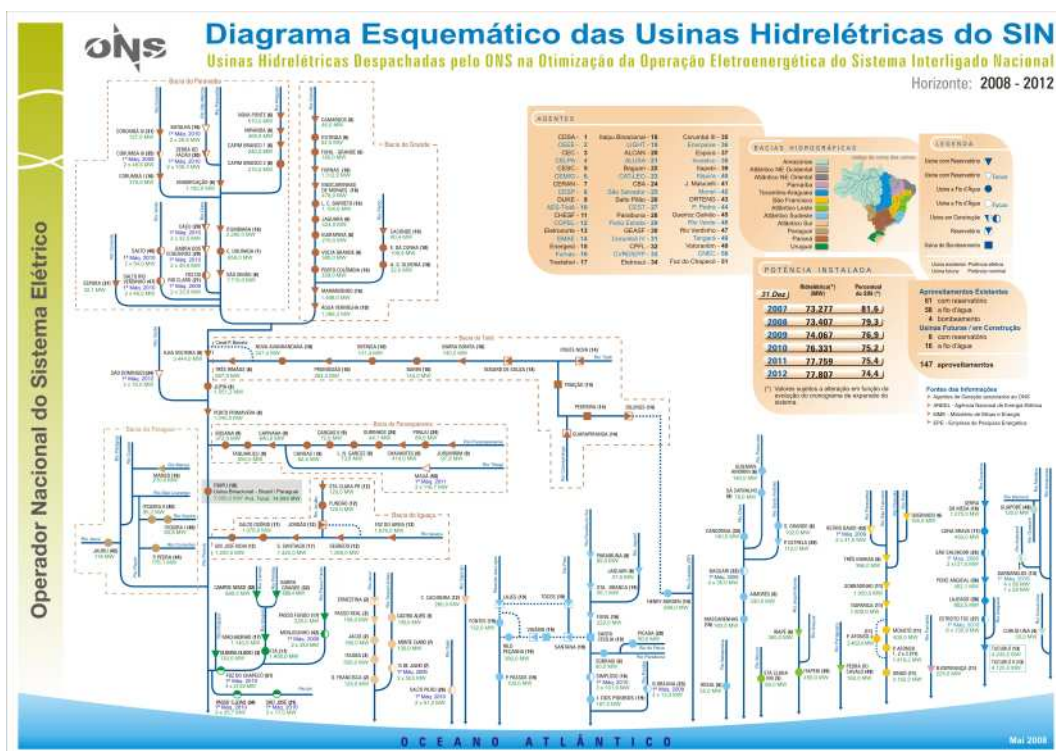


Figura 126: Diagrama esquemático das usinas hidroelétricas do SIN 2008-2012

O parque gerador hidroelétrico é agregado em quatro subsistemas equivalentes de energia, a saber, Sudeste/Centro-Oeste, Sul, Nordeste e Norte, conforme mostrado na Figura 127. Nesta figura também é possível visualizar os troncos de transmissão que interligam os subsistemas.

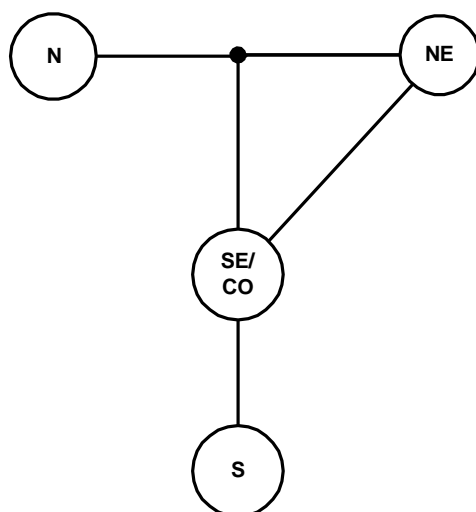


Figura 127: Representação dos subsistemas equivalentes

Para o cálculo da política ótima de operação são utilizados 200 cenários para o passo forward e para cada período são consideradas 20 aberturas (cenários backward). Após o cálculo da política mensal, é realizada uma simulação final utilizando-se 2000 cenários sintéticos de energia natural afluentes para obtenção dos índices de desempenho do sistema como os riscos de não atendimento à carga própria, valor esperado do custo de operação, gerações hidráulicas e térmicas de cada subsistema, armazenamentos, intercâmbios, vertimentos e etc. O conjunto de cenários utilizado na simulação final é gerado com amostragem aleatória simples e tem um horizonte de cinco anos. Este conjunto é utilizado na simulação final de todas as opções analisadas.

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

Ao todo foram estudados e analisados 11 diferentes casos de PMO. Esses casos foram escolhidos dentre um conjunto de 49 PMOs (PMO JAN/04 a PMO

JAN/08) de forma a representar a diversidade de resultados observados no conjunto completo de casos. Dentre os casos analisados existem, por exemplo, casos iniciando em período úmido e período seco, casos com tendência hidrológica (passado recente) favoráveis e desfavoráveis.

Neste capítulo não são analisados resultados de risco e custo marginal de operação para o subsistema Norte, pois devido às características hidrológicas deste subsistema e dado o seu pequeno porte frente ao SIN, a FCF construída pelo modelo de médio prazo pode não ficar bem detalhada para esse subsistema e, como consequência, ao simular a operação ótima do SIN podem ser encontrados altos índices de risco de déficit para o subsistema Norte (Penna et al, 2006). É importante ressaltar que o objetivo básico do planejamento da operação é obter, para cada etapa, as metas de geração de cada usina hidrelétrica e termelétrica do SIN de forma a atender a demanda e minimizar o valor esperado do custo de operação ao longo do período de planejamento. O pouco detalhamento da FCF para o subsistema Norte não implica que o processo de construção da função de custo futuro esteja ruim, uma vez que o objetivo do modelo, como já foi dito, é encontrar a operação ótima para o sistema interligado como um todo.

Neste capítulo são apresentados basicamente os resultados para o PMO FEV/07, os resultados obtidos para os demais PMO estudados podem ser consultados no Anexo C.

## **8.2. Avaliação da convergência**

A cada iteração do algoritmo de solução da PDDE são calculados os limites superior (ZSUP) e inferior (ZINF) da solução ótima. O limite inferior é obtido na solução do problema de primeiro período, e o limite superior é estimado a partir dos resultados de todos os períodos e cenários visitados durante a simulação forward. A incerteza em torno da estimativa de ZUP é utilizada como critério de convergência estatístico para o algoritmo. No modelo NEWAVE há um limite para o número de iterações, atualmente este limite é igual a 45.

O número de iterações necessárias para que o critério de convergência fosse alcançado é apresentado na Tabela 12 e na Tabela 13 para todas as opções descritas neste trabalho. Na última coluna da Tabela 13 é apresentado o número de casos que atingiram o número máximo de iterações considerado.

Tabela 12: Número de iterações para convergência

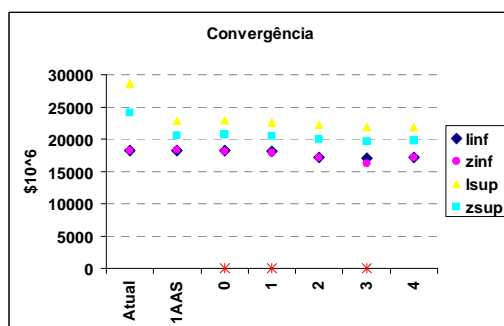
	2004-05	2004-09	2005-09	2006-06	2006-07	2007-01
<b>Atual</b>	24	45	14	17	28	16
<b>1AAS</b>	43	24	17	18	39	45
<b>0</b>	45	45	45	45	45	45
<b>1</b>	45	45	45	21	45	38
<b>2</b>	12	17	10	10	10	7
<b>3</b>	45	45	45	45	45	45
<b>4</b>	14	28	16	12	12	7

Tabela 13: Número de iterações para convergência (cont.)

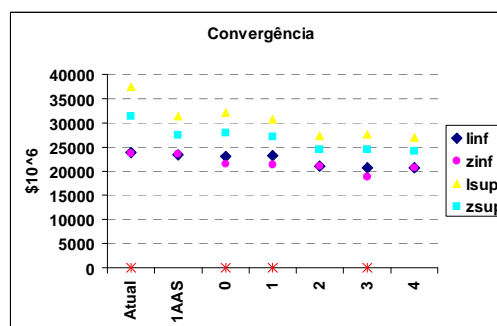
	2007-02	2007-05	2007-07	2007-08	2008-01	#45
<b>Atual</b>	10	16	34	19	10	1
<b>1AAS</b>	37	45	45	18	45	4
<b>0</b>	36	45	45	45	45	10
<b>1</b>	14	12	43	25	17	4
<b>2</b>	7	7	10	11	5	0
<b>3</b>	45	45	45	45	35	10
<b>4</b>	9	9	9	11	6	0

Pode-se observar nas opções Atual, 1, 2 e 4 uma maior facilidade para alcançar o critério de convergência estatístico. Os valores de ZINF e ZSUP estão bem próximos na maioria dos casos que chegam a 45ª iteração, conforme pode ser verificado na Figura 128.

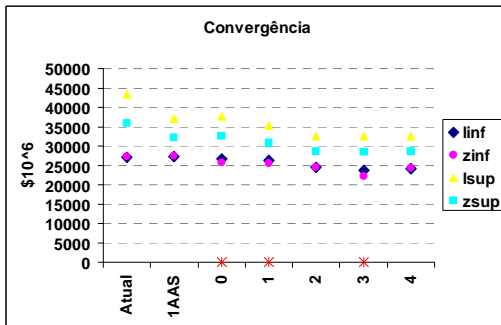
Da Figura 128a até a Figura 128k são apresentados os limites inferiores e superiores para o valor ótimo obtido na última iteração do algoritmo da PDDE.



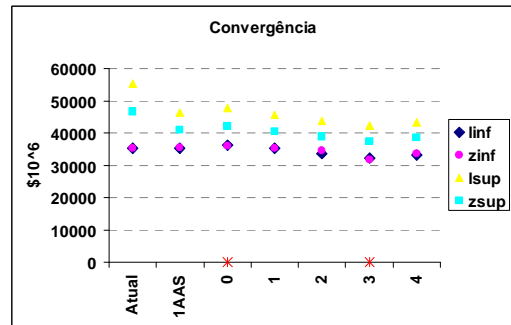
(a) PMO MAI/04



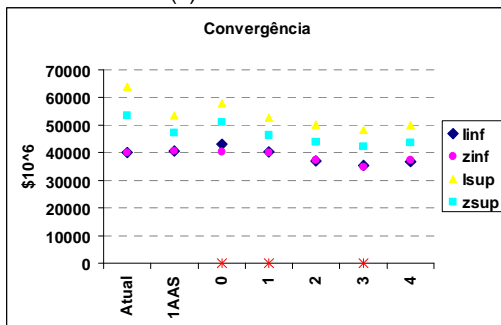
(b) PMO SET/04



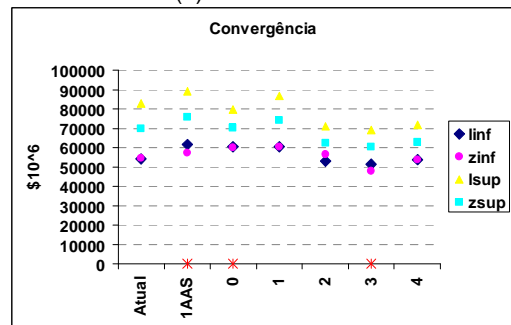
(c) PMO SET/05



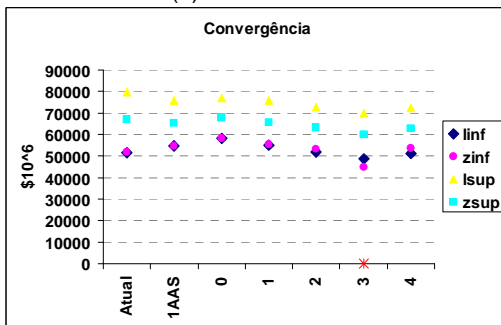
(d) PMO JUN/06



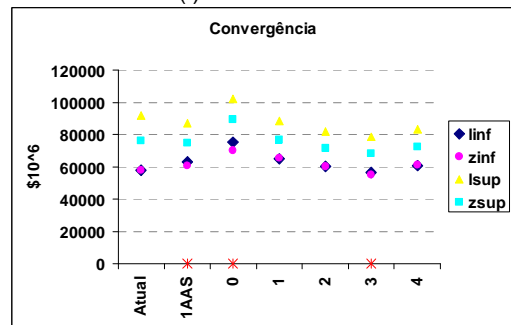
(e) PMO JUL/06



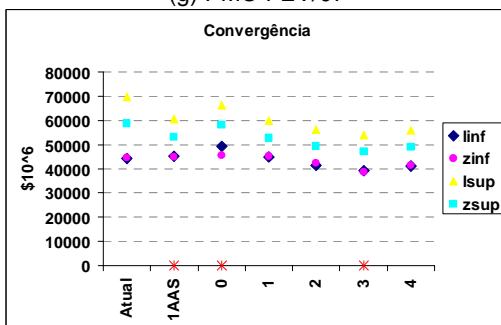
(f) PMO JAN/07



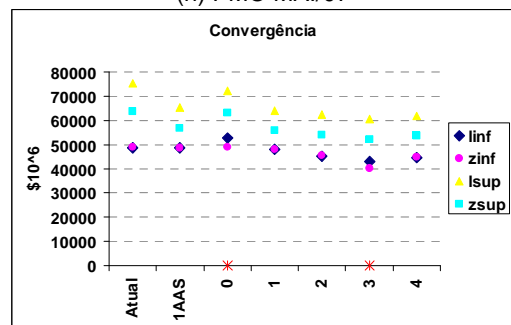
(g) PMO FEV/07



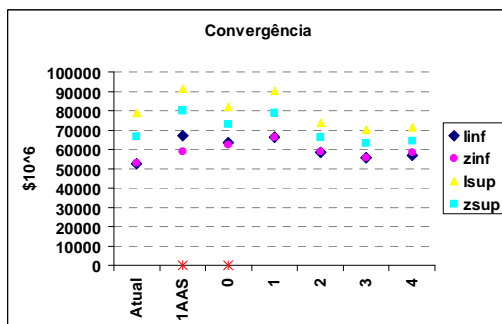
(h) PMO MAI/07



(i) PMO JUL/07



(j) PMO AGO/07



(k) PMO JAN/08

Figura 128: Convergência – Última Iteração

Como pode ser constatado na figura anterior, o intervalo de confiança para a solução ótima para opção Atual é maior do que os intervalos das demais opções e o custo de operação verificado no cálculo da política é, em geral, um pouco mais elevado. Já as opções que empregam técnicas de agregação para construção das amostras forward e backward, apresentam um custo de operação ligeiramente inferior que as demais opções.

Apesar de haver pequenas diferenças entre os custos de operação, o mais importante a ser verificado é a robustez dos resultados face a alterações na árvore de cenários. Desta forma, a seguir é analisada a alteração da árvore de cenários, isto é, são utilizadas diferentes sub-árvores obtidas a partir de diferentes amostras de ruídos.

### 8.3. Alteração da amostra de cenários

Neste item é analisado apenas o PMO FEV/07 considerando 200 cenários forward, 20 aberturas e amostra com tamanho igual a 2000 para as opções que utilizam agregação de cenários. Além da amostra de ruídos utilizada nos casos de PMO oficial, são utilizadas outras 7 amostras distintas. Os resultados da convergência da última iteração são apresentados na Figura 129. Os primeiros valores de ZINF e ZSUP de cada opção referem-se a amostra original de ruídos. Pode-se observar uma maior variabilidade principalmente com relação aos valores de ZINF nas opções que não utilizam processo de agregação. O impacto da alteração da amostra de ruídos tem um reflexo maior na amostra backward, pois seu tamanho é reduzido em comparação com a amostra forward. Cabe ressaltar que valores discrepantes, como aquele observado na opção Atual, têm maior chance de ocorrer nas opções que utilizam amostragem aleatória simples pura.

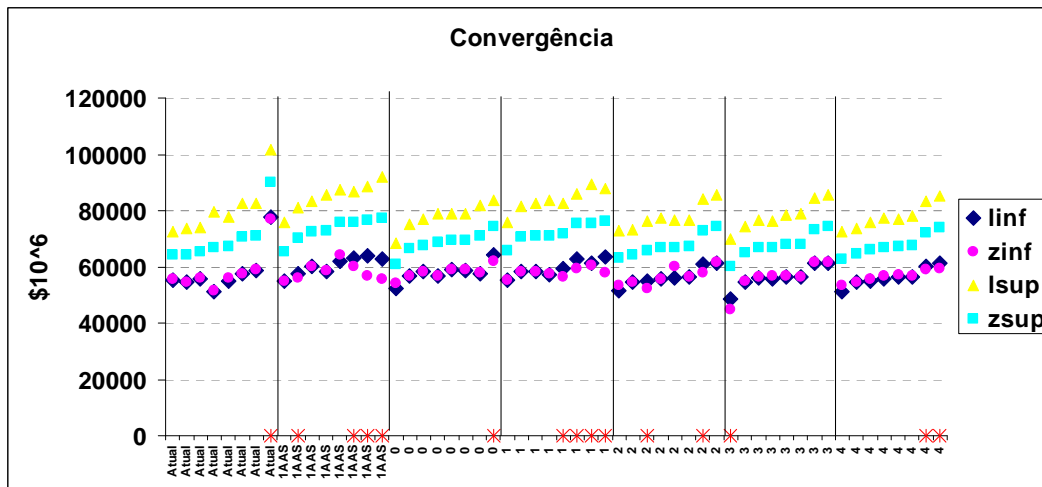


Figura 129: Convergência – Última Iteração – Variação Amostra – FEV/07

Os resultados observados para os PMO SET/04, JUL/06 e JAN/08 corroboram com as observações feita para o PMO FEV/07. Estes PMO apresentam uma variabilidade bem maior nos resultados, quando há alteração na amostra de ruídos. Estes resultados podem ser conferidos no Anexo C.

Na Figura 130 e na Figura 131 são apresentadas trajetórias de convergência utilizando duas amostras distintas de ruídos para a opção Atual e 4. São consideradas a Amostra 0 e aquela que apresentou o maior valor esperado do custo total de operação, Amostra 572 para opção Atual e Amostra 1000 para opção 4.

Para a opção Atual, o critério estatístico de convergência foi satisfeito na décima iteração para o caso Amostra 0, já o caso Amostra 45 atingiu o número máximo de iterações sem que o critério de convergência fosse atingido. Para a comparação da convergência dos dois casos é considerado o mesmo número de iterações. Nota-se que cada caso converge para uma solução distinta, e que as duas soluções são bem diferentes ente si. A diferença entre os valores de ZINF observados na 45ª iteração é de 41% e de ZSUP é de 38%.

Logo, a sub-árvore construída com esta opção não traz robustez ao processo de planejamento da operação energética, com relação a variação da amostra de ruídos tanto para o passo backward quanto para o passo forward.



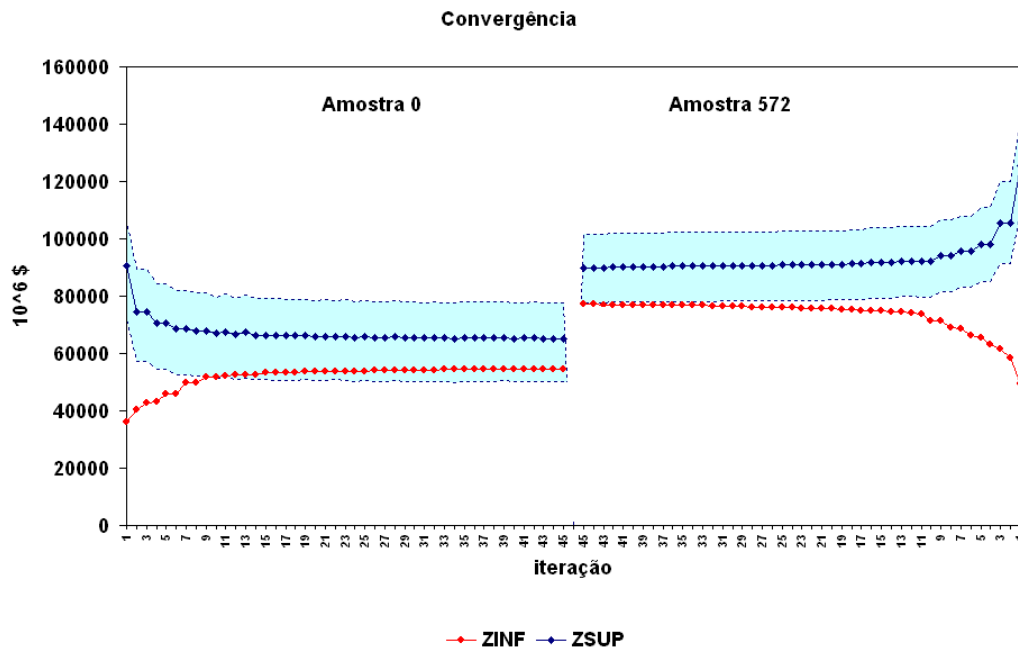


Figura 130: Trajetória Convergência – Variação Amostra – Opção Atual – FEV/07

Para efetuar a comparação das trajetórias para a opção 4, é considerado o mesmo número de iterações para ambas as amostras (no caso 45 iterações). Observe-se que o valor de ZINF pouco se altera com relação à amostra de ruídos utilizada. A diferença entre os valores de ZINF é de 4% e de ZSUP é de aproximadamente 24%. Note que há uma região em comum entre os intervalos de confiança das duas trajetórias, o que não ocorre para opção Atual.

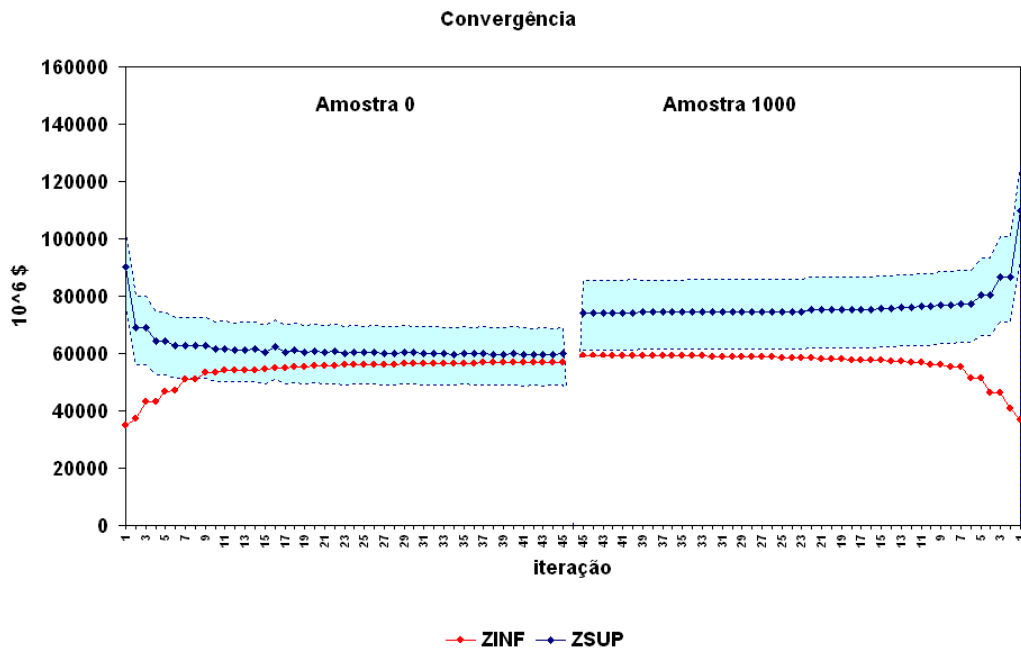


Figura 131: Trajetória Convergência – Variação Amostra – Opção 4 – FEV/07

O procedimento de reamostragem de cenários proposto no capítulo 6 é aplicado para as opções 0 e 4. Os resultados para a última iteração do processo de convergência estão ilustrados na Figura 132. Pode-se observar que a aplicação deste procedimento reduz a variabilidade dos resultados, com relação a variação da amostra de ruídos. Além disso, há uma redução no valor esperado do custo total de operação (ZSUP).

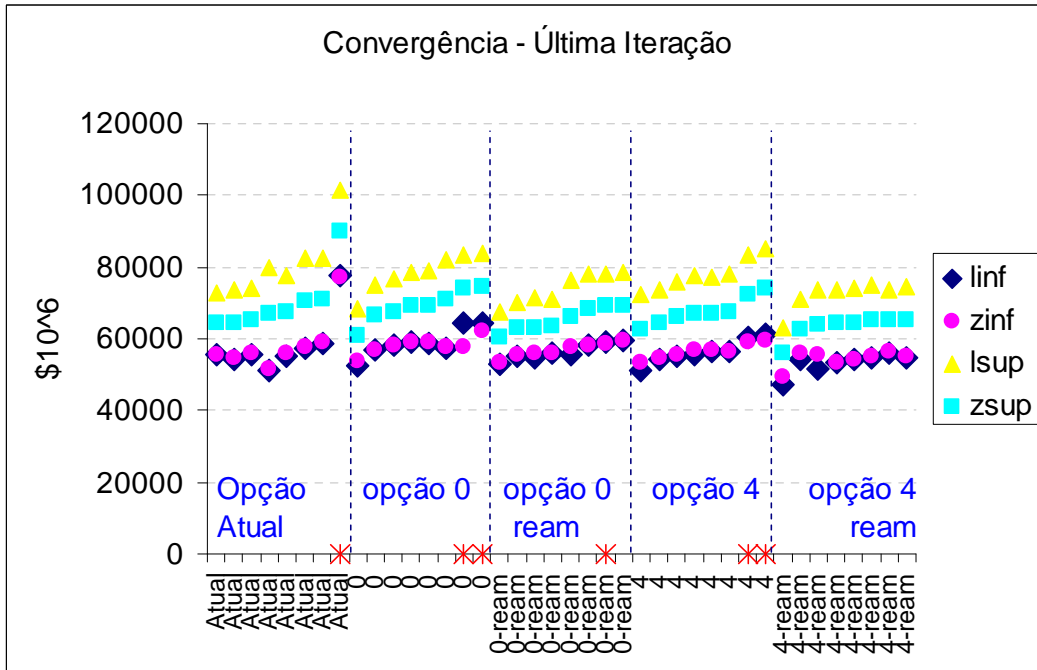


Figura 132: Convergência – Última Iteração – Variação Amostra – Com Reamostragem - FEV/07

Nesta análise, independentemente da opção adotada há um ganho com relação à opção Atual com relação a estabilidade nos resultados de CMO (Figura 133), do valor esperado do custo total de operação da simulação final (Figura 134), no risco anual médio de qualquer déficit (Figura 135) e no valor esperado de energia não suprida (Figura 136).

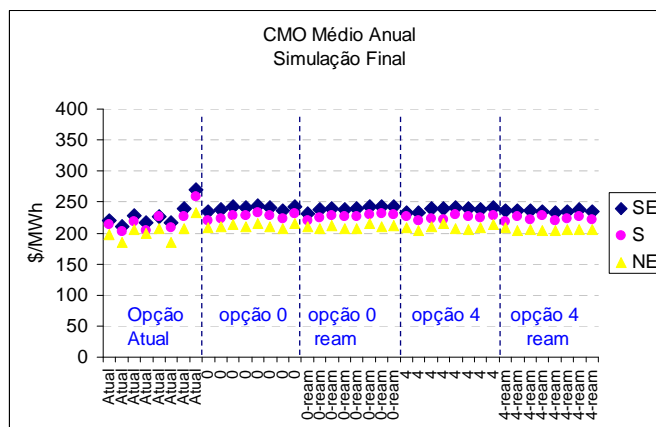


Figura 133: CMO – Variação Amostra – Com Reamostragem - FEV/07

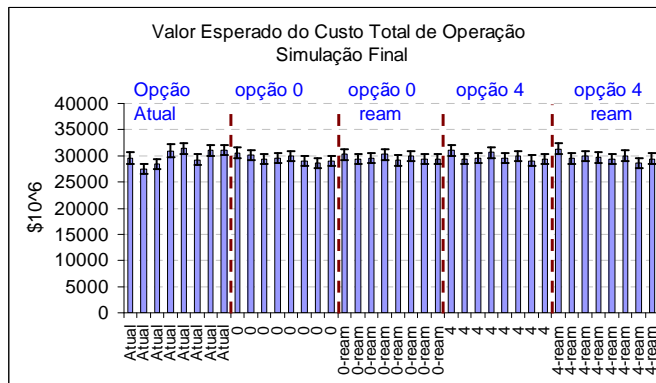


Figura 134: COPER – Variação Amostra – Com Reamostragem - FEV/07

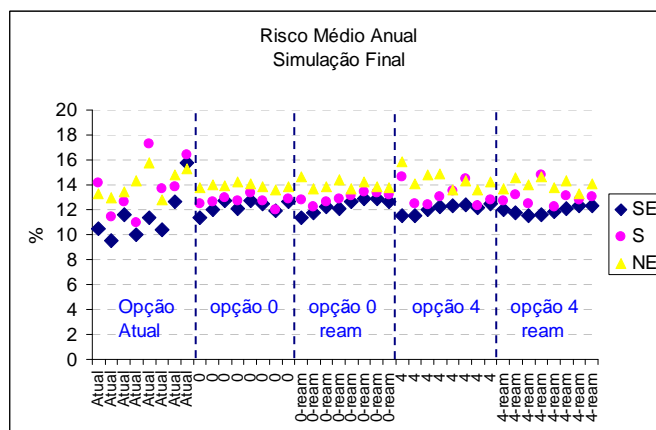


Figura 135: Risco – Variação Amostra – Com Reamostragem - FEV/07

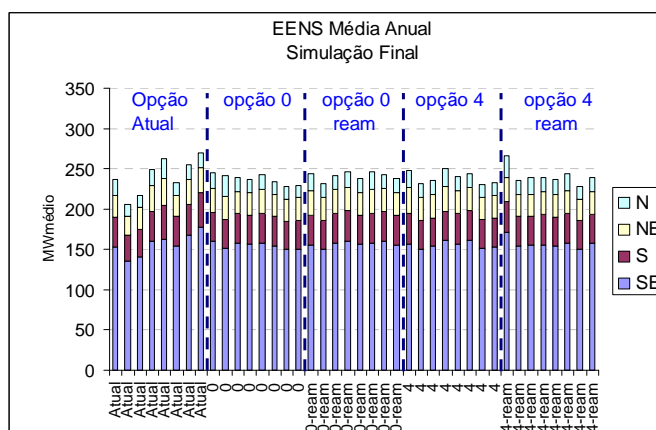


Figura 136: EENS – Variação Amostra – Com Reamostragem - FEV/07

#### 8.4. Alteração do número de cenários backward

Neste item é realizada uma sensibilidade com relação ao número de aberturas do passo backward. Atualmente, nos casos oficiais de PMO, são utilizadas 20

aberturas. Nesta sensibilidade são avaliadas árvores de cenários com 50 e 100 aberturas. Na Figura 137, são apresentados o valor esperado do custo total de operação (ZSUP), o seu valor estimado (ZINF) e o intervalo de confiança de 95% para a última iteração do processo de convergência.

Cabe ressaltar que para as opções Atual e 0, por construção, a alteração na amostra backward gera necessariamente uma modificação na amostra forward.

Observa-se que a variação dos valores de ZINF e ZSUP, para diferentes números de aberturas, é menor nas opções que utilizam o processo de agregação do que a variação dos valores obtidos com a opção Atual. Isto é um indício de que o uso de agregação no passo backward torna o processo mais robusto com relação ao número de aberturas. Ou seja, uma amostra com 20 cenários obtidos com agregação fornece na prática os mesmos resultados que uma amostra com 100 cenários obtidos com agregação, o que não ocorre quando não se utiliza agregação.

Outro ponto importante a ser observado é a redução no valor esperado do custo total de operação para opção 4 quando se adota o processo de reamostragem de cenários.

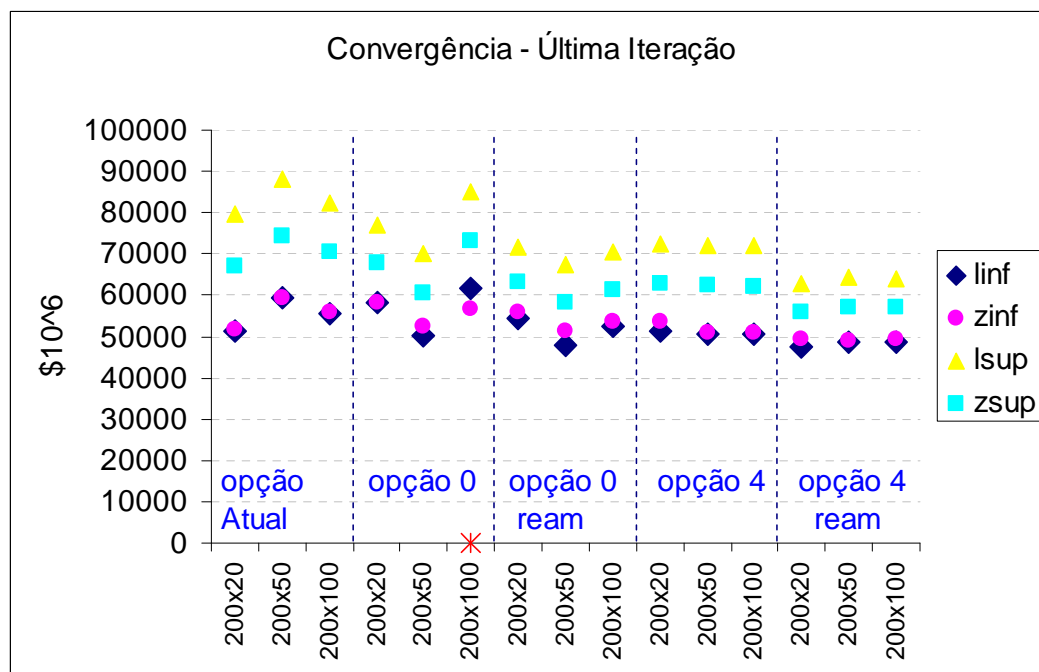


Figura 137: Convergência – Última Iteração – Variação Backward – Com Reamostragem - FEV/07

Apesar de o processo de geração da amostra backward ser o mesmo para as opções 0 e 4, na opção 0 há uma dependência forte entre as amostras backward e

forward. Como já mencionado anteriormente, uma alteração na backward, leva a uma alteração na amostra forward. Logo, a variação dos resultados observada na opção 0 é maior do que na opção 4.

Os impactos observados nos resultados obtidos durante a simulação final são apresentados da Figura 138 até a Figura 141. Note que não há uma variação significativa de resultados entre as versões 0 e 4, e com e sem aplicação do processo de reamostragem.

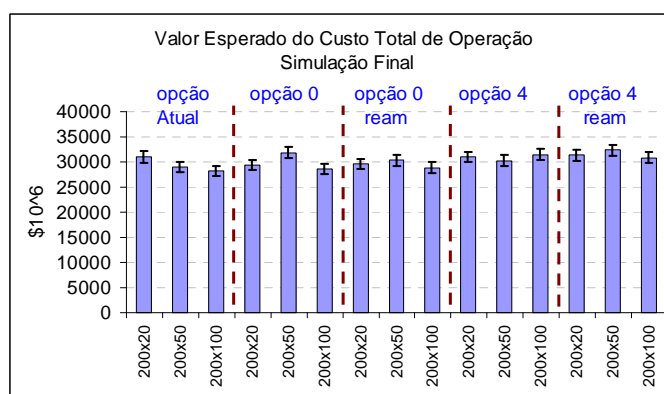


Figura 138: COPER – Variação Backward – Com Reamostragem - FEV/07

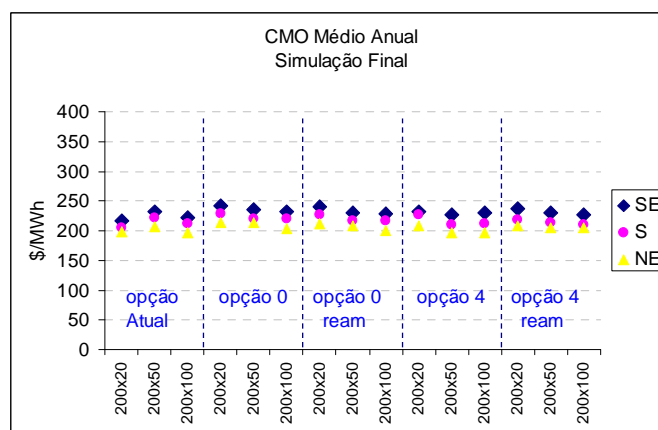


Figura 139: CMO – Variação Backward – Com Reamostragem - FEV/07

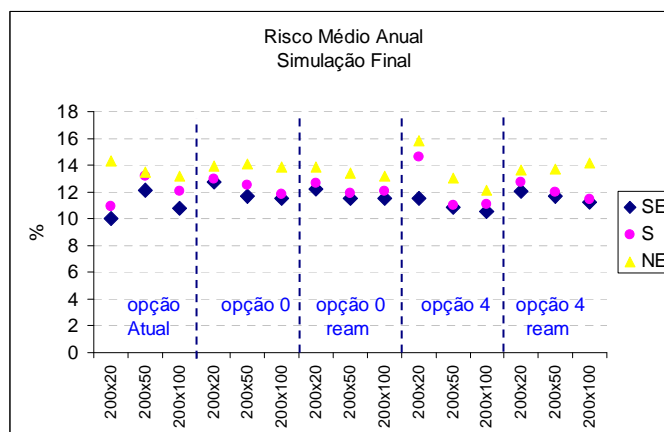


Figura 140: Risco – Variação Backward – Com Reamostragem - FEV/07

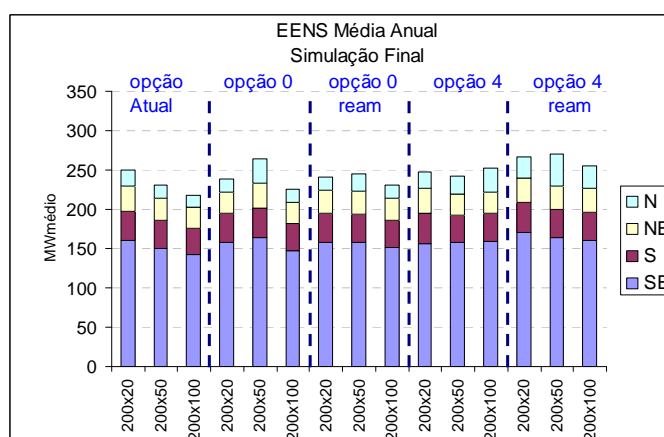


Figura 141: EENS – Variação Backward – Com Reamostragem - FEV/07

## 8.5. Alteração do número de cenários forward

Outra questão importante é verificar se os resultados são robustos com relação a uma sensibilidade no número de cenários utilizados no passo forward. Nos casos oficiais de PMO são utilizados 200 cenários forward. São avaliadas árvores de cenários com 250 e 300 aberturas.

Na Figura 142, são apresentados o valor esperado do custo total de operação (ZSUP), o seu valor estimado (ZINF) e o intervalo de confiança de 95% para a última iteração do processo de convergência. Note que para a opção 4 há uma pequena variação nos resultados à medida que se altera o número de cenários forward.

Conforme observado anteriormente, pode-se verificar na Figura 143, que a utilização do processo de reamostragem em geral reduz o valor de ZSUP.

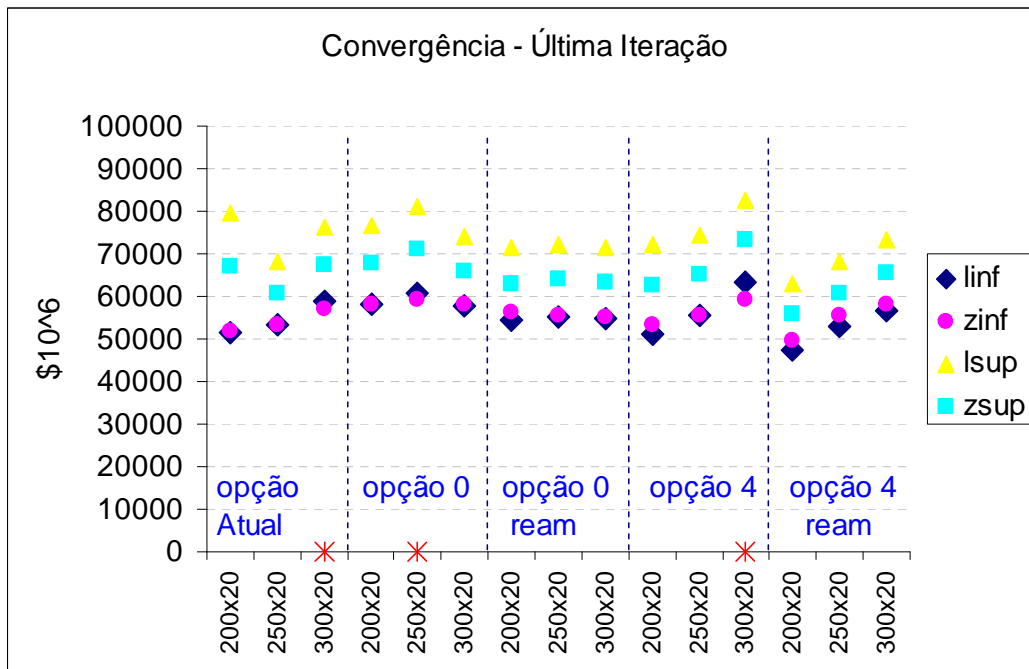


Figura 142: Convergência – Última Iteração – Variação Forward – Com Reamostragem - FEV/07

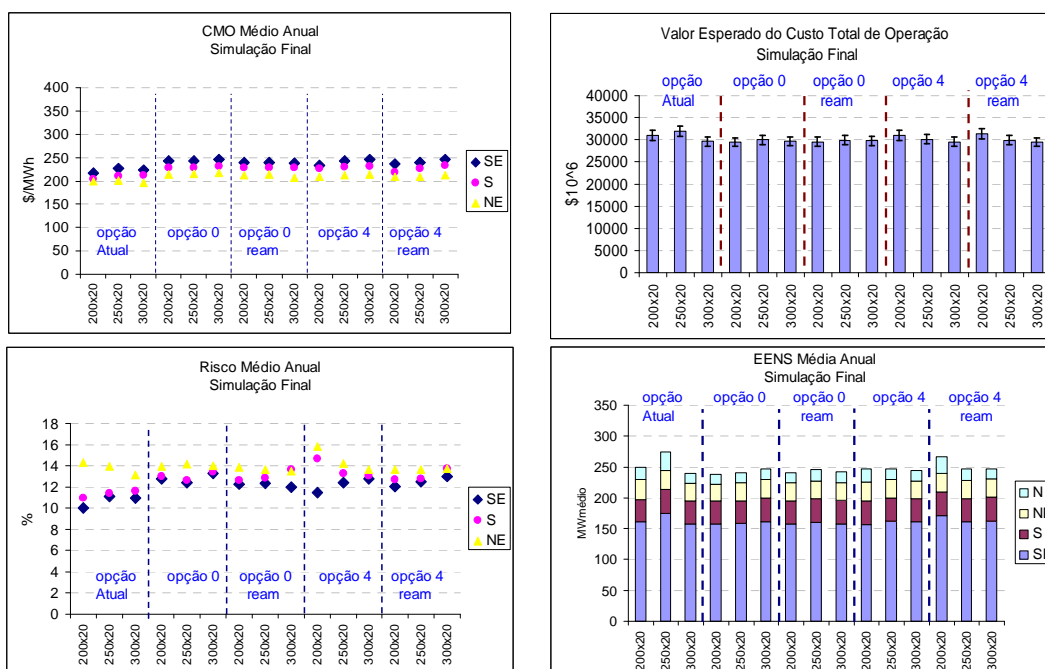


Figura 143: Resultados – Variação Forward – Com Reamostragem - FEV/07

Na análise anterior é considerada apenas a amostra original de ruídos para cada variante do número de cenários forward. Na Figura 144 os resultados são avaliados considerando variações no número de cenários forward e na amostra de ruídos utilizada, para a opção Atual. Pode-se verificar que mesmo considerando uma maior



amostra forward persiste a variação dos resultados quando a amostra de ruídos é alterada. O mesmo pode ser observado para os demais resultados analisados, Figura 145.

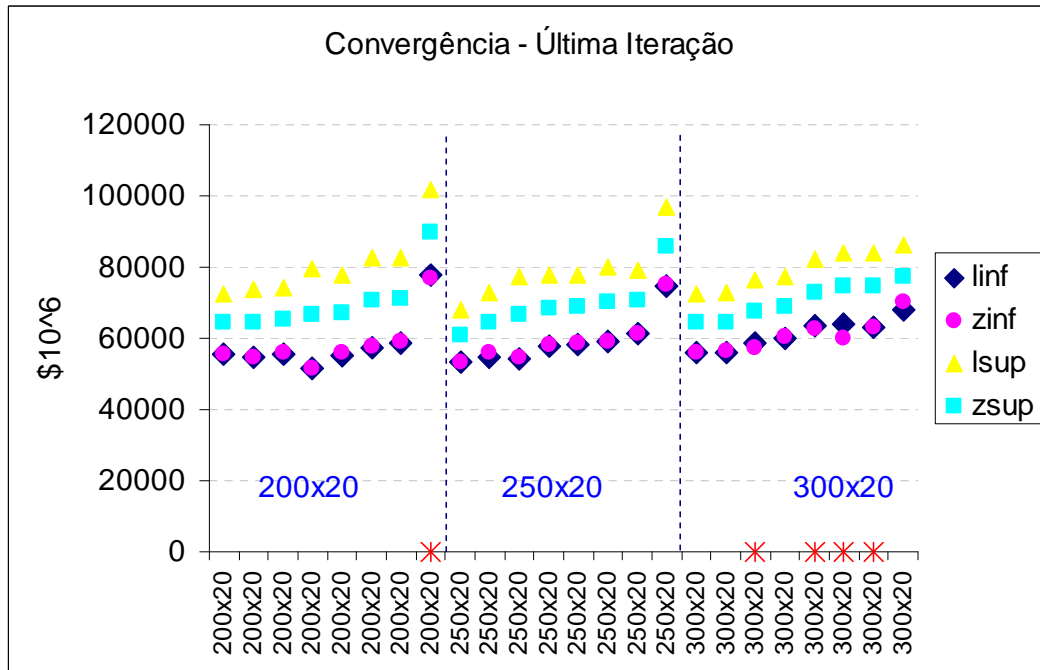


Figura 144: Convergência – Última Iteração – Variação Forward & Amostra – Opção Atual - FEV/07

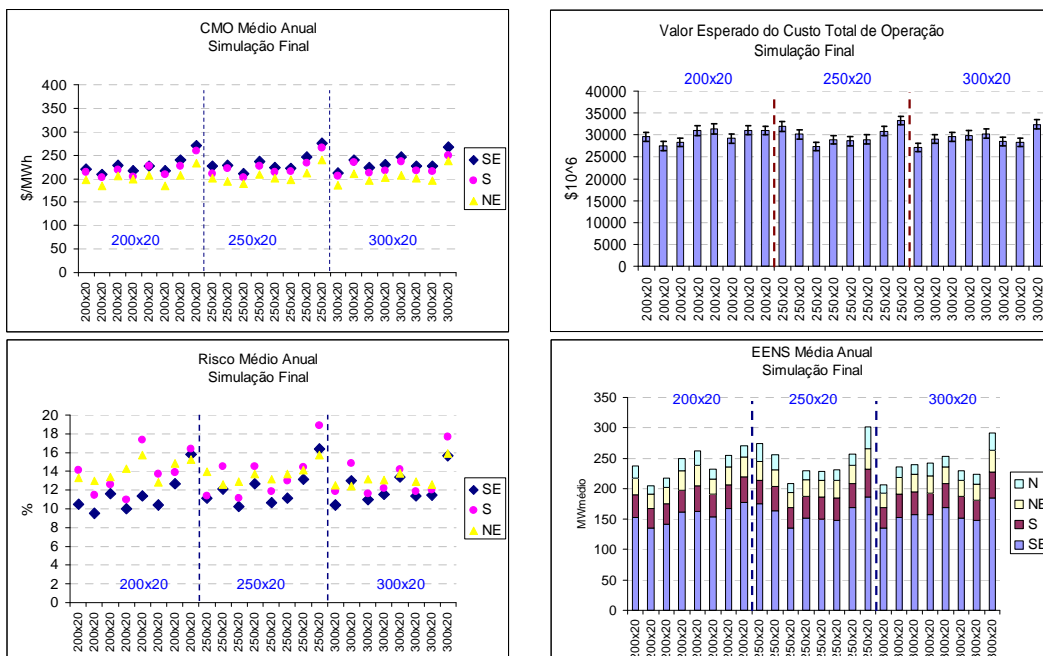


Figura 145: Resultados – Variação Forward & Amostra – Opção Atual - FEV/07

Na Figura 146 são apresentados os resultados da última iteração para a opção 0, variando tamanho da forward e amostra de ruídos. Com relação à opção Atual,

houve uma redução acentuada na variabilidade nos resultados. Isto ocorre, pois a amostra backward que dá origem às amostras forward, fica mais bem representada quando se utiliza métodos de agregação.

Os valores médios anuais do CMO, risco e EENS, além do custo total de operação da simulação final, também estão mais estáveis com relação à alteração da amostra de ruídos (sementes), conforme pode ser verificado na Figura 147.

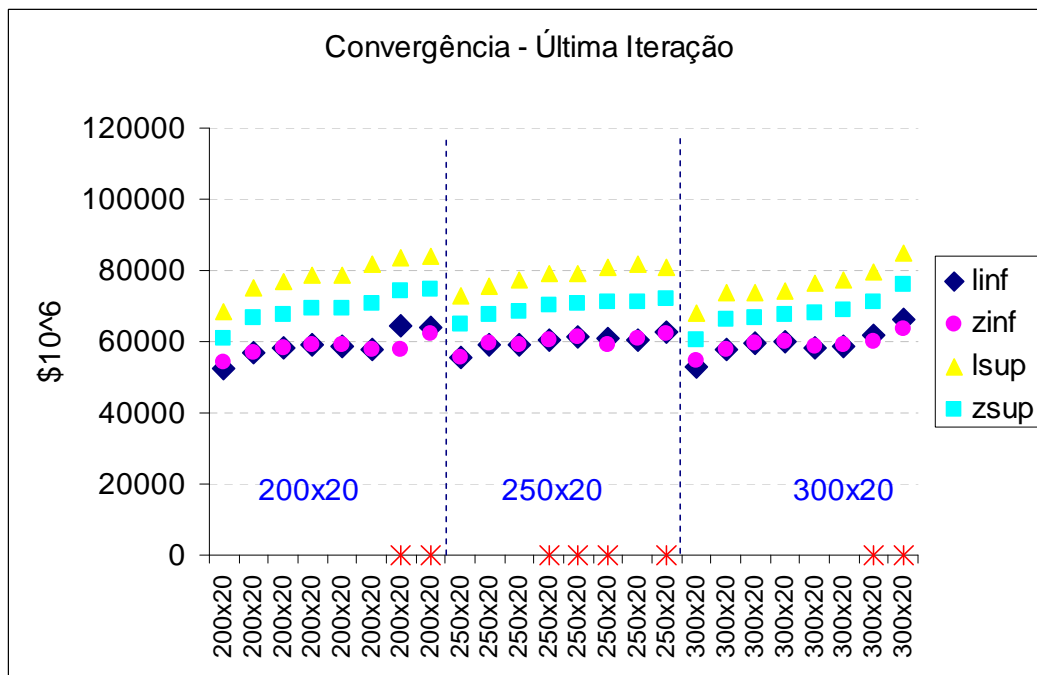
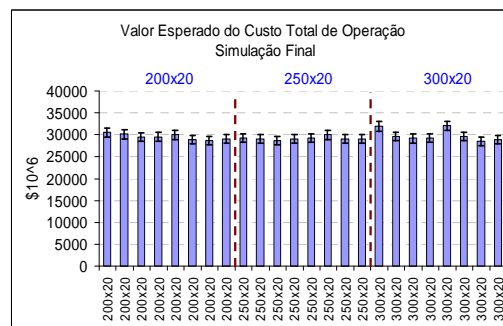
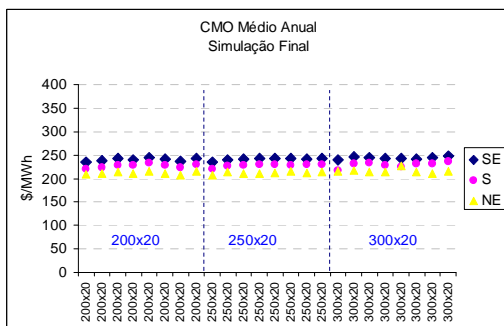


Figura 146: Convergência – Última Iteração – Variação Forward & Amostra – Opção 0 - FEV/07



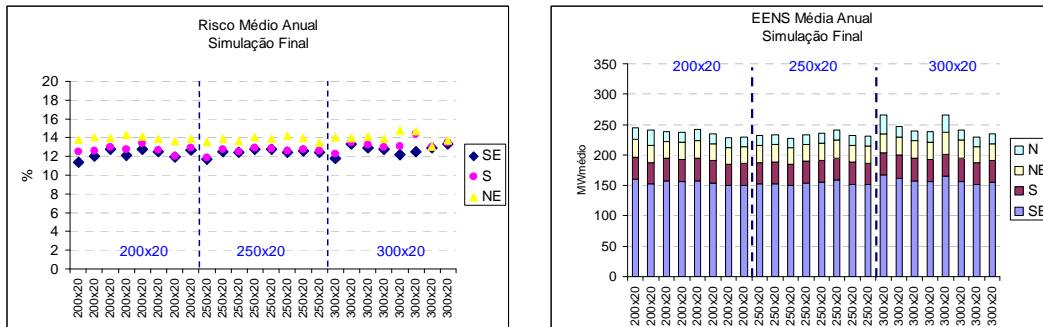


Figura 147: Resultados – Variação Forward & Amostra – Opção 0 - FEV/07

A opção 4 também é analisada com relação às alterações no número de cenários forward e na amostra original de ruídos que é utilizada no processo de agregação. Na Figura 148 são mostrados os valores de ZINF e ZSUP obtidos durante a última iteração do processo de cálculo da política. Há uma pequena flutuação nos valores com relação à variação da amostra de ruídos, mas essa variação é menor do que aquela observada com a opção Atual. O mesmo pode ser observado nos demais resultados analisados, Figura 149.

