

Cassio Roberto Vieira Tahan Junior

**APLICAÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS
NO PROBLEMA DO PLANEJAMENTO DA OPERAÇÃO
ENERGÉTICA DE MÉDIO PRAZO**

Dissertação submetida ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Sistemas de Energia.
Orientador: Prof. D. Eng. Erlon Cristian Finardi

Florianópolis
2017

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Tahan Junior, Cassio Roberto Vieira
Aplicação da análise de componentes principais no
problema do planejamento da operação energética de
médio prazo / Cassio Roberto Vieira Tahan Junior ;
orientador, Erlon Cristian Finardi, 2017.
147 p.

Dissertação (mestrado) - Universidade Federal de
Santa Catarina, Centro Tecnológico, Programa de Pós
Graduação em Engenharia Elétrica, Florianópolis, 2017.

Inclui referências.

1. Engenharia Elétrica. 2. Otimização
Estocástica. 3. Análise de Componentes Principais.
4. Modelo Autoregressivo Periódico. 5. Planejamento
da Operação Energética. I. Finardi, Erlon Cristian .
II. Universidade Federal de Santa Catarina.
Programa de Pós-Graduação em Engenharia Elétrica.
III. Título.

Cassio Roberto Vieira Tahan Junior

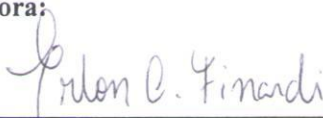
**APLICAÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS
NO PROBLEMA DO PLANEJAMENTO DA OPERAÇÃO
ENERGÉTICA DE MÉDIO PRAZO**

Esta dissertação foi julgada aprovada para obtenção do título de “Mestre em Engenharia Elétrica”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.

Florianópolis, 22 de novembro de 2017.

Prof. Marcelo Lobo Heldwein, Dr.
Coordenador do Curso de Pós-Graduação em Engenharia Elétrica

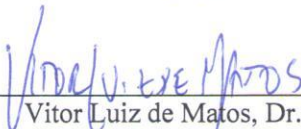
Banca Examinadora:



Prof. Erlon Cristian Finardi, Dr. Eng.
Orientador



Prof. Luiz Ricardo Nakamura, Dr.
UFSC



Vitor Luiz de Matos, Dr. Eng.
Plan4 Engenharia

AGRADECIMENTOS

Fico muito feliz em poder escrever, mesmo que de forma breve, toda minha gratidão às pessoas que contribuíram durante meu mestrado. Essa participação me fez crescer tanto profissionalmente quanto pessoalmente. Sem o apoio destas pessoas, com certeza não seria possível passar por mais essa etapa em minha vida.

Primeiramente gostaria de agradecer a Deus, que sempre me deu força e vontade para continuar, principalmente nos momentos mais difíceis dessa caminhada. Agradeço também:

À minha mãe, Maria Aparecida, e meu pai, Cassio, que ofereceram ajuda incondicional durante todo momento e me apoiaram em cada uma de minhas escolhas. Seu amor e suas atitudes são exemplos em todas as etapas de minha vida. Agradeço aos meus irmãos, Eduardo e Lívia por todo apoio demonstrado.

Ao meu orientador, Erlon Cristian Finardi, deixo minha admiração e gratidão por todo o suporte e ensinamento que reflete demais na minha formação acadêmica. Agradeço também pela excelente orientação e por acreditar em mim para fazer um trabalho que leva seu nome.

Ao amigo Paulo Larroyd, por toda dedicação e contribuição durante essa jornada. Seu conhecimento, boa vontade e apoio foram essenciais para a elaboração deste trabalho.

Aos meus queridos amigos Rogério Meneghetti e Bruci Cabrera por entrarem nessa caminhada comigo, pelo companheirismo e ajuda ao longo do percurso.

Aos professores do LabPlan e Labspot pela transmissão de conhecimento no primeiro ano do mestrado.

Aos meus amigos do LabPlan e Labspot que fizeram o período de pesquisa mais prazeroso e pela atenção nos momentos que precisei de ajuda.

Aos meus amigos da graduação, que mesmo de longe, tiveram influência determinante para a conclusão do mestrado. Obrigado pelas discussões, ideias e assistência. Vocês sabem da importância de cada um.

Finalmente, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro durante a pesquisa.

RESUMO

No caso Brasileiro, o problema do planejamento da operação energética de médio prazo (PMP) visa minimizar uma combinação convexa dada pela soma dos valores da geração termelétrica esperada e o *Conditional Value at Risk* dos cenários hidrológicos críticos ao longo de um horizonte de cinco anos. As principais restrições são dadas pelo atendimento à demanda, balanço hidráulico dos reservatórios e limites operativos das usinas. Matematicamente, o problema do PMP é representado por um modelo de programação estocástica linear multiestágio, cuja complexidade de solução é demandada por algoritmos que empregam algum tipo de decomposição. Nesse sentido, a Programação Dinâmica Dual Estocástica (PDDE) é a metodologia mais bem-sucedida na solução do PMP. Na PDDE, o problema é decomposto em uma série de subproblemas de um único mês e o objetivo consiste em construir iterativamente a Função de Custo Futuro (FCF). No caso Brasileiro, a FCF depende dos volumes dos reservatórios no início do mês e das afluições de meses anteriores. Com modelagem das usinas hidrelétricas individualizadas, objeto deste trabalho, a PDDE demanda um número elevado de iterações para convergir devido à dimensão elevada do espaço de estados da FCF. Dado que as afluições são as variáveis aleatórias do problema, este trabalho propõe o uso da técnica de Análise de Componentes Principais (ACP) com finalidade de reduzir o espaço de estados da PDDE, sem perda significativa da informação presente no comportamento estatístico original das afluições. Por isso, com o intuito de reduzir o esforço computacional e obter uma política operativa menos onerosa, a ACP inicialmente é empregada para construir um modelo autoregressivo das afluições com base em componentes principais (CP). Na sequência, é feita uma adaptação do algoritmo da PDDE para incluir as CP como variável de estado no lugar das afluições dos meses anteriores. Neste contexto, o principal objetivo deste trabalho é usar essa estratégia baseada na ACP e resolver o PMP em menor tempo sem comprometer a qualidade dos resultados, tendo como base a solução fornecida pelo problema sem CP. Para tanto, este trabalho faz uso de uma configuração hidrotérmica com 16 hidrelétricas localizadas nas cascatas dos rios Uruguai e Iguaçu. Por meio do modelo autoregressivo periódico, quatro modelos de geração de afluições são calculados: sem as CP, uma, três e 16 CP, nomeados SCP, CP1, CP3 e CP16, respectivamente. Com o uso da PDDE, são determinadas as FCF sendo, em seguida, as mesmas utilizadas em uma simulação da operação no histórico e em 10.000 cenários de afluições geradas em cada um dos modelos. No total, foram executados 1.200 problemas do PMP e os resultados indicam

que os modelos SCP e CP16 apresentam custos esperados de operação bastante similares. Por sua vez, o modelo CP1 apresentou em média custos 1,5% inferiores. Para um mesmo tempo computacional, o número de iterações do modelo CP1 é aproximadamente 10% maior em comparação aos modelos SCP e CP16. Por outro lado, os modelos com CP apresentaram em média reduções no tempo computacional por iteração da PDDE na ordem de 20%, estas especialmente verificadas nas últimas iterações do algoritmo. No geral, a técnica da ACP combinada com a PDDE necessita ainda de uma série de investigações, mas como um trabalho pioneiro na área tem-se uma expectativa positiva no seu uso para o problema do PMP do caso Brasileiro.

Palavras-chave: Otimização Estocástica, Análise de Componentes Principais, Modelo Autoregressivo Periódico, Planejamento da Operação Energética.

ABSTRACT

In Brazilian electric-energy system, the long-term energetic operation planning (EOP) problem aims to minimize a convex combination given by the sum of the expected thermoelectric generation and the Conditional Value at Risk from critical hydrological scenarios throughout a five-year horizon. The main constraints are given by the demand, hydraulic balance of the reservoirs and operating limits of the hydro plants. Mathematically, the EOP problem is represented by a multistage linear stochastic programming model whose solution complexity is demanded by algorithms that use some type of decomposition. In this sense, Stochastic Dual Dynamic Programming (SDDP) is the most successful methodology to the EOP solution. In SDDP, the problem is decomposed into a set of single-month subproblems, and the goal is to build the Future Cost Function (FCF) iteratively. In the Brazilian case, the FCF depends on the reservoirs volumes at the beginning of the month and the inflows from previous months. In a modeling of individual hydropower plants, SDDP demands many iterations to converge due to the high dimension of the state space of the FCF. Once the inflows are the random variables of the problem, this work proposes the use of the Principal Component Analysis (PCA) to reduce the SDDP state space without significant information loss of the original statistical behavior of the inflows. Therefore, in order to reduce the computational effort and to obtain a less costly operational policy, the PCA is initially used to build an inflows autoregressive model based on principal components (PC). Hereupon, an adaptation of the SDDP algorithm is made to include the PC as a state variable instead of the inflows from previous months. In this context, the main objective of this work is to use this strategy based on the PCA to solve the EOP in a shorter amount of time without compromising the quality of the results, which is based on the solution given by the problem without PC. This work uses a hydrothermal configuration with 16 hydroelectric plants from Uruguay and Iguazu rivers. Through periodic autoregressive model, four inflow generation models are calculated: without PC, one, three and 16 PC, named SPC, PC1, PC3 and PC16, respectively. By SDDP, the FCF are determined and then the same are used to simulate the operation in the historical and in 10,000 scenarios of inflows generated by each one of the models. In total, 1,200 long-term EOP problems were executed and the results indicate that the SPC and PC16 models have very similar expected operating costs. On the other hand, the PC1 model had on average 1.5% lower costs. For the same computational time, the number of iterations of the PC1 model is approximately 10% higher in comparison to the SPC and

PC16 models. The models with PC showed around 20% reductions in the computational time by iteration of the SDDP, these specially verified in the last iterations of the algorithm. In general, PCA combined with SDDP still requires a series of investigations, but as a pioneering work in the area there is a positive prospect in its use for the EOP problem in Brazilian case.

Keywords: Stochastic Optimization, Principal Component Analysis, Periodic Autoregressive Model, Energetic Operation Planning.

LISTA DE FIGURAS

Figura 1.1 - Processo de decisão de sistema hidrotérmicos.	23
Figura 1.2 – Funções de custo futuro e imediato.	23
Figura 2.1 - Função de Cota Montante da UHE de Machadinho.	35
Figura 2.2 - Cota Jusante da UHE de Machadinho.	36
Figura 2.3 - FPH UHE de Machadinho ($s=0$).	38
Figura 2.4 - Registro de aflúncias mensais de Salto Osório, y_h e y_p	41
Figura 2.5 - Comparação das aflúncias da usina de Salto Osório.	42
Figura 2.6 - FPH por <i>Convex Hull</i> UHE de Machadinho.	44
Figura 2.7 - Dois modelos da FPH de Machadinho ($s=0$).	45
Figura 3.1 - Séries aflúncia da UHE Salto Osório.	50
Figura 3.2 - Coeficientes da FACP e intervalo de confiança.	53
Figura 3.3 - Distribuição Lognormal positiva.	54
Figura 4.1 - Função de Custo Futuro.	63
Figura 4.2 - Árvore de cenários.	65
Figura 4.3- Criação dos Cortes de Benders.	68
Figura 4.4 - Esquemático das UHE para o exemplo numérico.	71
Figura 4.5 - Árvore de cenários para 5 meses (duas realizações/mês).	72
Figura 4.6 - Evolução do Z_{sup} e Z_{inf}	73
Figura 4.7 - Volume útil ao longo do tempo.	73
Figura 5.1 - Evolução do Z_{sup} e Z_{inf} (com CP).	89
Figura 5.2 - Volume útil ao longo do tempo.	90
Figura 5.3 - Comparação dos casos para a UHE de Barra Grande.	92
Figura 6.1 – Configuração hidráulica do sistema.	94
Figura 6.2 - Limite inferior dos problemas com 20 realizações por mês.	102
Figura 6.3 - Limite inferior dos problemas com 50 realizações por mês.	102
Figura 6.4 - Limite inferior dos problemas com 100 realizações por mês.	103
Figura 6.5 – Limite inferior da PDDE dos casos SCP.	105
Figura 6.6 – Valor esperado do limite superior do problema SCP.	105

Figura 6.7 - Brecha de otimalidade do modelo SCP.....	106
Figura 6.8 - Brecha de otimalidade do modelo CP16.	107
Figura 6.9 - Brecha de otimalidade do modelo CP3.	107
Figura 6.10 - Brecha de otimalidade do modelo CP1.	108
Figura 6.11 - Custo esperado das séries SCP (10 FCFs por modelo) – Limite de tempo de 10 minutos.	110
Figura 6.12 - Custo esperado das séries SCP (10 FCFs por modelo) – Limite de tempo de 30 minutos.	111
Figura 6.13 - Custo esperado das séries SCP (10 FCFs por modelo).	112
Figura 6.14 - Custo esperado das séries CP16 (10 FCFs por modelo)	113
Figura 6.15 - Custo esperado das séries CP3 (10 FCFs por modelo).....	113
Figura 6.16 - Custo esperado das séries CP1 (10 FCFs por modelo).....	114
Figura 6.17 - Custo esperado das séries históricas (10 FCFs por modelo).	115
Figura 6.18 – Percentil do custo total esperado - Séries SCP.	116
Figura 6.19 - Médias de Custo Total em Valor Presente (Milhões de Reais)..	117
Figura 6.20 – Percentil do custo total esperado - Séries CP16.	118
Figura 6.21 - Médias de Custo Total em Valor Presente (Milhões de Reais)..	118
Figura 6.22 - Percentil do custo total esperado - Séries históricas.....	119
Figura 6.23 - Médias de Custo Total em Valor Presente (Milhões de Reais)..	120
Figura 6.24 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 100% (mais caro) da operação por FCFs 1CP na série histórica.	121
Figura 6.25 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 75% da operação por FCFs 1CP na série histórica.	121
Figura 6.26 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 50% da operação por FCFs 1CP na série histórica.	121
Figura 6.27 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 25% da operação por FCFs 1CP na série histórica.	122
Figura 6.28 - Simulações de operação (FCFs de todos os modelos) no cenário mais barato da operação por FCFs 1CP na série histórica.	122
Figura 6.29 – Variação média de <i>Zinf</i> nas últimas três iterações.....	124

Figura 6.30 - Variação média de <i>Zinf</i> nas últimas cinco iterações.....	124
Figura 6.31 - Variação média de <i>Zinf</i> nas últimas 10 iterações.	125
Figura 6.32 - Número de iterações.....	126
Figura 8.1 - Limite inferior da PDDE dos casos CP16.	133
Figura 8.2 - Valor esperado do limite superior do problema CP16.....	133
Figura 8.3 - Limite inferior da PDDE dos casos CP3.	134
Figura 8.4 - Valor esperado do limite superior do problema CP3.	134
Figura 8.5 - Limite inferior da PDDE dos casos CP1.	135
Figura 8.6 - Valor esperado do limite superior do problema CP1.	135
Figura 9.1 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 100% (mais caro) da operação por FCFs 1CP em séries SCP.....	137
Figura 9.2 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 75% da operação por FCFs 1CP em séries SCP.	137
Figura 9.3 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 50% da operação por FCFs 1CP em séries SCP.	138
Figura 9.4 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 25% da operação por FCFs 1CP em séries SCP.	138
Figura 9.5 - Simulações de operação (FCFs de todos os modelos) no cenário mais barato da operação por FCFs 1CP em séries SCP.	138
Figura 10.1 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 100% (mais caro) da operação por FCFs 1CP em séries de 16 CP. ...	139
Figura 10.2 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 75% da operação por FCFs 1CP em séries de 16 CP.	139
Figura 10.3 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 50% da operação por FCFs 1CP em séries de 16 CP.	140
Figura 10.4 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 25% da operação por FCFs 1CP em séries de 16 CP.	140
Figura 10.5 - Simulações de operação (FCFs de todos os modelos) no cenário mais barato da operação por FCFs 1CP em séries de 16 CP.....	140
Figura 11.1 - Número de Iterações (20 realizações por mês).....	141

Figura 11.2 - Número de Iterações (50 realizações por mês).141
Figura 11.3 - Número de Iterações (100 realizações por mês).142

LISTA DE TABELAS

Tabela 2.1 - <i>Convex Hull</i> da FPH de Campos Novos.	48
Tabela 4.1 - Resíduos das UHEs de Barra Grande e Machadinho	72
Tabela 5.1 - Histórico de afluições incrementais	86
Tabela 5.2 - Autovalores do exemplo numérico	87
Tabela 5.3 - Autovetores do exemplo numérico	87
Tabela 5.4 - Histórico de componentes principais do exemplo numérico.....	88
Tabela 5.5 - Resíduos das UHEs de Barra Grande e Machadinho (com CP).....	89
Tabela 6.1 – Dados operativos das UHEs.	95
Tabela 6.2 – Usinas termelétricas do sistema.	96
Tabela 6.3- Porcentagem acumulada das CP.	96
Tabela 6.4 – Ordem do modelo PAR para o caso SCP.	97
Tabela 6.5 – Ordem do modelo PAR com CP.	98
Tabela 6.6 - Número de variáveis de estado.	100
Tabela 6.7 – Desempenho computacional do modelo SCP. <i>Zinf</i> e <i>Zsup</i> em milhões de reais.	104
Tabela 6.8 - Brecha de otimalidade (%) para o modelo SCP (10 minutos).	104
Tabela 6.9 - Número de realizações com menor brecha de otimalidade.	108
Tabela 6.10 – Tempo médio por iteração	127

LISTA DE ABREVIATURAS E SIGLAS

ACP	Análise de Componentes Principais
ANEEL	Agência Nacional de Energia Elétrica
CH	<i>Convex Hull</i>
CMO	Custo Marginal de Operação
CP	Componente(s) Principal(ais)
DA	Decomposição Aninhada
EMQ	Erro Médio Quadrático
FACP	Função de Autocorrelação Parcial
FCF	Função Custo Futuro
FCI	Função de Custo Imediato
FCJ	Função de Cota Jusante
FCM	Função de Cota Montante
FDP	Função de Distribuição de Probabilidade
FPH	Função de Produção Hidrelétrica
ONS	Operador Nacional do Sistema Elétrico
PAR	<i>Periodic Autoregressive Model</i>
PCA	<i>Principal Component Analysis</i>
PCP	Planejamento da Operação Energética de Curto Prazo
PDDE	Programação Dinâmica Dual Estocástica
PDE	Programação Dinâmica Estocástica
PDO	Programação Diária da Operação Eletroenergética
PL	Programação Linear
PMP	Planejamento da Operação Energética de Médio Prazo
SIN	Sistema Interligado Nacional
UHE	Usina Hidrelétrica
UTE	Usina Termelétrica

SUMÁRIO

1.	INTRODUÇÃO	21
1.1	REVISÃO BIBLIOGRÁFICA	27
1.2	OBJETIVOS DO TRABALHO	30
1.3	ESTRUTURA DA DISSERTAÇÃO	31
2.	COMPONENTES DO SISTEMA HIDROTÉRMICO	33
2.1	USINAS TERMELÉTRICAS	33
2.2	USINAS HIDRELÉTRICAS	34
2.2.1	Função de Produção Hidrelétrica	34
2.2.2	Balanco Hídrico nos Reservatórios	38
2.3	REPRESENTAÇÃO DAS AFLUÊNCIAS	40
2.4	LINEARIZAÇÃO DA FPH	42
2.4.1	Precisão do modelo linear por partes	45
3.	MODELO PARA GERAÇÃO DE AFLUÊNCIA	49
3.1	MODELO AUTOREGRESSIVO PERIÓDICO	49
3.2	MODELO PAR LOGNORMAL A TRÊS PARÂMETROS	50
3.2.1	Identificação da Ordem do Modelo Autoregressivo Periódico	51
3.2.2	Estimação dos Parâmetros	53
3.3	CORRELAÇÃO ESPACIAL	56
3.4	GERAÇÃO DAS SÉRIES SINTÉTICAS	58
4.	PROGRAMAÇÃO DINÂMICA DUAL ESTOCÁSTICA ...	59
4.1	DECOMPOSIÇÃO DE BENDERS	60
4.2	ÁRVORE DE CENÁRIOS	64
4.3	FORMULAÇÃO MATEMÁTICA	68
4.3.1	Modelo PAR Lognormal a Três Parâmetros por Usina	68
4.3.1.1	Função Objetivo	69
4.3.1.2	Restrição de Atendimento à Demanda	69
4.3.1.3	Restrições de Balanço Hídrico	70
4.3.1.4	Restrições de Dependência Temporal das Afluências	70

4.3.1.5 Função de Custo Futuro	70
4.3.1.6 Formulação Completa	70
4.4 EXEMPLO NUMÉRICO.....	71
4.5 SIMULAÇÃO DA OPERAÇÃO	75
4.5.1 Avaliação da Qualidade da Solução	75
4.5.1.1 Estimação da brecha de otimalidade para otimização estocástica	77
5. ANÁLISE DE COMPONENTES PRINCIPAIS	79
5.1 METODOLOGIA	79
5.1.1 Matriz de Dados	80
5.1.2 Matriz de Covariância.....	80
5.2 DETERMINAÇÃO DAS COMPONENTES PRINCIPAIS	81
5.3 NÚMERO DE COMPONENTES UTILIZADAS	82
5.4 TRANSFORMAÇÃO PARA AS VARIÁVEIS ORIGINAIS	82
5.5 MODELO DE GERAÇÃO DE AFLUÊNCIAS CONSIDERANDO COMPONENTES PRINCIPAIS.....	83
5.5.1 Modelo PAR Lognormal a Três Parâmetros considerando a ACP84	
5.6 EXEMPLO NUMÉRICO COM CP	85
5.7 SÉRIE SINTÉTICA PARA MODELOS SEM E COM CP	91
6. RESULTADOS COMPUTACIONAIS	93
6.1 SISTEMA HIDROTÉRMICO	94
6.1.1 Modelo Autoregressivo	96
6.2 DADOS ADICIONAIS PARA OTIMIZAÇÃO E SIMULAÇÃO	98
6.3 MODELOS ESTUDADOS	99
6.4 CRITÉRIO DE CONVERGÊNCIA.....	100
6.5 SIMULAÇÃO PARA ESTIMATIVA DO CUSTO SUPERIOR	100
6.6 ESTRATÉGIA DE COMPARAÇÃO	100
6.7 TEMPO COMPUTACIONAL DE 10 MINUTOS.....	101
6.7.1 Modelo SCP	101
6.8 RESULTADOS PARA TODOS OS TEMPOS COMPUTACIONAIS .	104
6.9 SIMULAÇÕES OPERATIVAS NAS SÉRIES SINTÉTICAS.....	109

6.9.1 Simulação das séries sintéticas SCP (10 minutos)	109
6.9.2 Simulação das séries sintéticas SCP (30 minutos)	110
6.9.3 Simulação das séries sintéticas SCP (todos os tempos)	111
6.9.4 Simulação das séries sintéticas CP16, CP3 e CP1.....	112
6.10 SIMULAÇÕES OPERATIVAS NAS SÉRIES HISTÓRICAS	114
6.11 CUSTOS DE GERAÇÃO TERMELÉTRICA E DÉFICIT	115
6.11.1 Sintética SCP	116
6.11.2 Custos Totais por Série Sintética CP16.....	117
6.11.3 Custos Totais por Série Histórica	119
6.12 DESPACHO TERMELÉTRICO.....	120
6.13 CONVERGÊNCIA E NÚMERO ITERAÇÕES	123
7. CONCLUSÕES E SUGESTÕES PARA TRABALHOS	
FUTUROS.....	129
7.1 SUGESTÕES PARA TRABALHOS FUTUROS	131
8. APÊNDICE A: Limite inferior e superior do problema	133
9. APÊNDICE B: Despacho termelétrico para a série SCP....	137
10. APÊNDICE C: Despacho termelétrico para a série CP16..	139
11. APÊNDICE D: Número de iterações por tamanho de árvore	
141	
REFERÊNCIAS	143

1. INTRODUÇÃO

A energia elétrica é um bem de extrema importância para a sociedade atual, visto que é uma das formas de energia mais utilizadas no mundo nas mais distintas áreas. Seu crescimento é um reflexo do ritmo de atividade dos setores industrial, comercial e de serviços. Então, a disponibilidade de geração de energia elétrica é um fator importante no processo de crescimento econômico e tecnológico de um país. Por isso, para um uso racional da energia elétrica se faz necessário realizar um planejamento preciso para perseguir a modicidade tarifária e continuidade de suprimento.

No Brasil, o parque gerador é predominantemente hidrelétrico, uma fonte bastante atrativa por ser barata, abundante e renovável. Segundo dados da ANEEL, em 2017 a hidroeletricidade cobre cerca de 61% da capacidade instalada do país, composta por grandes reservatórios com capacidade de regularização plurianual, principalmente aqueles localizados nas regiões nordeste e sudeste do país. Por sua vez, as unidades térmicas do país correspondem a quase 27% da capacidade instalada. Estas usinas são utilizadas em complementação ao despacho hidrelétrico, o que configura o Sistema Elétrico Brasileiro como de base hídrica com complementação térmica. Ainda há outras fontes de energia usadas no Brasil com menor relevância na matriz energética, mas suas projeções apontam para um aumento notável ao longo da próxima década, em especial as energias eólica e solar.

Pode-se dizer que a segurança energética de um sistema de base hídrica é expressivamente elevada com a construção de usinas de grande capacidade de regularização, isto é, grandes reservatórios para grandes deplecionamentos de água na geração de energia elétrica. Isso faz com que seja possível que a geração hídrica perca por períodos consideráveis de pouca chuva ou até seca. Outra alternativa para o aumento da segurança energética é a complementação com outras fontes, como as usinas termelétricas.

O sistema de produção e transmissão de energia elétrica do Brasil engloba as cinco regiões do país, o que pode considerá-lo único em âmbito mundial, sendo a coordenação técnica e operacional do sistema realizado pelo Operador Nacional do Sistema Elétrico (ONS). O sistema de transmissão deve ser bastante robusto, isto porque, dada sua extensão, as construções de unidades de geração hidráulica não necessariamente estão próximas dos centros de consumo. No caso brasileiro, o planejamento de um sistema com tais características é bastante complicado, uma vez que

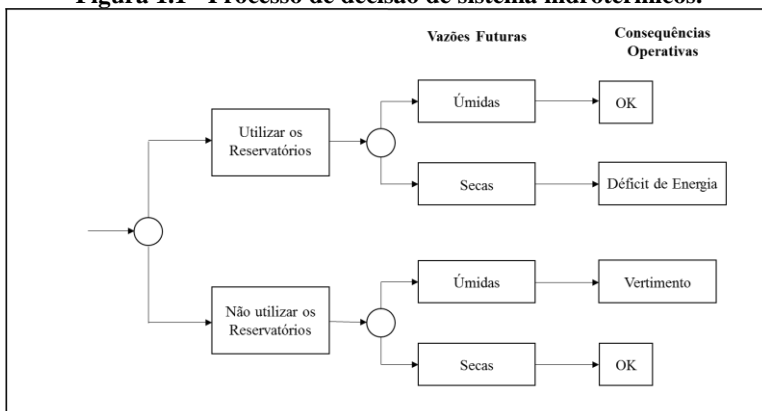
deve ser analisado o desempenho do sistema para horizontes de longo, médio e curto prazos (PEREIRA; PINTO, 1982).

Classicamente, o objetivo da operação energética de sistemas hidrotérmicos é proporcionar uma política para geração de cada usina hidrelétrica (UHE) e termelétrica (UTE) que minimize o custo operacional esperado de geração térmica para o período de estudo. Este processo é realizado por meio de uma sequência de decisões que garanta o atendimento da demanda, sempre respeitando as restrições operacionais.

Basicamente, a função objetivo considera o custo da geração das usinas termelétricas mais o custo de déficit, este relacionado ao não atendimento da demanda. Esta função é recorrentemente minimizada por meio da execução de modelos matemáticos em que, após cada período de execução (geralmente mensal), um plano de operação para cada unidade geradora é estabelecido, levando em consideração o valor da água para evitar possíveis déficits no futuro (GARCIA, 2005). No Brasil, atualmente o objetivo inclui, além do custo esperado de geração termelétrica, o Conditional Value at Risk (CVaR) associado com um percentual dos cenários hidrológicos mais secos.

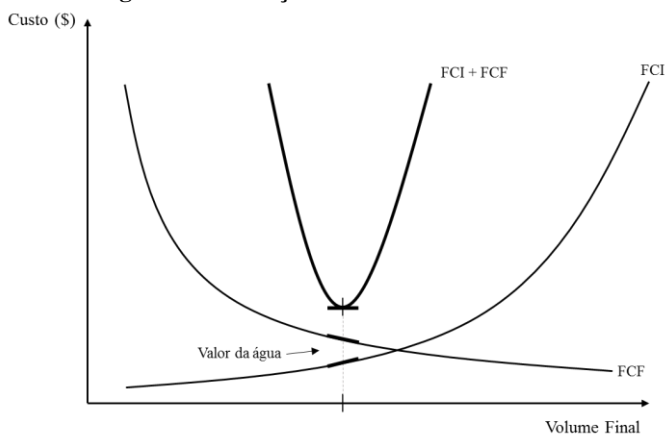
A configuração do Sistema Elétrico Brasileiro ainda é caracterizada pela predominância de geração hidrelétrica. Por isso, a decisão operativa é dependente diretamente das vazões afluentes e, então, decisões tomadas no presente têm impacto significativo nas decisões a serem tomadas no futuro. Desta forma, o operador deve considerar tanto o benefício do uso imediato da água estocada dos reservatórios, quanto o benefício futuro do seu armazenamento. Isso porque, apesar do custo da água ser considerado igual a zero no presente, um custo futuro está associado a cada decisão operativa no período atual. Portanto, se ao deplecionar grandes volumes no presente, e por questões climáticas ocorrer baixas afluições no futuro, provavelmente as usinas hidrelétricas não serão capazes de atender toda a demanda de energia, podendo ocasionar um racionamento com elevados custos para a sociedade. Por outro lado, pode-se optar por preservar as reservas de água no presente, concentrando maior uso de geração térmica. Contudo, se as afluições futuras forem altas, pode ser necessário ter que verter o excesso de água, ou seja, desperdício de energia ocasionando um aumento desnecessário do custo operativo (GARCIA, 2005). A Figura 1.1 ilustra este comportamento.

Figura 1.1 - Processo de decisão de sistema hidrotérmicos.



A Função de Custo Imediato (FCI) representa os custos de operação em um determinado estágio t e, como mostrado na Figura 1.2, aumenta quanto maior for o volume armazenado, ou seja, quando mais água é poupada. Já a Função de Custo Futuro (FCF) está ligada ao custo de geração pelas usinas térmicas a partir do estágio $t+1$ até o final do período de planejamento. Esta função é calculada por meio de simulações do sistema para diferentes níveis de armazenamento no início do estágio $t+1$. Conforme pode ser visto, a FCF diminui à medida que aumenta o armazenamento final.

Figura 1.2 – Funções de custo futuro e imediato.



O ponto ótimo corresponde ao mínimo da soma das duas curvas, isto é, quando o módulo das derivadas da FCI e FCF se igualam. Nele, o valor da água se iguala ao custo de geração da termelétrica mais cara utilizada.

Portanto, o principal propósito do planejamento da operação energética é determinar a melhor estratégia de uso dos estoques de água. É necessário decidir para cada instante de tempo do horizonte de estudo qual é a vantagem em armazenar a água, comparando-a com mantê-la em estoque para aproveitamento futuro.

O problema de planejamento da operação de sistemas hidrotérmicos é bastante complexo. Por isso, é necessária sua decomposição em problemas menores de forma que possa solucioná-los a partir de modelos matemáticos mais simplificados. O problema principal é subdividido em outros problemas com diferentes horizontes de planejamento e níveis de detalhamento. Os resultados, no caso Brasileiro a FCF, encontrados nos problemas menos detalhados (com maior horizonte) servem como entrada para os mais detalhados (com menor horizonte de planejamento). A medida que o horizonte de estudo diminui, nota-se redução nas incertezas sobre as afluências futuras e é necessário um maior nível de detalhamento das usinas e da rede de transmissão. Vale lembrar que quanto maior o número de componentes considerados do sistema hidrotérmico, maior será o número de variáveis, coeficientes e equações dos modelos matemáticos definidos a partir de funções não lineares e multivariáveis de alta complexidade (ROSENTHAL, 1981).

Especificamente citando o caso Brasileiro, as subdivisões do problema de planejamento da operação energética são:

- **Planejamento da operação energética de médio prazo (PMP).**

É nesta etapa que se encontra o foco deste trabalho, em que o horizonte de estudo é de até cinco anos, discretizado em etapas mensais. Aqui, as usinas hidrelétricas são agregadas em Reservatórios Equivalentes de Energia (REE). Os resultados das simulações desta etapa geram informações futuras sobre o uso dos reservatórios (por meio das FCFs), além de informações diversas sobre a política operativa do sistema, como o montante energético a ser produzido pelas hidrelétricas, os níveis dos REEs e os riscos de não suprimento energético. Além disso, também são obtidas sinalizações econômicas para o mercado de energia elétrica, como as projeções de custo marginal de operação (CMO) de cada submercado. É de grande importância

a representação das condições hidrológicas do sistema e da quantidade de água a ser estocada nos reservatórios durante o horizonte de estudo. Isso quer dizer que a representação da característica estocástica das aflúências é mais importante que a modelagem das não linearidades da função de produção das usinas hidrelétricas e dos custos de geração das usinas térmicas. Além dessas, outras simplificações podem ser necessárias para a viabilidade da solução pelo fato de haver uma grande quantidade de UHEs, de número de estágios e pela característica estocástica das aflúências (DE MATOS; FINARDI; LARROYD, 2012).

- **Planejamento da operação energética de curto prazo (PCP).** No problema de curto prazo, o horizonte de estudo é de dois meses com discretização semanal no primeiro mês e mensal no segundo. A incerteza relacionada às aflúências é representada por uma árvore de vazões e as usinas são representadas de forma individualizada. A função objetivo do PCP é a mesma que o problema de médio prazo.
- **Programação Diária da Operação Eletroenergética (PDO).** Aqui o horizonte é de 7 a 14 dias, com discretização em etapas horárias ou de meia hora nas primeiras 48 horas. Todo o parque hidrotérmico é representado de forma detalhada, atentando para as restrições das unidades geradoras, segurança e confiabilidade da operação do sistema. Com a FCF gerada pelo modelo do PCP, a programação diária é usada para definir os níveis de geração para cada unidade geradora (MARCATO, 2002).

Conforme supracitado, o foco deste trabalho é o modelo do PMP, sendo que o horizonte de planejamento é de quatro anos e, portanto, tem-se um problema essencialmente probabilístico. A única fonte de incerteza considerada no modelo é aquela relacionada com as vazões afluentes. Contudo, outros fatores contêm incertezas tais como a demanda, disponibilidade dos equipamentos, custos incrementais das UTEs, atrasos na entrada de operação de equipamentos, etc.

O modelo do PMP é modelado matematicamente por um problema de programação linear estocástica (PL) com acoplamento temporal e espacial. O acoplamento temporal acontece devido a participação das usinas hidrelétricas, em que a disponibilidade de água no futuro depende das aflúências futuras e da operação atual do reservatório, por causa da restrição de armazenamento de cada usina. Já o acoplamento espacial ocorre

pelo fato de que a água liberada em uma usina para a produção de energia afeta a operação das usinas que estão à jusante na cascata (DE MATOS; FINARDI; LARROYD, 2012).

Dentre todas as estratégias disponíveis para resolver o problema do PMP, a mais utilizada é o uso de árvore de cenários de afluência. Utiliza-se essa estratégia, pois é computacionalmente inviável considerar todo o espaço amostral que é definido pelas Funções de Distribuição de Probabilidade (FDP) das afluências. As árvores representam uma aproximação do problema utilizando essa distribuição (DE MATOS; FINARDI; LARROYD, 2012), as quais foram inseridas em simulações de séries estocásticas por Thomas e Fiering em meados da década de 60 (THOMAS; FIERING, 1962). No entanto, apenas ao fim da década de 80 que suas aplicações ganharam destaque no âmbito de despacho hidrotérmico (BEZERRA, 2005).

Com relação a solução do problema de PL do PMP, inicialmente este era resolvido por meio da complementação de energia térmica requerida, tendo como base os resultados de simulações do sistema com as piores séries de vazões afluentes registradas. Apesar de simples, essa estratégia não era muito adequada, visto que a série testada era de comprimento pequeno e não levava em consideração a estocasticidade, gerando resultados sem nenhuma medida de risco.

