



REPRESENTAÇÃO DE RESTRIÇÕES DE AVERSÃO A RISCO DE CVAR EM
PROGRAMAÇÃO DINÂMICA DUAL ESTOCÁSTICA COM APLICAÇÃO AO
PLANEJAMENTO DA OPERAÇÃO DE SISTEMAS HIDROTÉRMICOS

Luiz Carlos da Costa Júnior

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientador: Márcia Helena Costa Fampa

Rio de Janeiro
Novembro de 2013

REPRESENTAÇÃO DE RESTRIÇÕES DE AVERSÃO A RISCO DE CVAR EM
PROGRAMAÇÃO DINÂMICA DUAL ESTOCÁSTICA COM APLICAÇÃO AO
PLANEJAMENTO DA OPERAÇÃO DE SISTEMAS HIDROTÉRMICOS

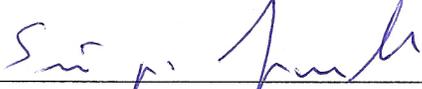
Luiz Carlos da Costa Júnior

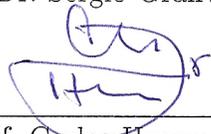
TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:


Prof. Márcia Helena Costa Fampa, D.Sc.


Dr. Mario Veiga Ferraz Pereira, D.Sc.


Dr. Sérgio Granville, D.Sc.


Prof. Carlos Humes Junior, D.Sc.


Prof. Nelson Maculan Filho, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
NOVEMBRO DE 2013

Costa Júnior, Luiz Carlos da

Representação de restrições de aversão a risco de CVaR em Programação Dinâmica Dual Estocástica com aplicação ao planejamento da operação de sistemas hidrotérmicos/Luiz Carlos da Costa Júnior. – Rio de Janeiro: UFRJ/COPPE, 2013.

XV, 144 p.: il.; 29, 7cm.

Orientador: Márcia Helena Costa Fampa

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2013.

Referências Bibliográficas: p. 137 – 144.

1. Aversão ao Risco. 2. Programação Dinâmica Dual Estocástica. 3. Relaxação Lagrangeana. I. Fampa, Márcia Helena Costa. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

*Àqueles que dão sentido à tudo:
meus pais, Luiz e Sônia,
meus irmãos, Ana e Cadu,
e minha esposa, Fernanda.*

Agradecimentos

Em primeiro lugar, agradeço aos meus pais, Luiz e Sônia, e aos meus irmãos, Ana Carolina e Carlos Eduardo, por terem sempre me apoiado e incentivado em todas as minhas escolhas. Seu amor, compreensão e exemplos são essenciais em todas as etapas de minha vida.

À Fernanda Thomé da Costa, amor da minha vida, por todo carinho, paciência e incentivo constantes, indispensáveis para obter o estímulo e motivação necessários para seguir adiante sempre.

À minha querida família: às minhas avós, tios, sogros, primos, cunhados e às minhas duas lindas afilhadas. Por todo o carinho e apoio demonstrado.

Ao meu orientador e amigo Mario Veiga F. Pereira, a minha admiração e gratidão especial por toda a motivação, confiança e orientação acadêmica impecável, fundamentais para a conclusão deste trabalho. Sua influência ao longo de diversos anos certamente tem um papel distinto em minha formação profissional.

Ao amigo Sérgio Granville, pela co-orientação, por toda dedicação, contribuições e constante boa vontade nas diversas discussões ao longo do trabalho. À amiga Nora Campodónico, pelos incentivos frequentes e pela valiosa ajuda dispensada sempre que necessário. Aos amigos André Dias, Bernardo Bezerra, Fernanda Thomé, Júlio Dias, Luiz Augusto Barroso, Luiz Maurício Thomé, Marcelo Cruz, Martha Carvalho, Pedro Avila, Rafael Kelman, Raphael Chabar e Silvio Binato por suas contribuições. Aos demais amigos da PSR pela colaboração e pela agradável convivência.

À minha orientadora professora Márcia Fampa pela confiança e importantes contribuições desde o curso mestrado, que certamente se refletem na minha formação acadêmica. Aos professores e funcionários do PESC por toda dedicação e ensinamentos, em especial aos queridos professor Nelson Maculan, Cláudia, Fátima, Solange e Sônia pela simpatia e incansável prontidão para ajudar sempre que foi necessário. À CAPES pelo auxílio a mim concedido na forma de bolsa de estudos.

Aos meus queridos amigos Bruno Gomes, Bruno Romanelli, Cecília Cabral, Diogo Gonçalves, Evelinne Varela, Felipe Gullo, Gabriela Saramago, Gustavo Almeida, Gustavo Audi, Gustavo Wo, Ingrid Delpy, João Luis Lobo, Juliana Noronha, Letícia Fonti, Marcos Vinícius Spiguel e Rafael Diniz pelo carinho e companheirismo, fundamentais para obter o ânimo necessário e alcançar mais esta etapa.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

REPRESENTAÇÃO DE RESTRIÇÕES DE AVERSÃO A RISCO DE CVAR EM
PROGRAMAÇÃO DINÂMICA DUAL ESTOCÁSTICA COM APLICAÇÃO AO
PLANEJAMENTO DA OPERAÇÃO DE SISTEMAS HIDROTÉRMICOS

Luiz Carlos da Costa Júnior

Novembro/2013

Orientador: Márcia Helena Costa Fampa

Programa: Engenharia de Sistemas e Computação

Esta tese descreve uma metodologia para incorporação de restrições de aversão a risco no problema de planejamento ótimo da operação hidrotérmica formulado como um modelo de programação linear estocástica multiestágio e solucionado através de uma extensão do algoritmo de Programação Dinâmica Dual Estocástica (PDDE). A minimização dos custos esperados no problema clássico de programação hidrotérmica multiestágio envolvendo o controle de risco é considerada uma tarefa desafiadora: a política operativa neutra ao risco não consegue captar eventos que têm baixa probabilidade de ocorrência, mas podem ter impactos importantes. O objetivo deste trabalho é propor a representação de aversão ao risco no problema de programação hidrotérmica estocástico através da consideração explícita de restrições que limitam o CVaR associado ao déficit de energia no modelo de otimização. É utilizada a técnica de relaxação Lagrangeana para permitir a decomposição do problema e é apresentada uma extensão do algoritmo para incorporar tais restrições de CVaR no ambiente da PDDE. Na proposta deste trabalho, estamos interessados em controlar os déficits de energia a um nível de risco pré-estabelecido, visando assim a conciliar os objetivos do planejamento da expansão e da operações (assegurando a minimização dos custos e o atendimento do critério de confiabilidade do fornecimento de energia). Além disto, mostra-se que essas restrições podem ser representadas como penalidades lineares por partes sobre os déficits de energia e, em seguida, desenvolve-se um algoritmo para calcular o coeficiente e a profundidade deste segmento. A metodologia será ilustrada com problemas de programação hidrotérmica estocástica realistas.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

REPRESENTING CVAR RISK AVERSION CONSTRAINTS ON STOCHASTIC
DUAL DYNAMIC PROGRAMMING WITH APPLICATION IN THE
HYDROTHERMAL OPERATION PLANNING

Luiz Carlos da Costa Júnior

November/2013

Advisor: Márcia Helena Costa Fampa

Department: Systems Engineering and Computer Science

This thesis describes a methodology to incorporate risk aversion constraints in the optimal hydrothermal scheduling problem, formulated as a multistage stochastic linear programming problem and solved through an extension of the Stochastic Dual Dynamic Programming algorithm (SDDP). The minimization of expected costs in the classic multistage scheduling problem involving risk aversion is considered a challenging task: the risk-neutral operation policy fails to capture events that have low probability of occurrence but might have important impacts. The objective of this work is to propose the representation of risk aversion in stochastic hydrothermal scheduling problem through the explicit consideration of constraints that limit the CVaR associated to energy deficit in the optimization model. The Lagrangian relaxation technique is applied in order to allow model decomposition and an algorithm extension is presented in order to incorporate such CVaR constraints in the SDDP framework. In our model we are interested in controlling the energy deficits at a predetermined risk level, thus aiming the reconciliation of the planning and operations objectives (ensuring cost minimization and supply reliability at a given risk level). Moreover, it is shown that these constraints can be represented as piecewise linear penalties on energy deficits and we develop an algorithm to calculate the coefficient and depth of these segments. The methodology will be illustrated with real stochastic hydrothermal scheduling problems.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xv
1 Introdução	1
1.1 Despacho de sistemas hidrotérmicos	1
1.1.1 Operação de sistemas térmicos	1
1.1.2 Operação de sistemas hidrotérmicos	3
1.2 Incertezas no problema do planejamento da operação	6
1.3 Segurança energética	10
1.4 Motivação	11
1.5 Objetivo e contribuições	17
1.6 Organização da tese	20
2 Programação estocástica multiestágio	22
2.1 Introdução	22
2.2 Formulação matemática	23
2.2.1 Formulação geral	23
2.2.2 Formulação linear	25
2.3 Representação das incertezas	27
2.3.1 Distribuição de probabilidade	27
2.3.2 Árvore de cenários	27
2.3.3 Formulação em árvore	31
2.4 Programação Dinâmica Dual Estocástica	31
2.4.1 Dependência temporal	32
2.4.2 Amostragem por Monte Carlo e <i>Sample Average Approximation</i>	35
2.4.3 Formulação matemática	36
2.4.4 Algoritmo	37
3 Planejamento da operação de sistemas hidrotérmicos	41
3.1 Usinas termelétricas	41
3.1.1 Custo de produção	42

3.1.2	Limites operativo	43
3.2	Usinas hidrelétricas	43
3.2.1	Balço hídrico	44
3.2.2	Limites no armazenamento e na defluência	46
3.2.3	Produção de energia	46
3.3	Demanda	46
3.3.1	Atendimento à demanda	46
3.3.2	Déficit de energia	47
3.4	Formulação do problema de uma etapa	47
4	Representação da aversão ao risco	49
4.1	Introdução	49
4.2	Risco	50
4.3	Valor esperado	53
4.4	Value-at-Risk (VaR)	54
4.4.1	Definição	54
4.4.2	Relação entre o risco de déficit e VaR	56
4.5	Coerência das medidas de risco	56
4.6	Conditional Value-at-Risk (CVaR)	58
4.6.1	Definição	58
4.6.2	Relação entre o valor esperado do corte de carga e o CVaR	60
4.6.3	Relação entre o VaR e o CVaR	61
5	Planejamento da operação com restrições de CVaR com formulação em árvore	64
5.1	Introdução	64
5.2	Formulação geral	64
5.3	Relaxação Lagrangeana	67
5.4	Algoritmo de solução	68
6	Planejamento da operação com restrições de CVaR em um estágio com PDDE	70
6.1	Introdução	70
6.2	Programação Dinâmica	70
6.2.1	Estágios posteriores ao estágio da restrição de CVaR	72
6.2.2	Estágio da restrição de CVaR	73
6.2.3	Estágios anteriores ao estágio da restrição de CVaR	74
6.2.4	Estágio de decisão estratégica	75
6.3	Programação Dinâmica Dual Estocástica	75
6.3.1	Estágios posteriores ao estágio da restrição de CVaR	76

6.3.2	Estágio da restrição de CVaR	78
6.3.3	Estágios anteriores à restrição de CVaR	79
6.3.4	Estágio de decisão estratégica	80
6.4	Algoritmo de solução	80
6.5	Interpretação econômica da restrição de CVaR de um estágio	81
6.5.1	Análise direta	81
6.5.2	Reformulação do problema	82
6.6	Resultados numéricos	84
6.6.1	Descrição do sistema	84
6.6.2	Caso BASE	85
6.6.3	Caso com restrição de CVaR de um estágio	87

7 Planejamento da operação com restrições de CVaR multiestágio com PDDE 93

7.1	Introdução	93
7.2	Formulação matemática para o déficit acumulado	94
7.2.1	Equivalente determinístico	94
7.2.2	Programação dinâmica	94
7.3	Planejamento da operação com restrições de CVaR no déficit acumulado - formulação direta	95
7.3.1	Estágio de decisão estratégica	96
7.3.2	Estágios anteriores ao intervalo de déficit acumulado	96
7.3.3	Estágio inicial do intervalo do déficit acumulado	97
7.3.4	Estágio intermediário do intervalo do déficit acumulado	97
7.3.5	Estágio final do intervalo do déficit acumulado	98
7.3.6	Estágios posteriores ao intervalo de déficit acumulado	99
7.4	Planejamento da operação com restrições de CVaR no déficit acumulado - formulação final	99
7.4.1	Estágio final do intervalo do déficit acumulado	101
7.4.2	Estágio intermediário do intervalo do déficit acumulado	104
7.4.3	Estágio inicial do intervalo do déficit acumulado	105
7.5	Algoritmo de solução	105
7.6	Resultados numéricos	105
7.6.1	Caso base	106
7.6.2	Caso com restrição de CVaR em um estágio	107
7.6.3	Caso com restrição de CVaR multiestágio	107
7.7	Análise comparativa entre as metodologias	111