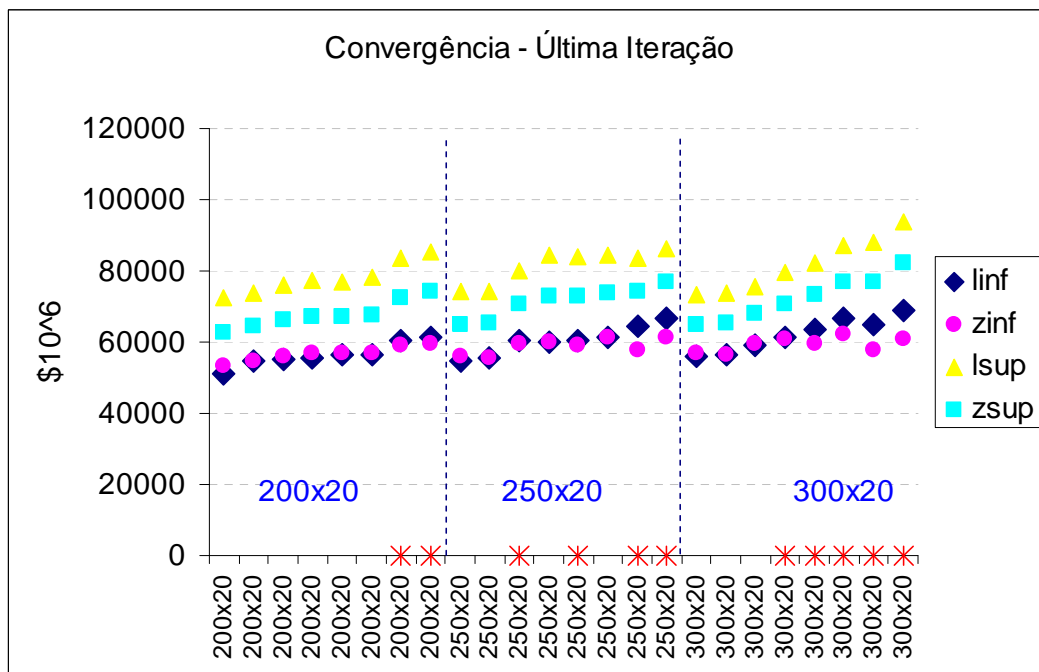


Figura 148: Convergência – Última Iteração – Variação Forward & Amostra – Opção 4 - FEV/07

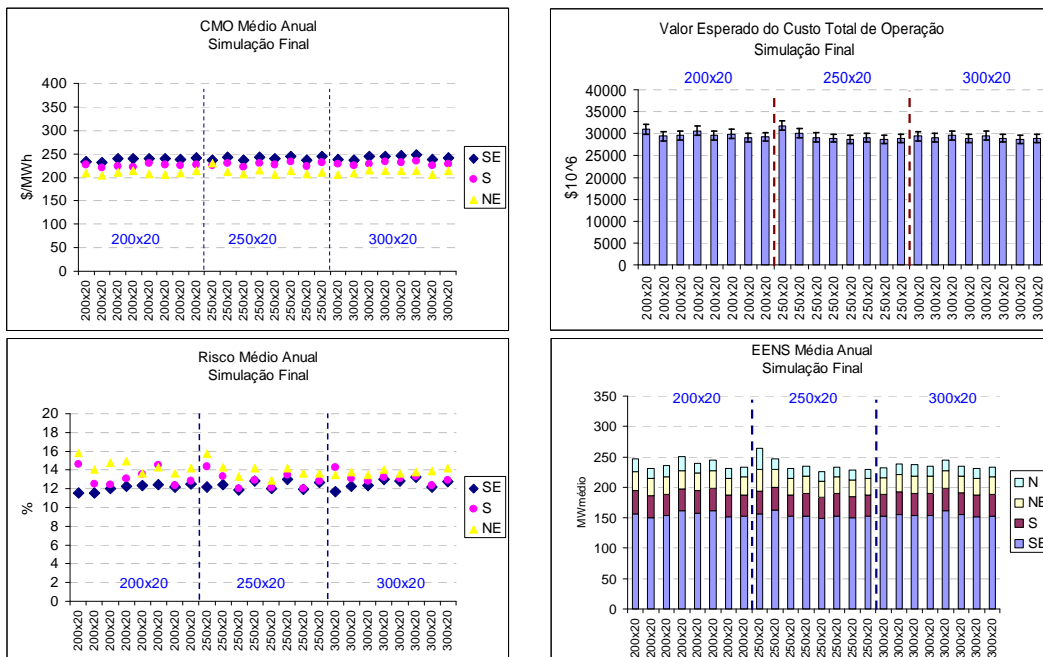


Figura 149: Resultados – Variação Forward & Amostra – Opção 4 - FEV/07

Em resumo, na sensibilidade com relação ao número de cenários forward, todas as opções analisadas apresentam menor variação nos resultados tanto na simulação final quanto no cálculo da política se comparadas com a opção Atual. Porém este ganho é menor do que o ganho observado na sensibilidade com relação ao número de cenários backward e à amostra de ruídos.

Para aumentar a robustez dos resultados quando é alterado o tamanho da amostra forward, deve-se tentar investigar outros métodos para construção dos cenários forward a serem visitados durante o cálculo da política.

## 8.6.

### Alteração do tamanho da amostra para o processo de agregação

É esperado que à medida que o tamanho da amostra original de ruídos utilizada no processo de agregação aumente menor seja a variação observada nos resultados. Porém, vale ressaltar que o tempo computacional gasto com a construção da árvore de cenários cresce consideravelmente com o aumento da amostra original de ruídos utilizada no processo de agregação.

Na Figura 150 são apresentados os resultados de ZSUP e ZINF para a última iteração, considerando a opção 4 utilizando diferentes amostras de ruídos e com tamanhos variando de 2 mil a 100 mil.

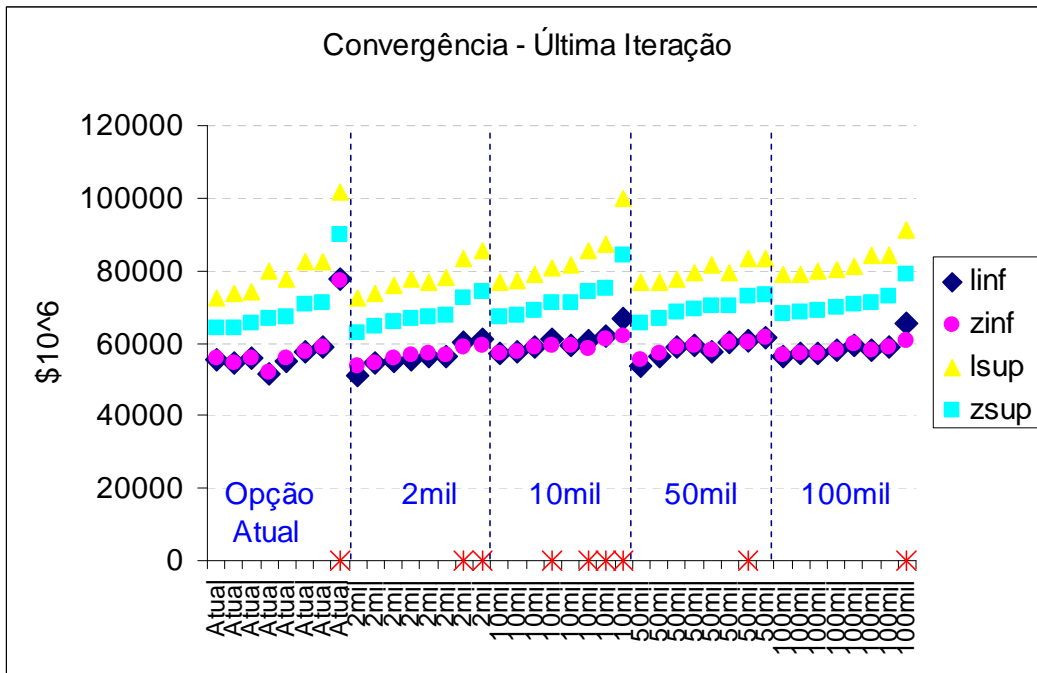


Figura 150: Convergência – Última Iteração – Variação Forward & Amostra – Opção 4 - FEV/07

Observando os resultados da simulação final, pode-se verificar que os valores obtidos com a amostra de tamanho 2 mil para o custo total de operação (Figura 151) e para o custo marginal de operação – CMO - (Figura 152) pouco se diferenciam dos valores obtidos com casos que utilizam uma amostra maior.

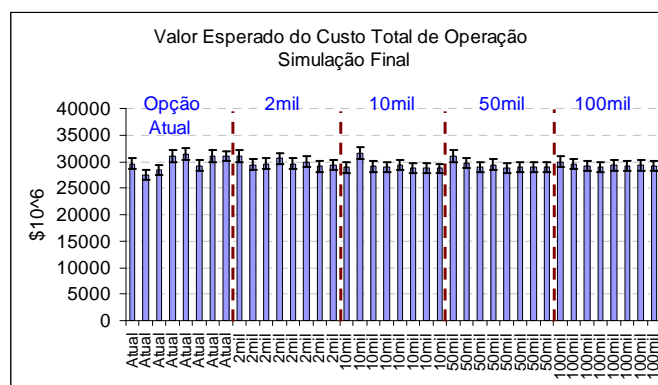


Figura 151: COPER – Variação Tamanho da Amostra - FEV/07

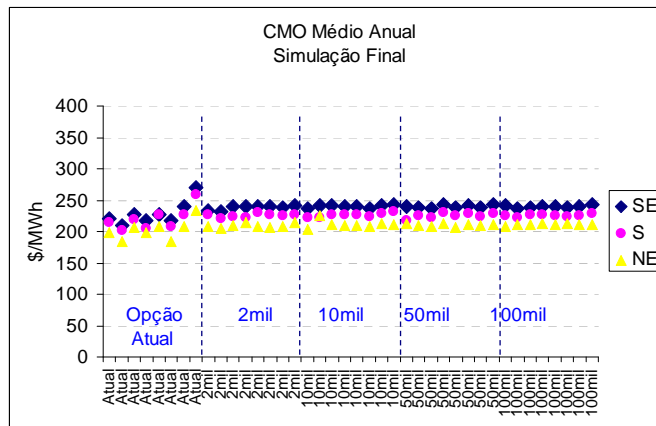


Figura 152: CMO – Variação Tamanho da Amostra - FEV/07

Como pode ser observado na Figura 153, há uma aparente estabilização nos valores do risco médio anual à medida que o tamanho da amostra cresce, porém isto não é verificado no valor esperado da energia não suprida, Figura 154. Como a variável risco de déficit não faz parte diretamente da função objetivo do problema de planejamento da operação, esta pode apresentar uma oscilação sem que haja necessariamente um aumento no montante de energia não suprida.

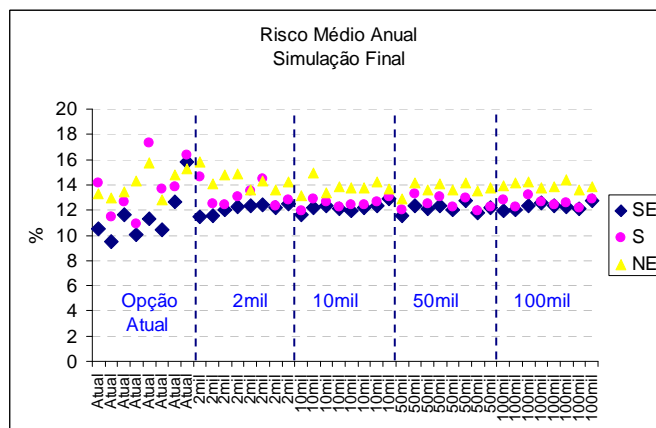


Figura 153: Risco – Variação Tamanho da Amostra - FEV/07

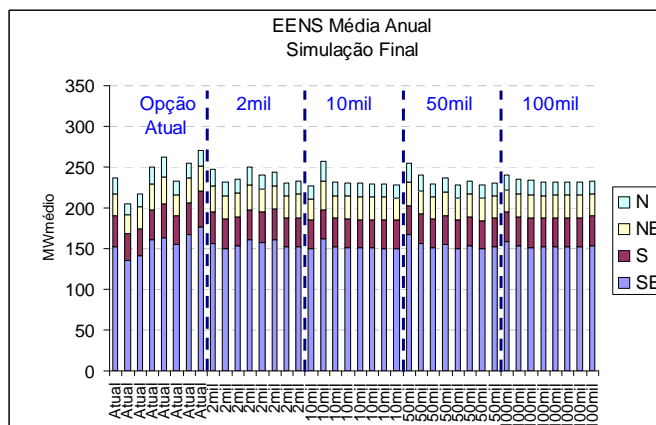


Figura 154: EENS – Variação Tamanho da Amostra - FEV/07

Os resultados observados neste item, em conjunto com as avaliações do tamanho da amostra realizadas no capítulo 7, indicam que uma amostra de tamanho 2mil é razoável para representar bem o processo estocástico multivariado. Esta afirmação é válida para as configurações utilizadas neste estudo. Para casos como o Plano Decenal de Expansão, que normalmente possuem um número maior de subsistemas equivalentes de energia, o dimensionamento da amostra a ser utilizado no processo de agregação deve ser reavaliado.

### **8.7. Alteração do método de amostragem**

Neste item é analisada a troca do método de amostragem aleatória simples utilizado na opção 1AAS pelos métodos hipercubo latino (LHC) e amostragem descritiva (AD) para gerar as amostras backward e forward. O caso analisado utiliza 200 cenários forward e 20 aberturas e são consideradas diferentes amostras de ruídos (variação de semente).

Na Figura 155 são apresentados os resultado da última iteração. Pode-se observar que o valor de ZSUP para os casos que utilizam LHC e AD é maior do que o observado com o método AAS, assim como sua a variância, representada pelo intervalo de confiança. Com relação ao valor de ZINF, este é mais estável e tem a mesma magnitude observada nos caso AAS.

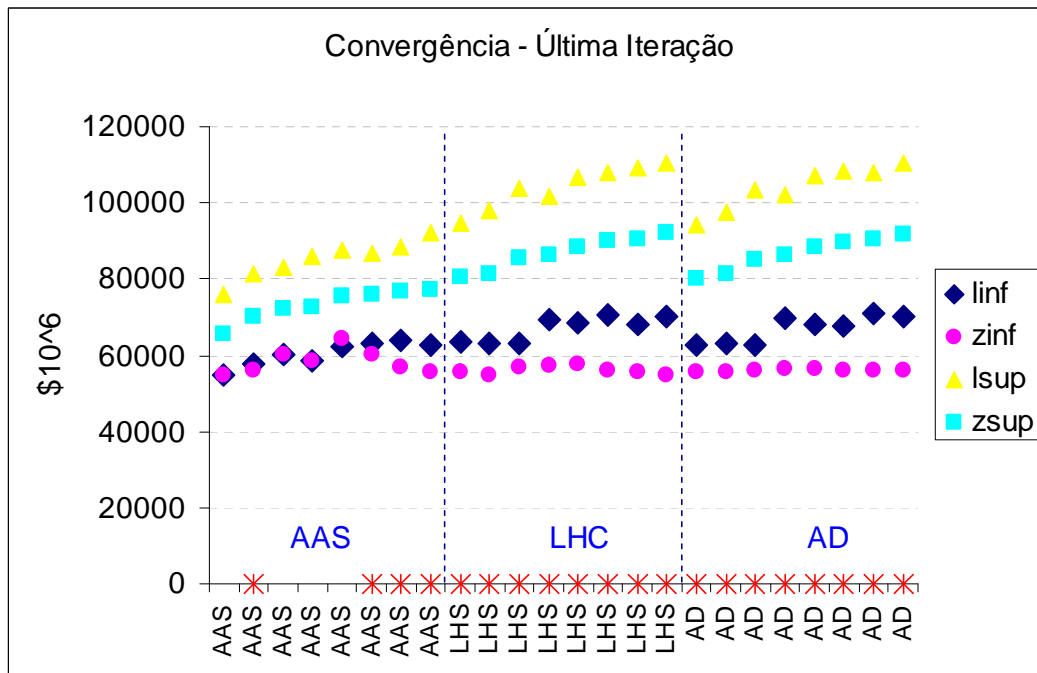


Figura 155: Convergência – Última Iteração – Variação Método de Amostragem - FEV/07

Os resultados de CMO, COPER, risco e EENS, apresentados na Figura 156, também apresentam uma maior estabilidade com relação aos mesmos resultados observados com o método AAS. Todos os casos LHC e AD atingem o número máximo de iterações sem que o critério de convergência estatístico tenha sido alcançado. O mesmo ocorre com alguns casos AAS.

Apesar dos casos AAS terem convergido em iterações diferentes, a oscilação de resultados é verificada mesmo se os casos forem forçados à 45ª iteração.

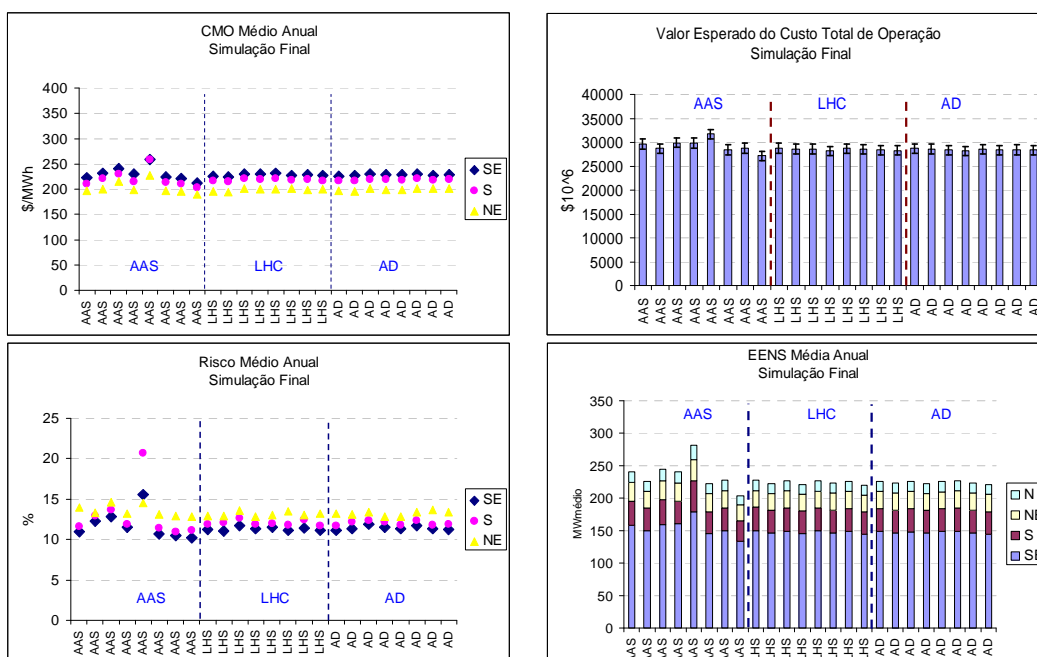




Figura 156: Resultados- Variação Método de Amostragem - FEV/07

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

Com relação à variação do número de cenários backward, pode-se observar na Figura 157 que não houve variações nos resultados de ZINF e ZSUP obtidos na última iteração. Porém, pode-se verificar novamente que o valor de ZSUP para os métodos de amostragem LHC e AD estabilizam em um patamar bem superior ao valor observado com AAS. Na Figura 158 são apresentados valores médios anuais obtidos para CMO, risco e EENS, além do custo total de operação da simulação final. Esses resultados são se alteram nos casos LHC e AD e nos casos AAS apresentam uma pequena variação.

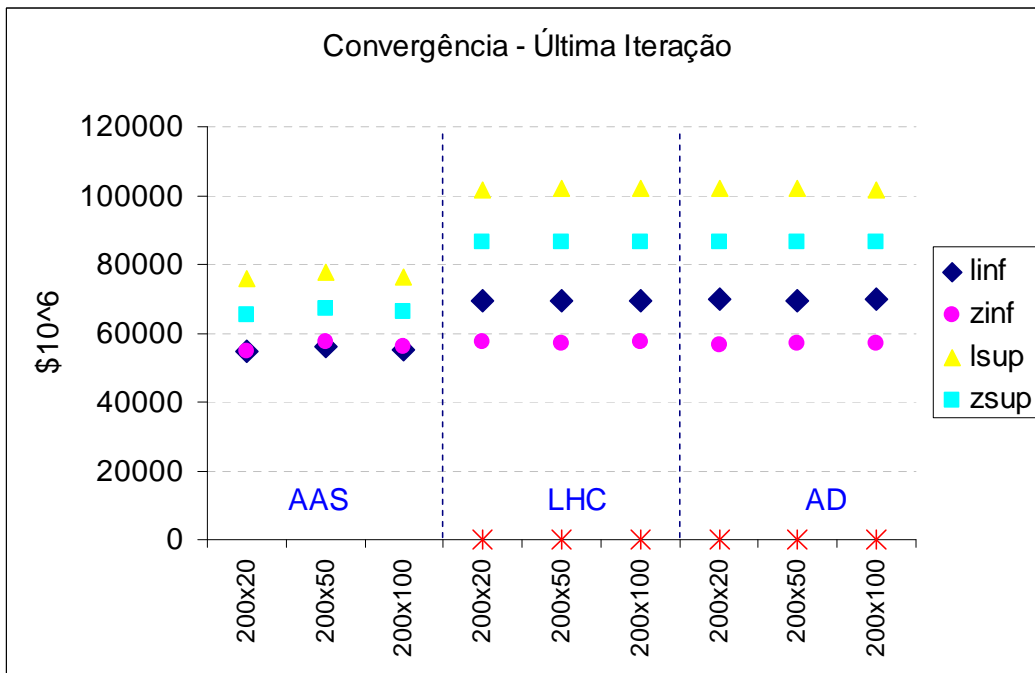


Figura 157: Convergência – Última Iteração – Variação Método de Amostragem e Backward - FEV/07

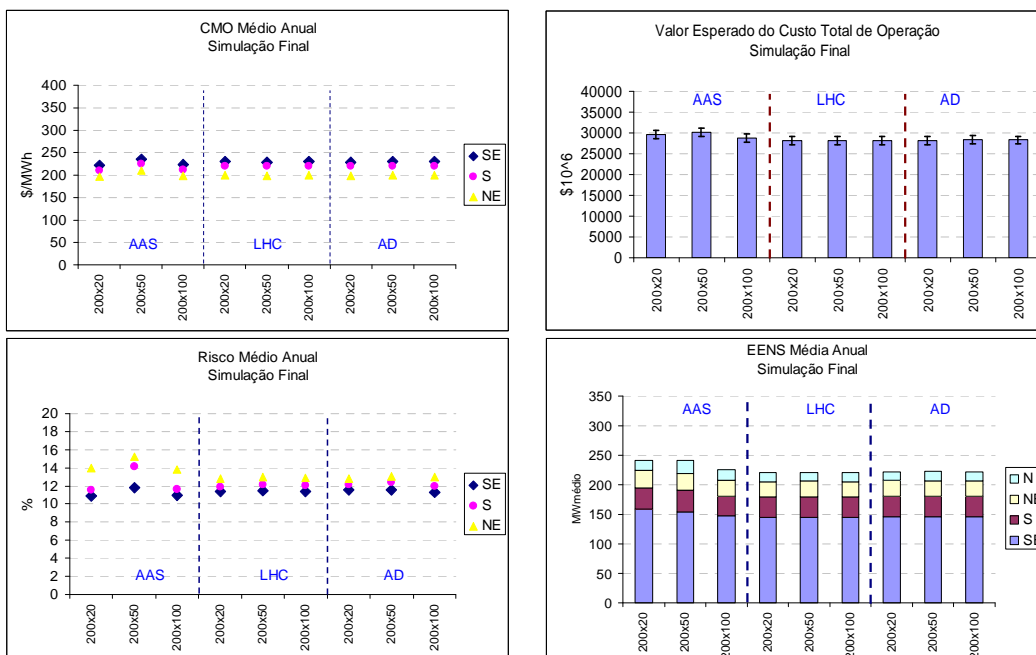


Figura 158: Resultados- Variação Método de Amostragem e Backward - FEV/07

Na Figura 159 e Figura 160 são mostrados os resultados de uma sensibilidade com relação ao número de cenários forward. Da mesma forma como observado anteriormente, os resultados não apresentam variação com a mudança do número de cenários forward, porém há um acréscimo no valor de ZUP para os casos LHC e AD.

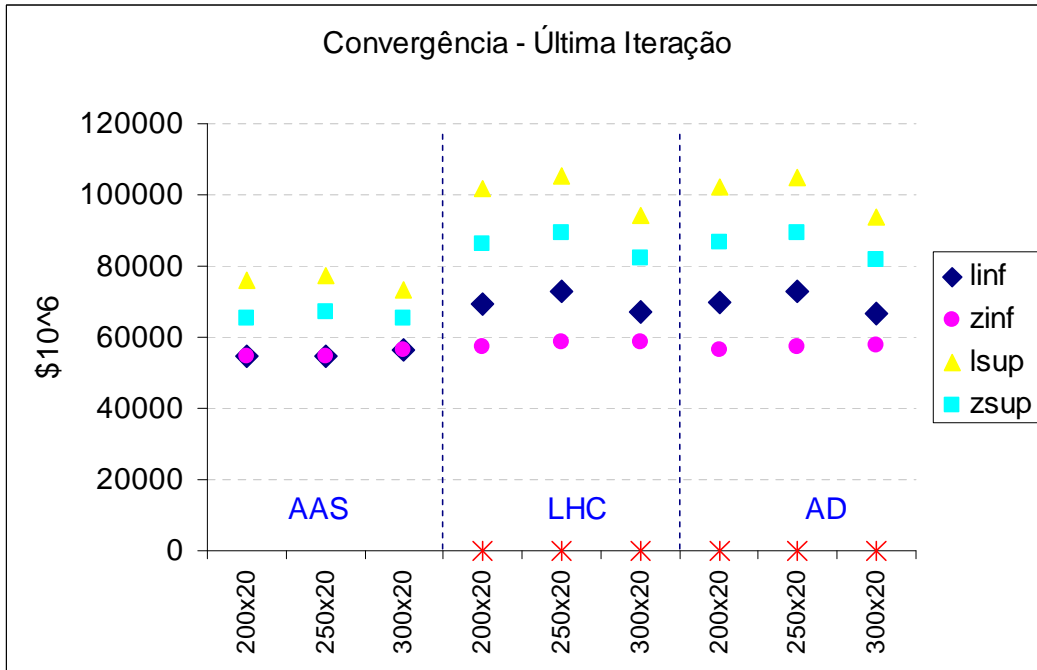


Figura 159: Convergência – Última Iteração – Variação Método de Amostragem e Forward - FEV/07



Figura 160: Resultados- Variação Método de Amostragem e Forward - FEV/07

## 8.8. Combinação de método de amostragem e agregação

Neste item são apresentadas três novas opções para construção da árvore de cenários, derivadas da aplicação em conjunto do método de amostragem LHC e das

técnicas de agregação. A opção 1+LHC é baseada na opção 1, a mudança ocorre somente na geração da amostra de ruídos para o passo forward, onde a amostragem aleatória é substituída pela amostragem LHC, conforme ilustrado na Figura 161a. A opção 1LHC também é derivada da opção 1, onde o método AAS é substituído pelo método LHC na geração da amostra forward e da amostra que será utilizada no processo de agregação, de acordo com a Figura 161b. Por último, o método AAS é trocado pelo método LHC na opção 4, dando origem à opção 4LHC, Figura 161c.

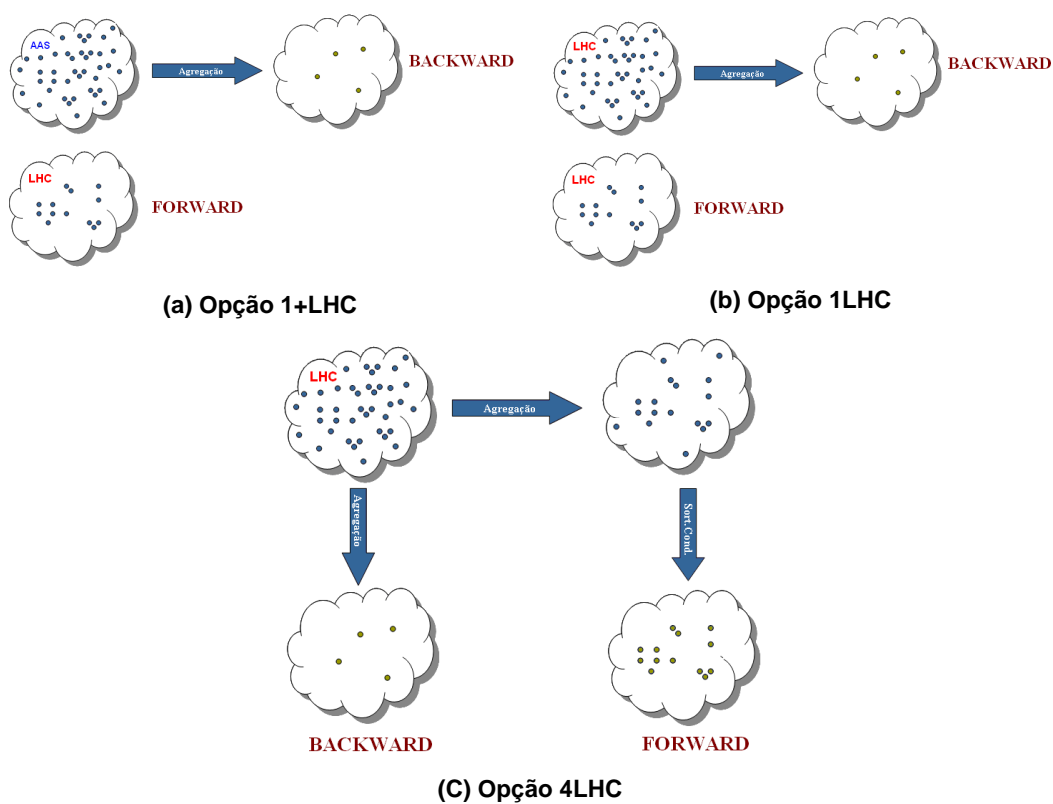


Figura 161: Opções com combinação de LHC com agregação

Na Figura 162 são apresentados os valores de ZINF e ZSUP obtidos durante a última iteração do processo de convergência. Pode-se observar que as opções que utilizam o método LHC diretamente para construir a amostra forward apresentam uma estimativa mais alta para o valor de ZSUP.

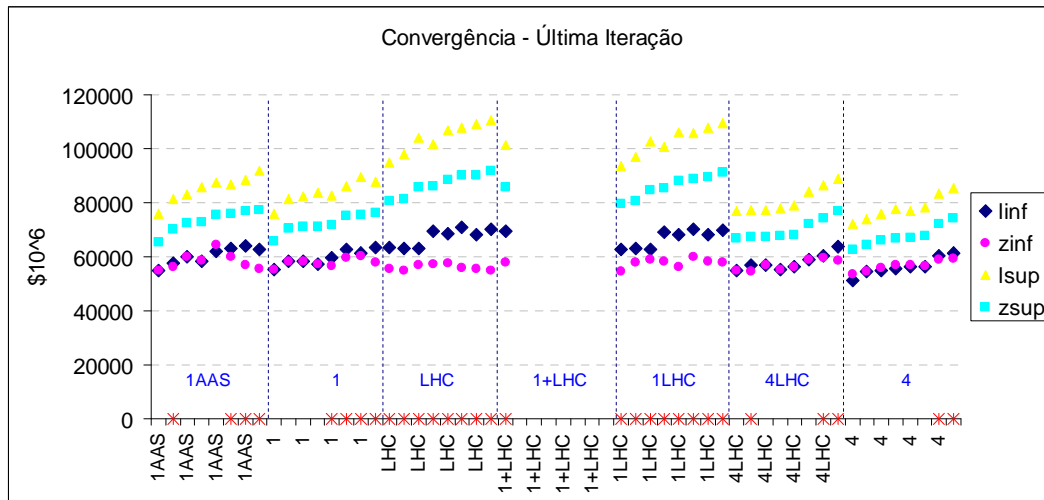


Figura 162: Convergência – Última Iteração – Variação Amostra – Combinação LHC e Agregação - FEV/07

Da Figura 163 a Figura 166 são mostrados os resultados da simulação final para todas as opções analisadas. Como pode ser observado, todas as opções que empregam o método de amostragem LHC apresentam uma menor variabilidade nos resultados, em especial as opções que a amostra original a ser agregada é gerada utilizando LHC (opções 1LHC e 4LHC).

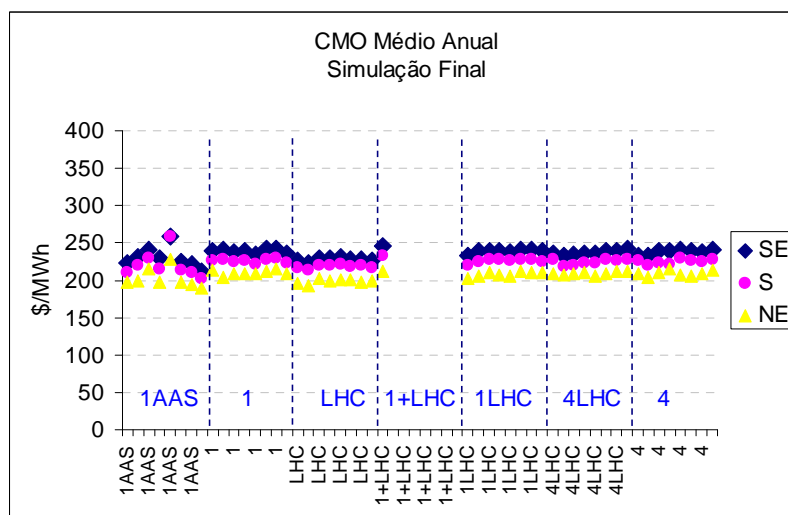


Figura 163: CMO - Variação Amostra – Combinação LHC e Agregação - FEV/07

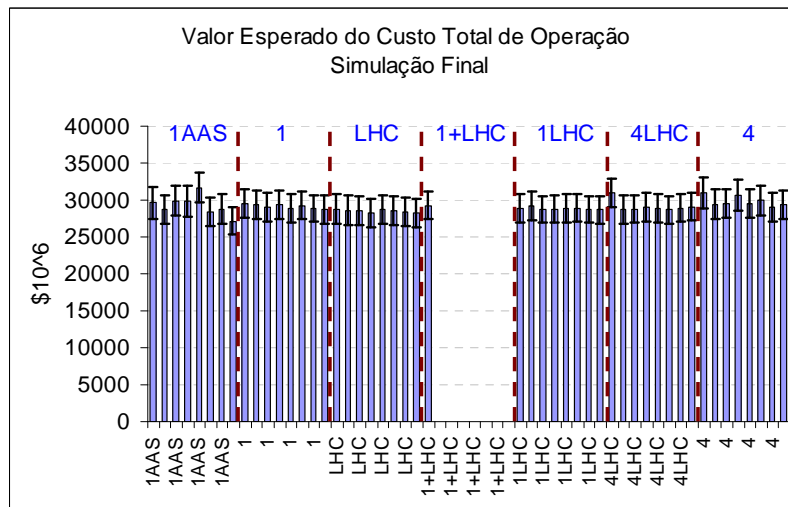


Figura 164: COPER - Variação Amostra – Combinação LHC e Agregação - FEV/07

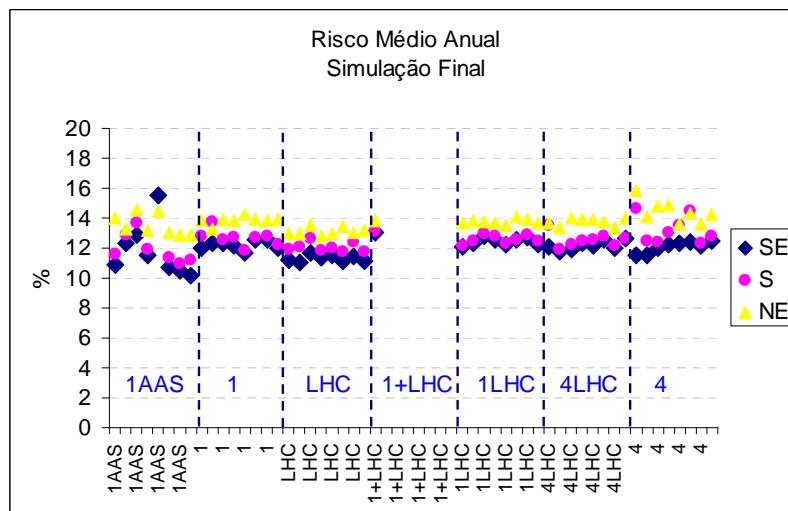


Figura 165: Risco - Variação Amostra – Combinação LHC e Agregação - FEV/07

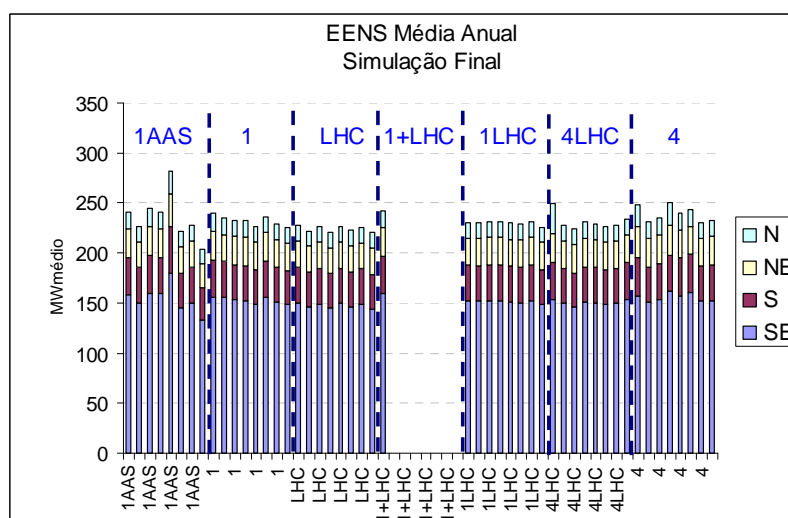


Figura 166: EENS - Variação Amostra – Combinação LHC e Agregação - FEV/07

Com relação à variação do número de cenários forward, pode-se verificar uma melhoria com relação à opção 4 quando se utiliza o método LHC para construção da amostra original. Para as opções derivadas da opção 1, verifica-se uma maior variabilidade nos valores de ZSUP. Os resultados estão ilustrados na Figura 167.

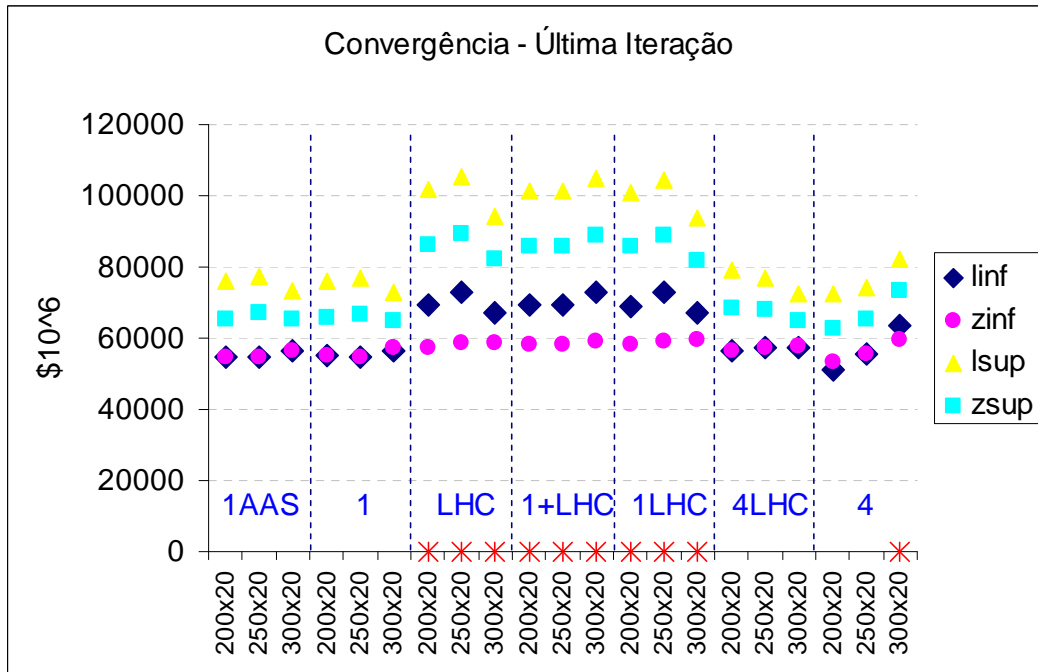
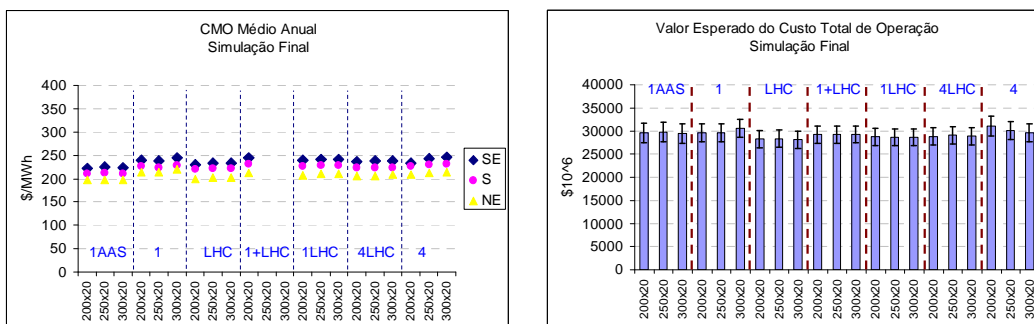


Figura 167: Convergência – Última Iteração – Variação Forward – Combinação LHC e Agregação - FEV/07

Na Figura 168 estão ilustrados os resultados obtidos durante a simulação final. Novamente a opção 4LHC traz algum ganho em comparação com a opção 4, principalmente com relação à variável risco de déficit.



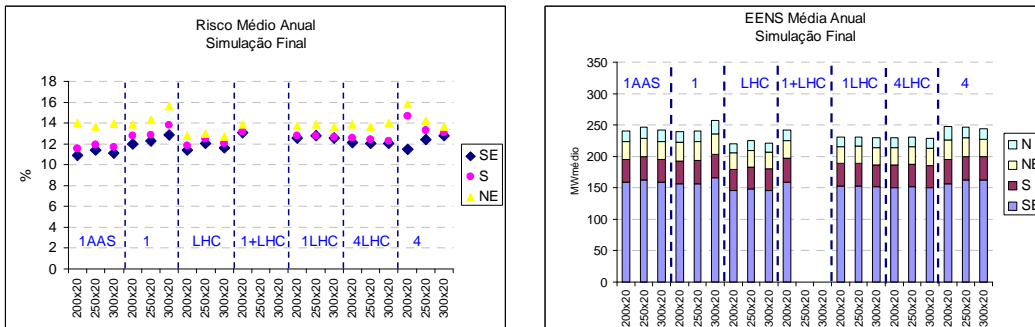


Figura 168: Resultados - Variação Forward – Combinação LHC e Agregação - FEV/07

No estudo de sensibilidade com relação ao número de aberturas, quase não há variação nos resultados de todas as opções analisadas, conforme pode ser constatado na Figura 169. Novamente, as opções que utilizam LHC para construção dos cenários forward apresentam um valor elevado de ZSUP.

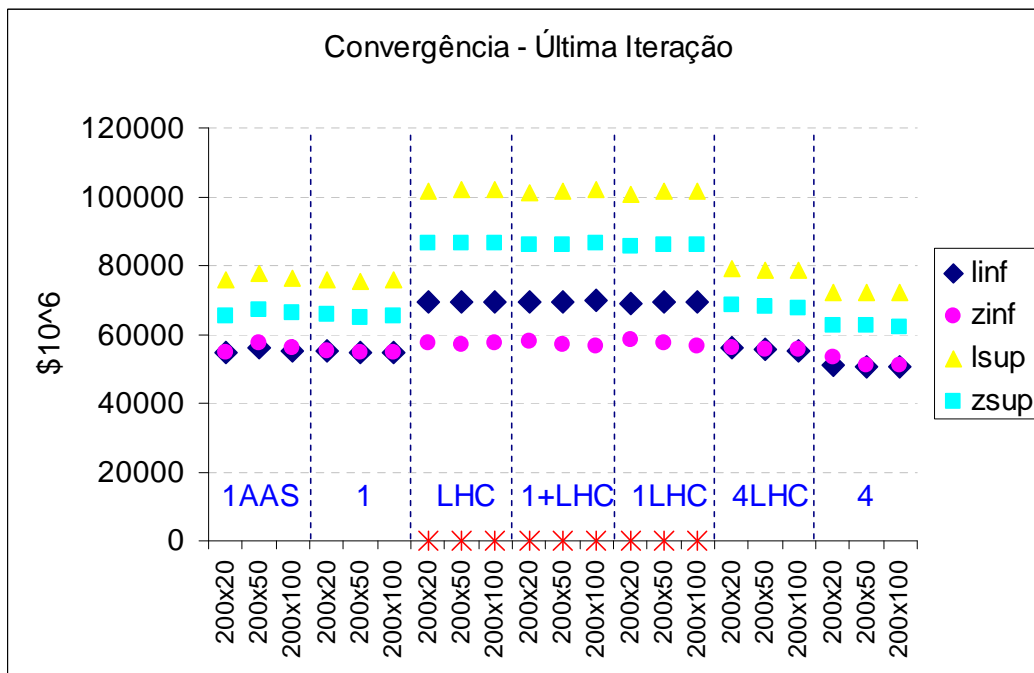


Figura 169: Convergência – Última Iteração – Variação Backward – Combinação LHC e Agregação - FEV/07

Em geral, as variações observada nos resultados da simulação final são pequenas, exceto para a variável risco de déficit, Figura 170.



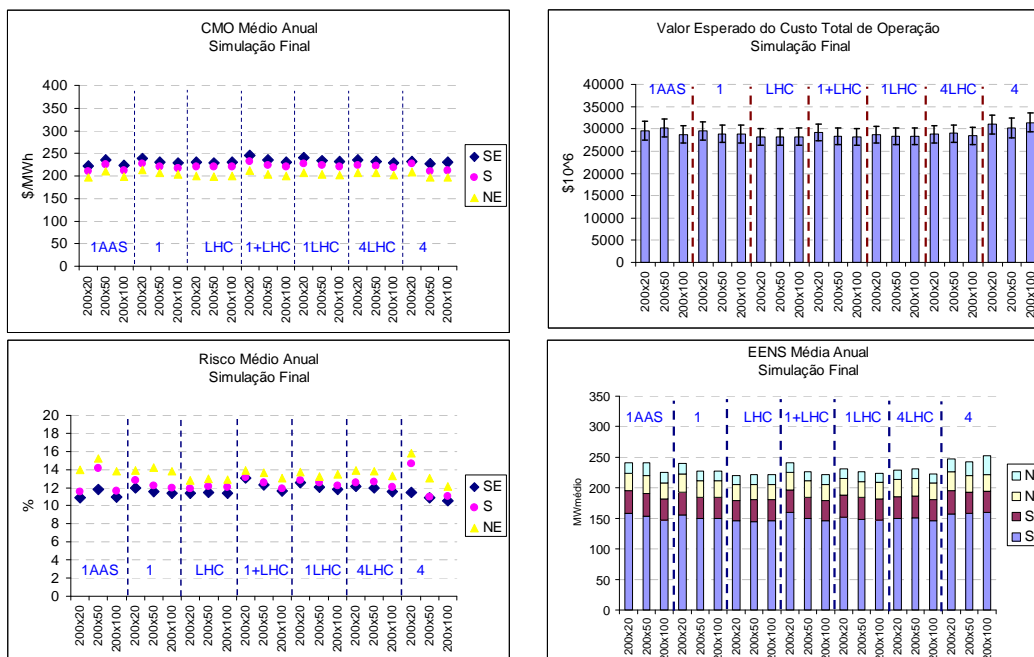


Figura 170: Resultados - Variação Backward – Combinação LHC e Agregação - FEV/07

### 8.9. Verificação da estabilidade dos resultados

É esperado que, após a convergência estatística do algoritmo de resolução do problema, a solução encontrada seja uma solução estável, isto é, que os resultados obtidos não se alterem caso sejam realizadas mais iterações no algoritmo de solução.

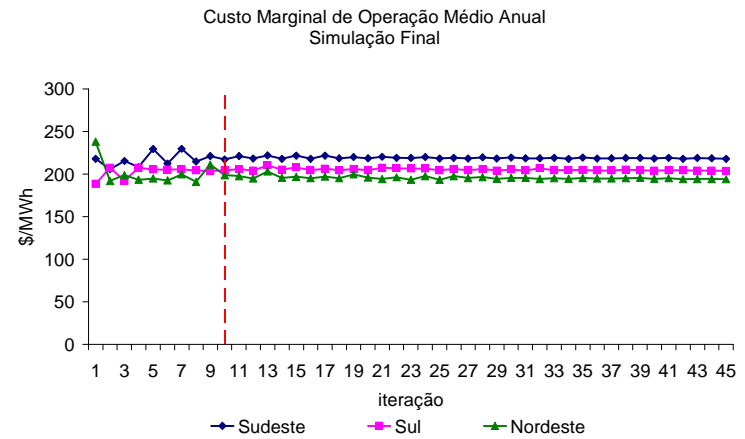
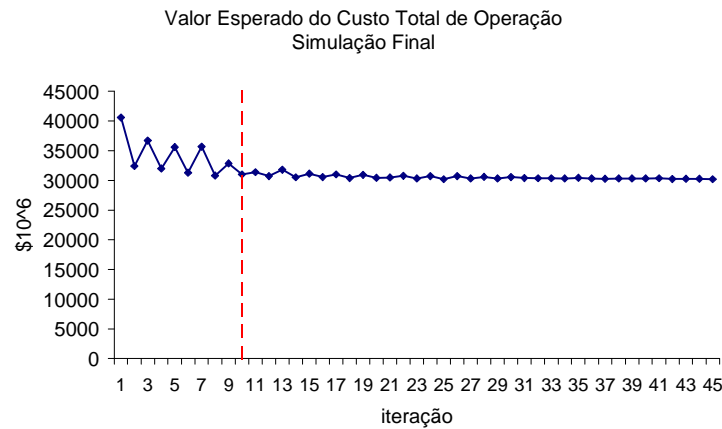
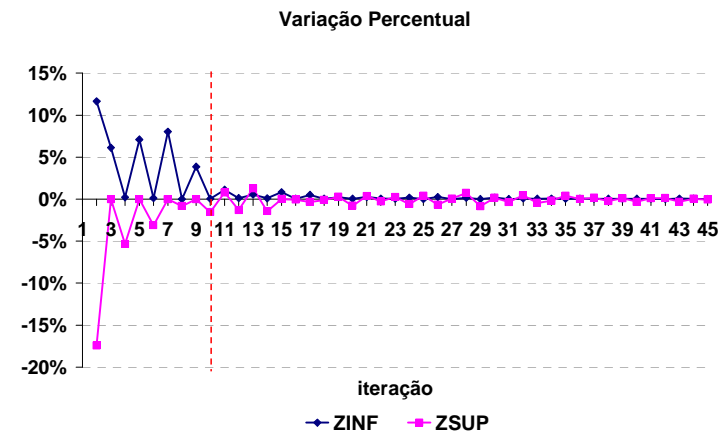
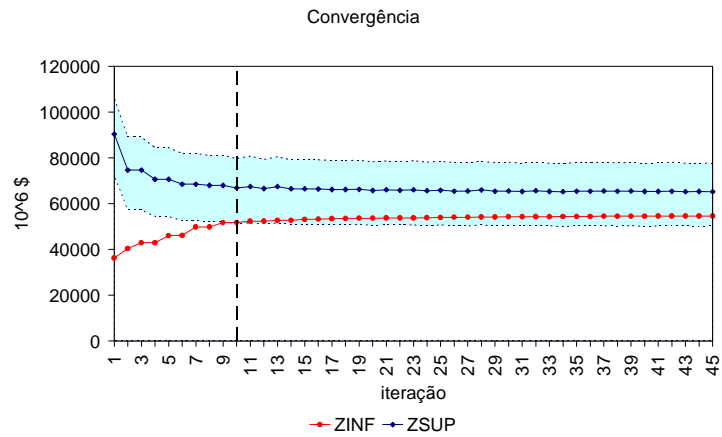
Para verificar a estabilidade da solução encontrada, deve-se analisar os resultados na simulação final considerando-se a função de custo futuro construída a cada iteração após a convergência estatística do processo. Para tanto, o algoritmo de solução é forçado a realizar tantas iterações quanto for o número máximo de iterações (45, neste caso). Logo após, são executadas 45 simulações da política de operação, de forma que a primeira simulação seja realizada utilizando a FCF construída até a primeira iteração, a segunda simulação seja realizada considerando a FCF construída até a segunda iteração, e assim por diante até que a 45ª simulação seja realizada com a FCF completa. Espera-se que os resultados obtidos após a iteração onde ocorre a convergência estatística não se alterem significativamente.

A seguir são apresentadas as análises para as opções Atual, 0 e 4 sem reamostragem de cenários e 0 e 4 com reamostragem, Figura 171, Figura 172, Figura 173, Figura 174 e Figura 175, respectivamente. Nas figuras a seguir, a iterações onde o caso converge está destacada por uma linha tracejada. São avaliados os seguintes

resultados: trajetória da convergência, variação dos valores de ZINF e ZSUP, custo total de operação, custo marginal de operação, risco e energia não suprida. Pode-se observar que após a convergência estatística do algoritmo de solução não há uma variação significativa nos resultados da política de operação.