As estratégias de solução foram se aperfeiçoando e a Programação Dinâmica Estocástica (PDE) se mostrou atraente, pois mantinha a característica estocástica das afluências. Porém, o algoritmo da PDE limitava seu uso para apenas um subsistema equivalente, devido à explosão combinatória do espaço de estados, já que o algoritmo demanda uma discretização do espaço de estados do problema. Nele, a decisão operativa depende de dois tipos de variáveis de estado: volumes armazenados nos reservatórios do sistema no início de cada período; e a “tendência hidrológica”, que é representada pelas vazões naturais afluentes de cada reservatório no mês anterior (SUANNO, 1995).

Na PDE, as decisões mensais são tomadas recursivamente no tempo, em que a política operativa é calculada em termos do montante total de armazenamento do sistema. Para este modelo, o esforço computacional é muito grande e cresce exponencialmente com o aumento do número de reservatórios, fenômeno conhecido como “maldição da dimensionalidade” da programação dinâmica, introduzido por Bellman (1961).

No entanto, para contornar essa dificuldade, no final da década de 90 (PEREIRA; PINTO, 1991) foi criada a Programação Dinâmica Dual Estocástica (PDDE), por meio da decomposição de Benders multiestágio. O uso da PDDE para a solução do problema obteve esse sucesso sobre a

PDE, pois utiliza a informação dual de um estágio para aproximar a FCF de cada estágio do problema. Para isso, utiliza conceitos da decomposição de Benders, sendo que cada aproximação da FCF é moldada por um corte de Benders. Então, com o aumento do número de iterações, mais cortes de Benders são adicionados à FCF, o que a deixa mais próxima da função real e a solução chega gradualmente mais próxima de sua solução ótima. Com este processo, a discretização em espaço de estados é contornada, já que a FCF é construída iterativamente e extrapolada no entorno das soluções encontradas.

Tanto para a PDE quanto para a PDDE, a estocasticidade do problema é representada pelas afluições futuras. Essa estocasticidade precisa ser configurada por meio de um modelo estatístico que represente mensalmente o comportamento sazonal de afluições em rios e bacias hidrográficas. Em problemas de planejamento energético, em que a incerteza hidrológica é grande, modelos de regressão linear são utilizados. No caso brasileiro, tal ação é feita com o uso de modelos autoregressivos periódicos de ordem p , ou modelos $PAR(p)$ (do inglês, *Periodic Autoregressive Model*). Esses modelos representam uma relação linear entre a vazão do estágio t com as vazões dos estágios anteriores ($t-1$, $t-2$, ..., $t-p$) (BEZERRA, 2005). Isso significa que a geração da árvore de cenários é realizada por meio de um modelo PAR (a ordem p varia de acordo com a usina hidrelétrica em questão) com valores encontrados com base no histórico de afluições mensais. Neste modelo, a afluição de um mês em questão é gerada com informações de p meses anteriores mais um resíduo correlacionado espacialmente. São esses resíduos que fornecem a aleatoriedade das afluições, já que a informações dos meses anteriores são conhecidas previamente (DE MATOS; FINARDI; LARROYD, 2012).

1.1 REVISÃO BIBLIOGRÁFICA

Esta seção tem como finalidade apresentar um breve histórico sobre o desenvolvimento das estratégias de solução para os problemas energéticos nos horizontes de curto e médio prazos. A barreira computacional tem sido sempre um fator complicador para o tratamento desses problemas. Por isso, ao longo dos anos, diversos autores criaram novas estratégias visando contornar ou minimizar este problema.

Sistemas com múltiplos reservatórios exigem decisões operativas simultâneas para cada reservatório, e para isso centenas ou milhares de variáveis de decisão são introduzidas no problema de otimização. Essa grande quantidade de variáveis eleva a complexidade da otimização, visto

que o espaço de estados cresce exponencialmente com o número das variáveis. Por isso, a redução da dimensionalidade dos problemas com múltiplos reservatórios se torna fundamental.

Neste contexto, Turgeon (1981) elaborou um método para a solução do problema com n reservatórios em série. Tal estratégia gerava $n-1$ problemas de duas variáveis de estado a serem resolvidos com a Programação Dinâmica (PD). Mesmo com o agravante que a correlação serial não fosse considerada, a técnica se mostrou atraente porque o tempo aumentava apenas linearmente com a adição do número de reservatórios, dado que apenas um problema de PD de duas variáveis de estado necessitava ser resolvido com a adição de um novo reservatório no sistema.

Na literatura, diversos trabalhos utilizam a PDE considerando a agregação de reservatórios para a solução do problema de médio ou curto prazos e uma consequente desagregação para quantidades obtidas em cada usina. Esses modelos são um equilíbrio entre o nível de detalhamento e a facilidade do uso e solução dos problemas, principalmente os de grande porte (ROGERS et al., 1991). Assim, um modelo de otimização de grande porte é substituído por um menor e mais manuseável, essencial para problemas em que a capacidade computacional é limitada. Dentre os trabalhos que utilizam métodos de agregação/desagregação, destacam-se Turgeon e Charbonneau (1998) e Valdes *et al.* (1992).

Neste cenário, uma ideia interessante foi desenvolvida por Saad *et al.* (1994) que, a partir de reservatórios equivalentes, propuseram uma estratégia efetiva de desagregação com o uso de redes neurais. Nele, os autores exibiram um método de desagregação não linear gerado pela rede neural baseado no “treinamento” de modelos de otimizações determinísticas.

Saad e Turgeon (1988) elaboraram um eficiente método utilizando a Análise de Componentes Principais (ACP) ou PCA (do inglês, *Principal Component Analysis*) a fim de modelar o problema de PDE com um número reduzido de variáveis de estado. Esta redução é baseada no nível de interdependência das variáveis e quanto maior essa dependência, melhor era a redução. Aplicado em um sistema de cascata com 5 reservatórios, a técnica se mostrou vantajosa, já que o problema original de 10 variáveis de estado foi reduzido para um com apenas quatro.

Em estudos realizados por Fu, Kapelan e Reed (2012), a Análise da Sensitividade Global foi a ferramenta utilizada para a redução da necessidade computacional. Embora não tendo sido usado especificamente no problema de planejamento energético, o uso das variáveis de decisão mais sensíveis reduziu radicalmente o tempo requerido. Johnson *et al.* (1993) usou uma técnica de funções polinomiais cúbicas por partes para

aproximar a função na solução numérica da programação dinâmica para reduzir o tempo computacional.

Todos trabalhos citados diminuem o número de variáveis do problema, e conseqüentemente sua complexidade. Isso pode fazer com que a precisão do processo de otimização seja reduzida. Por isso, Chen *et al.* (2016) propôs a transformação das variáveis de decisão do domínio do tempo para o domínio da frequência, sem a decomposição ou agregação do sistema. Neste procedimento é utilizado o Modelo de Otimização Espectral ou SOM (do inglês, *Spectral Optimization Model*) junto de um método de redução de dimensionalidade chamado expansão de Karhunen-Love (KL), que é truncado com poucos termos significativos. Em seus resultados, para um problema de 3360 variáveis de decisão, o modelo SOM com apenas cinco termos KL exibiu uma performance similar quando comparado ao método de solução convencional.

Também devido às restrições de tempo computacional e da grande dimensionalidade do problema completo, programas estocásticos contínuos no tempo muitas vezes necessitam que os estágios de decisão sejam agregados transformando problemas com muitos estágios para problemas com apenas poucos períodos (KUHN, 2008).

Já com a utilização da PDDE, algoritmo que não é afetado pela “maldição da dimensionalidade”, a busca pela redução do esforço computacional continua, visto que a árvore de cenários e os cenários percorridos são fatores determinantes para a qualidade da solução do problema. Com relação a esse ponto, foi proposta uma nova metodologia para a seleção da amostra do processo multivariado referente a possíveis influências futuras (DA COSTA; DE OLIVEIRA; LEGEY, 2006). A avaliação de métodos de redução de cenários também foi aplicada em pesquisas de Larsen, Doorman e Mo (2015). Em ambos os casos os resultados foram satisfatórios, mostrando que o problema pode ser solucionado em uma fração do tempo ao que seria necessário se fosse usado o problema original.

Para Dupačová e Kozmík (2016), de acordo com estudos realizados recentemente, a redução de cenários na etapa de *backward* da PDDE foi a investida para a redução do tempo computacional. Com pouco esforço, os procedimentos de solução existentes foram melhorados, visto que mostrou reduções significativas dos tempos computacionais mantendo uma boa aproximação com os casos originais. Adotando 1.000 cenários por estágio como sendo do problema original, a redução do tempo para 500 foi de 48%, enquanto que apenas 13% do tempo do caso base foi utilizado no caso de 100 cenários por estágio, sempre com resultados aceitáveis em relação ao original.

Heitsch e Römisch (2014) apresentaram um algoritmo que seleciona e remove nós da árvore de cenários com uma tolerância de erro pré-estabelecida. De Oliveira *et al.* (2008) empregaram uma metodologia chamada Redução Ótima de Cenários ou SOR (do inglês, *Scenario Optimal Reduction*) que minimiza a distância entre as medidas probabilísticas, aquelas relacionadas à diferentes árvores de cenários, para encontrar o conjunto ótimo de sequências sintéticas que representa a incerteza suficientemente. A métrica *Fortet-Mourier* foi usada para definir o problema SOR e selecionar, a partir de uma grande árvore de cenários, um subgrupo que reduza o tempo de otimização. Dentre os trabalhos que focam em um bom método de redução de cenários, pode-se citar Casey e Sen (2005) e Hochreiter e Pflug (2007).

O algoritmo de Decomposição Aninhada (DA) é de solução exata, já que é utilizada para resolver o problema relacionado à toda árvore (DE MATOS, 2012). Donohue e Birge (2006) incorporaram um processo amostral sobre a DA, chamando de Decomposição Aninhada Abreviada ou AND (do inglês, *Abridged Nested Decomposition*). Em seus resultados, o método superou a forma de solução tradicional com tempos computacionais cerca de 12 vezes menores.

Hindsbeger (2014) incorporou junto a PDDE e decomposição de Benders algumas das ideias do AND para um esquema de amostragem adequado, criando o que chamou de *Reduced Sampling* ou ReSa. Os testes foram simulados para o sistema hidrotérmico do sistema de potência nórdico e os resultados se mostraram melhores do que quando resolvidos por meio da PDDE e ligeiramente frente ao AND. Vale lembrar que os resultados encontrados pelo ReSa foram feitos para problemas com número reduzido de etapas e poucos valores aleatórios por estágio em comparação com o que é adotado no caso brasileiro.

1.2 OBJETIVOS DO TRABALHO

Este trabalho visa encontrar políticas operativas para o problema do PMP utilizando técnicas de ACP para reduzir o espaço de estados da PDDE associado com as vazões afluentes incrementais. O objetivo principal é analisar a qualidade das políticas fornecidas por essa estratégia em relação ao modelo utilizado atualmente no setor, isto é, a representação direta no algoritmo da PDDE do modelo PAR das aflúncias.

Em termos de modelagem, o problema do PMP implementado nesta dissertação representa a Função de Produção Hidrelétrica (FPH) das hidrelétricas por meio de aproximações lineares por partes para manter a convexidade do problema. Apesar de ser uma aproximação, essa forma

de representação é mais precisa quando comparada aos modelos com produtividade constante (FREDO, 2016). Ela é baseada na técnica de *Convex Hull* (ou Envoltória Convexa, em português).

O processo de geração de séries sintéticas de aflúências foi baseado no modelo desenvolvido por Larroyd (2016) e empregados no modelo de médio prazo individualizado, de forma que a árvore de cenários a ser resolvida via PDDE represente eficientemente as tendências hidrológicas passíveis de ocorrência no sistema estudado.

A técnica de ACP é um propósito de estudo e será introduzido no problema. A ACP é utilizada no processo de geração das séries sintéticas e incorporado no algoritmo da PDDE. O apelo para o uso de ACP é a redução do número de variáveis de estado do problema sem perda significativa do conjunto de dados originais de aflúências.

Assim, no problema de PMP utilizado neste trabalho, o modelo $PAR(p)$ será aplicado em:

- (i) nas aflúências incrementais das hidrelétricas (caso base);
- (ii) nas componentes principais (CP) estipuladas.

As variáveis aleatórias de ambos os modelos serão introduzidas na árvore de cenários da PDDE e as duas metodologias serão comparadas, sendo que para a ACP, a utilização de números reduzidos de componentes principais também será avaliada. O intuito é encontrar avanços sobre o caso base pela redução do esforço computacional ou melhora da qualidade da política operativa calculada.

1.3 ESTRUTURA DA DISSERTAÇÃO

Este trabalho está estruturado conforme descrito a seguir. Inicialmente, no Capítulo 2 serão apresentados os principais componentes do sistema hidrotérmico. Primeiramente é feita uma breve explicação sobre as unidades térmicas e, na sequência, sobre as hidrelétricas, nas quais são explicadas as equações matemáticas envolvidas na produção de energia elétrica. Ainda neste capítulo é exposto o mecanismo de representação de aflúência. Por fim, é mostrada a análise para a aplicação da técnica *Convex Hull* para desenvolver aproximações lineares em torno da FPH.

Na sequência, o Capítulo 3 trata da abordagem do modelo de geração de aflúências utilizados neste trabalho. Além da estimação dos parâmetros, é abordado o modo de representação por hidrelétricas e a correlação espacial existente no sistema.

O Capítulo 4 apresenta a formulação geral do algoritmo da PDDE na representação por UHEs para se obter a política operativa do sistema.

Os conceitos da decomposição de Benders no problema também são discutidos. Um exemplo de pequeno porte é introduzido a fim de explicar como o processo de solução do problema do PMP é feito. No final, é abordado a forma de obter o custo de operação do sistema e a estimação da qualidade da solução encontrada.

Por sua vez, o Capítulo 5 tem o objetivo de explicar de forma didática o conceito da técnica de ACP que será aplicada no problema do PMP. O capítulo trata sobre aspectos básicos dessa técnica e finaliza com um exemplo numérico correspondente a aquele apresentado no capítulo anterior.

No Capítulo 6, todos os conceitos anteriormente abordados são aplicados em um sistema hidrotérmico onde o recurso de geração hidrelétrica é dado pelas usinas das cachoeiras dos rios Uruguai e Iguaçu. Os resultados referentes às políticas operativas encontradas são apresentados e analisados, sempre comparando as respostas dos modelos com CP com o modelo sem essas componentes.

Finalmente, no Capítulo 7 são sintetizadas as conclusões dos resultados obtidos. As principais contribuições da proposta estudada também serão identificadas. Além disso, são propostas algumas diretrizes para trabalhos futuros.

2. COMPONENTES DO SISTEMA HIDROTÉRMICO

Este capítulo é referente à modelagem do problema de PMP. Inicialmente é feito um breve resumo sobre as unidades térmicas. Após isso, será mostrado como é expressa a função de produção de uma usina hidrelétrica. Sua utilização neste documento será feita por meio da agregação de todas as unidades que fazem parte da mesma UHE, formando assim uma unidade geradora equivalente. Os mecanismos de como será feita tal agregação será discutida neste capítulo, bem como sua formulação junto aos outros componentes do sistema hidrotérmico conforme a proposta do planejamento de médio prazo brasileiro.

Vale ressaltar que uma representação mais detalhada para o problema no âmbito de médio prazo não é computacionalmente viável devido ao grande número de elementos do sistema, tamanho do horizonte estudado e da natureza estocástica das afluições. Conforme poderá ser visto na sequência, todo comportamento não linear relativo a um determinado componente será abordado como um conjunto de aproximações lineares.

Além disso, será apresentado um breve resumo do comportamento das afluições das UHEs obtidos a partir do histórico e como tais informações são tratadas para poderem ser incorporadas no problema de planejamento.

2.1 USINAS TERMELÉTRICAS

As usinas termelétricas são constituídas de unidades geradoras que utilizam combustíveis fósseis, gás natural, biomassa ou fissão nuclear para a produção de energia. O calor utilizado na queima do combustível é utilizado para gerar vapor, que por sua vez movimentam as turbinas que estão acopladas no eixo do gerador.

De forma geral, o custo de produção de uma termelétrica é expressa por um polinômio de segunda ordem em função da quantidade de combustível utilizada durante os processos de partida/parada e operação nominal da unidade. Além disso, sua operação deve levar em consideração os limites de potência de saída, tempo mínimo e máximo de operação, bem como rampas de aquecimento e resfriamento (FINARDI, 1999). Com exceção aos limites de potência, as demais restrições descritas são desconsideradas, pois lidam com tempos muito curtos como horas ou dias, e o planejamento de médio prazo trata de decisões mensais. Além disso, para melhor aplicabilidade nos estudos relacionados a este trabalho, a representação do custo das térmicas será expressa na forma de uma ex-

pressão de primeira ordem em função de sua quantidade de potência gerada, as quais já incluem informações de gastos com partida e parada das unidades. Assim, o custo de uma unidade térmica é definido neste trabalho pela expressão:

$$FC_{CFt}^j(pT_t^j) = cT^j \cdot pT_t^j, \quad (2.1)$$

em que:

cT^j	Custo variável unitário da usina T (R\$/MWh);
pT_t^j	Potência gerada na usina T e estágio t (MWh);
FC_{CFt}^j	Custo de produção da usina T e estágio t (R\$).

2.2 USINAS HIDRELÉTRICAS

As usinas hidrelétricas são construções que aproveitam a água e os desníveis dos rios para gerar energia elétrica, buscando o melhor aproveitamento energético de acordo com as características hidrográficas e topológicas da região em questão (DE SOUZA SILVA, 2010). Os volumes operacionais dos reservatórios das hidrelétricas variam entre o volume máximo normal e o volume mínimo operativo. O volume máximo é definido por questões de construção e de segurança do reservatório e o volume mínimo pela menor queda de água possível para a geração de energia nas turbinas (SCARCELLI, 2012).

2.2.1 Função de Produção Hidrelétrica

A potência gerada de uma unidade geradora depende da altura líquida da queda d'água. Esta, por sua vez, é consequência de três fatores: altura de montante, no caso brasileiro definida por um polinômio de quarto grau em função do volume de montante; altura de jusante, também definida por um polinômio de quarto grau, mas dessa vez dependente do volume turbinado e vertido; e as perdas hidráulicas. A diferença entre altura montante e altura jusante é chamada de queda bruta. As perdas hidráulicas representam o efeito do atrito entre o material do conduto forçado e a massa d'água. A queda líquida pode ser definida por (2.2).

$$hl = fcm(v) - fcj(q, s) - ph(u), \quad (2.2)$$

em que,

$fc_m(v)$	Função cota de montante, em metros, dependente do volume armazenado no reservatório v (hm^3);
$fc_j(q, s)$	Função cota de jusante, em metros, dependente da vazão turbinada da usina q (m^3/s) e do vertimento s (m^3/s);
u	Vazão turbinada na unidade geradora (m^3/s);
$ph(u)$	Perdas hidráulicas na unidade geradora (m).

A função cota de montante é expressa pelo seguinte polinômio:

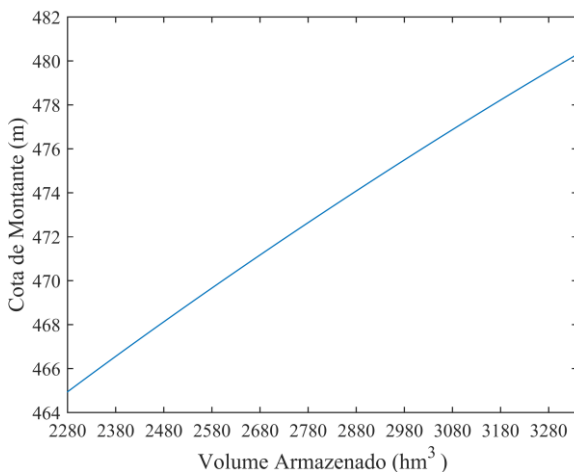
$$fc_m = HM_0 + HM_1 \cdot v + HM_2 \cdot v^2 + HM_3 \cdot v^3 + HM_4 \cdot v^4, \quad (2.3)$$

em que:

HM_0, \dots, HM_4 coeficientes do polinômio para cada usina.

De forma ilustrativa, a cota de montante em função do volume do reservatório da UHE de Machadinho é ilustrada na Figura 2.1.

Figura 2.1 - Função de Cota Montante da UHE de Machadinho.



A função cota de jusante é dada por:

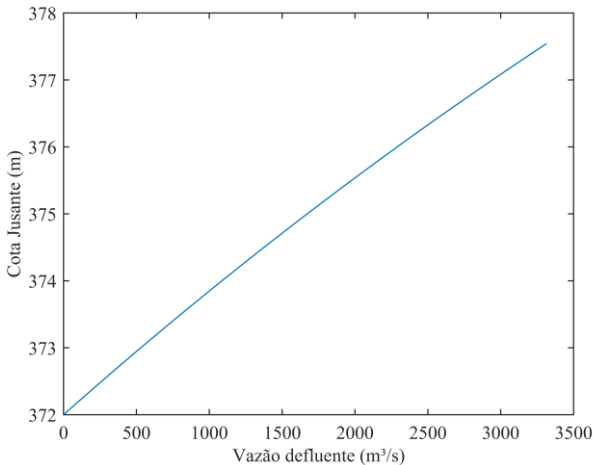
$$fcj(q, s) = HJ_0 + HJ_1 \cdot (q + s) + HJ_2 \cdot (q + s)^2 + HJ_3 \cdot (q + s)^3 + HJ_4 \cdot (q + s)^4, \quad (2.4)$$

sendo:

HJ_0, \dots, HJ_4 coeficientes da função cota jusante para cada usina.

A FCJ tem características próprias dependendo da usina, já que depende da vazão defluente e seus coeficientes são especificados de acordo com as características do leito defluente do rio, canal de restituição e vertedouro (LARROYD, 2016). Além disso, há alguns casos em que uma determinada usina possui mais de um polinômio porque pode depender também do volume do reservatório à jusante. Este é um caso que ocorre em minoria entre as UHEs pertencentes ao SIN, e caso haja usinas com tais características neste trabalho, serão consideradas apenas uma FCJ. A Figura 2.2 ilustra a FCJ da unidade de Machadinho.

Figura 2.2 - Cota Jusante da UHE de Machadinho.



Já as perdas hidráulicas são estimadas por meio de características de projeto das unidades geradoras. Elas são obtidas pelo produto do quadrado da vazão turbinada com uma constante característica do conduto forçado. Dessa forma, a expressão matemática da queda líquida é:

$$hl = HM_0 + HM_1 \cdot v + HM_2 \cdot v^2 + HM_3 \cdot v^3 + HM_4 \cdot v^4 - HJ_0 - \quad (2.5)$$

$$-HJ_1 \cdot (q+s) - HJ_2 \cdot (q+s)^2 - HJ_3 \cdot (q+s)^3 + HJ_4 \cdot (q+s)^4 - CH \cdot u^2,$$

em que:

CH Constante de perda hidráulica da unidade
(s²/m²).

Há perdas também na conversão eletromecânica do potencial hidráulico da água em energia elétrica. Estas perdas são agrupadas e representadas na produção hidrelétrica por meio de um rendimento global turbina-gerador obtido pelos fabricantes a partir de características de projeto e operação.

O rendimento hidráulico de uma turbina depende da vazão turbinada e queda líquida, sendo sua representação gráfica similar a uma curva colina. Neste trabalho, a aproximação do rendimento será por meio de um polinômio de segunda ordem, conforme mostra a equação (2.6).

$$\eta(u, hl) = R_0 + R_1 \cdot u + R_2 \cdot u^2 + R_3 \cdot hl + R_4 \cdot hl^2 + R_5 \cdot u \cdot hl, \quad (2.6)$$

em que:

R₀, ..., R₅ Coeficientes da função do rendimento.

Por fim, a FPH de uma unidade geradora que representa o comportamento da potência de saída pode ser encontrada a partir das características abordadas anteriormente. Ela é obtida pelo produto da vazão turbinada, queda líquida e rendimento turbina-gerador, conforme mostra a equação (2.7). Devido à natureza essa função é não linear e não convexa, o que torna um empecilho ao representá-la no PMP por meio da PDDE. Porém, mecanismos que fazem uma aproximação convexa da FPH serão utilizados e mais detalhes serão expostos no decorrer do trabalho.

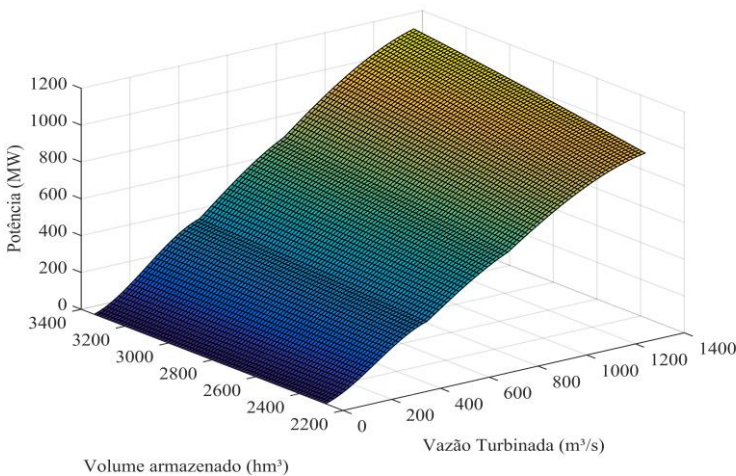
$$p = 9,81 \cdot 10^{-3} \cdot \eta(u, hl) \cdot hl \cdot u, \quad (2.7)$$

sendo:

p	Potência ativa da unidade geradora (MW);
9,81	Valor da aceleração da gravidade;
10^{-3}	Valor do peso específico da água.

A título de exemplo, a Figura 2.3 apresenta o comportamento da função de produção da usina de Machadinho. Note que a representação é agregada, isto é, considera-se a representação de uma única unidade equivalente (essa UHE possui três unidades idênticas). Nela, é possível observar as não linearidades no processo de geração de energia elétrica.

Figura 2.3 - FPH UHE de Machadinho (s=0).



2.2.2 Balanço Hídrico nos Reservatórios

O volume do reservatório de uma usina hidrelétrica depende de seu estado inicial, da vazão afluente que chega na usina e da defluência. Dessa forma, caso seja utilizada menos água para a produção de energia elétrica do que a quantidade afluente registrada em um mês, essa diferença pode ser utilizada para aumentar o nível do reservatório para utilizações futuras. Se o reservatório já está em sua capacidade máxima, este excesso

deve ser direcionado à jusante do rio pelos vertedouros. O vertimento pode ser necessário por outros motivos como, por exemplo, geração na UHE à jusante para atendimento da ponta. É bom salientar que no caso real, um rio pode conter inúmeras hidrelétricas em seu curso e, portanto, a vazão defluente em uma usina influencia no montante d'água da usina subsequente desta cascata. Como os períodos t são mensais, ressalta-se que a vazão defluente de uma usina m à montante de um determinado mês atinge a usina h ainda no mês vigente. Assim, desprezando-se os efeitos de evaporação e infiltração no reservatório, a equação de balanço hídrico é expressa da seguinte forma:

$$v_{t+1}^h - v_t^h + c \left(q_t^h + s_t^h - \sum_{m \in M_h} (q_t^m + s_t^m) \right) = c \cdot y_t^h, \quad (2.8)$$

$$h = 1, \dots, NH,$$

$$t = 1, \dots, NT,$$

em que:

M_h	Conjunto de UHEs localizadas imediatamente à montante da UHE h ;
m	Índice associado ao conjunto de M_h ;
v_{t+1}^h	Volume armazenado no reservatório da UHE h ao fim o estágio t (hm^3);
q_t^h	Vazão turbinada pela UHE h durante o estágio t (m^3/s);
q_t^m	Vazão turbinada pela m -ésima UHE do conjunto de M_h durante o estágio t (m^3/s);
s_t^h	Vazão vertida pela UHE h durante o estágio t (m^3/s);
s_t^m	Vazão vertida pela m -ésima UHE do conjunto de M_h durante o estágio t (m^3/s);
c	Fator de conversão da unidade de vazão (m^3/s) para a unidade de volume (hm^3) considerando a duração do estágio;

y_t^h	Vazão afluente incremental ao reservatório da UHE h no estágio t (m^3/s);
NH	Número de hidrelétricas h ;
NT	Número de estágios t .

2.3 REPRESENTAÇÃO DAS AFLUÊNCIAS

Conforme destacado anteriormente, as afluências incrementais representam a incerteza na solução do PMP. Algumas características relativas às vazões incrementais das UHEs obtidas por meio do histórico das afluências mensais médias são apresentadas nesta seção. Além disso, será mostrada a forma que estas estarão inseridas no algoritmo da PDDE. O processo aqui apresentado segue o que foi usado por Larroyd (2016).

O histórico de afluências utilizado nesta dissertação está disponível no site do Operador Nacional do Sistema Elétrico (ONS), sendo 1931 o ano inicial. Nele, as afluências de cada hidrelétrica estão na unidade de m^3/s e como são afluências naturais que chegam na usina h , o valor de afluência da usina diretamente à montante está somada com a afluência incremental da UHE em questão. O que se deve empregar na equação de balanço hídrico de uma usina h é sua afluência incremental yh_t^h , isto é, a diferença da vazão proveniente de chuvas, nascentes e rios afluentes com o montante evaporado infiltrado ou desviado do curso natural do rio. É utilizada a afluência incremental porque o interesse está no balanço hídrico de cada usina.

Tendo em conta o que foi mencionado anteriormente, o modelo lida apenas com a afluência incremental da usina h . Por isso é necessário retirar a quantia referente a usina diretamente à sua montante m , exceto quando a UHE h é a primeira usina da cascata (ou seja, não possui usina à sua montante). Assim sendo:

$$yh_t^h = yn_t^h - \sum_{m \in \mathcal{M}^h} yn_t^m, t \in \mathcal{T}_-, \quad (2.9)$$

em que:

yn_t^h	Afluência natural da hidrelétrica h em t (m^3/s);
\mathcal{T}_-	Conjunto de estágios t com registro de afluência natural.

É possível, mesmo esporadicamente, que yh_t^h assuma valores negativos devido ao caráter físico da afluência natural. Isso, por sua vez, se

transforma em uma dificuldade para a estimação dos parâmetros do modelo PAR usados para gerar os modelos de geração de aflúências incrementais para serem aplicados na PDDE. Informações mais precisas sobre tal procedimento irá ser abordado no Capítulo 3.

Para que esses valores negativos de aflúências incrementais sejam eliminados, a estratégia é a inserção de uma nova variável para o problema yp_t^h , denominada de aflúência do processo estocástico. Ela possui todos seus valores positivos e é a partir dela que o modelo PAR é aplicado. O processo para a determinação de yp é o seguinte: computa-se o menor valor de aflúência incremental de uma usina h presente no histórico em todo t . Caso este valor seja nulo ou inferior a zero, isto é, $yh_t^h \leq 0$, seu módulo acrescido de 1 estabelece o grau de liberdade do histórico incremental $\delta y h^h$. O valor do grau de liberdade é nulo caso não haja valores negativos de aflúência incremental na hidrelétrica h . Após determinado o grau de liberdade do histórico incremental, yp_t^h pode ser definido como:

$$yp_t^h \triangleq yh_t^h + \delta y h^h, t \in \mathcal{T}_-. \quad (2.10)$$

A Figura 2.4 e a Figura 2.5 ilustram o mecanismo feito e sua comparação com as aflúências incrementais para a UHE de Salto Osório. Elas exibem todos os valores de aflúência da série histórica de acordo com dados do ONS e o gráfico de caixas (Box-Plot) com separação mensal respectivamente.

Figura 2.4 - Registro de aflúências mensais de Salto Osório, yh e yp .

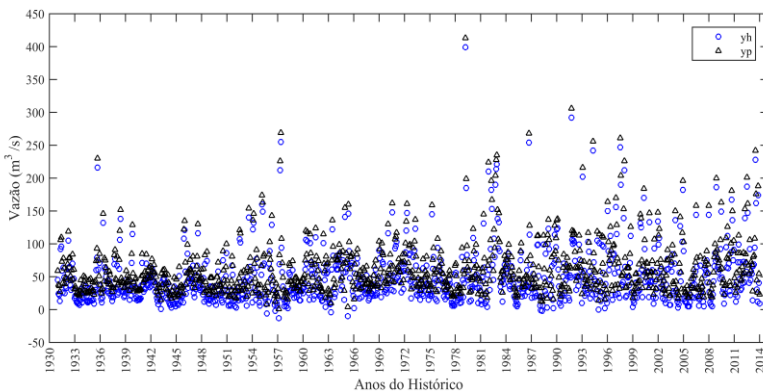
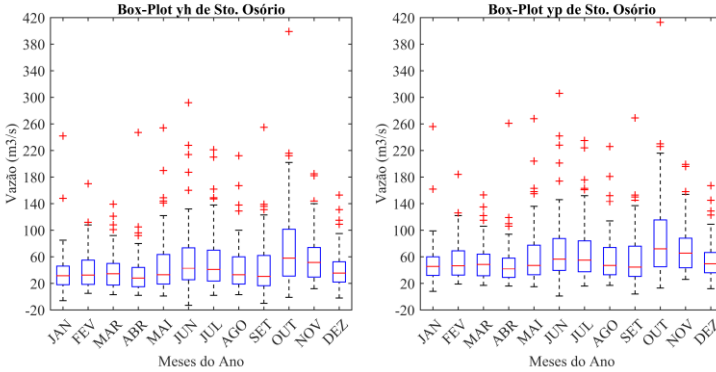


Figura 2.5 - Comparação das afluições da usina de Salto Osório.



Nota-se que a estatística do processo não é alterada, apenas há um deslocamento na ordem do grau de liberdade do histórico incremental δy_t^h .

Os valores de y_t^h que necessitam ser utilizados na equação de balanço hídrico são obtidos da seguinte forma:

$$y_t^h \triangleq y_p^h - \delta y_t^h. \quad (2.11)$$

2.4 LINEARIZAÇÃO DA FPH

O algoritmo para a solução do PMP, a PDDE, faz uma construção aproximada do custo futuro esperado em cada período de planejamento por meio de funções lineares por parte. Para isso, necessita-se que todas as equações envolvidas formem um conjunto convexo. Como foi possível ver, a FPH de uma usina hidrelétrica é do tipo não linear e não convexa, o que torna uma complicação neste tipo de abordagem.

Por isso, o conceito de *Convex Hull* (CH) é utilizado neste trabalho. Aqui, irá ser abordado o conceito de forma ampla, sendo que na literatura é possível encontrar informações adicionais como em Fredo (2016). O *Convex Hull* é usado para construir de maneira eficaz aproximações lineares com duas ou três variáveis, dependendo do tipo da usina hidrelétrica. Essas aproximações são inseridas na PDDE como desigualdades lineares do tipo menor igual. A seguir é apresentado o procedimento para gerar as aproximações lineares da FPH, que também tem como base os estudos apresentados por Diniz e Maceira (2008).

Inicialmente são escolhidos p pontos equidistantes entre os valores mínimo e máximo das variáveis Q (vazão turbinada), V (volume) e S (vertimento) para as UHEs. O volume é usado apenas nas usinas com reservatório. Isso gera uma grade de p^3 pontos no plano $Q \times V \times S$. Caso a usina seja fio d'água e tiver influência do vertimento, apenas Q e S são utilizados. Esses pontos, em conjunto do respectivo valor de potência, servem como dados de entrada para a função *convhulln* no *software Matlab*. Esta função retorna os planos que fazem parte da envoltória convexa.

Os planos escolhidos pelo CH são, matematicamente, expressos por:

$$p + PH_{0e} + PH_{1e} \cdot q + PH_{2e} \cdot v + PH_{3e} \cdot s \geq 0, \quad (2.12)$$

$$e = 1, \dots, NE,$$

sendo:

$PH_{0e}, PH_{1e}, PH_{2e}$ Coeficientes dos planos obtidos no CH;
 NE Número de planos obtidos no CH.

Nem todos os planos que a função retorna devem ser utilizados, por isso uma filtragem desses planos é feita por meio do algoritmo proposto por Diniz e Maceira (2008). O algoritmo que seleciona os planos que formam a envoltória superior da FPH é dado por:

Faça : $e = 1, \dots, NE$

Faça : $i = 1, \dots, NI$

Calcule $ph(v^i, q^i, s^i)$ e $ph^e(v^i, q^i, s^i)$;

Fim

Se : $\forall i \ ph^e(v^i, q^i, s^i) \geq ph(v^i, q^i, s^i)$ (2.1)

O plano $e \in \mathcal{Z}$;

$NZ = NZ + 1$;

Fim

Fim,

em que:

NI	Número de pontos pertencentes a algum plano do CH;
ph	Potência gerada de um ponto (v^i, q^i, s^i) pertencente a algum plano do CH;
ph^e	Potência gerada no ponto (v^i, q^i, s^i) do plano a ser testado;
Z	Conjunto de planos selecionados.