8	Estudo de caso	115
8.1	Implementação computacional	115
8.2	Descrição do sistema	115
8.3	Caso Base	116
8.4	CVaR associado ao déficit acumulado anual	119
9	Considerações finais e desenvolvimentos futuros	126
9.1	Conclusões	126
9.2	Desenvolvimentos futuros	132
9.2.1	Planejamento da expansão	132
9.2.2	Ajuste da Função de Custo Futuro	134
9.2.3	Aplicação da metodologia proposta em problemas relaciona- dos de outras áreas	136
	Referências Bibliográficas	137

Lista de Figuras

1.1	Processo de decisão operativa com incerteza nas vazões	4
1.2	Função de custo imediato e futuro em função do armazenamento . . .	5
1.3	Uso ótimo da água	5
1.4	Representação da incerteza hidrológica para um reservatório	6
2.1	Árvore de cenários: exemplo de 4 níveis (estágios) com 3 aberturas em cada nível, totalizando 27 cenários.	28
2.2	Árvore de cenários com independência temporal: exemplo de 4 níveis (estágios) com 3 aberturas em cada nível.	33
2.3	Amostragem de cenários na fase <i>forward</i> do algoritmo de PDDE . . .	38
2.4	Problemas solucionados para o cálculo dos cortes da FCF do penúltimo estágio na fase <i>backward</i> do algoritmo de PDDE	38
3.1	Esquema simplificado de geração de energia de uma usina térmica (a vapor)	42
3.2	Perfil típico de uma usina hidrelétrica	44
3.3	Balanco hídrico de usinas em cascata	45
4.1	Distribuição de probabilidade	51
4.2	Risco de déficit de energia	52
4.3	Função indicadora	52
4.4	Valor esperado da corte de carga	54
4.5	Função truncamento	54
4.6	VaR_α : <i>Value-at-Risk</i> associado ao corte de carga	55
4.7	Função indicadora trasladada	56
4.8	$CVaR_\alpha$: <i>Conditional Value-at-Risk</i> associado ao corte de carga	59
4.9	Função truncamento trasladada	60
4.10	Aproximação convexa da função indicadora	62
5.1	Árvore de cenários: exemplo de formulação com restrição de $CVaR$ associada ao estágio 4.	66

6.1	Diagrama do esquema de decomposição para o problema dualizado com restrição de CVaR no estágio τ	72
6.2	Curva de custo de déficit implícito	82
6.3	Sistema da Costa Rica (caso base): histograma do déficit de energia para a etapa 3	86
6.4	Sistema da Costa Rica - Caso CVaR(3): penalidade por iteração . . .	88
6.5	Sistema da Costa Rica: VaR _{8%} e CVaR _{8%} em função da penalidade .	88
6.6	Sistema da Costa Rica - gráfico da curva de custo de déficit de 2 segmentos	90
6.7	Sistema da Costa Rica - Caso CVaR(3): custo total em função da penalidade (Lagrangeano)	90
6.8	Sistema da Costa Rica - Caso CVaR(3): histograma do déficit de energia para a etapa 2	92
6.9	Sistema da Costa Rica - Caso CVaR(3): histograma do déficit de energia para a etapa 3	92
7.1	Diagrama do esquema de decomposição para o problema dualizado com restrição de CVaR para o déficit acumulado no intervalo \mathcal{T} . . .	96
7.2	Sistema da Costa Rica - Caso CVaR(1-3): penalidade por iteração . .	108
7.3	Sistema da Costa Rica - Caso CVaR(1-3): VaR _{8%} e CVaR _{8%} do déficit acumulado em função da penalidade	108
7.4	Sistema da Costa Rica - Caso CVaR(1-3): gráfico da curva de custo de déficit acumulado de 2 segmentos	110
7.5	Sistema da Costa Rica - Caso CVaR(1-3): custo total em função da penalidade (Lagrangeano)	110
7.6	Sistema da Costa Rica: comparação das medidas de risco associadas ao déficit da etapa 2	112
7.7	Sistema da Costa Rica: comparação das medidas de risco associadas ao déficit da etapa 3	113
7.8	Sistema da Costa Rica: comparação das medidas de risco associadas ao déficit acumulado	113
7.9	Sistema da Costa Rica: comparação do CVaR _{8%} associado ao déficit de cada etapa e ao déficit acumulado	114
8.1	Sistema Interligado Nacional - diagrama esquemático do sistema de transmissão (<i>Fonte: ONS</i>)	117
8.2	SIN - Penalidade por iteração	121
8.3	SIN - CVaR _{5%} associado ao déficit acumulado em função da penalidade	122
8.4	SIN - Custo total em função da penalidade (função Lagrangeana) . .	122
8.5	SIN - Evolução da energia armazenada para os casos BASE de CVaR	123

8.6	SIN - Comparação das funções densidade (aproximada) associadas ao déficit acumulado para os casos BASE de CVaR (ano 2014)	124
8.7	SIN - curva de custo de déficit acumulado de 2 segmentos	125
9.1	Planejamento da expansão: esquema de decomposição tradicional . .	133

Lista de Tabelas

1.1	Maldição da dimensionalidade	9
6.1	Sistema da Costa Rica - capacidade instalada e geração prevista (2013)	84
6.2	Sistema da Costa Rica - Caso BASE: geração média (MWmed)	85
6.3	Sistema da Costa Rica - Caso BASE: medidas de risco	87
6.4	Sistema da Costa Rica - Caso CVaR(3): tabela de convergência do método da bisseção	89
6.5	Sistema da Costa Rica - Caso CVaR(3): curva de custo de déficit de 2 segmentos	89
6.6	Sistema da Costa Rica - Caso CVaR(3): geração média (MWmed) . .	91
6.7	Sistema da Costa Rica - Caso CVaR(3): medidas de risco	91
7.1	Sistema da Costa Rica - Caso BASE: medidas de risco	106
7.2	Sistema da Costa Rica - Caso CVaR(3): medidas de risco	107
7.3	Sistema da Costa Rica - Caso CVaR(1-3): tabela de convergência do método da bisseção	109
7.4	Sistema da Costa Rica - Caso CVaR(1-3): curva de custo de déficit acumulado de 2 segmentos	109
7.5	Sistema da Costa Rica - Caso CVaR(1-3): geração média (MWmed) .	111
7.6	Sistema da Costa Rica - Caso CVaR(1-3): medidas de risco	111
8.1	SIN - Caso BASE: risco de déficit anual	118
8.2	SIN - Caso BASE: medidas de risco para o déficit acumulado anual .	119
8.3	SIN - Caso CVaR: tabela de convergência do método da bisseção . . .	120
8.4	SIN - Caso CVaR: medidas de risco para o déficit acumulado anual .	121
8.5	SIN - Caso CVaR: curva de custo de déficit acumulado de 2 segmentos	124

Capítulo 1

Introdução

1.1 Despacho de sistemas hidrotérmicos

1.1.1 Operação de sistemas térmicos

Em sistemas predominantemente térmicos, a cada intervalo de tempo, por exemplo uma hora ou um dia, cada produtor (gerador) oferece ao mercado o preço unitário pelo qual aceita vender a energia produzida e a quantidade que é capaz de produzir. Em seguida, aplica-se uma regra de despacho para determinar qual é a melhor maneira, sob o ponto de vista dos consumidores, de contratar a geração de energia de forma a assegurar o atendimento da demanda ao menor custo possível. De forma simplificada, este procedimento consiste em ordenar os produtores de acordo com os preços unitários ofertados e contratá-los, segundo esta ordem (também conhecida como ordem de mérito), até que a demanda seja atendida em sua totalidade. Para garantir a utilização dos recursos mais econômicos, todos os geradores despachados são contratados em sua capacidade máxima exceto o último gerador, que fica responsável por “completar” o atendimento à demanda.

Este último gerador recebe uma denominação especial e é chamado de gerador marginal. Em sistemas onde se aplica a teoria marginalista, cada produtor despachado é remunerado não pelo preço unitário que havia ofertado inicialmente e sim pelo “preço unitário marginal”, que corresponde à oferta do gerador marginal. Isto é, a receita bruta dos produtores despachados é dada pelo produto do preço marginal e sua quantidade de energia gerada enquanto que a receita líquida é dada pelo produto da diferença de preço entre o preço marginal e o preço ofertado pela energia gerada. Naturalmente, a receita bruta dos geradores não despachados será nula. A hipótese básica deste arranjo é que cada produtor decide sobre o preço e quantidade ofertados sem o conhecimento da decisão dos demais produtores.

Pode-se provar que, sob as condições do chamado “mercado perfeito”, a estratégia ótima de cada produtor é ofertar energia a um custo exatamente igual ao seu custo

unitário de produção. Intuitivamente, é simples perceber a racionalidade desta decisão pois caso o mesmo ofertasse a um preço inferior correria o risco de prejuízo caso fosse despachado. Por outro lado, caso ofertasse a um preço maior que seu custo de produção, correria o risco de deixar de lucrar caso a ordem de mérito fosse alterada e não fosse despachado.

Este mecanismo é bastante atraente do ponto de vista do consumidor pois seria capaz de eleger os recursos de menor custo para o atendimento à demanda e, ao mesmo tempo, assegurar a sustentabilidade do processo produtivo mesmo em caso de crescimento da demanda. O funcionamento deste mecanismo é previsto sob a hipótese de mercado perfeito porém, é sabido que em mercados imperfeitos onde alguns produtores tem capacidade de manipular os preços pode haver a tendência de que estes ofertem preços unitários maiores que seus custos de produção. Para evitar este comportamento, diversos países desenvolveram mecanismos de correção e procedimentos regulatórios com o objetivo de assegurar a concorrência entre os agentes.

Em suma, o problema do planejamento da operação de um sistema puramente térmico é resolvido colocando as usinas em ordem crescente de custo de produção de uma unidade de energia adicional (custo incremental) e ajustando a ordem de despacho das usinas com relação às variações da demanda. Existem diversos fatores adicionais que tornam a solução deste problema mais complexo como por exemplo limites de transmissão de energia pela rede, perdas de energia, restrições operativas dos geradores térmicos como limites nas taxas de variação de produção de energia, custos de partida para produção, entre outros. Entretanto, é possível estabelecer as seguintes características básicas do problema de operação térmica:

- é desacoplado no tempo, ou seja, a decisão operativa hoje não afeta o custo operativo da próxima etapa;
- as unidades geradoras tem custo direto de produção, isto é, os custos operativos dependem apenas do seu nível de geração e não da produção das outras usinas;
- o despacho econômico não depende do custo econômico de racionamento para sociedade (supondo que este custo exceda o da térmica mais cara). Como consequência, o risco de interrupção de suprimento de um sistema térmico depende somente da capacidade total das térmicas disponíveis e não da política operativa.

Estas características influenciaram os mecanismos propostos de desestatização e descentralização operativa nos países industrializados, quase todos à base térmica.

Entretanto, como será visto na seção a seguir, características como predominância hidrelétrica e a presença de grandes reservatórios, como é o caso do Brasil e

diversos outros países, fazem com que a operação de sistemas hidrotérmicos tenha características bastante diferentes.

1.1.2 Operação de sistemas hidrotérmicos

A característica mais evidente de um sistema composto por usinas hidrelétricas e termoelétricas é poder utilizar a energia “gratuita” que está armazenada nos reservatórios para atender à demanda, evitando, dessa maneira, gastos de combustível com as unidades térmicas. Entretanto, a disponibilidade de energia hidrelétrica está limitada pela capacidade de armazenamento nos reservatórios. Isto introduz uma dependência entre a decisão operativa de hoje e os custos operativos no futuro.

Em outras palavras, se deplecionarmos hoje as reservas de energia hidrelétrica com o objetivo de minimizar os custos térmicos e houver uma seca severa no futuro, poderá haver um racionamento de custo elevado para sociedade. Se, por outro lado, preservamos as reservas de energia hidrelétrica por meio de um uso mais intenso de geração térmica e as afluições futuras forem elevadas, poderá ocorrer um vertimento nos reservatórios do sistema, o que representa desperdício de energia e, conseqüentemente, um aumento nos custos operativos (neste caso fica evidente uma oportunidade perdida de ter economizado no passado).

A figura 1.1 ilustra de modo bem simplificado este processo de decisão. Neste exemplo, há apenas dois cenários de afluições futuras: alta e baixa. Portanto, teremos quatro cenários de decisão/consequência possíveis, resultantes da combinação das duas decisões operativas imediatas (acionar ou não as térmicas) e dos dois cenários de afluição futura. Por exemplo, se tomarmos a decisão de acionar as térmicas, e ocorrer uma afluição alta no futuro (primeiro ramo da figura acima), a consequência será um vertimento no sistema. Isto significa, “olhando pelo retrovisor”, que a geração térmica foi exagerada, que é uma consequência indesejável.

Por outro lado, se a afluição futura for baixa (segundo ramo da figura), a operação terá um resultado favorável, pois usaremos a energia hidrelétrica poupada com o acionamento térmico para evitar o uso de térmicas mais caras, ou mesmo um racionamento, no futuro. Em resumo: a primeira decisão, de acionar as térmicas, tem duas consequências possíveis, uma desfavorável e outra favorável.