Nas opções Atual, 0 e 4, o modelo NEWAVE considera sempre a melhor solução encontrada, neste caso o menor valor para o custo total de operação (ZSUP), durante o cálculo da política de operação. Porém, após atingir a convergência estatística, esta consideração deixa de ser válida e a partir desta iteração em diante passa a valer o valor de ZSUP obtido na iteração corrente.

a) Opção Atual



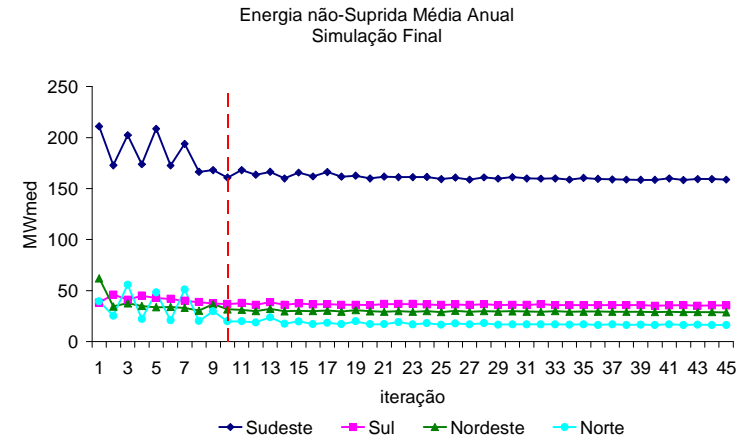
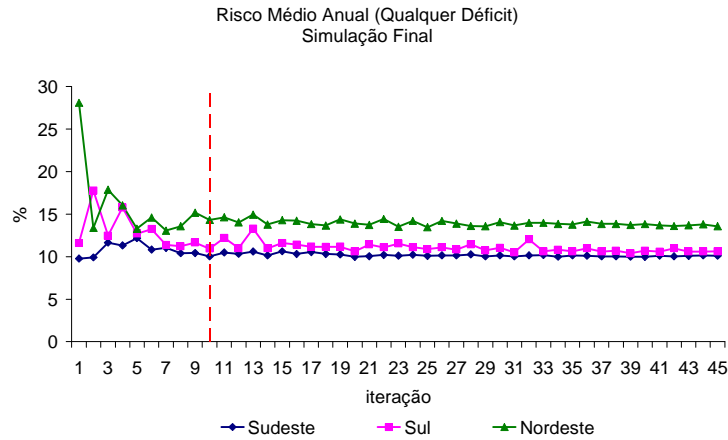
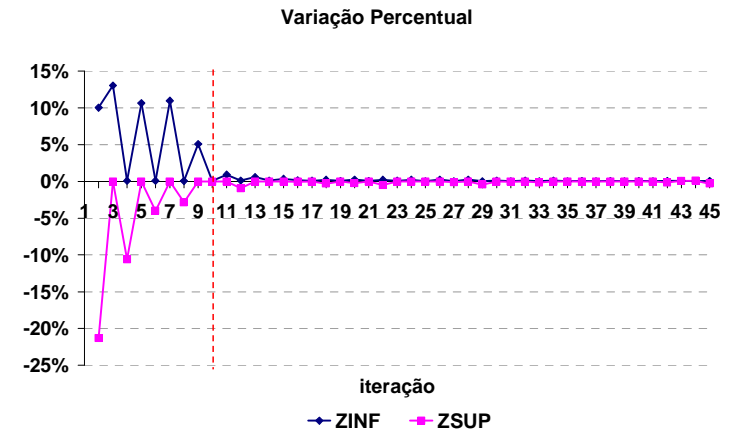
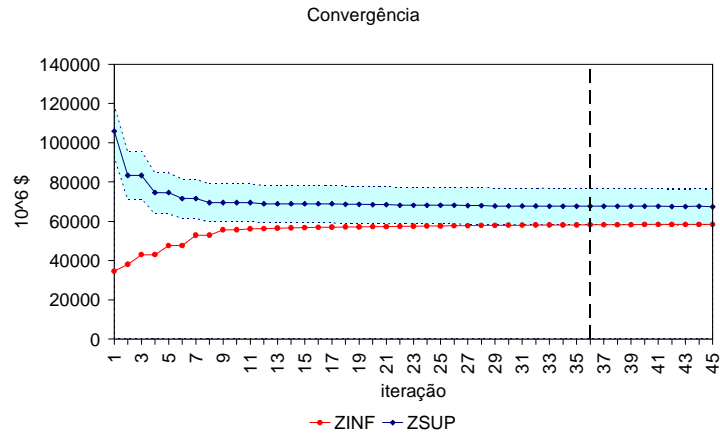


Figura 171: Resultados 45ª iteração – Opção Atual

b) Opção 0



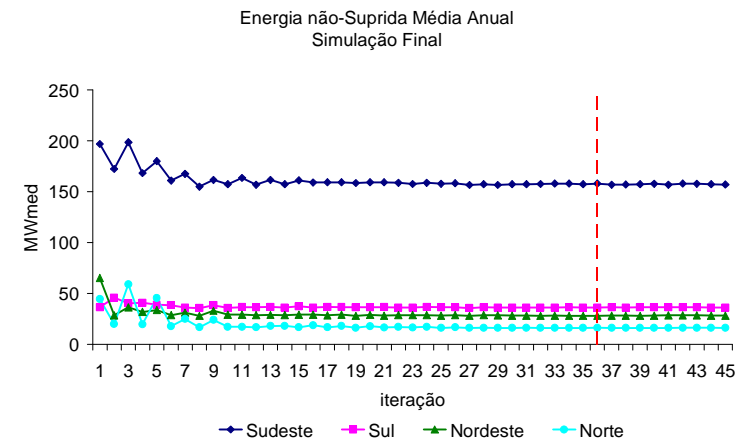
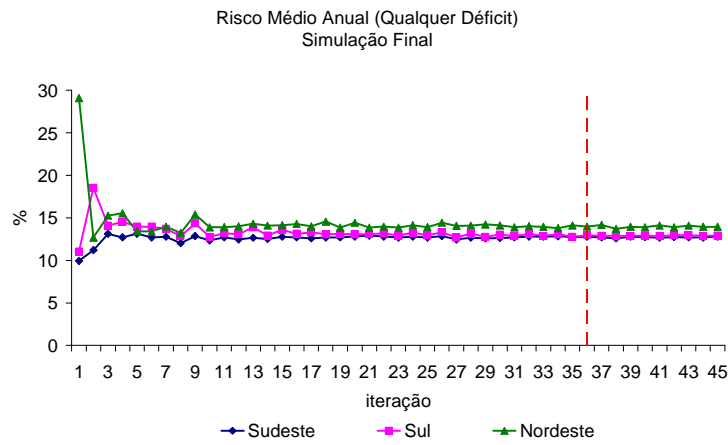
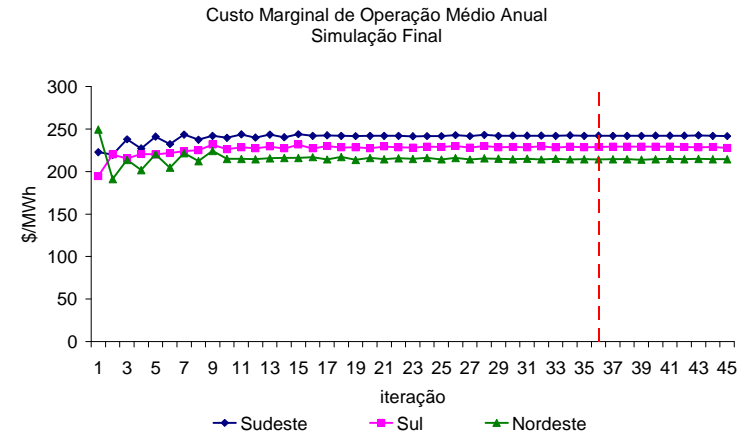
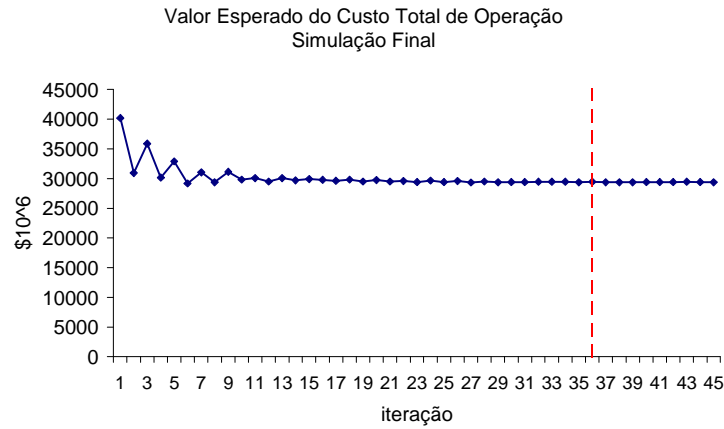
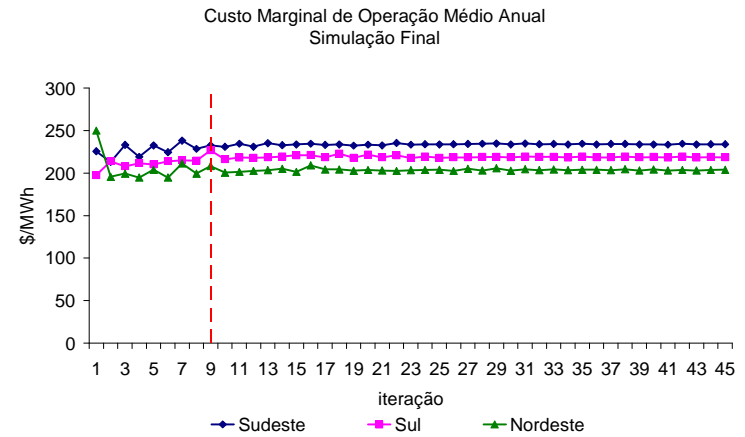
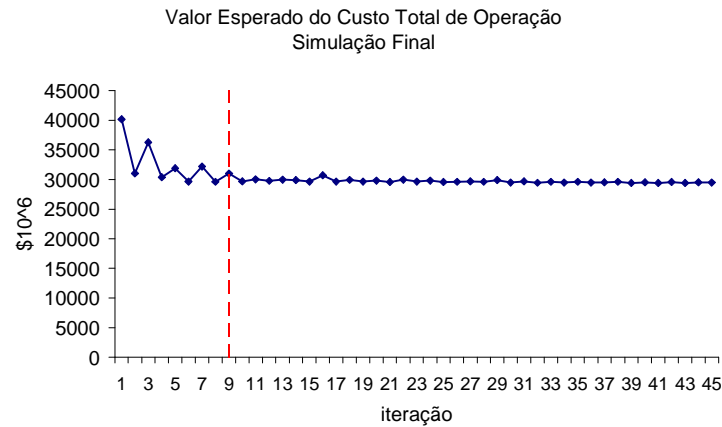
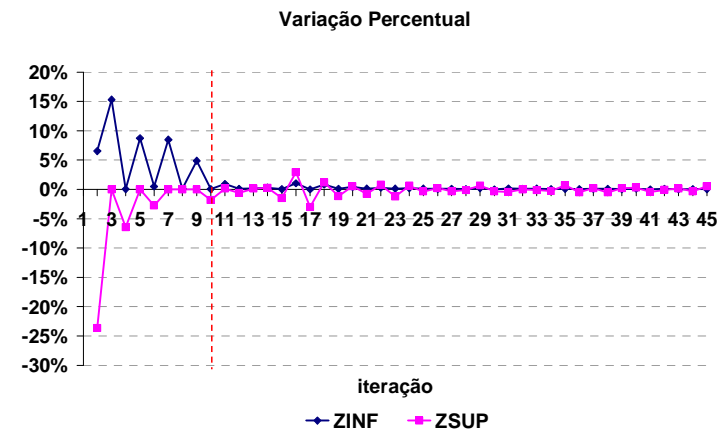
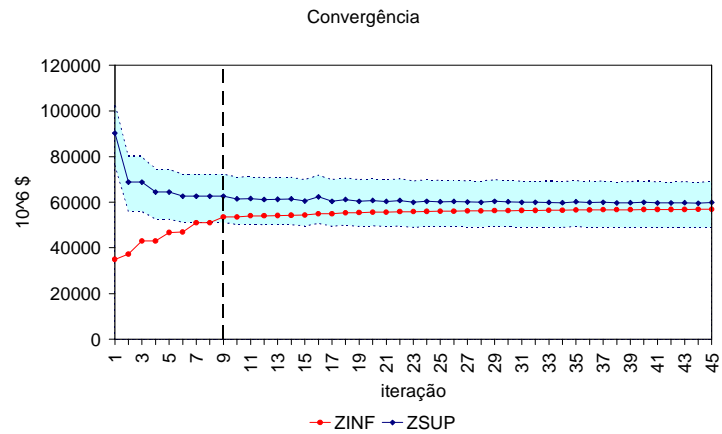


Figura 172: Resultados 45ª iteração – Opção 0

c) Opção 4



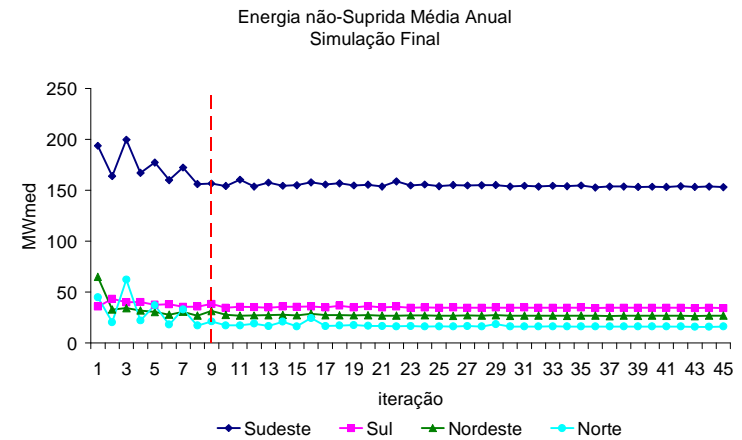
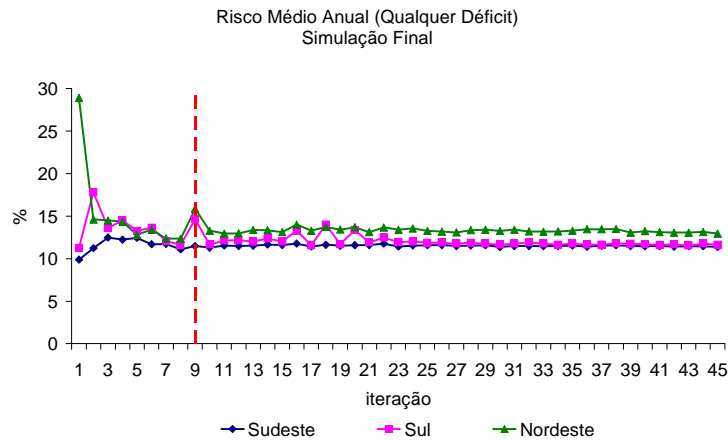
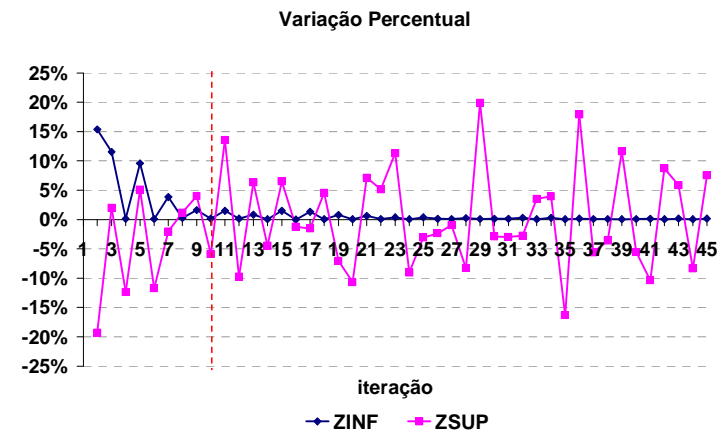
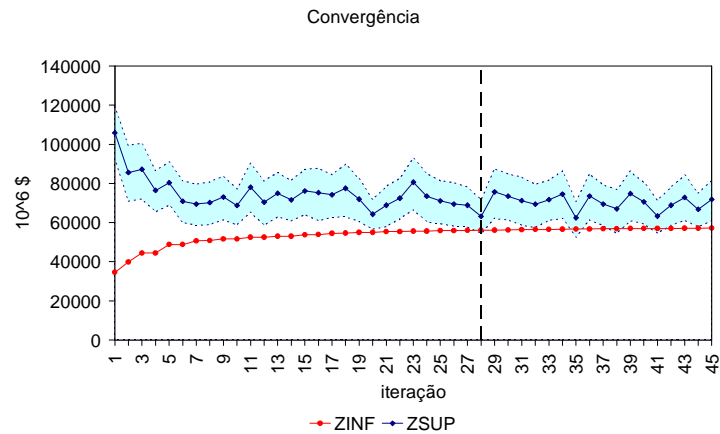


Figura 173: Resultados 45ª iteração – Opção 4

d) Opção 0 com reamostragem



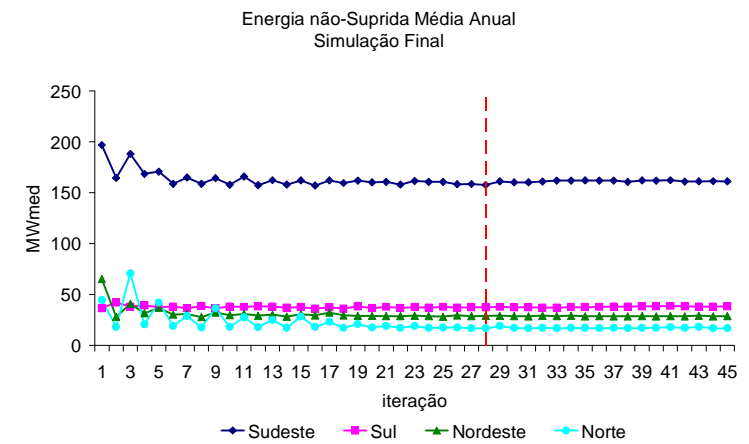
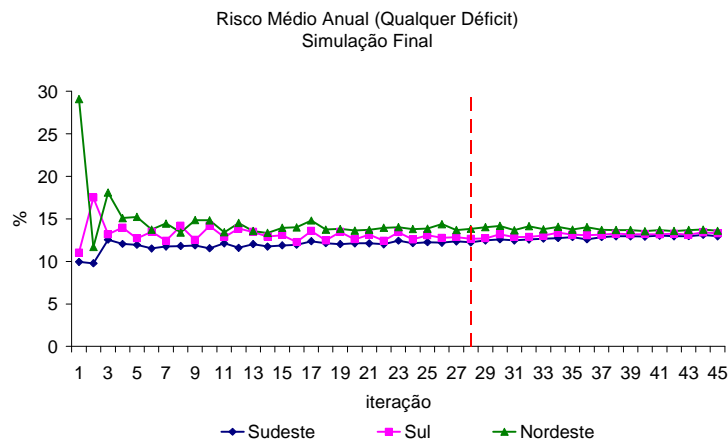
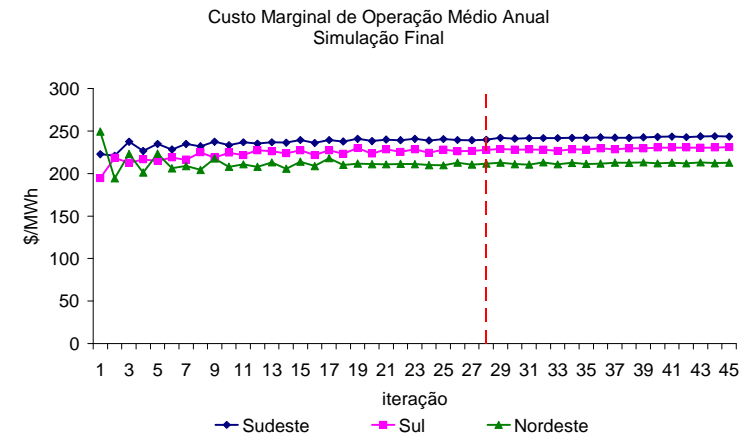
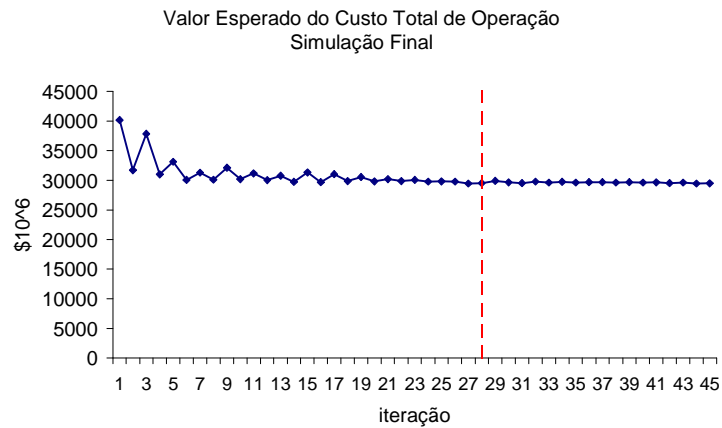
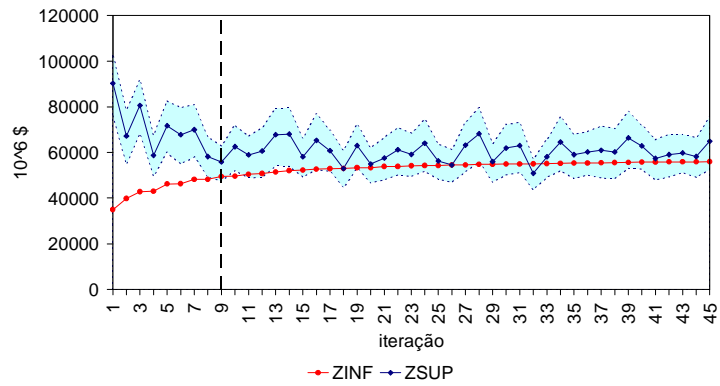


Figura 174: Resultados 45ª iteração – Opção 0 com reamostragem

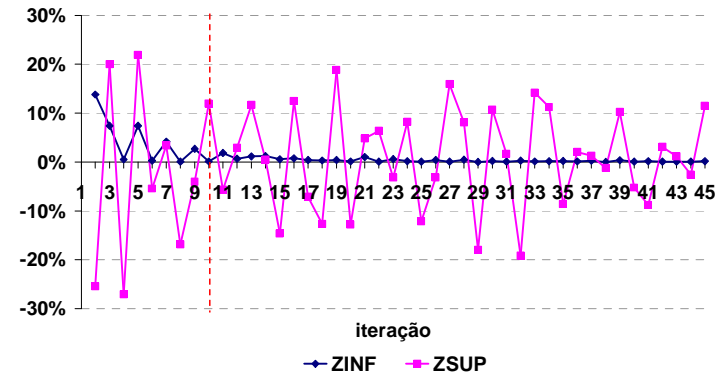
e) Opção 4 com reamostragem



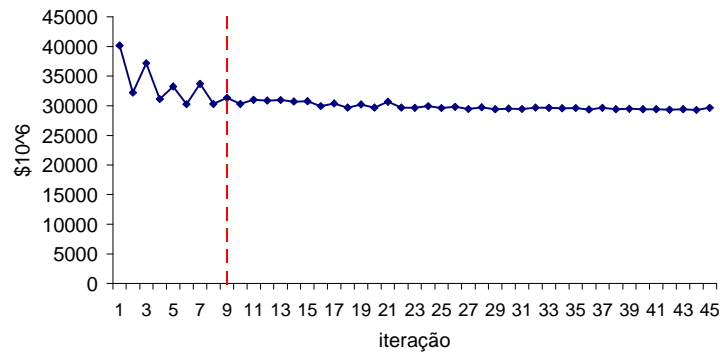
Convergência



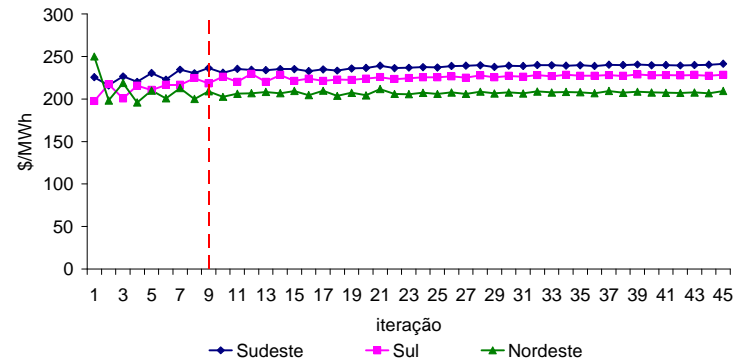
Variação Percentual



Valor Esperado do Custo Total de Operação  
 Simulação Final



Custo Marginal de Operação Médio Anual  
 Simulação Final



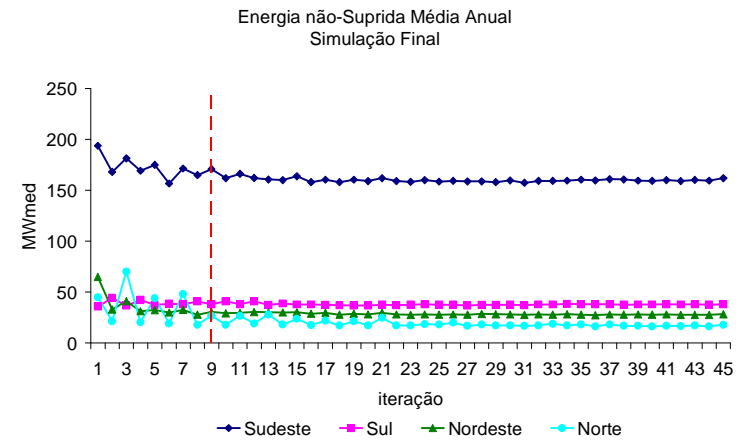
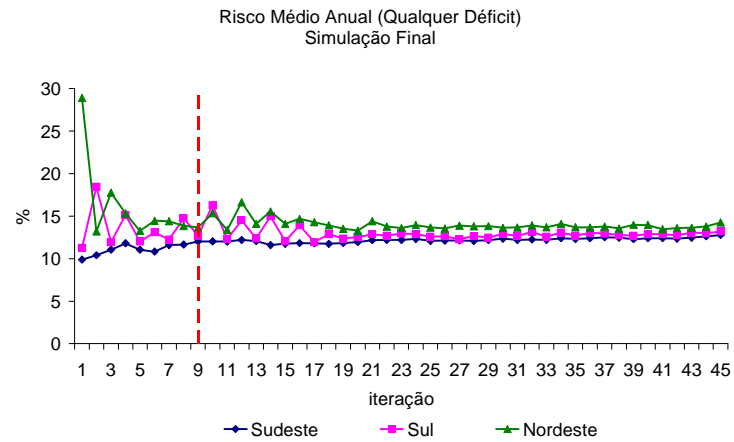


Figura 175: Resultados 45ª iteração – Opção 4 com reamostragem

## **8.10. Resumo**

Neste capítulo foi dado um maior destaque para a opção 4 pois esta apresentou melhores resultados nas análises preliminares. A opção 0 também foi estudada exaustivamente por ser a opção que mais se aproxima da opção Atual.

O aumento do tamanho da amostra backward tem um impacto significativo no tempo computacional de resolução do problema de planejamento da operação. Contudo não é possível usar uma amostra de tamanho muito reduzido, pois é necessário representar de forma adequada o processo estocástico do problema. Nas opções que utilizam as técnicas de agregação para geração da amostra backward, é possível conseguir resultados tão bons utilizando uma amostra de tamanho reduzido, quanto aqueles obtidos com amostras maiores.

As opções que utilizam técnicas de agregação apresentaram menor variação nos resultados se comparadas com a opção Atual, com relação às diversas sensibilidades efetuadas neste capítulo. Em geral, a aplicação do processo de reamostragem resultou em uma redução do custo total de operação durante o cálculo da política ótima de operação.

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

No próximo capítulo será apresentada uma retrospectiva das principais contribuições deste trabalho e serão indicadas sugestões para trabalhos futuros.

## 9 Conclusões e Sugestões para Trabalhos Futuros

O extenso horizonte do planejamento da operação e a estocasticidade das afluições futuras tornam o problema de planejamento da operação energética do sistema hidrotérmico brasileiro um problema de grande porte e de difícil solução, sendo necessária sua divisão em diversas etapas, de tal forma que cada etapa considere diferentes graus de detalhamento do sistema, horizonte de estudo e representatividade da incerteza hidrológica.

O NEWAVE, desenvolvido pelo CEPEL, é o modelo atualmente utilizado para definição das metas para o planejamento da operação de médio prazo. O método de solução adotado é a programação dinâmica dual estocástica (PDDE), proposto na década de 90 por Pereira & Pinto (1991). Neste algoritmo a construção da função de custo futuro é construída iterativamente através do princípio da decomposição de Benders, sem a necessidade de discretizar o espaço de estados, evitando desta forma a explosão combinatória decorrente da PDE. A incerteza com relação às afluições futuras é representada por uma árvore de cenários, construída utilizando amostragem aleatória simples.

Neste trabalho foram apresentadas propostas para a definição da árvore de afluições a ser visitada durante o cálculo da política ótima de operação. As opções elencadas fazem uso de ferramental estatístico (técnicas de agregação) multivariado capaz de agrupar um conjunto de objetos, baseado em suas características, em grupos de tal forma que objetos similares estejam alocados em um mesmo grupo, e que os grupos formados sejam bem diferentes entre si. Adicionalmente, foi proposto utilizar outros métodos de amostragem com o objetivo de reduzir a variabilidade causada naturalmente pela amostragem aleatória simples (AAS).

Todas as opções propostas neste trabalho foram aplicadas no modelo de geração de cenários sintéticos de energia natural afluyente. A geração dos cenários sintéticos é feita empregando-se modelos autorregressivos periódicos (PAR(p)) para modelar o processo estocástico das afluições para cada subsistema equivalente de energia. Os cenários sintéticos são gerados de forma multivariada e tomando-se o cuidado para que não sejam gerados valores negativos para as afluições. Para tanto,

é ajustada uma distribuição lognormal três parâmetros à parcela aleatória do modelo PAR(p).

Foi apresentado o método utilizado atualmente para geração da árvore de cenários a ser percorrida durante o cálculo da política, onde a geração das amostras de ruídos para o passo forward utiliza resorteios das amostras de ruídos do passo backward obtida por AAS. Esse método, chamado neste trabalho de opção Atual, se mostrou pouco robusto com relação à estabilidade de resultados quando são alterados parâmetros utilizados na construção da árvore. Adicionalmente, este método de geração apresenta grande instabilidade na reprodução das médias, desvios-padrão e correlações cruzadas históricas de ENA. Mesmo considerando um maior número de cenários forward, mostrou-se que o ganho é pequeno, analisando-se a representatividade da série de ruídos. Uma alternativa para melhorar a instabilidade referida acima seria aumentar o tamanho da amostra backward, o que acarretaria um aumento significativo no tempo computacional necessário para solução do problema.

Uma alternativa à opção Atual é a opção 1AAS, que utiliza amostragem aleatória simples para geração das amostras ruídos para os passos forward e backward de forma independente. Nos experimentos realizados neste trabalho, essa opção se mostrou mais robusta que a opção Atual, principalmente com relação à variação no número de cenários, porém também apresentou grande variabilidade nos resultados quando a amostra de ruídos é alterada.

Uma das propostas apresentadas neste trabalho foi substituir o método de amostragem aleatória simples por métodos de amostragem estratificada, tais como o LHC e AD, com o intuito de reduzir a variação nos resultados causada por alterações nas amostras de ruídos. Como era de se esperar, a variabilidade com relação aos cenários hidrológicos são bem pequenas. Porém, esses métodos de amostragem apresentaram grande dificuldade em reproduzir uma característica muito importante da série histórica: a correlação cruzada. Há uma tendência de geração de valores mais baixos de correlação cruzada para os cenários da backward em comparação com os cenários do passo forward, causando impactos na construção da FCF e tornando o valor ZSUP mais elevado.

Foram apresentadas opções que aplicam técnicas de agregação para a construção da amostra de ruídos utilizada somente no passo backward, opções 0 e 1. Os resultados observados mostram que a opção 0 causa uma degradação acentuada no desvio-padrão das amostras de ruídos e de ENA. Este efeito permanece o mesmo

se o tamanho da amostra for aumentado. A opção 1 apresentou resultados satisfatórios para média e correlação cruzada, porém, há uma ligeira perda no desvio-padrão. A degradação do desvio-padrão ocorreu em todas as outras opções testadas neste trabalho que utilizaram agregação de cenários.

As opções 2, 3 e 4 aplicaram as técnicas de agregação para obtenção tanto da amostra forward quanto da amostra backward, de forma independente (opção 4) ou dependente (opções 2 e 3). A opção 3 apresentou um desempenho bem ruim no processo iterativo de solução do modelo de planejamento, por isso foi abandonada. As opções 2 e 4 apresentaram bons desempenho com relação à reprodução das médias e correlações cruzadas das ENAs. A opção 0 também apresentou os mesmos problemas mencionados na opção 3, mas não foi abandonada por representar uma transição entre as opções Atual e as opções 2 e 4.

Adicionalmente, foi empregado, em conjunto com as opções 0 e 4, um método de reamostragem de cenários com o objetivo de percorrer uma maior porção da árvore completa de cenários. Observou-se que, quando esse método é aplicado há, em geral, uma redução na variabilidade nos resultados obtidos.

As opções 2 e 4 apresentaram menor variação nos resultados se comparadas com a opção Atual, com relação às diversas sensibilidades efetuadas neste trabalho. Dentre as opções analisadas, a que mostrou melhor desempenho foi a opção 4, onde o processo de agregação é aplicado de forma independente, a partir de uma mesma amostra original, para construção das amostras de ruídos para os passos backward e forward.

O aumento do tamanho da amostra backward tem um impacto significativo no tempo computacional de resolução do problema de planejamento da operação. Contudo não é possível usar uma amostra aleatória de tamanho muito reduzido, pois é necessário representar de forma estável o processo estocástico do problema. Nas opções que utilizam as técnicas de agregação para geração da amostra backward, é possível conseguir resultados tão bons na solução do problema utilizando uma amostra de tamanho reduzido, quanto àqueles obtidos com amostras de tamanhos maiores. Além disso, os resultados observados são mais robustos com relação à variação da amostra de ruídos utilizada.

Os resultados obtidos neste trabalho sinalizam que a utilização de técnicas de agregação melhora a qualidade da função de custo futuro. Essa melhoria pode ser maior se forem utilizados o método de amostragem LHC, para obtenção da amostra

de ruídos original empregada no processo de agregação, e o método de reamostragem de cenários forward.

### **9.1. Sugestões para trabalhos futuros**

Os estudos apresentados mostraram que a técnica de agregação é uma ferramenta que fornece uma melhoria na robustez dos resultados do modelo NEWAVE, se consideradas variações no tamanho da amostra backward e variação de amostra. Porém, esta melhoria é menor para variações no tamanho da amostra de cenários utilizada no passo forward. Desta forma, é interessante investir em outras técnicas para geração desta amostra, como por exemplo, técnicas de redução de cenários.

Uma das propostas iniciais deste trabalho era a aplicação destas técnicas para a construção dos cenários forward. A redução seria feita de forma local, porém condicionada aos demais cenários de ENA do período em questão. Porém, o tempo computacional gasto com esse procedimento foi excessivamente grande, aproximadamente 20 horas para uma árvore com 120 períodos, o que tornou inviável o seu estudo. Foi aventada a possibilidade de utilizar técnicas de processamento paralelo para solucionar o problema de tempo computacional, porém na geração dos cenários de ENA há dependência entre os períodos, o que impossibilita a divisão das tarefas entre diversos processadores.

Como não é possível a paralelização da aplicação das técnicas de redução, uma possibilidade a ser investigada é a paralelização do algoritmo utilizado nas técnicas de redução ótima de cenários.

Em relação à degradação da representatividade do desvio-padrão nas opções que utilizam a agregação de cenários, é sugerido a incorporação de alguma informação na medida de distância que faça com que o desvio-padrão seja preservado, como por exemplo, a distância do desvio calculado com relação ao seu valor teórico.

Os experimentos realizados neste trabalho consideraram apenas casos de PMO que utilizam um horizonte de 10 anos (5 anos de estudo mais 5 anos de pós estudo) e quatro subsistemas equivalentes de energia. Desta forma, sugere-se estender os experimentos para avaliar casos do Programa Decenal de Expansão, que consideram um horizonte maior, além de representar outros subsistemas.

Um ponto que deve ser investigado futuramente é a definição de outros critérios de parada para o algoritmo de solução, de forma a substituir ou complementar o critério de convergência atualmente utilizado.

Finalmente, outra sugestão é combinar o método de reamostragem de cenários com o método de amostragem LHC e AD. Porém, deve-se buscar, primeiramente, uma solução para a degradação observada na representação da correlação cruzada dos cenários gerados por estes métodos.



## 10 Referências Bibliográficas

ALDENDERFER, M.S., BLASHFIELD, R.K., 1984, **Cluster Analysis**, Beverly Hills, Sage Publications.

ANDERBERG, M.R., 1973, **Cluster Analysis for Applications**, New York, Academic Press.

BELLONI, A., LIMA, A L. D. S., MACEIRA, M. E. P., SAGASTIZÁBAL, C., **Bundle Relaxation and Primal Recovery in Unit Commitment Problems. The Brazilian Case**, Annals of Operations Research, Vol. 120, pp. 21 – 44, 2003.

BENDERS, J. F., 1962., **Partitioning Procedures for Solving Mixed Variabels Programming Problems**, Numerische Mathematik, v. 4, pp. 238-252.

BIRGE, J. R., 1985, **Decomposition and partitioning methods for multistage stochastic linear programs**. Operational Resource, vol 33, no. 5, pp. 989-1007.

BRADLEY, J., 1968, **Distribution Free Statistical Tests**, Prentice Hall Inc.

BOUROCHE, J.M., SAPORTA, G., 1980, **Análises de Dados**, Rio de Janeiro, Zahar Editores.

BOX, G.E.P., JENKINS, G.M., 1976, **Time Series Analysis – Forecasting and Control**, San Francisco, Holden – Day.

BUSSAB, W.O., MIAZAKI, E.S., ANDRADE, D.F., 1990, **Introdução à Análise de Agrupamentos**, 9º Simpósio Nacional de Probabilidade e Estatística, São Paulo, Associação Brasileira de Estatística.

CHARBENEAU, R.J., 1978, **Comparison of the Two- and Three-Parameter Log Normal Distributions Used in Streamflow Synthesis**, Water Resources Research, v. 14, n. 1 (feb), pp. 149-150.

CHEN, Z. L.; POWELL, W. B., 1999, **Convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse**. Journal of Optimization Theory and Applications, vol 102, no. 3, pp. 497-524.

COSTA, F.S.; DAMÁZIO, J.M.; MACEIRA, M.E.P.; SOUZA, M.D.S.; GUILHON, L.G. e SILVA, S.B., 2003, **Modelagem Estocástica de Previsão de Vazões Mensais: PREVIVAZM**. 15º Simpósio Brasileiro de Recursos Hídricos, Curitiba, Associação Brasileira de Recursos Hídricos.

COSTA, J. P., PRADO, S., BINATO, S., 1999, Modelo DECOMP – **Manual de Metodologia**, In: Relatório Técnico CEPEL no. DPP/PEL – 639/99.

COSTA, J.P., 2007, **Árvore reduzida de cenários para o planejamento da operação de médio prazo**, Tese de Doutorado, COPPE-UFRJ, Rio de Janeiro, Brasil.

DONOHUE, C. J.; BIRGE, J. R., 2006, **The abridged nested decomposition method for multistage stochastic linear programs with relatively complete recourse**. *Algorithmic Oper. Res.*, vol 1, no. 1, pp. 20-30.

DUPACOVA, J., GROW-KUSKA, N., ROMISCH, W., **Scenario Reduction in Stochastic Programming: An Approach Using Probability Metrics**, Web site: <http://citeseer.nj.nec.com>, acessado em 12/09/01.

DUPACOVA, J., CONSIGLI, G., WALLACE, S.W., 2000, **Scenarios for Multistage Stochastic Programs**, *Annals of Operations Research* 100, pp 25-53.

DURAN, B.S., ODELL, P.L., 1974, **Cluster Analysis – A Survey**, Berlin, Springer-Verlag.

FARREL, J.L., 1997, **Portfolio Management Theory & Application**, 2ed. Orlando, Mc Graw Hill.

FORTUNATO, L. A. M., NETO, T. A. A., ALBUQUERQUE, J. C. R., PEREIRA, M. V. F., 1990, **Introdução ao Planejamento da Expansão e Operação de Sistemas de Produção de Energia Elétrica**, Editora da Universidade Federal Fluminense.

FROTA, A. E. F. 2003, **Avaliação de Opções Americanas Tradicionais e Complexas**, Dissertação de Mestrado, PUC-Rio, Rio de Janeiro, Brasil

GEOFFRION, A., 1972, **Generalized Benders decomposition**, *Journal of Optimization Theory and Applications*, vol. 10, n. 4, pp. 237-260.

GROWE-KUSKA, N., HEISTSCH, H. ROMISCH, W. 2003, **Scenario reduction and scenario tree construction for power management problems**. In *IEEE Bologna Power Tech Proceedings*, vol 3, 7pp

HAIR Jr., J.F., ANDERSON, R.E., TATHAN, R.L., BLACK, W.C., 1998, **Multivariate Data Analysis**, New Jersey, Prentice Hall.

HARVEY, A.C., 1981, **Time Series Models**, London, P Allan.

HARTIGAN, J.A., 1975, **Clustering Algorithms**, New York, John Wiley & Sons.

HARTIGAN, J., WONG, M, 1979, **A K-Means Clustering Algorithm**, *Applied Statistics*, vol.28, no. 1, pp. 100-108.

HEITSH, H., ROMISH, W., 2001, **Scenario Reduction Algorithms in Stochastic Programming** In: *Proceedings of Humboldt – University Berlin, Institute of Mathematics*, 10099 Berlin, Germany.

HINDSBERGER, M.; PHILPOTT, A. B., 2001, **Resa: A method for solving multistage stochastic linear programs**. In: *Stochastic Programming '01*, Berlin, Germany

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

JARDIM, D.L.D., 2002, **Modelo de geração de séries sintéticas de vazões utilizando técnicas de agregação**, Dissertação de Mestrado, COPPE-UFRJ, Rio de Janeiro, Brasil

JIRUTITIJAROEN, P., SINGH, C., 2008, **Comparison of Simulation Methods for Power System Reliability Indexes and Their Distributions**, IEEE Transactions on Power Systems, vol 23, no. 2, pp. 486-493.

JOHNSON, R.A., WICHERN, D.W., 1998, **Applied Multivariate Statistical Analysis**, 4 ed. New Jersey, Prentice Hall.

KELMAN, J., PEREIRA, M.V.F., 1977, **Crítérios de Avaliação para Modelos de Séries Hidrológicas**. In: Proceedings of IV Seminário Nacional de Produção e Transmissão de Energia Elétrica, Rio de Janeiro, Set.

KELMAN, J., OLIVEIRA, G.C., MACEIRA, M.E.P., 1992, **Modelo de Séries Hidrológicas – Manual de Metodologia**. In: Relatório Técnico nº 242-92, CEPEL, Rio de Janeiro.

LATORRE, J.M., CERISOLA, S., RAMOS, A., 2007, **Clustering Algorithms for Scenario Tree Generation: Application to Natural Hydro Inflow**, European Journal of Operational Research 181 pp. 1339-1353.

LIMA, A.L.D.S., 2007, **Uma Estratégia de Decomposição por Relaxação Lagrangeana para a Otimização da Programação Diária da Operação de Sistemas Hidrotérmicos com Modelagem Detalhada da Rede Elétrica – Aplicação ao Sistema Brasileiro**, Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro, Brasil.

MARCATO, A.L.M., 2002, **Representação híbrida de sistemas equivalentes e individualizados para o planejamento da operação a médio prazo de sistemas de potencia de grande porte**, Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro. Brasil.

MACEIRA, M.E.P., 1989, **Operação Ótima de Reservatórios com Previsão de Afluências**. Dissertação de Mestrado, COPPE/UFRJ, Rio de Janeiro, Brasil.

MACEIRA, M.E.P., 1993, **Programação Dinâmica Dual Estocástica Aplicada ao Planejamento da Operação Energética de Sistemas Hidrotérmicos com Representação do Processo Estocástico de Afluências por Modelos Autorregressivos Periódicos**, Relatório Técnico CEPEL 237-1993.

MACEIRA, M.E.P., CUNHA, S.H.F., 1995, **Simulating the Energy Generation of Interconnect Hydro-Thermal Systems – SUISHI Model**. In: Proceedings of XIII Seminário Nacional de Produção e Transmissão de Energia Elétrica, Balneário de Camboriú, Brasil, Oct.

MACEIRA, M.E.P., BEZERRA, C.V., 1997, **Stochastic Streamflow model for Hydroelectric Systems** In: Proceedings of 5th International Conference on

Probabilistic Methods Applied to Power Systems, pp. 305-310, Vancouver, Canada, Sep.

MACEIRA, M.E.P., BEZERRA, C.V., GORENSTIN, B.G., CUNHA, S.H.F., SUANNO, C., SACRAMENTO, M.C., KLINGERMAN, A.S., 1998, **Aplicação do Modelo na Avaliação Energética da Interligação dos Sistemas Norte/Nordeste e Sul/Sudeste Brasileiros** In: Proceedings of VI Simpósio de Especialistas em Planejamento da Operação e Expansão Elétrica, Salvador, May.

MACEIRA, M.E.P., COSTA, A.P.C., MARCATO, A.L.M., 1999, **Newave Especificação Funcional**, In: Relatório Técnico DPP/PEL 349/99, CEPEL, Rio de Janeiro.

MACEIRA, M.E.P., JARDIM, D.L.D.D., 2000, **Modelo de Geração de Séries Sintéticas de Energias e Vazões – GEVAZP – Manual de Referência**, In: Relatório Técnico DPP/PEL 83/2000, CEPEL, Rio de Janeiro.

MACEIRA, M. E. P., TERRY, L. A., COSTA, F. S., DAMÁZIO, J. M., MELO, A. C. G., 2002, **Chain of Optimization Models for Setting the Energy Dispatch and Spot Price In the Brazilian System**, In: Proceedings of 14<sup>th</sup> PSCC, Sevilla, Espanha.

MACEIRA, M. E. P.; DUARTE, V. S.; PENNA, D.D.J.; MORAES, L.A. M.; MELO, A.C.G., **Ten years of application of stochastic dual dynamic programming in official and agent studies in brazil description of the NEWAVE**. In: Power Systems Computation Conference, 2008, Glasgow.

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

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

MOROKOFF, W.J., CAFLISCH, R.E., 1995, **Quase-Monte Carlo Integration**, Journal of Computation Physics, 122: 218-230.

NOONAN, F., GIGLIO, R.J., 1977, **Planning electrical power generation: a non-linear mixed integer model employing Bender's decomposition**, Management Science, vol. 23, no. 9, pp. 946-956.

OLIVEIRA G, 1992, **Otimização da operação energética de sistemas hidroelétricos com representação individualizada das usinas e aflúências deterministicas**. Tese de Mestrado, FEE, Unicamp, FAPESP

OLIVEIRA, W.L., 2007, **Redução ótima de cenários em programação estocástica. Aplicação às vazões afluentes aos aproveitamentos hidroelétricos**, Dissertação de Mestrado, IMPA, Rio de Janeiro, Brasil.

OLIVEIRA.W.L., SAGASTIZABAL, C., PENNA, D.D.J., MACEIRA. M.E.P., DAMAZIO, J.M., 2008, **Optimal Scenario Tree Reduction for Stochastic Streamflows in Power Generation Planning Problems**, In Optimization Methods and Software – Special issue – Engopt 2008.

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.

PENNA, D.D.J., 2009, **Definição da árvore de cenários de aflúncias para o planejamento da operação energética de médio prazo**, Tese de Doutorado, PUC-Rio, Rio de Janeiro, Brasil.

PEREIRA, M. V. F., 1989 **Optimal stochastic operations scheduling of large hydroelectric systems**. Electrical Power & Energy Systems, vol 11, no. 3, pp. 161-169.

PEREIRA, M.V.F., PINTO, L.M.V.G., 1991, **Multi-Stage Stochastic Optimization Applied to Energy Planning**. Mathematical Programming, v. 52, n. 3 (may), pp. 359-375.

PHILPOTT, A. B.; GUAN, Z., 2007, **On the convergence of sampling-based methods for multi-stage stochastic linear programs**. Electric Power Optimization Centre, EPOC publications.

RIPPL W., 1883, **Capacity of Storage Reservoirs for Water Supply**. In: Proceedings of Institution of Civil Engineers v. 71.

ROSENTHAL R. E, 1981, **A nonlinear network flow algorithm for maximization of benefits in a hydroelectric power system**, Operations Research, v29, n4, pp763-786.

ROTTING, T.A., GJELSVIK, A., 1992, **Stochastic dual dynamic programming for seasonal scheduling in the Norwegian power system**, IEEE Transactions on Power Systems, vol. 7, no. 1, pp. 273-279.

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

SALIBY,E., PACHECO, F., 2002, **An Empirical Evaluation of Sampling Methods in Risk Analysis Simulation: Quasi-Monte Carlo, Descriptive Sampling, and Latin Hypercube Sampling**, In: Proceedings of the 2002 Winter Simulation Conference, pp. 1606-1610.

SALLABERRY, C.J, HELTON, J.C., HORA, S.C, 2008, **Extensión of Latin Hypercube Samples with Correlated Variables**. Reliability Engineering & System Safety, vol 93, no.7, pp. 1047-1059.

SILVA, E.L., 2001, **Formação de Preços em Mercados de Energia Elétrica**, Editora Sagra Luzzatto.

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

STEINBACH, M., KARYPIS, G., KUMAR, V., 2000, **A Comparison of Document Clustering Techniques**. In: Technical Report # 00-034, Department of Computer Science and Engineering, University of Minnesota, Minnesota.

STEPHENS, M.A. 1974, **EDF Statistics for Goodness of Fit and Some Comparisons**, Journal of the American Statistical Association, v.69, n.347 (Sep), pp. 730-737.

SUANNO, C.M., MACEIRA, M.E.P., COSTA, J.P. da, 1995, **Representação da Variação da Produtibilidade no Algoritmo de Programação Dinâmica Dual Estocástica**, In: Relatório Técnico CEPEL no. DPP/PEN – 153/95.

SUANNO, C.M., 1995, **Operação Estocástica de Sistemas Hidrotérmicos com Representação da Produtividade Variável**, Dissertação de Mestrado, COPPE/UFRJ, Rio de Janeiro, Brasil.

TERRY, L. A., 1980, **Modelo a Sistema Equivalente - Descrição Geral**, In: Relatório Técnico CEPEL no. DPP/PEN – 1705/80.

TONG, C., 2006, **Refinement Strategies for Stratified Sampling Methods**, Reliability Engineering & System Safety 91: 1257-1265

VALENTIN, J.L., 2000, **Ecologia Numérica, Uma Introdução à Análise Multivariada de Dados Ecológicos**, Rio de Janeiro, Ed. Interciência.

VAN SLYKE, R., WETS, R. J. B., 1969, **L-shaped linear programs with applications to optimal control and stochastic linear programs**. SIAM Journal on Applied Mathematics, 17, pp.638-663.

VELASQUEZ, R.M.G., PESSANHA, J.F.M., JARDIM, D.L.D.D., MELO, S.L., MELO, A.C.G., 2001, **Técnicas de Classificação para Caracterização da Curva de Carga de Empresas de Distribuição de Energia Elétrica – Um Estudo Comparativo** In: Proceedings of the V Congresso Brasileiro de Redes Neurais, pp. 133-138, Rio de Janeiro, Apr.

XAVIER, N.R., LIMA, L., DINIZ, A.L. ; COSTA, F.S., MACEIRA, M. E. P. . **Aprimoramento da Modelagem da Função de Produção Energética das Usinas Hidroelétricas do Modelo Decomp: Metodologia e Resultados**. In: XVIII Seminário Nacional de Produção e Transmissão de Energia Elétrica, 2005, Curitiba. Anais do XVIII Seminário Nacional de Produção e Transmissão de Energia Elétrica. Rio de Janeiro : CIGRÉ, 2005.

YEH, W. W-G., BECKER, I., HUA, S-Q. et al, 1992, **Optimization of real-time hydrothermal system operation**, Journal of Water Resources Palnning ana Management, v. 118, n.6, pp 636-653

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

YU, H., CHUNG, C.Y, WONG, K.P., LEE, H.W., ZHANG, J.H., 2009, **Probabilistic Load Flow Evaluation With Hybrid Latin Hypercube Sampling and Cholesky Decomposition**, IEEE Transations on Power Systems, vol 24, no. 2, pp. 661-667.

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

ZIKMUND, W.G., 1999, **Exploring Marketing Research**, 6 ed. Singapore, Dryden Press.

## Apêndice A Modelo Newave

### A.1. Introdução

No planejamento da operação de médio e longo prazo, o horizonte de 5 a 10 anos com desagregação mensal permite estudar o efeito da capacidade de regularização plurianual dos reservatórios do sistema brasileiro, assim como o impacto de períodos hidrológicos desfavoráveis de longa duração, como o observado no registro histórico de novembro de 1951 a julho de 1956. Adicionalmente, é importante considerar todas as possíveis evoluções das aflúncias às usinas hidrelétricas ao longo deste horizonte, bem como, a ocorrência conjunta e a sua distribuição entre as diferentes bacias hidrográficas que compõem o sistema interligado brasileiro.

A política de operação para sistemas hidrotérmicos com predominância da geração hidroelétrica vem sendo calculada através de modelos baseados em programação dinâmica estocástica (PDE) e programação dinâmica dual estocástica (PDDE). Em geral, para qualquer modelo baseado em programação dinâmica, as variáveis que irão representar o estado do sistema devem afetar o custo futuro de operação. Para representar sistemas hidrotérmicos são necessários, ao menos, as seguintes variáveis: nível de armazenamento dos reservatórios e a tendência hidrológica do sistema, como por exemplo, as vazões afluentes às usinas hidrelétricas no período anterior. Isto significa que a estratégia de operação do sistema deve ser calculada para todas as possíveis combinações entre níveis de armazenamento dos reservatórios e tendência hidrológica. Por exemplo, suponha que o volume armazenado em cada reservatório seja representado por 20 valores discretos e a tendência hidrológica seja representada por apenas uma variável, discretizada também em 20 valores. Se o sistema possui  $n=4$  reservatórios, existem  $20^{2n}=20^8$  combinações possíveis entre os estados de armazenamento inicial dos reservatórios e tendência hidrológica.

Uma solução para reduzir este crescimento exponencial do número de combinações é usar uma representação agregada do conjunto de reservatórios que compõe o sistema. Essa representação é dada por um único reservatório e é



conhecida como representação por reservatório equivalente. A energia que pode ser produzida com o volume total de água armazenada no sistema depende da forma como está distribuída esta água. Sendo assim, convém representar o estado de armazenamento agregado do sistema, ou seja, o armazenamento do reservatório equivalente, através da energia que pode ser produzida pelo deplecionamento total do sistema de reservatórios, dado o armazenamento de água inicial.

Empregando-se um procedimento similar pode-se obter a energia correspondente às aflúncias às usinas hidrelétricas. Para tanto, obtém-se a representação do sistema hidrelétrico - dado por um reservatório equivalente de energia - que a cada intervalo de tempo sofre um deplecionamento correspondente à energia total gerada pelo sistema hidroelétrico e um reenchimento, dado pela energia correspondente às aflúncias no mesmo intervalo de tempo.

Atualmente no âmbito do planejamento da operação, o SIN é subdividido em quatro regiões denominadas: Sudeste, Sul, Nordeste e Norte. Além dos vínculos hidráulicos, os subsistemas são interligados eletricamente através de grandes troncos de interligação, constituindo desta forma um sistema interligado.

O modelo utilizado no planejamento da operação de médio prazo do sistema elétrico brasileiro é o modelo NEWAVE que será descrito a seguir.

## **A.2. Modelo Newave**

No modelo NEWAVE, o problema de planejamento da operação energética de médio prazo é representado por um problema de programação linear multiestágio, cujo objetivo é a minimização do custo total de operação (custo de combustível das unidades térmicas mais a penalidade de atendimento à demanda) ao longo do horizonte de planejamento. O modelo é baseado na técnica de PDDE e considerando uma representação agregada do parque hidrelétrico, o modelo NEWAVE pode considerar vários subsistemas interligados, permite a representação estática ou dinâmica da configuração do sistema, discretização da carga própria em até três patamares (por exemplo, pesada, média e leve), representação dos cortes no suprimento do mercado de energia elétrica em até quatro patamares de déficit, além da consideração de diversos cenários de energias afluentes, obtidos através de um modelo autorregressivo periódico de ordem  $p$ , PAR( $p$ ) (Maceira & Bezerra, 1997), que modela a aflúncia de um mês como sendo função das aflúncias dos  $p$  meses

anteriores. Através do uso do PAR(p), secas de longa persistência são capazes de serem reproduzidas.

Basicamente, o NEWAVE é composto pelos seguintes módulos:

- Módulo de cálculo dos sistemas equivalentes – Calcula os subsistemas equivalentes de energia: energias armazenáveis máximas, séries históricas de energias controláveis e energias fio d'água, parábolas de energia de vazão mínima, energia evaporada, capacidade de turbinamento, correção da energia controlável em função do armazenamento e perdas por limite de turbinamento nas usinas fio d'água.
- Módulo de cálculo do modelo estocástico de energias afluentes – Estima os parâmetros do modelo estocástico e gera séries sintéticas de energias afluentes que são utilizadas no módulo de cálculo da política de operação hidrotérmica e para geração de séries sintéticas de energias afluentes para análise de desempenho no módulo de simulação da operação.
- Módulo de cálculo da política de operação hidrotérmica – Determina 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 – Simula a operação do sistema ao longo do período de planejamento, para distintos cenários de seqüências hidrológicas, falhas dos componentes e variações da demanda. Calcula 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.

O relacionamento entre os módulos utilizados é apresentado na Figura 176.

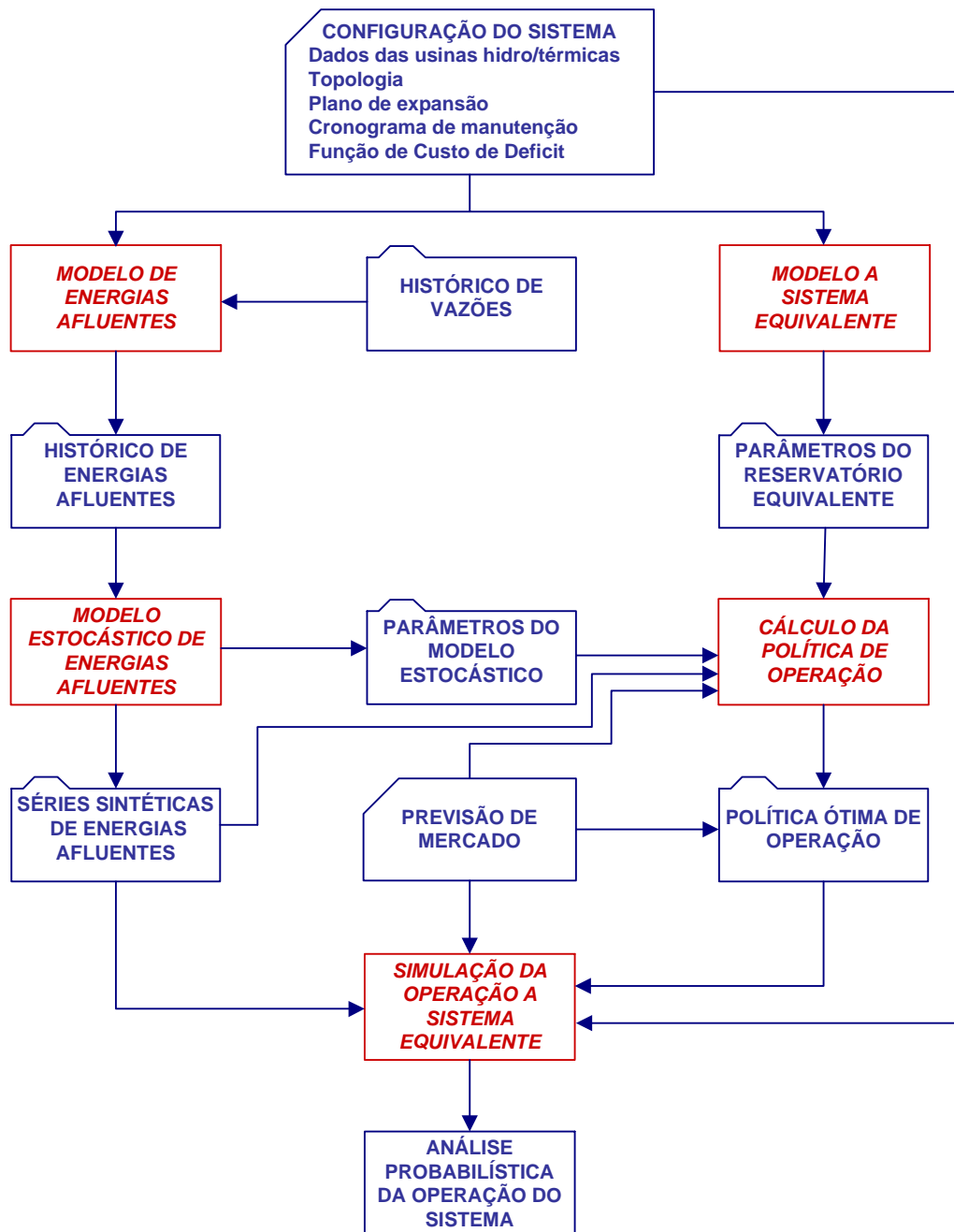


Figura 176: Esquema de relacionamento entre os módulos do programa NEWAVE

### A.2.1. Módulo de Cálculo do Sistema Equivalente

Mesmo com a utilização da PDDE, a estratégia ótima de operação deve ser calculada para diversos cenários hidrológicos. A dimensão do espaço de estados é dado pelo número de reservatórios somado à ordem máxima do modelo PAR(p). No caso do sistema elétrico brasileiro, com mais de 50 reservatórios, faz-se necessário

reduzir o número de variáveis de estado através da agregação dos diversos reservatórios de uma mesma região em um reservatório equivalente de energia, cujos principais parâmetros são descritos a seguir (Terry, 1980):

#### **a) Energia Armazenável Máxima**

A energia armazenável máxima,  $\overline{EARMf}$ , representa a capacidade de armazenamento do conjunto de reservatórios do sistema e é estimada pela energia produzida pelo esvaziamento completo dos reservatórios do sistema de acordo com uma política de operação estabelecida. Adota-se a hipótese de operação em paralelo, isto é, os armazenamentos e deplecionamentos são feitos paralelamente em volume. Sendo assim, a energia armazenada entre dois estados de armazenamento do sistema é definida como sendo a energia gerada ao se deplecionar paralelamente os reservatórios entre os estados inicial e final, sem considerar novas afluições.