Por fim, a representação matemática dos planos fornecidos pelo CH possui a seguinte forma:

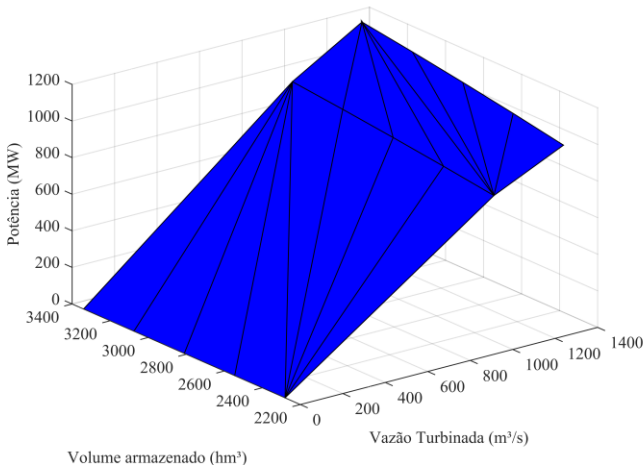
$$p + PH_{0z} + PH_{1z} \cdot q + PH_{2z} \cdot v + PH_{3z} \cdot s \geq 0, \quad (2.13)$$

$$z = 1, \dots, NZ.$$

Para as UHEs que não possuem reservatório, o mesmo procedimento é feito. A diferença é que a variável de volume armazenado v é ignorada.

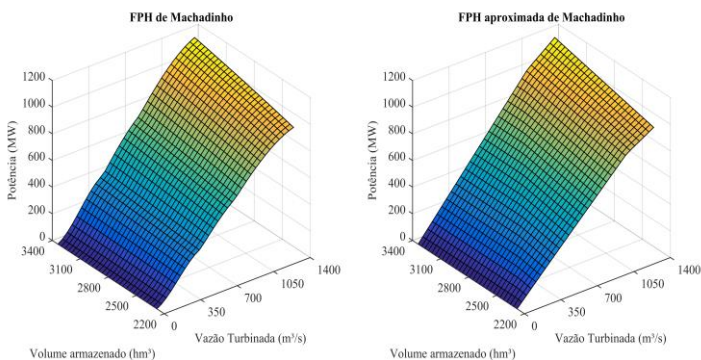
A Figura 2.6 ilustra os planos resultantes após a aplicação do CH e consequente execução do algoritmo (2.1). Lembrando que, para fins de representação, neste exemplo, a vazão vertida foi ignorada. O número de discretizações em V e Q foi de cinco.

Figura 2.6 - FPH por *Convex Hull* UHE de Machadinho.



Uma vez encontrados os planos que representam a FPH da UHE, é possível ilustrar a FPH não linear com sua versão linear por partes (Figura 2.7). A análise de precisão desta é feita na seção seguinte.

Figura 2.7 - Dois modelos da FPH de Machadinho (s=0).



2.4.1 Precisão do modelo linear por partes

Para que as FPHs sejam inseridas no PMP é preciso classificar quantitativamente seu nível de precisão. Isso porque, a quantidade de pontos discretizados em cada uma das variáveis para aplicar o CH não precisa ser necessariamente o mesmo, isto é, a quantidade de pontos aplicados para uma usina pode não ser tão interessante se for aplicada para outra.

À medida que um número maior de pontos é aplicado ao CH, maior é o número de hiperplanos, ou seja, aquela hidrelétrica terá mais restrições para representar sua FPH. Este fato não significa necessariamente que o erro em relação à FPH não linear original seja menor. Isto posto, esta seção visa mostrar a heurística utilizada a fim de encontrar o melhor número de pontos para cada variável em que o CH vai ser aplicado. O critério que irá mensurar a precisão no presente trabalho será o erro médio quadrático (HYNDMAN; KOEHLER, 2006) que matematicamente é expresso por (2.14).

$$EMQ = \frac{\sqrt{\sum_{g=1}^{NG} (FPH_g - FPHa_g)^2}}{\sqrt{NG}}, \quad (2.14)$$

em que:

EMQ	Erro médio quadrático (MW);
g	Índice associado aos pontos utilizados na discretização da grade de pontos para o cálculo do EMQ;
FPH _g	Valor da FPH não linear original avaliada no ponto operativo g (MW);
FPHa _g	Valor da FPH linear por partes avaliada no ponto operativo g (MW).

Esta heurística deve ser utilizada, pois aplicar o CH para uma usina com um número elevado de pontos, pode resultar em centenas ou até milhares de funções lineares por partes. Isso não é desejável visto que, como já mencionado, os hiperplanos são restrições no algoritmo da PDDE e, portanto, quanto maior o número de restrições, mais onerosa se torna a solução do PMP. Por outro lado, se esse número de pontos for pequeno demais, pode ser que apesar de resultar em uma pequena quantidade de planos, a representação da função original não linear da UHE não seja satisfatória. Por esse motivo, é útil que o número de planos e a precisão (baseada no EMQ) da FPH linear por partes encontrada sejam levados em consideração.

Assim sendo, para as hidrelétricas que a FPH dependa apenas de duas variáveis (como Q e S , por exemplo), a heurística se baseia em adotar um número p mínimo de 2 e máximo de 15 pontos equidistantes para cada uma das variáveis. Isso irá fornecer uma grade mínima de 4 (2^2) e máxima de 225 (15^2) pontos nos casos a serem executados. Seja nq o número de discretizações em Q , nv em V e ns em S , para cada uma das 196 combinações resultantes de nq e ns , isto é, nq e ns variando de 2 a 15, é executado o comando *convhulln* do *Matlab*. Em seguida o algoritmo (2.1) é aplicado e conseqüentemente computado o valor de EMQ. O número de discretizações em Q e em S empregado é aquele que fornecer o menor valor de EMQ.

Já para as UHEs em que a FPH depende de Q , V e S , o procedimento é um pouco diferente. Isto porque, quando se aplica a função *convhulln* nas três variáveis, a sensibilidade diante do aumento do número de discretizações principalmente sobre as variáveis Q e V é muito grande. Isto pode fazer que com o acréscimo do número de discretizações nessas dimensões o número de planos aumente consideravelmente e, se mais preciso, a diferença do EMQ para um caso com menos planos pode ser irrelevante.

Assim sendo, o processo é o mesmo para o caso com apenas as variáveis Q e S , com a diferença que 10 é o número máximo de pontos equidistantes para cada variável. Agora a grade de pontos é $nq \times nv \times ns$ (1000 pontos) e fornece um total de 729 combinações de nq , nv e ns (cada um varia de 2 a 10). Após rodar a função *convhulln* e aplicar o algoritmo (2.1) todos os 729 valores de EMQ são calculados. O caso com menor EMQ (base) e aqueles com valores de EMQ até 5% superior ao caso base são selecionados. Destes, o caso que apresentar o menor número de hiperplanos será o escolhido para representar a FPH da UHE em questão.

Para exemplificar o processo, será aplicada como o exemplo a usina de Campos Novos. A Tabela 2.1 apresenta um exemplo de 40 combinações de nq , nv e ns com seus respectivos erros e número de hiperplanos. Como a variável S possui menor significância na FPH, ela foi simulada até o valor de 5 discretizações. Destes, está assinalado em verde o caso com o menor erro, em amarelo os casos com EMQ até 5% superior ao menor global e por fim em azul a combinação escolhida. O EMQ desta tabela é dado em percentual com relação à máxima potência da unidade hidrelétrica.

Tabela 2.1 - Convex Hull da FPH de Campos Novos.

Discretizações em Q	Discretizações em V	Discretizações em S	EMQ (%)	Número de Hiperplanos
2	2	2	3,46316	7
2	2	5	3,46316	7
2	6	5	3,45954	24
2	10	5	3,45942	40
3	2	2	2,14594	14
3	5	2	2,14376	49
3	5	4	2,14376	49
3	10	2	2,14364	103
3	10	3	2,14364	106
4	4	3	1,50303	37
4	4	4	1,50303	39
5	2	3	1,21100	28
5	5	3	1,21042	68
5	5	4	1,21042	69
5	7	5	1,21039	93
5	8	2	1,21039	105
6	2	2	0,96908	17
6	3	4	0,96901	28
6	4	2	0,96899	38
6	4	5	0,96899	39
6	10	2	0,96898	96
7	3	3	1,01141	48
7	7	2	1,01143	125
7	7	3	1,01143	122
7	10	3	1,01143	196
7	10	4	1,01143	183
8	2	5	1,03294	26
8	6	5	1,03402	94
8	10	3	1,03405	166
9	2	5	1,05853	38
9	3	2	1,05889	61
9	6	3	1,05900	122
9	6	4	1,05900	116
9	10	5	1,05901	199
10	3	2	1,01520	68
10	4	5	1,01525	94
10	6	5	1,01528	137
10	9	5	1,01529	200
10	10	2	1,01529	211

3. MODELO PARA GERAÇÃO DE AFLUÊNCIA

A fim de calcular a política operativa ótima do sistema, é necessário obter a solução de modelos de otimização apropriados para este fim. Para isso, é preciso ter o conhecimento do modelo estocástico das afluições, pois a otimização deve ser feita com um número grande de cenários. Como estas não são variáveis conhecidas, devem ser estimadas por meio de um modelo baseado na série histórica. Nesta série constam registros das afluições nas UHEs do SIN desde 1931. Como tais medições começaram a ser feitas recentemente, ainda há poucos dados e por isso não há cenários suficientes para que o estudo do planejamento da operação energética seja feito de forma correta com os níveis de incerteza aceitáveis (DE SOUZA SILVA, 2010). Assim, são obtidas da série histórica os principais parâmetros estatísticos que, por sua vez, são utilizados para regular um modelo a ser usado no estudo. Este modelo pode ser diferente do histórico, mas possui as mesmas características estatísticas e consegue gerar quantos cenários forem necessários. São as chamadas séries sintéticas. É interessante entender qual é a diferença entre previsão e geração de séries sintéticas. A previsão ocorre quando se quer acertar a afluição que irá acontecer no futuro e geração de cenários é quando se quer acertar a distribuição do que pode acontecer. No PMP, a preocupação maior está no que pode acontecer do que no que realmente irá acontecer, já que a segunda opção não é realizável. No caso Brasileiro, o modelo estatístico é de natureza autoregressiva e periódica, ajustando para cada mês e usina uma função que indica a quantos meses anteriores as afluições incrementais de um determinado mês estão correlacionadas. A seguir, descreve-se com mais detalhe o modelo autoregressivo das afluições.

3.1 MODELO AUTOREGRESSIVO PERIÓDICO

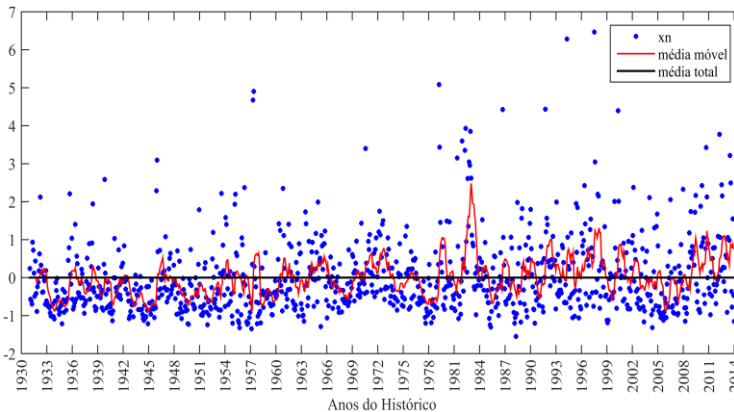
Séries hidrológicas com tempos curtos, como séries mensais, possuem um comportamento cíclico em suas características probabilísticas, tais como média, variância, assimetria e autocorrelação. Por isso, modelos autoregressivos com comportamento periódico podem ser utilizados para reproduzir esses tipos de séries. O modelo estocástico utilizado no Brasil para a geração das afluições que serão inseridas nos estudos para obter a política operativa do sistema estudado é o Modelo Autoregressivo Periódico de ordem p . A explicação dos passos para a geração da série sintética é feita com relação à afluição incremental de uma UHE.

Inicialmente, considere que yp^h é um valor de afluição que chega até a usina h . O número total de estágios do horizonte de planejamento é

denominado por ne , sendo que cada mês do ano associado a e (estágio do horizonte de planejamento) é representado por t e nt é o número total de estações do ano (i.e., $nt = 12$). Lembrando que yp^h não necessariamente é igual a afluência mensal da usina, mas sim um valor de afluência incremental em que todos os valores do histórico são deslocados a fim de se tornarem positivos. Isso é necessário porque as manipulações algébricas a serem feitas não admitem valores negativos, sendo o deslocamento representado por δy^h .

É essencial que as afluências tenham a condição de estacionaridade no tempo. Um processo estocástico é considerado estacionário quando sua distribuição de probabilidade em um determinado instante t é a mesma para qualquer outro instante t . A Figura 3.1 ilustra este fato.

Figura 3.1 - Séries afluência da UHE Salto Osório.



Nota-se que a série não possui componente de longo prazo e nem interferência. A média móvel não possui tendência de queda ou aumento ao longo dos anos. Isso é importante, pois possíveis componentes não estacionários presentes na série podem afetar o desempenho do modelo PAR, que é apresentado na próxima seção.

3.2 MODELO PAR LOGNORMAL A TRÊS PARÂMETROS

Os principais passos para a definição do modelo PAR de interesse para este trabalho estão detalhados nas próximas seções. Um modelo periódico autoregressivo pode ser definido como:

$$\frac{yp_t^h - \mu yp_t^h}{\sigma yp_t^h} = rl_t^h + \sum_{p=1}^{np} fl_{t,p}^h \cdot \left(\frac{yp_{t-p}^h - \mu yp_{t-p}^h}{\sigma yp_{t-p}^h} \right), \quad (3.1)$$

em que:

rl_t^h	Resíduo de distribuição Lognormal de média μrl_t^h e variância $(\sigma rl_t^h)^2$ em t ;
$fl_{t,p}^h$	Coefficiente linear de dependência de t com $t - p$;
μyp_t^h	Média dos valores observados em h nos estágios da estação t ;
σyp_t^h	Desvio padrão dos valores observados em h nos estágios da estação t .

3.2.1 Identificação da Ordem do Modelo Autoregressivo Periódico

Como as séries são produzidas mediante modelos autoregressivos periódicos, é necessário o cálculo da ordem de cada mês para cada UHE. Esta ordem representa o número de períodos de dependência a ser utilizado no modelo. Eles são gerados por meio do cálculo dos coeficientes de autocorrelação e autocorrelação parcial dos dados com suas respectivas defasagens.

A autocorrelação estimada $\rho yp_{t,k}^h$ da usina h e mês t em relação a k meses anteriores é expressa por:

$$\rho yp_{t,k}^h = \frac{1}{N} \sum_{t=1}^N \left[\left(\frac{yp_{t,k}^h - \mu yp_t^h}{\sigma yp_t^h} \right) \cdot \left(\frac{yp_{t-k}^h - \mu yp_{t-k}^h}{\sigma yp_{t-k}^h} \right) \right]. \quad (3.2)$$

Em que μyp_t^h e σyp_t^h são os estimadores de média e desvio padrão das afluições yp para o mês t e N é o número de observações da série.

Com os valores de autocorrelação de (3.2), é possível formar as equações de Yule-Walker, as quais são matricialmente expressas por:

$$\begin{bmatrix} 1 & \rho y p_{t-2,1}^h & \cdots & \rho y p_{t-np_t^h,1}^h \\ \rho y p_{t-1,2}^h & 1 & \rho y p_{t-3,2}^h & \rho y p_{t-np_t^h,2}^h \\ \vdots & \vdots & \ddots & \vdots \\ \rho y p_{t-1,np_t^h}^h & \rho y p_{t-2,np_t^h}^h & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} fl_{t,1}^h \\ fl_{t,2}^h \\ \vdots \\ fl_{t,np_t^h}^h \end{bmatrix} = \begin{bmatrix} \rho y p_{t,1}^h \\ \rho y p_{t,2}^h \\ \vdots \\ \rho y p_{t,np_t^h}^h \end{bmatrix}. \quad (3.3)$$

A partir de (3.3), quando $t = p = 1$ (defasagem), a correlação é realizada entre os meses 1 (janeiro) e 12 (dezembro) do histórico, uma vez que a defasagem é de um mês anterior. Caso $p=2$, a correlação se daria entre as observações de janeiro e de novembro. Pode-se notar que os elementos da diagonal principal são unitários, isso porque a autocorrelação ocorre entre os mesmos estágios (mês). Verifica-se também que a ordem dos índices não altera o valor da autocorrelação.

Neste trabalho, a estratégia utilizada para a identificação da ordem parte-se da avaliação da Função de Autocorrelação Parcial (FACP). A ordem do modelo pode variar de 1 até a máxima que em teoria é 11, porque se trabalha com períodos mensais. A FACP é composta pelo conjunto $fl_{t,p}^h$, sendo estes nunca nulos, porém sempre próximos de zero. Por isso, é necessário definir um critério de escolha para determinar quando esse valor é significativo. Com base em estudos de Hipel e McLeod (1994), o método de escolha da ordem do modelo é definido pelo último coeficiente da FACP a ficar fora do intervalo de confiança. Tal intervalo é determinado pela expressão (3.4):

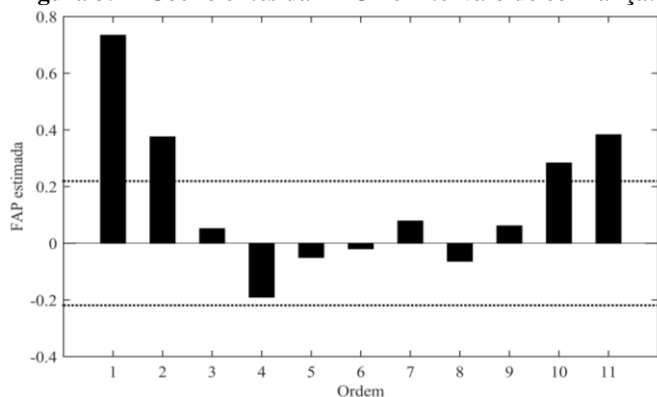
$$IC = \frac{1}{\sqrt{N}} \cdot 1,96. \quad (3.4)$$

Para decidir a ordem do modelo, a execução das equações de Yule-Walker deve ser feita de maneira gradual, ou seja, inicialmente adota-se como a ordem máxima sendo 1 e após a execução da equação, armazena-se o valor de $fl_{t,1}^h$. Caso este valor seja inferior a IC, a usina não possui coeficientes autoregressivos para aquele mês. Porém, se o valor de $fl_{t,1}^h$ é superior ao encontrado em (3.4), a ordem 2 deve ser testada. Se os valores do conjunto $[fl_{t,1}^h, fl_{t,2}^h]$ forem maiores que o valor do intervalo de confiança, a ordem 3 deve ser testada e assim sucessivamente até a ordem máxima de 11. Caso algum valor do conjunto $fl_{t,p}^h$ resultante encontrar-se dentro do intervalo de confiança, a ordem determinada para aquele mês é a anterior.

Pode ser que coeficientes autorregressivos negativos produzam coeficientes positivos indesejáveis nos cortes de Benders. Diante deste fato, mesmo que este esteja fora da zona de confiança considerada, nos estudos realizados nessa dissertação nenhum coeficiente negativo é considerado.

A Figura 3.2 ilustra um caso que a ordem verificada foi de 2 e, mesmo sem necessidade, as equações de Yule-Walker foram efetuadas para p igual a 1 até p igual a 11 a fim de deixar o exemplo mais completo.

Figura 3.2 - Coeficientes da FACP e intervalo de confiança.



3.2.2 Estimação dos Parâmetros

Após definida a ordem do modelo por meio das equações de Yule-Walker, o próximo passo consiste em efetivamente encontrar os resíduos de distribuição Lognormal rl_t^h . Diz-se que uma variável aleatória tem uma distribuição Lognormal se seu logaritmo natural é normalmente distribuído. Dessa forma, os resíduos do modelo PAR Lognormal, definido pela variável rx_t^h são encontrados aplicando o logaritmo natural em rl_t^h . A equação (3.5) consegue representar o comportamento natural da variável aleatória dos resíduos.

$$rx_t^h = rb_t^h \cdot \sigma rx_t^h + \mu rx_t^h. \quad (3.5)$$

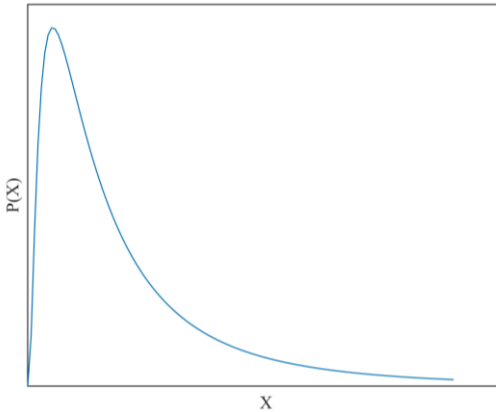
A partir de (3.5) e partindo do histórico original de afluições, pode-se estimar os valores dos resíduos de rl_t^h relativos ao histórico. Com isso, obtém-se um “histórico” de rl_t^h . Ao aplicar o logaritmo natural neste histórico, encontra-se os valores de rx_t^h e assim é possível estimar os parâmetros estatísticos como média e desvio padrão do histórico de rx_t^h . De

posse de $\mu r x_t^h$ e $\sigma r x_t^h$, os resíduos de $r l_t^h$ podem efetivamente ser calculados a partir de realizações de ruído branco (rb), conforme é mostrado em (3.6).

$$r l_t^h = \exp\left(r b_t^h \cdot \sigma r x_t^h + \mu r x_t^h\right). \quad (3.6)$$

Um ponto importante sobre o comportamento de uma distribuição Lognormal é a assimetria em relação ao seu valor médio. Numa distribuição assimétrica positiva, caso das afluências incrementais, há uma alta concentração de dados nos valores mais baixos. A cauda mais longa da distribuição situa-se na direita, apontando para uma menor frequência de valores altos (Figura 3.3). Por isso, a fim de obter uma melhor aproximação do comportamento da distribuição Lognormal é importante estimar o coeficiente de assimetria da amostra.

Figura 3.3 - Distribuição Lognormal positiva.



Com a introdução do coeficiente de assimetria estimado a partir dos valores do histórico de $r l_t^h$, uma nova formulação denominada de Resíduos a Três Parâmetros é utilizada, a qual foi inicialmente introduzida por Charbeneau (1978). Por isso, toda a parametrização da distribuição Lognormal é alterada, e pode ser definida como:

$$r l_t^h = \exp\left(r b_t^h \cdot \sigma r x_t^h + \mu r x_t^h\right) + \delta r x_t^h. \quad (3.7)$$

De acordo com a formulação acima, percebe-se que os parâmetros como $\mu r x_t^h$, $\sigma r x_t^h$ e $\delta r x_t^h$ são calculados para cada mês e para cada usina.

Além disso, observa-se na equação que os parâmetros $\mu r x_t^h$ e $\sigma r x_t^h$ também devem ser utilizados, sendo que $\delta r x_t^h$ proporciona um grau de liberdade extra que ajusta o modelo da Lognormal com os dados históricos (CHARBENEAU, 1978). Porém, com a inserção do termo $\delta r x_t^h$, os outros precisam ser redefinidos novamente. Os passos para o cálculo de cada um desses parâmetros serão mostrados adiante.

Por definição, o coeficiente de assimetria de uma distribuição Lognormal é dada por:

$$\gamma r l_t^h = \frac{\left\{ \exp \left[3 \cdot (\sigma r x_t^h)^2 \right] - 3 \exp \left[(\sigma r x_t^h)^2 \right] + 2 \right\}}{\left\{ \exp \left[(\sigma r x_t^h)^2 \right] - 1 \right\}^{3/2}}. \quad (3.8)$$

Uma vez definido o valor da assimetria $\gamma r l_t^h$ e definir $\Phi_t^h = \exp (\sigma r x_t^h)^2$, a equação (3.8) pode ser reescrita como:

$$(\Phi_t^h)^3 + 3 \cdot (\Phi_t^h)^2 - \left[4 + (\gamma r l_t^h)^2 \right] = 0. \quad (3.9)$$

O modelo Lognormal a Três Parâmetros é aplicado apenas para distribuições com coeficiente de assimetria positivo, sendo Φ_t^h sempre maior ou igual a 1. Entretanto, deve-se excluir o caso de $\Phi_t^h = 1$, pois o coeficiente de assimetria é nulo, o que aponta para uma distribuição simétrica em relação à média. Assim, para $\gamma r l_t^h < 0$, a equação (3.9) possui apenas uma raiz real (ABRAMOWITZ; STEGUN; MILLER, 1964), dada por:

$$\Phi_t^h = \left\{ \left[1 + \frac{(\gamma r l_t^h)^2}{2} \right] + \left[(\gamma r l_t^h)^2 + \frac{(\gamma r l_t^h)^4}{4} \right]^{1/2} \right\}^{1/3} + \left\{ \left[1 + \frac{(\gamma r l_t^h)^2}{2} \right] - \left[(\gamma r l_t^h)^2 + \frac{(\gamma r l_t^h)^4}{4} \right]^{1/2} \right\}^{1/3} - 1. \quad (3.10)$$

De posse do valor de Φ_t^h , os parâmetros de rx são facilmente encontrados:

$$\sigma r x_t^h = \sqrt{\ln(\Phi_t^h)}, \quad (3.11)$$

$$\mu r x_t^h = \frac{1}{2} \cdot \ln \left[\frac{(\sigma r l_t^h)^2}{\Phi_t^h \cdot (\Phi_t^h - 1)} \right], \text{ e} \quad (3.12)$$

$$\delta r x_t^h = \mu r l_t^h - \left[\frac{(\sigma r l_t^h)^2}{(\Phi_t^h - 1)} \right]^{\frac{1}{2}}. \quad (3.13)$$

Neste modelo, ainda segundo Charbeneau (1978), há alguns valores de Φ improváveis na prática que torna este modelo indefinido. A distribuição Gamma deve ser utilizada nestes casos e não a Lognormal. Mas neste trabalho, para fins de simplificações, caso aconteça será utilizado a formulação Lognormal com apenas dois parâmetros (3.6).

3.3 CORRELAÇÃO ESPACIAL

O modelo PAR discutido até agora define os coeficientes autoregressivos e estimam os parâmetros necessários para gerar as séries sintéticas, ou seja, tratam apenas com a dependência temporal. Entretanto, as usinas hidrelétricas possuem uma certa distância geográfica uma das outras e isso faz com que o comportamento hidrológico em algumas usinas seja semelhante. Por isso é muito importante representar a correlação espacial entre as hidrelétricas (PEREIRA; PINTO, 1985).

Já que a aleatoriedade do modelo PAR considera apenas a correlação temporal entre as hidrelétricas, a representação do ruído branco para a geração das séries temporais será feita por meio de sua ponderação (MACEIRA et al., 2006). Portanto, o ruído branco espacialmente correlacionado é dado por:

$$rc_t = D_t \cdot r b_t, \quad (3.14)$$

em que:

rc_t	Vetor de ruídos correlacionados de dimensão (nh, I) ;
rb_t	Vetor de ruídos brancos de dimensão (nh, I) ;
D_t	Matriz de carga triangular inferior de dimensão (nh, nh) .

Calcula-se o valor esperado, multiplicando o transposto da matriz de ruídos correlacionados nos dois lados da equação (3.14). Assim:

$$E[rc_t \cdot rc_t^T] = E[D_t \cdot rb_t \cdot rc_t^T]. \quad (3.15)$$

Substituindo o transposto de rc_t de (3.14) em (3.15), obtém-se:

$$E[rc_t \cdot rc_t^T] = E[D_t \cdot rb_t \cdot rb_t^T \cdot D_t^T]. \quad (3.16)$$

Os ruídos brancos são independentes entre eles e normalmente distribuídos, isto é, com média zero e desvio padrão unitário. Por isso, $E[rb_t \times rb_t^T]$ resulta na matriz identidade. Dessa maneira:

$$E[rc_t \cdot rc_t^T] = D_t \cdot D_t^T. \quad (3.17)$$

Uma vez que rc_t representa a matriz que considera a correlação cruzada das UHEs, ao fazer $E[rc_t \times rc_t^T]$ é obtido como resposta uma estimativa da própria matriz de correlação espacial mensal entre as hidrelétricas ρD_t .

Assim, para descobrir a matriz de carga D_t , que é uma matriz triangular inferior, deve-se fazer a inspeção dos elementos da matriz de correlação estimada.

$$D_t \cdot D_t^T = \rho D_t. \quad (3.18)$$

Uma vez encontrada a matriz de carga, esta é utilizada para correlacionar os ruídos brancos que são gerados de forma independente (equação (3.14)), resultando na matriz de ruídos correlacionados espacialmente.

3.4 GERAÇÃO DAS SÉRIES SINTÉTICAS

A partir dos valores de coeficientes autoregressivos estimados f_t^h para $p=1, \dots, np$ obtidos por meio de (3.3), e o valor do resíduo definido em (3.7), a partir de um ruído branco correlacionado, o modelo PAR Lognormal a Três Parâmetros utilizado neste trabalho pode ser determinado como:

$$yp_t^h \triangleq rp_{t,r}^h + \sum_{p=1}^{np} fp_{t,p}^h \cdot yp_{t-p}^h, t \in \mathcal{T}_+, \quad (3.19)$$

em que:

$$fp_{t,p}^h = fl_{t,p}^h \cdot \left(\frac{\sigma yp_t^h}{\sigma yp_{t-p}^h} \right), p \leq np_t^h, \quad (3.20)$$

$$fp_{t,p}^h = 0, \quad np_t^h \leq p \leq np, \quad (3.21)$$

$$rp_{t,r}^h = \mu yp_t^h + \sigma yp_t^h \cdot rl_{t,r}^h - \sum_{p=1}^{np} fp_{t,p}^h \cdot \mu yp_{t-p}^h. \quad (3.22)$$

4. PROGRAMAÇÃO DINÂMICA DUAL ESTOCÁSTICA

A Programação Dinâmica Dual Estocástica é uma ferramenta eficaz e amplamente utilizada para a solução do problema do PMP. O método utiliza a informação dual para a construção da extrapolação da função de custo futuro utilizando o princípio da decomposição de Benders e não faz o uso da discretização de espaço de estados do problema. A solução ótima do algoritmo é obtida iterativamente por meio de uma árvore de cenários, mas não é necessário percorrer todos os nós da árvore.

A PDDE é largamente utilizada no sistema brasileiro para modelos de reservatório equivalente. Mesmo fornecendo respostas satisfatórias, esta simplificação faz com que o parque gerador não seja totalmente aproveitado, já que com o reservatório equivalente de energia não é possível representar as características individuais de cada usina. Por isso, a aplicação da PDDE de forma individualizada é ainda mais interessante, visto que, apesar do maior esforço computacional, promove maior eficiência da geração de energia oriunda destas usinas.

Teoricamente, é possível a formulação de um grande problema de programação linear estocástica, o chamado equivalente determinístico e assim solucioná-lo com algoritmos de programação linear. Este método elimina a representação explícita da estocasticidade das afluições das usinas, obtendo uma solução aproximada ao substituir as variáveis estocásticas por seus valores esperados (GORENSTIN, 1991).

Porém, mesmo podendo parecer interessante em teoria, na prática não é possível representar as restrições e variáveis em um único problema de otimização de forma que as incertezas referentes ao processo estocástico sejam razoavelmente representadas. Sua aplicação deve ser feita em estudos com número reduzido de estágios de tempo e poucas realizações (LARROYD, 2016).

O período de estudo de planejamento de médio prazo que almeja encontrar o volume ótimo dos reservatórios considerando a característica de estocasticidade das afluições não deve ser muito curto, isso porque ao final do período de planejamento os reservatórios tendem a esvaziar. Por isso, muitos estudos visam calcular a política operativa para um período maior que o requerido. Assim, ao final do período de estudo, o armazenamento dos reservatórios não deve ficar baixo.

4.1 DECOMPOSIÇÃO DE BENDERS

Esta dissertação visa a aplicação do modelo autoregressivo periódico por afluições incrementais e por Componentes Principais para a solução do PMP. A parte referente às CP será abordada no próximo capítulo. O problema é solucionado com a PDDE, e por isso é necessário apresentar alguns aspectos referentes a este processo. Como no presente contexto a proposta é de caráter introdutório e não profundo, nem todos os detalhes referentes à PDDE serão mostrados.

Inicialmente, será apresentado o procedimento de criação dos cortes de Benders, isto é, conjunto de restrições que formam a FCF para o processo de PDDE com resíduos originados pelo modelo independente de geração de afluição. As mudanças deste para o modelo com dependência temporal serão mostradas logo depois.

O uso da decomposição de Benders na PDDE permite construir, iterativamente aproximações da função de custo futuro do estágio t por meio da solução do problema do estágio $t + 1$. Para o entendimento do processo de criação dos cortes de Benders, é necessário a descrição dos processos que são os pilares da metodologia: as recursões diretas (*forward*) e inversas (*backward*).

Considerando que o estudo do planejamento possui T estágios, na etapa *forward* resolve-se o conjunto de problemas percorrendo desde o estágio 1 até o estágio $T - 1$. Para cada um dos problemas, é necessário armazenar os valores ótimos das variáveis de estado e custo imediato associado de cada estágio. Por sua vez, a recursão inversa segue o caminho oposto ao anterior, ou seja, inicia-se no estágio T e finaliza no segundo estágio. A cada novo estágio de *backward* uma nova restrição a ser colocada no problema do estágio anterior é gerada.

A seguir, é apresentada a formulação do método da PDDE. Lembrando que este caso é para eventos independentes, ou seja, as afluições de um estágio qualquer não dependem das afluições de estágios anteriores. O algoritmo considera que cada um dos T estágios do horizonte de planejamento possuem m cenários com as realizações da variável estocástica. Assim, o problema inicial que visa determinar a estratégia que minimiza o valor do custo esperado é expresso da seguinte forma:

$$\begin{aligned}
\min f &= c_1 x_1 + (P_1 c_2 x_{2,1} + P_1 c_2 x_{2,2} + \dots + P_1 c_2 x_{2,m}) \\
s.a. \\
A_1 x_1 &\geq B_1 \\
E_1 x_1 + A_2 x_{2,1} &\geq B_{2,1} \\
E_1 x_1 + A_2 x_{2,2} &\geq B_{2,2} \\
\dots \\
E_1 x_1 + A_2 x_{2,m} &\geq B_{2,m}.
\end{aligned} \tag{4.1}$$

O vetor $[B_{2,1}, B_{2,2}, \dots, B_{2,m}]$ está relacionado com as variáveis (P_1, P_2, \dots, P_m) , sendo que $\sum_{i=1}^m P_i = 1$. O problema corresponde ao seguinte processo de decisão:

1º estágio: Determinar uma solução viável x_1^* , tal que $A_1 x_1^* \leq b_1$;

2º estágio: Encontrar um vetor $[x_{21}^*, x_{22}^*, x_{23}^*, \dots, x_{2m}^*]$ que é a solução do problema:

$$\begin{aligned}
z &= \min (p_1 c_2 x_{2,1} + p_2 c_2 x_{2,2} + \dots + p_m c_2 x_{2,m}) \\
s.a. \\
A_2 x_{2,1} &\geq b_{2,1} - E_1 x_1^* \\
A_2 x_{2,2} &\geq b_{2,2} - E_1 x_1^* \\
\dots \\
A_2 x_{2,m} &\geq b_{2,m} - E_1 x_1^*.
\end{aligned} \tag{4.2}$$

O problema de otimização apresentado em (4.2) pode ser decomposto em m subproblemas independentes (4.3), em que suas soluções são ponderadas de acordo com as probabilidades p_1, p_2, \dots, p_m , respectivamente.

$$\begin{aligned}
& \min f = p_1 c_2 x_{2,1} \\
& \text{s.a.} \\
& A_2 x_{2,1} \geq B_{2,1} - E_1 x_1^* \\
& \dots \\
& \min f = p_m c_2 x_{2,m} \\
& \text{s.a.} \\
& A_2 x_{2,m} \geq B_{2,m} - E_1 x_1^*.
\end{aligned} \tag{4.3}$$

A solução de cada subproblema de segundo estágio é uma função da decisão x_1 do problema de primeiro estágio. Dessa forma, o problema em (4.1) pode ser reescrito como:

$$\begin{aligned}
z &= \min c_1 x_1 + \bar{\alpha}(x_1) \\
& \text{s.a.} \\
Ax_1 &\geq b_1.
\end{aligned} \tag{4.4}$$

Em que $c_1 x_1$ representa o custo imediato e $\bar{\alpha}_2(x_1)$ representa o valor esperado do custo futuro e as soluções dos subproblemas representados em (4.3).