Analisando agora as consequências da segunda decisão, não acionar as térmicas, há novamente uma consequência favorável e outra desfavorável. Porém, desta vez, as consequências são trocadas: se ocorrerem vazões altas, a operação terá sido boa, pois evitamos geração térmica e reenchemos os reservatórios; se ocorrerem vazões baixas, haverá um *déficit de suprimento* (acionamento).

Portanto, em contraste com os sistemas puramente térmicos, cuja operação é desacoplada no tempo, a operação de um sistema hidrotérmico é um problema aco-

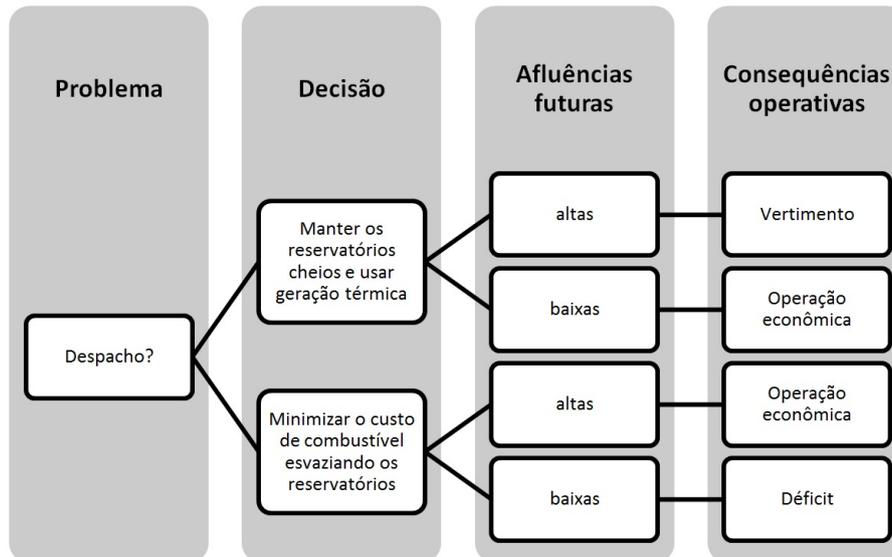


Figura 1.1: Processo de decisão operativa com incerteza nas vazões

plado no tempo, isto é, uma decisão operativa hoje afeta o custo operativo futuro [49]. É possível perceber claramente que existe uma relação de compromisso entre a utilização de água para geração de energia elétrica no presente e seu armazenamento para o futuro exigindo que o operador de um sistema compare, a cada intervalo de tempo, o benefício imediato do uso da água com o benefício futuro de seu armazenamento. Tipicamente, representa-se este *trade-off* através de duas funções, conforme ilustrado na figura 1.2.

A função de custo imediato (FCI) mede, principalmente, os custos de geração térmica no estágio corrente. Observa-se que o custo imediato aumenta à medida que diminui a energia hidrelétrica disponível no estágio, ou seja, à medida que aumenta o volume armazenado final. Por sua vez, a função de custo futuro (FCF) está associada ao custo esperado de geração térmica e racionamento do final do estágio (início do estágio seguinte) até o término do período de estudo. Esta função diminui à medida que aumenta o volume armazenado final, pois haverá mais energia disponível no futuro.

O uso ótimo da água corresponde ao ponto que minimiza a soma dos custos imediato e futuro. Como ilustrado na figura 1.3, o ponto de mínimo custo global também corresponde ao ponto em que as derivadas da função de custo imediato e da função de custo futuro se igualam em módulo. Esta derivada é conhecida como *valor da água*.

Uma questão bastante importante e que deve ser considerada nos estudos de planejamento da operação é a presença de diversas fontes de incerteza envolvidas neste processo. É possível citar, por exemplo, incertezas macroeconômicas que afetam o crescimento do país e, com isso, a demanda energética, disponibilidade dos

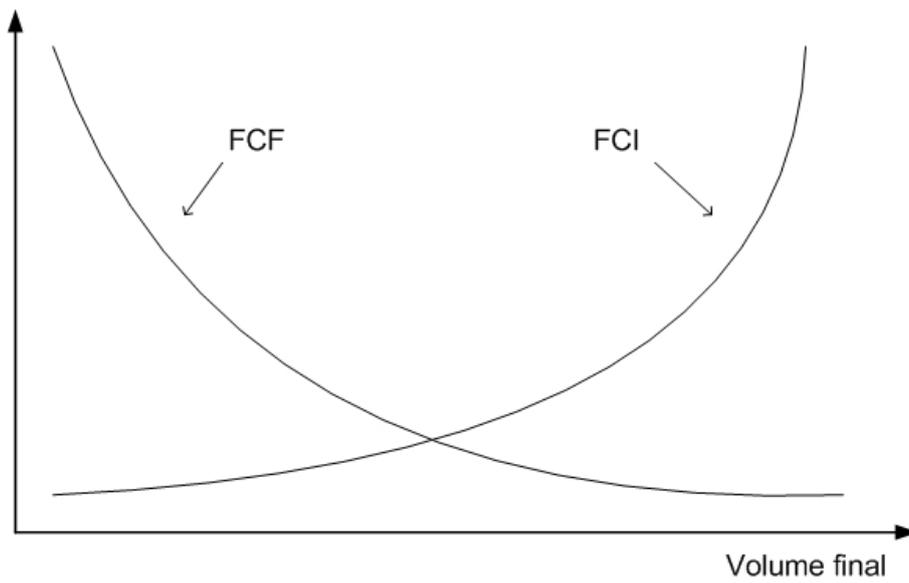


Figura 1.2: Função de custo imediato e futuro em função do armazenamento

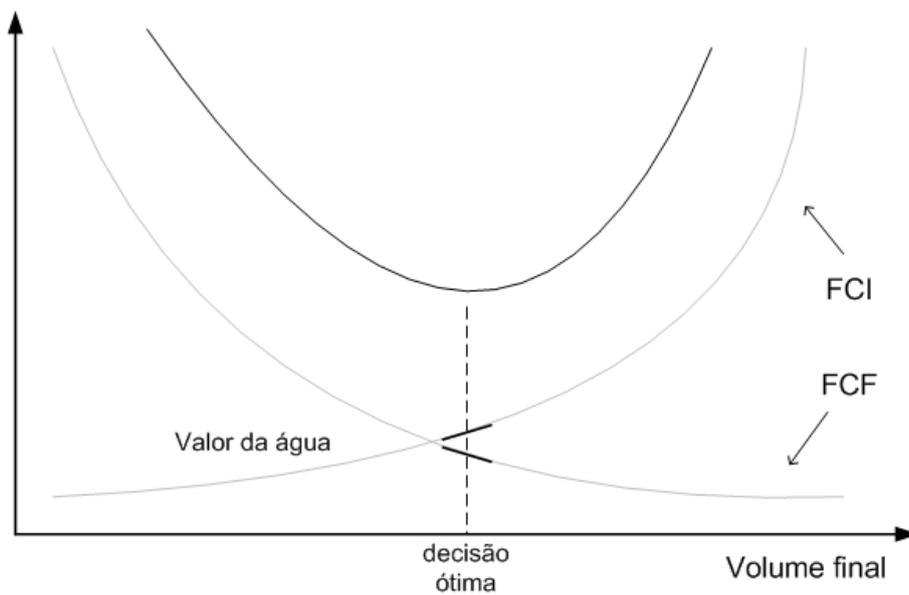


Figura 1.3: Uso ótimo da água

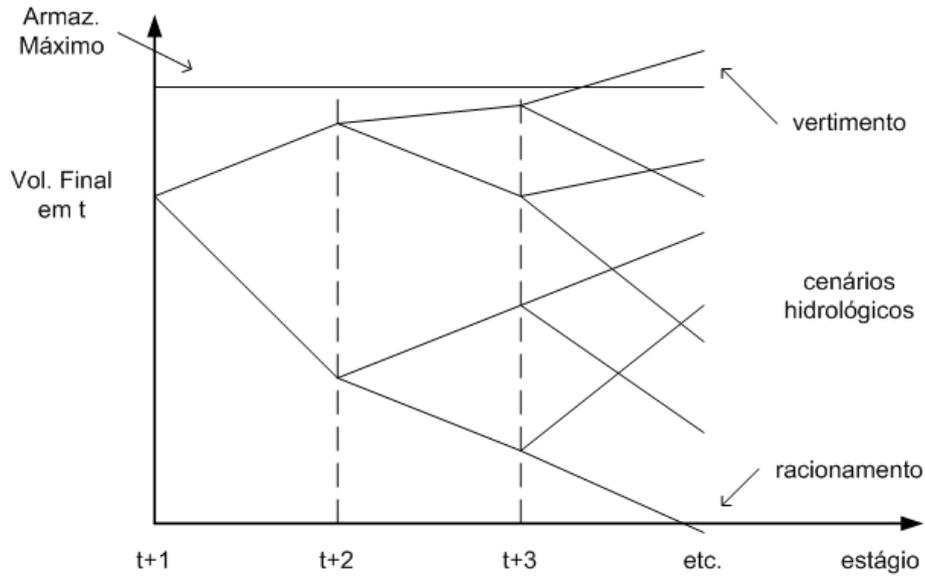


Figura 1.4: Representação da incerteza hidrológica para um reservatório

recursos, custo de combustíveis, afluências hidrológicas, entre outras.

Um exemplo da influência destas incertezas é a influência sobre o custo de produção de energia, que está diretamente associado à disponibilidade dos recursos hídricos. Em sistemas hidrotérmicos, como é o caso do Brasil, as incertezas mais significativas estão associadas principalmente aos recursos representados pelas afluências futuras das usinas hidroelétricas, em especial nos horizontes de médio e longo prazos. Estas incertezas devem ser consideradas nos estudos do planejamento da operação e, tradicionalmente, são realizadas simulações operativas do sistema para diversas hipóteses de afluências futuras. A figura 1.4 ilustra a possível evolução do nível de armazenamento de uma usina hidrelétrica para diferentes cenários de afluências.

Uma característica crucial do problema operativo é que a decisão deve ser tomada sob incerteza, isto é, sem saber se as afluências futuras são favoráveis ou desfavoráveis: apenas se conhece sua probabilidade de ocorrência (esta probabilidade é estimada a partir de análises estatísticas dos registros históricos das vazões). Este importante aspecto será discutido a seguir.

1.2 Incertezas no problema do planejamento da operação

Como foi observado, o problema de planejamento da operação de médio e longo prazo deve levar em consideração as incertezas associadas aos parâmetros de entrada do problema, em especial quanto às vazões afluentes. Isto ocorre em geral

porque os dados podem ser inerentemente aleatórios ou existem erros inevitáveis no processo de medição. Pode-se citar o trabalho de MASSÉ [40], em 1946, como uma das primeiras referências encontradas na literatura que sugere que os modelos matemáticos aplicados à gerência de recursos naturais deveriam levar em consideração parâmetros incertos. Em 1955, DANTZIG [24] e BEALE [11] identificaram que mesmo uma diferença relativamente pequena nos valores usados como dado de entrada poderiam comprometer bastante a qualidade da solução encontrada para um problema.

Desde então, diversas abordagens foram desenvolvidas para lidar com dados de entrada incertos em problemas de programação matemática envolvendo decisões imediatas e, de certa maneira, irreversíveis. É possível classificá-las em duas vertentes principais conhecidas como otimização robusta [13, 14] e programação estocástica [17]. Estes métodos diferem basicamente na maneira em que são consideradas as incertezas associadas aos dados e parâmetros do modelo.

Apesar de, na atualidade, não existirem modelos capazes de prever as afluências futuras com a precisão desejada para o horizonte em questão, é sabido que o comportamento das vazões não é completamente aleatório, apresentando padrões sazonais e uma tendência que depende das vazões ocorridas anteriormente. Esta característica favorece a utilização de métodos de programação estocástica, que se baseiam na hipótese de que a distribuição de probabilidade “real” das variáveis aleatórias pode ser razoavelmente aproximada por uma distribuição conhecida.

Métodos de análises de cenários representativos para dados estocásticos tem sido bastante utilizados em métodos de otimização de problemas estocásticos implícitos ou explícitos. Uma revisão detalhada da literatura sobre o tema de geração de cenários pode ser encontrada no trabalho de DUPACOVA *et al.* [28].

De fato, técnicas de programação estocástica tem sido utilizadas por mais de trinta anos em problemas de planejamento de sistemas elétricos e, atualmente, são o padrão adotado no setor. Como será visto a seguir, diferentes abordagens foram desenvolvidas e aplicadas ao longo desses anos com o objetivo de solucionar este problema.

A versão clássica para este modelo representa as incertezas considerando que o processo aleatório pode ser discretizado e representa sua evolução através da construção de uma árvore de cenários que possui ramificações a cada estágio. À cada nó desta árvore, está associado um problema de otimização que deve ser solucionado baseado nas condições iniciais provenientes do problema antecessor e cuja solução servirá de condição inicial para os próximos problemas (ramificações). A árvore de cenários é capaz de capturar a dinâmica do processo decisório uma vez que as decisões vão sendo ajustadas com base na informação conhecida até o momento. É possível encontrar diferentes métodos propostos para criar árvores de cenários

aderentes às propriedades estatísticas das variáveis estocásticas como métodos de clusterização [50], não-lineares [32], heurísticos [33] assim como métodos que consideram múltiplos cenários [56].