#### **b) Energia Controlável**

A energia controlável, EC, pode ser obtida a cada mês, a partir da soma da vazão natural a cada reservatório multiplicada pela sua produtividade equivalente somada às produtibilidades das usinas a fio d'água a jusante até o próximo reservatório exclusivo ou, alternativamente, pode também ser obtida pela soma das vazões incrementais a cada reservatório valorizadas pela produtividade equivalente em todas as usinas a jusante do mesmo.

#### **c) Energia a Fio D'Água**

A energia a fio d'água, EFIO, corresponde às vazões incrementais às usinas a fio d'água e conseqüentemente não são passíveis de armazenamento. A determinação da vazão incremental é feita em cada usina a partir da afluição natural, da qual são descontadas as vazões naturais às usinas de reservatório imediatamente a montante. O máximo de vazão incremental que pode ser transformado pela usina em energia é limitado pelo engolimento máximo das turbinas.

#### **d) Separação da Energia Controlável da Energia Natural Afluente**

As séries sintéticas geradas correspondem ao valor total de energia afluente, isto é, energia controlável mais energia fio d'água bruta, EFIOB. Para obter-se a energia controlável a partir da energia afluente total é necessário ter-se calculado previamente,

a partir do histórico, a participação média da energia controlável na energia afluyente total.

#### **e) Correção da Energia Controlável**

A energia controlável é calculada supondo que os reservatórios do sistema em estudo estejam com seus níveis de armazenamento correspondentes às alturas equivalentes,  $H_{eq}$ . Tendo em vista que durante a simulação da operação os níveis de armazenamento se modificam, o valor da energia controlável inicialmente calculado deve ser corrigido em função dessas variações, uma vez que estas implicam em mudanças no valor do coeficiente de produção utilizado (Suanno et al, 1995), e conseqüentemente, no valor da disponibilidade total de energia controlável ao sistema. Essas variações de altura de queda podem ser traduzidas através de um fator de correção, FC (Suanno, 1995).

#### **f) Energia de Vazão Mínima**

A energia de vazão mínima, EVMIN, independe da série hidrológica considerada, dependendo tão somente da configuração do sistema. Seu valor máximo é calculado multiplicando-se a descarga mínima obrigatória de cada usina com reservatório pela soma da produtibilidade associada à altura equivalente, e as de todas as usinas fio d'água existentes entre o reservatório e o próximo reservatório a jusante.

#### **g) Energia Evaporada**

O valor máximo da energia evaporada, EVAP, é calculado multiplicando-se o coeficiente de evaporação de cada reservatório pela área correspondente à altura equivalente e pelo produto da produtibilidade, associado também à altura equivalente, de todas as usinas existentes (com reservatório e fio d'água) entre o reservatório e a última usina da cascata.

#### **h) Geração Hidráulica Máxima**

A capacidade de geração hidráulica deve levar 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. A geração hidráulica máxima,  $\overline{GH}$ , independe da série

hidrológica considerada, dependendo tão somente da configuração e do nível de armazenamento.

Assim, para cada mês do horizonte de estudo, 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 de segundo grau, a partir da qual se obtém a geração hidráulica máxima em função da energia armazenada no mês.

### **A.2.2. Módulo de Cálculo da Política de Operação**

O planejamento da operação visa calcular estratégias de operação para as usinas hidrelétricas e termelétricas de forma a minimizar o custo de operação no horizonte de planejamento.

Assim, o problema de operação ótima de um sistema hidrotérmico consiste em determinar uma estratégia de operação que, a cada estágio do período de planejamento, conhecido o estado do sistema no início do estágio, forneça as metas de geração de cada unidade geradora e intercâmbio de energia entre elas. Desta forma, para cada estágio, deseja-se minimizar o custo presente mais o custo futuro de operação. O primeiro é representado pelos custos dos combustíveis das usinas termelétricas mais o custo associado à interrupção do fornecimento de energia. O segundo representa o impacto futuro de uma decisão presente, ou seja, o custo de operação desde o estágio seguinte até o final do horizonte de estudo.

O estado do sistema é composto por variáveis que podem influir no resultado da operação. No caso do planejamento energético, o nível de armazenamento no início do estágio e as energias afluentes aos estágios anteriores são as variáveis componentes do estado.

O problema de operação hidrotérmica pode ser resolvido por Programação Dinâmica Estocástica, PDE. Contudo, para aplicações envolvendo múltiplos reservatórios, tem-se recorrido à Programação Dinâmica Dual Estocástica, PDDE, apresentada na próxima seção.

### A.2.2.1.

#### Programação Dinâmica Dual Estocástica Aplicada ao Planejamento da Operação Hidrotérmica

Com o objetivo principal de evitar a explosão combinatória decorrente de um algoritmo baseado em Programação Dinâmica Estocástica, a PDDE se apresenta como uma alternativa viável para resolver o problema, com um esforço computacional moderado. Na PDDE, a função de custo futuro é construída analiticamente através do 'Princípio de Decomposição de Benders' (Benders, 1962).

Para demonstrar o processo de construção da função de custo futuro, inicialmente apresenta-se uma versão determinística da PDDE, aqui denominada de Programação Dinâmica Dual Determinística (PDDD). Em seguida é feita a extensão para o caso estocástico (Pereira & Pinto, 1991 e Silva, 2001).

### A.2.2.1.1.

#### Programação Dinâmica Dual Determinística

Considere o seguinte problema de programação linear de dois estágios:

$$\text{Min } (f) = c_1x_1 + c_2x_2 \quad (\text{A.1})$$

sujeito a:

$$\begin{aligned} A_1x_1 &\geq b_1 \\ E_1x_1 + A_2x_2 &\geq b_2 \end{aligned} \quad (\text{A.2})$$

O problema apresentado em (A.1) e (A.2) pode ser interpretado como um processo de tomada de decisão seqüencial de dois estágios:

- 1º estágio: escolhe-se uma decisão viável,  $x_1^*$ , tal que  $A_1x_1^* \geq b_1$ ;
- 2º estágio: com  $x_1^*$  fixado, resolve-se o problema de otimização do 2o estágio, dado por (A.3) e (A.4).]

$$\text{Min } c_2x_2 \quad (\text{A.3})$$

sujeito a:

$$A_2x_2 \geq b_2 - E_1x_1^* \quad (\text{A.4})$$

Como  $x_1^*$  é conhecido, passa para o lado direito do conjunto de restrições do problema do 2o estágio.

Os vetores  $x_1^*$  e  $x_2^*$ , do primeiro e do segundo estágio, respectivamente, são compostos pelas variáveis hidráulicas e térmicas: volumes armazenados ao final do período, volumes turbinados e vertidos, e níveis de geração térmica. Os custos do primeiro estágio são representados por  $c_1x_1$  e as restrições de operação do sistema (balanço hídrico, limites superior e inferior de volumes) são representadas por  $A_1x_1 \geq b_1$ . Analogamente,  $c_2x_2$  representa o custo de operação do segundo estágio e  $A_2x_2 \geq b_2 - E_1x_1^*$ , as respectivas restrições operativas. O objetivo do problema consiste em minimizar o custo total de operação composto por  $c_1x_1 + c_2x_2$ .

Desta forma, o problema de decisões seqüenciais (A.1) e (A.2) pode ser resolvido por Programação Dinâmica. As expressões recursivas ficam:

2º estágio:

$$\begin{aligned} \alpha_2(x_1) &= \min c_2x_2 \\ \text{s.a.} \\ A_2x_2 &\geq b_2 - E_1x_1 \end{aligned} \quad (\text{A.5})$$

1º estágio:

$$\begin{aligned} \alpha_1(x_0) &= \min c_1x_1 + \alpha_2(x_1) \\ \text{s.a.} \\ A_1x_1 &\geq b_1 \end{aligned} \quad (\text{A.6})$$

O valor de  $x_0$  pode ser interpretado como um estado inicial do sistema. A função  $\alpha_2(x_1)$  fornece informações sobre as conseqüências nos estágios futuros da decisão  $x_1$ . Caso esta função esteja disponível, o problema dado por (A.1) pode ser resolvido sem a representação explícita das restrições do 2º estágio.

De um modo geral, entretanto, a função de custo futuro, generalizada para qualquer estágio  $t$ ,  $\alpha_{t+1}(x_t)$ , não é conhecida. Na recursão usual da Programação Dinâmica esta função é calculada para valores discretos de  $x_t$ . A Figura 177 ilustra o cálculo para  $\alpha_2(x_1)$ . O eixo horizontal  $x_1$ , é discretizado em  $N$  pontos  $(x_1^j, j = 1, \dots, N)$  e  $\alpha_2(x_1)$ , sendo caracterizada pelo conjunto  $\{\alpha_2(x_1^j), j = 1, \dots, N\}$ .



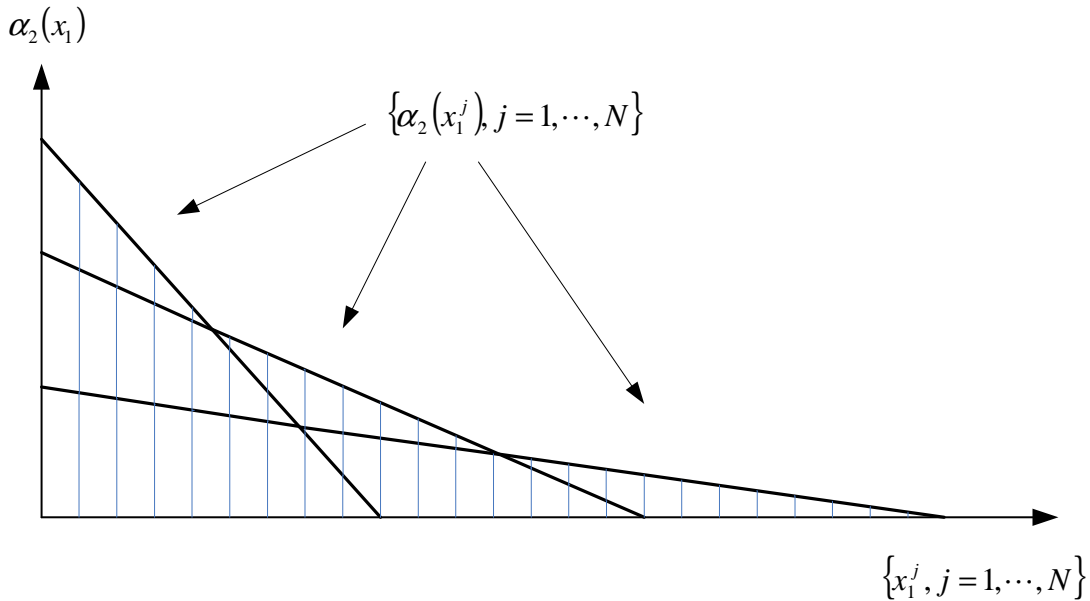


Figura 177: Cálculo da Função de Custo Futuro através da Programação Dinâmica

O princípio de ‘Decomposição de Benders’ é uma técnica que permite construir, iterativamente, aproximações para a função  $\alpha_2(x_1)$  do problema dado pela por (A.5), baseada na solução do problema do primeiro estágio (A.6).

Considerando que há um problema dual associado a qualquer problema de Programação Linear, tem-se que o dual do problema (A.5) pode ser representado por:

$$\begin{aligned}
 \alpha_2(x_1) &= \text{Max } \pi_2(b_2 - E_1x_1) \\
 \text{s.a.} & \\
 A_2\pi_2 &\leq c_2
 \end{aligned}
 \tag{A.7}$$

onde  $\pi_2$  é o vetor com as variáveis duais associadas ao problema do 2o estágio.

Assim, através da análise dual do problema do segundo estágio, dada pelo conjunto de equações (A.7), a restrição transferida ao problema do 1o estágio pode ser escrita da forma:

$$\alpha_2(x_1) + \pi_2^* E_1 x_1 \geq \omega_2^* + \pi_2^* E_1 x_1^*
 \tag{A.8}$$

A vantagem deste processo é que não há necessidade da discretização do espaço de estados. A cada iteração, uma nova aproximação da função de custo futuro é gerada em torno do ponto obtido a partir da solução do problema do 1º estágio,  $x_1^*$ .

Quanto a escolha dos pontos  $x_1^*$ , tem-se um processo baseado em iterações que consistem na seleção de uma série de recursões diretas, denominada *forward*, e inversas, *backward*, para todo o período de estudo.

A Programação Dinâmica Dual é facilmente aplicada à problemas multi-estágios. Considerando que um problema possua  $T$  estágios, e sendo  $k$  o contador de iterações, inicialmente resolve-se uma seqüência de problemas, percorrendo desde o estágio 1 até o estágio  $T-1$ . De cada problema, obtém-se o valor ótimo das variáveis,  $x_t^{*k}$ , e o custo imediato associado a cada estágio,  $c_t x_t^{*k}$ . Esta etapa do algoritmo é denominada de processo *forward*.

Ao chegar ao último estágio, começa-se a recursão inversa, denominada processo *backward*, do estágio  $T$  até o 2º estágio. Para cada estágio  $t$ , obtém-se os valores de  $\omega_t^{*k}$  e  $\pi_t^{*k}$ , que em conjunção com o vetor  $x_{t-1}^{*k}$ , calculados no processo *forward*, monta-se a seguinte restrição, generalizada para qualquer estágio  $t$ :

$$\alpha_t(x_{t-1}) + \pi_t^{*k} E_{t-1} x_{t-1} \geq \omega_t^{*k} + \pi_t^{*k} E_{t-1} x_{t-1}^{*k} \quad (\text{A.9})$$

Essa restrição (Cortes de Benders) é então passada para o estágio anterior.

O processo iterativo termina quando, a cada estágio  $t$ , o custo previsto no estágio  $t-1$  para o estágio  $t$ , iguala-se ao custo efetivo do estágio  $t$ ,  $\omega_t^{*k}$ . Com isso, tem-se que o custo total do 1º estágio,  $\omega_1^*$ , iguala-se ao valor da soma composta por  $c_1 x_1^* + c_2 x_2^* + \dots + c_T x_T^*$ . Ou seja, a soma dos custos efetivos de todos os estágios em uma iteração define o limite superior do problema, sendo que o limite inferior é obtido no 1º estágio,  $\omega_1^*$ . Desta forma, o problema converge quando a diferença entre o limite inferior e o superior é menor do que uma tolerância especificada.

### A.2.2.1.2 Programação Dinâmica Dual Estocástica

A extensão do algoritmo de Programação Dinâmica Dual para problemas estocásticos com multi-estágios segue o mesmo princípio descrito anteriormente, onde  $\alpha_{t+1}(x_t)$  é agora expresso como o valor esperado das derivadas dos custos futuros da decisão  $x_t$  tomada sobre todos os  $m$  possíveis cenários dos estágios subseqüentes.

Assim, a nova expressão para o Corte de Benders relativo ao caso estocástico, em um estágio  $t$  qualquer, fica da seguinte forma:

$$\overline{\alpha}_t(x_{t-1}) + \overline{\pi}_t^* E_{t-1} x_{t-1} \geq \overline{\omega}_t^* + \overline{\pi}_t^* E_{t-1} x_{t-1}^* \quad , \quad (\text{A.10})$$

onde

$\overline{\alpha}_t(x_{t-1})$ : é o valor esperado das aproximações da função de custo futuro relacionadas com os  $m$  cenários;

$\overline{\omega}_t^*$ : é o valor esperado dos custos de operação para todos os  $m$  cenários utilizados;

$\overline{\pi}_t^*$ : é o valor esperado dos multiplicadores simplex pra todos os  $m$  cenários utilizados.

A cada iteração é possível obter um limite superior para o valor esperado do custo futuro dado por (A.11):

$$\overline{z} = E \left\{ \sum_{i=1}^T c_i x_i \right\} \quad , \quad (\text{A.11})$$

e um limite inferior, dado pelo custo esperado total aproximado do 1º estágio  $\alpha_1(x_0)$ .

### A.2.2.2 Despacho de Operação Hidrotérmica em Sistemas Equivalentes de Energia

O problema de minimização de um estágio que descreve o cálculo da política ótima de operação hidrotérmica em sistemas equivalentes de energia é dado por :

- Função Objetivo:

$$z_t = \min \left[ \sum_{NSIS} \sum_{NPMC} \left[ \sum_{TCLISIS} \psi_T g_T + \sum_{NPDF} \psi_D def \right] + \frac{1}{1+\beta} \alpha_{t+1} \right] \quad (\text{A.12})$$

sujeito a:

- Equações de Balanço Hídrico – EBH (uma equação para cada subsistema):

$$earmf_t + \sum^{NPMC} g_H + evert = EARMf_{t-1} + FC(EARMf_{t-1}) \cdot \gamma \cdot EAF_t -$$

$$- M_{EVMIN}(EARMf_{t-1}) - EVAP(EARMf_{t-1}) - EMORTO \quad (A.13)$$

- Equações de Atendimento à Carga Própria – EAD (uma equação para cada subsistema e para cada patamar):

$$g_H + \sum^{TCLISIS} g_T + \sum^{NPDF} def + \sum i_{recebido} - \sum i_{exportado} - exc =$$

$$[DEMLIQ - (1 - \gamma)EAF_t - M_{EVMIN}(EARMf_{t-1})]FPENG \quad (A.14)$$

- Restrições que representam a Função de Custo Futuro – ECOR (tantos quantos forem os Cortes de Benders (BENDERS, 1962):

$$\alpha_{t+1} \geq z_t + \sum_{isis=1}^{NSIS} \left( \begin{array}{l} \pi_V^{isis} (earmf_t^{isis} - EARMf_t^{* isis}) + \\ + \pi_{A1}^{isis} (EAF_{t-1}^{isis} - EAF_{t-1}^{* isis}) + \pi_{A2}^{isis} (EAF_{t-2}^{isis} - EAF_{t-2}^{* isis}) + \\ + \dots + \\ + \pi_{Ap}^{isis} (EAF_{t-p}^{isis} - EAF_{t-p}^{* isis}) \end{array} \right) \quad (A.15)$$

onde

(\*): estado da série em questão no período t.

Seja

$$W = z_t - \sum_{isis=1}^{NSIS} (\pi_V^{isis} EARMf_t^{* isis} + \pi_{A1}^{isis} EAF_{t-1}^{* isis} + \pi_{A2}^{isis} EAF_{t-2}^{* isis} + \dots + \pi_{Ap}^{isis} EAF_{t-p}^{* isis}) \quad (A.16)$$

então a equação do corte construído no período t pode ser escrita da forma:

$$\alpha_{t+1} \geq W + \sum_{isis=1}^{NSIS} (\pi_V^{isis} earmf_t^{c isis} + \pi_{A1}^{isis} EAF_{t-1}^{c isis} + \pi_{A2}^{isis} EAF_{t-2}^{c isis} + \dots + \pi_{Ap}^{isis} EAF_{t-p}^{c isis}) \quad (A.17)$$

- Restrições de Geração Hidráulica Máxima Controlável – EEH (para cada subsistema e para cada patamar):

$$g_H \leq [\overline{GH}(EARMf_{t-1}) - (1 - \gamma)EAF_t - M_{EVMIN}(EARMf_{t-1})]FPENG \quad (A.18)$$

- Equações de Nó – EFIC (para cada nó fictício, para cada patamar de carga):

$$\sum_{\forall j \neq k} i_{j,k} - \sum_{\forall j \neq k} i_{k,j} = 0 \quad (A.19)$$

- Limites:

$$0 \leq i_{i,k} \leq \overline{i_{i,k}} \quad (\text{A.20})$$

$$0 \leq g_T \leq \overline{g_T} \quad (\text{A.21})$$

$$\underline{earmf_t} \leq earmf_t \leq \overline{earmf_t} \quad (\text{A.22})$$

Na equação (A.19), chama-se de nó fictício aquele nó que não tem carga, armazenamento ou geração.

Na função objetivo (A.12)  $\alpha_{t+1}$  é o valor esperado do custo de operação do estágio  $(t+1)$  até o final do horizonte de estudo.

A energia afluyente,  $EAF_t$ , presente em (A.13) e (A.14) é composta da energia controlável, afluyente às usinas hidrelétricas com reservatório, e da energia a fio d'água, afluyente às usinas a fio d'água.

Na equação (A.13), o fator de separação,  $\gamma$ , é aplicado à energia afluyente para a obtenção da energia controlável. De acordo com o modelo PAR( $p$ ), a energia afluyente de um estágio  $t$  é função de  $p$  energias afluyentes passadas, onde  $p$  é a ordem do modelo, mais uma parte aleatória:

$$EAF_t = \phi_1^t EAF_{t-1} + \phi_2^t EAF_{t-2} + \dots + \phi_p^t EAF_{t-p} + a_t \quad (\text{A.23})$$

Na parcela da energia afluyente correspondente à energia controlável está contabilizada a energia de vazão mínima, logo em (A.13) essa parcela é abatida na equação do balanço hídrico.

Como a energia de vazão mínima é uma geração hidráulica da qual não se tem controle, ela é abatida diretamente da demanda (A.14), assim como a geração a fio d'água.

O conjunto de equações (A.17) representa a função de custo futuro, que no caso da programação dinâmica dual estocástica, PDDE, é representada por uma função linear por partes e construída iterativamente pelos cortes de Benders.

Em cada estágio  $t$  e para cada estado, os coeficientes do corte de Benders associados à energia armazenada no início do estágio,  $\pi_v$ , e às energias afluyentes passadas,  $\pi_{A_i}, i = 1, \dots, p$ , são obtidos por (A.24) e (A.25), durante as recursões *backward* do algoritmo de PDDE.

$$\begin{aligned} \pi_V^{isis} = & \eta_{BH} \frac{\partial}{\partial EARMf_{t-1}} EBH + \sum^{NPMC} \eta_{AD} \frac{\partial}{\partial EARMf_{t-1}} EAD + \\ & + \sum^{NPMC} \eta_{EH} \frac{\partial}{\partial EARMf_{t-1}} EEH \end{aligned} \quad (A.24)$$

$$\begin{aligned} \pi_{Ai}^{isis} = & \eta_{BH} \frac{\partial}{\partial EAF_{t-i}} EBH + \sum^{NPMC} \eta_{AD} \frac{\partial}{\partial EAF_{t-i}} EAD + \\ & + \sum^{NCOR} \eta_{COR} \frac{\partial}{\partial EAF_{t-i}} ECOR + \sum^{NPMC} \eta_{EH} \frac{\partial}{\partial EAF_{t-i}} EEH \end{aligned} \quad (A.25)$$

Sejam as equações abaixo que representam as parábolas do fator de correção da energia controlável, da meta de energia de vazão mínima, energia evaporada e geração hidráulica máxima, em função da energia armazenada inicial no estágio  $t$  ( $EARMf_{t-1}$ ):

$$FC(EARMf_{t-1}) = a_{EC} EARMf_{t-1}^2 + b_{EC} EARMf_{t-1} + c_{EC} \quad (A.26)$$

$$M_{EVMIN}(EARMf_{t-1}) = a_{EZ} EARMf_{t-1}^2 + b_{EZ} EARMf_{t-1} + c_{EZ} \quad (A.27)$$

$$EVAP(EARMf_{t-1}) = a_{EV} EARMf_{t-1}^2 + b_{EV} EARMf_{t-1} + c_{EV} \quad (A.28)$$

$$\overline{GH}(EARMf_{t-1}) = a_{EH} EARMf_{t-1}^2 + b_{EH} EARMf_{t-1} + c_{EH} \quad (A.18)$$

Das equações (A.24) e (A.25), os coeficientes dos cortes de Benders (MACEIRA *et al*, 1999) associados ao armazenamento inicial e às  $p$  energias afluentes passadas no estágio  $t$  são dados por (A.30) e (A.31), respectivamente:

$$\begin{aligned} \pi_V^{isis} = & \eta_{BH} \left[ (1 + \gamma) EAF_t (2a_{EC} EARMf_{t-1} + b_{EC}) - (2a_{EV} EARMf_{t-1} + b_{E} \right. \\ & \left. - (2a_{EZ} EARMf_{t-1} + b_{EZ}) \right) \\ & - \eta_{AD} (2a_{EZ} EARMf_{t-1} + b_{EZ}) + \\ & + \eta_{EH} \left[ (2a_{EH} EARMf_{t-1} + b_{EH}) - (2a_{EZ} EARMf_{t-1} + b_{EZ}) \right] \end{aligned} \quad (A.30)$$

$$\begin{aligned} \pi_{Ai}^{isis, t} = & \eta_{BH} \left[ (a_{EC} EARMf_{t-1}^2 + b_{EC} EARMf_{t-1} + c_{EC}) \gamma \phi_i^t \right] - \\ & - \eta_{AD} \left[ (1 - \gamma) \phi_i^t \right] + \\ & + \eta_{COR} \left( \sum^{NCOR} \pi_{Ai}^{isis, t-1} \right) - \\ & - \eta_{EH} \left[ (1 - \gamma) \phi_i^t \right] \end{aligned} \quad (A.31)$$

Dessa forma, o módulo de cálculo da política de operação é responsável pela construção iterativa da função de custo futuro através do algoritmo de PDDE.

O critério de convergência do modelo NEWAVE é alcançado quando a função de custo esperado futuro real estiver bem representada através da função linear por partes definida pelos cortes de Benders.

Dado que o processo de otimização da política convergiu, podem ser obtidos indicadores estatísticos da operação futura do sistema através do módulo de simulação da operação descrito na próxima seção.

### **A.2.3. Módulo de Simulação da Operação**

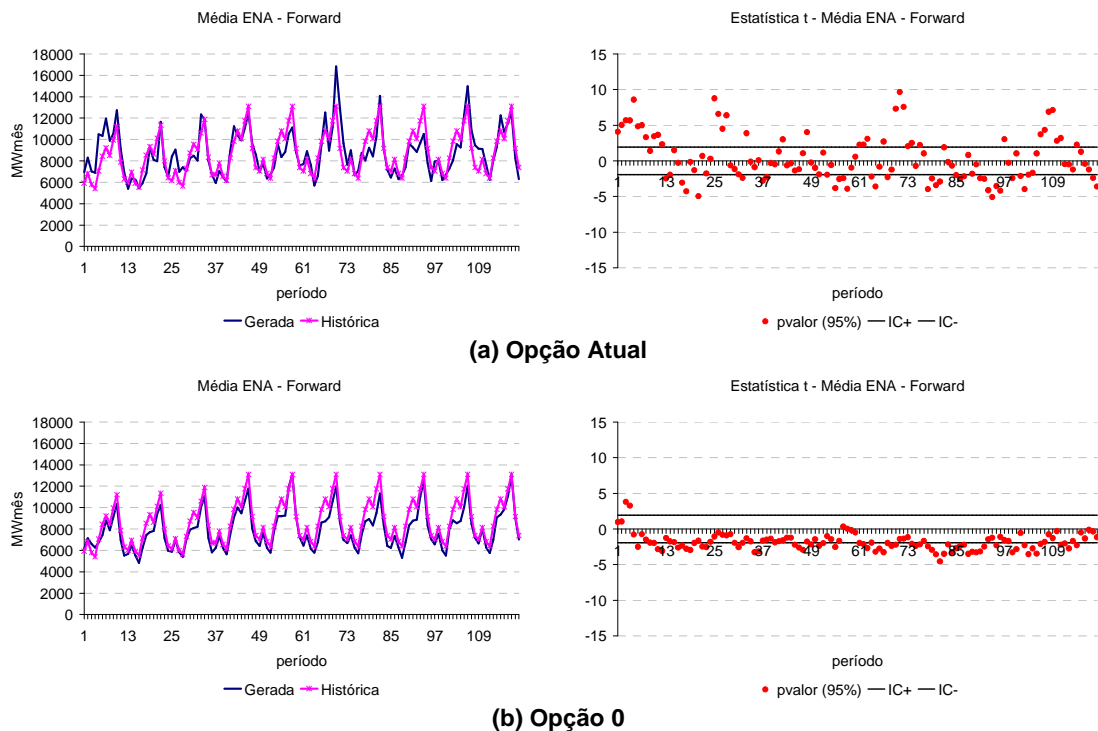
Neste módulo é realizada a simulação da operação do sistema para diversos cenários de energias naturais afluentes, gerados pelo modelo PAR( $p$ ), ou para a série histórica. Nesta simulação são usadas as funções de custo futuro que foram calculadas previamente no módulo de cálculo da política de operação. Com as soluções obtidas para cada cenário e período do horizonte de estudo, é possível calcular os riscos de não atendimento à carga própria, valor esperado do custo de operação, gerações hidráulicas e térmicas de cada subsistema, armazenamentos, intercâmbios, vertimentos entre outros parâmetros.

## Apêndice B Análise dos Cenários Gerados de Energia Natural Afluyente (Complementação)

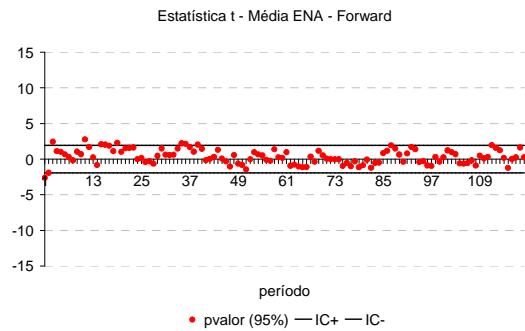
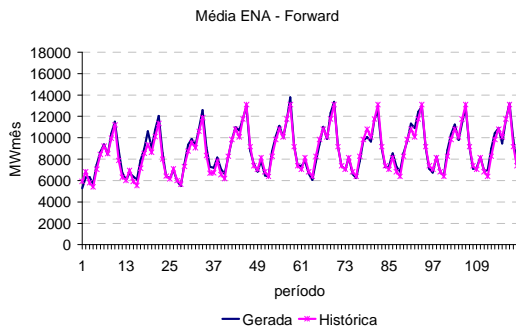
### B.1 Análise dos Cenários Gerados para Passo Forward

#### B.1.1 Média dos Cenários

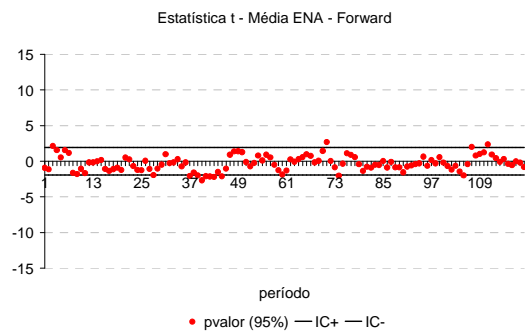
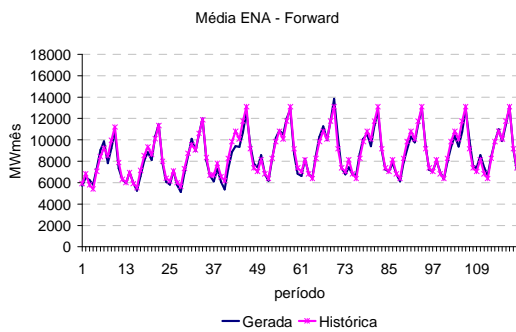
A avaliação da média dos cenários utilizados nos passo forward para os subsistemas Sul, Nordeste e Norte é apresentada na Figura 178, Figura 179 e Figura 180, respectivamente.



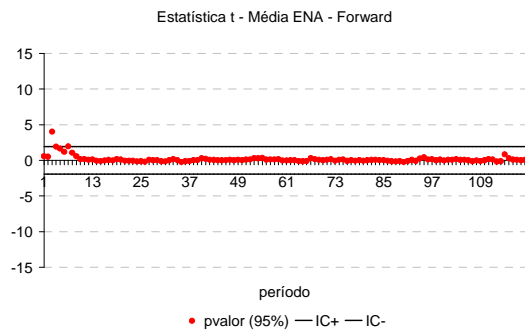
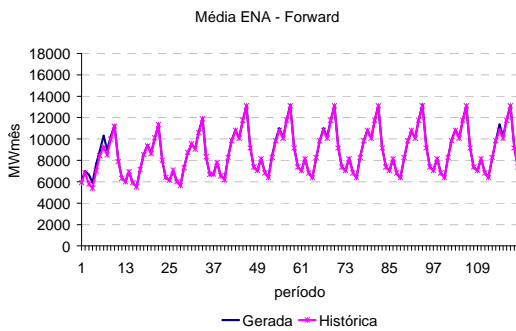




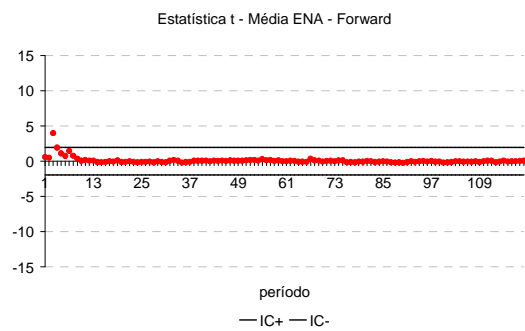
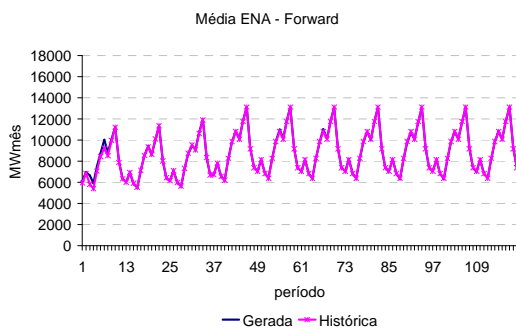
**(c) Opção 1 e 1AAS**



**(d) Opção 2, 3 e 4**

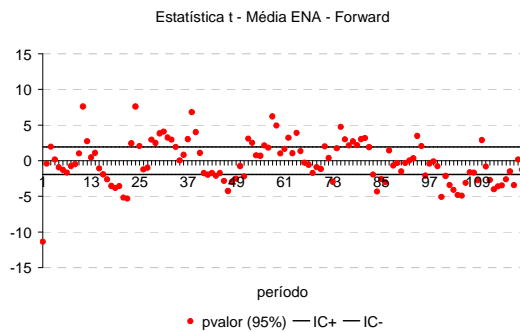
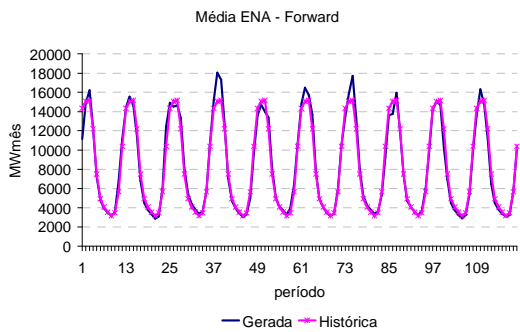


**(e) Opção LHC**

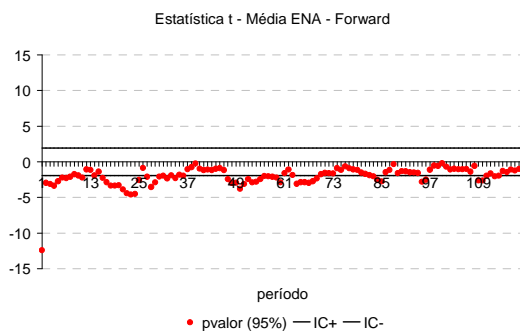
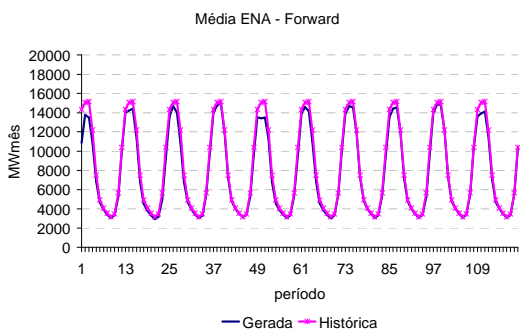


**(f) Opção AD**

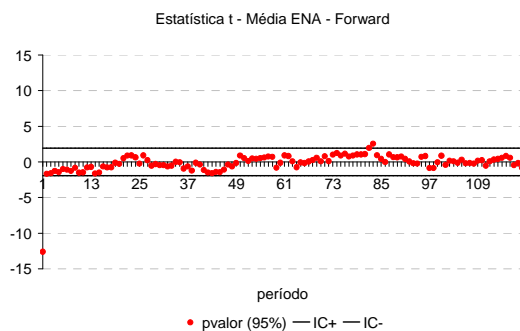
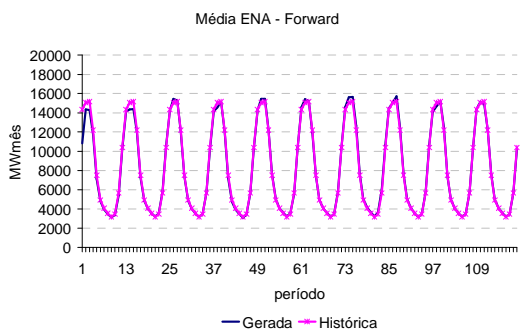
Figura 178: Média Cenários Forward – Sul



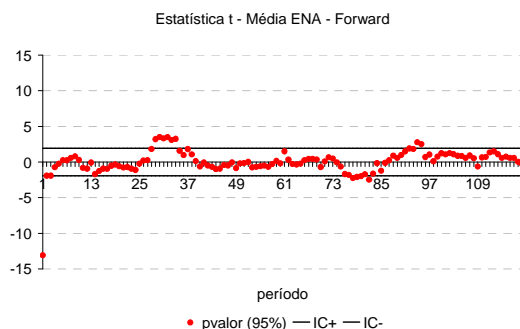
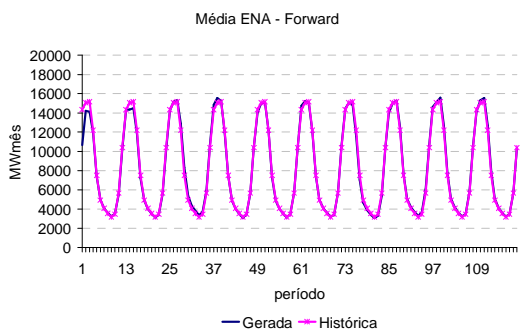
**(a) Opção Atual**



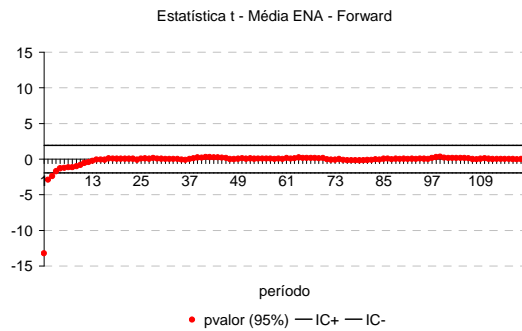
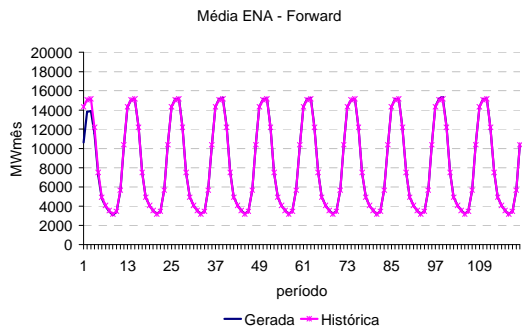
**(b) Opção 0**



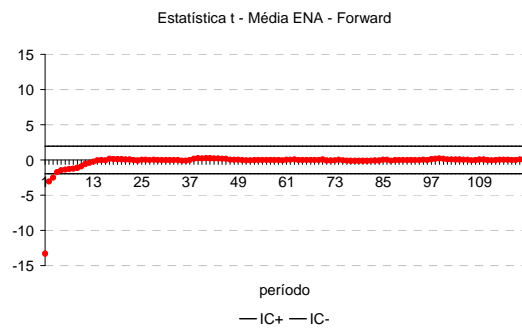
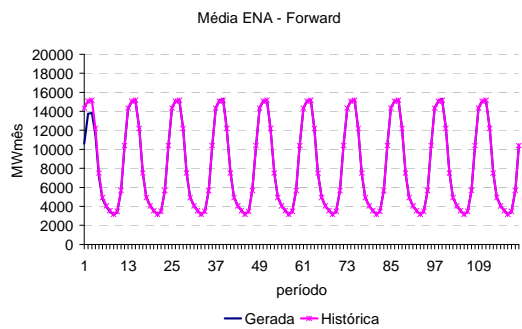
**(c) Opção 1 e 1AAS**



**(d) Opção 2, 3 e 4**

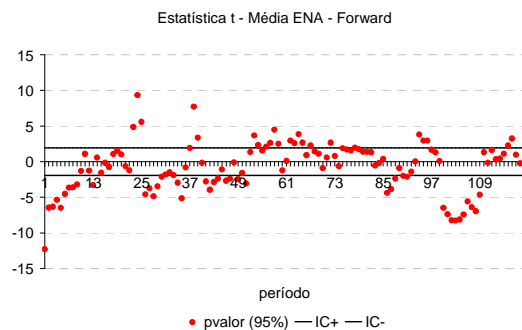
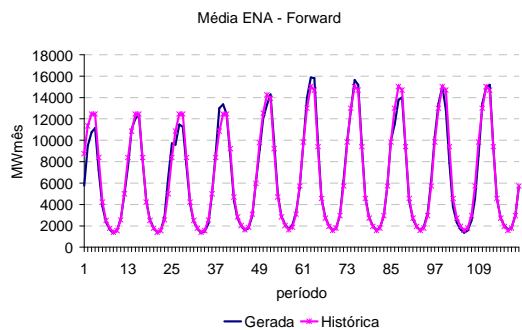


**(e) Opção LHC**

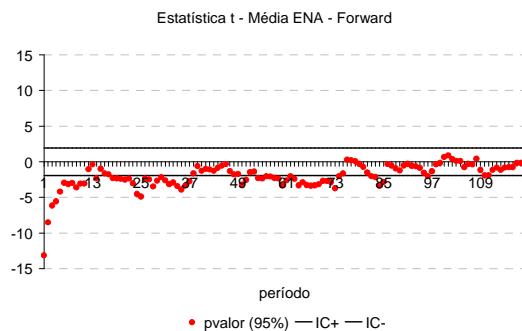
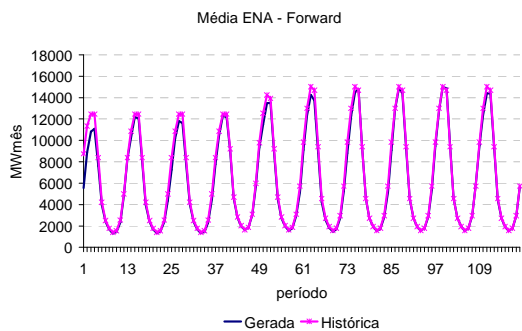


**(f) Opção AD**

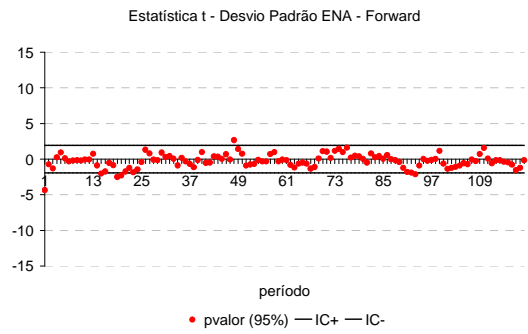
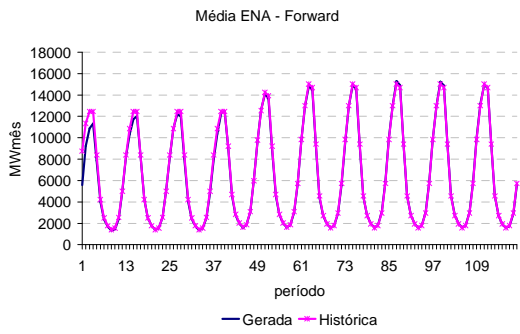
Figura 179: Média Cenários Forward – Nordeste



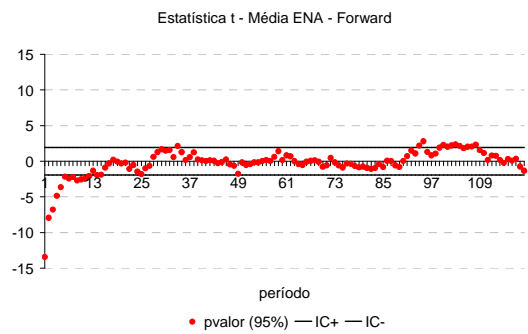
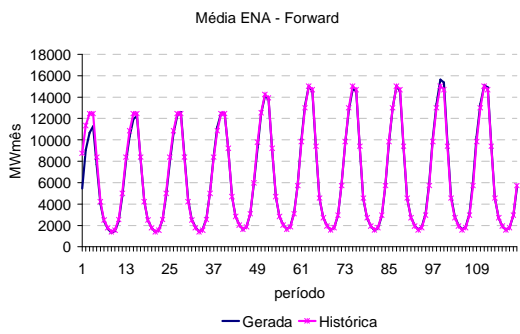
**(a) Opção Atual**



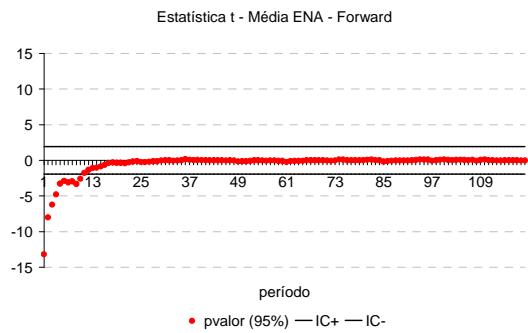
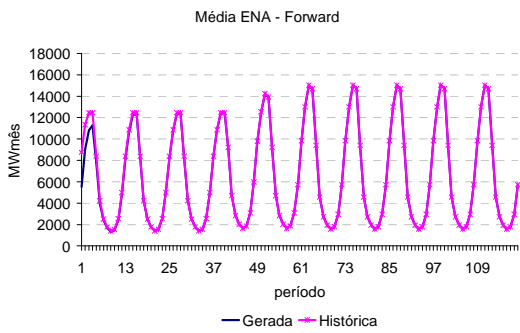
**(b) Opção 0**



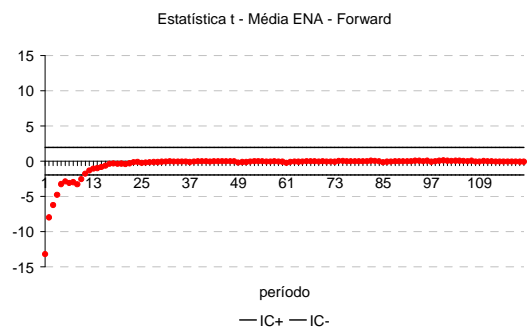
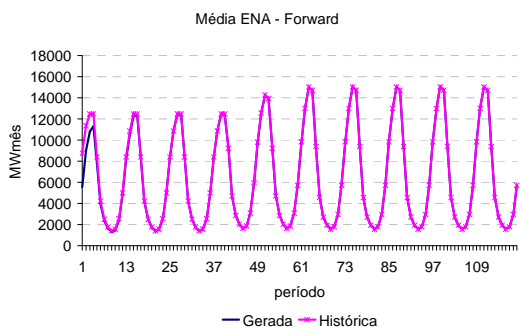
**(c) Opção 1 e 1AAS**



**(d) Opção 2, 3 e 4**



**(e) Opção LHC**

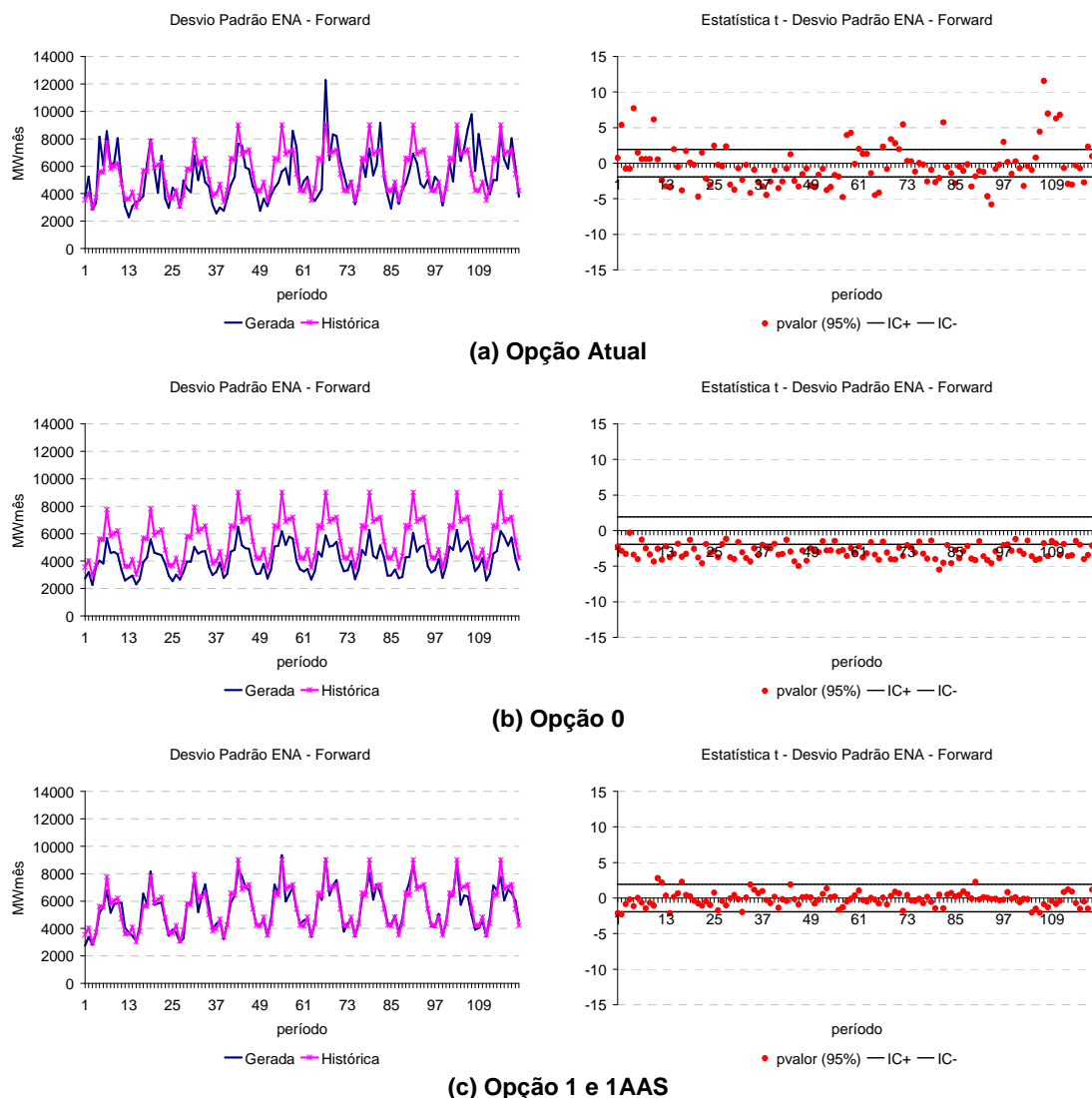


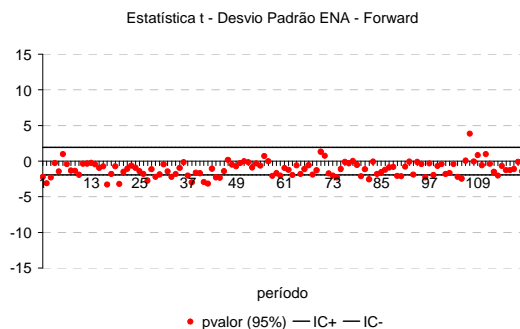
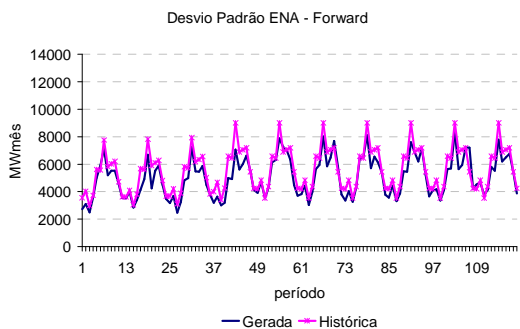
**(f) Opção AD**

Figura 180: Média Cenários Forward – Norte

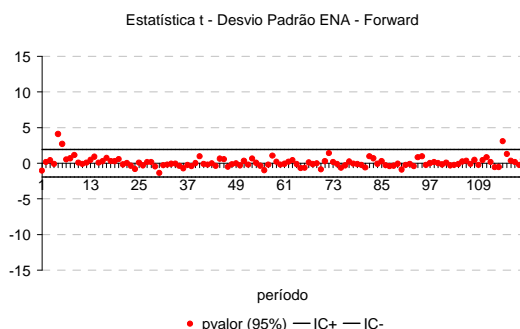
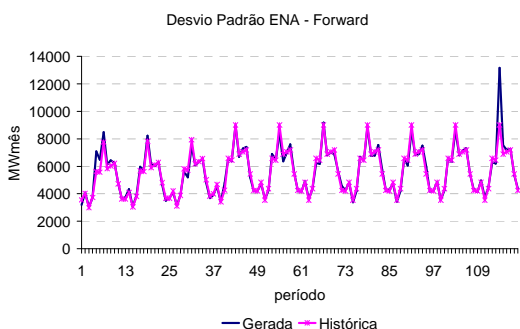
### B.1.2 Desvio-padrão dos Cenários

Na Figura 181 são apresentados os resultados obtidos para o subsistema Sul. As mesmas observações feitas para o subsistema Sudeste se aplicam neste caso. Novamente as opções Atual e 0 não conseguem representar de forma adequada a estimativa do desvio padrão.

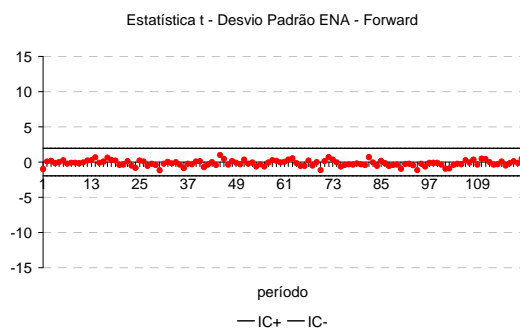
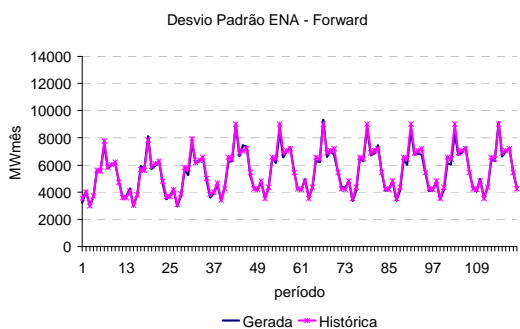




**(d) Opção 2, 3 e 4**



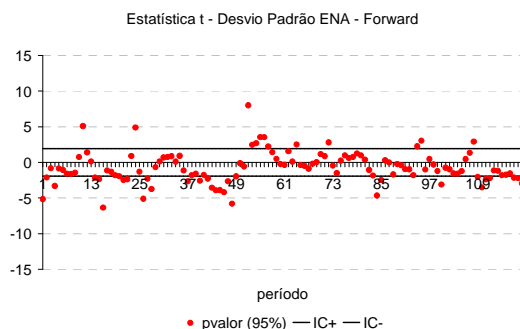
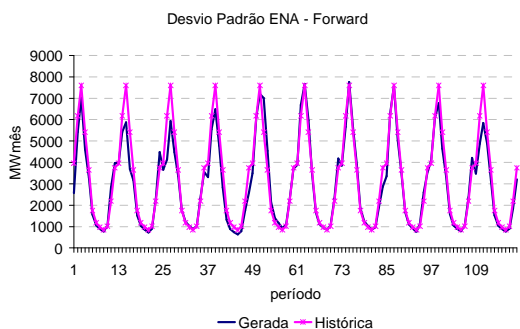
**(e) Opção LHC**



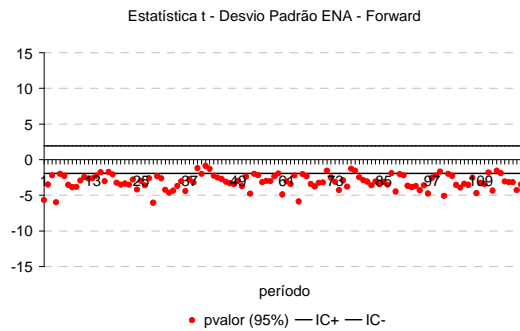
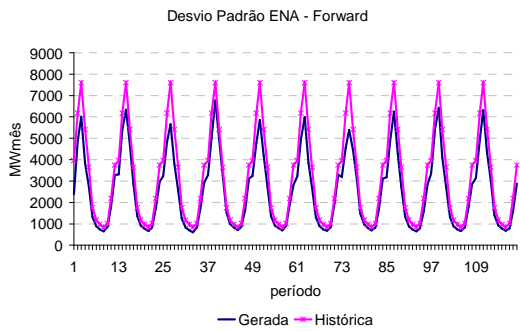
**(f) Opção AD**

Figura 181: Desvio-Padrão Cenários Forward – Sul

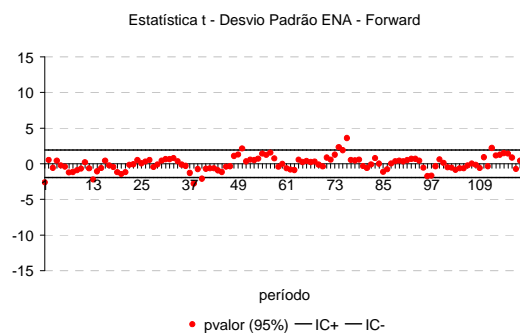
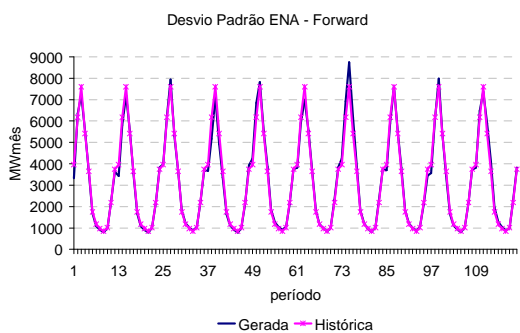
Analisando a Figura 182 e a Figura 183, para o subsistema Nordeste e Norte, respectivamente, pode-se observar o mesmo comportamento descrito para o subsistema Sudeste.