Para cada um dos m cenários, seja $(\pi_1, \pi_2, \dots, \pi_m)$ os multiplicadores simplex das equações de (4.3) e $(\omega_1^*, \omega_2^*, \dots, \omega_m^*)$ os valores das respectivas soluções ótimas, o corte de Benders associado ao problema (4.4) é expresso por:

$$p_1 \pi_1 (b_{2,1} - E_1 x_1) + p_2 \pi_2 (b_{2,2} - E_1 x_1) + \dots + p_m \pi_m (b_{2,m} - E_1 x_1) \leq \bar{\alpha}. \tag{4.5}$$

Ou, substituindo o vetor b :

$$\begin{aligned}
& p_1 \left(\omega_1^* + \pi_1 E_1 (x_1^* - x_1) \right) + p_2 \left(\omega_2^* + \pi_2 E_1 (x_1^* - x_1) \right) + \dots \\
& + p_m \left(\omega_m^* + \pi_m E_1 (x_1^* - x_1) \right) \leq \bar{\alpha}.
\end{aligned} \tag{4.6}$$

Agrupando os dados, o corte de Benders pode ser expresso como:

$$\bar{\omega}^* + \bar{\pi}^* E_1 (x_1^* - x_1) \leq \bar{\alpha}, \tag{4.7}$$

em que:

$$\bar{\omega}^* = p_1\omega_1^* + p_2\omega_2^* + \dots + p_m\omega_m^* \quad (4.8)$$

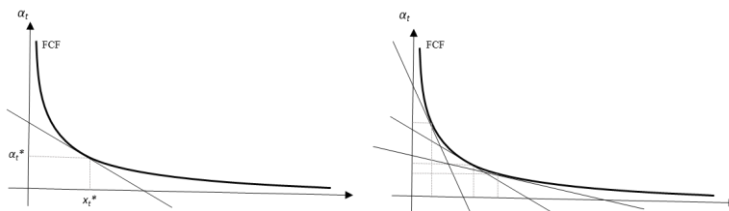
e

$$\bar{\pi}^* = p_1\pi_1 + p_2\pi_2 + \dots + p_m\pi_m. \quad (4.9)$$

Tem-se $\bar{\omega}^*$ como sendo a média do valor ótimo das PLs na etapa *backward* e $\bar{\pi}^*$ é o multiplicador simplex associado à equação de balanço hidráulico, que mede a variação do custo operativo em relação a variações marginais nas variáveis de estado do sistema.

Graficamente, uma FCF pode ser ilustrada conforme a Figura 4.1. É possível perceber que se trata de uma aproximação formada por uma função linear por partes, sendo um corte criado a cada iteração. Por construção, a FCF resultante formada pelos diversos cortes de Benders é uma aproximação inferior, já que é formada a partir da derivada da função real do custo esperado ou de uma aproximação inferior desta.

Figura 4.1 - Função de Custo Futuro.



Conforme mostrado anteriormente, os estudos realizados neste trabalho consideram a geração de afluições segundo o modelo PAR, o que significa que a afluição incremental em determinado estágio T sofre influência das afluições de estágios anteriores. Assim, a afluição incremental passa a ser uma variável de estado e por isso deve ser representada no corte de Benders. Em outras palavras, um termo a mais deve ser adicionado na expressão (4.6).

Segundo Soares (2015), as decisões fornecidas pelo algoritmo sofrem maior influência das afluições passadas em comparação com o volume dos reservatórios. Por isso, é necessário que as informações referentes às afluições passadas sejam representadas nos cortes de Benders.

O novo equacionamento do corte de Benders, com o termo associado à nova variável de estado é dado por:

$$\alpha_{t+1} \geq \bar{Z}_{t+1} + \bar{\pi} v_{t+1}^h \cdot (v_{t+1}^h - v_{t+1}^{h*}) + \sum_{h=1}^{N_h} \sum_{p=1}^{N_p} \left[\bar{\pi} y p_{t+1-p}^h \cdot (y p_{t+1-p}^h - y p_{t+1-p}^{h*}) \right]. \quad (4.10)$$

Os termos com o sobrescrito “*” são dados por valores conhecidos, encontrados durante o processo iterativo.

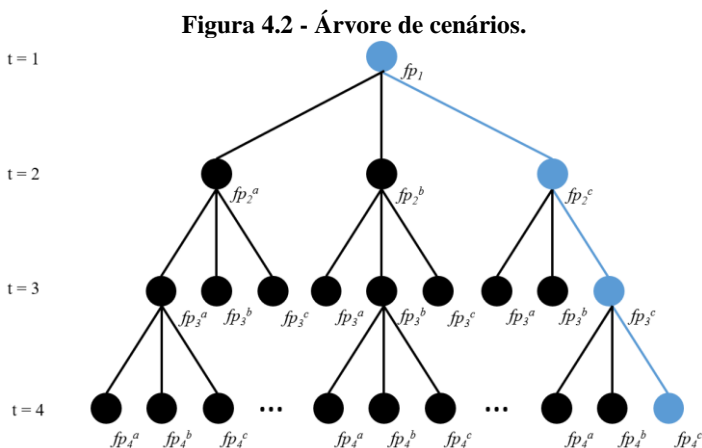
4.2 ÁRVORE DE CENÁRIOS

Devido às incertezas associadas ao problema de planejamento de médio prazo, é importante traçar estratégias que garantam com que o processo da PDDE encontre boas aproximações das FCFs.

Existe um processo estocástico de afluências da natureza. Este processo é contínuo e não conhecido. Os únicos dados do processo contínuo da natureza que se tem conhecimento estão presentes no histórico de afluências naturais. Quando se observa os dados de vazões presentes neste histórico, é possível verificar seu comportamento sazonal. Portanto, para representar o processo estocástico da natureza, é possível modelar um processo estocástico contínuo com variáveis aleatórias que possuem a estatística observada nos dados históricos (Capítulo 3). A representação dessa variável aleatória é feita de maneira discreta no planejamento de médio prazo. Dessa forma, um conjunto de realizações da variável aleatória pode ser gerado a partir de algum mecanismo de amostragem que se baseia na distribuição de probabilidade associada (LARROYD, 2016). Uma das maneiras mais usuais de representar o processo estocástico das afluências é por meio da amostragem dos dados para a geração de uma árvore de cenários. Isso porque a abertura em árvore reflete diretamente no processo decisório do problema.

Após selecionar randomicamente uma amostra com quantidade suficiente de realizações da variável aleatória, sua estatística deve representar satisfatoriamente o processo estocástico que a originou. Esta é uma das razões que faz a construção da árvore de cenários ser uma importante tarefa na solução do modelo multiestágio estocástico. É sempre desejável que o número de cenários construídos não seja muito grande, de tal forma que a solução do PL associado seja resolvido com um esforço computacional moderado.

A figura a seguir apresenta o modelo de uma árvore de cenários com três realizações por período.



Em uma árvore de cenários, os nós definem o momento da tomada de decisão em um estágio t e os traços que ligam um nó a outro indicam as conexões temporais possíveis entre as realizações. Como pode-se perceber pela da Figura 4.2, cada nó do estágio $t \geq 1$ possui apenas 1 nó antecessor e cada nó do estágio $t < T$ tem um conjunto de nós sucessores, que correspondem ao número de aberturas por período, sendo cada abertura composta por uma realização fp . O número m de aberturas de cada estágio é o mesmo, bem como o número de aberturas dos nós descendentes, embora seja possível implementar o modelo com número diferente de aberturas para os nós posteriores. Além disso, ressalta-se que o primeiro estágio (ou nó raiz) é considerado determinístico, ou seja, considera-se que o valor de sua realização é previamente conhecido.

A partir da mesma figura, nota-se também que a árvore de cenários além de possuir o mesmo número de aberturas para todos os estágios, apresenta amostras comuns. Isto quer dizer que as realizações fp descendentes de cada nó são iguais para o mesmo estágio t . É este tipo de configuração de árvore de cenários que será utilizada nas simulações deste trabalho.

O primeiro ponto deste processo consiste na definição do número de aberturas que farão parte das simulações, isto é, o número de possibilidades referentes aos cenários de afluência para cada estágio do horizonte de planejamento. Estas aberturas são consideradas no processo da simulação *backward* para a criação dos cortes e também formam os caminhos

para a simulação *forward*, que serão explicados com mais detalhes no decorrer do capítulo.

Além da qualidade do processo utilizado para a construção da árvore, o número de aberturas por estágio é muito importante. Isso porque ele se relaciona diretamente com a precisão e velocidade do processo iterativo. Cada um dos possíveis cenários é retirado do processo amostral gerado conforme exibido na Seção 3.2 e, diferentemente do que acontece no modelo independente, eles não representam diretamente a afluência incremental, mas sim um resíduo que foi gerado usando os coeficientes autoregressivos de cada usina. Estes resíduos, por sua vez, são usados para fornecer o valor da afluência incremental. Os vetores de resíduos são retirados da Função de Densidade de Probabilidade que o descreve.

Outra questão é que nem todos os possíveis cenários dos resíduos escolhidos podem ser visitados. Por exemplo, considera-se um problema de quatro anos com discretização mensal (ou 48 estágios) e apenas cinco possibilidades de resíduos por estágio. Assim de acordo com a formulação da árvore de cenários há 5^{48} problemas de otimização para serem calculados. Mesmo fornecendo uma solução ótima, o esforço computacional para solucionar o problema é muito grande o que inviabiliza sua aplicação.

O esquemático deste problema com estrutura de “árvore” que contém essa ampla quantidade de cenários foi mostrado na Figura 4.2. Felizmente não é necessário simular todo o espaço amostral, já que a partir de uma quantidade reduzida de cenários é possível estimar uma solução com precisão e tempos de simulação aceitáveis.

O conjunto total de afluências de cada usina faz parte de um processo contínuo e por isso possui infinitas possibilidades. Assim, os processos de geração de afluências promovem um conjunto finito, ainda muito grande, de afluências. Por isso, a pergunta que surge é: “Qual é o número razoável de cenários que se deve escolher para que a solução seja próxima da solução ótima do problema original?”. Se por um lado o número de cenários precisa ser tal que o esforço computacional não seja muito grande, por outro, a árvore de cenários deve fornecer uma solução com boa qualidade para o problema. Uma estratégia interessante que concilia esse objetivo é fazer o uso de algum tipo de processo de amostragem. Neste trabalho o processo de amostragem é feito de forma que as realizações possuam a mesma probabilidade (equiprováveis) e sejam independentes das demais.

No caso do modelo PAR Lognormal a Três Parâmetros utilizado para esta dissertação, a variável a ser sorteada é o resíduo, que por sua vez é transformado em afluência incremental. É interessante que a PDDE seja

executada mais de uma vez em árvores de mesmo tamanho para que mais dados possam ser considerados e a resposta final seja mais justa.

Conforme já abordado nas seções anteriores, o processo iterativo é composto por duas etapas: as recursões diretas e inversas. Inicialmente o processo de *forward* é executado com o intuito de calcular para uma amostra de cenários $\{b_{t,s}; s=1, \dots, S\}$ os valores ótimos dos volumes finais das hidrelétricas $\{x_{t,s}^*; s=1, \dots, S\}$, bem como as variáveis de afluência yp_t^h e yp_t^p . Estes valores são repassados para o problema primal subsequente. Lembrando que o problema do primeiro estágio é determinístico, portanto é necessário ter o conhecimento dos valores iniciais do problema, como volumes dos reservatórios e afluências do primeiro mês e meses anteriores (caso o primeiro mês possua coeficientes autoregressivos).

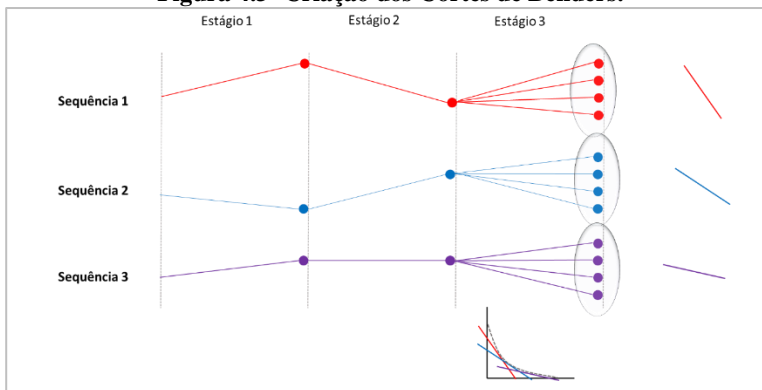
Assim, para cada estágio t e para um cenário hidrológico s resolve um subproblema de operação que minimiza os custos de operação do estágio em questão mais o valor aproximado do custo a partir de $t + 1$ até o último estágio de estudo. Ainda na recursão progressiva, a partir do segundo estágio, os cenários podem ser visitados paralelamente a fim de deixar o processo mais rápido. No entanto, tal procedimento não será aplicado no presente trabalho, isto é, os cenários serão visitados individualmente.

Já no processo *backward*, é necessário calcular os valores esperados dos custos futuros e custos marginais $\{\omega_t^*$ e π_t^* , $t = T, T - 1, \dots, 2\}$ para as m realizações de afluências, calculadas em torno do ponto $\{x_{t,s}^*; s = 1, \dots, S\}$ da recursão direta. Nota-se que a cada iteração uma nova função é atribuída para cada estágio, o que deixa o problema mais restrito. Lembra-se que todas as restrições calculadas na etapa *backward* no estágio t podem ser aplicadas para qualquer outro cenário no estágio $t - 1$. Este compartilhamento de cortes para os cenários de um mesmo estágio faz com que o processo iterativo atinja a convergência mais rápido, pois caso contrário apenas uma pequena quantidade de cenários seria visitada, produzindo um número reduzido de cortes.

Segundo Marcató (2002), o critério de parada do algoritmo que indica a convergência do processo iterativo, ocorre quando a diferença do *Zinf* e *Zsup*, calculado ao fim de cada iteração seja menor que uma tolerância previamente estabelecida. Nas simulações executadas neste trabalho, a parada do processo iterativo ocorrerá por tempos máximos de execução, o que garante um resultado não enviesado com o mesmo número para todos os modelos executados. A precisão encontrada em cada caso será posteriormente avaliada por meio da estabilização do *Zinf*. Isto porque nos modelos oficiais do setor elétrico, utiliza-se como critério de parada a estabilização do *Zinf*.

A Figura 4.3 apresenta, na etapa *backward* como é feita a construção dos cortes de Benders para o penúltimo estágio, representado no exemplo como estágio 2. Nota-se que o corte é feito com os valores médios das variáveis para todos os possíveis cenários do período seguinte (estágio 3), considerando os valores finais do estágio anterior. Por exemplo, os valores finais das variáveis de volumes dos reservatórios e afluência do estágio 2 são os valores iniciais dessas variáveis no estágio 3.

Figura 4.3- Criação dos Cortes de Benders.



4.3 FORMULAÇÃO MATEMÁTICA

Nas seções e capítulos anteriores, foi discutido sobre a representação dos principais componentes do sistema hidrotérmico e modelo de geração de afluência no âmbito do PMP para usinas individualizadas. Esta seção se trata da formulação do PMP para este trabalho.

4.3.1 Modelo PAR Lognormal a Três Parâmetros por Usina

Os modelos de afluências incrementais do modelo Lognormal a Três Parâmetros são originados a partir de ruídos brancos correlacionados espacialmente e de parâmetros estatísticos do histórico. Além disso, estes modelos são definidos considerando a dependência temporal entre as afluências passadas. Assim, as variáveis referentes às afluências do estágio anterior são interpretadas como variáveis de estado do modelo de planejamento de médio prazo. Por essa razão, para uma política de operação

ótima do sistema a ser aplicado, sua FCF deve depender tanto das variáveis de volume armazenado dos reservatórios, vh_t^h , quanto das variáveis de afluência para np estágios anteriores, $yp_{t-p}^h, \dots, yp_{t-np}^h$.

Na restrição de balanço hídrico pode ser que, dependendo do valor estado vh_t^h , o problema de otimização se torne inviável para valores negativos de afluências incrementais. Por isso, uma variável de afluência complementar yc_t^h é adicionada como uma folga nessa restrição de maneira que o problema possa sempre ser viável e que não comprometa a utilização do modelo Lognormal. Essa variável, também requerida na função objetivo do problema, possui sempre um valor de “penalidade” yc_t^h mais alto que o de déficit, o que faz que ela seja acionada unicamente quando for necessário. Por fim, vale salientar que, conforme o que foi explicado no Capítulo 2, a FPH está aproximada em uma função linear por partes e é inserida no problema por meio de restrições.

4.3.1.1 Função Objetivo

A função objetivo do problema apresenta a finalidade dos problemas da operação energética, que é atender a carga requerida com o menor custo de operação esperado. Então, a função objetivo de um determinado estágio t corresponde ao custo operativo do estágio presente somado a um custo futuro esperado, representado pelas funções de custo futuro que se inicia a partir do estágio $t+1$ até o fim do horizonte de planejamento. Atualmente, além deste custo a função objetivo inclui o *Conditional Value at Risk* dos cenários hidrológicos críticos. Porém, a metodologia proposta neste documento independe deste risco.

Nas simulações realizadas, a taxa de desconto que referencia os custos de estágios posteriores para o presente não é utilizada.

4.3.1.2 Restrição de Atendimento à Demanda

A restrição do atendimento à demanda visa assegurar que toda a energia requerida, em MWm, pela demanda D_t seja atendida em todos os estágios de tempo pelas unidades hidrelétricas ph_t e termelétricas pT_t . Pode ser que as UHEs e UTEs não sejam suficientes para atender a carga. Nesses casos, o déficit que representa o corte de carga e possui capacidade infinita de geração será despachado. Seu custo para cada estágio t é designado por cd_t e sua quantidade por pd_t . No modelo de interesse deste trabalho, a equação de demanda tem apenas um patamar de carga.

4.3.1.3 Restrições de Balanço Hídrico

As restrições de balanço hídrico devem obedecer ao princípio da conservação da massa da água. Por isso, o volume de água que aflui ao aproveitamento no período t deve ser igual à soma dos volumes que correspondem à variação do nível do reservatório e liberados pelas turbinas e vertedouros. O acoplamento espacial entre as UHEs de uma cascata também é considerado, já que a água liberada em uma usina por meio da defluência e vertimento afeta na operação das usinas à jusante. Aqui, as equações desconsideram os efeitos de evaporação e infiltração nos reservatórios.

4.3.1.4 Restrições de Dependência Temporal das Afluências

Essas restrições têm por objetivo incorporar o modelo PAR à formulação do problema. Para isso, é necessário definir os coeficientes do modelo autoregressivo periódico em função do estágio, visto que o modelo foi definido para cada usina em cada mês do ano.

4.3.1.5 Função de Custo Futuro

A função de custo futuro num problema de planejamento da operação energética visa comparar o custo de operação considerando o uso imediato da água armazenada com o custo de poupar essa água ao longo do horizonte de estudo. No problema de planejamento, a FCF é formada por uma função linear por partes.

4.3.1.6 Formulação Completa

A formulação no contexto da PDDE para uma realização de resíduos das afluências em um dado mês é apresentada abaixo.

Função Objetivo

$$\min \sum_{j=1}^{NT} cT_i^j \cdot pT_i^j + cd_i \cdot pd_i + \sum_{h=1}^{Nh} cyc_i^h \cdot yc_i^h + \alpha_{t+1}.$$

s.a.:

Restrição de Atendimento à Demanda

$$\sum_{j=1}^{NT} pT_i^j + \sum_{h=1}^{Nh} ph_i^h + pd_i = D_i.$$

Restrição de Balanço Hídrico das UHEs

$$v_{t+1}^h + q_t^h + s_t^h - \sum_{m=1}^{Nm} (q_t^m + s_t^m) - y h_t^h - y c_t^h = v_t^h.$$

Restrições de Dependência Temporal das Afluências

$$\begin{aligned} y h_t^h + y c_t^h &\geq 0, \\ y h_t^h - y p_t^h &= -\delta y h_t^h, \\ y p_t^h &= r p_{t,r}^h + \sum_{p=1}^{Np} f p_{t,p}^h \cdot y p_{t-p}^h. \end{aligned} \quad (4.11)$$

Cortes de Benders – Função de Custo Futuro

$$\begin{aligned} \alpha_{t+1} &\geq \bar{Z}_{t+1} + \bar{\pi} v_{t+1}^h \cdot (v_{t+1}^h - v_{t+1}^{h*}) + \\ &+ \sum_{h=1}^{Nh} \sum_{p=1}^{Np} \left[\bar{\pi} y p_{t+1-p}^h \cdot (y p_{t+1-p}^h - y p_{t+1-p}^{h*}) \right]. \end{aligned}$$

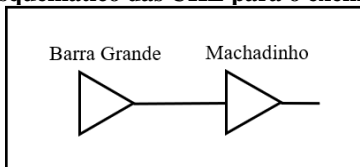
Restrição de Limite das Variáveis

$$\begin{aligned} \underline{v}_t^h \leq v_{t+1}^h \leq \bar{v}_t^h, \quad p T_t^j \leq p T_t^j \leq p T_t^j, \quad p d_t \leq p d_t \leq p d_t, \\ 0 \leq q_t^h \leq \bar{q}_t^h, \quad s_t^h \geq 0, \quad y c_t^h \geq 0. \end{aligned}$$

4.4 EXEMPLO NUMÉRICO

Esta seção visa mostrar um caso com valores numéricos para ilustrar os conceitos que foram apresentados sobre a PDDE com o modelo autoregressivo. Será usado como exemplo os dados de duas usinas conectadas em cascata: Barra Grande e Machadinho (Figura 4.4). O exemplo, que pretende mostrar como o problema de planejamento é realizado, tem horizonte de 5 meses e duas realizações por estágio (16 cenários no total).

Figura 4.4 - Esquemático das UHE para o exemplo numérico



Inicialmente, é necessário encontrar a ordem do modelo autoregressivo para cada mês e para cada usina. Tendo como base os valores do

histórico das aflúncias incrementais, ele é encontrado aplicando o exposto no Capítulo 3.

A árvore de cenários para este exemplo pode ser vista na Figura 4.5. Para as duas UHEs, os valores de cada resíduo são encontrados estão na Tabela 4.1. O sistema possui cinco UTEs e a demanda constante a ser atendida tem valor de 1.000 MW.

Figura 4.5 - Árvore de cenários para 5 meses (duas realizações/mês)

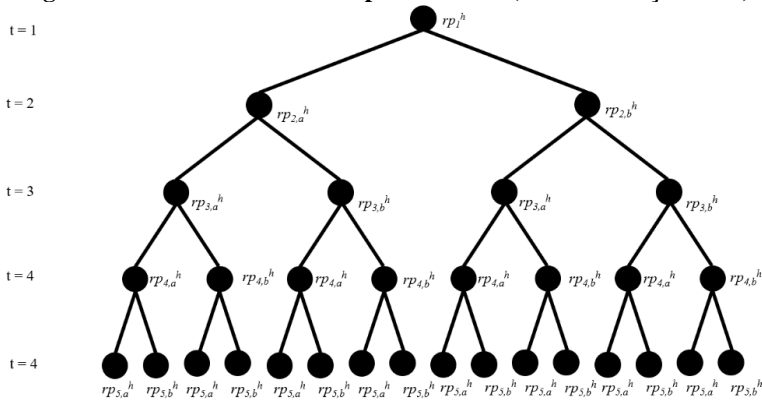
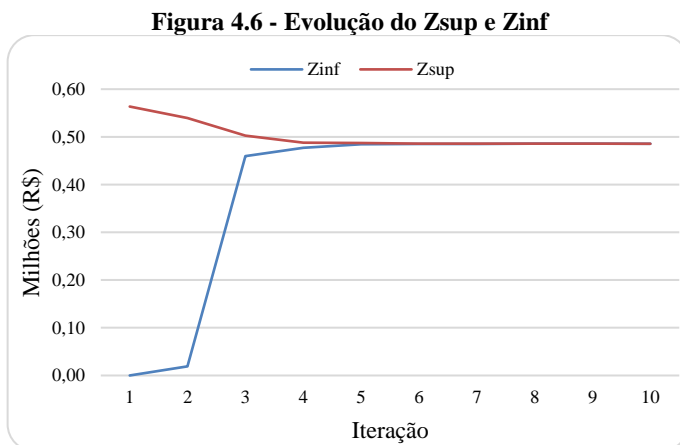


Tabela 4.1 - Resíduos das UHEs de Barra Grande e Machadinho

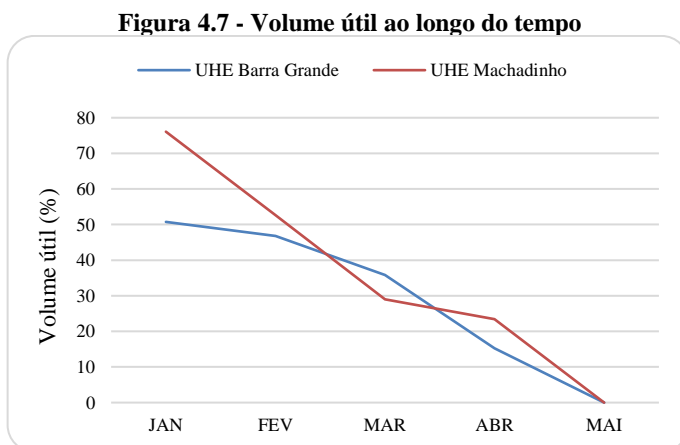
Estágio	Barra Grande		Machadinho	
	$rp_{t,a}^1$	$rp_{t,b}^1$	$rp_{t,a}^2$	$rp_{t,b}^2$
1	180,91		89,15	
2	-183,48	83,58	-61,07	128,83
3	-36,64	9,57	2,13	-1,99
4	-22,66	-98,97	4,25	-58,72
5	129,61	-158,85	82,61	-37,81

Este processo visa resolver toda a árvore de cenários. Na etapa *forward*, todos os cenários são percorridos e os valores de custo de otimização, aflúncias e volumes armazenados são salvos a cada iteração. O mesmo é feito na etapa *backward*, em que os cortes da FCF são construídos resolvendo todos os nós da árvore de cenários. Neste caso, o primeiro mês terá um corte de Benders por iteração, o segundo terá sua função acrescida de dois cortes por iteração. Já o terceiro estágio será acrescido de quatro restrições e o penúltimo mês de oito cortes a cada iteração.

Neste exemplo, o limite inferior é representado pelo custo de primeiro estágio, enquanto que o limite superior é o custo médio de operação de todos os cenários. A Figura 4.6 mostra a evolução desses limites.



Como sistema possui poucas variáveis e realizações, o limite superior encontra-se com o inferior em apenas 10 iterações com custo de R\$ 485.599,08. O valor médio de volume útil para os 16 cenários ao longo dos cinco meses é exibido na Figura 4.7. Conforme esperado, nota-se que os reservatórios se esvaziam conforme se aproxima do estágio final do planejamento.



O problema de programação linear de um estágio pode ser visto na equação (4.12).

$$\min z = 40gt_i^1 + 60gt_i^2 + 150gt_i^3 + 90gt_i^4 + \\ + 200gt_i^5 + 2000 \cdot yc_i^1 + 2000 \cdot yc_i^2 + \alpha_{t+1}$$

s.a.:

$$\begin{aligned} gt_i^1 + gt_i^2 + gt_i^3 + gt_i^4 + gt_i^5 + ph_i^1 + ph_i^2 &= 1000 \\ v_{t+1}^1 + q_t^1 + s_t^1 - yh_t^1 - yc_i^1 &= v_t^{1*} \\ v_{t+1}^2 + q_t^2 + s_t^2 - q_t^1 - s_t^1 - yh_t^2 - yc_i^2 &= v_t^{2*} \\ yh_t^1 - yp_t^1 &= -\delta yh_t^1 \\ yh_t^2 - yp_t^2 &= -\delta yh_t^2 \\ yh_t^1 + yc_i^1 &\geq 0 \\ yh_t^2 + yc_i^2 &\geq 0 \\ yp_t^1 &= rp_{t,r}^1 + \sum_{p=1}^{Np} fp_{t,p}^1 \cdot yp_{t-p}^1 \\ yp_t^2 &= rp_{t,r}^2 + \sum_{p=1}^{Np} fp_{t,p}^2 \cdot yp_{t-p}^2 \end{aligned} \tag{4.12}$$

Cortes de Benders – Função de Custo Futuro

$$\begin{aligned} \alpha_{t+1} &\geq \bar{z}_{t+1} + \bar{\pi} v_{t+1}^h \cdot (v_{t+1}^h - v_{t+1}^{h*}) + \\ &+ \sum_{h=1}^2 \sum_{p=1}^{Np} \left[\bar{\pi} yp_{t+1-p}^h \cdot (yp_{t+1-p}^h - yp_{t+1-p}^{h*}) \right] \end{aligned}$$

Restrição de Limite das Variáveis

$$\begin{aligned} 2712 \leq v_{t+1}^1 \leq 4905, 2283 \leq v_{t+1}^2 \leq 3340, 0 \leq q_t^1 \leq 516, \\ 0 \leq q_t^2 \leq 1311, s_t \geq 0, yc_i \geq 0, 0 \leq gt_i^1 \leq 200, 0 \leq gt_i^2 \leq 300, \\ 0 \leq gt_i^3 \leq 500, 0 \leq gt_i^4 \leq 100, 0 \leq gt_i^5 \leq 300 \end{aligned}$$

4.5 SIMULAÇÃO DA OPERAÇÃO

Uma vez encontradas as FCFs para cada árvore na etapa de otimização, é necessário utilizá-las em uma outra etapa para que diversos parâmetros operativos do sistema possam ser analisados. Esses parâmetros podem ser: evolução de armazenamento do sistema, geração termelétrica, déficit de energia, entre outros. Para isso, lança-se mão do uso de um simulador que consiste na resolução de diversas recursões progressivas.

Esse procedimento realiza as recursões progressivas da PDDE com cenários de vazões afluentes obtidas por meio da função de probabilidade original e não cenários sorteados dentro da árvore em que a PDDE foi executada, isto é, são sorteios *out-of-sample*. O número de cenários de afluência utilizado para a simulação da operação é de 10.000. Lembra-se que o sorteio desses cenários é feito nos resíduos do modelo PAR que geram as possíveis realizações de afluências no mês de estudo. No caso de afluências incrementais, esses resíduos são criados pelo modelo PAR com base no histórico das vazões incrementais das usinas. Já para as CP (que será detalhado no capítulo seguinte), o histórico das componentes principais serve de base para o modelo autoregressivo.

4.5.1 Avaliação da Qualidade da Solução

A avaliação da política é de grande importância, uma vez que a partir dela é possível mensurar a qualidade da solução. Além disso, é possível obter uma orientação para a escolha do tamanho da amostra e critérios de parada dos algoritmos (SHAPIRO; RUSZCZYNSKI, 2009).

O critério para avaliar a qualidade da solução obtida é feito pelo cálculo da brecha (ou *gap*) de otimalidade entre a solução da PDDE encontrada com árvore de cenários de nr_t realizações por estágio, com a solução do problema original obtido pelo processo estocástico contínuo. O primeiro caracteriza o limite inferior, enquanto o segundo, superior. Dessa forma, obtém-se um intervalo com um certo nível de confiança no qual o *gap* não será maior que esse valor com um determinado nível de confiança. Os conceitos utilizados nesta seção seguem o proposto por Shapiro e Ruszczyński (2009). Seja a brecha de otimalidade dada por:

$$gap(x_1') = z(x_1') - z^*. \quad (4.13)$$

Na equação (4.13), z^* é o valor ótimo do problema real e $z(x_1')$ tem seu valor estimado pela média das amostras que são geradas de forma

independente e igualmente distribuídas da função de densidade de probabilidade original em que x_1' é um ponto viável. A variância da amostra é representada conforme o seguinte equacionamento:

$$\sigma_{Nu}^2(x_1') = \frac{1}{Nu} \sum_{\omega=1}^{Nu} [q(x_1', \xi^\omega) - q_{Nu}(x_1')]^2. \quad (4.14)$$

Sendo ξ^ω , $\omega=1, \dots, Nu$ as realizações, é possível determinar os valores de $q(x_1', \xi^\omega)$. Em (4.14), $q_{Nu}(x_1')$ corresponde à média dos valores de $q(x_1', \xi^\omega)$.

É viável usar um tamanho de amostra relativamente alto para Nu , já que o cálculo de $q(x_1', \xi^\omega)$ envolve a solução de problemas de segundo estágio individuais. O limite superior de $z(x_1')$ com $100(1-\alpha)\%$ de confiança é expresso por:

$$U_{Nu}(x_1') = q_{Nu}(x_1') + v_{1-\alpha} \frac{\sigma_{Nu}(x_1')}{\sqrt{Nu}}. \quad (4.15)$$

O valor de $v_{1-\alpha} = \Phi(1-\alpha)$, em que $\Phi(1-\alpha)$ é a Função de Distribuição Acumulada da Distribuição Normal Padrão.

O próximo passo consiste em encontrar o valor do limite inferior para z^* . Seja z_N o valor ótimo de um problema de aproximação pela média amostral calculado de uma árvore de cenários de tamanho N . Este valor é considerado aleatório, uma vez que é obtido por meio de amostras aleatórias. Sendo que a árvore de cenários é um subconjunto do problema original para o qual se encontrou a solução ótima, é possível compreender que $z^* \geq E[z_N(x)]$.

Pode-se estimar $E[z_N(x)]$ resolvendo M vezes os problemas baseados em amostras geradas independentes de tamanho N e calculando a média de seus valores ótimos. Dessa forma, seja z_N^1, \dots, z_N^M , as soluções ótimas de todos esses problemas, sua média e variância estimadas são expressas por:

$$z_{nM} = \frac{1}{M} \sum_{j=1}^M z_n^j \quad (4.16)$$

$$\sigma_{nM}^2 = \frac{1}{M} \sum_{j=1}^M (z_n^j - \bar{z}_{nM})^2. \quad (4.17)$$

Assim, o limite inferior, a um nível de confiança $100(1-\alpha)\%$ é dado por:

$$L_{nM} = \bar{z}_{nM} - t_{\alpha, v} \frac{\sigma_{nM}}{\sqrt{M}}. \quad (4.18)$$

Em que $v=M-1$ e $t_{\alpha, v}$ é o valor de α crítico da distribuição *t-student* com v graus de liberdade. As aplicações costumam utilizar pequenas quantidades para M , como $M=5$ ou $M=10$.

4.5.1.1 Estimação da brecha de otimalidade para otimização estocástica

Uma vez determinada a maneira de se encontrar os valores dos limites superior e inferior, pode-se construir um intervalo de confiança para a brecha de otimalidade, mostrado em (4.19).

$$GAP = U_{Nu}(x_1') - L_{nM}. \quad (4.19)$$

Ou em porcentagem:

$$GAP = \frac{U_{Nu}(x_1') - L_{nM}}{L_{nM}} \cdot 100\%. \quad (4.20)$$

5. ANÁLISE DE COMPONENTES PRINCIPAIS

A Análise de Componentes Principais (ACP) introduzida por Hotelling (1933) é um procedimento matemático capaz de transformar um grande conjunto de dados correlacionados entre si em outro conjunto menor que contenha a maior parte da informação original. São as chamadas componentes principais. Neste caso, as componentes principais não possuem correlação entre si (i.e., são ortogonais), sendo dadas por combinações lineares das variáveis originais.

Seja z o número total de variáveis de um conjunto de dados. Considere ainda que há necessidade de z componentes para reproduzir a variabilidade total desses dados. Porém, grande parte dessa variabilidade pode ser expressa por um menor número k (tal que $z > k$) de componentes principais. Isso porque, em geral é nas primeiras componentes que a maior parcela da variância original dos dados está concentrada. Assim, as últimas ($z - k$) componentes podem ser ignoradas ou substituídas, por exemplo, pelos seus valores médios sem grande efeito na solução do problema.

Neste trabalho, o principal intuito do emprego das Componentes Principais é diminuir o número de variáveis presente no algoritmo da PDDE. Essas variáveis, que serão discutidas com mais detalhes no decorrer do capítulo, estão relacionadas às afluições das usinas hidrelétricas para o PMP.