A representação explícita da árvore de cenários permite que o problema resultante seja solucionado diretamente, por exemplo, como um único problema de programação linear comumente denominado de *equivalente determinístico* [68]. Entretanto, devido às características de sua construção, o número de nós da árvore cresce exponencialmente com o número de estágios e, mesmo para um problema de poucas etapas, torna-se difícil a solução do problema completo. Técnicas de redução da árvore de cenários foram aplicadas como, por exemplo, a metodologia proposta por COSTA [19] que emprega uma técnica especializada para a criação de uma árvore de cenários reduzida, porém representativa, para cenários multivariados de afluências voltada para utilização no problema de planejamento da operação de médio prazo de sistemas hidrotérmicos.

Entretanto, esta abordagem ainda se mostra limitada a problemas com poucas etapas. O principal problema é que, em esquemas de solução baseados em árvore de cenários, o número de nós desta árvore explode exponencialmente com o número de estágios e o tamanho do problema a ser solucionado é proporcional ao número de nós da árvore. Como alternativa, foram propostos métodos de amostragem de cenários baseados em técnicas de Monte Carlo [38] em troca de se admitir uma solução aproximada frente à “árvore original”.

Em seguida, foram propostos métodos de decomposição do problema baseados em programação dinâmica [12]. Esta técnica implementa a decomposição temporal do problema, o que permite a solução de problemas com diversos estágios. A consideração de incertezas foi tratada com a extensão desta técnica para programação dinâmica estocástica (PDE ou, em inglês, SDP). Em termos gerais, a PDE trabalha com uma representação aproximada da função de custo futuro, que é calculada recursivamente desde o último estágio até o estágio inicial. Durante o processo de recursão, a FCF de cada etapa é avaliada para diferentes estados e, na etapa anterior, é utilizada uma aproximação baseada na interpolação da solução do problema para os estados avaliados.

O esquema PDE foi utilizado por vários anos em muitos países com sistemas hidroelétricos. Porém, apesar de permitir a solução de problemas maiores a PDE possui uma limitação importante pois o processo de recursão requer a enumeração de todas, ou de um número razoável, de combinações dos valores das variáveis de estado¹ para representar de maneira adequada a FCF. Como consequência, o esforço computacional aumenta exponencialmente com o número de variáveis de estado.

¹No caso do problema de sistemas hidrotérmicos os estados são o armazenamento inicial e as afluências anteriores em cada reservatório.

Tabela 1.1: Maldição da dimensionalidade

Número de reservatórios	Número de Estados
1	$20^2 = 400$
2	$20^4 = 160$ mil
3	$20^6 = 64$ milhões
4	$20^8 \approx 25$ bilhões
5	$20^{10} \approx 10$ trilhões

Este crescimento de esforço computacional, conhecido como a “maldição da dimensionalidade” da programação dinâmica, está ilustrado na tabela 1.1, que mostra o crescimento do número de estados com o aumento do número de reservatórios, supondo que os níveis dos reservatórios e as afluências são discretizadas por 20 valores cada.

Em 1982, YAKOWITZ [69] fez uma revisão sobre as técnicas de programação dinâmica utilizadas até então para a gerência de recursos hídricos e relata severas limitações computacionais que restringiam a aplicação de programação dinâmica à problemas de pequeno porte.

Na prática, para utilizar tais métodos é necessário recorrer a agregações ou decomposições do problema antes de solucioná-lo. Algumas heurísticas relativamente boas foram desenvolvidas para contornar este problema e aplicações importantes são apresentadas por TERRY *et al.* [64]. Uma abordagem um pouco diferente para lidar com problemas multi reservatório foi proposta por TURGEON [66] e SHERKAT *et al.* [62].

Apesar das heurísticas propostas viabilizarem a solução aproximada dos problemas, métodos que conseguissem lidar com problemas multidimensionais com diversas variáveis de estado ainda se mostravam necessários. Em meados do anos 80, pesquisadores da Universidade de Stanford começaram a abordar o problema utilizando técnicas que estendiam o trabalho original de BENDERS [15] chamadas de decomposição de Benders multiestágio (aninhada) [16]. Em seguida, estes métodos foram utilizados e refinados nos trabalhos de PEREIRA e PINTO [47], JACOBS *et al.* [36] e MORTON [42].

Neste contexto, o algoritmo proposto por PEREIRA e PINTO [48] despertou bastante interesse na comunidade de otimização de sistemas energéticos. Batizado de Programação Dinâmica Dual Estocástica (PDDE ou SDDP, do termo em inglês) ele é capaz de combinar de maneira eficaz as características da PDE e os benefícios da decomposição de Benders multiestágio.

Em termos gerais, o algoritmo de PDDE emprega um procedimento iterativo

onde, a cada iteração, soluciona-se o problema com uma aproximação subestimada da FCF e utiliza-se a informação dual para refiná-la, através da adição de cortes de Benders. Este procedimento evita a discretização do espaço de estados pois a FCF é construída iterativamente e extrapolada no entorno das soluções encontradas.

Ele representa uma evolução importante do algoritmos de decomposição de Benders multiestágio utilizando dois conceitos importantes: o compartilhamento de cortes e a amostragem de cenários. Existem diversas implementações deste algoritmos utilizadas comercialmente em larga escala na atualidade [53].

1.3 Segurança energética

Além dos custos incorridos devido ao consumo de combustível das térmicas para geração, outro aspecto muito importante para o planejamento da operação é a garantia do suprimento de energia dentro de critérios de qualidade e segurança previamente estabelecidos. Em sistemas sujeitos à incertezas sobre as condições futuras, este tema ganha importância adicional pois faz-se necessário o monitoramento constante dos déficits de energia previstos, assim como a aplicação de medidas corretivas para evitá-los e garantir a minimização dos impactos para a sociedade.

Do ponto de vista matemático, o déficit de energia é representado como uma “capacidade de geração” adicional que pode ser acionada caso os demais recursos sejam escassos ou, em situações menos típicas, caso seja interessante do ponto de vista econômico. Este gerador complementar é responsável por garantir a viabilidade do modelo em caso de escassez pois é capaz de “fechar” a equação de atendimento à demanda².

Há dois enfoques básicos para estabelecer a solução de compromisso entre os custos operativos e a qualidade no suprimento de energia no planejamento da expansão e operação do sistema. No primeiro, supõe-se conhecido o valor econômico do não atendimento ao mercado de energia e minimiza-se o custo total de operação do sistema. Os parâmetros operativos que refletem a qualidade do suprimento são obtidos como consequências dos custos de geração, investimento, combustíveis e do custo de déficit pré-fixado.

Nesta linha, o valor econômico dos déficits de energia, ou custo social do déficit, é determinado através de estudos macroeconômicos constantemente atualizados que avaliam o impacto causado pela perda da produção agregada (PIB) nas diferentes atividades econômicas no caso de um racionamento. Para a avaliação global dos custos associados ao país, emprega-se um método baseado na matriz de inter-relações setoriais (matriz insumo-produto), enquanto que para avaliações regionais usem-se métodos baseados em regressões econométricas.

²Note que o déficit é limitado, em princípio, à demanda total do sistema.

Embora conceitualmente perfeita, esta abordagem tem sua aplicação dificultada pela necessidade de se dispor de dados econômicos complexos, confiáveis e atualizados para a avaliação do custo social do déficit [41]. Sendo assim, tais estudos são bastante complexos e, na prática, verifica-se indiretamente se o parâmetro custo de déficit adotado foi suficiente para garantir o atendimento do mínimo índice de qualidade de suprimento aceitável para a política estratégia.

Outro fator que merece atenção é a dependência da qualidade do suprimento com relação aos custos de combustível das usinas termelétricas. Como matematicamente o déficit pode ser visto como uma térmica adicional, seu acionamento depende de seu custo quando comparado ao custo das demais fontes de energia. Caso haja um aumento generalizado dos custos de combustível, o déficit passa a ser relativamente menos oneroso e, por esta razão, a operação do sistema estaria sujeita a acioná-lo mais vezes, acarretando em maiores riscos para sociedade.

Além disto, é importante notar que mesmo que os custos de déficit sejam determinados de forma bastante acurada nenhum planejador aceitaria níveis de risco acima do estabelecido como resultado da simulação da política operativa obtida e, certamente, revisaria os critérios adotados para o cálculo destes custos caso isto acontecesse.

No outro enfoque, adota-se um índice pré-fixado de qualidade de suprimento considerado adequado para o planejamento e a operação do sistema e determina-se a política operativa considerando-se a minimização dos custos sujeito à restrições que garantem que o índice mínimo de qualidade não será violado. Neste caso, é possível calcular um *custo de déficit implícito*, associado à consequência do índice de qualidade (risco) pré-fixado, dos custos relativos de geração e combustíveis.

1.4 Motivação

Em 2001, o Brasil sofreu séria crise no abastecimento de energia elétrica que culminou com o racionamento do consumo devido ao elevado risco de ocorrência de um apagão sistêmico. A insuficiência de chuvas experimentada naquele período, precedida por algumas interrupções imprevistas no fornecimento de energia elétrica no final da década de 1990, foi considerada apenas como fator de antecipação dos problemas de abastecimento que ocorreriam cedo ou tarde, em razão do progressivo “desestoque” dos reservatórios das usinas hidrelétricas. Tal fato teria sido provocado, principalmente, pela queda dos investimentos no setor a partir da década de 1980 que causou atrasos e a não-realização de obras fundamentais à devida expansão das atividades de geração e de transmissão em níveis compatíveis com o desenvolvimento econômico e social do país.

A ocorrência de racionamentos acarreta prejuízos diversos a um país, como os

registrados no Brasil em 2001. Naquele ano, os níveis de demanda de energia elétrica caíram ao patamar registrado em 1997 (redução em cerca de 20%). Parte dessa queda no consumo esteve associada ao uso mais eficiente da energia elétrica, outra parte, porém, deu-se em razão da redução forçada da atividade produtiva do país. Além da diminuição global no consumo, o impacto sobre os consumidores foi direto, com elevação das tarifas.

Diante disto, é importante avaliar o problema sob a ótica dos tomadores de decisão. Em geral, os modelos aplicados ao problema do planejamento da operação de sistemas hidrotérmicos eram e, de certa forma, ainda são *neutros ao risco* no sentido de que eles buscam por políticas que minimizam o valor esperado do custo de operação. Como visto até o momento, em sistemas hidrotérmicos estes custos são oriundos dos gastos com combustível térmicos e penalidades, como as aplicadas à défits no suprimento.

Sob esta visão, um operador do sistema racional aceitaria défits ocasionais se estes levassem à minimização dos custos com combustíveis térmicos no longo prazo. Na prática, défits de energia não ocorrem com frequência mas, quando eles ocorrem causam um transtorno tão grande para a sociedade que em geral políticos e operadores do sistema preferem por evitá-los. Sendo assim, é primordial buscar por modelos que possam encontrar políticas operativas *aversas ao risco* de racionamento.

Como discutido anteriormente, o algoritmo de otimização assume que existe uma penalidade para falhas no suprimento da demanda. Idealmente, os custos associados à racionamentos de energia deveriam refletir a redução nas atividades econômicas causadas pela interrupção do atendimento da demanda. Entretanto, estes indicadores macroeconômicos são extremamente difíceis de serem estimados [57]. Como consequência, a adequação de uma estratégia operativa é geralmente mensurada pela confiabilidade no suprimento de energia que ela é capaz de assegurar [25]. Em outras palavras, um caminho viável para os estudos de planejamento da operação é a busca de estratégias operativas que minimizem o valor esperado do custo operativo térmico e satisfaçam uma *restrição* de confiabilidade pré-estabelecida.

No Brasil, que é um sistema com forte predominância hidrelétrica, este tema tem uma relevância ainda maior. Como as termelétricas correspondem a apenas 20% da capacidade instalada, é inútil acioná-las somente quando os níveis dos reservatórios estão extremamente reduzidos. A única opção é acioná-las com antecedência de vários meses, para que a contribuição acumulada da energia gerada pelas termelétricas seja relevante. Por outro lado, quanto maior a antecedência do acionamento das térmicas, maior a incerteza quanto às vazões afluentes futuras e, com isso, a probabilidade de acionar térmicas desnecessariamente.

Após o racionamento de 2001-2002, foi incorporado ao processo decisório um critério de segurança de suprimento, conhecido como Curva de Aversão ao Risco

(CAR) [18]. A CAR baseia-se na adição de curvas bianuais de segurança (para os primeiros dois anos de estudo) que penalizam a utilização da água nos reservatórios caso a energia armazenada no sistema fique inferior à curva.