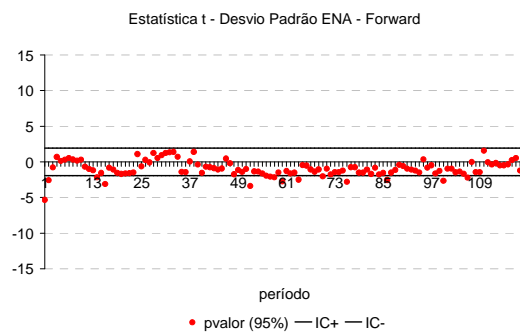
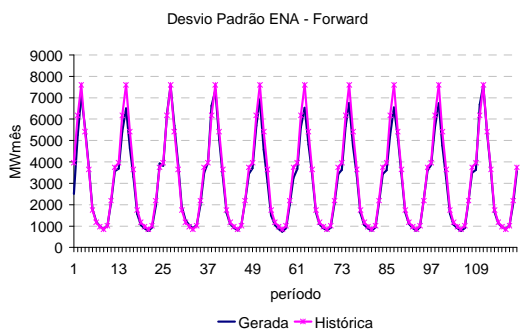
**(a) Opção Atual**



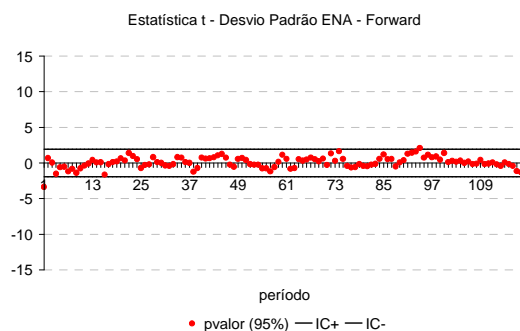
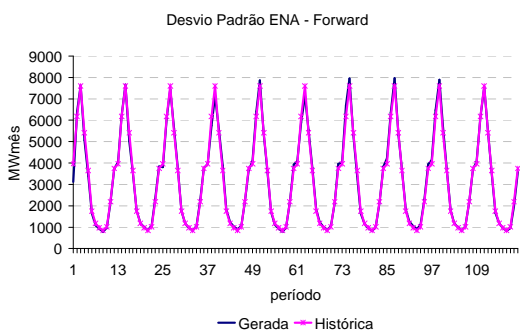
**(b) Opção 0**



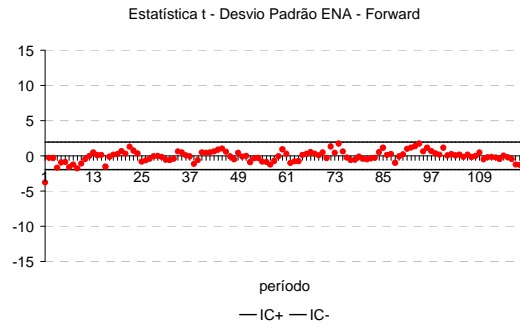
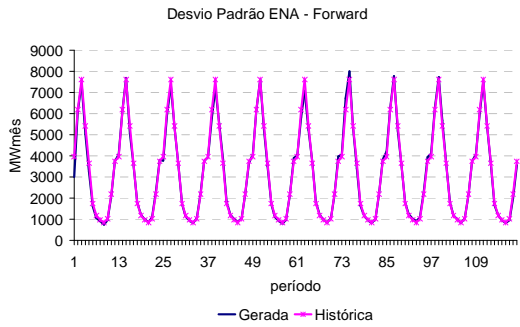
**(c) Opção 1 e 1AAS**



**(d) Opção 2, 3 e 4**

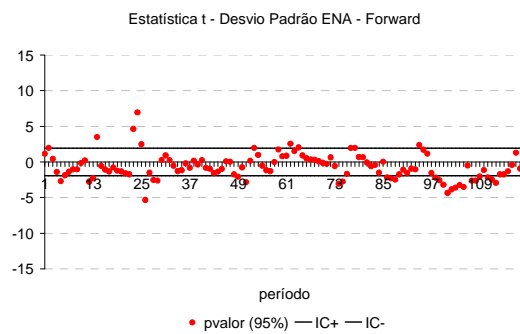
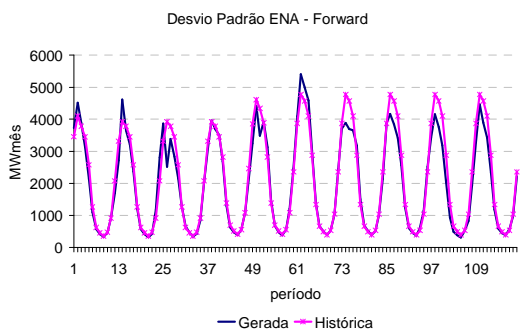


**(e) Opção LHC**

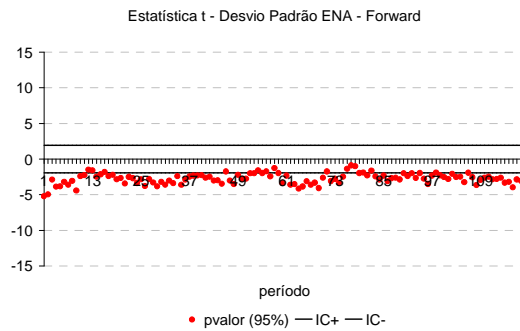
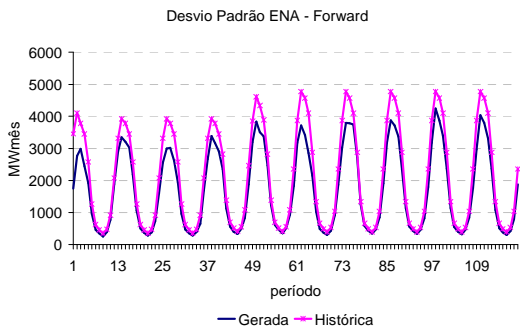


**(f) Opção AD**

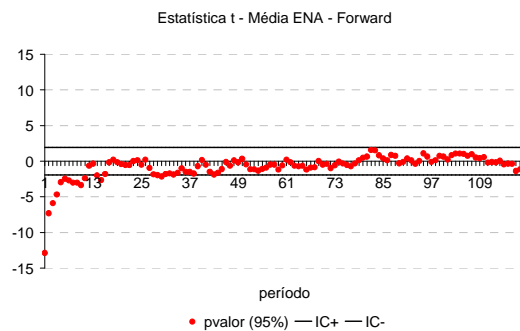
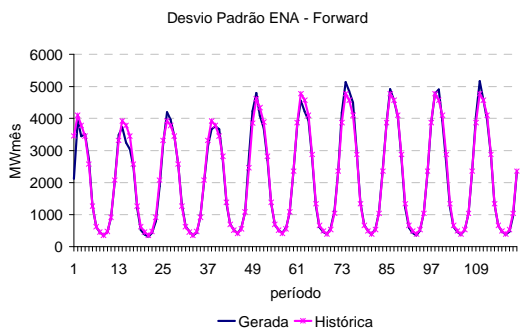
Figura 182: Desvio-Padrão Cenários Forward – Nordeste



**(a) Opção Atual**

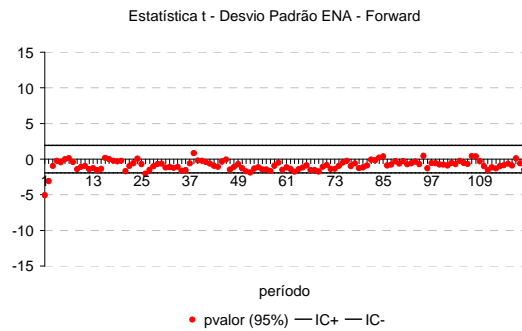
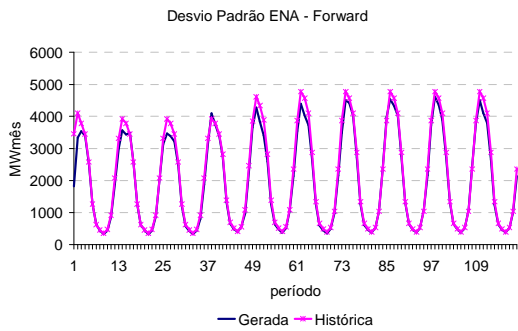


**(b) Opção 0**

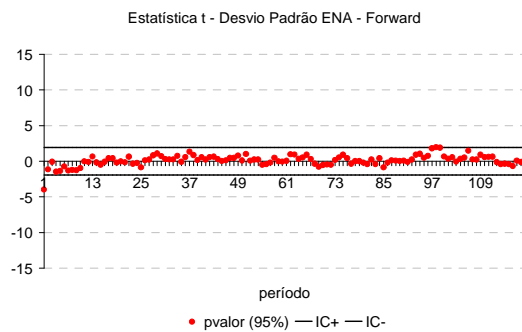
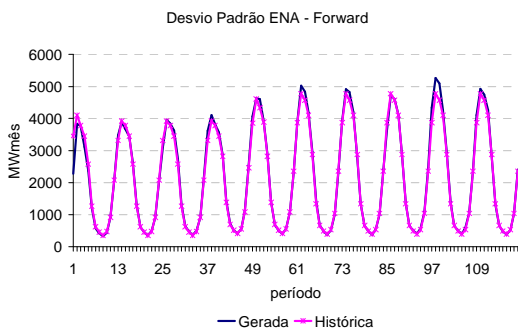


**(c) Opção 1 e 1AAS**

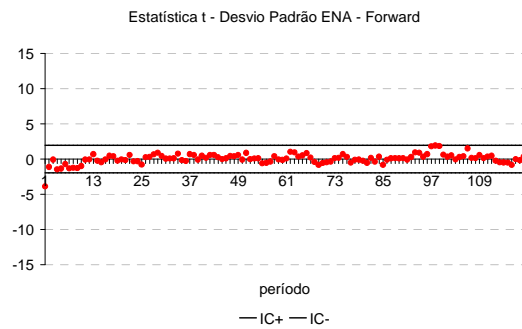
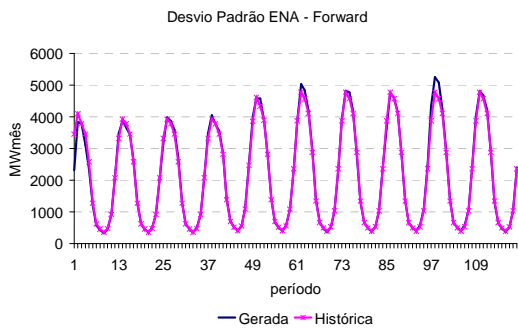




(d) Opção 2, 3 e 4



(e) Opção LHC



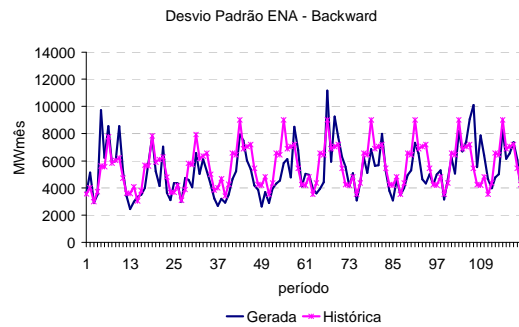
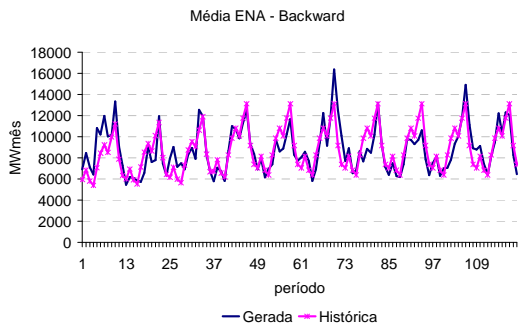
(f) Opção AD

Figura 183: Desvio-Padrão Cenários Forward – Norte

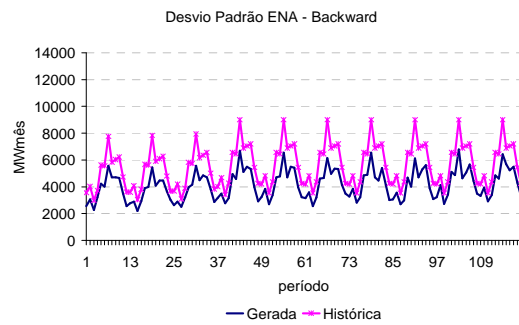
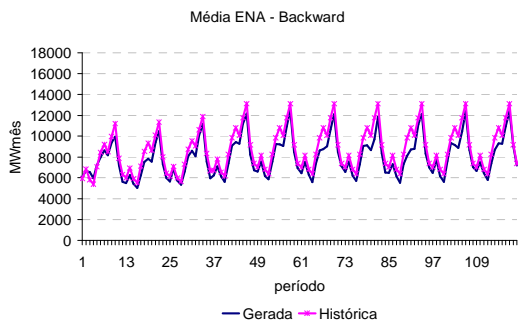
## B.2. Análise dos Cenários Gerados para Passo Backward

### B.2.2 Testes não Condicionados para Média e Desvio-padrão

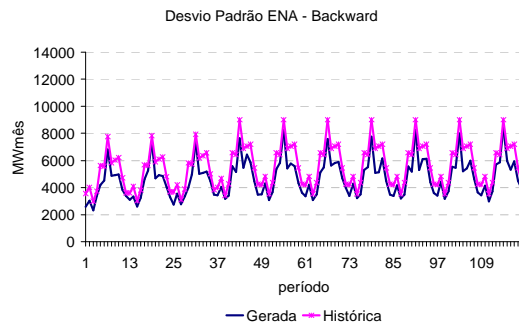
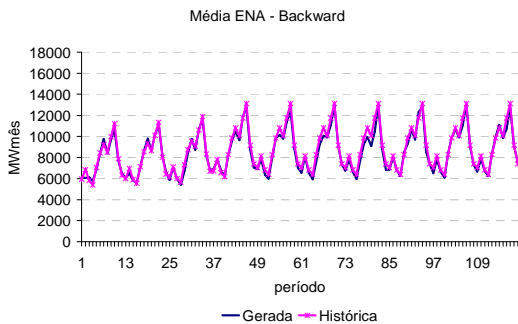
A avaliação da média e do desvio-padrão dos cenários utilizados nos passo backward para os subsistemas Sul, Nordeste e Norte é apresentada na Figura 184, Figura 185 e Figura 186, respectivamente.



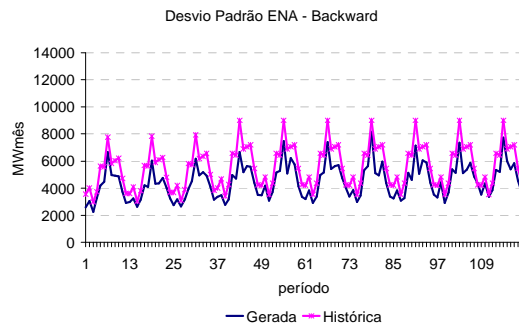
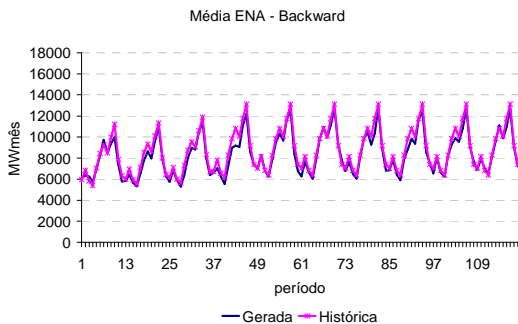
**(a) Opção Atual**



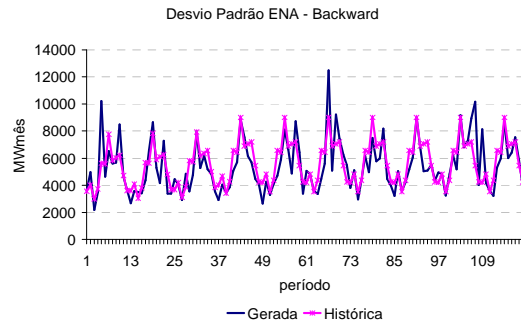
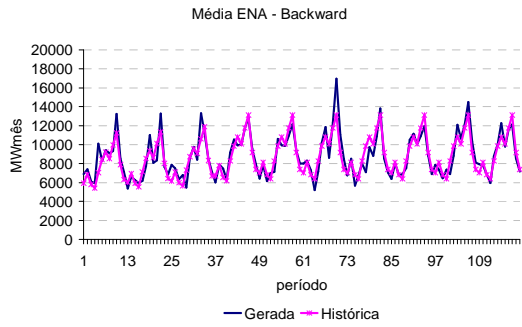
**(b) Opção 0**



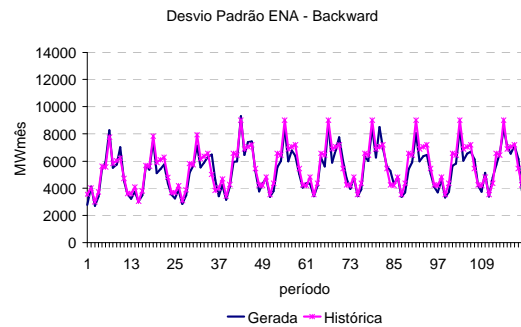
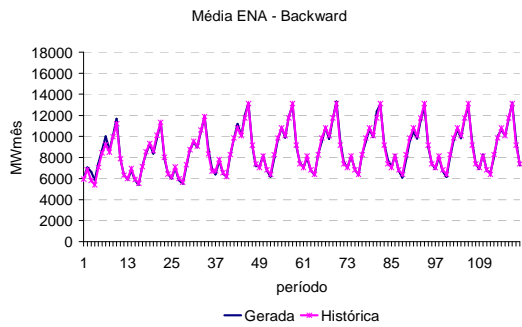
**(c) Opção 1**



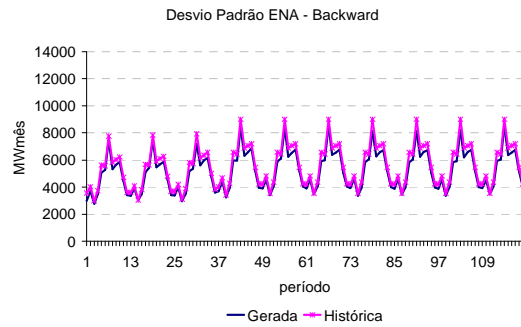
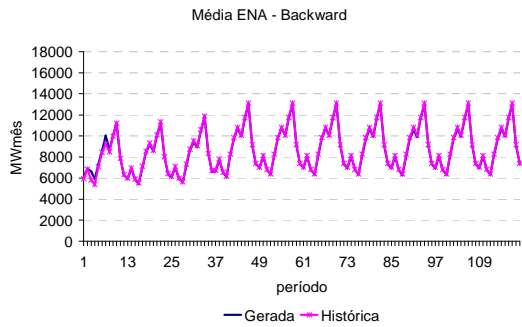
**(d) Opção 4**



**(e) Opção 1AAS**

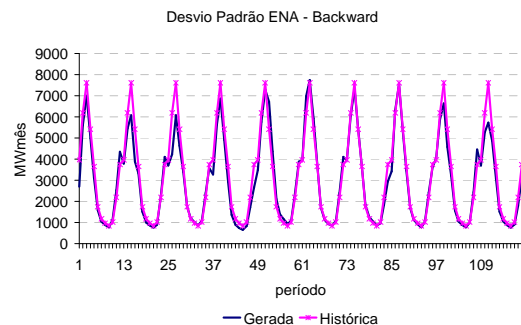
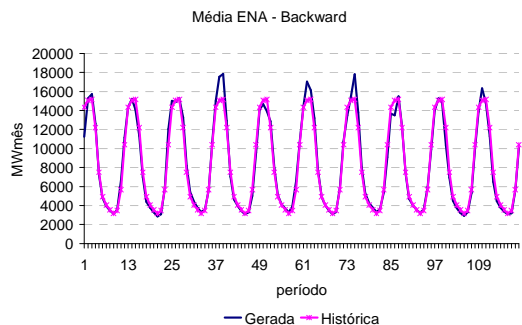


**(f) Opção LHC**

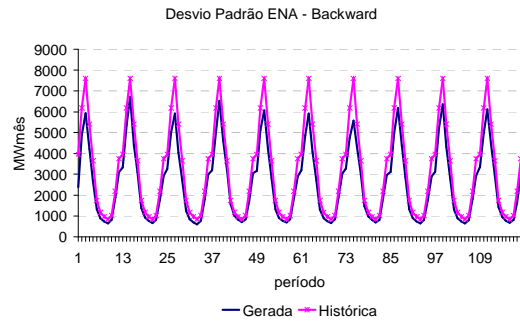
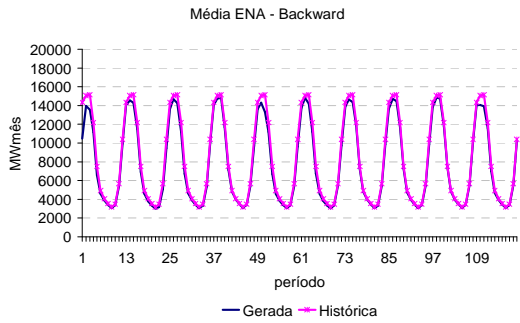


**(g) Opção AD**

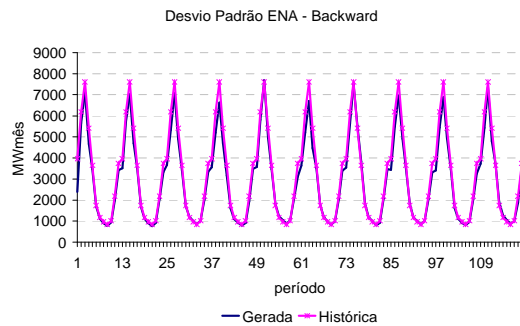
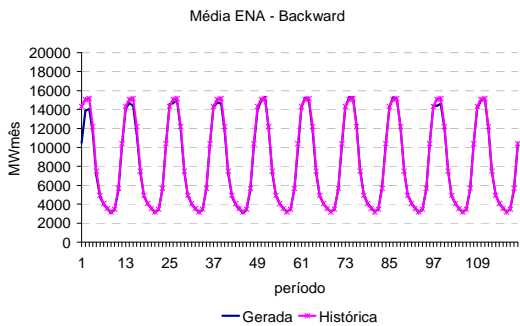
Figura 184: Média Cenários Backward – Sul



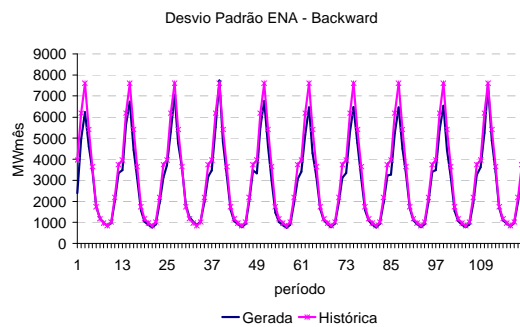
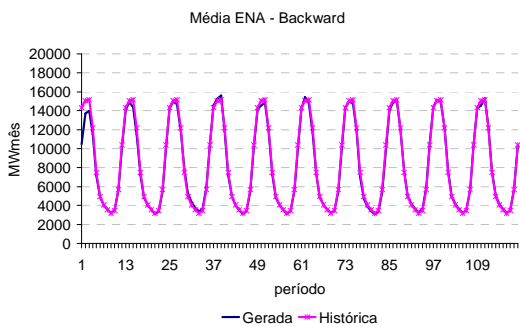
**(a) Opção Atual**



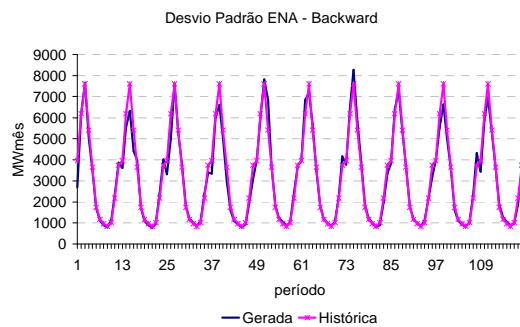
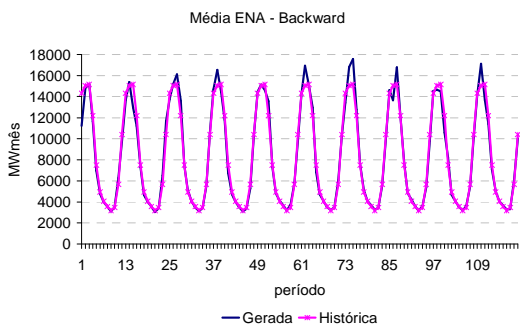
**(b) Opção 0**



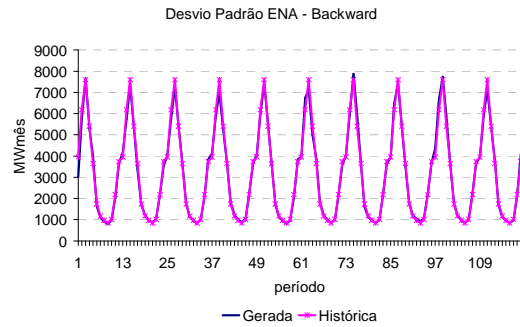
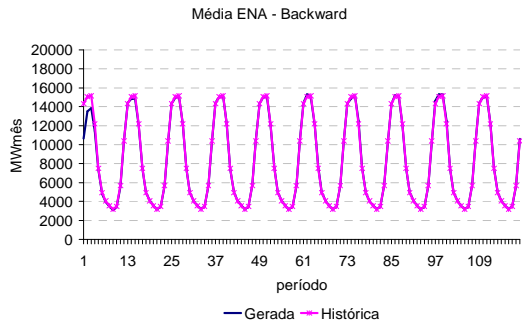
**(c) Opção 1**



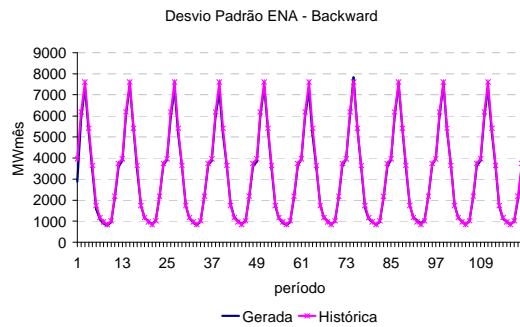
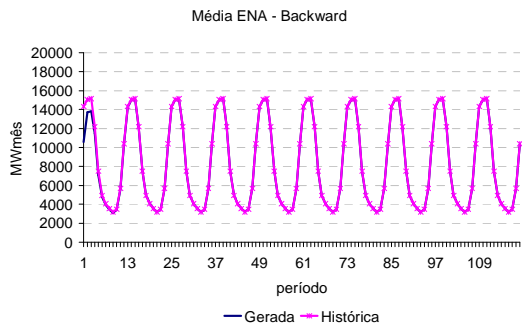
**(d) Opção 4**



**(e) Opção 1AAS**

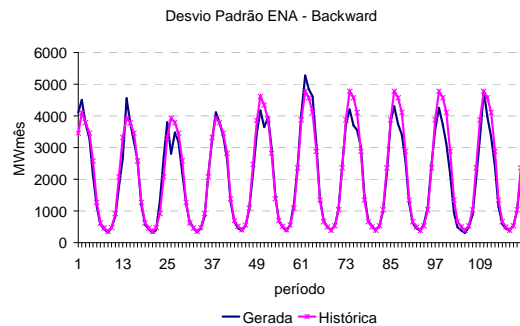
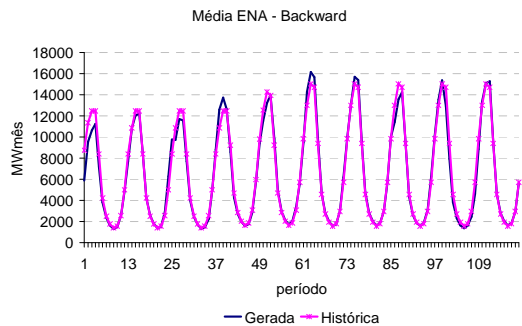


**(f) Opção LHC**

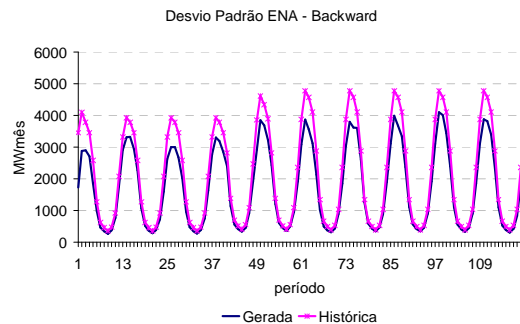
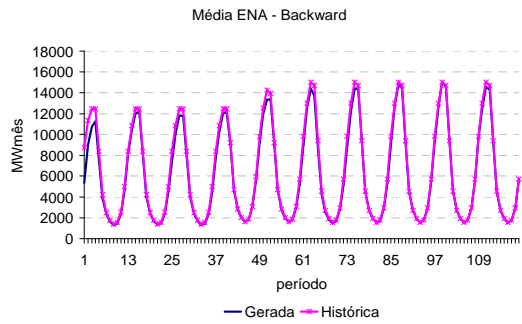


**(g) Opção AD**

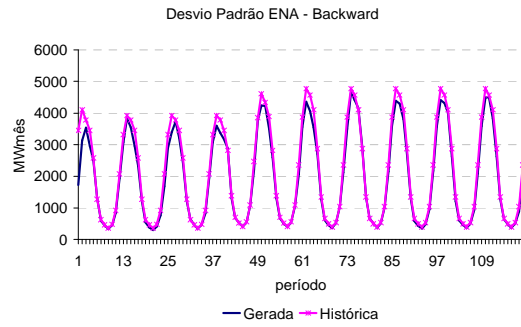
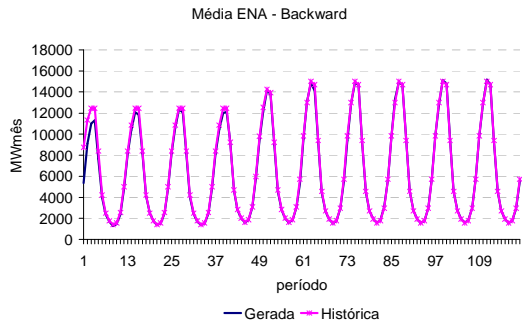
Figura 185: Média Cenários Backward – Nordeste



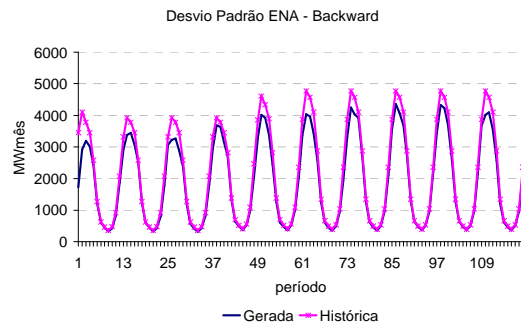
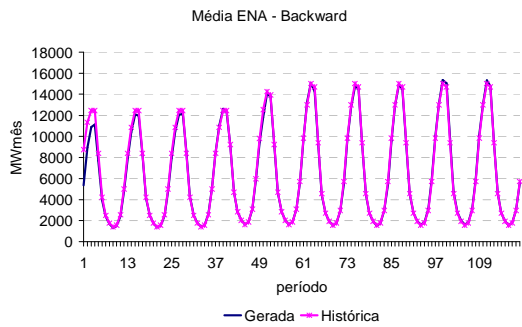
**(a) Opção Atual**



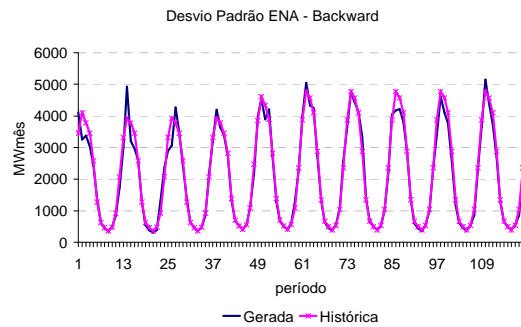
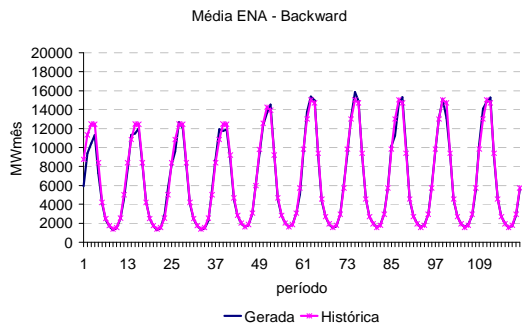
**(b) Opção 0**



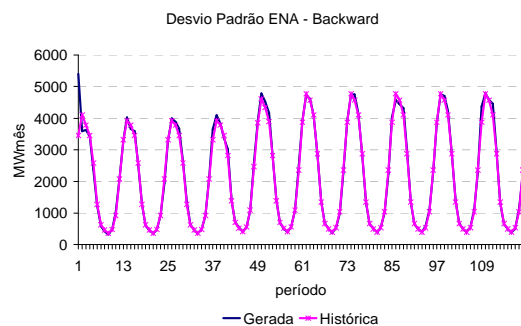
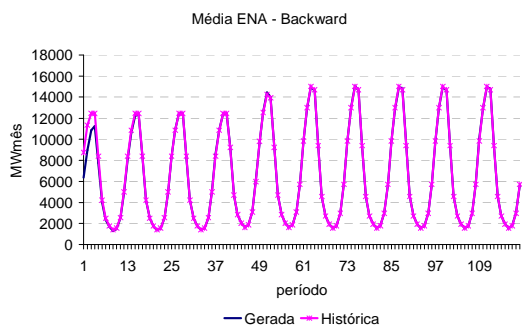
**(c) Opção 1**



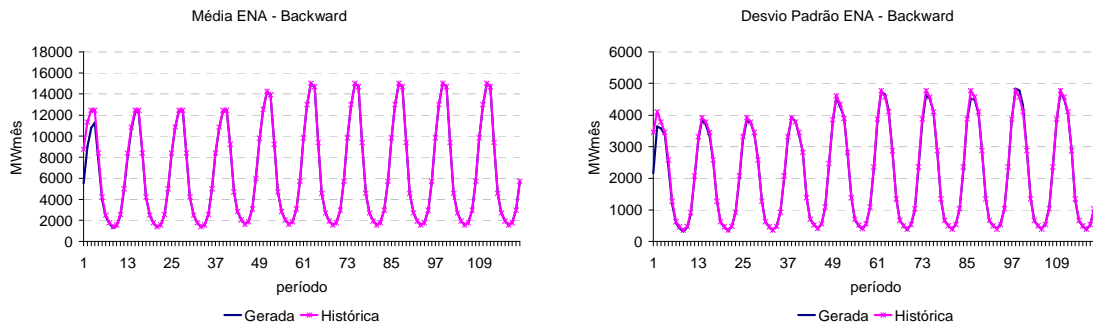
**(d) Opção 4**



**(e) Opção 1AAS**



**(f) Opção LHC**

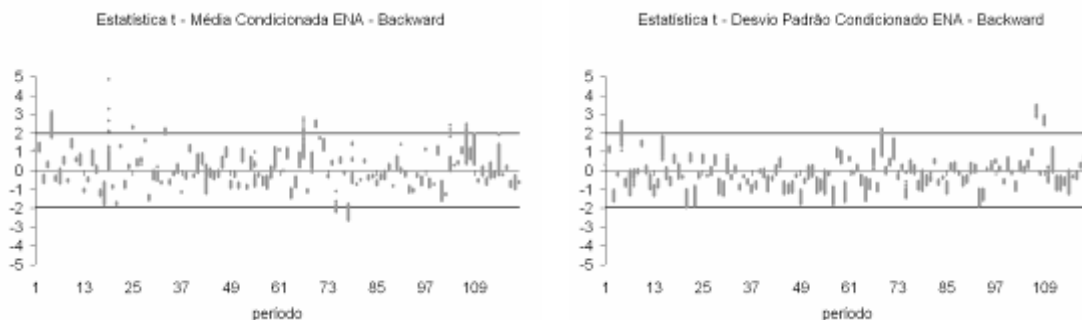


**(g) Opção AD**

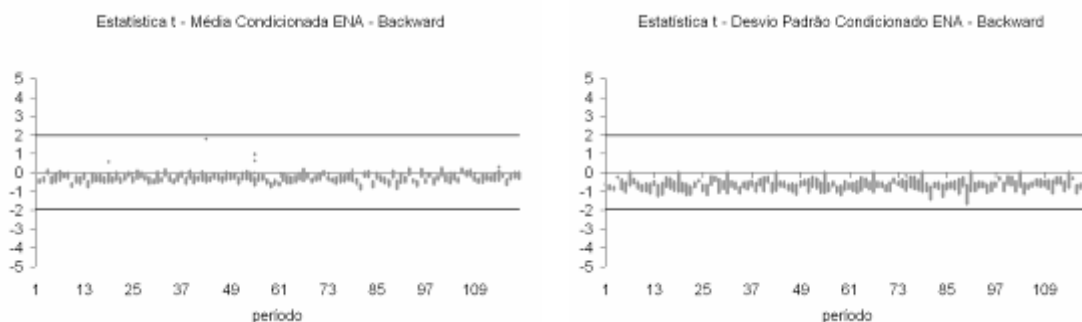
Figura 186: Média Cenários Backward – Norte

## B.2.2 Testes Condicionados para Média e Desvio-padrão

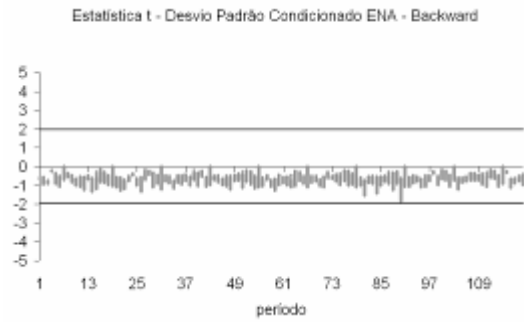
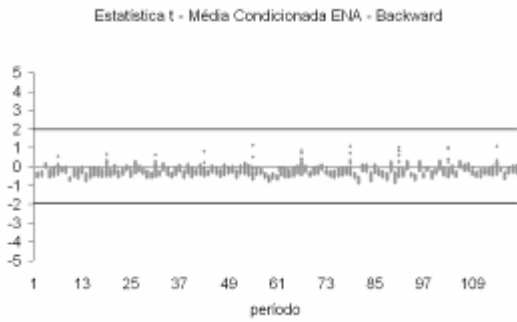
A avaliação da média e do desvio-padrão dos cenários utilizados nos passo backward para os subsistemas Sul, Nordeste e Norte é apresentada na Figura 187, Figura 188 e Figura 189, respectivamente.



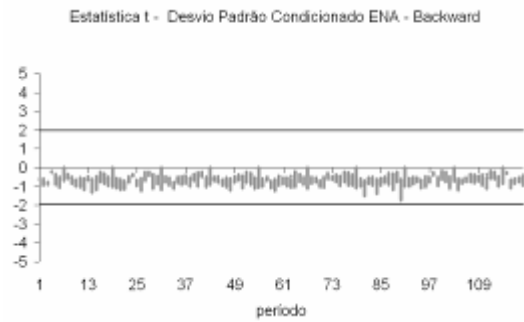
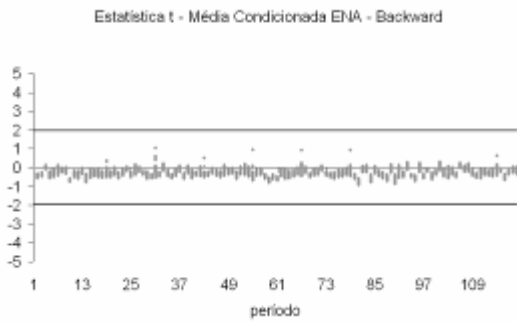
**(a) Opção Atual**



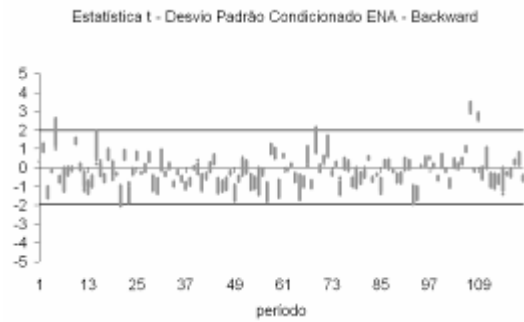
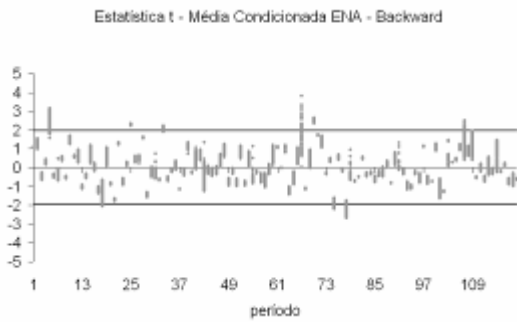
**(b) Opção 0**



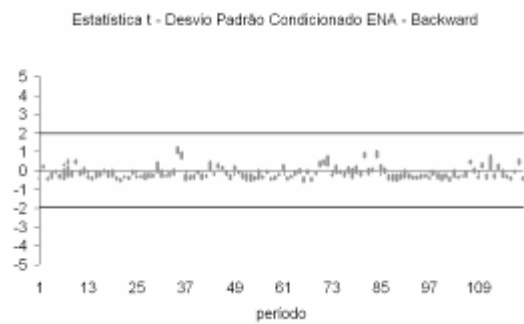
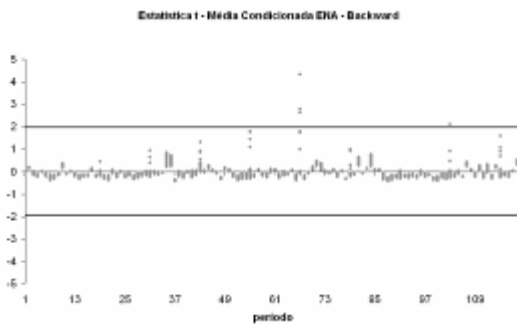
**(c) Opção 1**



**(d) Opção 4**

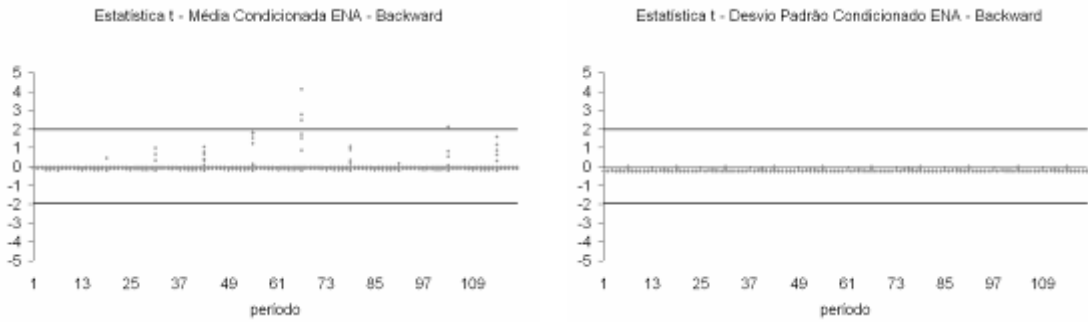


**(e) Opção 1AAS**



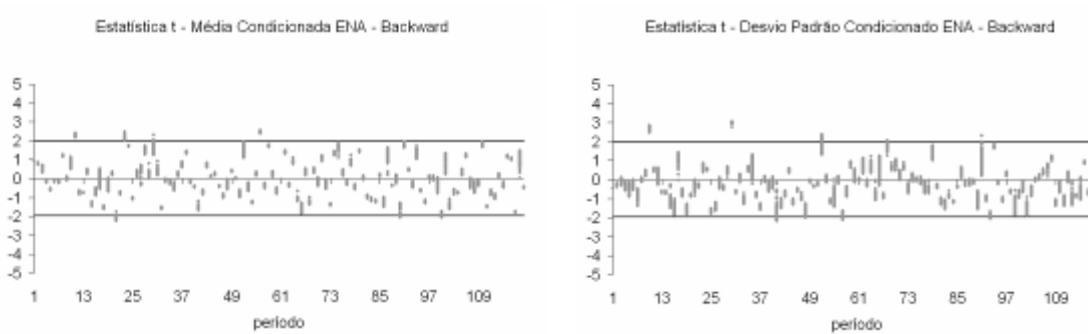
**(f) Opção LHC**



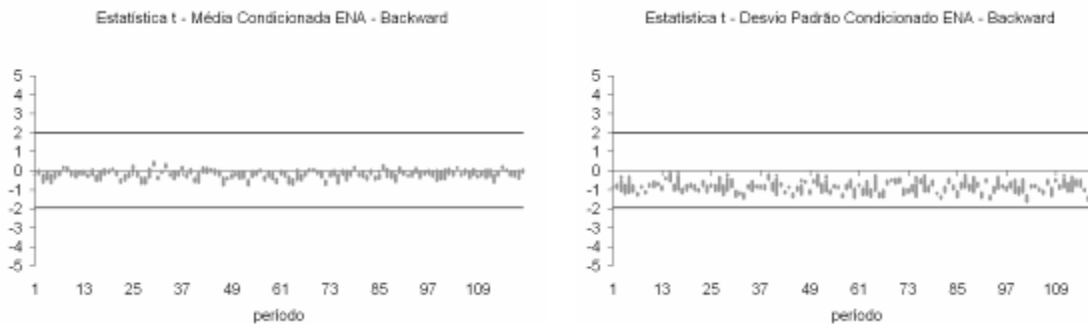


**(g) Opção AD**

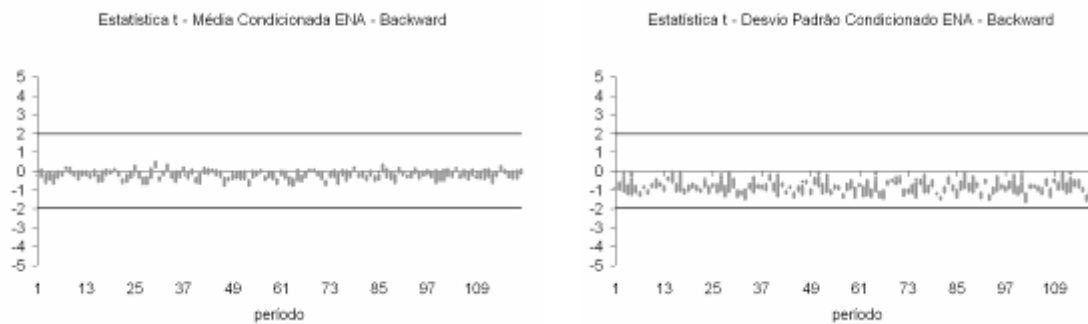
Figura 187: Estatística t Média e Desvio-Padrão – Cenários Backward – Sul



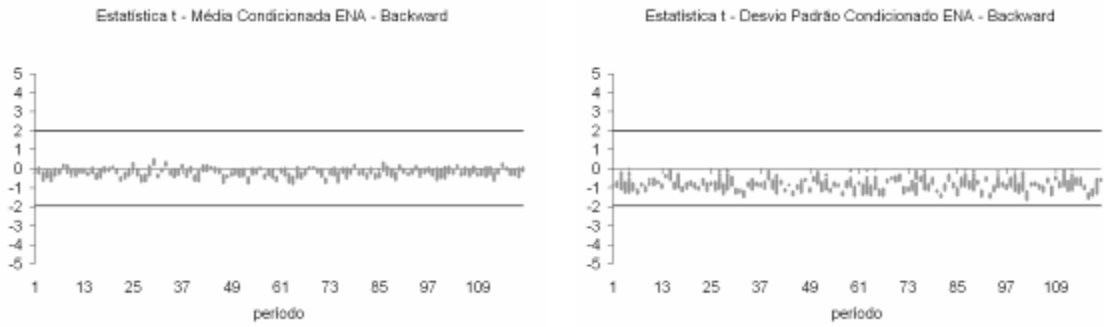
**(a) Opção Atual**



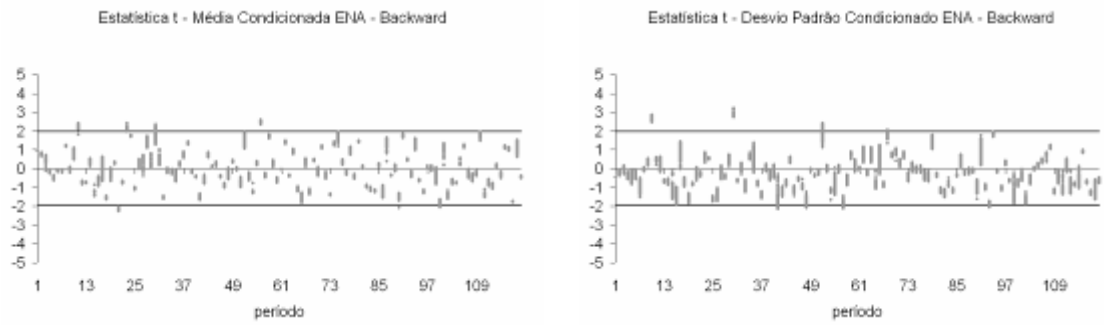
**(b) Opção 0**



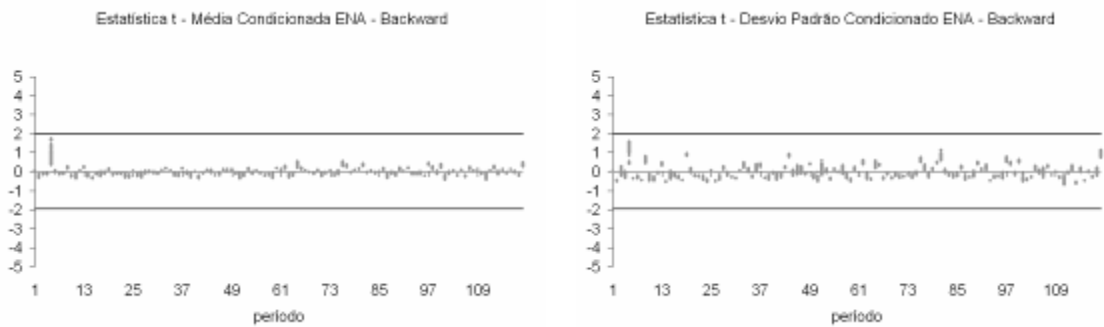
**(c) Opção 1**



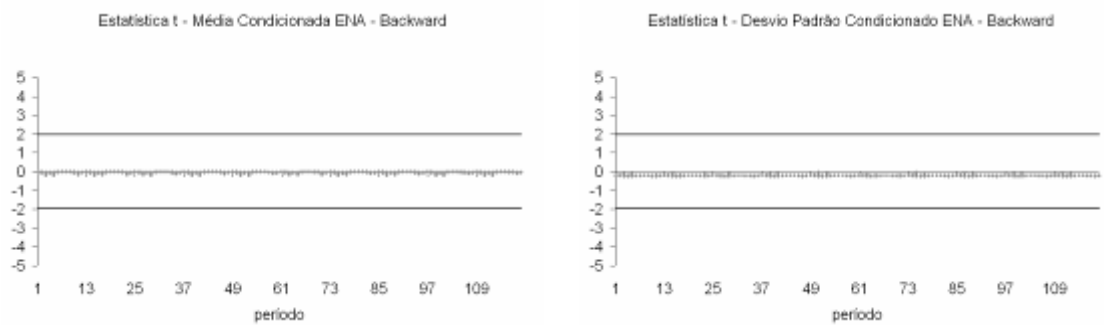
**(d) Opção 4**



**(e) Opção 1AAS**

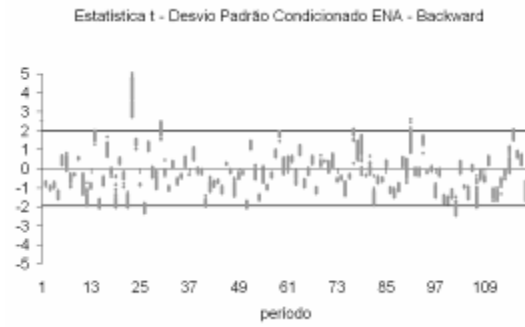
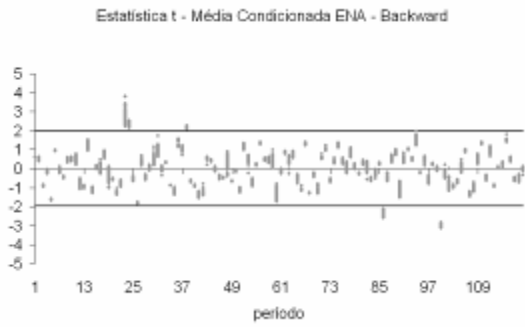


**(f) Opção LHC**

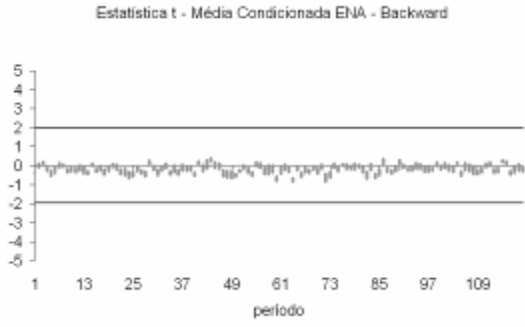


**(g) Opção AD**

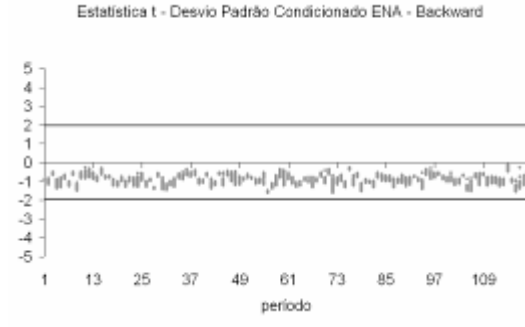
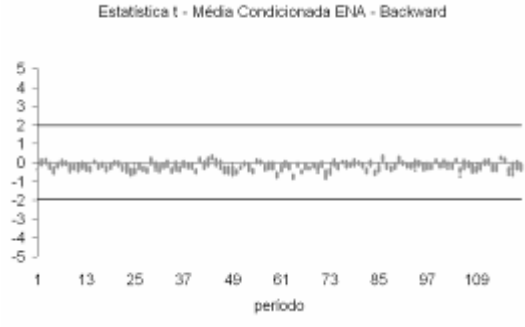
Figura 188: Estatística t Média e Desvio-Padrão – Cenários Backward - Nordeste



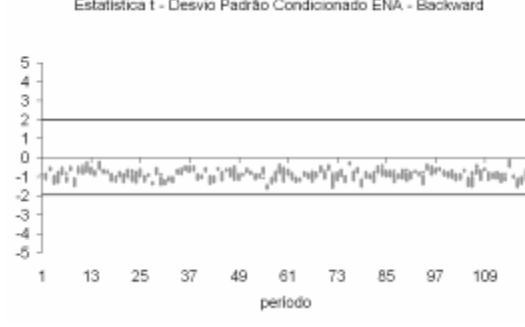
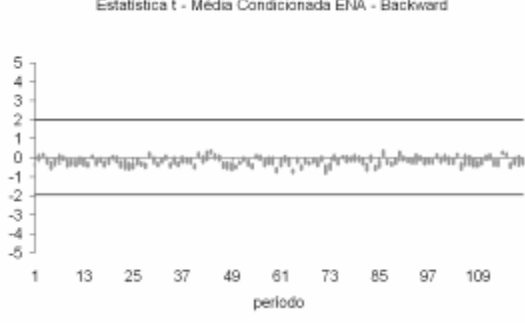
**(a) Opção Atual**



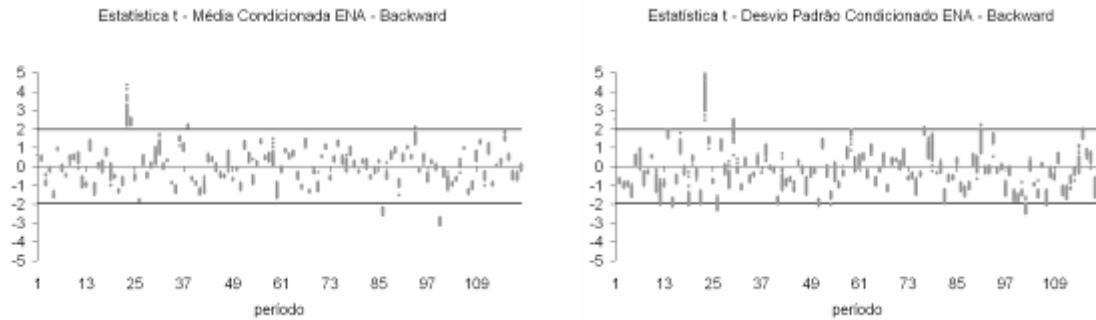
**(b) Opção 0**



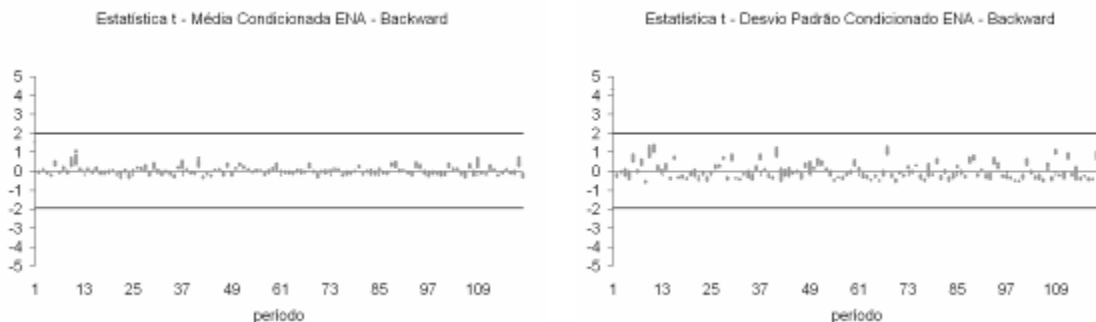
**(c) Opção 1**



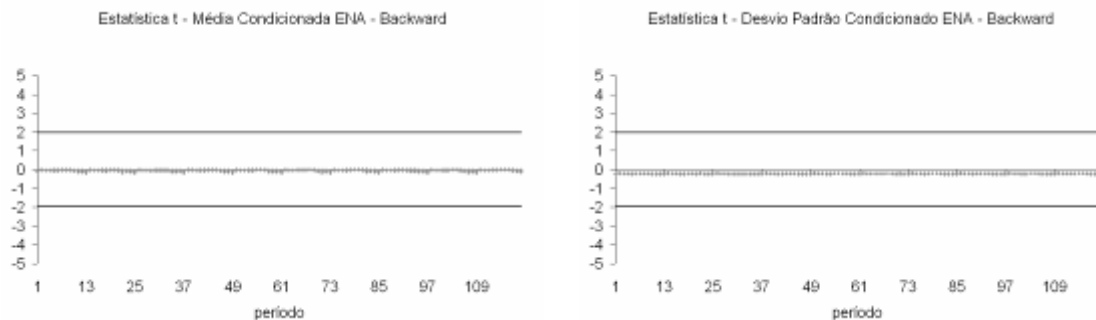
**(d) Opção 4**



(e) Opção 1AAS



(f) Opção LHC

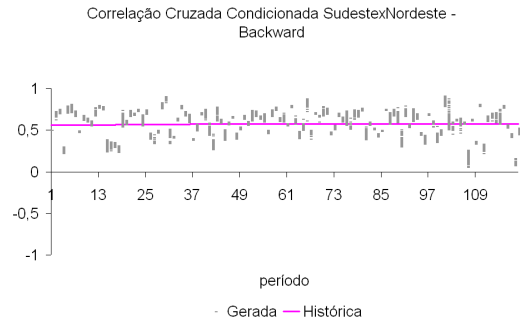
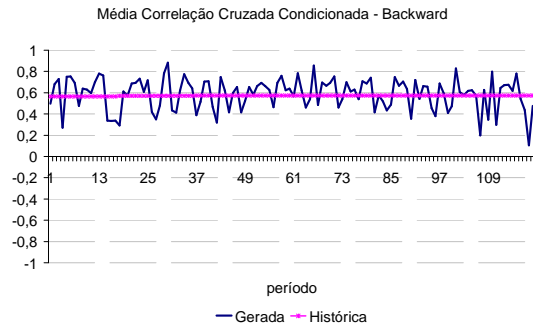


(g) Opção AD

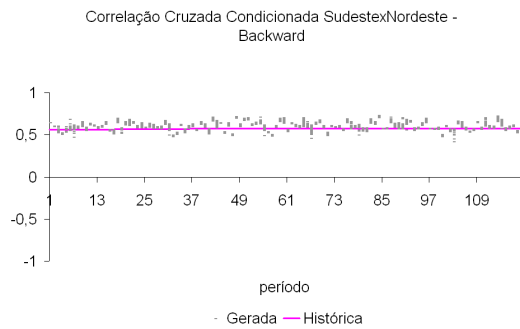
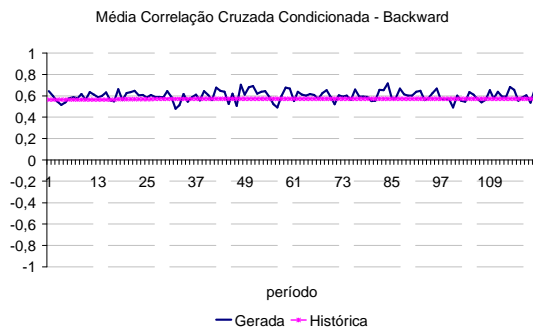
Figura 189: Estatística t Média e Desvio-Padrão – Cenários Backward - Norte

### B.2.3 Testes Condicionados para Correlação Cruzada

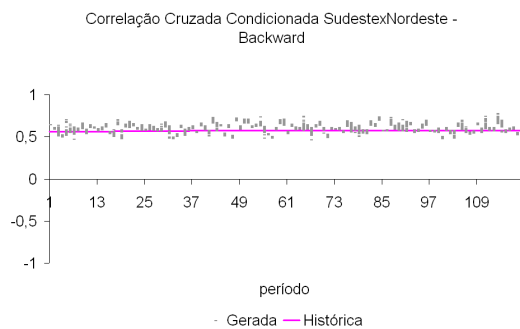
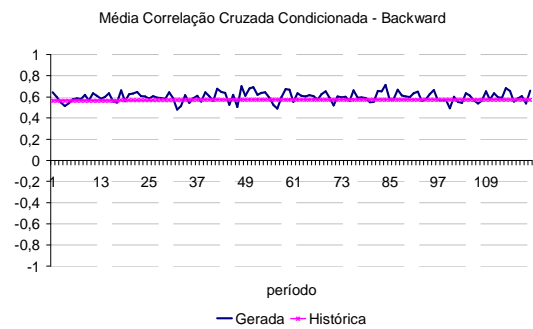
Da Figura 190a até a Figura 190g pode-se analisar o comportamento da correlação cruzada entre os subsistemas Sudeste e Nordeste. As correlações cruzadas obtidas com as opções LHC e AD continuam, em geral, abaixo do valor histórico, porém as diferenças relação ao histórico são bem menores. Conforme observado anteriormente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico.