Este capítulo aborda a explicação das componentes principais e o passo-a-passo de como são obtidas. Como exemplo, essa técnica será aplicada nas afluições incrementais de um sistema com duas usinas hidrelétricas e posterior discussão sobre os resultados obtidos.

5.1 METODOLOGIA

Uma das vantagens da ACP é que se pode conseguir reduzir o número de dimensões de um conjunto de dados com mínima perda de informação. A quantidade de informação perdida depende do nível de exigência da aplicação que será feita. As componentes principais são dispostas em ordem decrescente de importância, de modo que a primeira componente principal é aquela com variância máxima, e assim por diante para as demais componentes. Inicialmente, serão apresentados os passos para a obtenção das componentes principais a partir de um conjunto de dados genérico.

5.1.1 Matriz de Dados

Considere z características de n indivíduos de uma população. As características são representadas pelas variáveis $\{x_{11}, x_{21}, x_{31}, \dots, x_{ij}, i=1, n; j=1, z\}$ e a matriz de dados de ordem $n \times z$ é denominada aqui como X .

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1z} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2z} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3z} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \cdots & x_{nz} \end{bmatrix}. \quad (5.1)$$

A estrutura de interdependência entre as variáveis é representada pela matriz de covariância S ou matriz de correlação R . Para o caso de R , há a necessidade de fazer uma padronização dos dados, a qual pode ser feita com média zero e variância 1 ou variância 1 e média qualquer. Desta forma, os elementos da diagonal principal desta matriz possuirão valor unitário. O emprego da matriz de correlação promove uma ponderação das variáveis originais, dando a elas o mesmo nível de importância, não importando os valores relativos à variância de cada uma.

A escolha entre as matrizes de covariância ou correlação não é uma tarefa simples. Neste trabalho, a matriz de covariância será utilizada, já que a ACP será empregada para as afluências incrementais das hidrelétricas, sendo importante a distinção entre a quantidade de chuvas de cada usina.

5.1.2 Matriz de Covariância

Partindo da matriz de dados, é possível encontrar a matriz de covariância, que é simétrica e de ordem $z \times z$. A construção da matriz é apresentada em (5.2), em que n representa o número total de termos e \bar{X} e \bar{Y} representam valores médios (HOTELLING, 1933).

$$S = \begin{bmatrix} \text{Var}(x_1) & \text{Cov}(x_1x_2) & \text{Cov}(x_1x_3) & \cdots & \text{Cov}(x_1x_z) \\ \text{Cov}(x_2x_1) & \text{Var}(x_2) & \text{Cov}(x_2x_3) & \cdots & \text{Cov}(x_2x_z) \\ \text{Cov}(x_3x_1) & \text{Cov}(x_3x_2) & \text{Var}(x_3) & \cdots & \text{Cov}(x_3x_z) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(x_zx_1) & \text{Cov}(x_zx_2) & \text{Cov}(x_zx_3) & \cdots & \text{Var}(x_z) \end{bmatrix}, \quad (5.2)$$

sendo:

$$\text{Var}(X) = \frac{\sum_{i=1}^n (X_i - \bar{X}) \cdot (X_i - \bar{X})}{(n-1)},$$

$$\text{Cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{(n-1)}.$$

5.2 DETERMINAÇÃO DAS COMPONENTES PRINCIPAIS

As componentes principais são encontradas ao resolver a equação característica da matriz S , conforme é mostrada abaixo. O termo I representa a matriz identidade.

$$\det |S - \lambda I| = 0. \quad (5.3)$$

É importante salientar que o número de raízes (autovalores) da matriz S é igual a z , caso a mesma possua posto completo. Outro ponto importante é que, em aplicações convencionais, o valor de n indivíduos deve ser no mínimo igual a $z+1$.

Ao resolver (5.3), encontram-se as raízes da equação característica $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_z$, tal que $\lambda_1 > \lambda_2 > \lambda_3, \dots > \lambda_z$. Para cada autovalor λ_i , há um autovetor \tilde{a}_i associado, dado por:

$$\tilde{a}_i = \begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{iz} \end{bmatrix}, i = 1, \dots, z. \quad (5.4)$$

Uma vez que o autovetor foi calculado pela equação (5.4), é possível separar as componentes principais. Deste modo, a i -ésima componente principal é dada por:

$$Y_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{iz}X_z. \quad (5.5)$$

A primeira componente principal (Y_1) terá variância igual ao valor do primeiro autovalor, e assim por diante. Isto quer dizer que $\text{Var}(Y_1) > \text{Var}(Y_2) > \dots > \text{Var}(Y_z)$. Outra propriedade importante assegura que o total de variância presente nas variáveis originais iguala-se ao somatório dos autovalores. Este somatório, por sua vez, também é igual ao total de variância das componentes principais. A contribuição de cada componente principal é expressa em porcentagem, cujo cálculo pode ser feito ao dividir a variância Y_i pelo somatório das variâncias.

5.3 NÚMERO DE COMPONENTES UTILIZADAS

Uma tarefa fundamental consiste em definir quantas componentes devem ser consideradas em uma dada aplicação, de forma que a porcentagem da variação dos dados originais explicada seja satisfatória. Uma das técnicas sugere plotar um gráfico em que a variação explicada por cada componente esteja representada nas ordenadas e a porcentagem das componentes em ordem decrescente nas abscissas, de modo que se deve considerar as componentes que se encontram anteriores ao ponto de inflexão da curva. Tal critério foi exposto por Cattell (1966) e visa incluir um número elevado de componentes. Outro modo de seleção inclui apenas as componentes que seus valores próprios sejam superiores à média. Esse critério, proposto por Kaiser (1960) tende a considerar um número reduzido de componentes. Entretanto, de forma geral, utiliza-se ao menos o número de componentes que compreendem uma variância acumulada em torno de 70% do total.

5.4 TRANSFORMAÇÃO PARA AS VARIÁVEIS ORIGINAIS

Uma vez encontradas as componentes principais, é possível que dependendo da aplicação, o usuário necessite retornar para as variáveis originais. Se esta transformação for feita com todas as componentes, isso não é muito atrativo, pois não irá ter uma redução de variáveis e o resultado encontrado será o mesmo que os dados que originaram tais componentes. Por isso, é de interesse eliminar determinadas componentes antes de fazer essa transformação, como é o caso da aplicação nesse trabalho.

A conversão das variáveis em âmbito de componentes principais para variáveis originais é feita de maneira bem simples. Como já foi exposto, matricialmente as componentes são expressas por $Y=A \cdot X$. Isso quer dizer que $A^{-1} \cdot Y=X$, e dado que matriz dos autovalores A é ortonormal, então $A^{-1}=A^T$ e, portanto, $X=A^T \cdot Y$.

5.5 MODELO DE GERAÇÃO DE AFLUÊNCIAS CONSIDERANDO COMPONENTES PRINCIPAIS

Em termos de representação dos recursos hidráulicos, o problema do PMP pode ser modelado por meio de reservatórios equivalentes de energia, por bacias hidrográficas ou por outro meio que possa diminuir o número de reservatórios do sistema e, conseqüentemente, deixar a solução do problema mais rápida quando comparada com a representação por usina. Essas são alternativas interessantes, pois como o Sistema Interligado Nacional é de grande porte e se todas as usinas fossem representadas, o espaço amostral seria muito grande e sua solução bastante onerosa.

Entretanto, no tocante a este trabalho, a representação por hidrelétrica será adotada, visto que os estudos são iniciais e feitos para uma quantidade máxima de 16 usinas hidrelétricas (pertencentes as cascatas dos rios Iguaçu e Uruguai). Outro ponto importante deste trabalho é o emprego do conceito de ACP para comparar com o caso base de hidrelétricas.

Como mencionado anteriormente, a ACP compreende uma transformação de x variáveis correlacionadas para outras x variáveis não correlacionadas.

Se no caso de representação de afluências incrementais por hidrelétricas os valores de yp_t^h são obtidos a partir dos dados do histórico de afluências incrementais das usinas, para as componentes os valores são determinados a partir de um histórico de CP. É possível construir este histórico por meio dos valores do histórico incremental original com os seus respectivos autovetores. Com isso, o processo de representação de afluências será feito de forma distinta daquele com as afluências incrementais das usinas hidrelétricas.

Para as CP, o processo em relação ao caso da geração a partir das afluências incrementais varia pouco, mas algumas colocações devem ser comentadas. Agora, o modelo autoregressivo explicado no Capítulo 3 é aplicado a um histórico de componentes principais. Este histórico é obtido com base no histórico das afluências incrementais e calculado por meio

de z autovetores de dimensão $z \times I$ para cada mês, sendo z o número de UHEs.

Após a aplicação da ACP para um determinado mês, todas as afliências estão correlacionadas. Desta forma, o procedimento da Seção 3.3 que pretende encontrar uma matriz de carga para correlacionar as usinas não é necessário. Vale lembrar que quanto menor o número de componentes principais consideradas no problema, maior será a correlação espacial entre as usinas.

Uma diferença em relação ao caso da aplicação do modelo PAR por hidrelétrica está na regressão para o contexto de afliência incremental, já que yp não está relacionada às afliências incrementais de uma hidrelétrica, mas sim às componentes principais, sendo estas relacionadas às afliências das hidrelétricas.

Por isso, a regressão para se gerar a série sintética de afliências incrementais a partir da série das componentes principais deve ser feita de forma inversa ao que foi feito no início para a formação do histórico de componentes. Antes disso, deve-se reduzir os valores da série sintética δcp^{CP} unidades e depois multiplicar o resultado encontrado pelo transposto do autovetor de cada mês obtido no início do processo.

5.5.1 Modelo PAR Lognormal a Três Parâmetros considerando a ACP

O ponto principal deste trabalho consiste na aplicação de modelos autoregressivos periódicos para a geração de afliências e aplicá-lo no PMP por meio da PDDE. O uso desse procedimento deve ser comparado com a aplicação do modelo PAR nas séries contendo as variáveis de afliência incremental transformadas: as componentes principais. Assim, alguns ajustes devem ser feitos na PDDE para a entrada da ACP. Os problemas tratam de variáveis de naturezas distintas no que diz respeito à afliência do processo estocástico yp . A formulação de yp permanece a mesma, sendo que a mudança consiste em criar uma nova variável auxiliar denominada $yp2_t^{CP}$ a qual é encarregada de reduzir o valor do grau de liberdade δcp^{CP} de yp_t^{CP} . E é com a variável auxiliar que a transformação para afliências incrementais yh_t^h é feita. Isso é necessário, pois ambos os procedimentos não podem ser realizados por meio de uma única equação. Conforme mostrado no presente capítulo, o termo A_t é o autovetor no estágio t .

Com base nessas considerações, a formulação no contexto da PDDE para uma realização de resíduos das afluições em um dado mês considerando o uso de CP é dado por:

Função Objetivo

$$\min \sum_{j=1}^{NT} cT_t^j \cdot pT_t^j + cd_t \cdot pd_t + \sum_{h=1}^{Nh} cyc_t^h \cdot yc_t^h + \alpha_{t+1}.$$

s.a.:

Restrição de Atendimento à Demanda

$$\sum_{j=1}^{NT} pT_t^j + \sum_{h=1}^{Nh} ph_t^h + pd_t = D_t.$$

Restrição de Balanço Hídrico das UHEs

$$v_{t+1}^h + q_t^h + s_t^h - \sum_{m=1}^{Nm} (q_t^m + s_t^m) - yh_t^h - yc_t^h = v_t^h.$$

Restrições de Dependência Temporal das Afluições

$$yh_t^h + yc_t^h \geq 0,$$

$$yh_t^h - A_t \cdot yp2_t^{cp} = 0,$$

$$-yp1_t^{cp} + yp2_t^{cp} = -\delta cp^{cp},$$

(5.6)

$$yp1_t^{cp} = rp_{t,r}^{cp} + \sum_{p=1}^{Np} fp_{t,p}^{cp} \cdot yp1_{t-p}^{cp}.$$

Cortes de Benders – Função de Custo Futuro

$$\alpha_{t+1} \geq \bar{Z}_{t+1} + \bar{\pi} v_{t+1}^h \cdot (v_{t+1}^h - v_{t+1}^{h*}) + \\ + \sum_{cp=1}^{Ncp} \sum_{p=1}^{Np} \left[\bar{\pi} yp1_{t+1-p}^{cp} \cdot (yp1_{t+1-p}^{cp} - yp1_{t+1-p}^{h*}) \right].$$

Restrição de Limite das Variáveis

$$v_t^h \leq v_{t+1}^h \leq \bar{v}_t^h, pT_t^j \leq pT_t^j \leq pT_t^j, pd_t \leq pd_t \leq pd_t \\ 0 \leq q_t^h \leq \bar{q}_t^h, s_t^h \geq 0, yc_t^h \geq 0.$$

5.6 EXEMPLO NUMÉRICO COM CP

O mesmo procedimento do feito no exemplo da Seção 4.4 é feito aplicando os conceitos envolvendo as componentes principais mostradas nas últimas seções. Os coeficientes do modelo autoregressivo também são

calculados utilizando os valores do histórico. Mas neste caso o processo é feito com o histórico das CP. Este novo histórico é encontrado ao aplicar a equação $X = A^T \cdot Y$ com os valores do histórico das afluições incrementais. Neste caso prático, z corresponde ao número total das usinas a serem estudadas (duas neste caso), enquanto n é relativo ao número de anos disponível no histórico de afluições das usinas, i.e., 80.

Para ilustrar os passos deste processo para encontrar o histórico CP, o mês de janeiro será usado como exemplo, sendo que o procedimento para o cálculo das componentes principais para os demais meses é o mesmo.

A Tabela 5.1 apresenta os dados de afluições incrementais, de ordem $n \times z$. Os valores de afluições são expressos em m^3/s .

Tabela 5.1 - Histórico de afluições incrementais

Ano	Barra Grande	Machadinho
1931	440	229
1932	482	320
1933	89	41
1934	179	91
1935	135	60
1936	847	485
⋮	⋮	⋮
2006	65	67
2007	134	147
2008	247	190
2009	56	38
2010	351	220

De posse da matriz de dados, o próximo passo é calcular a matriz de covariância S ($z \times z$), exibida em (5.7).

$$S = \begin{bmatrix} 13208,0 & 5969,2 \\ 5969,2 & 3727,9 \end{bmatrix} \quad (5.7)$$

A partir da equação característica é possível encontrar os autovalores e, conseqüentemente, os autovetores. Neste caso serão dois autovalores e para cada autovetor, um autovalor, conforme mostrado a seguir.

Tabela 5.2 - Autovalores do exemplo numérico

	Autovalor	Percentual
λ_1	16090,26	95,01
λ_2	845,58	4,99

Tabela 5.3 - Autovetores do exemplo numérico

	\tilde{a}_1	\tilde{a}_2
	0,9005	-0,4348
	0,4348	0,9005

Após determinados esses parâmetros, as duas componentes principais podem ser encontradas. Os coeficientes da primeira componente (que detém a maior parte da variabilidade dos dados iniciais) são extraídos do autovetor relativo ao primeiro autovalor. Já os coeficientes da segunda, do segundo autovetor. As duas componentes são mostradas em ((5.8).

$$\begin{aligned}
 Y_1 &= 0,9005X_1 + 0,4348X_2 \\
 Y_2 &= -0,4348X_1 + 0,9005X_2
 \end{aligned}
 \tag{5.8}$$

Uma vez calculadas todas as componentes, dependendo da aplicação pode-se excluir algumas CP, fixar no valor médio as componentes com menor variância ou até considerar todas as componentes. Neste último, apesar do problema continuar com as mesmas variáveis, o uso de um outro conjunto de variáveis, dessa vez ortogonal, pode ser interessante dependendo da aplicação.

Neste trabalho, o próximo passo consiste em construir um histórico de componentes principais partindo do histórico de afluências incrementais. Para isso, basta substituir em X_1 e X_2 os valores referentes às afluências incrementais das usinas Barra Grande e Machadinho. Assim, obtém-se uma nova matriz $n \times z$, agora com valores de CP.

Caso o intuito seja voltar ao âmbito de afluências incrementais, o processo é o inverso. Basta multiplicar cada linha (ano) do histórico pela matriz de autovetores transposta, conforme já explicado. Caso o propósito

seja diminuir o número de componentes antes de voltar ao histórico, pode-se fazer o mesmo processo, porém fixando a média do histórico para componentes menos importantes. Assim, apenas as primeiras componentes (que possuem maior relevância) entrarão no algoritmo da PDDE como variáveis de estado do problema.

A tabela a seguir apresenta alguns valores presentes no histórico de componentes principais para o mês de janeiro.

Tabela 5.4 - Histórico de componentes principais do exemplo numérico

Ano	CP 1	CP 2
1931	172,19	0,14
1932	107,79	0,15
1933	60,74	6,21
1934	153,47	-4,14
1935	76,02	1,05
1936	178,81	6,94
⋮	⋮	⋮
2006	169,15	-38,36
2007	137,20	40,36
2008	134,62	32,72
2009	181,38	14,58
2010	486,67	-4,01

De acordo com tabela, nota-se que a segunda componente possui alguns valores negativos. Antes de prosseguir para o cálculo do modelo PAR para as CP é necessário deslocar δ_{cp}^2 unidades. Este valor é dado por usina, por isso o mesmo para todos os meses deve ser feito. Para o presente sistema de duas UHEs, este valor é de 261,17. O procedimento para o cálculo dos coeficientes autoregressivos ($fp_{t,p}^{cp}$) é feito segundo o exibido na Seção 3.4.

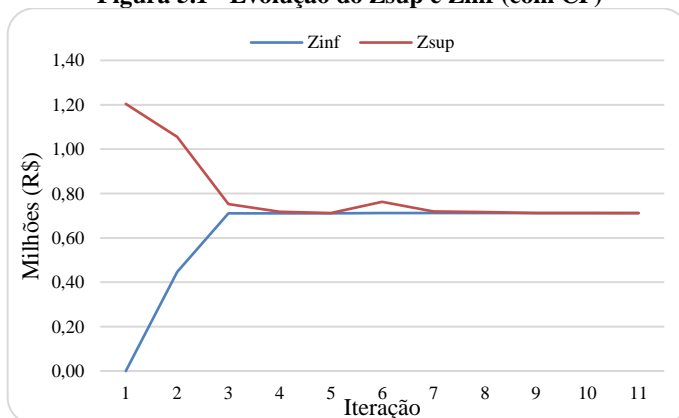
Após a aplicação da ACP para um determinado mês, todas as afluições estão correlacionadas não sendo necessário encontrar uma matriz de carga para correlacionar os valores entre as usinas. Vale lembrar que quanto menor o número de componentes principais consideradas no problema, maior será a correlação espacial entre as usinas. De posse desses dados, pode-se encontrar os resíduos $rp_{t,r}^{cp}$ para a geração das séries sintéticas yp_t^{cp} . Para este exemplo, os resíduos são exibidos na tabela a seguir.

Tabela 5.5 - Resíduos das UHEs de Barra Grande e Machadinho (com CP)

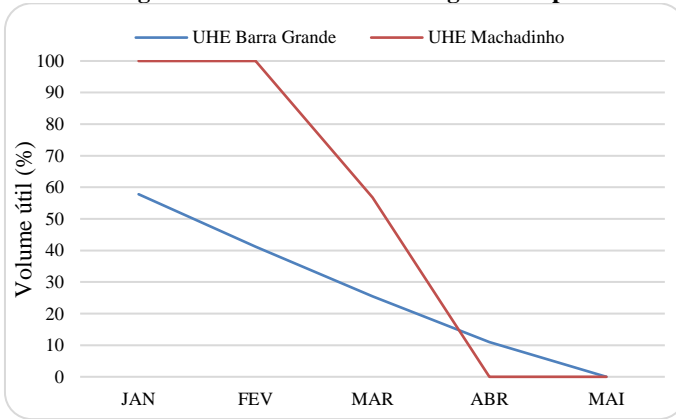
Estágio	Barra Grande		Machadinho	
	$rp_{t,a}^1$	$rp_{t,b}^1$	$rp_{t,a}^2$	$rp_{t,b}^2$
1		202,62		259,83
2	-303,75	-154,28	185,33	168,15
3	-16,15	-40,09	182,74	106,22
4	-27,84	18,81	220,20	198,56
5	348,28	-5,08	129,02	69,66

A Figura 5.1 ilustra os limites inferior e superior após a aplicação da PDDE para o caso com as duas CP.

Figura 5.1 - Evolução do Zsup e Zinf (com CP)



O processo convergiu em 11 iterações e custo ótimo de R\$ 711.886,55. A Figura 5.2 exibe o valor médio de volume útil das duas usinas para os 16 cenários. Da mesma forma que no caso com as afluições incrementais, os reservatórios se esvaziam conforme se aproxima do estágio final do planejamento. Salienta-se que nenhum dos resultados encontrados com as CP podem ser comparados com os obtidos no exemplo anterior feito com as afluições incrementais, já que o sistema utilizado, horizonte de planejamento e número de realizações são pouco representativos.

Figura 5.2 - Volume útil ao longo do tempo

O problema de programação linear de um estágio para o caso com CP é mostrado em (5.9).

$$\min z = 40gt_i^1 + 60gt_i^2 + 150gt_i^3 + 90gt_i^4 + \\ + 200gt_i^5 + 2000 \cdot yc_i^1 + 2000 \cdot yc_i^2 + \alpha_{t+1}$$

s.a.:

$$\begin{aligned} gt_i^1 + gt_i^2 + gt_i^3 + gt_i^4 + gt_i^5 + ph_i^1 + ph_i^2 &= 1000 \\ v_{i+1}^1 + q_i^1 + s_i^1 - yh_i^1 - yc_i^1 &= v_i^{1*} \\ v_{i+1}^2 + q_i^2 + s_i^2 - q_i^1 - s_i^1 - yh_i^2 - yc_i^2 &= v_i^{2*} \\ yh_i^1 - Av_i^1 \cdot yp2_i^1 &= 0 \\ yh_i^2 - Av_i^2 \cdot yp2_i^2 &= 0 \\ -yp1_i^1 + yp2_i^1 &= -\delta y_i^1 \\ -yp1_i^2 + yp2_i^2 &= -\delta y_i^2 \\ yp1_i^1 &= rp_{i,r}^1 + \sum_{p=1}^{Np} fp_{i,p}^1 \cdot yp1_{i-p}^1 \\ yp1_i^2 &= rp_{i,r}^2 + \sum_{p=1}^{Np} fp_{i,p}^2 \cdot yp1_{i-p}^2 \end{aligned} \quad (5.9)$$

$$yh_t^1 + yc_t^1 \geq 0$$

$$yh_t^2 + yc_t^2 \geq 0$$

$$yp_t^1 = rp_{t,r}^1 + \sum_{p=1}^{Np} fp_{t,p}^1 \cdot yp_{t-p}^1$$

$$yp_t^2 = rp_{t,r}^2 + \sum_{p=1}^{Np} fp_{t,p}^2 \cdot yp_{t-p}^2$$

Cortes de Benders – Função de Custo Futuro

$$\alpha_{t+1} \geq \bar{z}_{t+1} + \bar{\pi} v_{t+1}^h \cdot (v_{t+1}^h - v_{t+1}^{h*}) + \\ + \sum_{cp=1}^2 \sum_{p=1}^{Np} \left[\bar{\pi} yp_{t+1-p}^{cp} \cdot (yp_{t+1-p}^{cp} - yp_{t+1-p}^{cp*}) \right]$$

Restrição de Limite das Variáveis

$$2712 \leq v_{t+1}^1 \leq 4905, 2283 \leq v_{t+1}^2 \leq 3340, 0 \leq q_t^1 \leq 516,$$

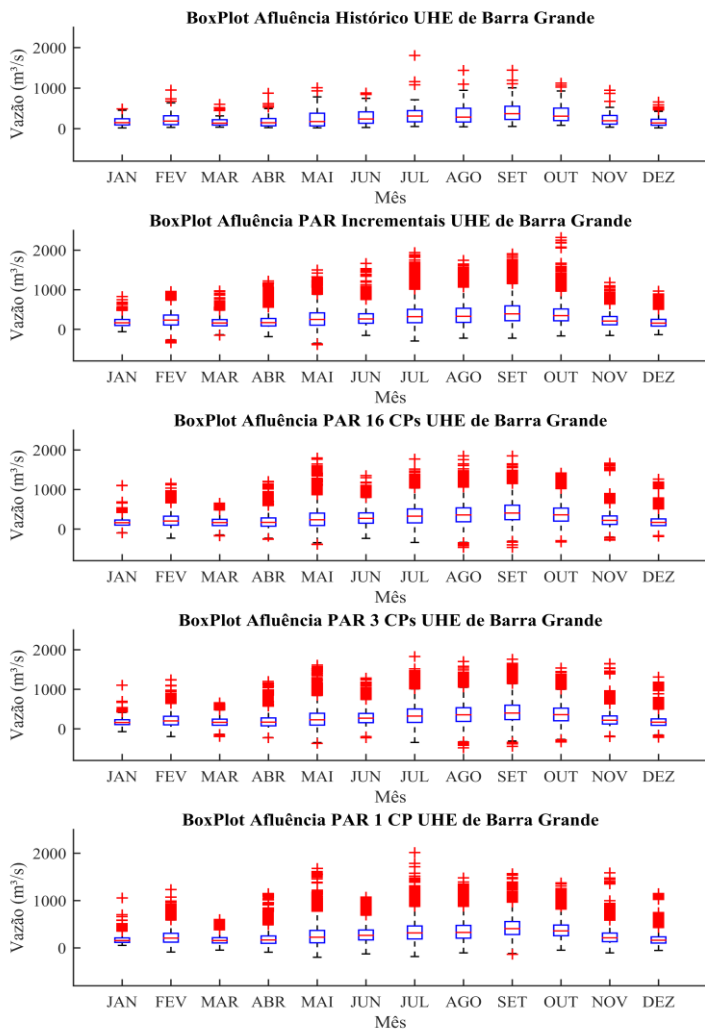
$$0 \leq q_t^2 \leq 1311, s_t \geq 0, yc_t \geq 0, 0 \leq gt_t^1 \leq 200, 0 \leq gt_t^2 \leq 300,$$

$$0 \leq gt_t^3 \leq 500, 0 \leq gt_t^4 \leq 100, 0 \leq gt_t^5 \leq 300$$

5.7 SÉRIE SINTÉTICA PARA MODELOS SEM E COM CP

Para exemplificar a eficácia do modelo Lognormal a Três Parâmetros, os gráficos da Figura 5.3 ilustram o comportamento da afluência da usina de Machadinho gerada por meio das equações apresentadas neste capítulo e no capítulo do modelo para as afluências incrementais. Nesta figura, destacam-se os modelos com PAR com as afluências incrementais de 16 usinas da cascata do Uruguai e Iguaçu, PAR com 100% das componentes principais, PAR com 3 componentes e PAR com apenas uma componente principal. Todas podem ser vistas em comparação com os valores retirados do histórico. Para o modelo Lognormal foram utilizadas 10.000 séries sintéticas. Os gráficos de tipo Box-Plot presentes na Figura 5.3 mostram que nos 4 modelos estudados, o comportamento das vazões afluentes incrementais e dos dados do histórico são bastante similares. Também é possível notar que o modelo autoregressivo fornece alguns valores de afluências negativas. Mas isso não compromete na visão geral dos dados com médias e desvios padrão parecidos.

Figura 5.3 - Comparação dos casos para a UHE de Barra Grande.



6. RESULTADOS COMPUTACIONAIS

Este capítulo tem como foco apresentar os resultados obtidos nos experimentos computacionais baseados nas formulações e metodologias propostas nos capítulos anteriores. Inicialmente são apresentados os principais parâmetros e dados do sistema hidrotérmico utilizado nas simulações computacionais.

Um ponto já discutido que tem fundamental importância no tempo computacional é o número de realizações de afluições de cada mês. Deste modo, é necessário que seja arbitrado um valor com uma representatividade estatística adequada ao modelo e seja compatível com as exigências de tempo e precisão esperados. Nesse sentido, este capítulo apresenta a avaliação da qualidade da solução obtida com a PDDE para problemas de números distintos de realizações por mês perante o problema original (infinitas realizações). Esta avaliação é feita por meio da estimativa da brecha de otimalidade e envolve um considerável esforço computacional, conforme detalhado na Seção 6.6.

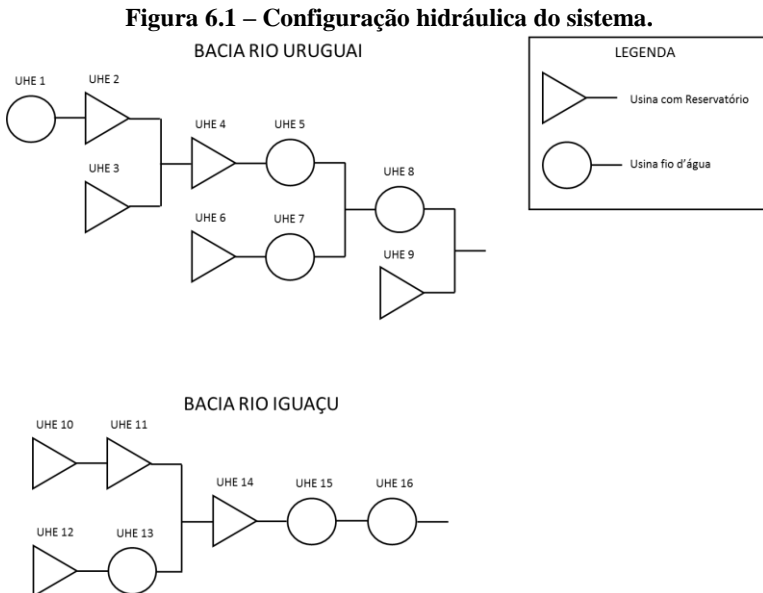
Outro aspecto importante diz respeito ao critério de parada do algoritmo da PDDE. Esse é um assunto ainda em aberto, sendo que não existe um critério disponível que seja sólido metodologicamente na avaliação da solução obtida perante o problema original. Para fins deste trabalho, utiliza-se o critério de limite de tempo de execução do algoritmo. Algumas avaliações do critério de estabilização do limite inferior são feitas na Seção 6.13, isso porque o mesmo tem sido utilizado em maior escala na atualidade, mesmo sem observar índices qualitativos do problema original.

No final do capítulo é mostrado a comparação dos resultados do modelo de geração de afluições por usina em relação aos modelos com componentes principais, citando as limitações e pontos positivos de cada uma das metodologias. Todos os resultados foram obtidos por meio de um computador com a seguinte configuração: processador Intel core i7, 8 GB de RAM e com sistema operacional Windows 10. Os estudos referentes ao *Convex Hull*, modelo de geração de afluições e implementação do modelo do PMP foram feitos em linguagem *Matlab*. Para resolver cada problema de PL, foi utilizado o pacote computacional Gurobi 7.5.1.

6.1 SISTEMA HIDROTÉRMICO

O sistema hidrotérmico utilizado neste trabalho é composto por 16 UHEs e cinco UTEs. Todas as usinas estão conectadas em barra única para atender um patamar único de carga constante em um horizonte de quatro anos, discretizado mensalmente. Não é utilizada a atualização para valor presente dos custos futuros. Uma termelétrica fictícia com capacidade de suprir toda a demanda e CVU de R\$ 1.000,00/MWh representa o déficit.

No tocante ao sistema hidrelétrico são utilizadas as principais usinas locadas nas bacias hidrográficas dos rios Uruguai e Iguazu, as quais totalizam 16 UHEs. A Figura 6.1 mostra a configuração hidráulica das usinas, destacando o acoplamento espacial entre as mesmas, bem como as respectivas características de regularização. Conforme pode ser visto, nove usinas possuem capacidade de regularização e sete são modeladas como usinas fio d'água.



Por sua vez, a Tabela 6.1 apresenta as características operativas das UHEs.

Tabela 6.1 – Dados operativos das UHEs.

UHE	Nome	Volume (hm ³) mín/máx	Potência (MW)	Vazão máxima (m ³ /s)	FPH NL/EMQ
1	Garibaldi	296 / 296	192	501	12 / 0,40
2	Campos Novos	1320 / 1477	880	558	14 / 0,97
3	Barra Grande	2712 / 4904	690	516	20 / 1,86
4	Machadinho	2283 / 3340	1140	1311	18 / 1,60
5	Itá	5100 / 5100	1450	1590	7 / 0,62
6	Passo Fundo	337 / 796	226	102	5 / 0,73
7	Monjolinho	151 / 151	74	134	6 / 1,27
8	Quebra-Queixo	111 / 137	120	114	13 / 0,94
9	Foz do Chapecó	1428 / 1428	855	1888	8 / 0,57
10	Foz do Areia	1974 / 5779	1676	1376	17 / 2,04
11	Segredo	2562 / 2950	1260	1268	35 / 0,85
12	Santa Clara	169 / 431	120	162	18 / 2,45
13	Fundão	35 / 35	120	152	6 / 1,87
14	Salto Santiago	2662 / 6775	1420	1576	26 / 1,91
15	Salto Osório	1124 / 1124	1078	1784	16 / 2,04
16	Salto Caxias	3573 / 3573	1240	2100	8 / 0,63

A última coluna da tabela refere-se ao número de aproximações lineares por parte utilizadas em cada UHE e o erro quadrático médio associado, em %. Especificamente neste sistema, todas as usinas fio d'água tem sua FPH dependente da vazão turbinada e do vertimento. Por sua vez, nas usinas com regularização a FPH depende ainda do volume armazenado, ou seja, é quadridimensional. Para o propósito deste trabalho foi escolhido uma grade de discretização do CH que resultasse em um conjunto de planos com um erro médio quadrático ao redor de 2%.

Com relação às usinas termelétricas, neste trabalho as mesmas são hipotéticas. Ao todo se tem cinco usinas com custos variáveis unitários e capacidades máximas apresentadas na Tabela 6.2. Restrições como tempos mínimos de partida e parada, limites de rampa não são considerados, tampouco aquelas envolvendo interrupções para manutenção e geração mínima. Todos os limites mínimos de geração são nulos.

Tabela 6.2 – Usinas termelétricas do sistema.

UTE	Limite Máximo (MWm)	CVU (R\$ MWm)
1	1000	40
2	650	200
3	450	250
4	620	450
5	305	600

6.1.1 Modelo Autoregressivo

A afliências em um determinado mês dependem das afliências ocorridas nos meses anteriores. Essa informação está presente no problema devido aos coeficientes autoregressivos do modelo PAR. No tocante a esta dissertação, estes coeficientes são encontrados para as usinas de forma individualizada. Neste caso, então denomina-se aqui modelo PAR sem CP, abreviado por SCP.

O mesmo acontece quando o processo é feito quando se considera no modelo as CP. Como todas as CP possuem uma parcela da afliência incremental de cada usina, é possível eliminar aquelas menos importantes sob ponto de vista estatístico e ainda assim transformá-las em afliências por meio do processo de recursão.

Tabela 6.3- Porcentagem acumulada das CP.

CP	Média acumulada de todos os meses (%)	CP	Média acumulada de todos os meses (%)
1	83,2566	8	99,5732
2	92,5123	9	99,7114
3	95,9530	10	99,8224
4	97,4059	11	99,8978
5	98,3270	12	99,9471
6	98,9303	13	99,9769
7	99,3564	14	99,9973
8	99,5732	15	99,9999
9	99,7114	16	100,0000

A Tabela 6.3 ilustra o percentual acumulado das CP calculado com base nos autovalores da matriz de covariância. Nesta tabela estão considerados os valores médios para todos os meses.

Nota-se que a primeira CP detém cerca de 80% da variabilidade dos dados originais, muito superior sobre as demais componentes. As últimas possuem praticamente zero, o que as torna, de certo modo, descartáveis. É possível que se utilize quantidades variadas de CP para cada mês. No entanto, pela proximidade da quantidade que cada componente possui, no contexto desse trabalho será utilizado o mesmo número de CP em todos os meses, tendo a média total como base de medida desse valor.

As tabelas a seguir mostram a ordem do modelo PAR nos casos SCP e com CP.

Tabela 6.4 – Ordem do modelo PAR para o caso SCP.

UHE	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1	0	1	2	2	1	1	1	1	1	1	1	1
2	0	1	1	2	1	1	2	1	1	1	1	1
3	0	1	1	2	1	1	2	1	1	1	1	1
4	0	1	1	2	1	1	2	1	1	1	1	1
5	0	1	1	2	1	1	2	1	1	1	1	1
6	0	1	1	1	1	1	2	1	1	1	1	1
7	0	1	1	1	1	1	2	1	1	2	1	1
8	0	1	1	1	1	1	2	1	1	1	1	1
9	0	1	1	1	2	1	2	1	1	1	1	1
10	0	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	2	1	1	1	1
12	0	1	1	1	1	1	2	0	1	1	1	1
13	0	1	1	1	1	1	1	0	1	1	1	1
14	0	1	1	1	2	1	2	1	1	1	1	1
15	0	1	1	1	1	1	1	1	1	1	1	1
16	0	1	1	2	2	1	2	1	1	1	1	1

Tabela 6.5 – Ordem do modelo PAR com CP.