Entretanto, este mecanismo de penalização não é capaz de sinalizar corretamente o risco de déficit futuro. A justificativa para isto consiste no fato de os modelos estocásticos de otimização tradicionais utilizados buscam o despacho que minimiza os custos de suprimento da demanda sobre a *média* dos cenários hidrológicos considerados. Tais modelos são caracterizados como “neutros com relação ao risco” e, como a influência da violação das restrições é refletida no valor esperado do custo total, não há garantia de proteção para os eventos de maior arrependimento que correspondem à ocorrência de séries críticas e à violação de um nível de segurança previamente estabelecido.

Devido à possível vulnerabilidade desta metodologia, o Operador Nacional do Sistema (ONS) propôs uma medida adicional de segurança, conhecida como Procedimentos Operativos de Curto Prazo (POCP). Esta medida é composta pelos seguintes passos:

- (i) calcula-se o despacho através da solução do problema de planejamento ótimo da operação;
- (ii) realiza-se uma simulação determinística considerando que todos os recursos termoeletrônicos estão despachados nos próximos meses;
- (iii) calcula-se o montante de térmica a ser despachada além da decisão ótima do problema de planejamento para que o nível dos reservatórios no final do período seco (novembro) não seja inferior a um nível meta pré-estabelecido.

Em suma, o POCP é um recurso utilizado pelo ONS para evitar que o sistema atinja estados considerados inseguros e que o modelo de planejamento da operação não é capaz de capturar.

Em 2009, COSTA JR. *et al.* [22] publicaram um estudo com a avaliação da eficiência econômica da aplicação do POCP ao despacho do sistema brasileiro. Através de uma análise comparativa entre duas simulações, sem e com a representação do POCP, foi estimado um índice custo/benefício dado pela razão entre o custo associado ao despacho térmico adicional e o déficit de energia poupado com a aplicação do POCP. Os resultados mostraram este índice era superior ao próprio custo de déficit utilizado pelo ONS, o que indicou que a decisão de acionar térmicas não se justificava sob o ponto de vista econômico.

Este exemplo real do caso brasileiro ressalta a importância de se considerar restrições de risco no modelo de planejamento, considerando de maneira efetiva a busca de uma solução ótima que reflita os critérios de aversão ao risco do operador.

Neste caso, fica evidente o descasamento entre o valor considerado para o déficit e o valor associado à restrição indiretamente imposta pelo POCP.

Em 1974, ASKEW [9] propôs uma abordagem alternativa onde utilizava-se a função de custo de déficit como um parâmetro para assegurar que a estratégia operativa resultante satisfizesse uma meta pré estabelecida para um determinado índice de confiabilidade. Em seu procedimento de “penalização aumentada”, a política operativa era calculada através de PDE para diferentes funções de custo de déficit, que eram ajustadas iterativamente até que o índice de confiabilidade fosse igual, dada uma tolerância, ao valor pré estabelecido. Um procedimento similar a este foi utilizado há muitos anos no sistema brasileiro utilizando-se uma função de custo de déficit linear e índice de confiabilidade igual à probabilidade de ocorrência de qualquer déficit (risco de déficit).

A busca pela função de custo de déficit ideal fica evidente quando se descreve o problema de planejamento da operação da seguinte maneira:

“Escolha, entre todas as possíveis estratégias operativas, aquela que minimiza o valor esperado dos custos de geração e satisfaça as restrições de confiabilidade desejadas.”

Dada esta ótica, é possível questionar: que função de custo de déficit é mais adequada para encontrar a estratégia operativa ótima? Qual seria a forma desta função? Linear? Não linear?

ROSSMAN [55] demonstrou através de teoria de relaxação Lagrangeana que a restrição de confiabilidade considerada por ASKEW poderia ser incorporada na função objetivo de modo similar ao procedimento de penalização aumentada e que o ajuste de apenas um parâmetro na função penalidade era suficiente para garantir a otimalidade do problema. É importante ressaltar que a aplicação deste procedimento depende do índice de risco adotado. Por exemplo, é sabido que o índice de risco probabilidade de ocorrência de déficit possui características não convexas o que pode limitar a utilização de técnicas de solução que dependam da convexidade do problema para garantir a otimalidade da solução encontrada.

Outro aspecto importante é que, em termos de restrições de confiabilidade, não importa em que etapa uma restrição está sendo violada. Isto implica que a penalidade associada à violação desta restrição não deveria ser afetada pela taxa de desconto tipicamente utilizada para avaliar o valor esperado dos benefícios futuros em termos de valor presente. Em outras palavras, verificou-se que a taxa de desconto diminui a importância de uma violação futura não sinalizando, desta forma, que a política operativa deve ser alterada para poupar água nos reservatórios com a antecedência necessária.

Os estudos de ROSSMAN também mostraram um resultado inesperado ao alterar o algoritmo para não considerar a taxa de desconto para a função de custo

de déficit. A política operativa obtida antecipou os déficits de energia permitidos (abaixo do critério pré estabelecido) pois, devido à taxa de desconto aplicada aos custos das termoeletricas, é mais benéfico deixar de evitá-los nos primeiros estágios do que nos estágios ao final do horizonte. Claramente, esta política operativa não é aceitável em termos de estudos de planejamento da operação e este resultado ilustra é necessário ser cuidadoso na utilização da função custo de déficit para controlar indiretamente os índices de risco desejados.

Em 1979, SNIEDOVICH [63] utilizou técnicas de programação dinâmica para considerar problemas com restrições de risco de qualquer racionamento ao longo do horizonte de planejamento da operação. Também foi demonstrado que é possível utilizar o procedimento de penalização aumentada para solucionar este problema com o auxílio de uma variável de estado adicional que representava o número de falhas no atendimento à demanda até a etapa corrente.

Em 1985, ARARIPE NETO *et al.* [5] apresentaram um modelo de planejamento da operação solucionado por PDE onde era considerado um conjunto de problemas aninhados que, por sua vez, implicavam em restrições de risco de racionamento aninhadas. Mais recentemente, outros trabalhos como o de SHAPIRO [58] e PHILPOTT e DE MATOS [51] também utilizaram esta formulação de restrições aninhadas. A justificativa para este tipo de abordagem consiste em reproduzir a lógica da tomada de decisões ao longo do tempo.

Entretanto, o uso de restrições aninhadas onde as restrições são redefinidas a cada estado introduz dificuldades na interpretação e na definição dos critérios de risco adotados. Para contornar este problema, ARARIPE NETO *et al.* [5] propõe a utilização de uma heurística para definição do critério confiabilidade que baseia-se na multiplicação de um fator de redução aplicado ao risco resultante a partir da solução da política obtida com a formulação neutra ao risco. Neste caso, o critério de risco fica condicionado às condições iniciais do problema assim como às características dos dados. Porém, intuitivamente, o critério de risco é uma decisão de planejamento e não deveria haver esta dependência.

Diferentes medidas de risco tem sido utilizadas no setor elétrico, além do próprio risco de racionamento e do valor esperado da energia não suprida (EENS³). O *Value at Risk* (VaR) tem se destacado e se estabeleceu como padrão no setor bancário e demais áreas de aplicação [37] e é uma medida de risco importante utilizada no setor elétrico, principalmente com relações à decisões financeiras. Entretanto, desde que ROCKAFELLAR e URYASEV [54] introduziram sua variante chamada *Conditional Value at Risk* (CVaR) esta medida ganhou força entre os gestores de risco do setor financeiro. Propriedades matemáticas como convexidade e coerência [8] fizeram do CVaR uma medida de risco bastante atraente para aplicações em problemas

³do termo em inglês: *Expected Energy Not Supplied*

de energia, especialmente no contexto do planejamento da operação de sistemas hidrotérmicos [31].

Outros autores também desenvolveram implementações que levam em consideração tais medidas de risco. Em 2005, JABR [35] apresentou um modelo utilizando CVaR para a solução de problemas de despacho em sistemas descentralizados com incertezas no preços de energia. Em 2007, ILIADIS *et al.* [34] descreve um modelo de despacho hidrotérmico que considera o valor em risco condicionado da perda de receita acumulada ao final do horizonte de planejamento.

Em 2008, COSTA JR. [21] propôs uma metodologia para incorporar restrições de confiabilidade no problema do planejamento ótimo da expansão de sistema elétricos. Em sua proposta, eram adicionadas restrições de risco que limitavam o valor esperado da potência não suprida (EPNS) e o CVaR relativo ao corte de carga (potência) do sistema sob a hipótese de que as incertezas do estado operativo dos equipamentos não afetavam a operação da etapa seguinte (independência temporal). Para solucionar o problema utilizou-se a técnica de decomposição de Benders através da solução de um problema mestre de investimento e subproblemas para avaliação da política operativa e confiabilidade.

GUIGUES e SAGASTIZÁBAL [30] estudaram um modelo de horizonte rolante que soluciona e implementa a solução do problema de um estágio com restrições em probabilidade. GUIGUES e RÖMISCH [29] apresentaram uma abordagem genérica para a extensão de medidas de risco poliédricas no contexto da PDDE. A medida de risco apresentada utiliza um espaço de estados aumentado por um vetor de custos representando o possível histórico até a etapa corrente.

Mais recentemente, os trabalhos de SHAPIRO [59], SHAPIRO *et al.* [61] e PHILPOTT e DE MATOS [51] propuseram a utilização da medida de risco CVaR associada ao custo operativo onde em cada etapa do problema, minimiza-se uma combinação convexa do valor esperado e do CVaR associados aos custos futuros. De uma maneira simplificada, esta formulação é análoga a aumentar o peso relativo dos cenários de afluências mais desfavoráveis (que estão associados aos cenários de custos mais elevados e, portanto, são componentes do CVaR). Em seguida, DINIZ *et al.* [27] propuseram uma implementação que simplifica o processo de cálculo dos cortes da FCF na recursão backward da PDDE. As simulações apresentadas com este procedimento aumentam a segurança de suprimento, como esperado. No entanto, há discussões sobre a interpretação dos resultados e sobre o processo de convergência do algoritmo.

Os modelos propostos consideram a minimização de medidas de risco associadas a um valor econômico na função objetivo do problema. Como foi visto, esta abordagem pode diluir a importância das restrições de risco devido aplicação da taxa de desconto à uma penalidade definida à priori. SHAPIRO [59] aponta como uma vantagem

desta abordagem o fato de não possuir o inconveniente de causar inviabilidades ao problema de otimização. Por outro lado, a representação explícita de uma restrição associada à uma medida de risco pode ter uma explicação mais direta e intuitiva do que a atribuição arbitrária de uma penalidade à violação da mesma.

1.5 Objetivo e contribuições

Como foi visto, a complexidade do problema de planejamento operativo resulta de três características: (i) a decisão ótima para um estágio depende dos estágios seguintes (acoplamento temporal), pois é possível armazenar a água nos reservatórios para uso futuro; (ii) toda decisão operativa é tomada sob incerteza, isto é, não sabemos quais são as afluências futuras; só conhecemos a distribuição de probabilidade das mesmas; e (iii) há uma interdependência entre o despacho econômico e a confiabilidade de suprimento: a operação de maior segurança energética é, evidentemente, a que aciona permanentemente as térmicas em sua capacidade máxima; no entanto, esta operação mais segura também é a mais cara. E vice-versa: a operação mais econômica corresponde a nunca acionar as térmicas, o que obviamente resulta na pior segurança de suprimento.

As características (i) e (ii) podem ser representadas e solucionadas através do algoritmo de programação dinâmica estocástica dual (PDDE). O resultado da PDDE é uma política operativa que minimiza o valor esperado do custo operativo (basicamente relacionado com os gastos de combustível das usinas térmicas). Para representar a característica (iii), o usuário deve informar o peso relativo das entre falhas de suprimento (cortes de carga, ou déficits) com relação aos custos operativos.

Uma maneira de representar este peso relativo é calcular o custo econômico do déficit. Isto pode ser feito através de uma matriz insumo-produto da economia, onde basicamente se calcula a perda da produção agregada (PIB) no caso de um racionamento. O resultado é uma função custo x profundidade do déficit. Se a função for linear, ela é caracterizada por um único parâmetro, o custo unitário do déficit. Por exemplo, o custo unitário do déficit utilizado pela Empresa de Pesquisa Energética (EPE) em seus estudos é R\$ 3.100 /MWh. Uma vez conhecida a função de custo de déficit, o objetivo da PDDE passa a ser o de minimizar o valor esperado da soma dos custos operativo e de déficit.

No entanto, o uso do custo econômico do déficit tem duas limitações importantes: (i) incerteza nos parâmetros da matriz insumo-produto, que impacta a precisão do cálculo deste custo; e (ii) a confiabilidade de suprimento é um *resultado* do processo de otimização da PDDE, não um parâmetro de entrada para o mesmo. Isto significa que, se por alguma razão, houver uma mudança na aversão ao risco da sociedade, não seria possível representar a mesma no cálculo da política operativa.

Como incorporar a aversão ao risco no processo de decisão? As diferentes metodologias de aversão ao risco podem ser classificadas em três categorias: (i) adição de um “prêmio de risco” ao custo econômico do déficit; (ii) uso de uma reserva energética de segurança que evite a ocorrência de déficits caso ocorra um cenário desfavorável; e (iii) aumento do peso relativo dos cenários de vazão desfavoráveis no cálculo do custo futuro.