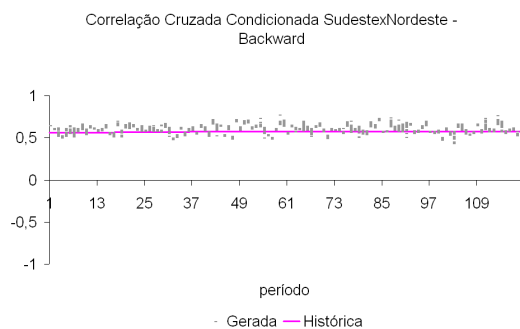
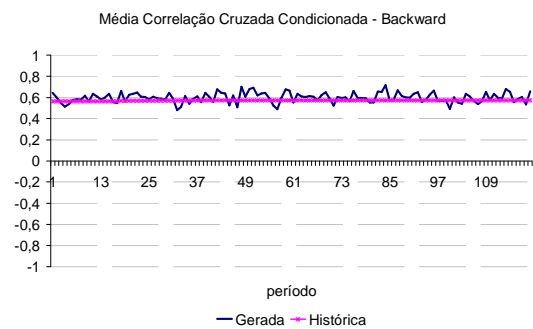
**(a) Opção Atual**



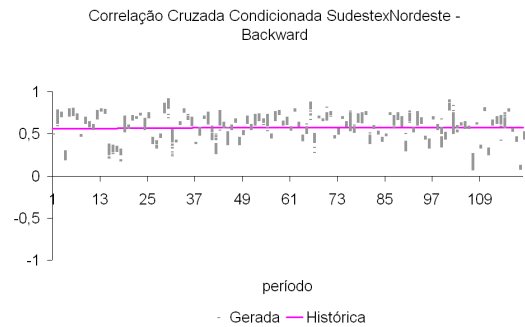
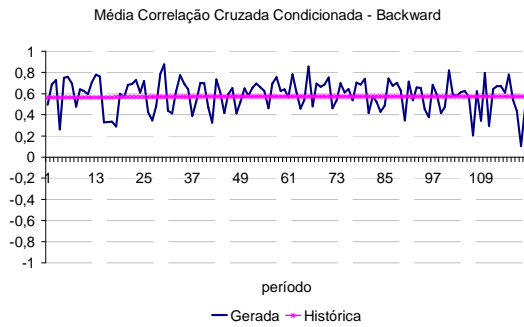
**(b) Opção 0**



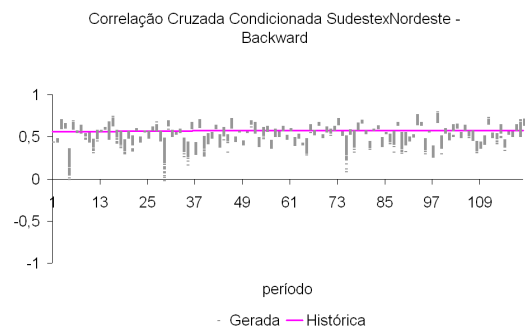
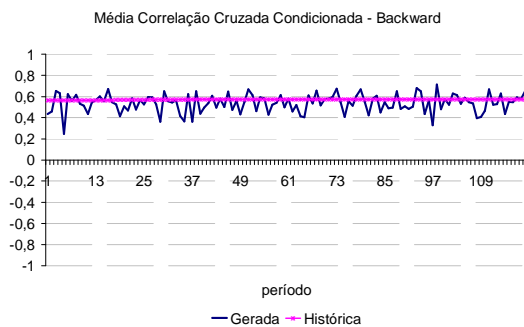
**(c) Opção 1**



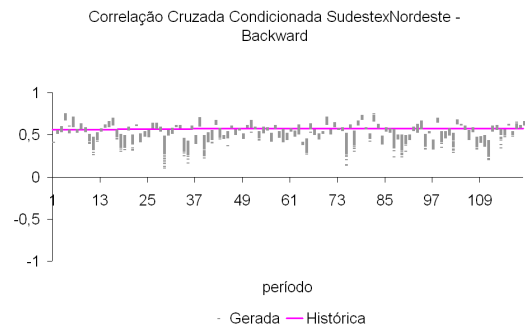
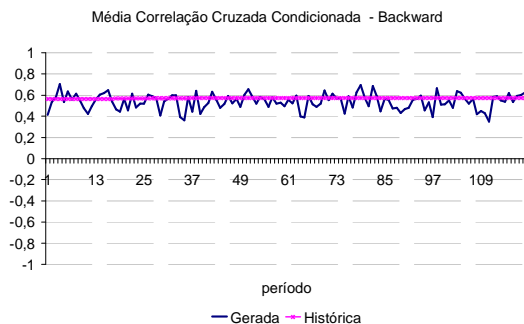
**(d) Opção 4**



**(e) Opção 1AAS**



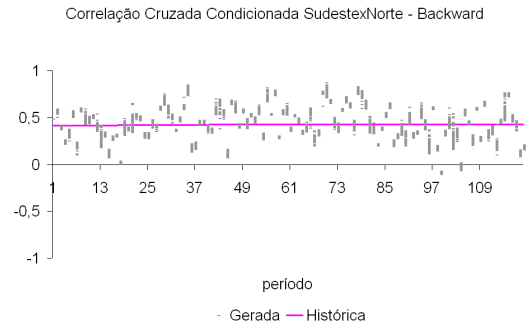
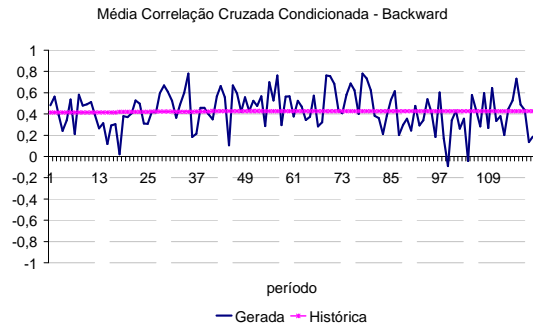
**(f) Opção LHC**



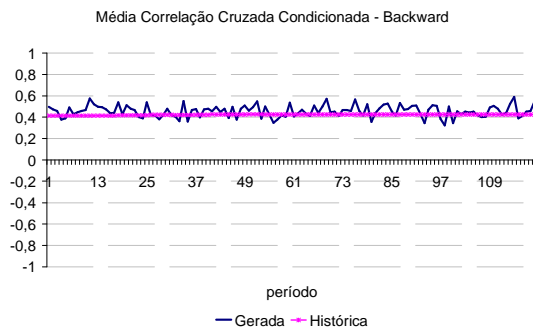
**(g) Opção AD**

Figura 190: Correlação Cruzada Cenários Backward – Sudeste x Nordeste

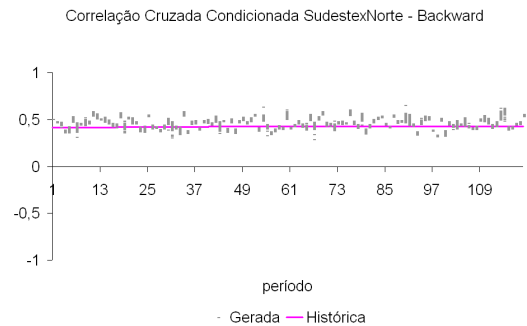
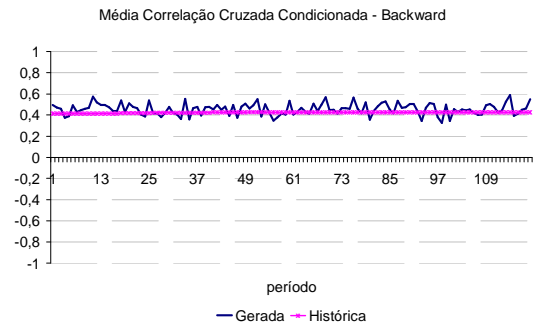
Da Figura 191a até a Figura 191g é apresentada a evolução temporal da correlação cruzada entre os subsistemas Sudeste e Norte. Conforme observado anteriormente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico. A correlação cruzada dos cenários gerados com a opção 0 fica um pouco mais dispersa em relação ao valor histórico. As correlações cruzadas obtidas com as demais opções se aproximam razoavelmente do valor observado no histórico.



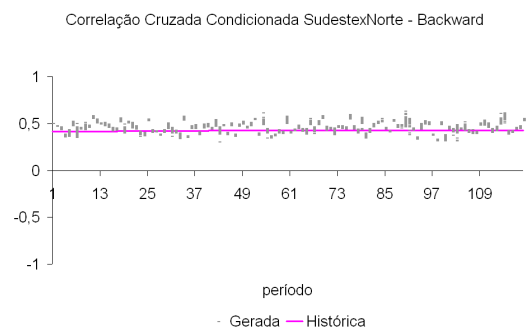
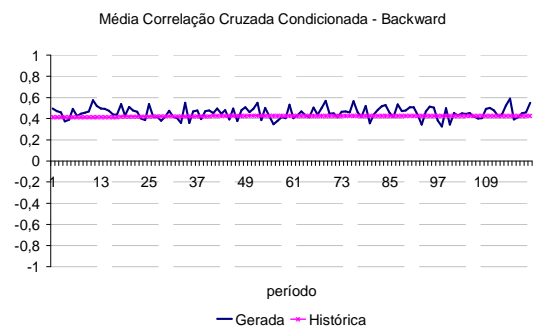
**(a) Opção Atual**



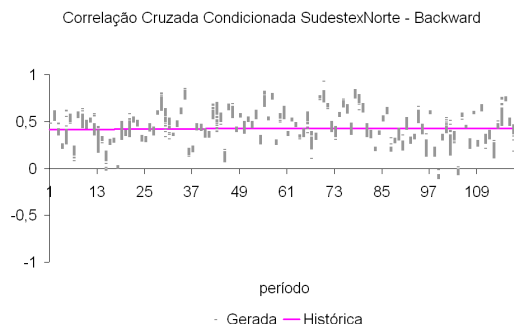
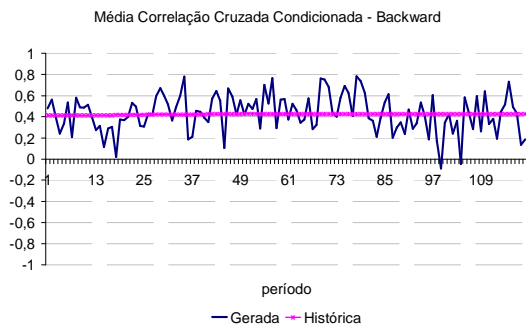
**(b) Opção 0**



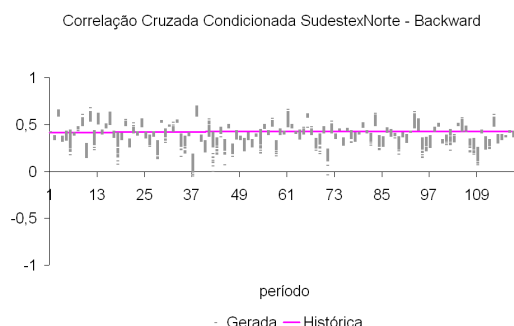
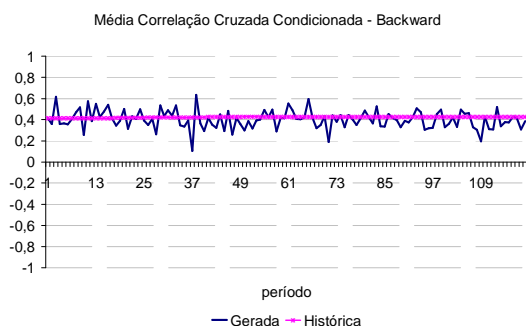
**(c) Opção 1**



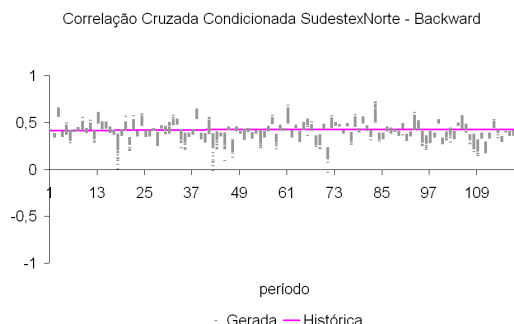
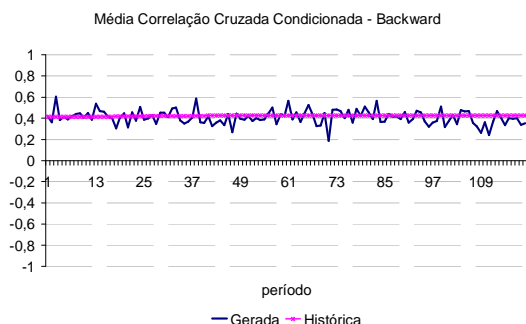
**(d) Opção 4**



**(e) Opção 1AAS**



**(f) Opção LHC**

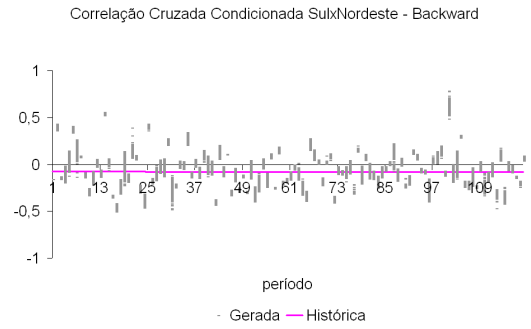
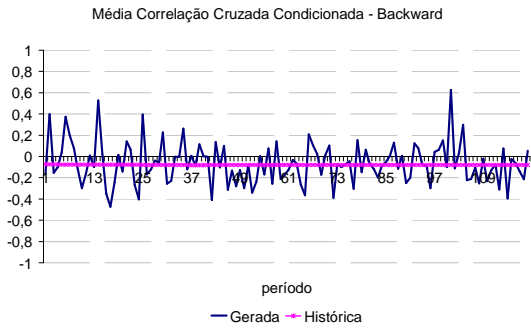


**(g) Opção AD**

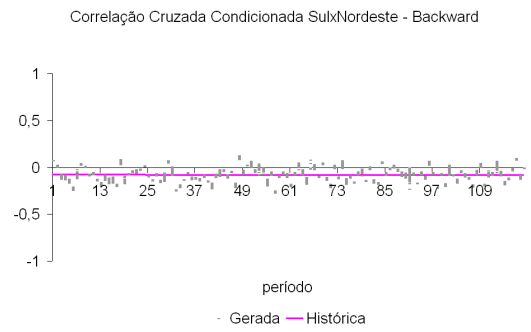
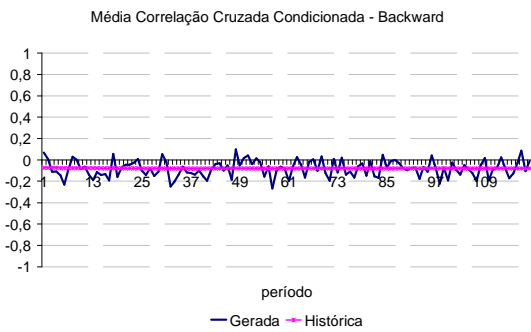
Figura 191: Correlação Cruzada Cenários Backward – Sudeste x Norte

Da Figura 192a até a Figura 192g pode-se analisar o comportamento da correlação cruzada entre os subsistemas Sul e Nordeste. As correlações cruzadas obtidas com as opções LHC e AD continuam, em geral, abaixo do valor histórico em valores absolutos. Conforme observado anteriormente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico.

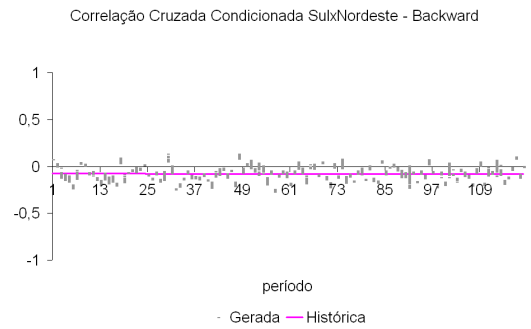
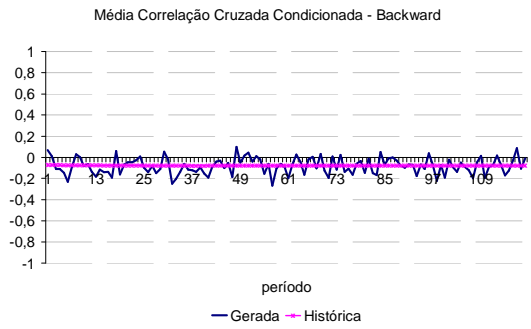




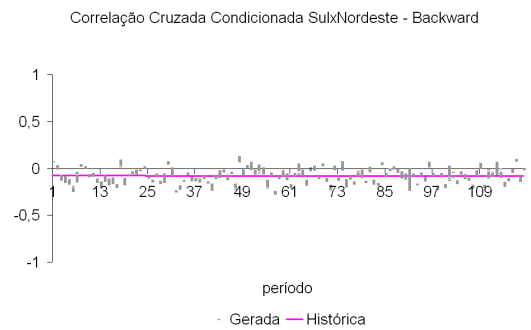
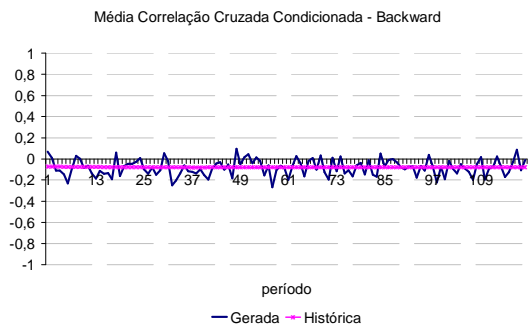
**(a) Opção Atual**



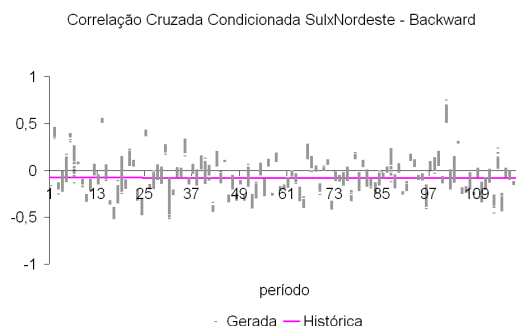
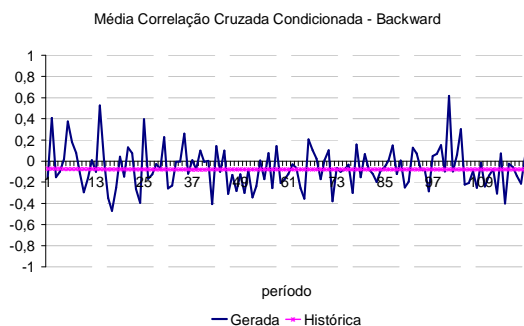
**(b) Opção 0**



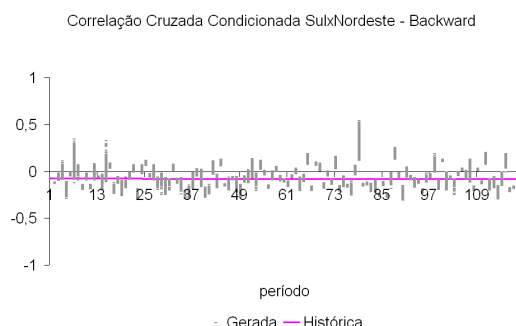
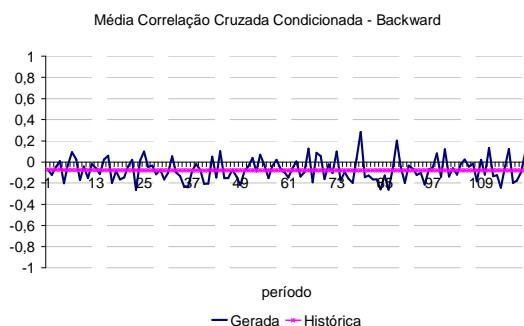
**(c) Opção 1**



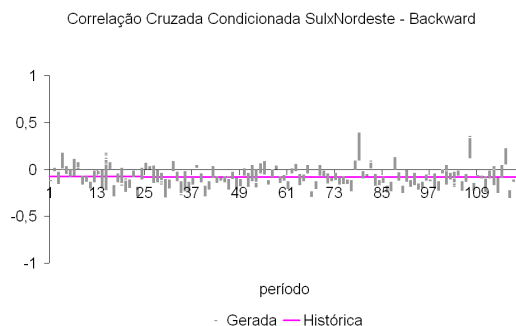
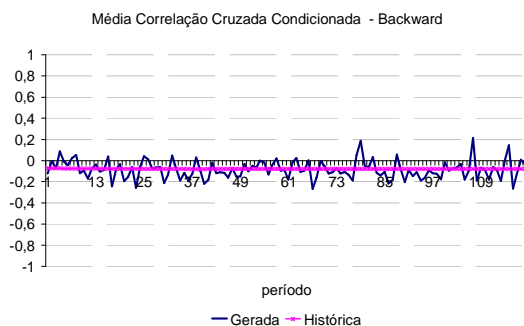
**(d) Opção 4**



**(e) Opção 1AAS**



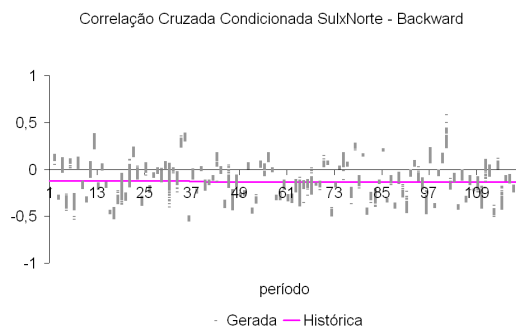
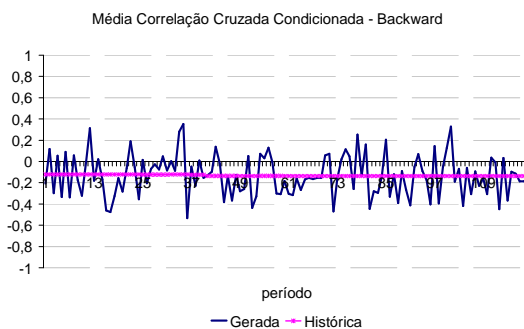
**(f) Opção LHC**



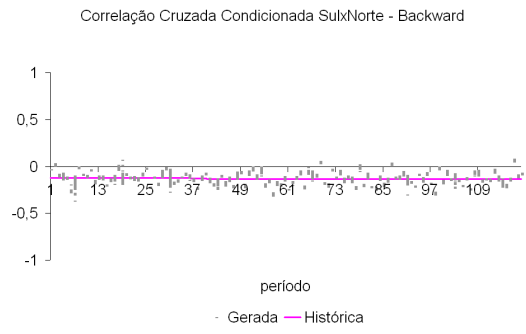
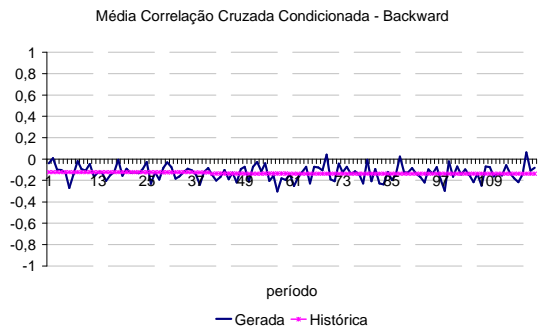
**(g) Opção AD**

Figura 192: Correlação Cruzada Cenários Backward – Sul x Nordeste

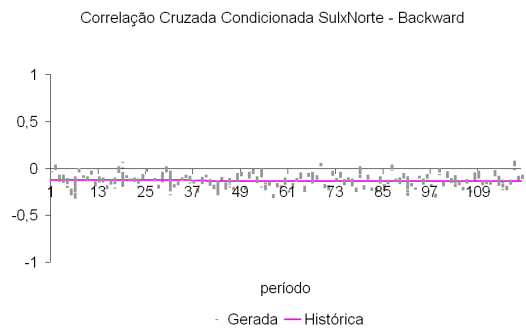
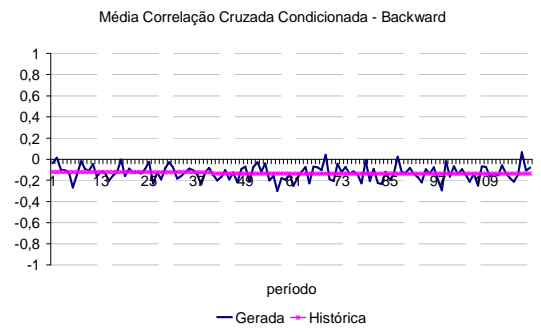
As mesmas observações feitas na análise da correlação cruzada entre Sul e Nordeste valem para a correlação cruzada entre Sul e Norte, Figura 193a até Figura 193g.



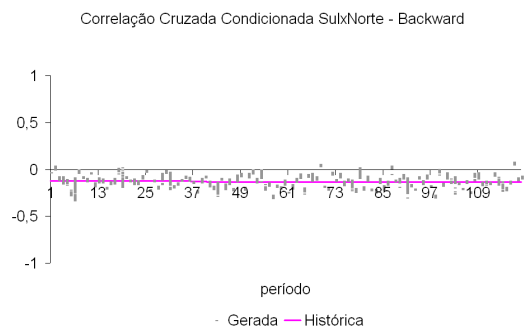
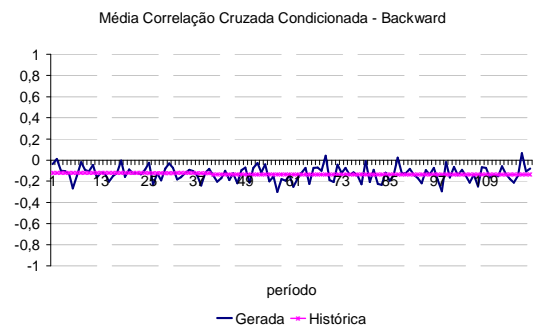
**(a) Opção Atual**



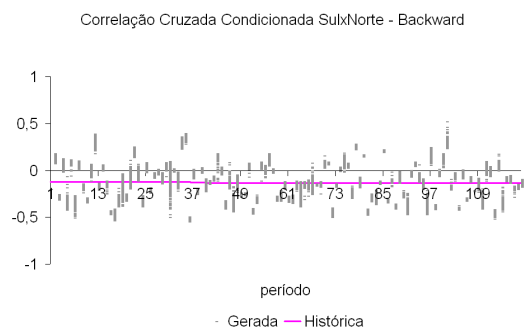
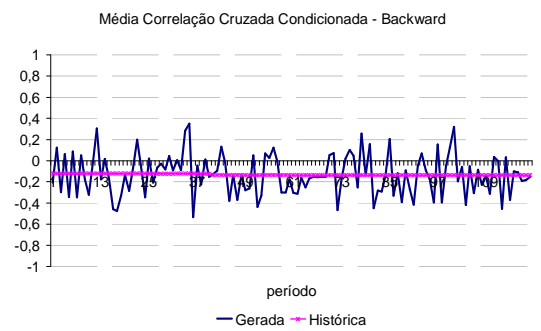
**(b) Opção 0**



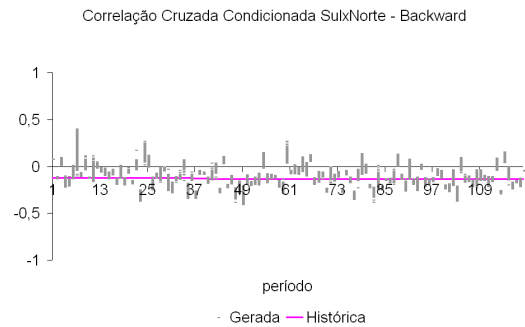
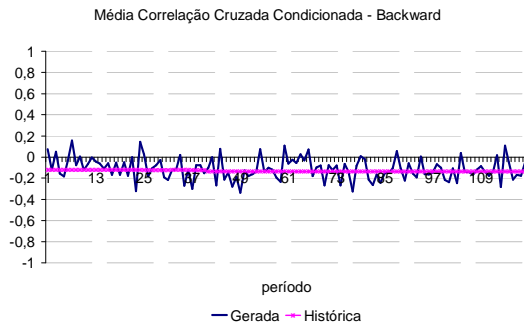
**(c) Opção 1**



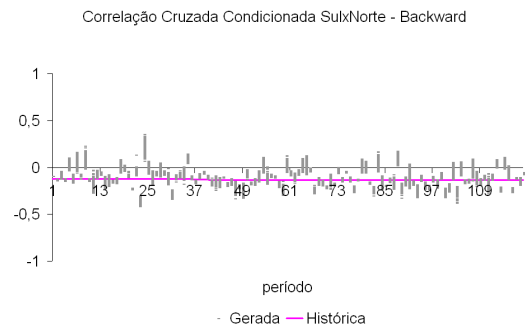
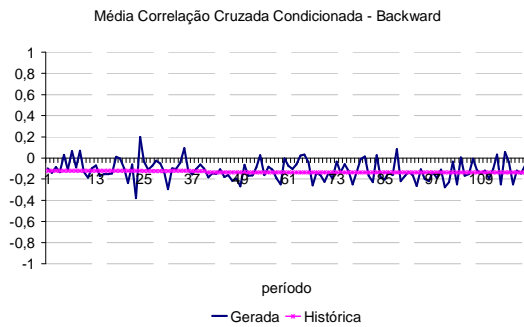
**(d) Opção 4**



**(e) Opção 1AAS**



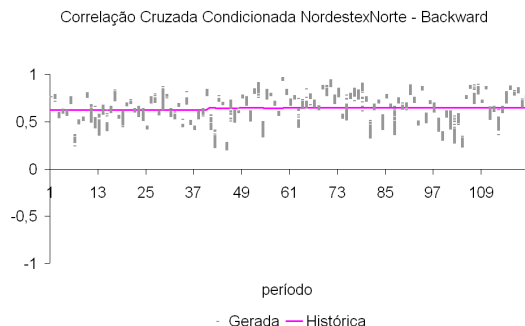
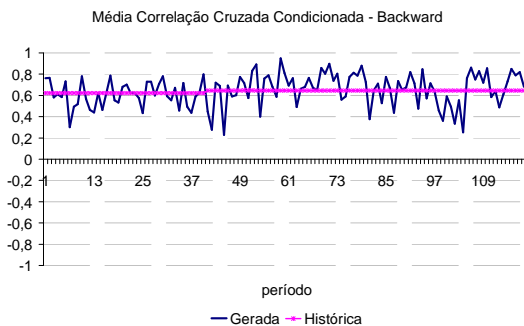
**(f) Opção LHC**



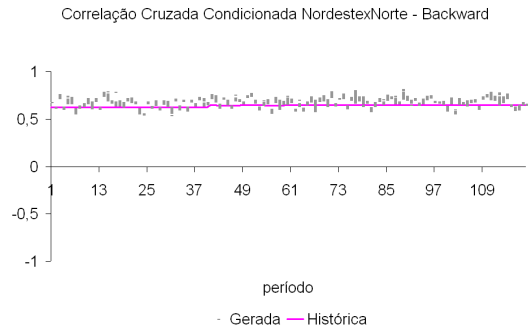
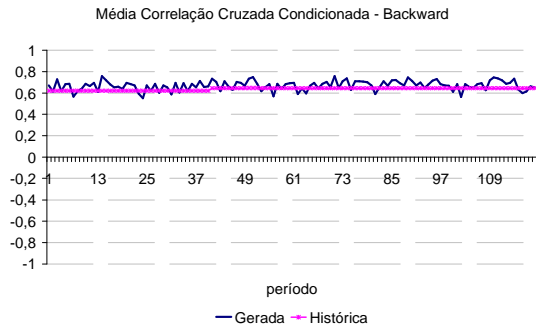
**(g) Opção AD**

Figura 193: Correlação Cruzada Cenários Backward –Sul x Norte

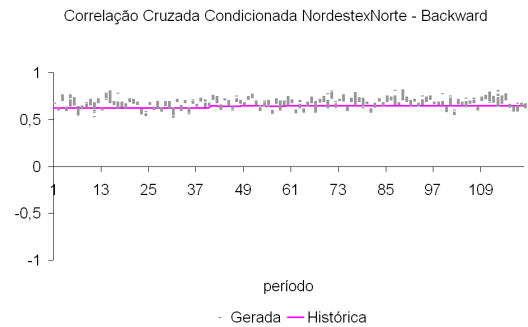
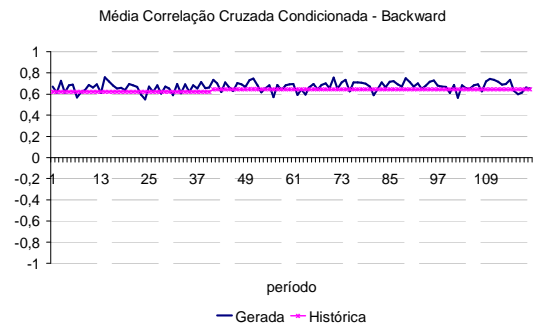
Da Figura 194a até a Figura 194g é comparada evolução da correlação cruzada entre os subsistemas Nordeste e Norte com o valor observado no histórico. As correlações cruzadas obtidas com as opções LHC e AD permanecem, em geral, abaixo do valor histórico, porém as diferenças relação ao histórico são bem menores. Novamente, a opção Atual apresenta uma variabilidade acentuada em relação ao valor histórico.



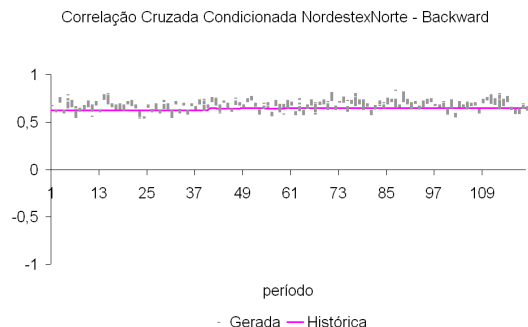
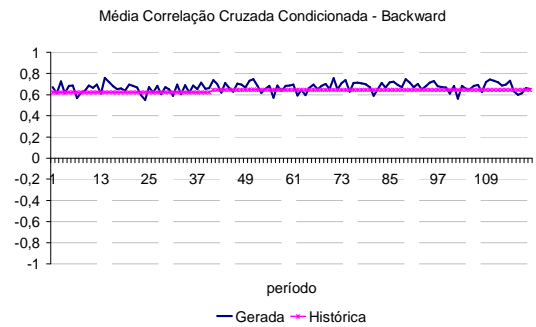
**(a) Opção Atual**



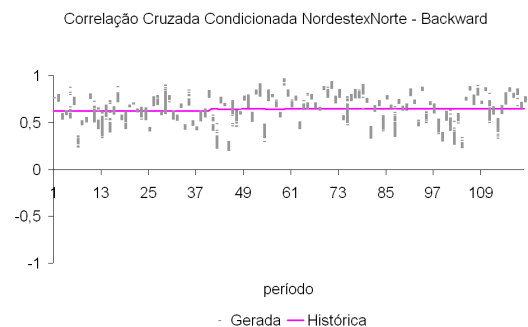
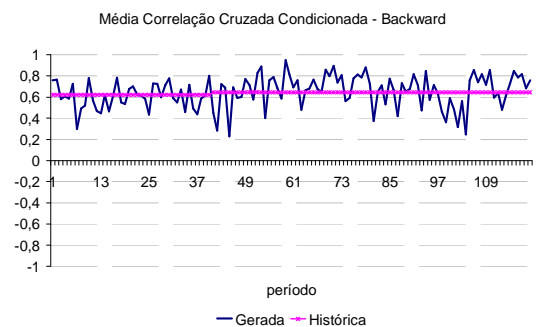
**(b) Opção 0**



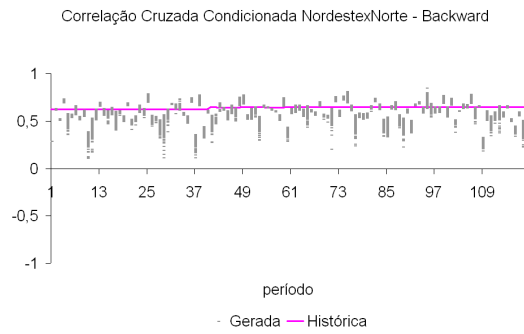
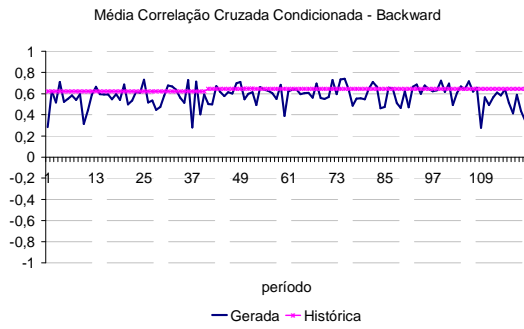
**(c) Opção 1**



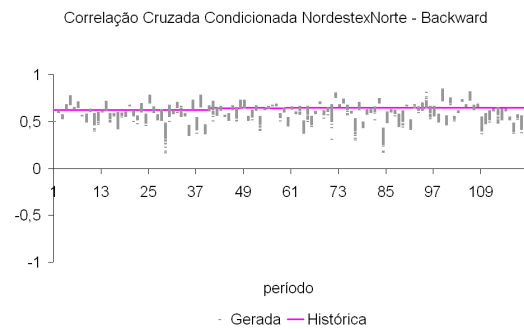
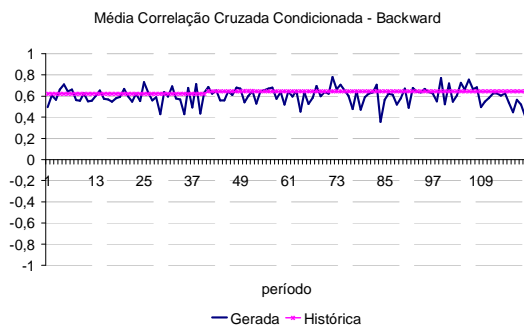
**(d) Opção 4**



**(e) Opção 1AAS**



**(f) Opção LHC**



**(g) Opção AD**

Figura 194: Correlação Cruzada Cenários Backward – Nordeste x Norte

## Apêndice C Avaliação dos Resultados do Problema de Planejamento da Operação (Complementação)

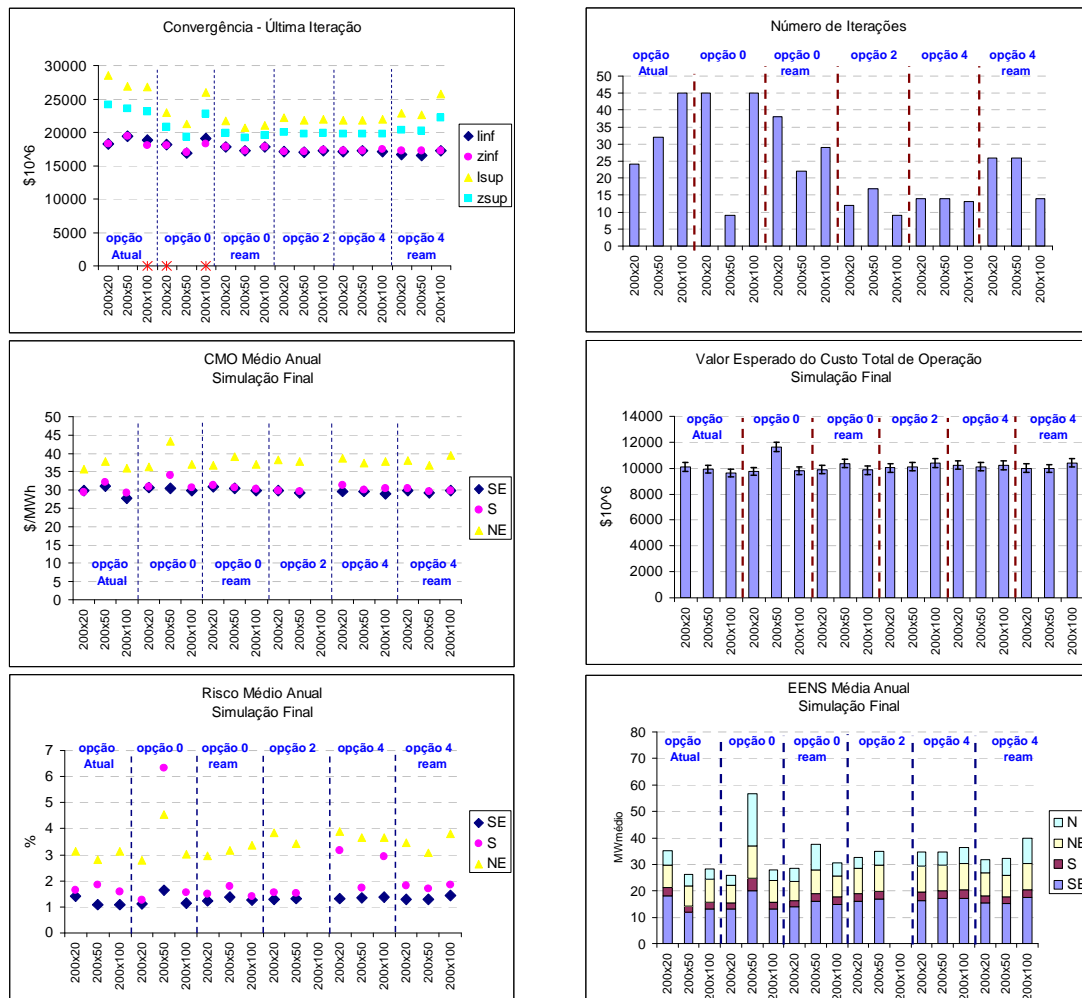


Figura 195: Resultados- Variação Backward - PMO Mai/04

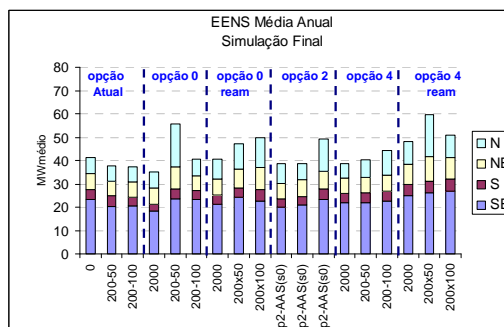
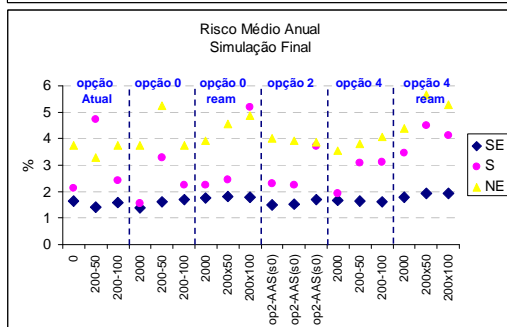
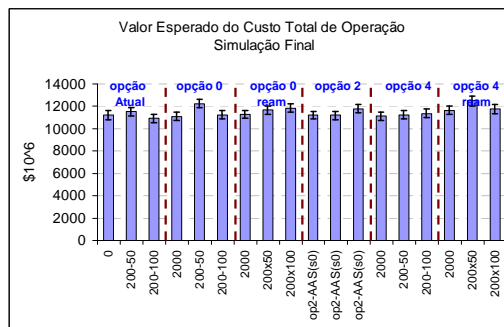
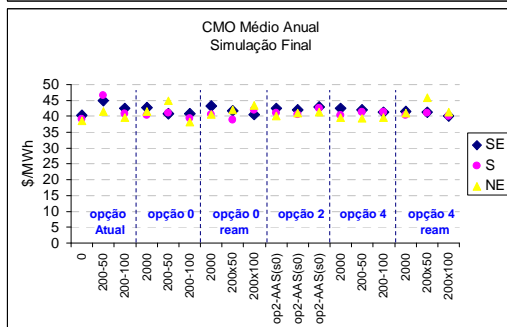
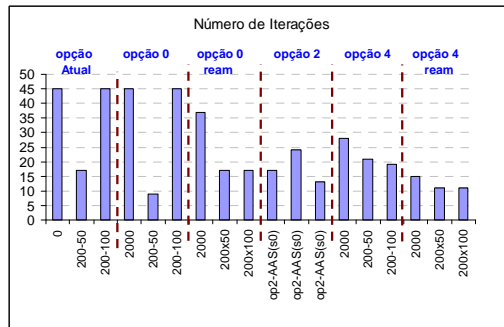
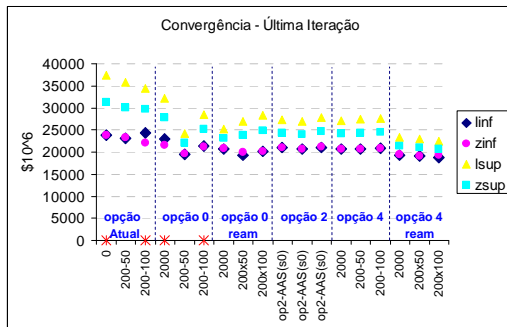
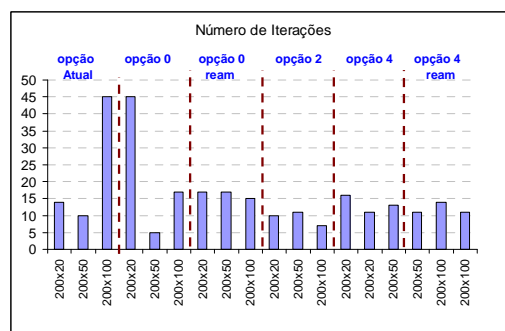
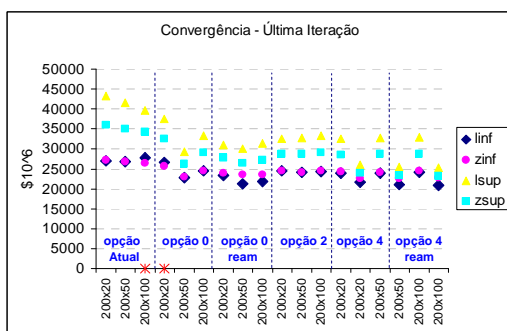