CP	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1	0	1	1	2	1	1	2	1	1	1	1	1
2	0	1	0	0	0	0	0	0	0	1	0	0
3	0	0	0	0	1	0	0	0	1	0	0	1
4	0	1	1	0	0	0	0	0	1	1	0	0
5	0	1	0	1	0	0	0	1	0	0	0	0
6	1	0	0	1	1	2	0	0	0	1	0	0
7	0	0	0	1	0	1	1	0	0	1	0	0
8	0	0	0	0	0	0	1	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	1	0	0	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0	1	0
12	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	1	2	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	0	0	1	1	0	0	1	0	0	0
16	1	0	0	2	1	0	1	1	0	2	0	1

Neste sentido, o trabalho irá analisar o desempenho de três variantes com afliências representadas por CP: com uma, três e 16 componentes. Esses casos são nomeados aqui como CP1, CP3 e CP16, respectivamente. Em conjunto com SCP, os modelos CP1, CP3 e CP16 representam o universo de modelo de afliências investigado nos experimentos computacionais.

6.2 DADOS ADICIONAIS PARA OTIMIZAÇÃO E SIMULAÇÃO

Conforme previamente citado, o horizonte de planejamento do problema é de 48 meses. A demanda elétrica do sistema é de 7.000 MWm constante para todos os meses, o que corresponde a aproximadamente 60% da capacidade instalada total da configuração de hidrelétricas. De acordo com a Tabela 6.2, a capacidade instalada UTEs é de 3.015 MWm, ou seja, cerca de 40% da demanda. O déficit é aplicado quando o sistema hidrotérmico não consegue atender toda a demanda. Seu custo é superior ao de todas as unidades térmicas (R\$ 1.000/MWh), independentemente

da quantidade de corte de carga. Com relação ao estado inicial de armazenamento do sistema, foi arbitrado um valor de 50% dos volumes úteis das usinas com capacidade de regularização.

As variáveis aleatórias que compõem a árvore de cenários na PDDE são os resíduos do modelo PAR. Dado que o primeiro mês da árvore de cenários é composto por um único nó, foi adotado para cada UHE uma média de 4.000 valores de resíduos de janeiro. Isso garante com que as afluições do primeiro mês para todos os casos fiquem equivalentes. Quando o mês inicial possui dependência temporal de meses anteriores, a afluição de dezembro do último ano do histórico será empregada. Se a dependência for maior que um, o mesmo para os meses anteriores é aplicado. Já para dependências do segundo mês ou posterior, a afluição aplicada é a já obtida pelo processo iterativo. Caso a dependência temporal for superior ao já obtido, aplica-se o histórico também.

6.3 MODELOS ESTUDADOS

Conforme já citado, as execuções do algoritmo da PDDE estão divididas em quatro modelos: SCP, CP16, CP3 e CP1. O modelo autoregressivo do primeiro deles é feito a partir do histórico de afluições incrementais e os demais com o histórico das componentes principais.

O número de realizações da variável aleatória em cada mês é um fator relevante no processo. Quanto mais elevado este número, mais precisa é a solução, porém o tempo requerido por iteração tende a ser maior. Por isso, foram arbitrados três valores: 20, 50 e 100.

A Tabela 6.6 mostra os quatro tipos de configuração no tocante ao número e tipo de variáveis de estado (as variáveis de estado referentes às afluições incrementais e CP podem variar dependendo da ordem do modelo). Nos modelos em que as CP são utilizadas, a geração dos valores de resíduos é feita com todas as componentes. Nas execuções da PDDE os resíduos referentes às componentes menos significativas são eliminados.

Para se ter uma representação estatística adequada, em cada caso serão utilizadas 10 sementes diferentes, ou seja, em cada caso da Tabela 6.6 tem-se 10 árvores diferentes. Considerando ainda que 20, 50 e 100 realizações de afluições por mês são avaliadas, tem-se ao final 120 execuções do algoritmo da PDDE nesta análise comparativa.

Tabela 6.6 - Número de variáveis de estado.

Modelo	Variáveis de estado
SCP	9 Volumes + 16 Afluências incrementais
CP16	9 Volumes + 16 CP
CP3	9 Volumes + 3 CP
CP1	9 Volumes + 1 CP

6.4 CRITÉRIO DE CONVERGÊNCIA

Neste trabalho, foram feitas algumas simulações iniciais da PDDE para servirem como indicador de base de tempo para as simulações. Para isso, foi adotado um rigoroso critério de estabilização do custo inferior da PDDE: média do crescimento relativo dos custos nas últimas 10 iterações deve ser menor que 0,02%. Dentre todas as 120 execuções do algoritmo, o maior valor de tempo encontrado foi de 5.252,1 segundos (ou 87min32s). Tal valor foi obtido em um caso com 100 realizações por mês no caso CP16. Por isso, o tempo de 90 minutos foi utilizado como o maior tempo de critério de parada da PDDE. Os demais tempos de parada foram estipulados em 10, 30, 50, e 70 minutos.

6.5 SIMULAÇÃO PARA ESTIMATIVA DO CUSTO SUPERIOR

Em conformidade ao explicado na Seção 4.5.1, para encontrar uma estimativa do custo superior são utilizadas as FCFs obtidas na PDDE. Para tanto, são sorteados 10.000 cenários (*out-of-sample*) sintéticos por mês. Após, é realizada uma iteração *forward* considerando os 10.000 cenários para obter um valor esperado do custo superior do problema.

6.6 ESTRATÉGIA DE COMPARAÇÃO

Tendo como tempos de parada para a execução da PDDE 10, 30, 50, 70 e 90 minutos, modelos de afluências SCP, CP1, CP3 e CP16, tamanhos de abertura por mês de 20, 50 e 100 realizações e 10 sementes, ao todo 600 conjuntos de aproximações de FCF foram simulados nos

10.000 cenários sintéticos de afluências incrementais para o cálculo das brechas de otimalidade. Embora a avaliação da brecha de otimalidade não sirva de parâmetro na comparação entre os distintos modelos de afluência utilizados, a mesma serve para identificar a melhor configuração de número de aberturas a ser adotada em cada modelo para cada um dos tempos de parada especificados.

6.7 TEMPO COMPUTACIONAL DE 10 MINUTOS

Para elucidar o procedimento utilizado para o cálculo da brecha de otimalidade, na subseção seguinte tal procedimento será descrito detalhadamente para as execuções envolvendo o modelo SCP no menor tempo estipulado, 10 minutos. Em seguida, na Seção 6.8, todos os resultados de brecha de otimalidade serão exibidos para todas as demais execuções da PDDE.

6.7.1 Modelo SCP

Esta seção tem como finalidade mostrar os resultados encontrados das execuções da PDDE e simulação nos cenários sintéticos de afluências incrementais para o caso SCP. Neste sentido, a Figura 6.2 mostra a evolução do limite inferior das 10 sementes ao longo do processo iterativo da PDDE para 20 realizações/mês. A Figura 6.3 e a Figura 6.4 exibem tal comportamento para 50 e 100 realizações por mês, respectivamente.

Figura 6.2 - Limite inferior dos problemas com 20 realizações por mês.

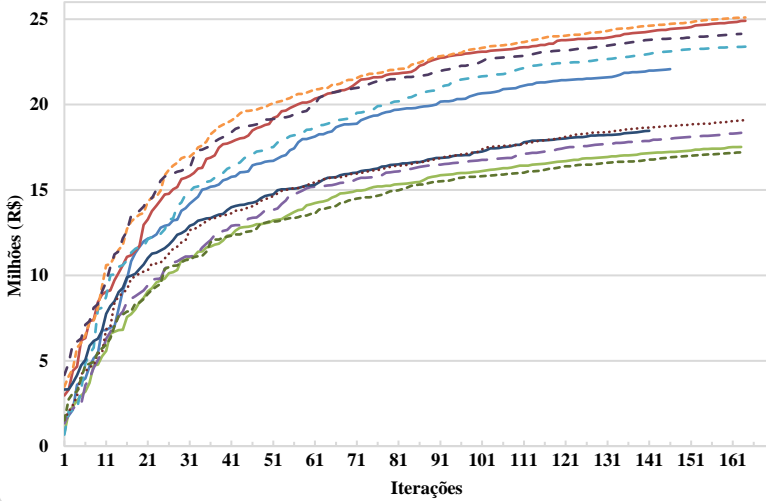


Figura 6.3 - Limite inferior dos problemas com 50 realizações por mês.

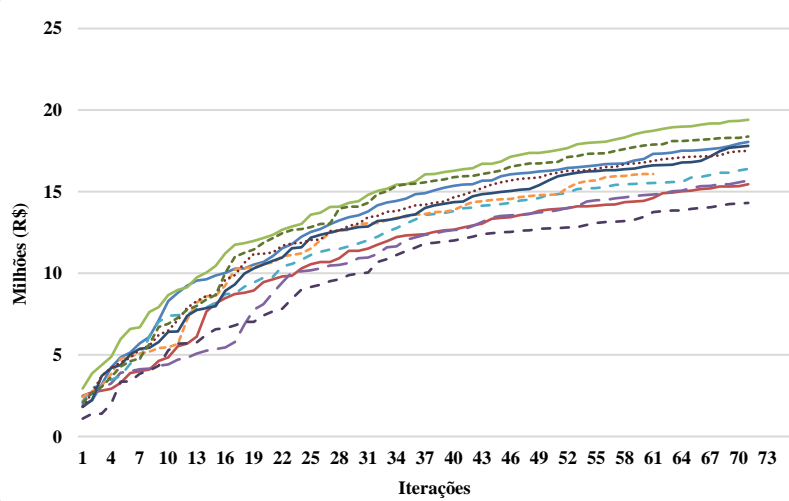
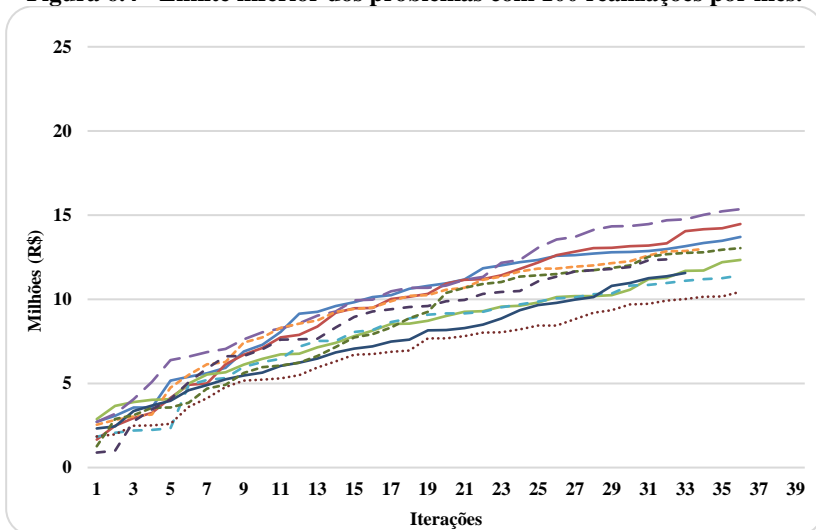


Figura 6.4 - Limite inferior dos problemas com 100 realizações por mês.



Na Tabela 6.7 foi feita uma sumarização dos valores do custo inferior e o respectivo número de iterações necessárias para serem obtidos em 10 minutos, bem como o custo esperado (ou Z_{sup}) com base na simulação *forward* com 10.000 cenários de afluências *out-of-sample*.

Vale salientar que os custos dos limites inferior e superior são associados ao custo da função objetivo dos subproblemas de otimização, não representando efetivamente o custo de operação do sistema, pois penalizações são utilizadas. O custo efetivo de operação será apresentado no momento da discussão da operação do sistema.

De forma geral, constata-se que os valores de Z_{sup} são consideravelmente maiores que os valores de Z_{inf} . Outro aspecto é o fato que a dispersão dos Z_{sup} em cada grupo com mesmo número de realizações é significativamente menor que do Z_{inf} .

De posse dos valores apresentados na Tabela 6.7, é possível calcular a brecha de otimalidade conforme mostrado na Seção 4.5.1 ou nos trabalhos de De Matos (2012) e Larroyd (2016). Na Tabela 6.8 são apresentados os valores médios e desvio padrão dos limites, bem como a brecha de otimalidade estipulada para cada árvore.

Tabela 6.7 – Desempenho computacional do modelo SCP. *Zinf* e *Zsup* em milhões de reais.

20 realizações										
Árvore	1	2	3	4	5	6	7	8	9	10
Zsup	29,5	29,3	29,4	29,5	29,3	29,2	29,5	29,6	29,5	29,3
Zinf	22,1	24,9	17,5	18,3	23,4	25,1	18,5	19,1	17,2	24,1
Iterações	146	164	163	163	164	164	141	163	163	163
50 realizações										
Árvore	1	2	3	4	5	6	7	8	9	10
Zsup	29,5	30,0	29,4	29,8	29,7	29,8	29,6	29,5	29,4	30,3
Zinf	18,1	15,5	19,4	15,7	16,4	16,1	17,8	17,5	18,4	14,3
Iterações	71	71	71	71	71	61	71	71	71	71
100 realizações										
Árvore	1	2	3	4	5	6	7	8	9	10
Zsup	30,1	29,8	29,8	30,1	30,3	30,2	30,5	30,6	30,0	30,6
Zinf	13,7	14,5	12,3	15,4	11,4	13,0	11,6	10,4	13,0	12,4
Iterações	36	36	36	36	36	34	33	36	36	32

Tabela 6.8 - Brecha de otimalidade (%) para o modelo SCP (10 minutos).

Realizações	20	50	100
Média <i>Zinf</i> (Mi. R\$)	21,02	16,91	12,77
Desvio Padrão <i>Zinf</i> (Mi. R\$)	3,20	1,57	1,47
Média <i>Zsup</i> (Mi. R\$)	29,41	29,69	30,21
Desvio Padrão <i>Zsup</i> (Mi. R\$)	1,30	2,86	2,92
Brecha de Otimalidade (%)	36,64	47,52	61,92

De acordo com a Tabela 6.8, pode-se observar que para o modelo SCP e um tempo de execução de 10 minutos a menor brecha é encontrada quando se executa a PDDE com 20 realizações/mês.

6.8 RESULTADOS PARA TODOS OS TEMPOS COMPUTACIONAIS

Na seção anterior foram exibidos os resultados dos *Zinf* e *Zsup* para o caso SCP com a PDDE limitada em 10 minutos. Esta seção mostra os principais resultados com diferentes limites de tempo para a parada da PDDE: 30, 50, 70 e 90 minutos.

As ilustrações das Figura 6.5 e Figura 6.6 mostram os resultados das 10 árvores simuladas para cada tempo e número de realizações.

Figura 6.5 – Limite inferior da PDDE dos casos SCP.

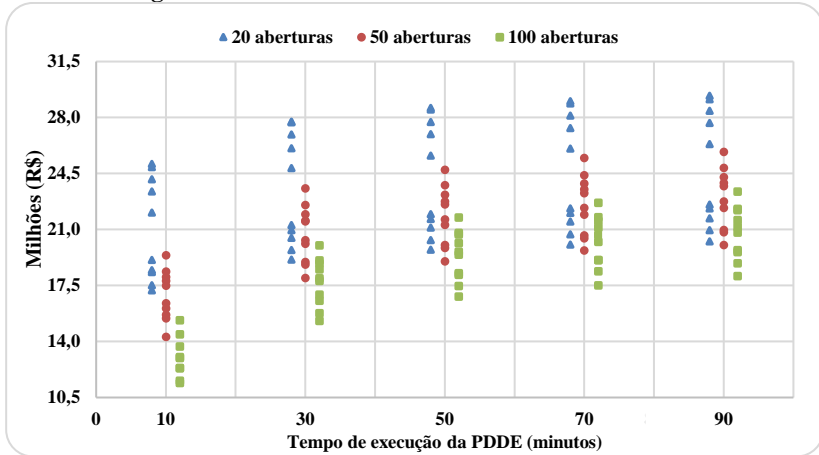
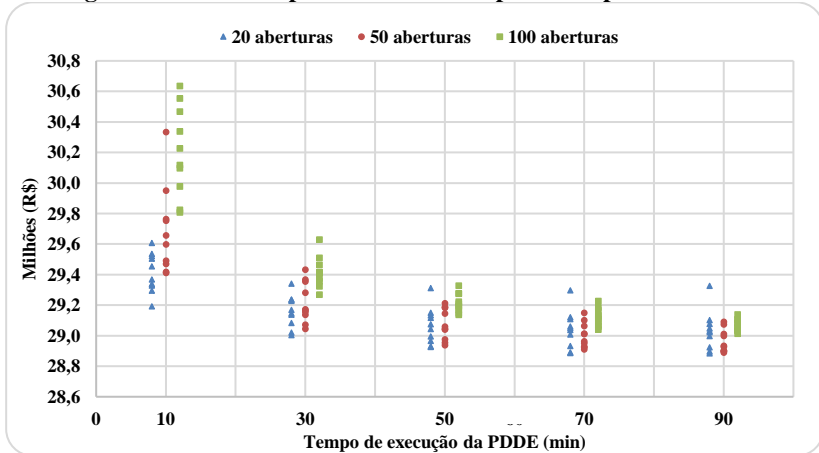


Figura 6.6 – Valor esperado do limite superior do problema SCP.



Por meio da Figura 6.5 nota-se que os valores de *Zinf* aumentam de forma parecida ao longo do tempo entre as árvores com mesmo número de realizações. Essa variação é maior nos casos com menor tempo disponível. Entre os casos com 50 minutos e 90 minutos, os *Zinf* já não possuem grandes variações. Em relação a estimativa do *Zsup*, seu valor diminui

quanto mais refinadas são as FCFs (maior tempo). Também se verifica uma diminuição da dispersão.

Todos os experimentos computacionais apresentados até aqui também foram realizados para os casos com componentes principais. Entretanto, como são muitos, os gráficos relativos aos casos CP16, CP3 e CP1 não são mostrados aqui para não comprometer a continuidade na leitura do documento. Contudo, os mesmos podem ser vistos no Apêndice A. Os custos de otimização dos modelos CP16 e CP3 são maiores que estes, tanto para o limite inferior quanto superior. Enquanto que o modelo CP1 apresenta custos bastante similares ao SCP. Assim, tem-se posse de 600 valores de Z_{inf} e estimativas de Z_{sup} .

Para cada tempo, a brecha de otimalidade é calculada com as médias e desvios padrões do Z_{inf} e Z_{sup} dos grupos das 10 árvores do mesmo modelo e mesmo número de realizações. Dessa forma, a quantidade de aberturas/mês que apresentarem os menores valores para a brecha de otimalidade serão aqueles simulados nas etapas posteriores. As figuras a seguir exibem esse valor para os quatro modelos.

Figura 6.7 - Brecha de otimalidade do modelo SCP.

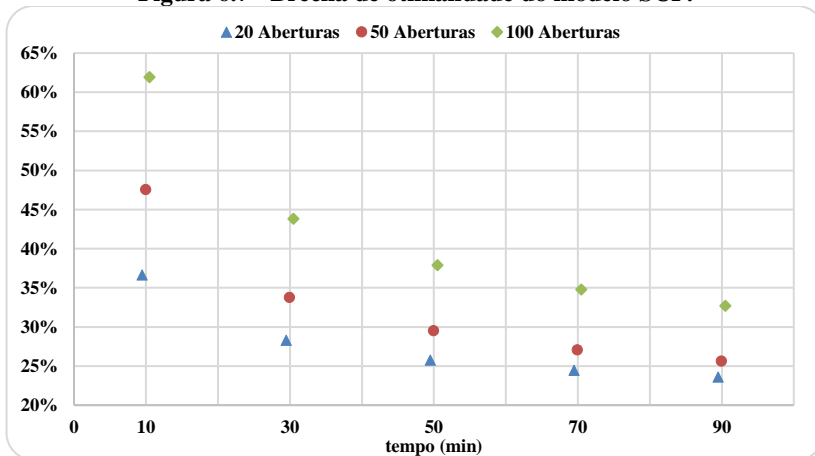


Figura 6.8 - Brecha de otimalidade do modelo CP16.

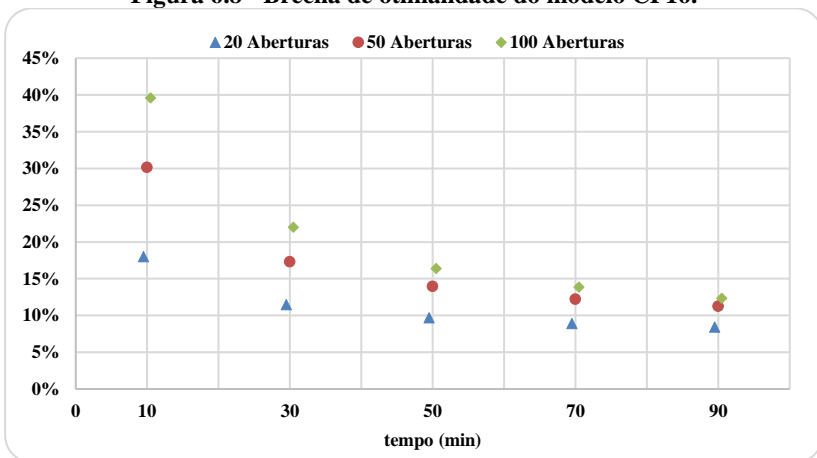


Figura 6.9 - Brecha de otimalidade do modelo CP3.

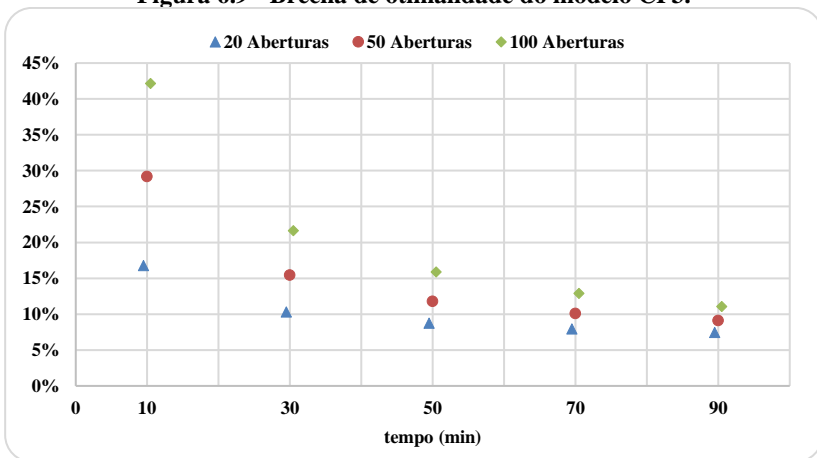
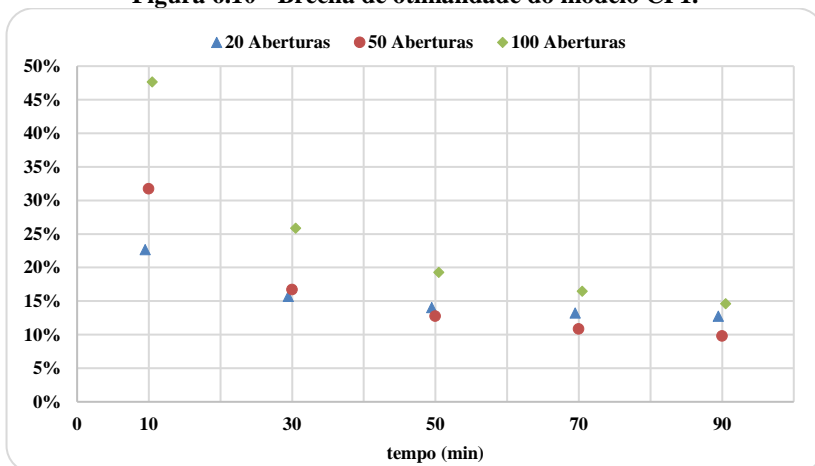


Figura 6.10 - Brecha de otimalidade do modelo CP1.



Vale observar que os valores percentuais da brecha tendem a diminuir conforme aumenta o tempo de execução, isto é, quando aumenta o número de aproximações da FCF.

Como os casos SCP e CP16 possuem um maior número de variáveis aleatórias, é de se esperar que elas necessitem de uma árvore de cenários mais significativa do que o caso com 1 CP. Porém, de acordo com a brecha de otimalidade encontrada, o tamanho de 20 aberturas por mês sempre é melhor para aqueles casos, enquanto que para esse, o tamanho de 50 aberturas é mais adequado a partir de 50 minutos. Um dos possíveis caminhos que levam a isso pode ser pelo fato de que quanto menor o número de variáveis aleatórias, seu corte será mais efetivo e levará mais informação para o problema em comparação aos casos com mais variáveis aleatórias como SCP e CP16.

De posse das brechas apresentadas, a Tabela 6.9 apresenta qual configuração de árvore será utilizada em cada modelo para cada tempo nas simulações posteriores.

Tabela 6.9 - Número de realizações com menor brecha de otimalidade.

Tempo (min.)	10	30	50	70	90
SCP	20	20	20	20	20
CP16	20	20	20	20	20
CP3	20	20	20	20	20
CP1	20	20	50	50	50

Nota-se que a configuração de 20 realizações foi amplamente dominante para o sistema estudado. Apenas nos casos com 50, 70 e 90 minutos do modelo CP1, o uso de árvores com 50 realizações por mês apresentou o menor valor da brecha de otimalidade.

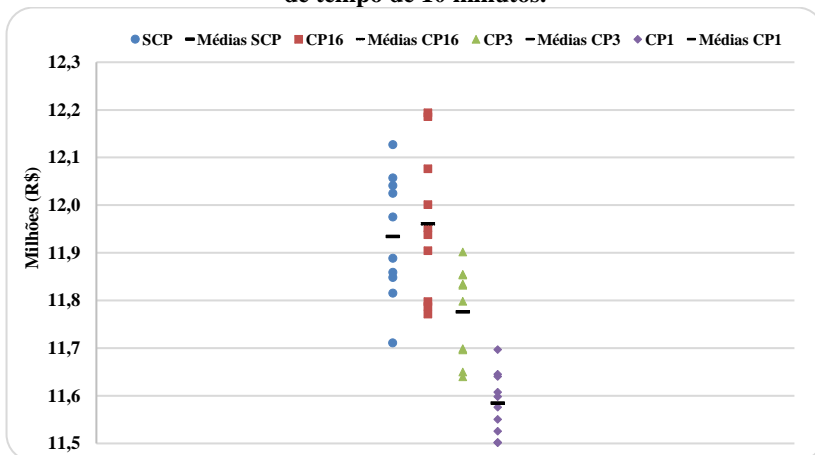
6.9 SIMULAÇÕES OPERATIVAS NAS SÉRIES SINTÉTICAS

As simulações realizadas até então visaram encontrar os melhores tamanho de árvore para cada um dos modelos e tempos de execução. Agora, são realizados sucessivos *forwards* com as FCFs dos casos escolhidos na Tabela 6.9. Os 10.000 cenários de afluência gerados pelos modelos SCP e CP16 (modelos mais detalhados, que tendem a melhor representar o processo estocástico das afluências) são simulados com as FCFs dos demais modelos. Nestes gráficos, é mostrado o custo de operação do sistema e não mais o custo no modelo da PDDE, ou seja, a parcela do custo referente a variável Y_c é retirada do custo da PDDE. Lembra-se que este custo, dependente apenas das vazões afluentes, é o mesmo para todas as sementes. O objetivo é encontrar a política operativa do sistema e verificar as diferenças em comparação com a política do modelo em que as afluências foram geradas.

6.9.1 Simulação das séries sintéticas SCP (10 minutos)

Inicialmente são utilizadas as 10.000 séries sintéticas geradas com o modelo SCP, logo as simulações da operação são procedidas com o cálculo do custo de operação adotando as FCFs dos modelos SCP, CP1, CP3 e CP16 obtidas após 10 minutos de execução da PDDE. A Figura 6.11 apresenta os resultados encontrados para essa condição.

Figura 6.11 - Custo esperado das séries SCP (10 FCFs por modelo) – Limite de tempo de 10 minutos.

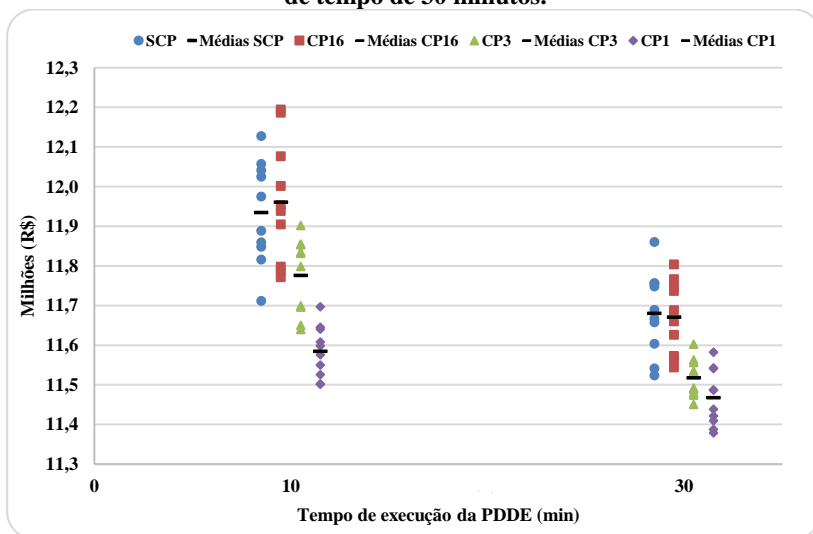


De acordo com a Figura 6.11, observa-se que para um mesmo modelo há uma dispersão significativa entre os custos obtidos com diferentes sementes. Isso ocorre devido ao curto tempo de execução imposto à PDDE. Os modelos SCP e CP16 apresentaram custos de operação muito parecidos, diferentemente dos outros dois modelos que apresentam valores médios bem inferiores.

6.9.2 Simulação das séries sintéticas SCP (30 minutos)

O próximo passo consiste na simulação das mesmas 10.000 séries SCP de aflúências nas FCFs obtidas com 30 minutos de execução da PDDE. Os resultados, juntamente com os valores da Figura 6.11, são exibidos na Figura 6.12.

Figura 6.12 - Custo esperado das séries SCP (10 FCFs por modelo) – Limite de tempo de 30 minutos.

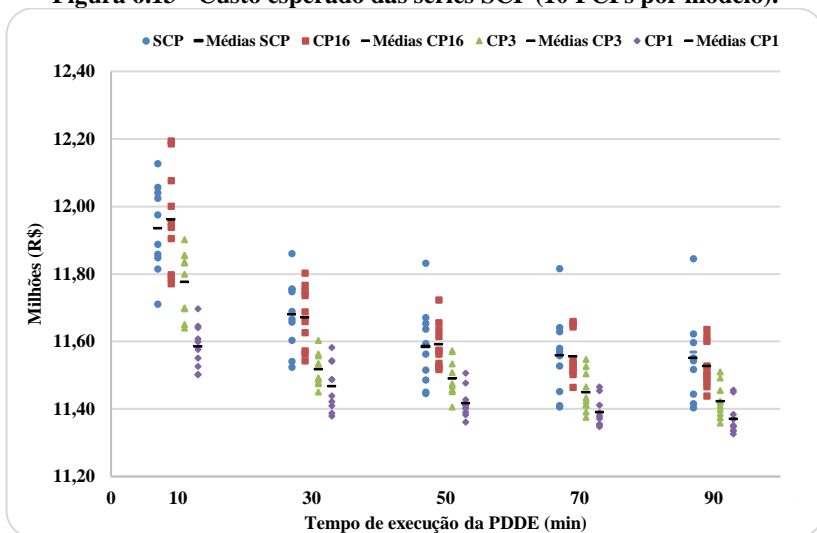


De acordo com a figura acima, é perceptível a diminuição dos custos de operação de 10 para 30 minutos de simulação. Isso ocorre pelo aumento do número de aproximações lineares da FCF. Por exemplo, a média do modelo SCP diminuiu de R\$ 11,93 milhões para R\$ 11,68 milhões. No modelo CP16 a mudança foi de R\$ 11,96 milhões para R\$ 11,67 milhões, enquanto que no CP3 o custo mudou de R\$ 11,77 milhões para R\$ 11,52 milhões. Por fim, para no modelo CP1 a variação foi de R\$ 11,58 milhões para 11,47 milhões de reais.

6.9.3 Simulação das séries sintéticas SCP (todos os tempos)

O mesmo que foi feito para as sementes de 10 e 30 minutos é feito para 50, 70 e 90 minutos. O gráfico dos custos encontrados pode ser visto na Figura 6.13.

Figura 6.13 - Custo esperado das séries SCP (10 FCFs por modelo).



Na comparação da performance entre os modelos utilizados, constata-se que a operação em FCFs do modelo SCP apresentou custos maiores, mesmo com as séries simuladas terem sido originadas em seu modelo. O modelo CP1 apresentou menores custos na maioria dos casos. Tal fato leva a acreditar que, caso o modelo SCP fosse o apontado como o melhor modelo a representar o processo estocástico das afluições, o modelo a ser utilizado na PDDE para prover o menor custo de operação do sistema seria o CP1.

6.9.4 Simulação das séries sintéticas CP16, CP3 e CP1

Em caso de se assumir que ou o modelo CP16, ou o CP3 ou o CP1 melhor representa o processo estocástico das afluições, o mesmo processo de simulação da operação feito com as séries SCP (apresentado na seção anterior) foi feito com as 10.000 séries sintéticas geradas agora com cada um dos respectivos modelos. A Figura 6.14 exibe os custos esperados de operação encontrados para as séries CP16. Enquanto que a Figura 6.15 e Figura 6.16 mostram para as séries CP3 e CP1, respectivamente.

Figura 6.14 - Custo esperado das séries CP16 (10 FCFs por modelo) .

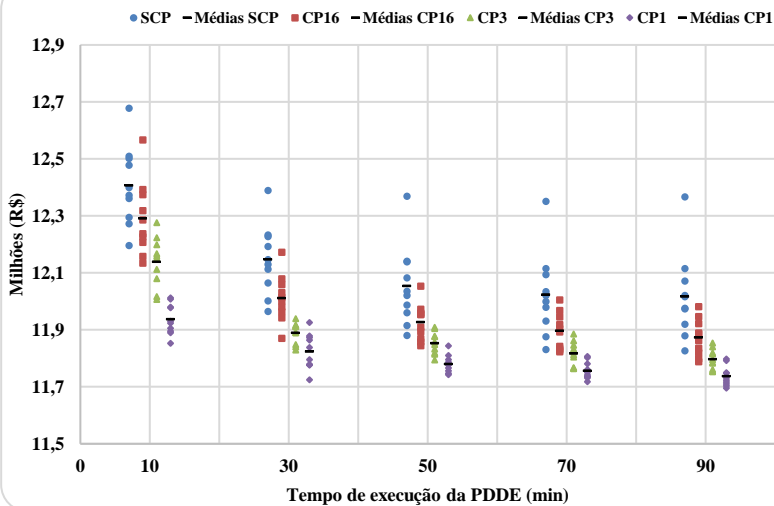


Figura 6.15 - Custo esperado das séries CP3 (10 FCFs por modelo).