A metodologia (i) (prêmio de risco) foi usada na década de 1980, com o nome “custo implícito do déficit”, pelo planejamento da expansão do setor (GCPS). Por sua vez, a metodologia (ii) (reserva de segurança) embasa os procedimentos de Curva de Aversão a Risco, Procedimento Operativo de Curto Prazo (POCP) e, mais recentemente, da Superfície de Aversão a Risco (SAR). Finalmente, a metodologia (iii) (maior peso aos cenários de vazões desfavoráveis) é equivalente, em termos matemáticos e resultados, à metodologia que considera uma medida de risco de CVaR associados aos custos na função objetivo.

Dado que esta mudança no perfil de aversão ao risco vem ocorrendo na operação do sistema brasileiro, como evidenciado pela introdução da Curva de Aversão a Risco CAR em 2002; do Procedimento Operativo de Curto Prazo (POCP) em 2008; e pelo maior acionamento das usinas térmicas desde outubro de 2012, passa a ser interessante analisar a possibilidade de determinar uma política operativa onde o objetivo é minimizar o valor esperado do custo operativo, sujeito a uma *restrição* na confiabilidade de suprimento.

O objetivo deste trabalho é mostrar como restrições relacionadas com a confiabilidade de suprimento podem ser incorporadas no algoritmo de PDDE. Em particular, será mostrado que o índice de risco correspondendo ao CVaR associado ao déficit de energia apresenta uma série de características vantajosas: adequação à percepção de aversão ao risco; facilidade de interpretação dos resultados para os usuários e sociedade em geral; simplicidade de implementação no modelo computacional; e eficiência computacional no cálculo da política operativa com restrição de CVaR.

Embora seja possível encontrar na literatura propostas para a consideração de restrições de risco em problemas estocásticos, ainda não foi proposta uma metodologia que considere explicitamente o risco, em forma de restrição, incorporado na metodologia de Programação Dinâmica Dual Estocástica de PEREIRA e PINTO [47, 48], utilizada largamente para a solução do problema de planejamento da operação de sistemas hidrotérmicos.

Em termos gerais, o algoritmo de PDDE é um processo que evita a conhecida “maldição da dimensionalidade” pois basea-se na representação linear por partes da função de custo futuro através de hiperplanos e, por esta razão, evita a necessidade da discretização do espaço de estados; e trata o crescimento exponencial da árvore de cenários assumindo que é possível compartilhar cortes entre os estados além

de empregar um processo de amostragem de cenários o que permite, efetivamente, solucionar a árvore de incertezas implicitamente.

A característica principal do PDDE é a decomposição do problema estocástico multiestágio original em diversos problemas menores de um estágio e um único cenário, solucionados sucessivamente. Este enfoque traz uma grande flexibilidade pois permite que problemas de grandes dimensões sejam abordados, utilizando inclusive computação paralela, com o compromisso de aceitar uma tolerância de convergência, tão pequena quanto desejada, para a solução encontrada.

Como será visto adiante, a adição de uma restrição de risco para um dado estágio do problema envolve a consideração de variáveis de decisão associadas à todos os cenários em uma única restrição. A primeira vista, esta característica impede a decomposição do problema original em problemas menores, inviabilizando a aplicação do algoritmo de PDDE. Para evitar este problema, alguns trabalhos propõem que seja considerado um custo para estas restrições na função objetivo. Entretanto, além de alterar o problema original, este tipo de abordagem exige a definição a priori de penalidades o que dificulta o entendimento sobre os parâmetros do modelo e a sua interpretação.

A proposta deste trabalho consiste em aplicar uma estratégia de relaxação Lagrangeana dualizando a restrição de CVaR. Esta abordagem tem como objetivo se beneficiar do fato de que a dualização da restrição de risco permite manter o esquema original de separação do problema permitindo, então, a aplicação do algoritmo original.

Em um método de relaxação Lagrangeana tradicional é definida uma função Lagrangeana baseada na dualização de um conjunto de restrições e, em termos gerais, os métodos de solução buscam encontrar o conjunto de multiplicadores que maximizam a função Lagrangeana⁴. Estes métodos solucionam, iterativamente, o problema relaxado e, a partir de sua solução obtém indicadores para novos multiplicadores. Propõe-se utilizar um método de busca binária para a penalização da restrição dualizada onde, a cada iteração, é solucionado um problema de PDDE até a otimalidade.

Além de contribuir com uma proposta original de algoritmo para a solução de problemas estocásticos com restrição de risco no algoritmo da PDDE, é apresentada a formulação para inclusão desta restrição no problema de planejamento da operação.

Finalmente, há uma grande contribuição conceitual neste trabalho que consiste em mostrar que o modelo proposto é, na verdade, uma formulação matemática original para a metodologia de ajuste da curva de custo de déficit utilizada no passado. Com isso, é possível delinear uma importante interpretação econômica clara que associa diretamente a definição de uma curva de déficit de dois segmentos

⁴ considerando um problema original de minimização.

à uma restrição de confiabilidade de CVaR associada aos déficits de energia.

Será mostrado que é possível ajustar a profundidade e o prêmio de risco associados ao segundo segmento de déficit, que podem ser utilizados para representar a aversão a risco do operador. Esta contribuição é importante pois traduz os parâmetros e resultados da metodologia em quantia que são facilmente interpretadas pelos agentes do setor elétrico, em contraste com as metodologias da SAR e CVaR associado aos custos esperados, que necessitam de parâmetros como uma penalização e fator de ponderação que não têm interpretação física direta.

Do ponto de vista do planejamento do setor elétrico, tem-se como objetivo adicional a utilização da metodologia proposta como instrumento para a unificação dos critérios de planejamento da expansão, atualmente baseado em uma meta de risco, e da operação, que baseia-se na penalização dos déficits de energia ao longo do horizonte.

1.6 Organização da tese

Esta tese está organizada da seguinte forma: o capítulo 2 apresenta o arcabouço matemático necessário para a formulação de modelos de programação linear estocástica multiestágio. São apresentados os principais conceitos que envolvem a representação de incertezas e as premissas adotadas na formulação utilizada. Além disto, também é apresentado o algoritmo tradicional de Programação Dinâmica Dual Estocástica, que será utilizado como base para algoritmo proposto neste trabalho.

No capítulo 3 é apresentada a formulação típica do problema de planejamento da operação de sistemas hidrotérmicos, onde são consideradas restrições de atendimento à demanda total, balanço hídrico por usina hidrelétrica (considera-se uma representação individualizada das usinas) e limites operativos para hidrelétricas (volume e turbinamento) e termelétricas (geração). A função objetivo típica considerada para este problema é a minimização do valor esperado do custo de operação, composto pelos custos de geração térmica e penalizações devido à déficits de energia.

Em seguida, o capítulo 4 descreve medidas de risco que podem ser utilizadas para mensurar o comportamento de variáveis aleatórias. As medidas apresentadas são o risco (probabilidade de ocorrência de algum evento), valor esperado, *Value at Risk* (VaR) e *Conditional Value at Risk* (CVaR) juntamente com a respectiva formulação matemática para a sua incorporação em modelos de otimização. Além disto, também são exploradas as propriedades e as relações matemáticas entre estas medidas de risco.

O capítulo 5 apresenta uma proposta de extensão do modelo de planejamento da operação tradicional através da representação explícita de uma restrição que limita o CVaR associado à variável aleatória déficit de energia em um estágio pré-

determinado. A adição desta restrição torna o modelo não separável e é utilizada a técnica de relaxação Lagrangeana da restrição de risco para permitir a separação do problema e a sua utilização do algoritmo de PDDE para sua solução. Além disto, é proposto um algoritmo iterativo para a obtenção da solução ótima.

O capítulo 6 apresenta uma proposta para a formulação do problema de planejamento da operação de sistemas hidrotérmicos com restrição de CVaR aplicada ao déficit de energia de determinado estágio. São descritas as equações recursivas e os coeficientes dos cortes que descrevem a função de custo futuro calculada pela PDDE. Ao final, são apresentados resultados numéricos para um caso simplificado da Costa Rica que ilustram esta formulação.

No capítulo 7, é apresentada uma nova proposta de formulação que tem como objetivo superar as limitações da formulação proposta no capítulo 6. A formulação é alterada para considerar um limite de CVaR associado ao déficit incorrido em um período de tempo e uma proposta de algoritmo de solução para que seja considerada mais de uma restrição. Esta proposta permite limitar o CVaR associado *déficit acumulado* ao longo de um período e tem características atraentes que serão discutidas. Esta nova formulação é aplicada ao mesmo caso da Costa Rica e os resultados são comparados com a formulação anterior.

O capítulo 8 apresenta a aplicação da metodologia proposta no capítulo 7 a um estudo de caso de planejamento ótimo da operação do sistema brasileiro. É realizada uma análise comparativa dos resultados obtidos com o caso neutro ao risco e são apontadas as principais características da metodologia proposta.

Finalmente, o capítulo 9 apresenta as considerações finais onde são destacados os objetivos atingidos e propostos nesta tese. Adicionalmente, são apresentadas algumas propostas para desenvolvimentos futuros.

Capítulo 2

Programação estocástica multiestágio

2.1 Introdução

Desde os primórdios, planejadores identificaram a importância de se incorporar o efeito de incertezas nos modelos matemáticos. WALLACE e FLETEN [67] destacam comentários no livro de MASSÉ [40], em 1946, sobre sua experiência e preocupações a respeito do otimismo dos modelos determinísticos utilizados para a operação de reservatórios durante o período da Segunda Guerra Mundial. Apesar de naquela época os conceitos de programação estocástica ainda não estarem formalmente definidos, Massé já identificara que a decisão sobre a liberação da água dos reservatórios deveria basear-se na comparação do benefício imediato de seu uso e no *valor esperado* futuro da água. Suas declarações são mais abrangentes e estabelecem que, na verdade, os usos de todos os recursos naturais envolvem a avaliação do *trade-off* entre utilizá-los agora ou em um futuro estocástico.

Em seu primeiro trabalho sobre planejamento sob incerteza, DANTZIG [24] resalta a importância de se considerar demandas incertas no problema de alocação da frota de aviões às rotas aéreas para melhor atendimento da demanda. Ao longo dos anos, outros pesquisadores também destacaram esta importância.

Uma constatação interessante realizada por MASSÉ [40] para ilustrar o uso de estatística no tratamento do futuro faz referência ao fato de que uma pessoa interessada em saber se estará viva amanhã irá verificar suas condições atuais de saúde. Entretanto, a mesma pessoa recorrerá à estatística caso deseje saber se estará viva em 30 anos. Esta conclusão sugere que a utilização de modelos determinísticos sejam mais adequados à problemas de curto prazo enquanto problemas relacionados ao longo prazo devem ser tratados essencialmente por modelos estocásticos. Nas próximas seções, será realizada uma revisão da formulação de problemas estocástico multiestágio.

2.2 Formulação matemática

Em problemas estocásticos multiestágio considera-se que os dados, representados pelos vetores ξ_1, \dots, ξ_T , estão sujeitos à incertezas e são revelados gradualmente ao longo do tempo, em T estágios. Por sua vez, as decisões são representadas pelos vetores x_1, \dots, x_T e devem ser adaptadas a este processo. Sendo assim, o processo de decisão tem o seguinte aspecto

$$\begin{aligned} \text{decisão}(x_1) \rightsquigarrow \text{realização}(\xi_2) \rightsquigarrow \text{decisão}(x_2) \rightsquigarrow \\ \dots \rightsquigarrow \text{realização}(\xi_T) \rightsquigarrow \text{decisão}(x_T). \end{aligned}$$

A sequência de vetores $\xi_t \in \mathbb{R}^{d_t}$, $t = 1, \dots, T$ pode ser vista como um *processo estocástico*, ou seja, uma sequência de variáveis aleatórias com uma distribuição de probabilidade específica. Será utilizada a notação $\xi_{[t]} := (\xi_1, \xi_2, \dots, \xi_t)$ para denotar a história do processo até o estágio t .

Os valores do vetor de decisões x_t , definidos na etapa t , devem depender apenas da informação $\xi_{[t]}$ disponível até o estágio t , mas não dos resultados das observações futuras. Esta propriedade de causalidade é referida como *não-anticipatividade*. Como x_t baseia-se somente em $\xi_{[t]}$, a sequência de decisões também é um processo estocástico.

O processo estocástico é dito ter *independência temporal* quando ξ_t é estocasticamente independente de $\xi_{[t-1]}$, $t = 2, \dots, T$. Em particular, o processo é chamado de Markoviano se, para todo $t = 2, \dots, T$, a distribuição condicional de ξ_t dado $\xi_{[t-1]}$ é a mesma que a distribuição condicional de ξ_t dado ξ_{t-1} . Naturalmente, processos estocásticos com independência temporal são também Markovianos.