Figura 196: Resultados- Variação Backward - PMO Set/04





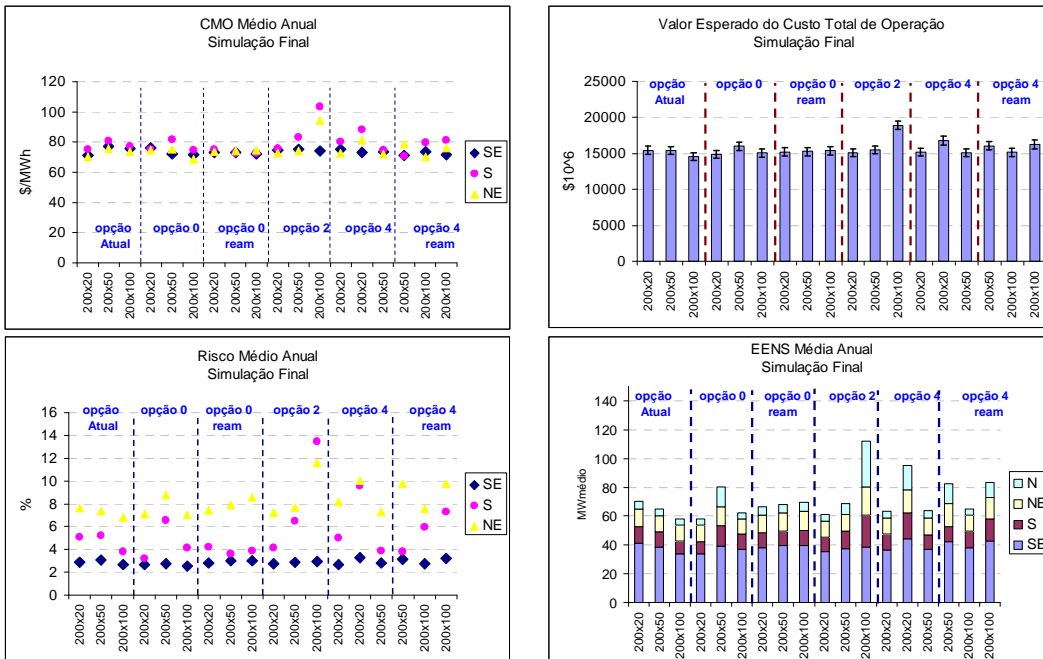


Figura 197: Resultados- Variação Backward - PMO Set/09

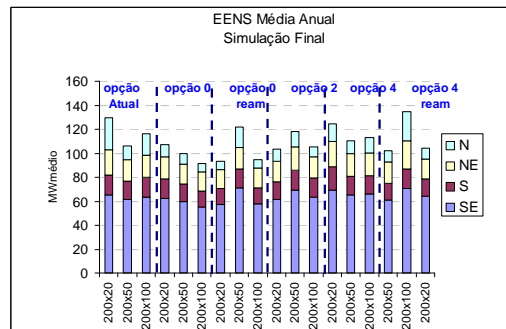
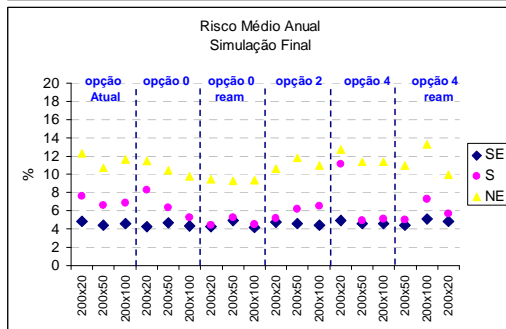
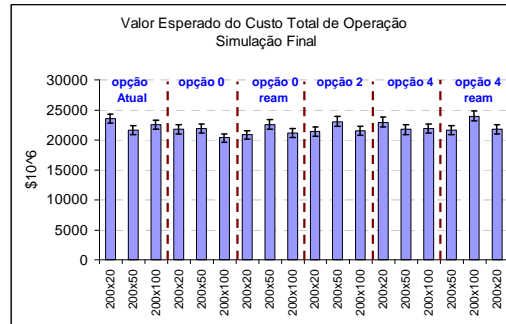
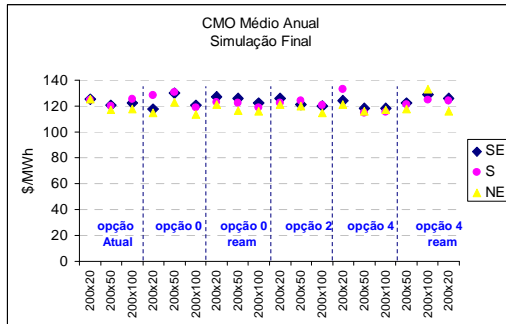
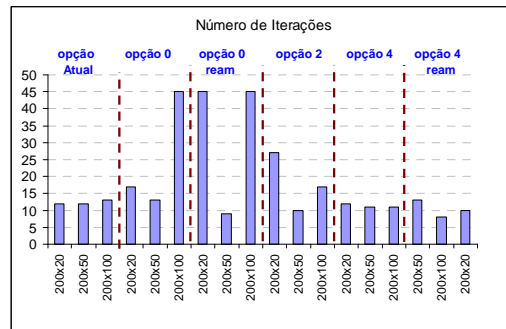
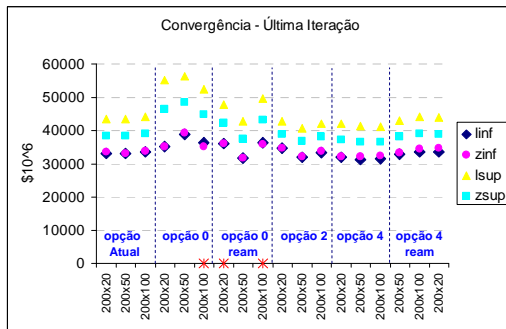
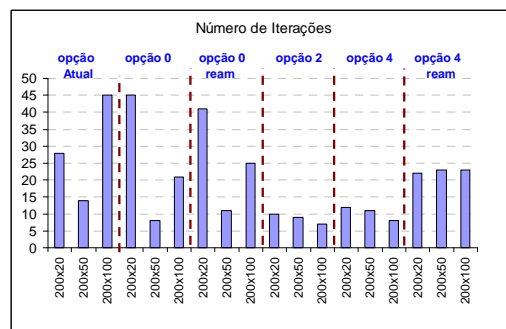
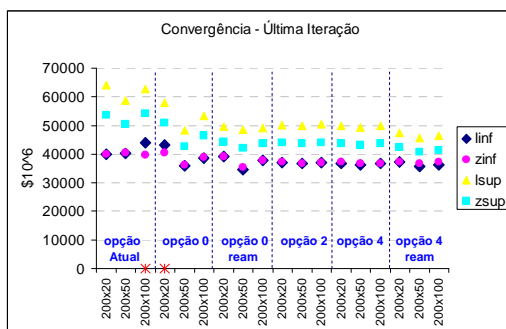


Figura 198: Resultados- Variação Backward - PMO Jun/06



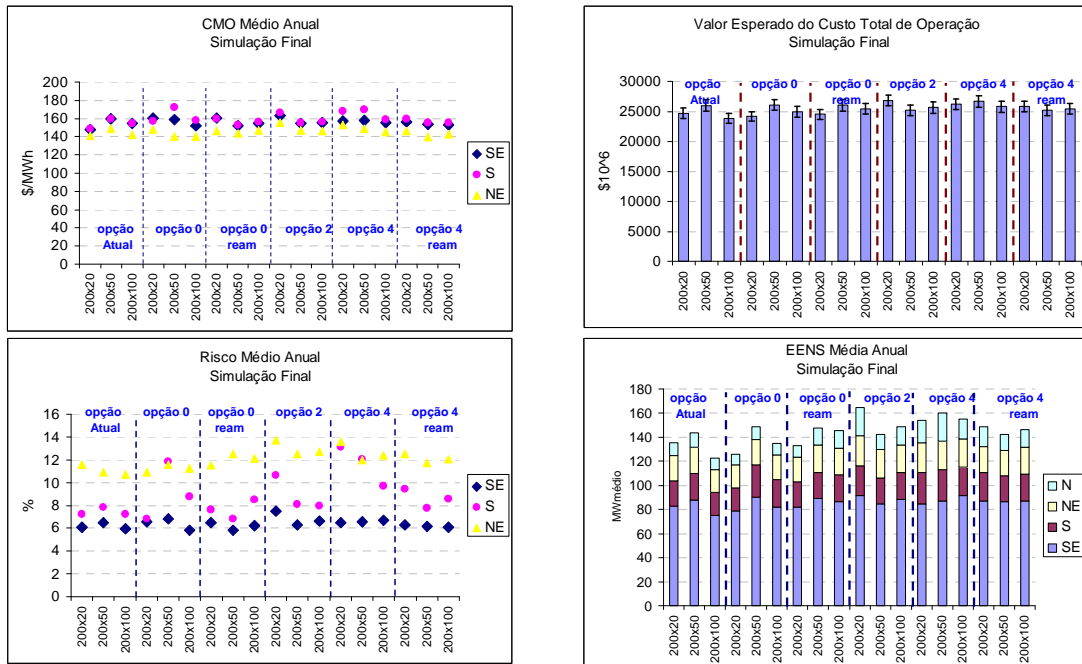
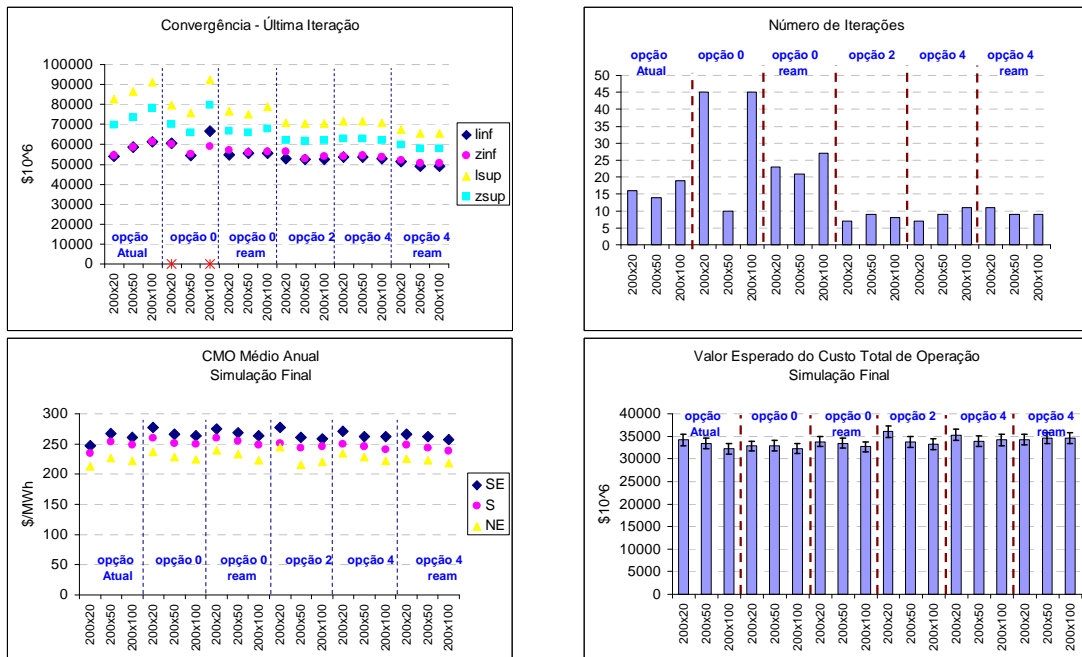


Figura 199: Resultados- Variação Backward - PMO Jul/06



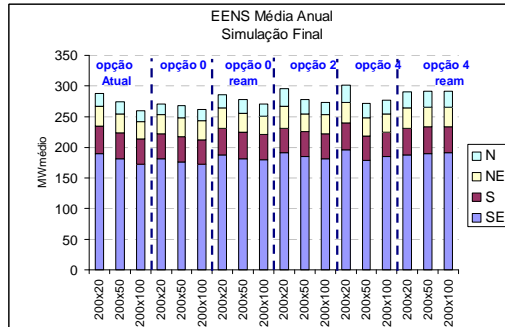
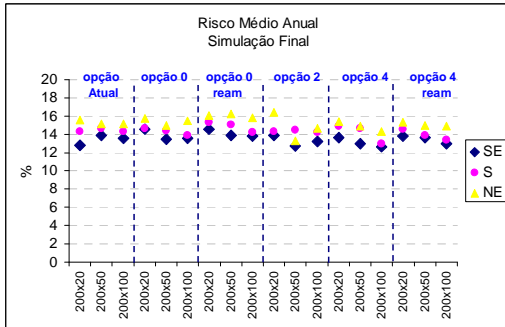


Figura 200: Resultados – Variação Backward - PMO Jan/07

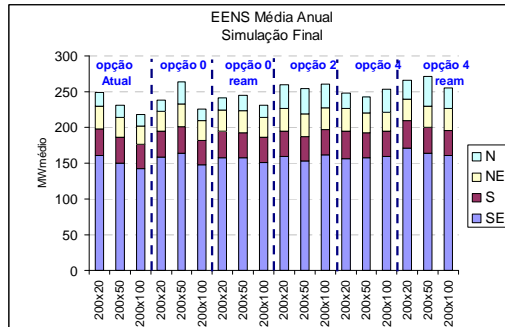
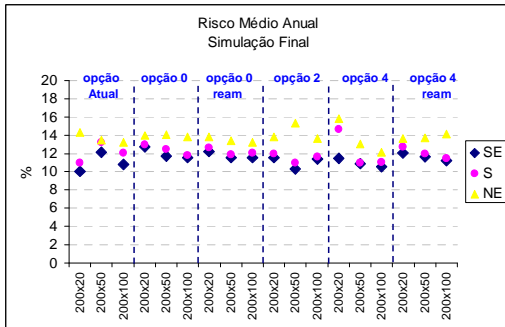
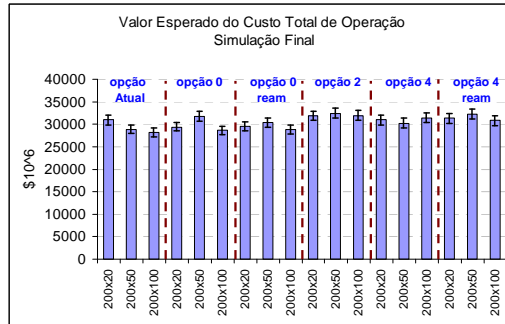
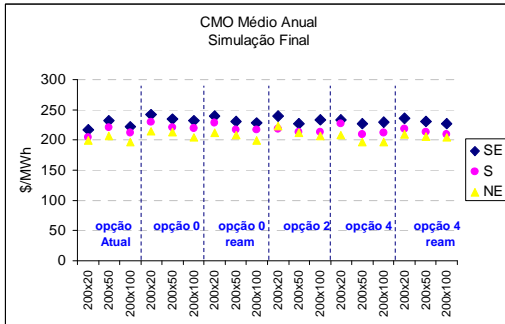
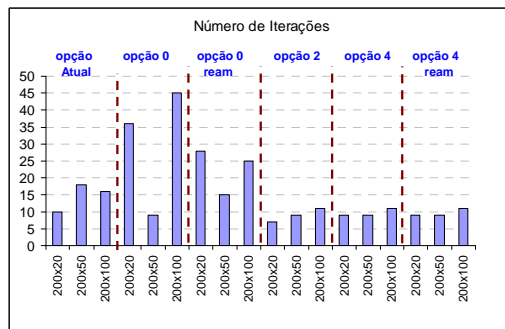
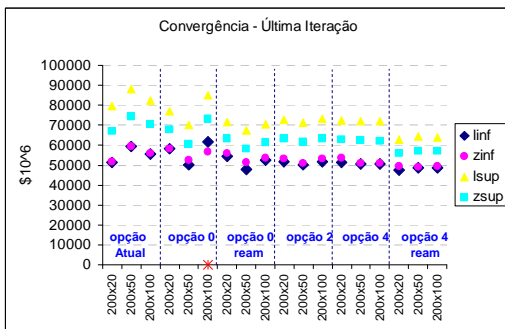


Figura 201: Resultados – Variação Backward - PMO Fev/07

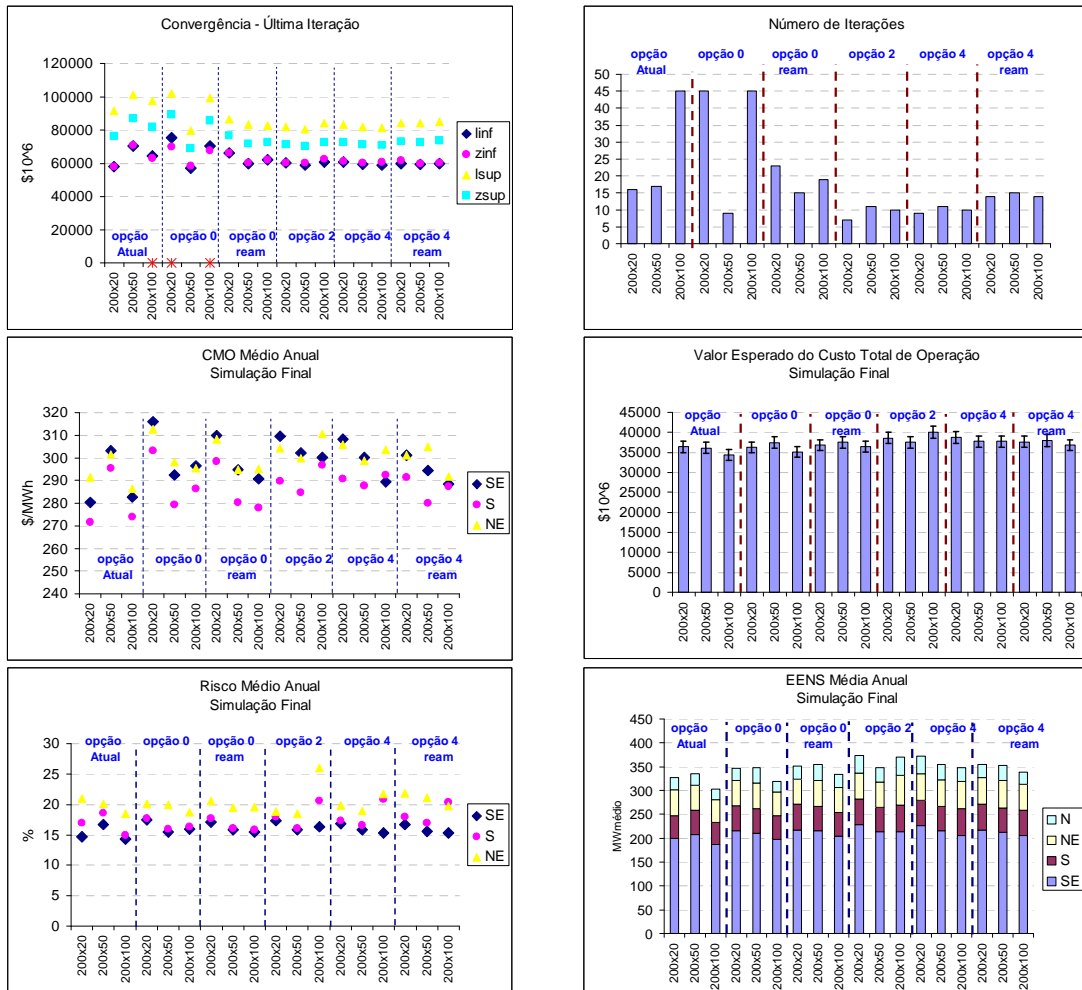
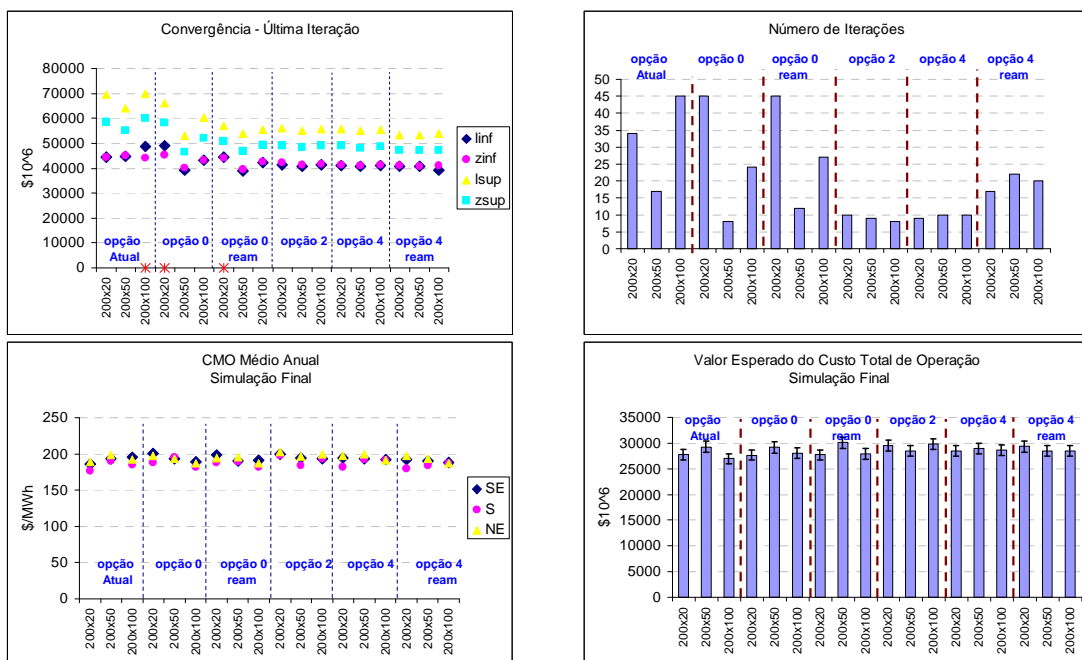


Figura 202: Resultados – Variação Backward - PMO Mai/07



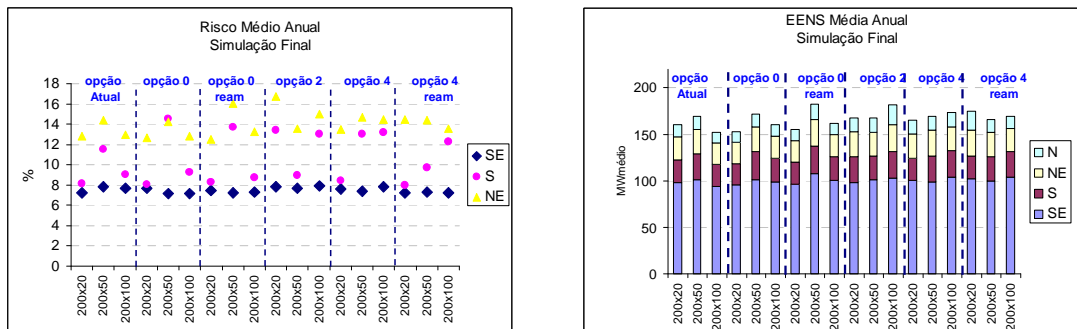


Figura 203: Resultados- Variação Backward - PMO Jul/07

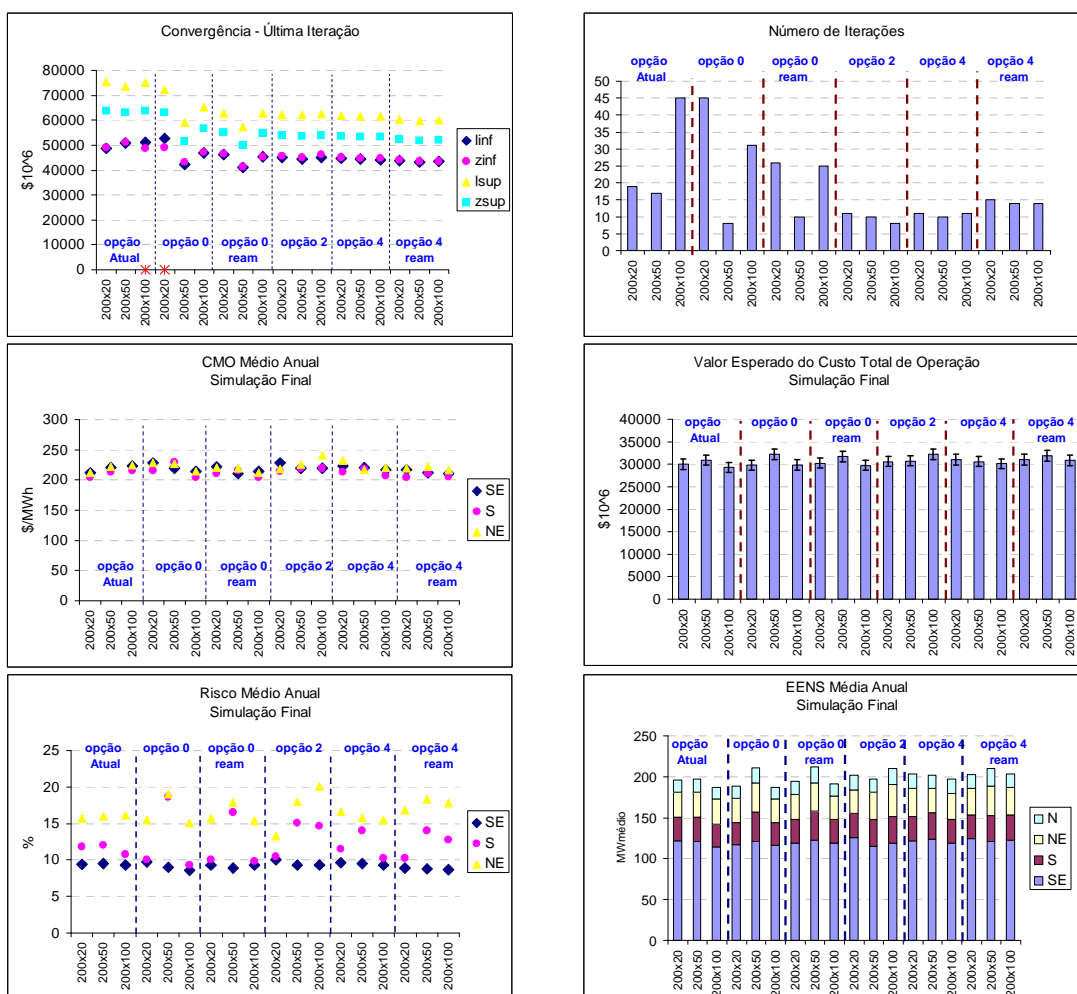


Figura 204: Resultados- Variação Backward - PMO Ago/07

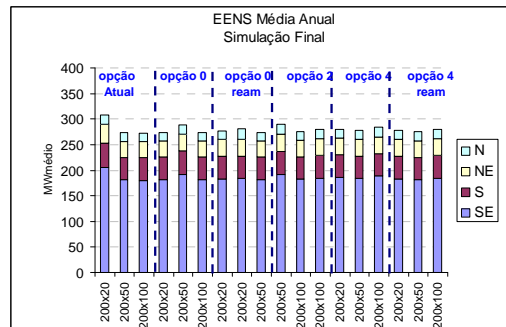
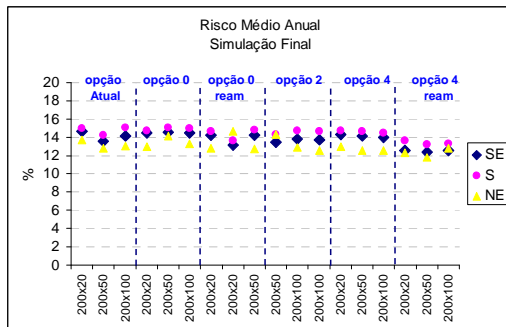
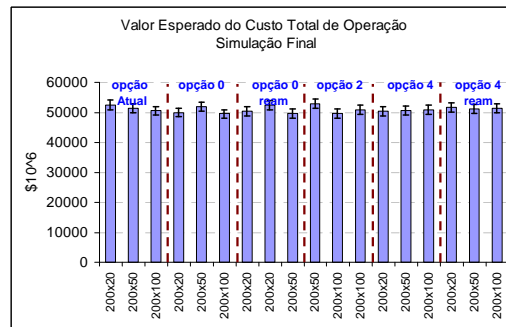
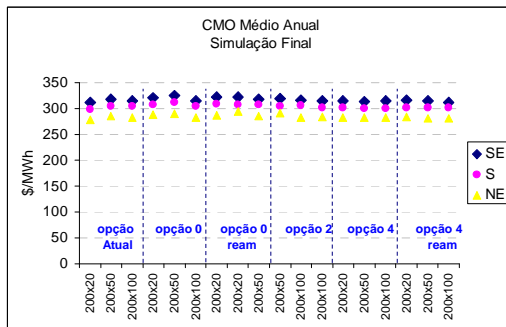
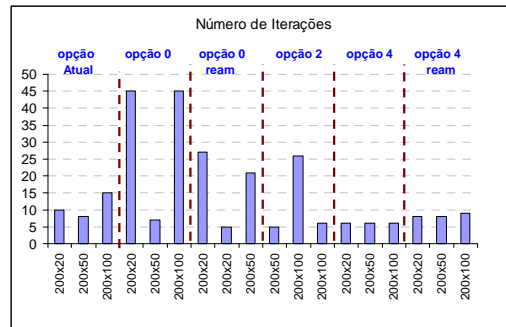
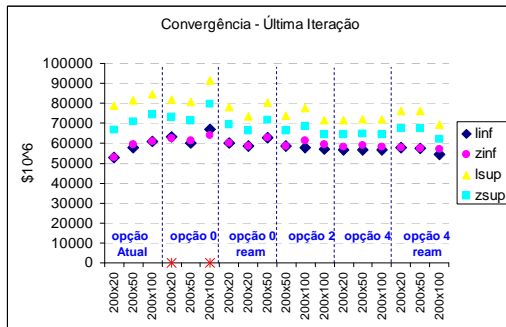
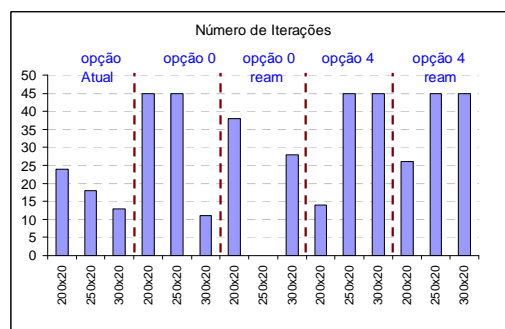
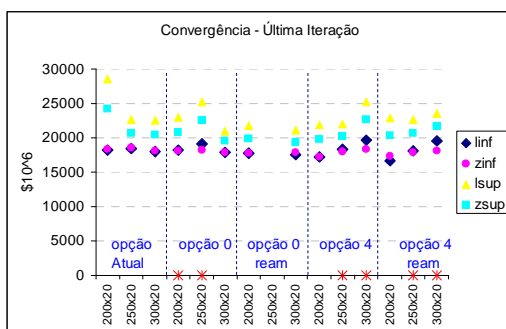


Figura 205: Resultados - Variação Backward - PMO Jan/08



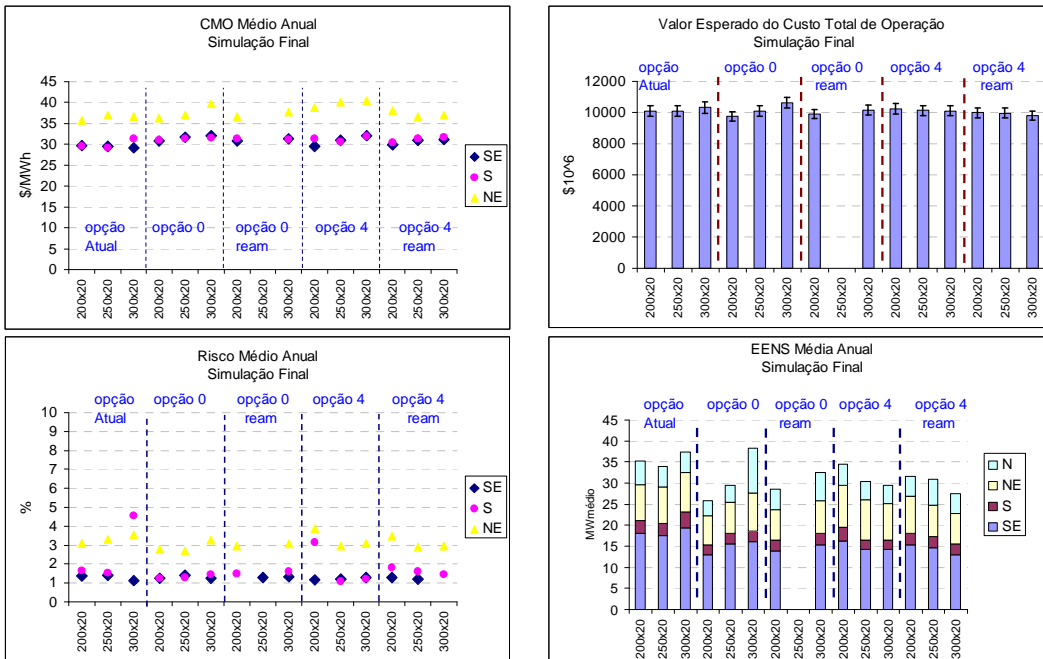
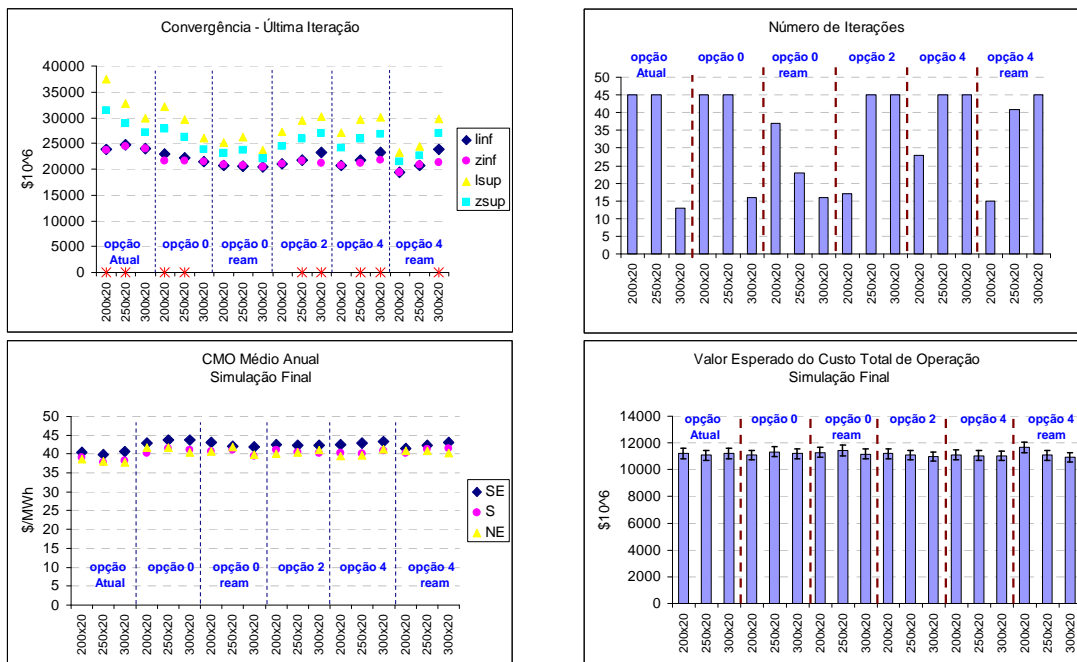


Figura 206: Resultados – Variação Forward - PMO Mai/04





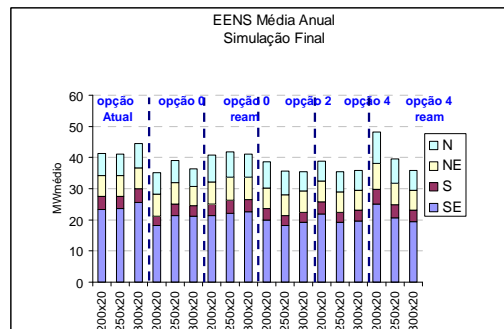
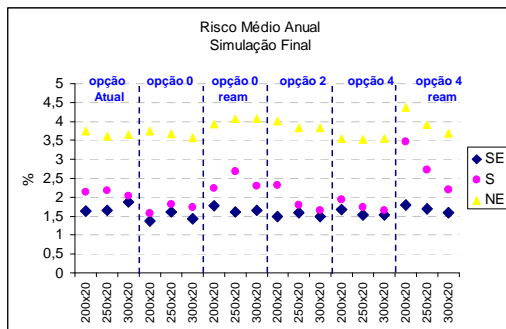


Figura 207: Resultados – Variação Forward - PMO Set/04

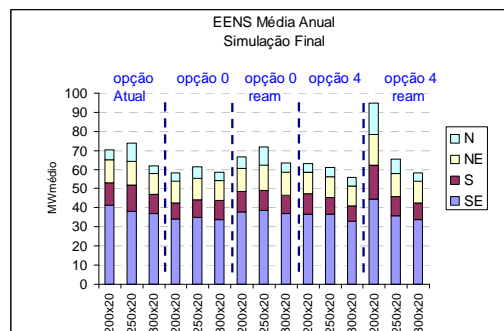
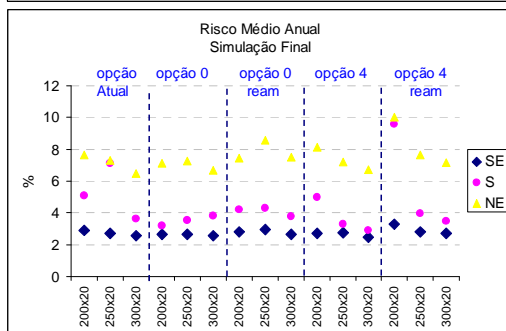
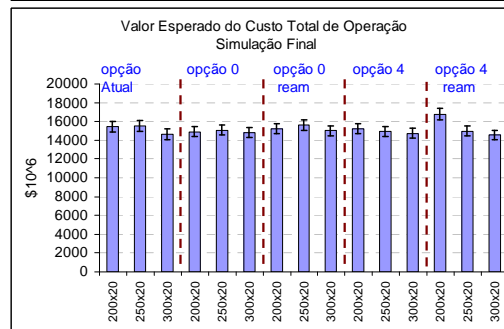
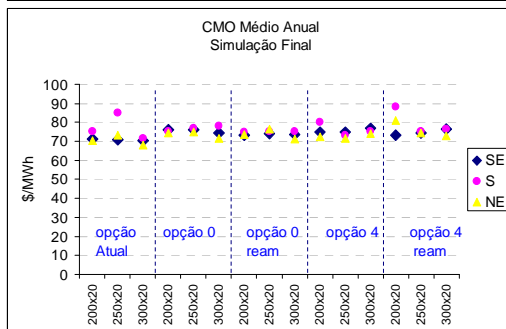
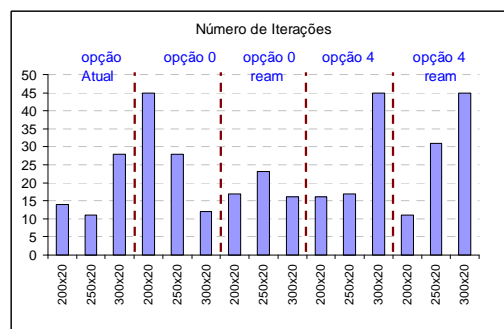
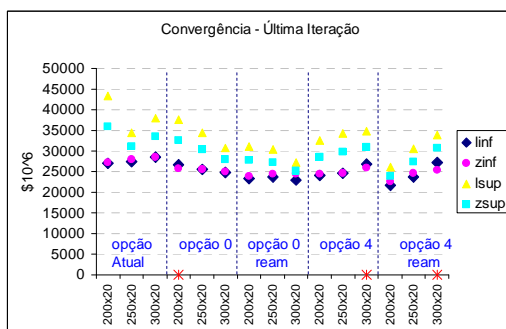


Figura 208: Resultados – Variação Forward - PMO Set/09

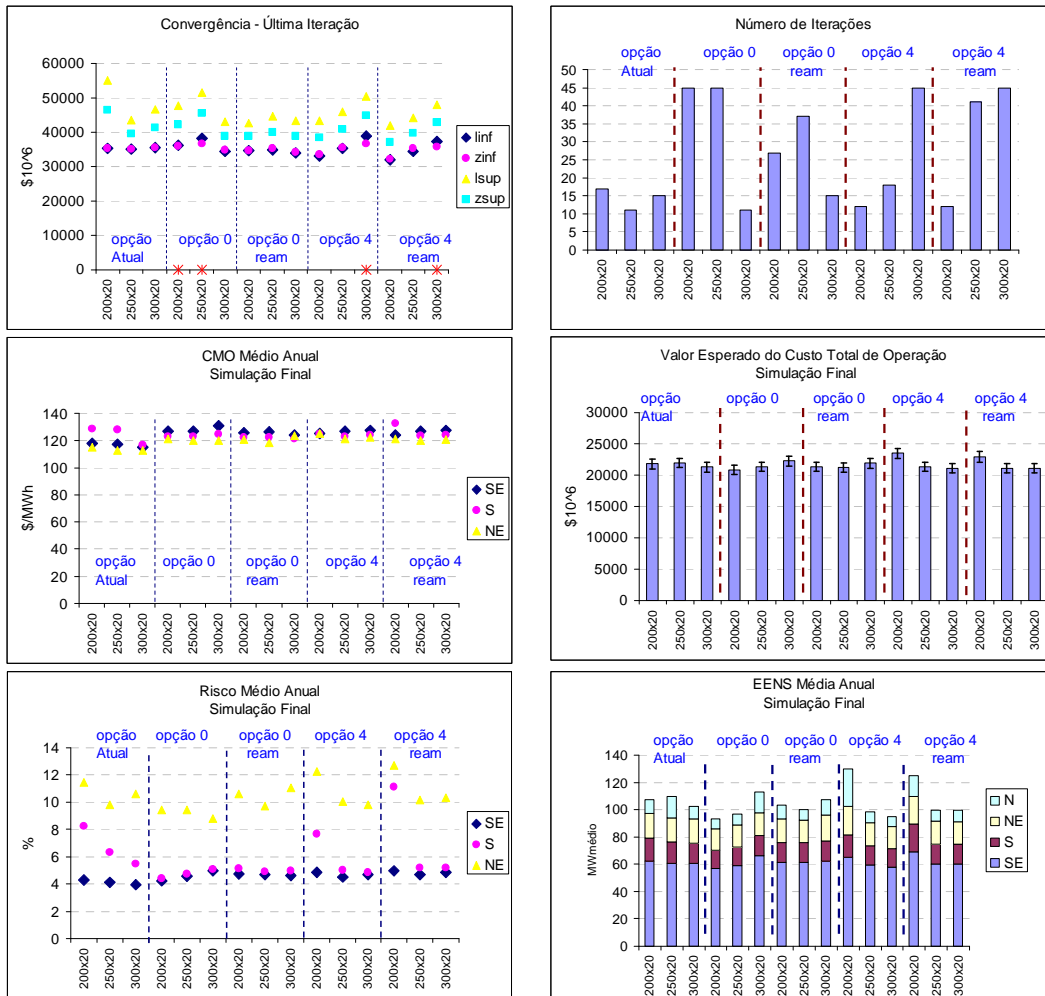
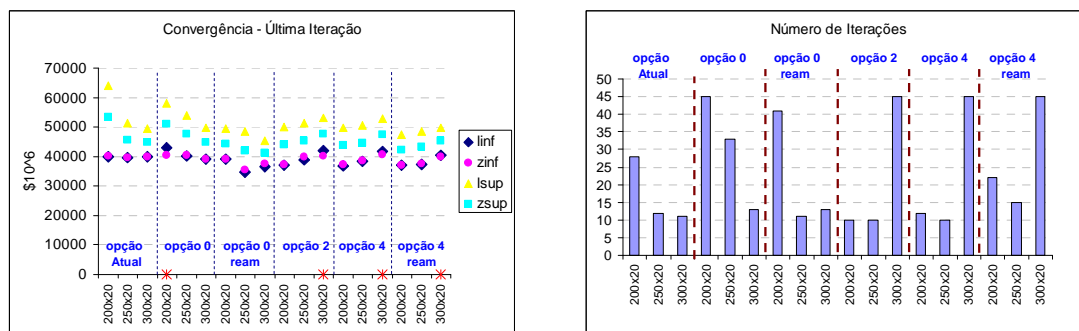


Figura 209: Resultados- Variação Forward - PMO Jun/06



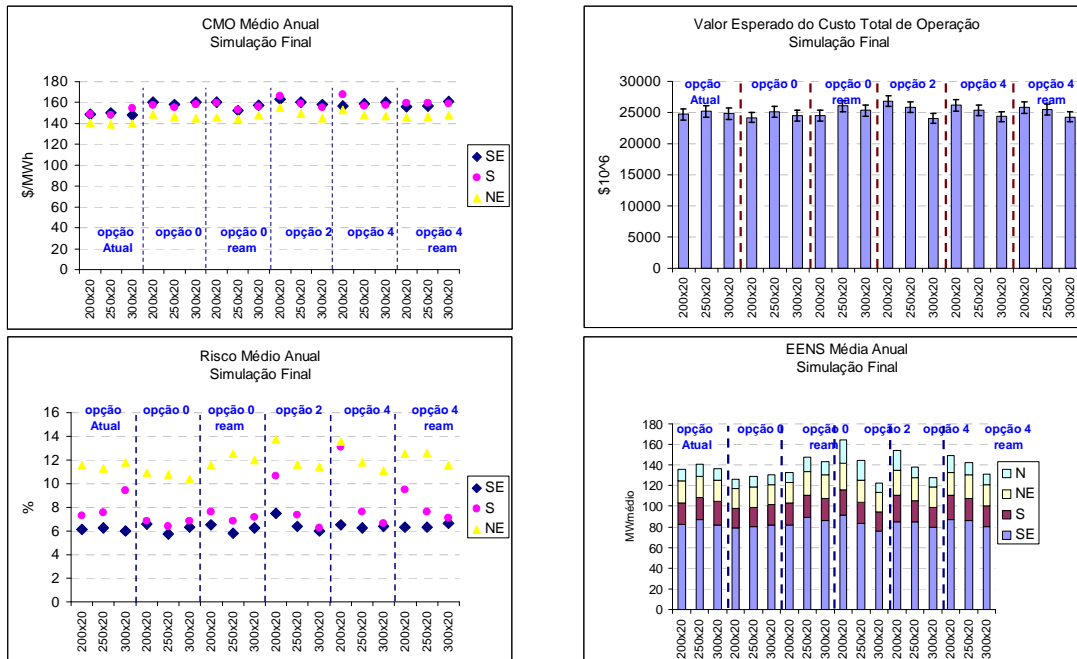
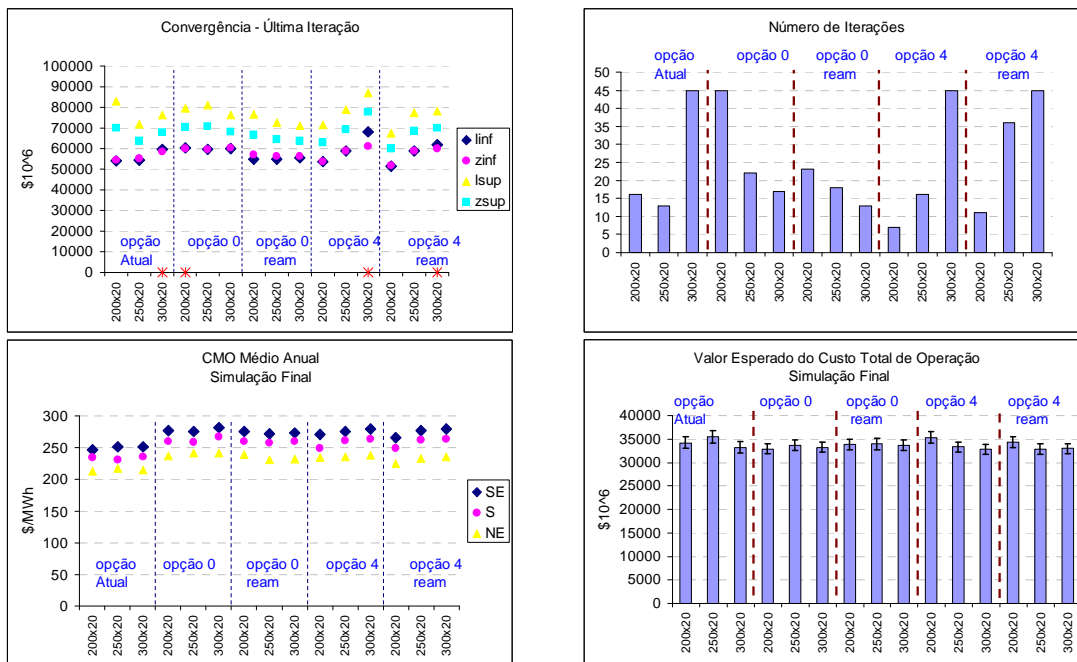


Figura 210: Resultados- Variação Forward - PMO Jul/06



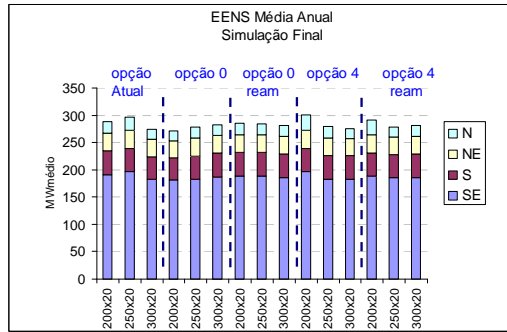
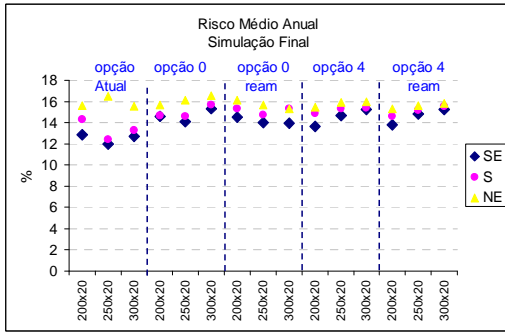


Figura 211: Resultados- Variação Forward - PMO Jan/07

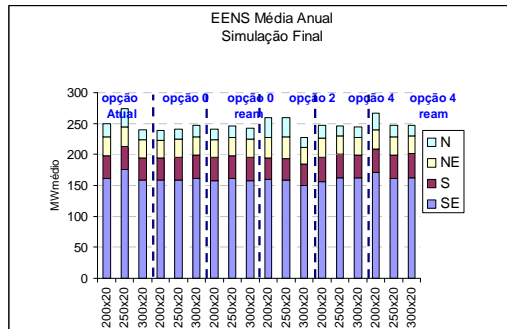
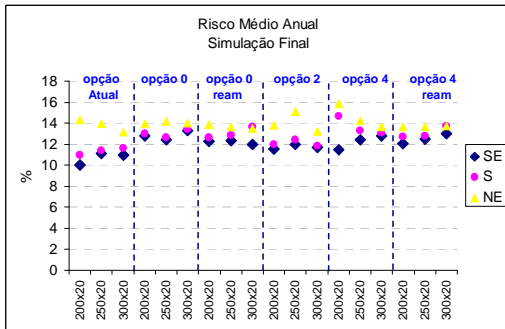
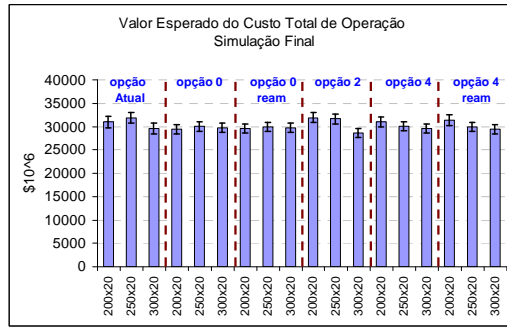
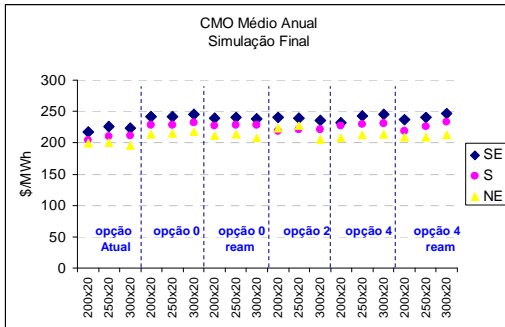
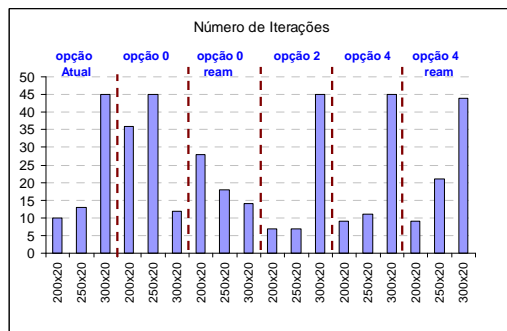
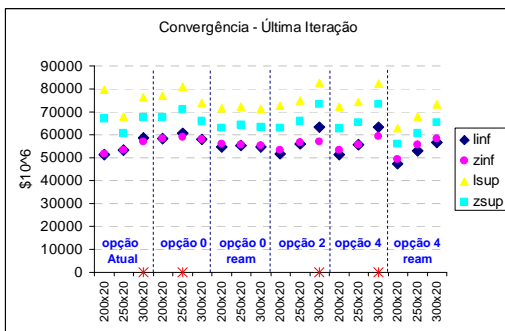
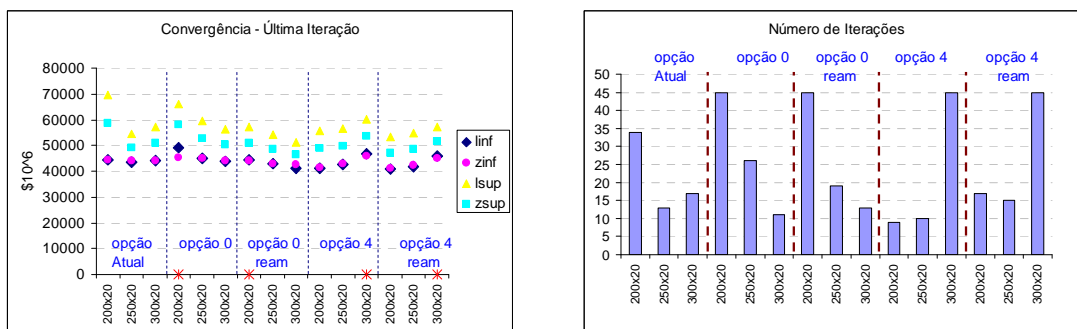


Figura 212: Resultados- Variação Forward - PMO Feb/07



Figura 213: Resultados- Variação Forward - PMO Mai/07



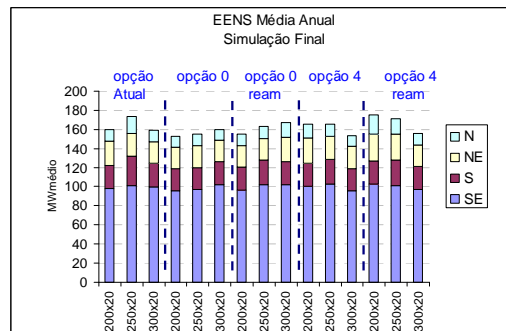
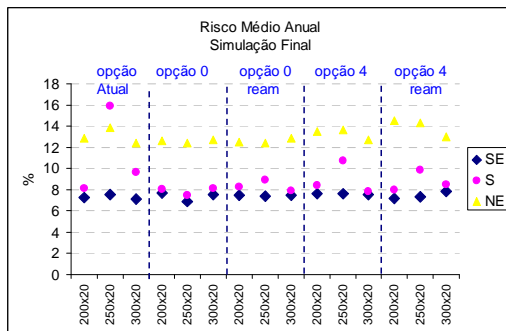
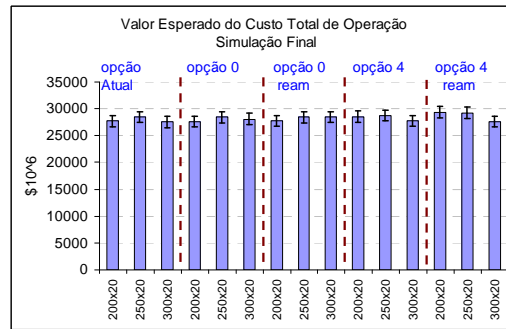
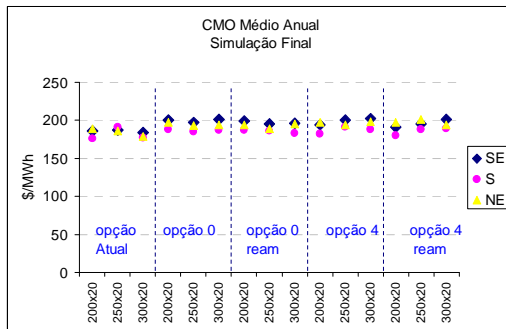
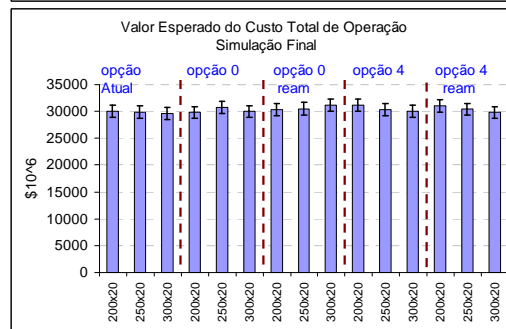
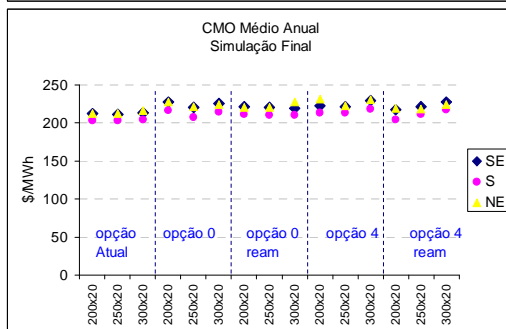
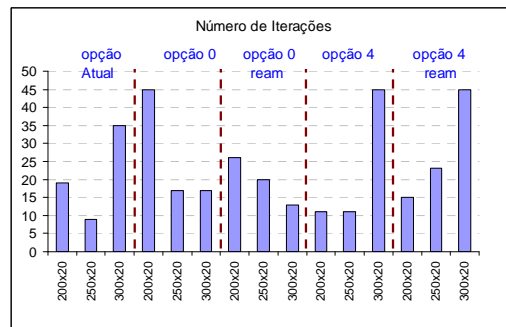
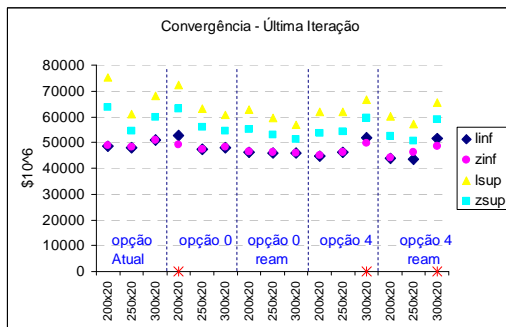