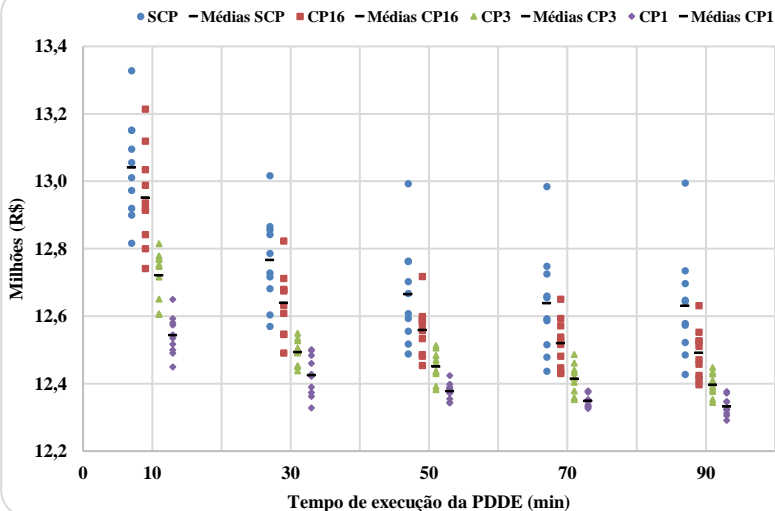
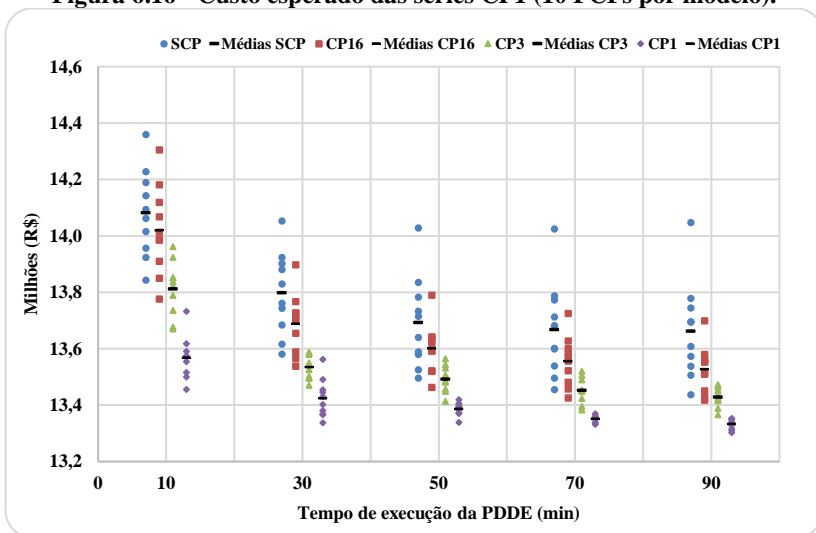


Figura 6.16 - Custo esperado das séries CP1 (10 FCFs por modelo).

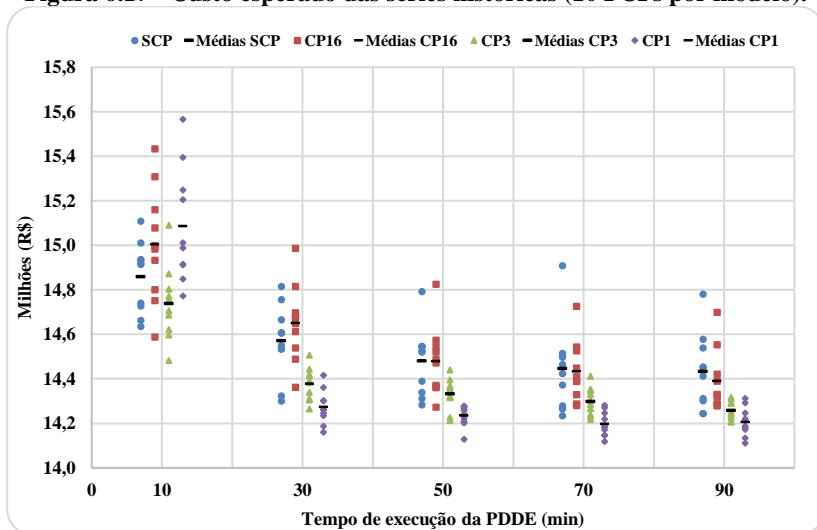


Para estas condições, o “desenho” segue a mesma tendência do caso anterior da Seção 6.9.3, pelos mesmos fatores. Em todas as simulações envolvendo as componentes principais (CP16, CP3 e CP1) o modelo SCP, em geral, apresentou maiores custos de operação. Os menores custos esperados de operação quase sempre foram obtidos com o modelo CP1 na PDDE, independentemente de qual modelo proveu as séries utilizadas na simulação.

6.10 SIMULAÇÕES OPERATIVAS NAS SÉRIES HISTÓRICAS

As afluições encontradas por cada modelo de série temporal deste trabalho foram geradas a partir dos registros históricos. Por esta razão, é interessante realizar as mesmas simulações com as afluições históricas, mesmo com um número bastante reduzido de cenários. O histórico sintético utilizado é de janeiro de 1931 a dezembro de 2010, totalizando 80 anos. Porém, como o estudo é de quatro anos, é possível formar 77 cenários com este histórico. A Figura 6.17 mostra os resultados encontrados quando a série histórica é aplicada aos modelos.

Figura 6.17 - Custo esperado das séries históricas (10 FCFs por modelo).



Observa-se que com 10 minutos, o maior custo de operação tende a ser provido pelo modelo CP1 na PDDE. A partir de 30 minutos, esse modelo se torna o mais barato quando se analisam a média e dispersão dos custos. O modelo SCP perde vantagem para os demais conforme mais tempo se passa executando a PDDE e, conseqüentemente, mais aproximações da FCF são calculadas.

6.11 CUSTOS DE GERAÇÃO TERMELÉTRICA E DÉFICIT

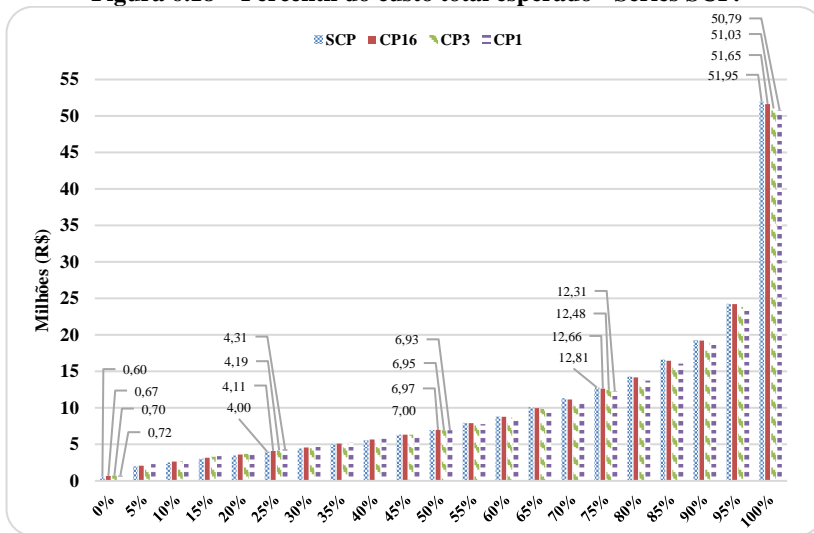
Os resultados referentes aos custos totais de operação médios, relativos a um certo tempo de simulação foram mostrados nas seções anteriores. Porém, é importante também expor alguns resultados sobre a dispersão dos custos ao longo das séries, caracterizando o comportamento da operação frente à variação dos cenários encontrados. Para os quatro modelos, são apresentados resultados referentes aos custos de geração térmica e déficit passados para valor presente de cada um dos 10.000 cenários.

Para cada modelo, será escolhido um caso e sua respectiva melhor abertura com 90 minutos de execução da PDDE (20 realizações/mês para SCP, 16 CP e 3 CP, e 50 realizações/mês para uma CP). Como são 10 FCFs, a média dos 10 custos de cada uma das séries simuladas é que será apresentada.

6.11.1 Custos Totais por Série Sintética SCP

Para eficiente visualização dos resultados, os custos dos referentes à simulação dos quatro modelos nos 10.000 cenários de afluência gerados pelo modelo SCP são ordenados em ordem crescente de valor. Na Figura 6.18, o valor 100% é referente ao custo no cenário mais caro.

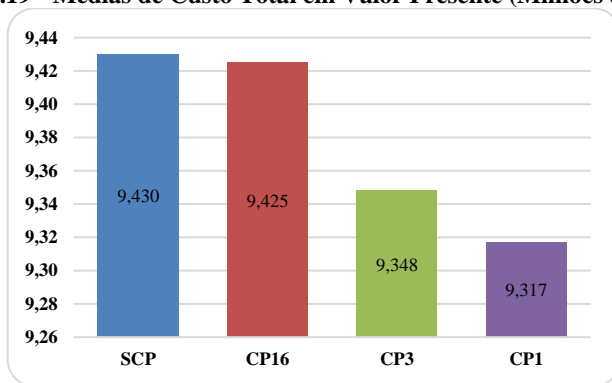
Figura 6.18 – Percentil do custo total esperado - Séries SCP.



De acordo com a Figura 6.18, é possível notar que os custos são muito parecidos entre os modelos. Nos cenários mais baratos, o modelo SCP é levemente menos oneroso. Já nos cenários mais caros, este é o que leva desvantagem perante os demais modelos. O 5000º cenário mais caro (mediana ou 50% no gráfico) mostra os casos com custos bastante parecidos, sendo o caso CP1 com o menor custo.

Nota-se também que os 5% cenários mais caros crescem de forma mais intensa em comparação aos cenários mais baratos. Todos os modelos apresentaram o maior custo no cenário 6.327.

Uma outra forma de comparação é com a média geral dos custos. A ilustração a seguir (Figura 6.19) mostra essa informação.

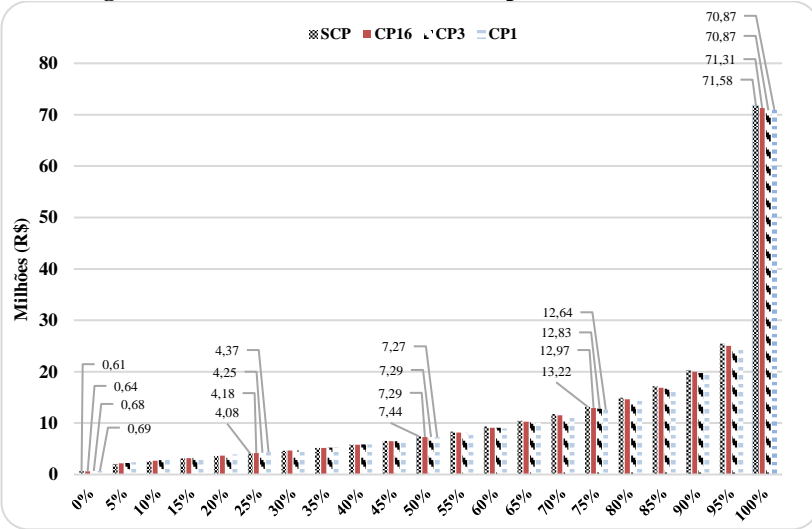
Figura 6.19 - Médias de Custo Total em Valor Presente (Milhões de Reais).

A figura mostra que os modelos de SCP e CP16 possuem médias equivalentes. As médias dos outros dois casos com componentes são mais baixas, com destaque para o modelo de CP1, menor entre todos.

6.11.2 Custos Totais por Série Sintética CP16

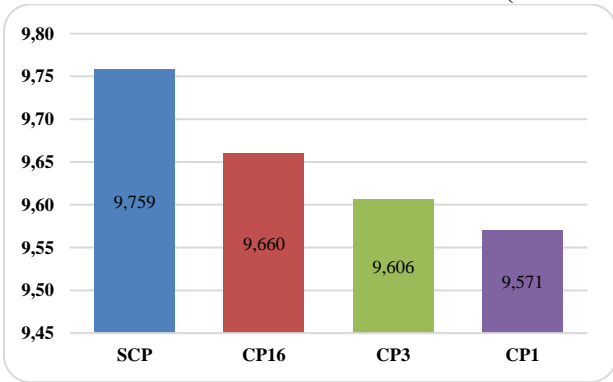
O mesmo procedimento da seção anterior é feito, com a diferença que agora os cenários são feitos usando o modelo com 16 CP. O gráfico da Figura 6.20 exhibe os custos ordenados em ordem crescente.

Figura 6.20 – Percentil do custo total esperado - Séries CP16.



Os custos dos modelos para estes cenários também são bastante similares. O comportamento permanece da mesma forma ao anterior, com melhor performance para o modelo SCP nos cenários mais baratos, enquanto para os cenários mais caros o modelo CP1 leva vantagem. A melhor mediana permanece no modelo de 1 CP, embora os valores estejam muito próximos. O cenário de número 6.481 foi o mais oneroso para todos os modelos.

Figura 6.21 - Médias de Custo Total em Valor Presente (Milhões de Reais).

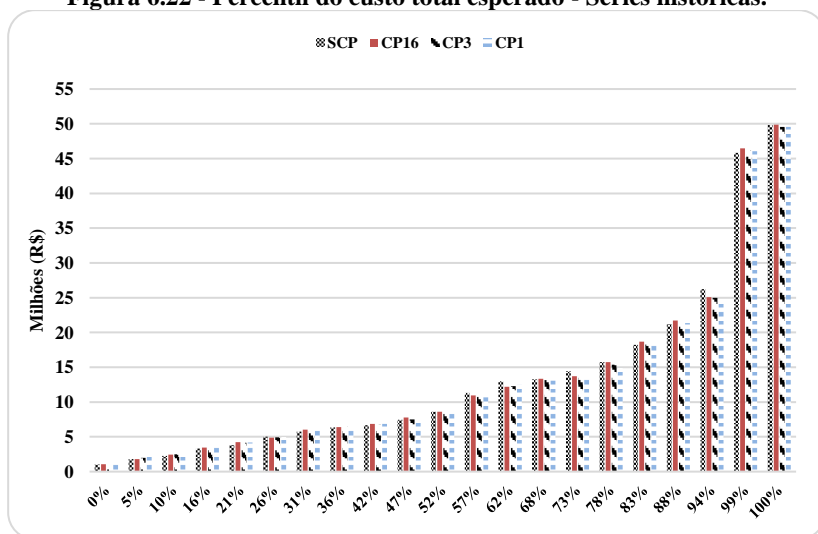


De acordo com a Figura 6.21, a maior média dos custos também ocorre no modelo SCP e diminui gradativamente até o modelo CPI. A diferença entre as médias do modelo sem componentes principais e com uma CP é de cerca de 2%.

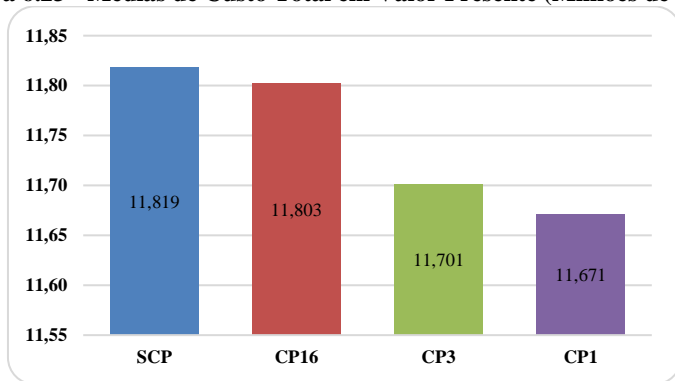
6.11.3 Custos Totais por Série Histórica

Por fim, as FCFs das 10 árvores dos quatro modelos para 90 minutos de execução de PDDE e com as respectivas configurações de abertura por mês são simulados para os cenários do histórico. Neste caso, o número de cenários é bastante reduzido em comparação com as séries sintéticas geradas. Os resultados do percentil podem ser vistos na Figura 6.22 e os custos médios na Figura 6.23.

Figura 6.22 - Percentil do custo total esperado - Séries históricas.



Mesmo com número consideravelmente inferior de cenários, o desempenho dos modelos parece não variar muito. Todos os modelos apresentam custos próximos.

Figura 6.23 - Médias de Custo Total em Valor Presente (Milhões de Reais).

Com relação aos valores médios, novamente o modelo de uma componente, mesmo que ligeiramente, apresenta menor valor. Assim como o visto na última e penúltima seção, o cenário mais caro é o mesmo em todos os modelos (cenário 13).

6.12 DESPACHO TERMELÉTRICO

Os resultados referentes aos custos totais de operação, relativos a um certo tempo de simulação e a análise de custos de operação foram mostrados nas seções anteriores. Porém, é importante também expor alguns resultados relativos à geração térmica e déficit entre os modelos.

Como a série sintética é muito grande, será escolhido como base os cenários do modelo que obteve melhor performance dos resultados da Seção 6.11. De forma geral, em todas as análises o modelo com uma componente principal registrou menores custos.

Dessa forma, inicialmente o cenário mais caro (de cada uma das séries obtidas por meio dos modelos SCP, CP16 e histórica) encontrado pelo modelo de uma CP será o cenário base, ou aquele com custo 100%. O cenário que apresentar o custo mais próximo a 75% deste será o cenário 2. O cenário 3 apresentará o custo referente à 50% e o quarto cenário será referente à 25%. O cenário mais barato obtido no modelo de uma componente principal será o cenário 5.

Os gráficos a seguir ilustram os valores de geração térmica + déficit em cenários presentes no histórico.

Figura 6.24 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 100% (mais caro) da operação por FCFs 1CP na série histórica.

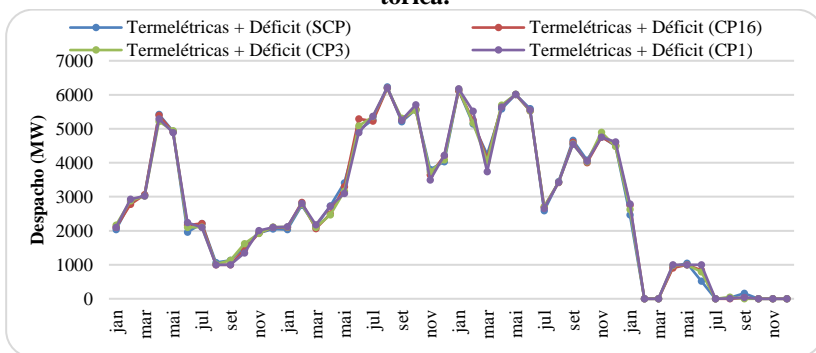


Figura 6.25 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 75% da operação por FCFs 1CP na série histórica.

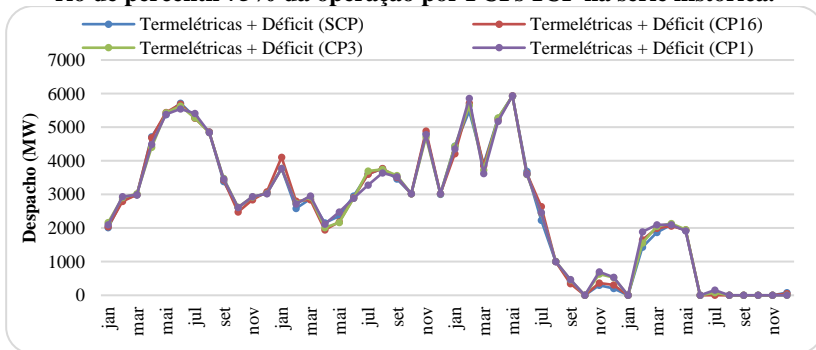


Figura 6.26 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 50% da operação por FCFs 1CP na série histórica.

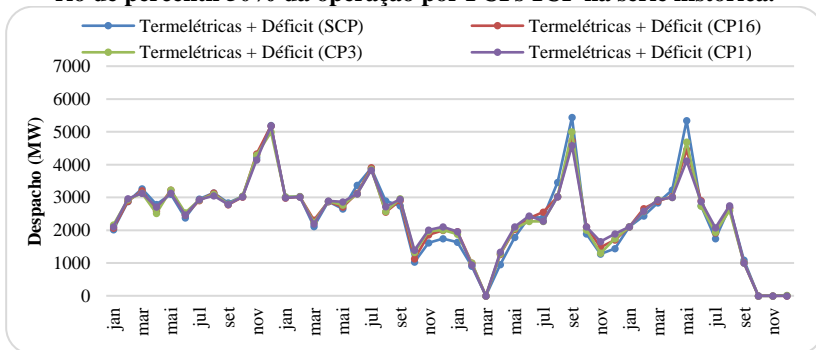


Figura 6.27 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 25% da operação por FCFs 1CP na série histórica.

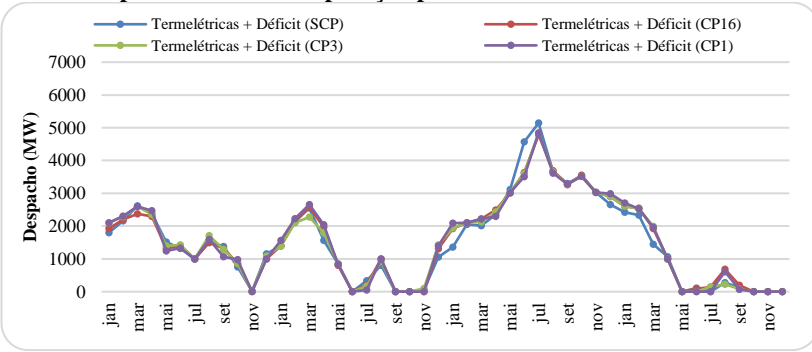
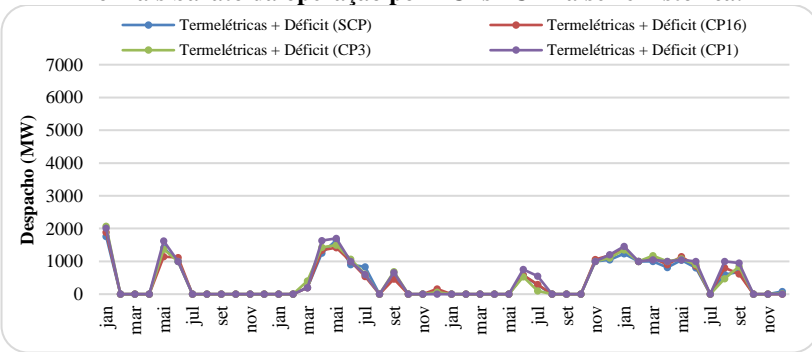


Figura 6.28 - Simulações de operação (FCFs de todos os modelos) no cenário mais barato da operação por FCFs 1CP na série histórica.



De acordo com as figuras anteriores, observa-se entre os modelos despacho próximo ao longo dos 48 meses. O modelo sem componentes principais é o mais caro, pois de forma geral utiliza um pouco mais o déficit nos piores cenários. Como é de se esperar, o cenário mais caro apresenta maior despacho das térmicas do que os outros cenários abordados.

Por fim, os gráficos referentes às simulações nas séries obtidas pelos modelos SCP e CP16 encontram-se no Apêndice B e C respectivamente.

6.13 CONVERGÊNCIA E NÚMERO ITERAÇÕES

Todos os principais resultados e performance de cada modelo frente às condições propostas foram apresentados. No processo, o critério de parada foi baseado no tempo de execução do algoritmo e não sobre um valor de tolerância para o algoritmo. Por isso, o próximo passo consiste em mostrar características sobre a convergência das árvores em que o algoritmo da PDDE foi executado.

Dado que cada um dos modelos tem quantidades diferentes de variáveis, o número de iterações varia para um mesmo tempo e de modelo para modelo. A finalidade é descobrir qual a convergência atingida por cada uma das árvores ao fim de cada tempo. Porém, um processo com maior número de iterações que outra não terá, necessariamente, melhor convergência.

Aqui, a análise de convergência é feita com relação a estabilização do custo inferior. Esta estabilização leva em conta o crescimento percentual do custo de uma iteração k com a iteração anterior. Por isso, o procedimento utilizado consiste em verificar para as 10 árvores na execução de 10, 30, 50, 70 e 90 minutos da PDDE, a média do crescimento percentual nas últimas 3, 5 e 10 iterações. Ressalta-se que foi utilizado o *Zinf* das árvores com melhor configuração de realizações por mês para cada tempo citado.

O gráfico da Figura 6.29 apresenta para os quatro modelos os resultados das médias do crescimento percentual das últimas 3 iterações. A Figura 6.30 das últimas cinco e a Figura 6.31 das últimas 10 iterações. Em seus estudos, o setor elétrico utiliza como critério de parada a estabilização do *Zinf* com crescimento inferior a 0,2% nas últimas 3 iterações.

De acordo com esses gráficos, para todos os testes de convergência quanto menor o tempo, maior é o percentual. Isso se deve ao fato de haver um número menor de iterações e o *Zinf* tende a crescer mais intensamente nas iterações iniciais. Por isso que esse valor diminui gradativamente conforme o tempo de execução da PDDE aumenta. Nos modelos CP16 e CP3 este valor é menor que 0,005% para 70 e 90 minutos.

Para o modelo CP1 e tempo superior ou igual a 50 minutos, o crescimento percentual e desvio padrão é maior frente aos outros modelos. A razão para isto acontecer é pelo fato de este operar com 50 realizações por mês (todos os demais são com 20 aberturas/mês). E cada iteração com 50 realizações é mais demorada em comparação a 20 realizações. O ponto positivo de ter mais aberturas por estágio é que as aproximações lineares da FCF são melhor refinadas.

Figura 6.29 – Variação média de *Zinf* nas últimas três iterações.

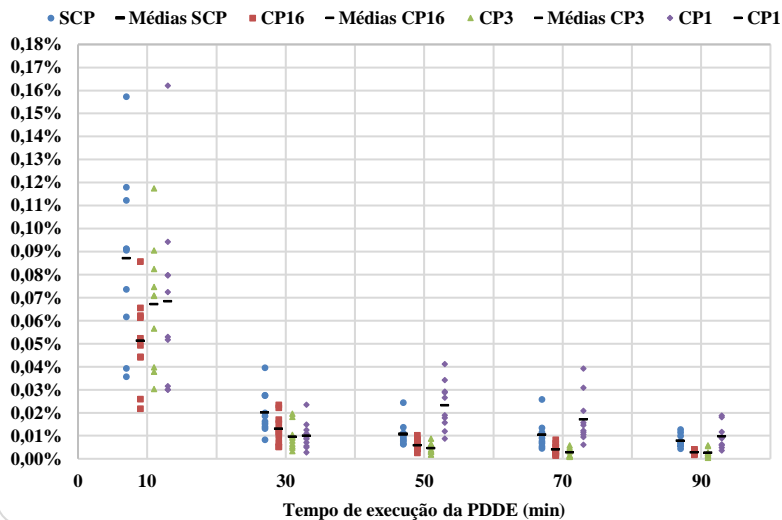


Figura 6.30 - Variação média de *Zinf* nas últimas cinco iterações.

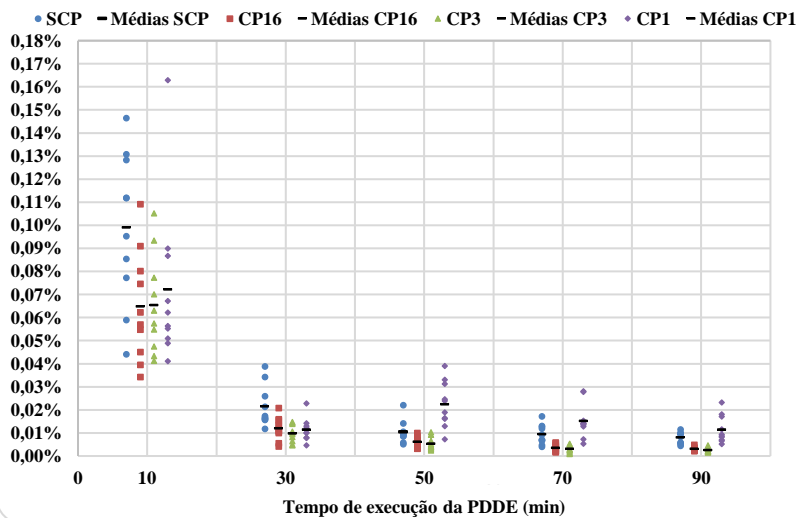
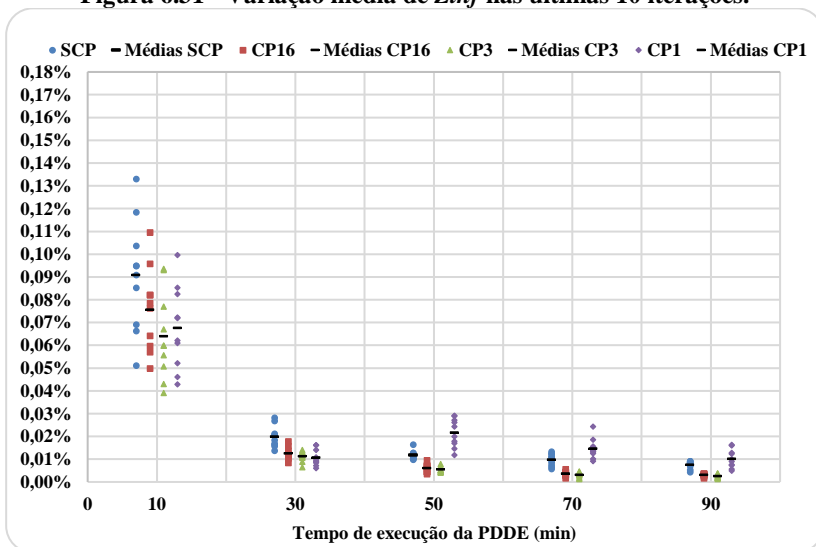


Figura 6.31 - Variação média de *Zinf* nas últimas 10 iterações.

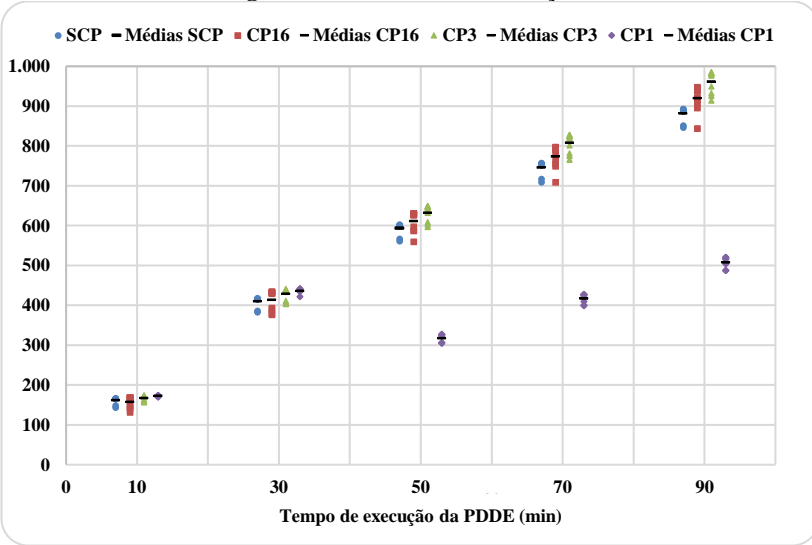


No setor elétrico utiliza-se um critério de parada porque ao atender a este critério, acredita-se que a política encontrada é suficientemente boa para tomar uma decisão.

Ao adotar esse valor (crescimento relativo do *Zinf* inferior a 0,2% nas últimas 3 iterações) percebe-se que para este problema, 10 minutos já seria um tempo que atendesse a este critério. Porém, os resultados mostram que para este problema o custo superior *Zsup* continua diminuindo substancialmente para maiores tempos de simulação (Figura 6.13 a Figura 6.17). Isso quer dizer que, se o critério de parada adotado no sistema elétrico fosse utilizado para este problema, o algoritmo seria interrompido de maneira precoce, pois há um ganho satisfatório ao simular com tempos maiores a 10 minutos.

As informações do número de iterações da PDDE para cada modelo e realizações por mês são mostradas a seguir.

Figura 6.32 - Número de iterações.



Observa-se que a quantidade de iterações é maior quanto menor é o número de aberturas. Como na etapa de *backward* menos cenários são percorridos, menor será o tempo por iteração.

Na comparação entre os modelos, o modelo sem componentes principais obteve menos iterações. Este fato é mais visível para tempos de execução mais longos. Os modelos com número reduzido de CP apresentaram quantidades maiores de iterações, pois possuem menos variáveis de estado, e portanto, menor tempo por iteração.

Entretanto, da mesma forma com o que ocorreu nos gráficos da convergência, os resultados do modelo CP1 para tempo igual ou superior a 50 minutos parece pior em comparação aos demais modelos. O motivo continua o mesmo e reforça que mesmo com menor número de iterações e “pior” convergência, esse modelo com 50 realizações propiciou menor brecha de otimalidade quando comparado às outras quantidades de realizações. O Apêndice D mostra o número de iterações obtido por tamanho de árvore.

Vale analisar também o tempo de cada iteração para cada um dos modelos. A Tabela 6.10 exibe o tempo médio das 10 sementes nas últimas iterações para os 4 modelos e número de realizações. Como cada árvore de cada modelo percorre uma quantidade diferente de iterações ao longo

dos 90 minutos do processo iterativo, um valor comum entre todas as sessões foi escolhido. Isto é feito para que se possa comparar casos com o mesmo número de aproximações lineares da FCF.

Tabela 6.10 – Tempo médio por iteração

#Iteração	Realizações	SCP	CP16	CP3	CP1
840	20	9,04	8,20	7,59	7,17
450	50	14,96	13,72	13,42	12,90
250	100	22,44	20,73	20,33	20,54

Nota-se que os tempos médios dos casos com CP são sempre inferiores ao caso sem as componentes principais. Fato que se mostra percentualmente maior quanto mais iterações o algoritmo percorre. A diferença ultrapassa 20% na comparação entre os modelos SCP e CP1 com 20 realizações por estágio.

7. CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

O problema de planejamento da operação energética de médio prazo tem o objetivo de encontrar, para o sistema estudado, a melhor estratégia de solução. Isto é, a política operativa que minimize o custo esperado de geração termelétrica para todo o período de planejamento. Este trabalho utiliza uma técnica estatística, a Análise de Componentes Principais para encontrar essa solução. Seu uso almeja encontrar avanços frente a performance do modelo de usinas individualizadas.

Para sistemas hidrotérmicos de grande porte, como é o caso Brasileiro, esse trabalho se torna bastante complexo. Por isso, como explicado no Capítulo 1, é necessária a discretização das etapas de planejamento. E dependendo de qual etapa se trabalha, algumas características possuem maior relevância, tendo que ser satisfatoriamente modeladas, enquanto outras não.

No caso do PMP, a característica estocástica das afluições é um fator que deve ser muito bem modelado, pois estes têm influência direta na construção das FCFs. Tem-se que as não linearidades em respeito à função de produção hidrelétrica caracterizam o problema como não linear. Por essa razão, conforme é mostrado na Seção 2.4, estratégias de linearização utilizando o algoritmo de *Convex Hull* são adotadas. Com discretização em Q, V e S, aproximações lineares são geradas para representar as UHEs. Sua qualidade foi avaliada calculando o Erro Médio Quadrático. O valor médio de EMQ das UHEs apresentadas na dissertação foi de 1,3%. Este valor mostra um bom nível de precisão, tendo em vista o baixo número de aproximações lineares relacionadas a cada usina.

Como a natureza das afluições do problema é estocástica, faz-se necessário o emprego de técnicas que abordem essa característica. Por isso, o algoritmo utilizado foi a PDDE, que não faz a discretização do espaço de estados.

Conforme discutido no Capítulo 3, o modelo de geração de afluições sintéticas é o PAR Lognormal a 3 Parâmetros. Ele produz FCFs válidas e consegue reproduzir de maneira satisfatória as correlações temporal e espacial das afluições. Como os modelos agregados de usinas não exploram plenamente os recursos do parque hidráulico, a representação individualizada das usinas foi utilizada. Paralelamente, a geração de séries sintéticas com o uso da ACP foi adotada para consequente comparação. Os modelos com componentes principais representaram convenien-

temente quando se observa em paralelo a estatística observadas nos registros históricos. Lembra-se que a correlação espacial entre as hidrelétricas é afetada quanto menor é o número de CP adotadas.

O número de CP utilizadas pode mudar conforme a necessidade requerida, variando de uma até o número total de hidrelétricas do sistema. Quanto menor este número, menor será o número de variáveis do problema. Então, adotou-se nas simulações três modelos distintos com componentes principais. Um modelo com a quantidade total de componentes principais, um modelo com 3 e outro com apenas uma CP.

Uma questão de grande importância no trabalho se diz respeito ao tempo computacional. Como o sistema proposto não é muito grande, o tempo máximo de execução da PDDE para a construção dos cortes de Benders foi de 1,5 horas. A qualidade da solução dos diferentes modelos testados também foi observada para valores de tempo menores que esse.

Outro ponto de bastante relevância no tocante à qualidade da solução é o número de realizações por estágio. No trabalho, adotou-se três tamanhos diferentes (20, 50 e 100), em que o mais conveniente é escolhido dependendo do modelo empregado e do tempo de execução do algoritmo. Tal valor é escolhido mensurando com 95% de confiança o *gap* de otimalidade de cada conjunto de árvores de cenários. Este conjunto é composto por árvores de 10 sementes distintas.

Com relação aos resultados, eles foram obtidos por meio da simulação de 10.000 cenários sintéticos elaborados por cada um dos modelos adotados. Além das simulações com dados sintéticos, os valores reais presentes no histórico de afluições também foram empregados.