2.2.1 Formulação geral

Um problema estocástico genérico de T estágios pode ser escrito de acordo com a seguinte formulação aninhada

$$\min_{x_1 \in \mathcal{X}_1} f_1(x_1) + \mathbb{E} \left[\inf_{x_2 \in \mathcal{X}_2(x_1, \xi_2)} f_2(x_2, \xi_2) + \mathbb{E} \left[\dots + \mathbb{E} \left[\inf_{x_T \in \mathcal{X}_T(x_{T-1}, \xi_T)} f_T(x_T, \xi_T) \right] \right] \right] \quad (2.1)$$

voltada para o processo aleatório $\xi_1, \xi_2, \dots, \xi_T$. Nesta formulação, $x_t \in \mathbb{R}^{n_t}$, $t = 1, \dots, T$, são variáveis de decisão, $f_t : \mathbb{R}^{n_t} \times \mathbb{R}^{d_t} \mapsto \mathbb{R}$ são funções contínuas e $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \mathbb{R}^{d_t} \mapsto \mathbb{R}^{n_t}$, $t = 2, \dots, T$, são funções mensuráveis que representam o conjunto de soluções viáveis no estágio t . Os dados do primeiro estágio, dados por ξ_1 , pela função $f_1 : \mathbb{R}^{n_1} \mapsto \mathbb{R}$ e pelo conjunto $\mathcal{X}_1 \subset \mathbb{R}^{n_1}$ são determinísticos.

Uma maneira de tornar esta formulação mais precisa é considerar as variáveis $x_t = \mathbf{x}_t(\xi_{[t]})$, $t = 1, \dots, T$ como funções do processo estocástico $\xi_{[t]}$ até o estágio

t . Esta sequência de mapeamentos $\mathbf{x}_t : \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_t} \mapsto \mathbb{R}^{n_t}$ é chamada de uma *política implementável* (ou simplesmente uma política). Uma política é dita viável se ela satisfaz as restrições de viabilidade, ou seja,

$$\mathbf{x}_t(\xi_{[t]}) \in \mathcal{X}_t(\mathbf{x}_t(\xi_{[t-1]}), \xi_t), t = 2, \dots, T, \text{ com probabilidade } 1. \quad (2.2)$$

Note que o processo de otimização em (2.2) é executado através de políticas implementáveis e viáveis e que $\mathbf{x}_2, \dots, \mathbf{x}_T$ são funções do processo estocástico e x_1 é um vetor determinístico. Na maioria das aplicações nas quais deseja-se solucionar o problema numericamente, assume-se a hipótese que o processo ξ_1, \dots, ξ_T pode ser corretamente aproximado por amostras, com um número finito de realizações. Caso contrário, a formulação (2.2) levaria a um problema de otimização de dimensão infinita.

Um modo bastante comum de representar este problema, como será abordado neste trabalho, é através das equações de BELLMAN [12], também conhecidas como *equações de programação dinâmica*. O objetivo da utilização das equações de Bellman é descrever o valor das decisões em um determinado estágio como um problema que é uma função de decisões iniciais e determinam, como consequência, as decisões para o estágio corrente. A equação de Bellman é derivada do princípio de otimalidade também formulado por Bellman: “De qualquer ponto de uma trajetória ótima, a trajetória remanescente é ótima para o problema correspondente que se inicia naquele ponto e termina no mesmo ponto do problema anterior”. Esta forma permite decompor o problema em subproblemas menores e explorar métodos de solução eficientes, como será abordado neste trabalho.

Considere então o problema associado ao último estágio T ,

$$Q_T(x_{T-1}, \xi_T) = \min f_T(x_T, \xi_T) \quad (2.3a)$$

$$\text{s.a.: } x_T \in \mathcal{X}_T(x_{T-1}, \xi_T). \quad (2.3b)$$

O valor ótimo deste problema, denotado por $Q_T(x_{T-1}, \xi_T)$, depende do vetor de decisões x_{T-1} e dos dados aleatórios ξ_T . Para os estágios $t = 2, \dots, T-1$ o problema é formulado como

$$Q_t(x_{t-1}, \xi_{[t]}) = \min f_t(x_t, \xi_t) + \mathbb{E} [Q_{t+1}(x_t, \xi_{[t+1]}) | \xi_{[t]}] \quad (2.4a)$$

$$\text{s.a.: } x_t \in \mathcal{X}_t(x_{t-1}, \xi_t). \quad (2.4b)$$

onde $\mathbb{E} [\cdot | \xi_{[t]}]$ denota o valor esperado condicional. O valor ótimo deste problema, depende da decisão x_{t-1} do estágio anterior e da realização do processo $\xi_{[t]}$, e é denotado por $Q_t(x_{t-1}, \xi_{[t]})$. A idéia básica da programação dinâmica é calcular as funções de acoplamento entre os estágios, $Q_t(x_{t-1}, \xi_{[t]})$, também conhecidas como

funções de custo futuro (FCF), recursivamente, percorrendo os estágios no sentido inverso.

Finalmente, no primeiro estágio tem-se o seguinte problema

$$\min f_1(x_1) + \mathbb{E}[Q_2(x_1, \xi_2)] \quad (2.5a)$$

$$\text{s.a.: } x_1 \in \mathcal{X}_1. \quad (2.5b)$$

As equações recursivas de programação dinâmica correspondentes à este problema são definidas como

$$Q_t(x_{t-1}, \xi_{[t]}) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} \{f_t(x_t, \xi_t) + Q_{t+1}(x_t, \xi_{[t]})\}, \quad (2.6)$$

onde $Q_{t+1}(x_t, \xi_{[t]}) := \mathbb{E}[Q_{t+1}(x_t, \xi_{[t+1]}) | \xi_{[t]}]$.

Se o processo ξ_1, \dots, ξ_T é Markoviano, então as distribuições condicionais nas equações acima, dado $\xi_{[t]}$, são as mesmas que as respectivas distribuições condicionais dado ξ_t . Neste caso, a função de custo futuro depende apenas de ξ_t , ao invés de depender de toda a história do processo $\xi_{[t]}$ e pode ser escrita como $Q_t(x_{t-1}, \xi_t)$. Se, além disto, houver a condição de independência temporal, então a função valor esperado Q_t não depende das realizações do processo aleatório e pode ser escrita simplesmente como $Q_t(x_{t-1})$.

2.2.2 Formulação linear

Um caso especial do problema (2.1) é o problema multiestágio *linear* estocástico, em que as funções objetivo e as restrições são lineares. Em uma formulação linear típica tem-se que

$$\begin{aligned} f_t(x_t, \xi_t) &:= c_t^\top x_t, \quad \mathcal{X}_1 := \{x_1 : A_1 x_1 = b_1, x_1 \geq 0\}, \\ \mathcal{X}_t(x_{t-1}, \xi_t) &:= \{x_t : B_{t-1} x_{t-1} + A_t x_t = b_t, x_t \geq 0\}, \quad t = 2, \dots, T. \end{aligned}$$

Os componentes dos vetores c_t e b_t e das matrizes A_t e B_T são modelados como variáveis aleatórias¹ e formam o processo estocástico $\xi_t = (c_t, A_t, B_t, b_t)$, $t = 2, \dots, T$ enquanto é considerado que $\xi_1 = (c_1, A_1, b_1)$ possui componentes com valores conhecidos (determinísticos).

¹Obviamente, nem todos os componentes de ξ_t precisam ser aleatórios. Tipicamente, temos que só o vetor de recursos b_t possui componentes aleatórias enquanto os demais componentes de ξ_t são conhecidos.

Desta forma, temos que o problema para o último estágio T pode ser escrito como

$$Q_T(x_{T-1}, \xi_T) = \min c_T^\top x_T \quad (2.7a)$$

$$\text{s.a.: } B_T x_{T-1} + A_T x_T = b_T \quad (2.7b)$$

$$x_T \geq 0. \quad (2.7c)$$

A solução ótima do problema (2.7) é dada por $Q_T(x_{T-1}, \xi_T)$ e depende do vetor de decisões do estágio anterior $x_{T-1} \in \mathbb{R}^{n_{T-1}}$ e dos dados $\xi_T = (c_T, B_T, A_T, b_T)$.

Para o estágio $T-1$ temos x_{T-2} e $\xi_{[T-1]}$ são conhecidos. Entretanto, como ainda não são conhecidas as realizações das incertezas para o estágio seguinte, o problema para este estágio é definido pelo seguinte problema de otimização estocástica

$$Q_{T-1}(x_{T-2}, \xi_{[T-1]}) = \min c_{T-1}^\top x_{T-1} + \mathbb{E} [Q_T(x_{T-1}, \xi_T) | \xi_{[T-1]}] \quad (2.8a)$$

$$\text{s.a.: } B_{T-1} x_{T-2} + A_{T-1} x_{T-1} = b_{T-1} \quad (2.8b)$$

$$x_{T-1} \geq 0. \quad (2.8c)$$

A solução ótima do problema acima depende das variáveis de decisão $x_{T-2} \in \mathbb{R}^{n_{T-2}}$ e dos dados $\xi_{[T-1]}$ e é denotada por $Q_{T-1}(x_{T-2}, \xi_{[T-1]})$. Para um estágio qualquer $t = 2, \dots, T-1$, tem-se o seguinte problema

$$Q_t(x_{t-1}, \xi_{[t]}) = \min c_t^\top x_t + Q_{t+1}(x_t, \xi_{[t]}) \quad (2.9a)$$

$$\text{s.a.: } B_t x_{t-1} + A_t x_t = b_t \quad (2.9b)$$

$$x_t \geq 0. \quad (2.9c)$$

onde

$$Q_{t+1}(x_t, \xi_{[t]}) := \mathbb{E} [Q_{t+1}(x_t, \xi_{t+1}) | \xi_{[t]}] \quad (2.10)$$

é função de custo que representa o futuro a partir de $t+1$.

Finalmente, tem-se o problema para o primeiro estágio onde são definidas as variáveis de decisão $x_1 \in \mathbb{R}^{n_1}$

$$\min c_1^\top x_1 + Q_2(x_2) \quad (2.11a)$$

$$\text{s.a.: } A_1 x_1 = b_1 \quad (2.11b)$$

$$x_1 \geq 0 \quad (2.11c)$$

onde $Q_2(x_2) := \mathbb{E} [Q_2(x_1, \xi_2)]$.

Note que todos os estágios subsequentes $t = 2, \dots, T-1$ estão incorporados e representados através do valor esperado da função $Q_2(x_2)$ e que esta função não

depende de ξ_1 pois este é não aleatório.

Pode-se então apresentar a *formulação aninhada* para o problema de otimização estocástico multiestágio linear como

$$\min_{\substack{A_1 x_1 = b_1 \\ x_1 \geq 0}} c_1^\top x_1 + \mathbb{E} \left[\inf_{\substack{B_2 x_1 + A_2 x_2 = b_2 \\ x_2 \geq 0}} c_2^\top x_2 + \mathbb{E} \left[\dots + \mathbb{E} \left[\inf_{\substack{B_T x_{T-1} + A_T x_T = b_T \\ x_T \geq 0}} c_T^\top x_T \right] \right] \right] \quad (2.12)$$

Finalmente, as equações recursivas de programação dinâmica para o caso linear são definidas como

$$\mathcal{Q}_t(x_{t-1}, \xi_{[t]}) = \inf_{\substack{B_t x_{t-1} + A_t x_t = b_t \\ x_t \geq 0}} \{c_t^\top x_t + \mathcal{Q}_{t+1}(x_t, \xi_{[t]})\}. \quad (2.13)$$

2.3 Representação das incertezas

2.3.1 Distribuição de probabilidade

A formulação (2.13) assume que o processo estocástico representado por ξ_t , $t = 1, \dots, T$, pode ser modelado como um vetor aleatório com distribuição de probabilidade conhecida (e não simplesmente aleatório). Esta hipótese é justificável em muitas situações.

Por exemplo, o comportamento de ξ_t pode ser derivado a partir dos dados históricos considerando-se que a distribuição de probabilidade não varia significativamente ao longo do período de tempo considerado. Nestes casos, é possível estimar com relativa confiança esta distribuição de probabilidade e a otimização do valor médio é justificável pela Lei dos Grandes Números.

Outro exemplo é que os valores de ξ_t podem ser realizações de um modelo de simulação cujos resultados são estocásticos. A distribuição de probabilidade empírica desta amostra pode ser utilizada como uma aproximação da distribuição real, porém desconhecida.

2.3.2 Árvore de cenários

Para solucionar problemas estocásticos numericamente, muitas vezes é necessário aproximar o vetor aleatório ξ_t e *discretizá-lo* em um número finito de possíveis realizações. Em problemas multiestágio, considera-se que o processo aleatório e a sequência de realizações dos dados, quando discretizado, assume a forma de uma *árvore de cenários*.