Figura 214: Resultados – Variação Forward - PMO Jul/07



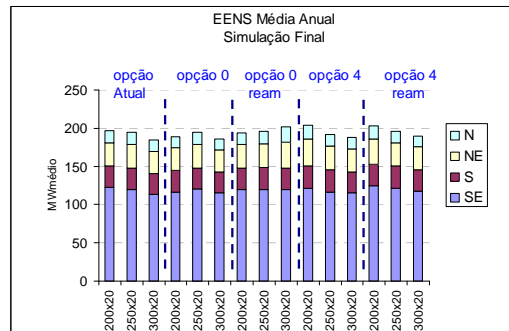
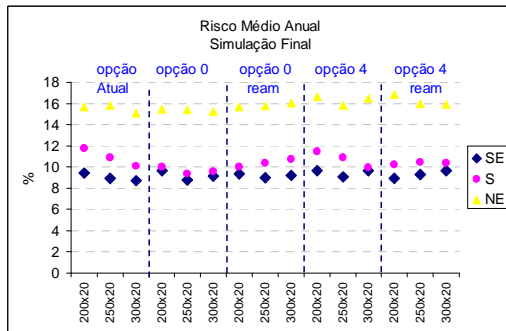


Figura 215: Resultados– Variação Forward - PMO Ago/07

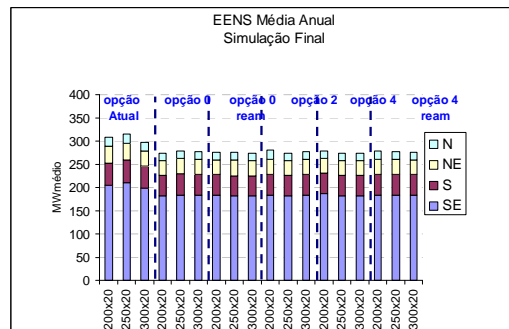
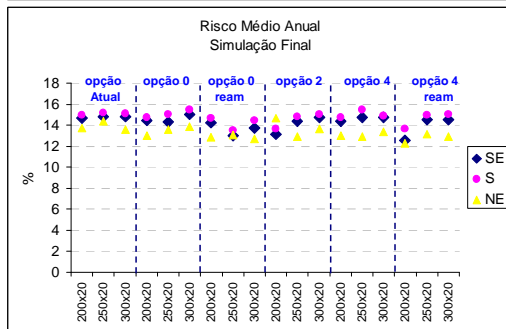
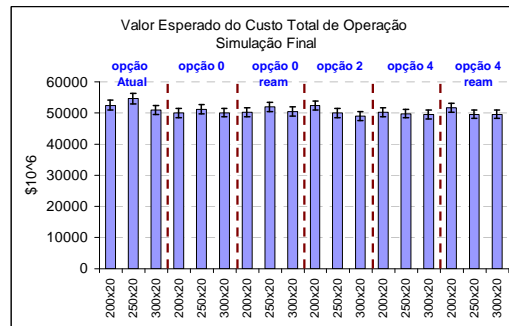
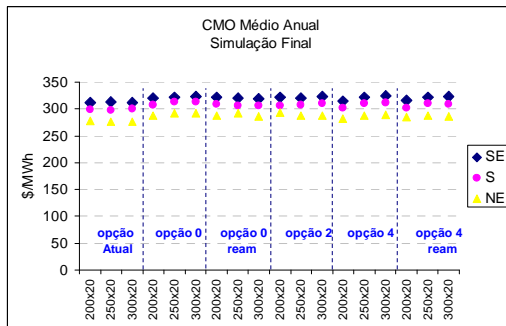
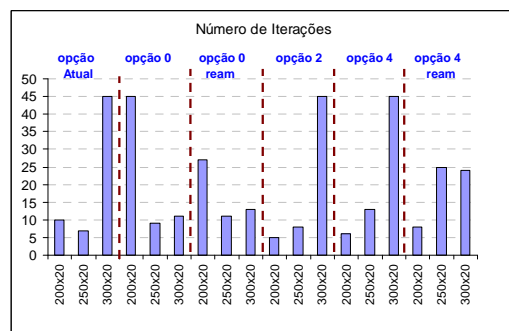
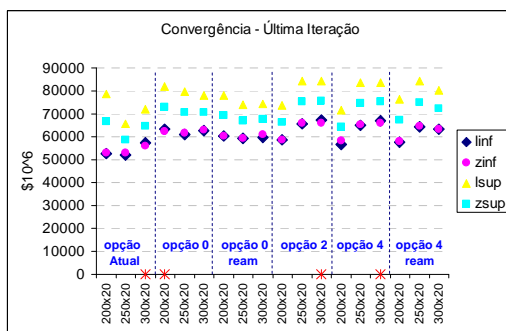


Figura 216: Resultados– Variação Forward - PMO Jan/08

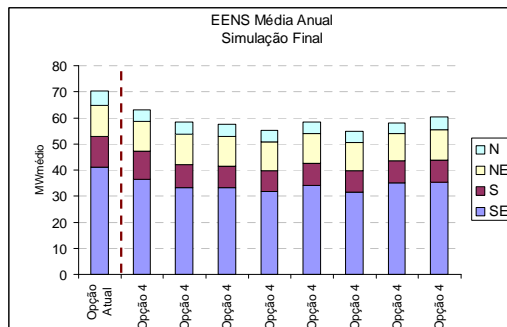
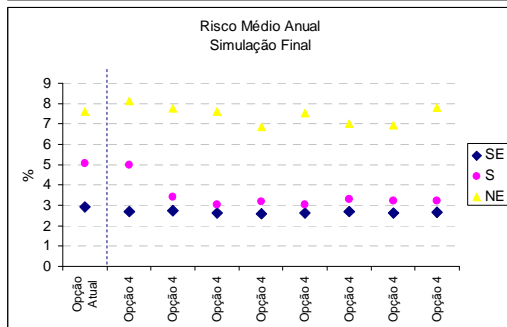
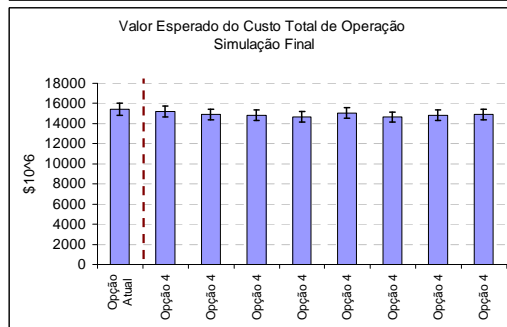
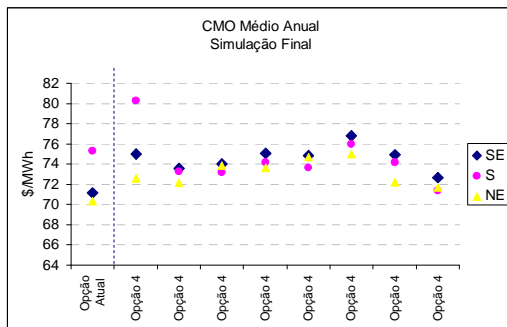
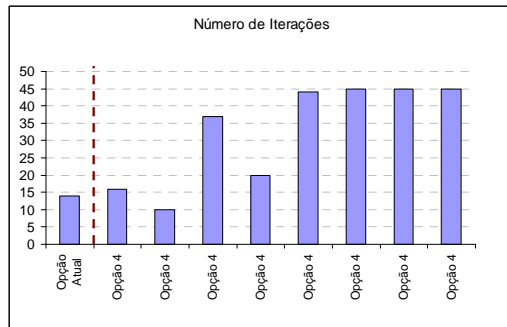
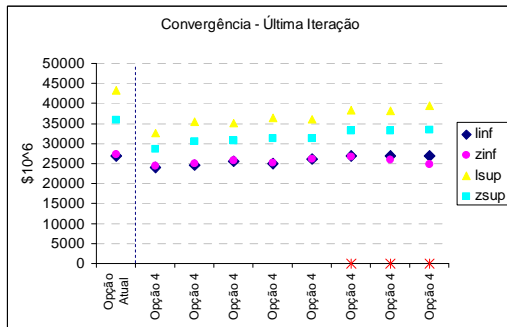
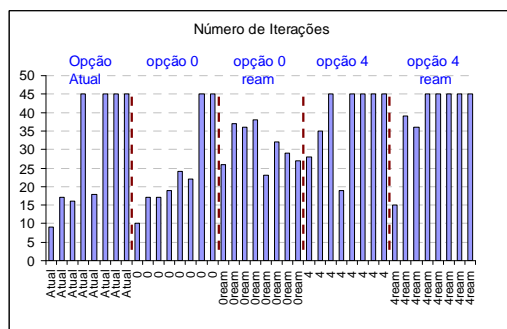
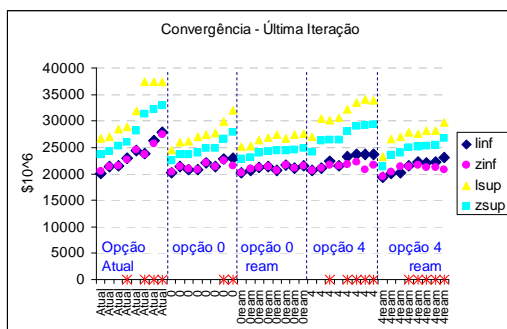


Figura 217: Resultados – Variação Semente – Opção 4 - PMO Set/05





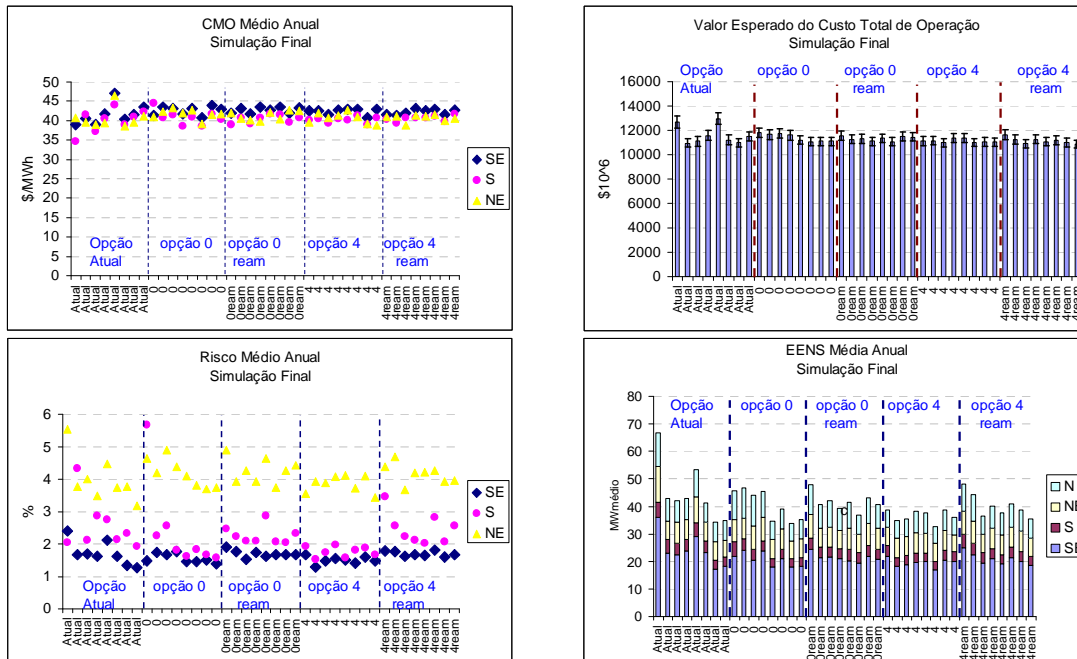
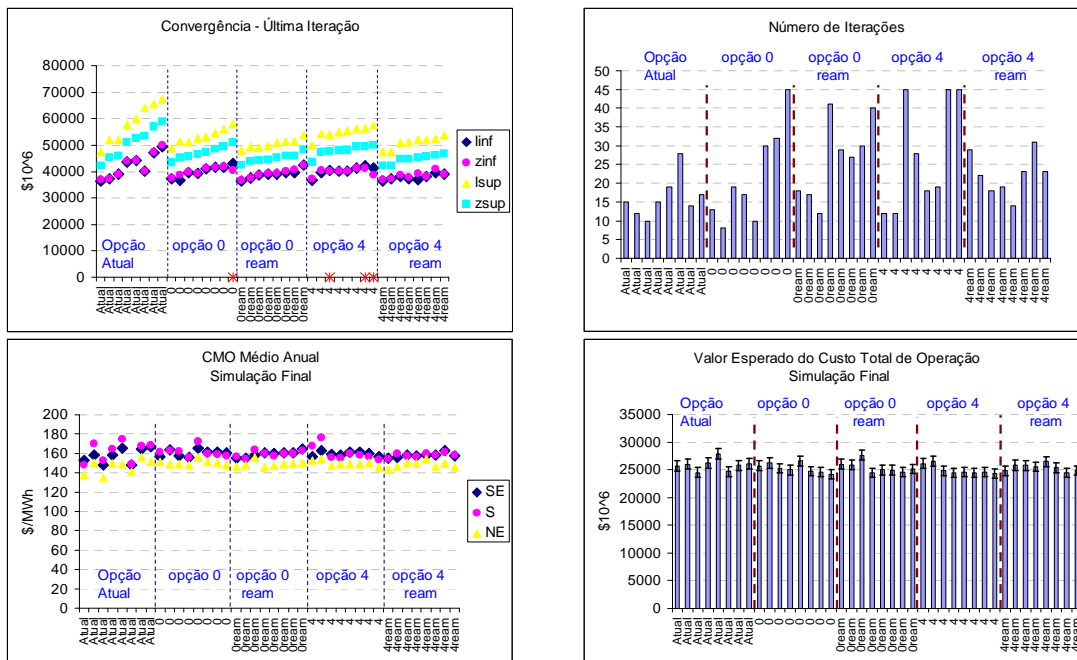


Figura 218: Resultados- Variação Semente – PMO Set/04



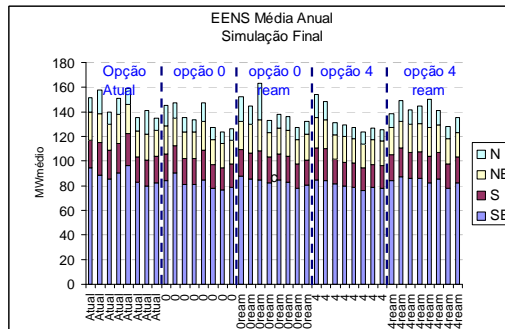
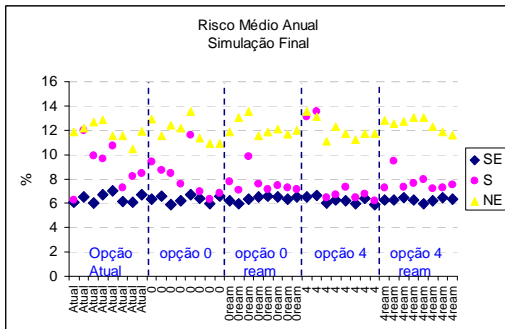


Figura 219: Resultados- Variação Semente – PMO Jul/07

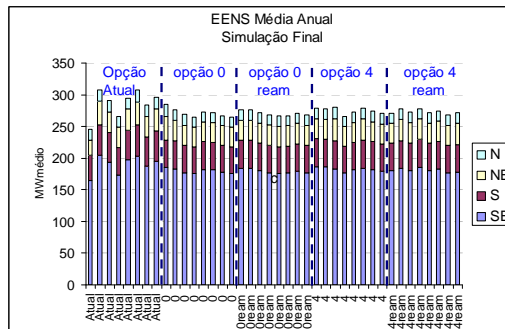
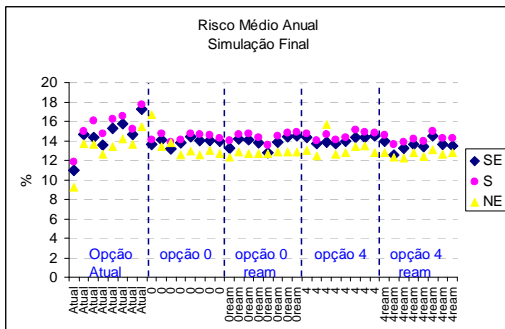
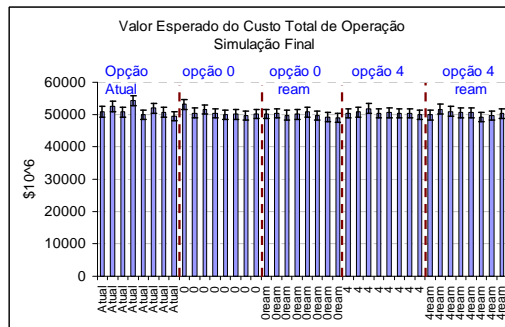
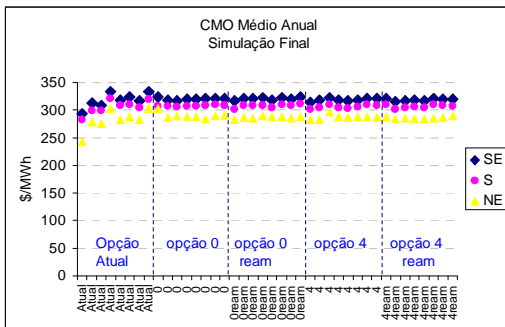
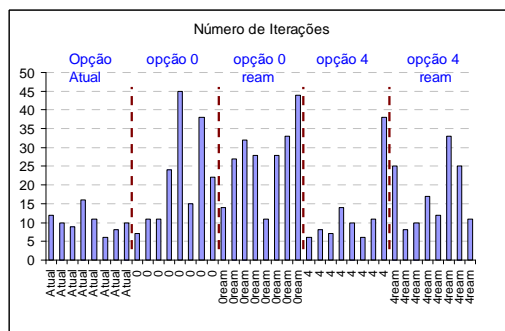
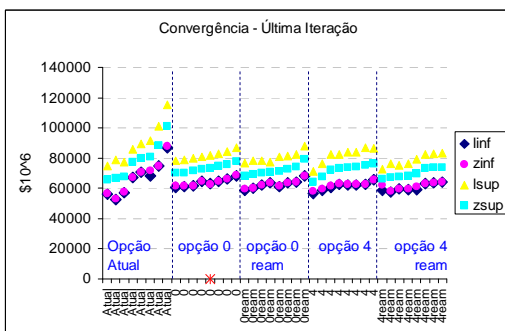


Figura 220: Resultados- Variação Semente – PMO Jan/08

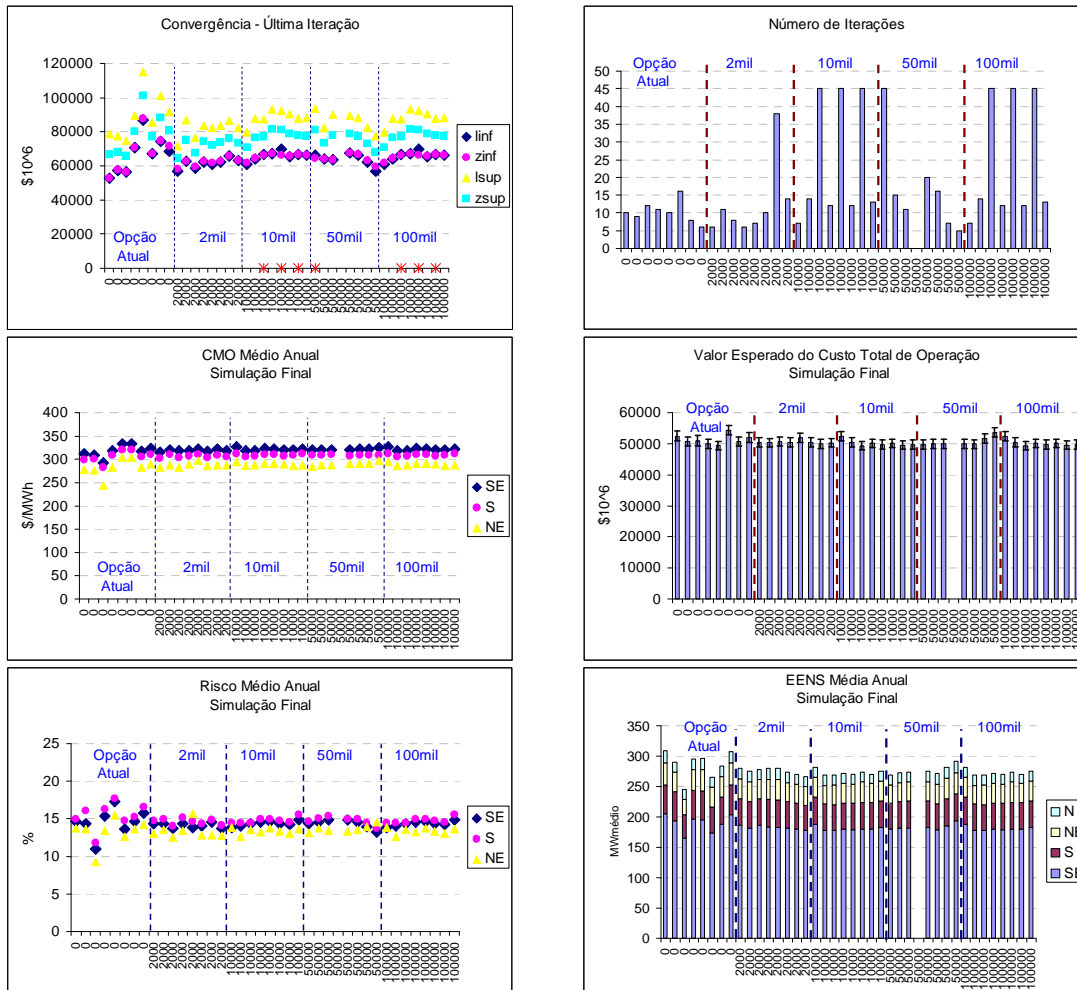
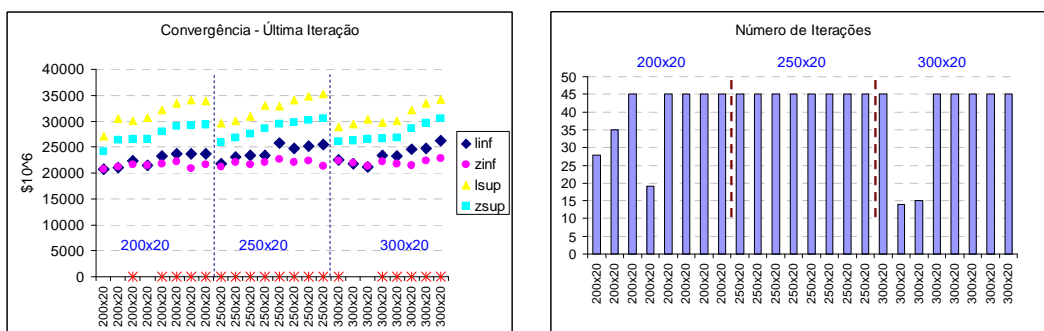


Figura 221: Resultados – Variação Amostra & Semente – Opção 4 - PMO Jan/08



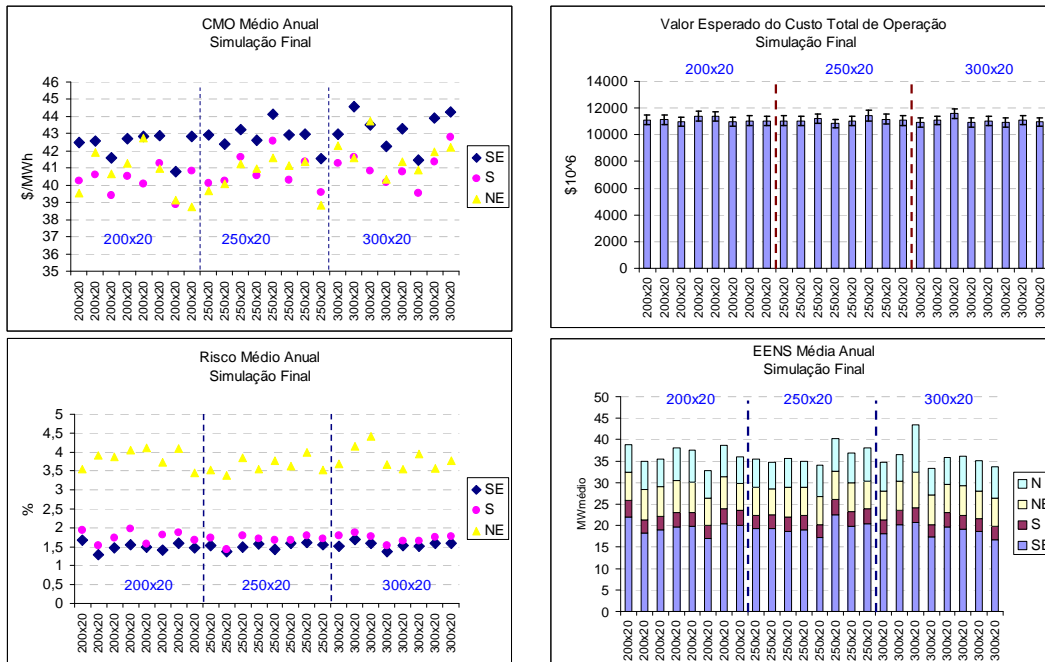


Figura 222: Resultados– Variação Forward & Semente – Opção 4 - PMO Set/04



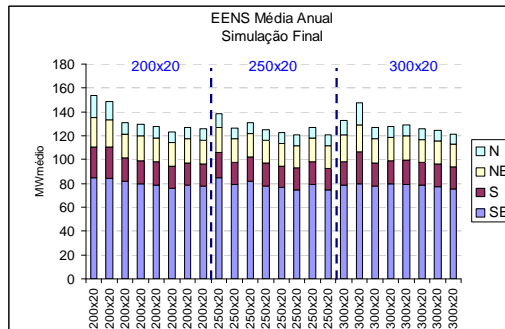
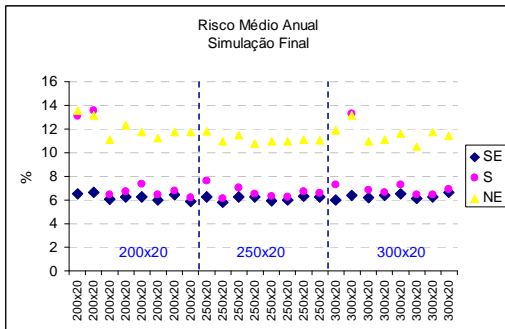


Figura 223: Resultados- Variação Forward & Semente – Opção 4 - PMO Jul/07

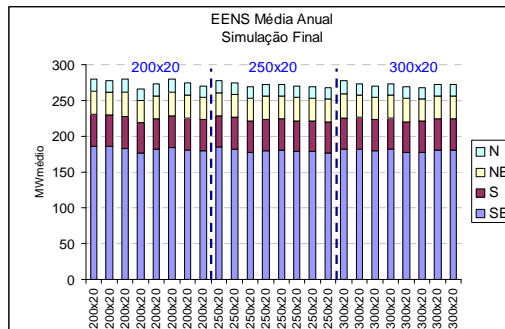
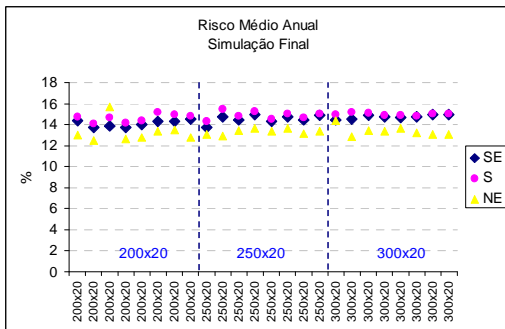
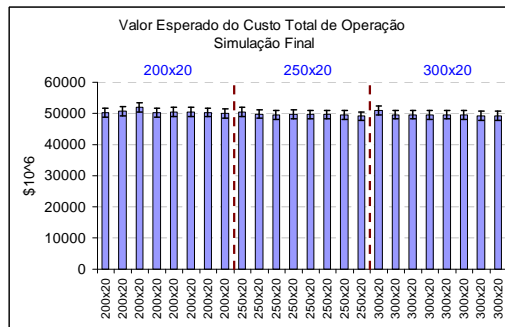
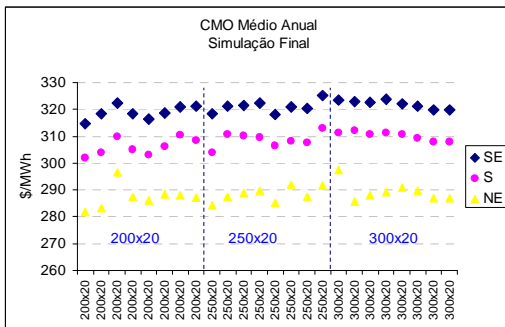
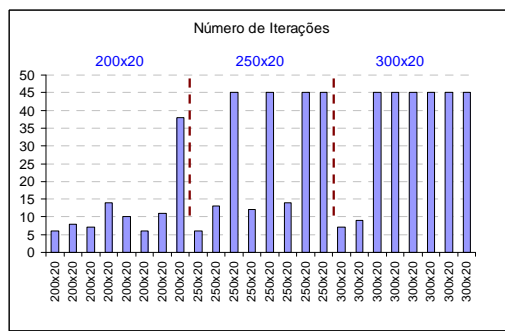
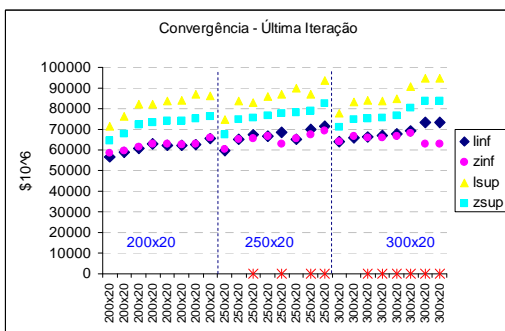


Figura 224: Resultados- Variação Forward & Semente – Opção 4 - PMO Jan/08

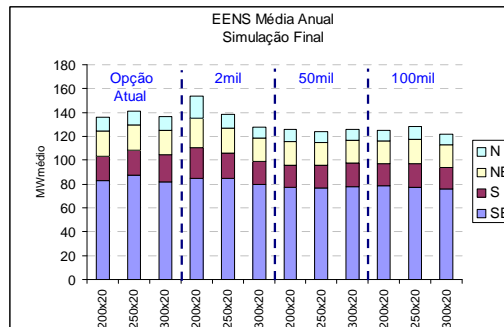
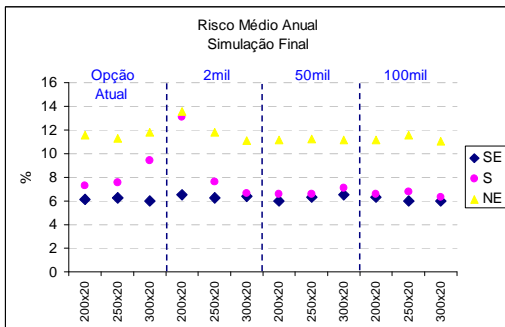
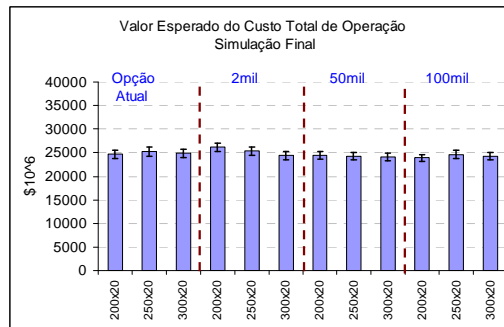
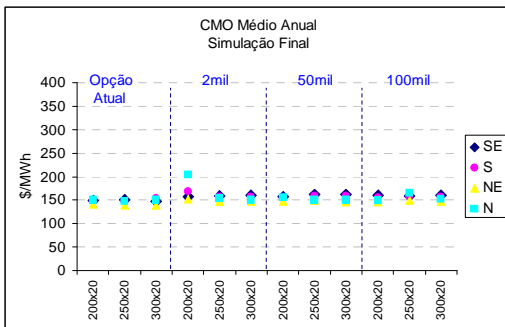
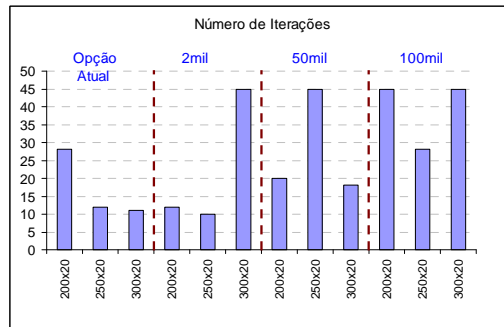
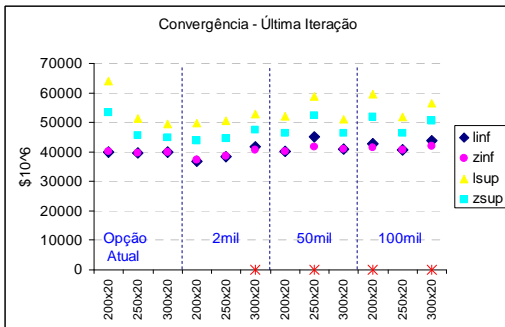
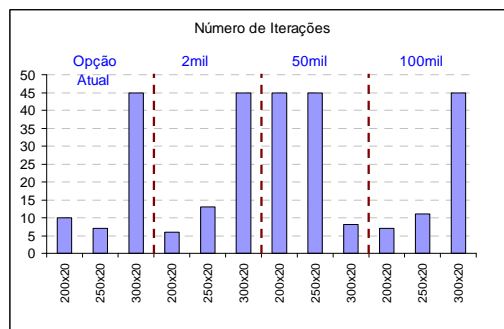
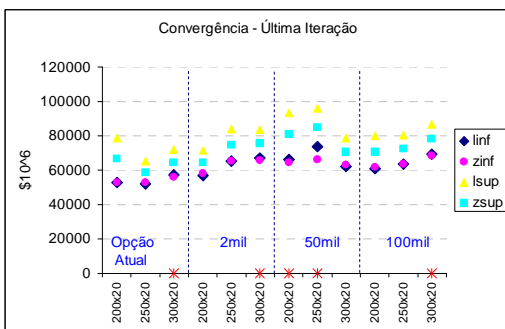


Figura 225: Resultados – Variação Forward & Amostra – Opção 4 - PMO Jul/07



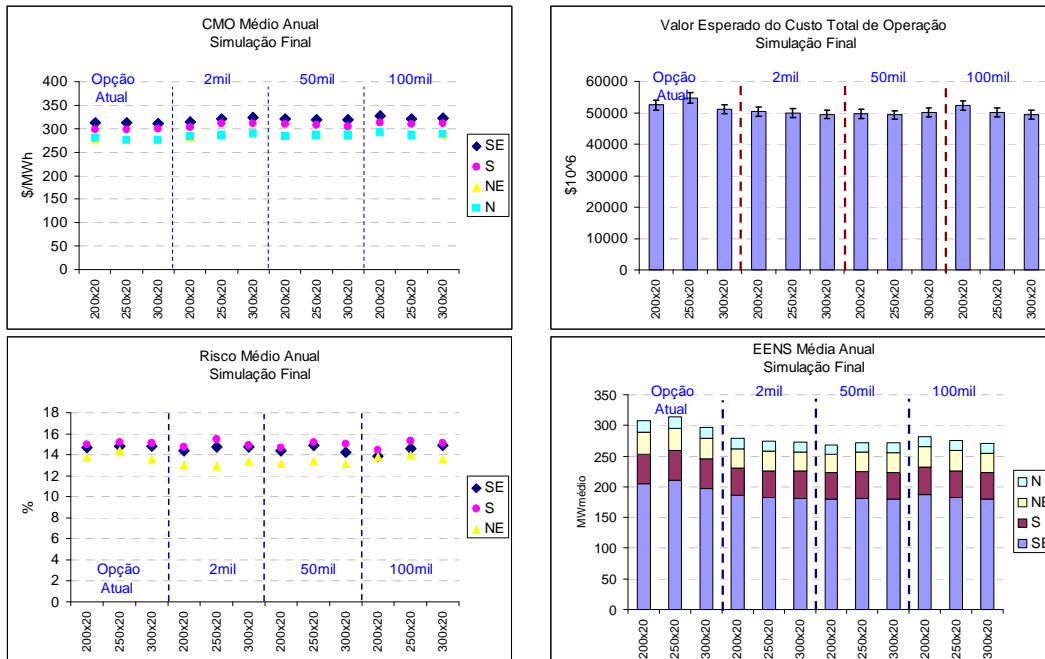
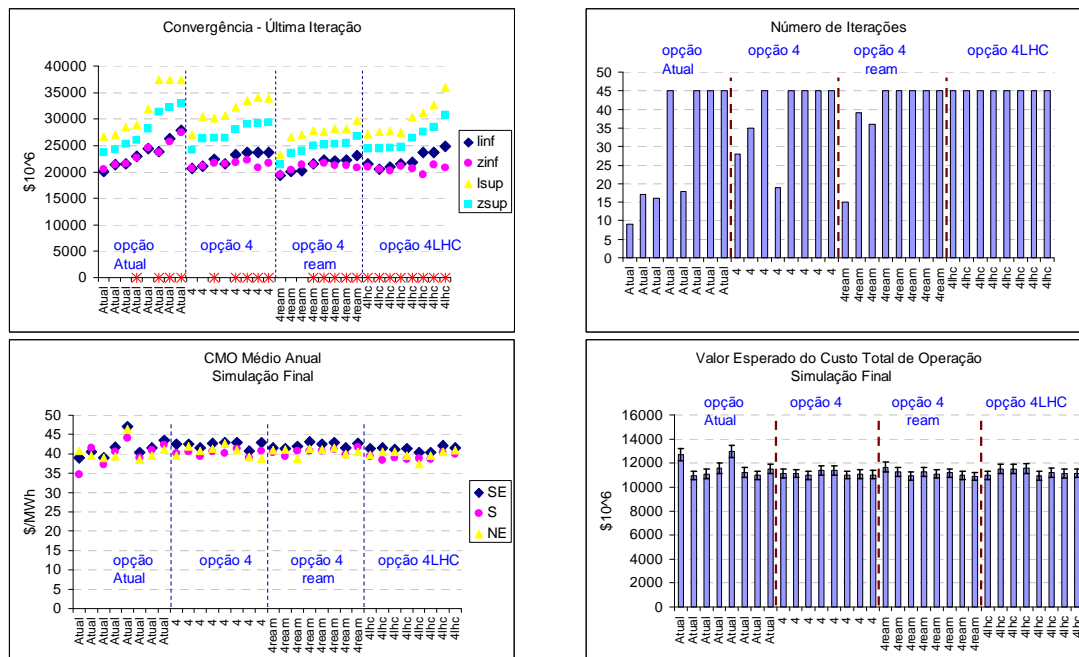


Figura 226: Resultados- Variação Forward & Amostra – Opção 4 - PMO Jan/08



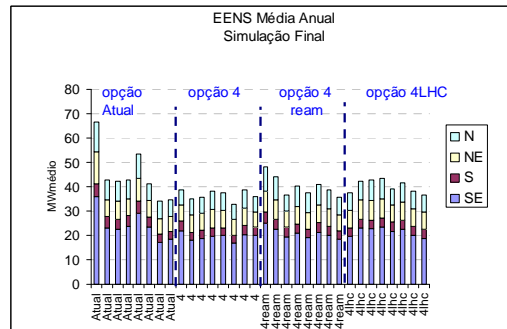
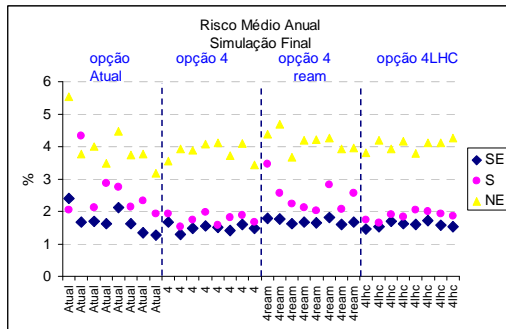


Figura 227: Resultados– Variação Semente – Opção 4 & Variantes - PMO Set/04

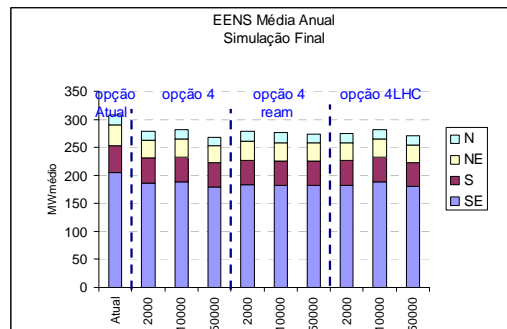
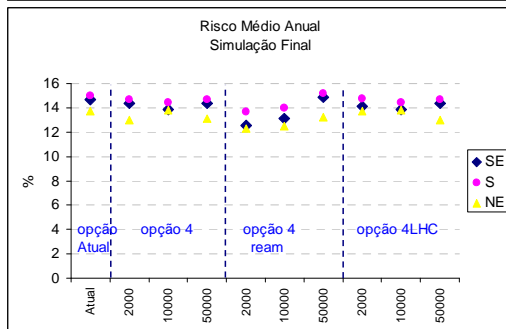
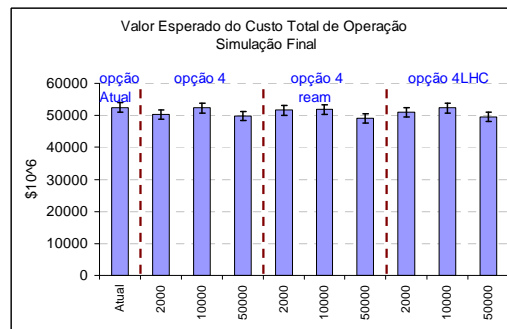
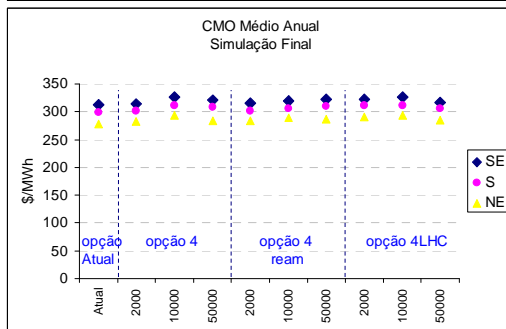
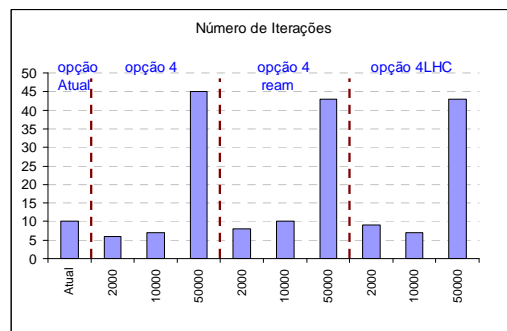
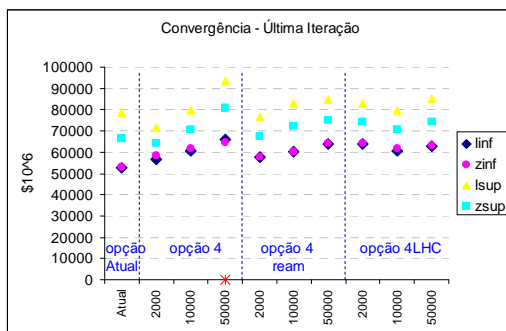
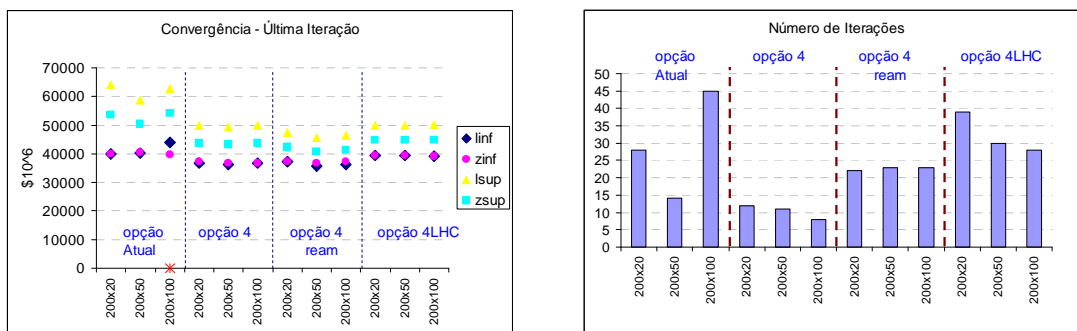


Figura 228: Resultados– Variação Amostra – Opção 4 & Variantes - PMO Jan/08





Figura 229: Resultados- Variação Backward – Opção 4 & Variantes - PMO Set/04



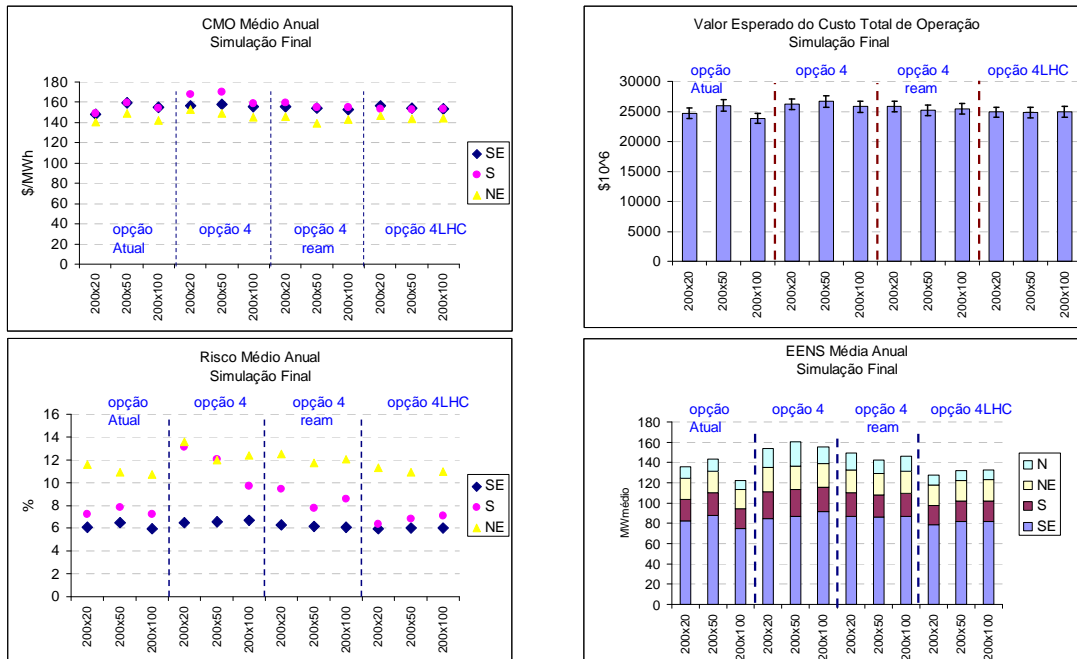
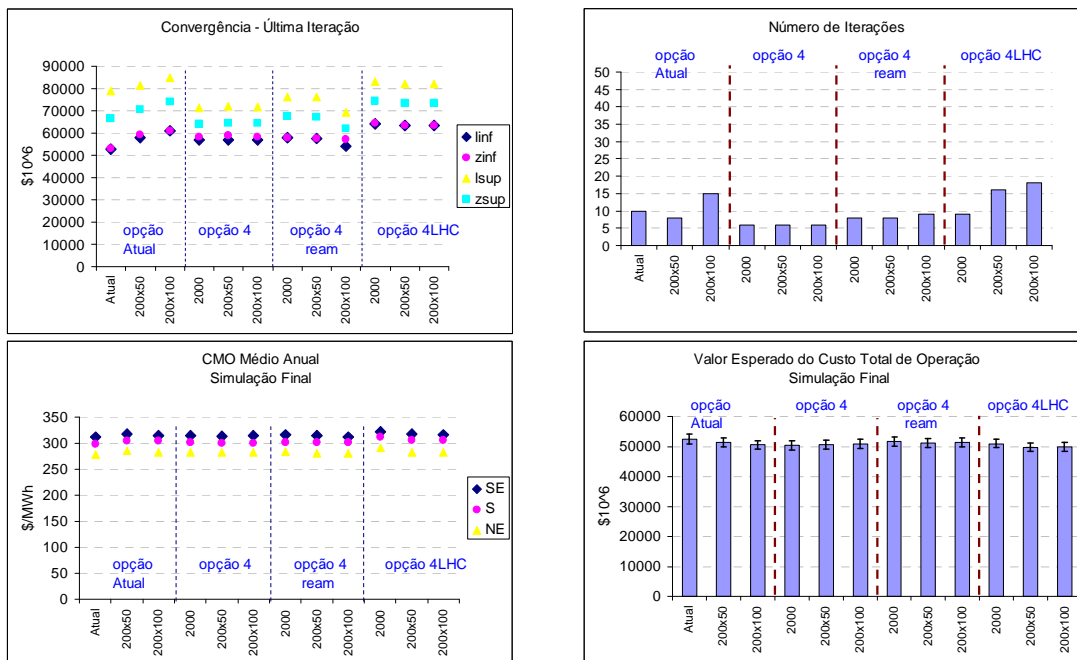


Figura 230: Resultados- Variação Backward – Opção 4 & Variantes - PMO Jul/06



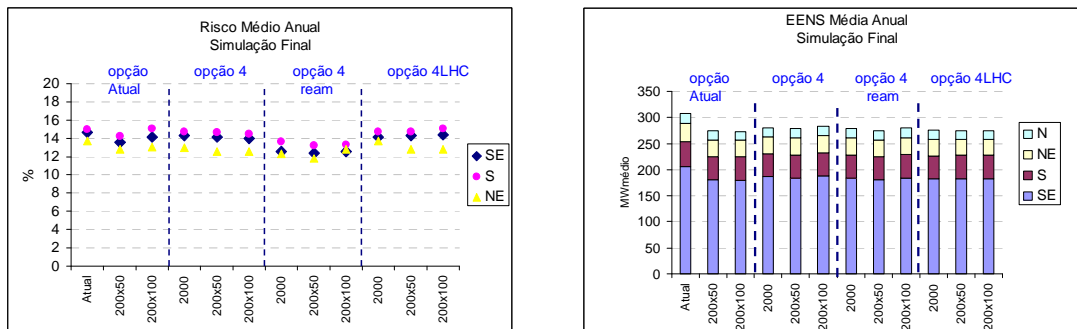


Figura 231: Resultados- Variação Backward – Opção 4 & Variantes - PMO Jan/08



Figura 232: Resultados- Variação Forward – Opção 4 & Variantes - PMO Set/04

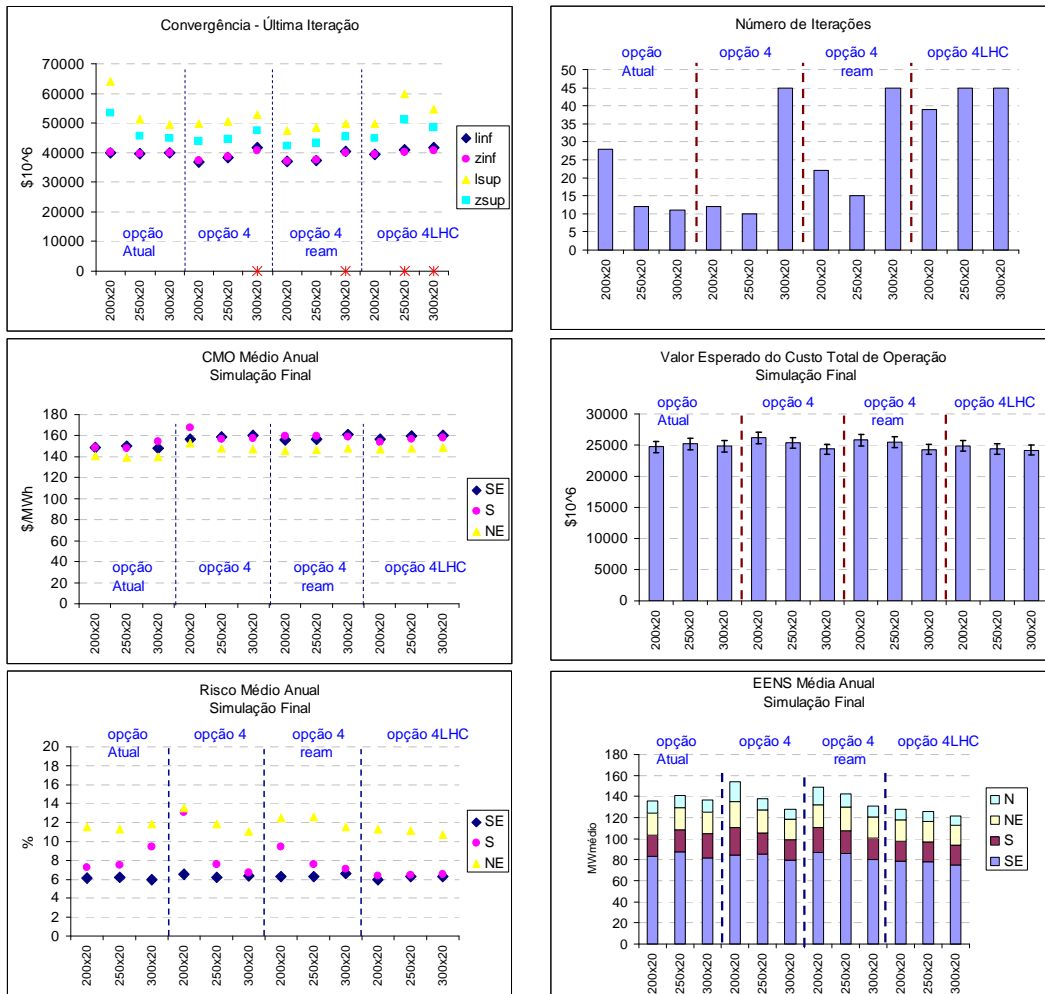
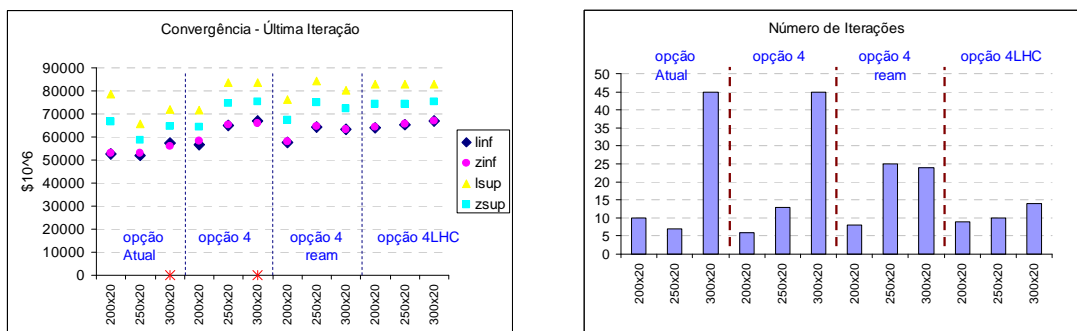


Figura 233: Resultados – Variação Forward – Opção 4 & Variantes - PMO Jul/06



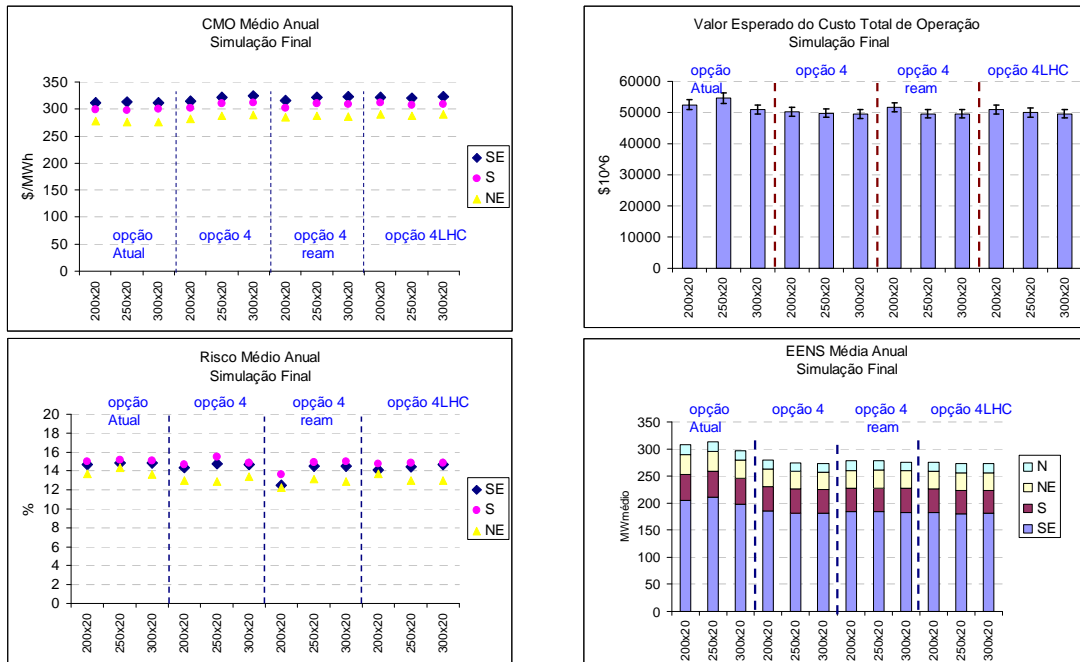


Figura 234: Resultados – Variação Forward – Opção 4 & Variantes - PMO Jan/08