Para todas as simulações realizadas os custos mostram vantagem, ainda que pequena, para o modelo com uma componente principal. Os custos esperados entre os modelos SCP e CP16 são praticamente equivalentes quando simulados na série sintética do modelo sem CP. Para essa mesma série os valores de custo médio encontrados nos modelos CP3 e CP1 são ligeiramente menores, com destaque para o modelo com apenas a primeira componente principal. Já para a simulação com as séries sintéticas geradas pelos modelos com as CP, os custos médios decrescem do modelo SCP para o modelo CP1.

Quando a simulação é realizada para a série de afluições incrementais do histórico, o modelo CP1 ainda possui o menor custo médio, exceto para o tempo de 10 minutos. Isso mostra mais uma vez sua superioridade frente aos outros modelos.

O custo gerado pelas termelétricas do sistema e déficit também foi avaliado. Nas afluições sintéticas do modelo com afluições incrementais, até cerca de 40% dos cenários mais baratos, o modelo sem CP aponta

menores custos médios. Em contrapartida, é o mais caro nos piores cenários, o que faz ele ser o modelo mais oneroso. Já para os cenários do histórico de incrementais, um comportamento similar é identificado.

Por fim, o número de iterações após a execução da PDDE para cada modelo, número de realizações por mês e para cada um dos tempos de estudos são averiguados. Verifica-se que a quantidade de iterações dos métodos com poucas componentes (1 e 3) e consequentemente com menos variáveis, obtiveram mais iterações em relação ao método sem e com todas as componentes principais. Há casos que essa superioridade ultrapassa 10%.

7.1 SUGESTÕES PARA TRABALHOS FUTUROS

Este trabalho abordou muitas questões acerca do PMP. Uma é relativa à qualidade da política operativa em relação ao número de realizações por estágio de cada árvore de cenários. Mas a contribuição principal refere-se à introdução da Análise de Componentes Principais no problema. Contudo, muitas outras possibilidades podem ser exploradas como uma possível sequência dos estudos realizados.

- Inclusão de aversão a risco no modelo do PMP. Decisões tomadas com aversão a risco provocam um aumento da segurança do sistema, já que o gerenciamento dos recursos hídricos no presente tem grande influência no atendimento do sistema no futuro;
- A ACP foi utilizada no trabalho para diminuir o número de variáveis de estado relacionadas às afluições incrementais das hidrelétricas. Porém, esta técnica também pode ser empregada nos volumes dos reservatórios, que são variáveis de estado do problema;
- Os coeficientes autoregressivos para o sistema abordado neste trabalho possuem em média ordem 1. É interessante aplicar a técnica de ACP para outros sistemas hidrotérmicos com ordem maior que essa, pois a técnica com componentes principais pode se mostrar ainda mais atrativa;
- Utilizar seleção de cortes na PDDE, já que o tempo computacional cresce ao longo do processo iterativo com o aumento da quantidade de cortes. Portanto, ao selecionar um

conjunto reduzido de cortes, sempre atentando para a qualidade da solução encontrada, o tempo computacional diminuirá;

- Combinar modelos com diferentes quantidades de CP em um mesmo problema da PDDE;
- Nas simulações desta dissertação é percorrido apenas um *forward* por iteração. Para acelerar o processo é interessante um maior número de cenários por iteração nas recursões progressivas.

8. APÊNDICE A: Limite inferior e superior do problema

Figura 8.1 - Limite inferior da PDDE dos casos CP16.

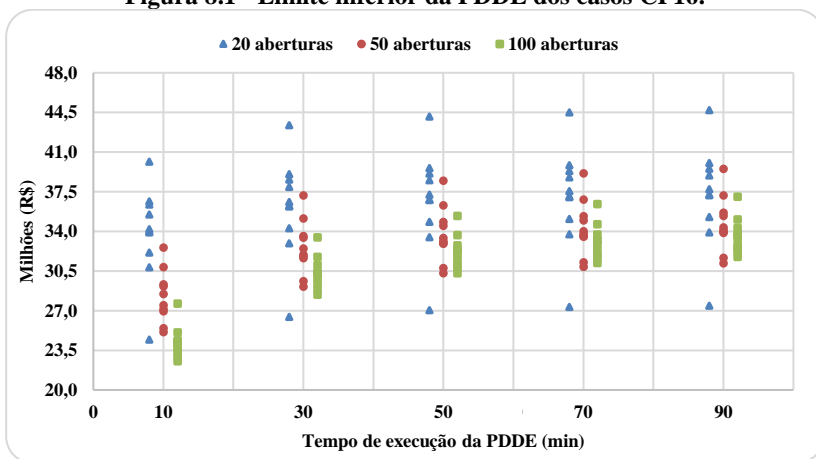


Figura 8.2 - Valor esperado do limite superior do problema CP16.

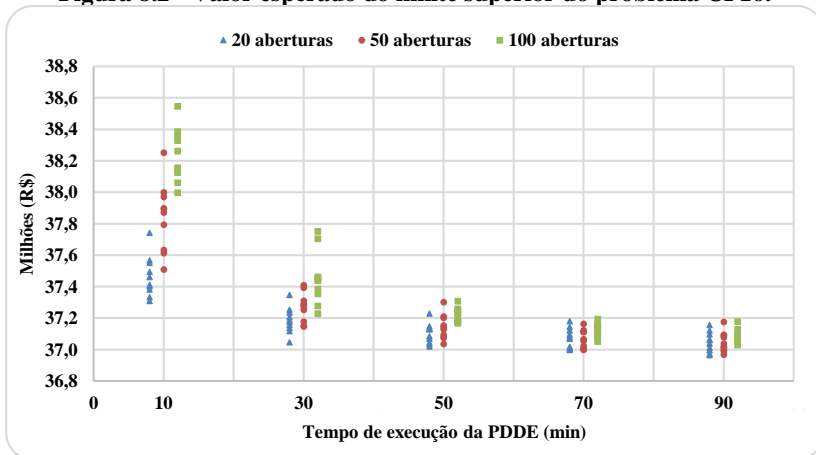


Figura 8.3 - Limite inferior da PDDE dos casos CP3.

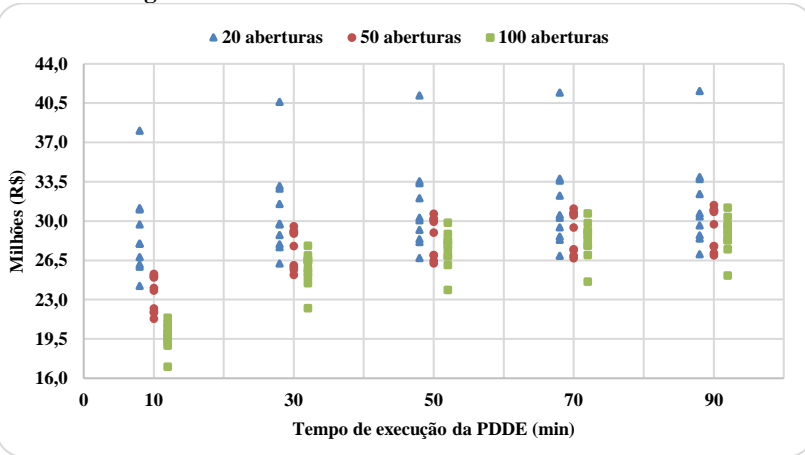


Figura 8.4 - Valor esperado do limite superior do problema CP3.

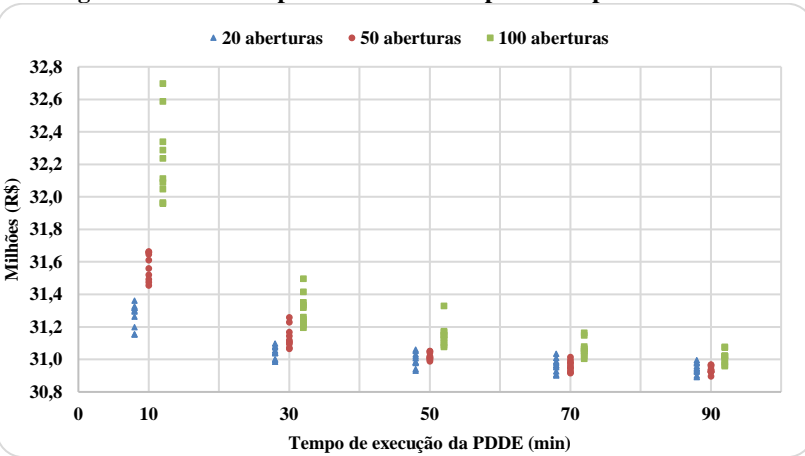


Figura 8.5 - Limite inferior da PDDE dos casos CPI.

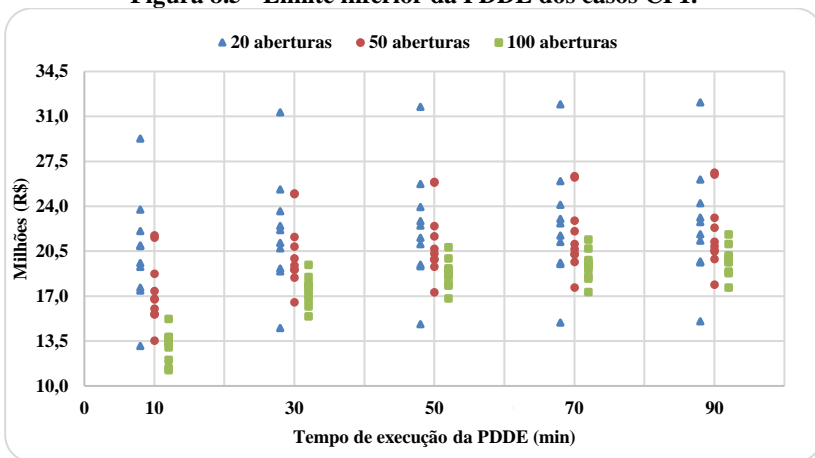
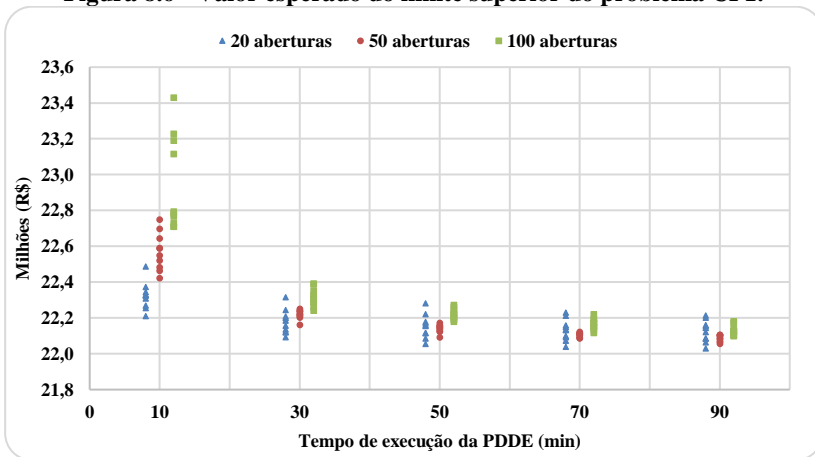


Figura 8.6 - Valor esperado do limite superior do problema CPI.



9. APÊNDICE B: Despacho termelétrico para a série SCP

Figura 9.1 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 100% (mais caro) da operação por FCFs 1CP em séries SCP.

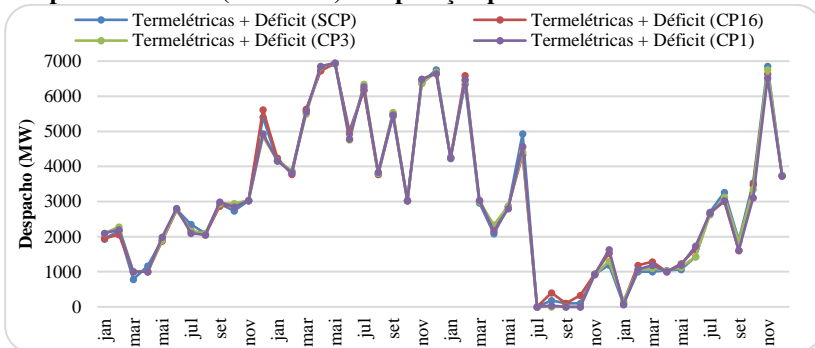


Figura 9.2 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 75% da operação por FCFs 1CP em séries SCP.

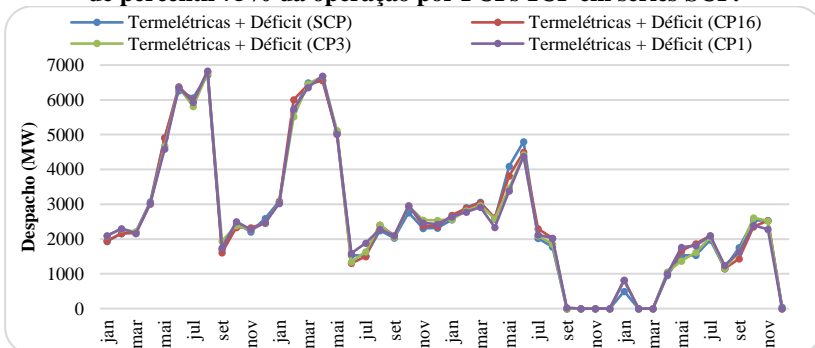


Figura 9.3 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 50% da operação por FCFs 1CP em séries SCP.

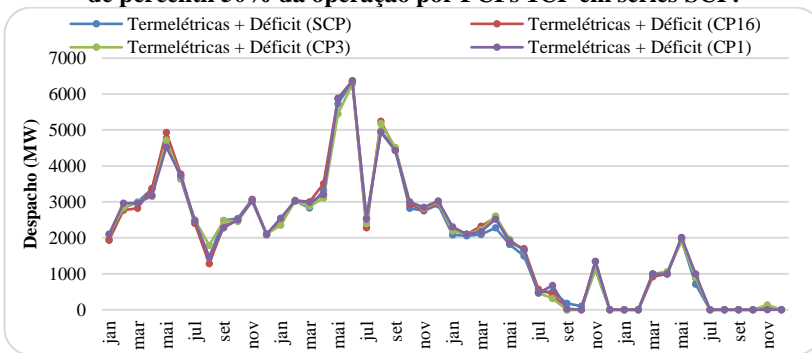


Figura 9.4 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 25% da operação por FCFs 1CP em séries SCP.

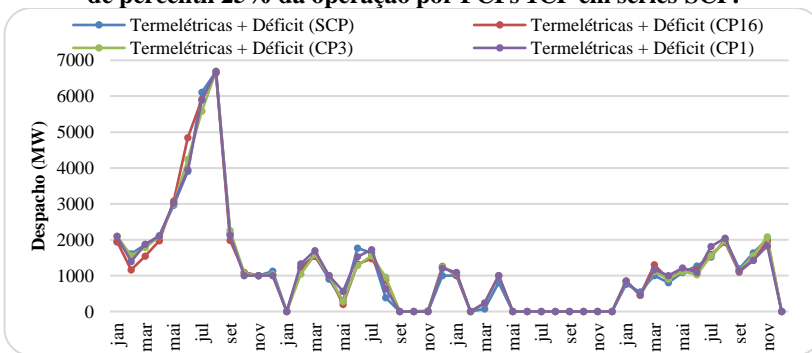
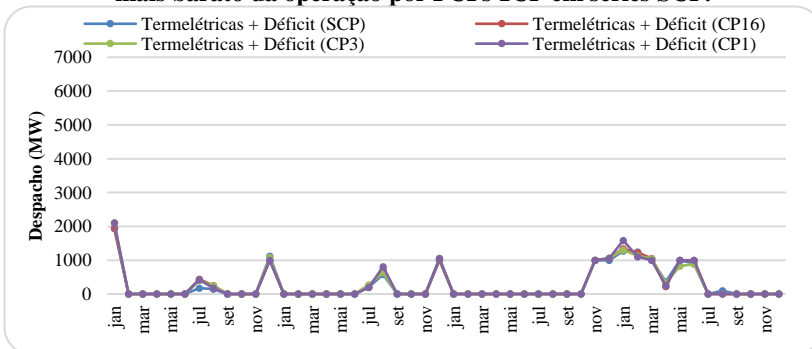


Figura 9.5 - Simulações de operação (FCFs de todos os modelos) no cenário mais barato da operação por FCFs 1CP em séries SCP.



10. APÊNDICE C: Despacho termelétrico para a série CP16

Figura 10.1 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 100% (mais caro) da operação por FCFs 1CP em séries de 16 CP.

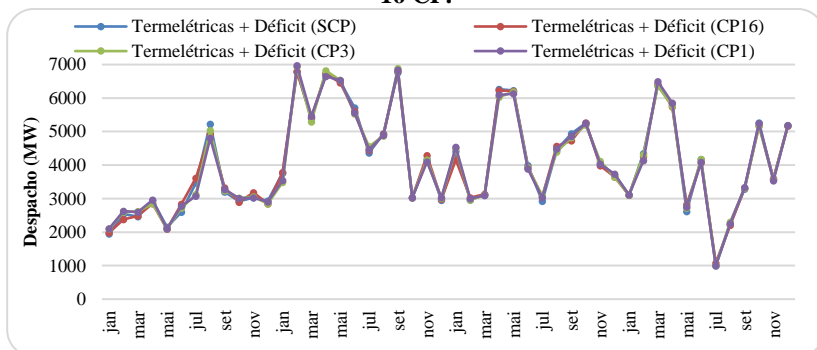


Figura 10.2 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 75% da operação por FCFs 1CP em séries de 16 CP.

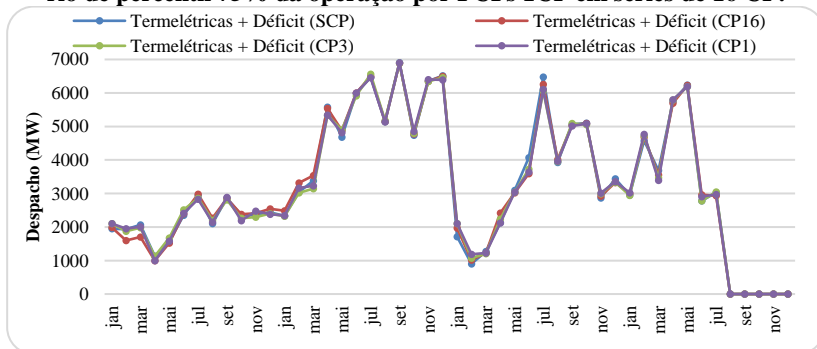


Figura 10.3 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 50% da operação por FCFs 1CP em séries de 16 CP.

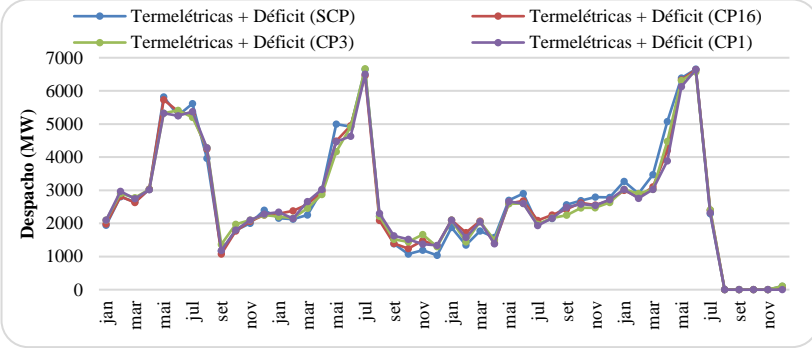


Figura 10.4 - Simulações de operação (FCFs de todos os modelos) no cenário de percentil 25% da operação por FCFs 1CP em séries de 16 CP.

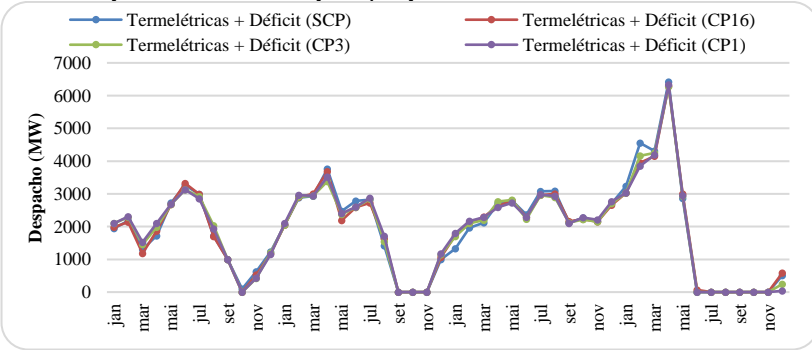
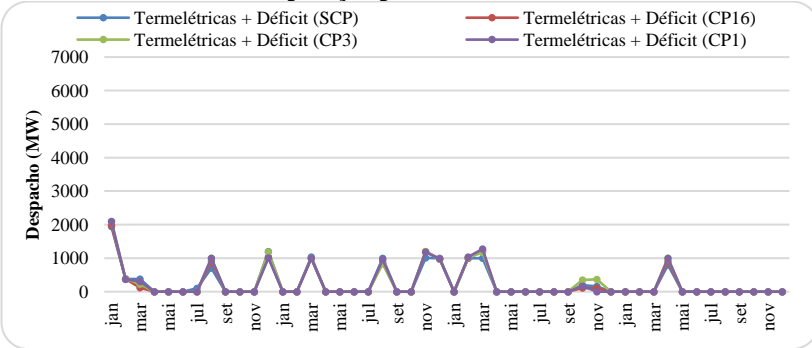


Figura 10.5 - Simulações de operação (FCFs de todos os modelos) no cenário mais barato da operação por FCFs 1CP em séries de 16 CP.



11. APÊNDICE D: Número de iterações por tamanho de árvore

Figura 11.1 - Número de Iterações (20 realizações por mês).

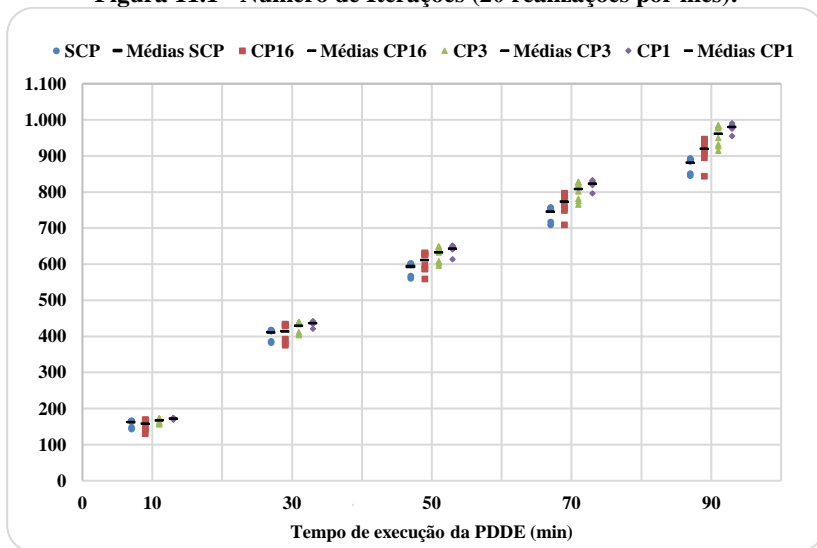


Figura 11.2 - Número de Iterações (50 realizações por mês).

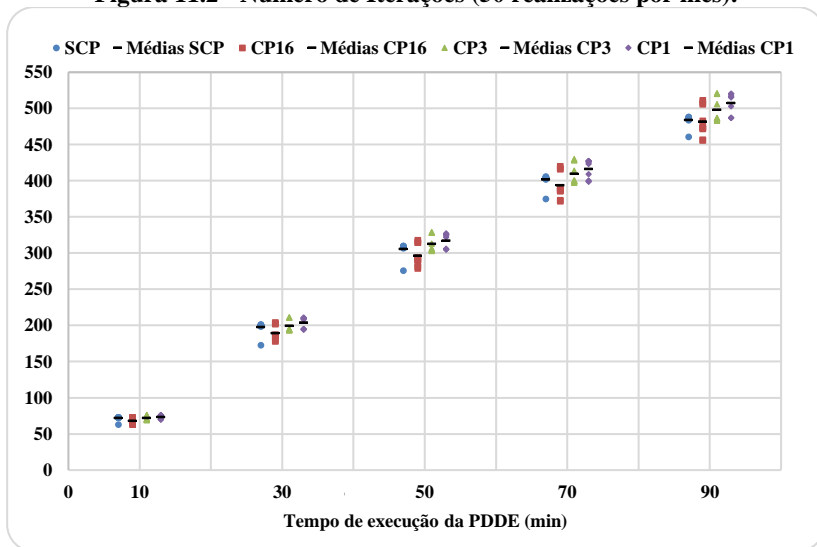
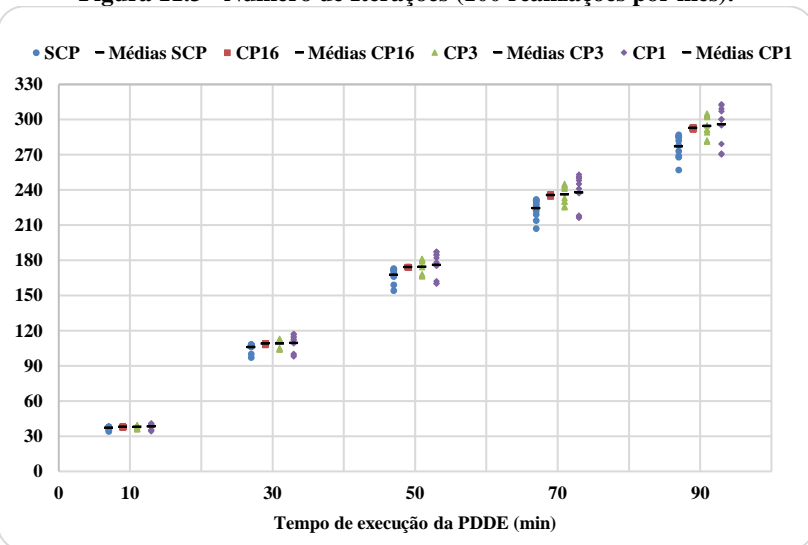


Figura 11.3 - Número de Iterações (100 realizações por mês).



REFERÊNCIAS

ABRAMOWITZ, M.; STEGUN, I. A.; MILLER, D. **Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables**. 10. ed. [s.l.] Courier Corporation, 1964.

BELLMAN, R. Adaptive Control Processes: A Guided Tour. In: PRINCETON UNIVERSITY PRESS (Ed.). [s.l.: s.n.]. p. 94.

BEZERRA, B. V. **Incorporação da Incerteza dos Parâmetros do Modelo Estocástico de Vazões na Política Operativa do Despacho Hidrotérmico**. [s.l.] Tese de Doutorado - Pontifícia Universidade Católica do Rio de Janeiro, 2005.

CASEY, M. S.; SEN, S. The Scenario Generation Algorithm for Multistage Stochastic Linear Programming. **Mathematics of Operations Research**, v. 30, n. 3, p. 615–631, 2005.

CATTELL, R. B. The Scree Test For The Number Of Factors. **Multivariate Behavioral Research**, p. 245–276, 1966.

CHARBENEAU, R. Comparison of the Two- and Three-Parameter Log Normal Distributions Used in Streamflow Synthesis. **Water Resources Research**, v. 14, n. Fevereiro, p. 149–150, 1978.

CHEN, D. et al. Dimension reduction of decision variables for multireservoir operation: A spectral optimization model. **Water Resources Research**, v. 52, n. Janeiro, p. 36–51, 2016.

DA COSTA, J. P.; DE OLIVEIRA, G. C.; LEGEY, L. F. L. **Reduced Scenario Tree Generation for Mid-Term Hydrothermal Operation Planning**. 9th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS. **Anais...Estocolmo, Suécia: 2006**

DE MATOS, V. L. **Um Modelo para o Planejamento Anual da Operação Energética Considerando Técnicas Avançadas de Otimização Estocástica**. [s.l.] Tese de Doutorado - Universidade Federal de Santa Catarina, 2012.

DE MATOS, V. L.; FINARDI, E. C.; LARROYD, P. V. **Uma Medida Coerente de Risco como Métrica para o Planejamento Anual da Operação Energética**. XII Simpósio de Especialistas em Planejamento da Operação e Expansão Elétrica. **Anais...Rio de Janeiro /RJ: 2012**

DE OLIVEIRA, W. L. et al. **Optimal Scenario Tree Reduction for Stochastic Streamflows in Power Generation Planning Problems**. EngOpt 2008 - International Conference on Engineering Optimization. **Anais...**Rio de Janeiro, Brasil: 2008

DE SOUZA SILVA, T. **APLICAÇÃO DO MÉTODO AMOSTRAGEM SELETIVA GLOBAL NA CONSTRUÇÃO DA ÁRVORE DE CENÁRIOS DE AFLUÊNCIAS UTILIZADA NO PLANEJAMENTO DA OPERAÇÃO ENERGÉTICA DE MÉDIO PRAZO**. [s.l.] Dissertação de Mestrado - Universidade Federal do Rio de Janeiro, 2010.

DINIZ, A. L.; MACEIRA, M. E. P. A Four-Dimensional Model of Hydro Generation for the Short-Term Hydrothermal Dispatch Problem Considering Head and Spillage Effects. **IEEE Transactions on Power Systems**, v. 23, n. 3, p. 1298–1308, 2008.

DONOHUE, C. J.; BIRGE, J. R. The Abridged Nested Decomposition Method for Multistage Stochastic Linear Programs with Relatively Complete Recourse. **Algorithmic Operations Research**, v. 1, p. 20–30, 2006.

DUPAČOVÁ, J.; KOZMÍK, V. SDDP for multistage stochastic programs: preprocessing via scenario reduction. **Computational Management Science**, 2016.

FINARDI, E. C. **Planejamento da Operação de Sistemas Hidrotérmicos Utilizando Computação de Alto Desempenho**. [s.l.] Dissertação de Mestrado - Universidade Federal de Santa Catarina, 1999.

FREDO, G. L. M. **Análise de Diferentes Representações da Função de Produção Hidroelétrica no Problema de Planejamento da Operação Energética de Médio Prazo**. [s.l.] Dissertação de Mestrado - Universidade Federal de Santa Catarina, 2016.

FU, G.; KAPELAN, Z.; REED, P. Reducing the Complexity of Multiobjective Water Distribution System Optimization through Global Sensitivity Analysis. **Journal of Water Resources Planning and Management**, v. 138, p. 196–207, 2012.

GARCIA, A. G. N. **Representação do Processo Estocástico de Energias Afluentes por Modelos Auto-Regressivos Periódicos no Planejamento de Sistemas Hidrotérmicos**. [s.l.] Dissertação de Mestrado - Universidade Federal de Santa Catarina, 2005.

GORENSTIN, B. G. **Otimização Estocástica de Sistemas Hidrotérmicos Incluindo Rede de Transmissão**. [s.l.] Dissertação de Mestrado - Universidade Federal do Rio de Janeiro, 1991.

HEITSCH, H.; RÖMISCH, W. **Scenario tree reduction for multistage stochastic programs** Berlin, Alemanha Humboldt-University Berlin, Institute of Mathematics, , 2014.

HINDSBERGER, M. ReSa: A method for solving multistage stochastic linear programs. **Journal of Applied Operational Research**, v. 6, n. 1, p. 2–15, 2014.

HIPEL, K. W.; MCLEOD, A. I. Time series modelling of water resources and environmental systems. In: ELSEVIER (Ed.). . 1st. ed. Amsterdam: [s.n.]. p. 1013.

HOCHREITER, R.; PFLUG, G. C. Financial scenario generation for stochastic multi-stage decision processes as facility location problems. **Annals of Operations Research**, v. 152, p. 257–272, 2007.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. **Journal of Educational Psychology**, v. 24, n. 6, p. 417–441, 1933.

HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. **International Journal of Forecasting**, v. 22, p. 679–688, 2006.

JOHNSON, S. A. et al. Numerical Solution of Continuous-State Dynamic Programs Using Linear and Spline Interpolation. **Operations Research**, v. 41, n. 3, p. 484–500, 1993.

KAISER, H. F. The Application of Electronic Computers to Factor Analysis. **Educational and Psychological Measurement**, v. XX, n. 1, p. 141–151, 1960.

KUHN, D. Aggregation and discretization in multistage stochastic programming. **Mathematical Programming**, v. 113, p. 61–94, 2008.

LARROYD, P. V. **Um Modelo de Otimização Estocástica com Representação Individualizada das Usinas Hidrelétricas no Planejamento de Médio Prazo da Operação Hidrotérmica**. [s.l.] Tese de Doutorado - Universidade Federal de Santa Catarina, 2016.

LARSEN, C. T.; DOORMAN, G. L.; MO, B. **Evaluation of scenario reduction methods for stochastic inflow in hydro scheduling models.** IEEE Eindhoven PowerTech, PowerTech 2015. **Anais...**2015

MACEIRA, M. E. P. et al. Geração de Cenários Sintéticos de Energia e Vazão para o Planejamento da Operação Energética. **CADERNOS DO IME – Série Estatística**, v. 21, p. 11–35, 2006.

MARCATO, A. L. M. **Representação Híbrida de Sistemas Equivalentes e Individualizados para o Planejamento da Operação de Médio Prazo de Sistemas de Potência de Grande Porte.** [s.l.] Pontifícia Universidade Católica do Rio de Janeiro, 2002.

PEREIRA, M. V. F.; PINTO, L. M. V. G. A Decomposition Approach to the Economic Dispatch of Hydrothermal Systems. **IEEE Transactions on Power Apparatus and Systems**, v. PAS-101, n. 10, p. 3851–3860, 1982.

PEREIRA, M. V. F.; PINTO, L. M. V. G. Stochastic Optimization of a Multireservoir Hydroelectric System: A Decomposition Approach. **Water Resources Research**, v. 21, n. 6, p. 779–792, 1985.

PEREIRA, M. V. F.; PINTO, L. M. V. G. Multi-stage stochastic optimization applied to energy planning. **Mathematical Programming**, v. 52, p. 359–375, 1991.

ROGERS, D. F. et al. Aggregation and Disaggregation Techniques and Methodology in Optimization. **Operations Research**, v. 39, n. 4, p. 553–582, 1991.

ROSENTHAL, R. E. A Nonlinear Network Flow Algorithm for Maximization of Benefits in a Hydroelectric Power System (1981) A Nonlinear Network Flow Algorithm for Maximization of Benefits in a Hydroelectric Power System. **Operations Research**, v. 29, n. 4, p. 763–786, 1981.

SAAD, M. et al. Learning disaggregation technique for the operation of long-term hydroelectric power systems. **Water Resources Research**, v. 30, n. 11, p. 3195–3202, 1994.

SAAD, M.; TURGEON, A. Application of Principal Component Analysis to Long-Term Reservoir Management. **Water Resources Research**, v. 24, n. Julho, p. 907–912, 1988.

SCARCELLI, R. DE O. C. **Programação Dinâmica Aplicada à Otimização Individualizada e Desacoplada das Usinas Hidrelétricas de Sistemas Hidrotérmicos.** [s.l.] Dissertação de Mestrado - Universidade de São Paulo, 2012.

SHAPIRO, A.; RUSZCZYNSKI, A. **Lectures On Stochastic Programming Modeling And Theory.** Philadelphia, PA - USA: Society for Industrial and Applied Mathematics and the Mathematical Programming Society, 2009.

SOARES, M. P. **On the Solution Variability Reduction of Stochastic dual Dynamic Programming Applied to Energy Planning.** [s.l.] Tese de Doutorado - Pontifícia Universidade Católica do Rio de Janeiro, 2015.

SUANNO, C. M. **Operação Estocástica de Sistemas Hidrotérmicos com Representação da Produtibilidade Variável.** [s.l.] Dissertação de Mestrado - Universidade Federal do Rio de Janeiro, 1995.

THOMAS, H. A.; FIERING, M. B. Mathematical Synthesis of Streamflow Sequences for the Analysis of River Basins by Simulation. In: MASS ET AL., H. U. P. (Ed.). **Design of Water Resource Systems.** [s.l: s.n.]. p. 460–493.

TURGEON, A. A Decomposition Method for the Long-Term Scheduling of Reservoirs in Series. **Water Resources Research**, v. 17, n. 6, p. 1565–1570, 1981.

TURGEON, A.; CHARBONNEAU, R. An aggregation-disaggregation approach to long-term reservoir management. **WATER RESOURCES RESEARCH**, v. 34, n. 12, p. 3585–3594, 1998.

VALDES, J. B. et al. Aggregation-Disaggregation Approach to Multireservoir Operation. **Journal of Water Resources Planning and Management**, v. 118, p. 423–444, 1992.