Esta árvore tem *nós* que são organizado em níveis que correspondem aos estágios $\mathbb{T} = \{1, \dots, T\}$. No nível $t = 1$ existe apenas um nó, denominado como *nó raiz*, e é associado à ele o valor de ξ_1 (que é considerado conhecido para $t = 1$). No nível

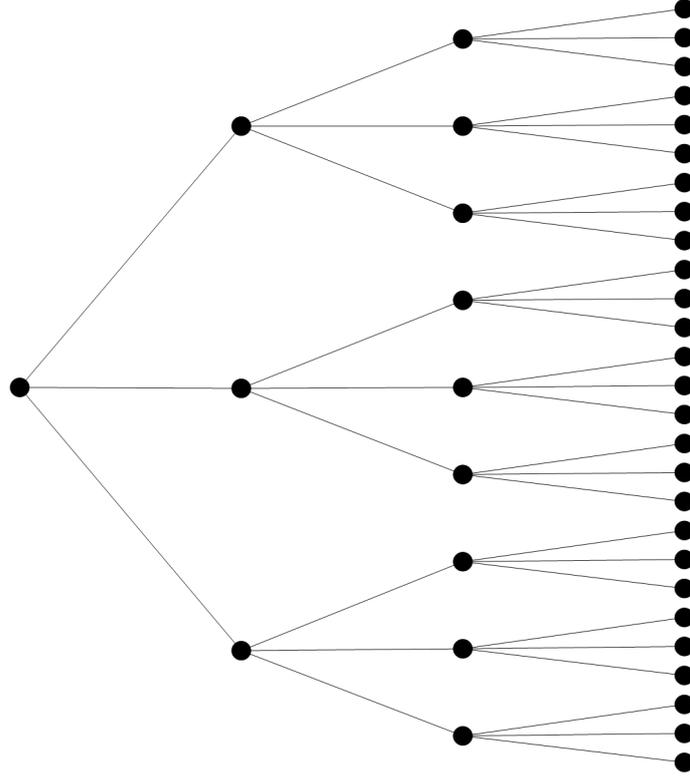


Figura 2.1: Árvore de cenários: exemplo de 4 níveis (estágios) com 3 aberturas em cada nível, totalizando 27 cenários.

$t = 2$ existem quantos nós quantas forem as possíveis realizações de ξ_2 . Cada um destes nós é conectado ao nó raiz por um arco. Para cada nó n do nível $t = 2$ (que corresponde à uma realização particular ξ_2^n de ξ_2 são criados pelo menos tantos nós para o nível 3 quantos forem os diferentes valores de ξ_3 que possam ocorrer a partir de ξ_2^n e estes nós são conectados com o nó n , e assim por diante. A figura 2.1 exemplifica uma árvore de cenários.

Em geral, os nós de um determinado nível t correspondem à possíveis valores de ξ_t que podem ocorrer. Cada um destes nós está conectado a um único nó no nível $t - 1$, chamado de *nó ancestral*, que corresponde às primeiras $t - 1$ partes idênticas do processo $\xi_{[t]}$. Além disto, este nó também está conectado à nós no nível $t + 1$, chamados de *nós filhos*, que correspondem à possíveis continuações da história de $\xi_{[t]}$. Em geral, as incertezas ξ_t^n são vetores e é possível que algumas destas realizações sejam iguais em um determinado nível t . Mesmo nestes casos é importante ressaltar que estas realizações podem ser representadas por nós diferentes pois podem corresponder à diferentes histórias do processo estocástico.

Denota-se como Ω_t o conjunto de todos os nós do estágio $t \in \mathbb{T}$. Em particular, Ω_1 consiste em um único nó raiz, Ω_2 possui tantos elementos quantos forem as possíveis realizações de ξ_2 e assim por diante. Para um nó $n \in \Omega_t$ denota-se como $C_n \subset \Omega_{t+1}, t = 1, \dots, T - 1$ o conjunto de todos os nós filhos de n e como $a(n) \in$

$\Omega_{t-1}, t = 2, \dots, T$ o nó ancestral (nó pai) de n . Observe que $\Omega_{t+1} = \bigcup_{n \in \Omega_t} C_n$ e que os conjuntos C_n são disjuntos, ou seja, $C_n \cap C_{n'} = \emptyset$ se $n \neq n'$.

Um *cenário* é definido como um caminho entre o nó raiz no estágio $t = 1$ e um nó no último estágio T . Cada cenário representa uma história do processo ξ_1, \dots, ξ_T . Por construção, existe uma relação unívoca entre os cenários e o conjunto Ω_T e, desta forma, o número total de cenários \mathcal{L} é igual à cardinalidade do conjunto Ω_T , ou seja, $\mathcal{L} = |\Omega_T|$.

Uma vez definida a estrutura da árvore de cenários, é possível definir a sua distribuição de probabilidade. De forma a lidar com a estrutura aninhada do processo de decisão, é necessário especificar a distribuição de probabilidade condicional de ξ_{t+1} dado $\xi_{[t]}, t = 1, \dots, T - 1$. Define-se então como $\varrho_{n\nu}$ a probabilidade de o processo estocástico evoluir do nó $n \in \Omega_t$ para um nó $\nu \in C_n$. Note que $\varrho_{n\nu} \geq 0$ e que $\sum_{\nu \in C_n} \varrho_{n\nu} = 1$ e que as probabilidades $\varrho_{n\nu}$ têm uma relação unívoca e direta com os arcos da árvore de cenários. Estas probabilidades $\varrho_{n\nu}, \nu \in C_n$, representam a distribuição de probabilidade condicional de ξ_{t+1} dado o caminho do processo estocástico ξ_1, \dots, ξ_t que termina no nó n .

Cada cenário pode ser definido por sua sequência de nós n_1, \dots, n_T organizados em ordem cronológica, ou seja, n_2 está conectado ao nó raiz, por sua vez n_3 está conectado ao nó n_2 e assim por diante. A probabilidade de um dado cenário é dada pelo produto $\varrho_{n_1 n_2} \varrho_{n_2 n_3} \dots \varrho_{n_{T-1} n_T}$, ou seja, um conjunto de probabilidades condicionais definem a distribuição de probabilidade de um conjunto de cenários. De maneira análoga, também é possível definir probabilidades condicionais a partir da distribuição de probabilidade dos cenários $p_l, l = 1, \dots, \mathcal{L}$, conforme definido a seguir. Seja $\mathcal{S}^{(n)}$ o conjunto de cenários que passam pelo nó n , em um determinado nível t , da árvore de cenários e seja $p^{(n)} := \Pr[\mathcal{S}^{(n)}]$, ou seja, $p^{(n)}$ é definido pela soma das probabilidades de todos os cenários que passar pelo nó n . Outra maneira de se calcular a probabilidade de atingir o nó n é a partir da história do processo até o nó n . Se a sequência de nós para atingir um determinado nó $n = n_t$ é dada por n_1, n_2, \dots, n_t então a probabilidade de se atingir este nó é dada pelo produto das probabilidades condicionais $p^{(n)} = \varrho_{n_1 n_2} \varrho_{n_2 n_3} \dots \varrho_{n_{t-1} n_t}$.

Se o processo ξ_1, \dots, ξ_T possui independência temporal, então a distribuição de probabilidade condicional de ξ_{t+1} dado ξ_t é a mesma que a distribuição de probabilidade incondicional de ξ_{t+1} , para $t = 1, \dots, T - 1$. Neste caso, a cada estágio $t = 1, \dots, T - 1$, cada nó $n \in \Omega_t$ está associado ao um conjunto idêntico de nós filhos, com o mesmo conjunto de probabilidades condicionais e os mesmo valores numéricos.

Um dado processo estocástico $Z_t, t = 1, 2, \dots$, que pode assumir um número finito $\{z_1, \dots, z_L\}$ de L valores diferentes é considerado como uma *cadeia de Markov*

se

$$\Pr \{Z_{t+1} = z_j | Z_t = z_i, Z_{t-1} = z_{i_{t-1}}, \dots, Z_1 = z_{i_1}\} = \Pr \{Z_{t+1} = z_j | Z_t = z_i\}$$

para todos os estados $z_{i_{t-1}}, \dots, z_{i_1}, z_i, z_j$ e para todos os estgios $t = 1, 2, \dots$. Pode-se ento definir

$$p_{ij} := \Pr \{Z_{t+1} = z_j | Z_t = z_i\}, \quad i, j = 1, \dots, L.$$

Em algumas situaes,  natural modelar o processo estocstico como uma cadeia de Markov com um espao de estados correspondente $\{\zeta^1, \dots, \zeta^L\}$ e probabilidades p_{ij} de mover-se do estado ζ^i para o estado ζ^j , para $i, j = 1, \dots, L$. Este processo pode ser modelado como uma rvore de cenrios onde considera-se que no estgio $t = 1$ existe um no raiz para o qual  atribudo um dos valores do espao de estados como, por exemplo, ζ^i . No estgio $t = 2$ existem L nos para os quais so atribudos (vetores de) valores ζ^1, \dots, ζ^L com respectivas probabilidades p_{i1}, \dots, p_{iL} . No estgio $t = 3$, por sua vez, existem L^2 nos onde cada um desses nos est conectado um dos L nos ancestrais do estgio $t = 2$. Para os estgio $t = 4$ existem L^3 nos e assim por diante. De modo geral, para um determinado estgio t de um processo de Markov de T estgios existem L^{t-1} nos e o vetor de variveis aleatrias ξ_t pode assumir os valores de um dos estados ζ^1, \dots, ζ^L . As respectivas probabilidades podem ser calculadas a partir da histria do processo at o estgio t e o nmero total de cenrios  igual  L^{T-1} .

Na formulao da cadeia de Markov apresentada acima, a rvore de cenrios representa a histria completa do processo estocstico e o nmero de cenrios cresce exponencialmente com o nmero de estgios. Se abordarmos o problema decompondo-o em estgios atravs da representao da funo de custo futuro $Q_t(x_{t-1}, \xi_t)$,  possvel notar que no  necessrio guardar toda a histria do processo. Isso ocorre pois, para todo estgio t , a funo de custo futuro $Q_t(\cdot, \xi_t)$ depende apenas do estado (realizao) corrente $\xi_t = \zeta^i, i = 1, \dots, L$ do processo estocstico. Por outro lado, se deseja-se escrever o problema de otimizao correspondente (no caso de um nmero finito/tratvel de cenrios) como um grande problema de programao linear, a formulao em rvore ainda  necessria. Esta  a diferena bsica entre as abordagens de programao estocstica e programao dinmica para o problema.

A abordagem considerada pela programao estocstica no se baseia necessariamente na estrutura Markoviana do processo, tornando-a mais genrica ao custo de considerar, possivelmente, um nmero muito grande de cenrios.

2.3.3 Formulação em árvore

Uma vez definida a estrutura matemática para representação de uma árvore de incertezas, é possível representar o problema de programação estocástica multiestágio linear (2.1) através de uma formulação *em árvore* ou *arborescente*, de acordo com o problema (2.14).

$$\min \quad c_1^\top x_1 + \sum_{t=2}^T \sum_{n \in \Omega_t} p^{(n)} c_t^\top x_t^{(n)} \quad (2.14a)$$

$$\text{s.a.:} \quad A_1 x_1 = b_1 \quad (2.14b)$$

$$B_t x_{t-1}^{a(n)} + A_t x_t^{(n)} = b_t^{(n)} \quad n \in \Omega_t, \quad t = 2, \dots, T \quad (2.14c)$$

$$x_t^{(n)} \geq 0 \quad n \in \Omega_t, \quad t = 1, \dots, T \quad (2.14d)$$

onde $x_1 = x_1^{(n)}$ para $t = 1$, pois Ω_1 possui apenas o nó raiz.

Na formulação apresentada acima, considera-se que apenas b_t é aleatório. Por esta razão, ele é indexado de acordo com o nó da árvore de incertezas e é representado como $b_t^{(n)}$. Todavia, esta formulação pode ser facilmente estendida para considerar incertezas nas demais componentes de ξ_t .

A função objetivo (2.14a) representa diretamente os custos associados às decisões em cada nó da árvore de decisões, considerando que os custos associados aos nós dos estágios $t = 2, \dots, T$ devem ser ponderados pelas respectivas probabilidades de ocorrência. O conjunto viável do problema é definido pelo conjunto viável associado ao nó raiz, representado por (2.14b). Para os demais estágios, há uma dependência com relação às decisões do estágio anterior associadas ao nó antecessor correspondente, conforme indicado na equação (2.14c). Finalmente, as restrições de não negatividade (2.14d) são representadas para as variáveis de decisão associadas a todos os estágios e nós da árvore.

2.4 Programação Dinâmica Dual Estocástica

Como foi visto na seção anterior, para solucionar numericamente um problema de programação estocástica multiestágio é necessário recorrer à discretização do processo estocástico e a formulação resultante pode ser representada através de uma árvore de cenários.

Em teoria, é possível formular o problema como um grande problema de programação estocástica multiestágio de programação linear, também conhecido como equivalente determinístico [68] e solucioná-lo com algoritmos convencionais de programação linear. Entretanto, como comentado, devido às características de sua construção, o número de nós da árvore de cenários cresce exponencialmente com o

número de estágios e, mesmo para um problema de poucos estágios, torna-se difícil a solução do problema completo.

Esta limitação motivou o desenvolvimento de esquemas de solução com base em decomposição em espaço de estados. O algoritmo apresentado por PEREIRA e PINTO [48], conhecido como Programação Dinâmica Dual Estocástica, também baseia-se na decomposição do problema em estágio, assim como a Programação Dinâmica Estocástica.

Entretanto, PEREIRA e PINTO [48] combinaram de maneira elegante a hipótese de independência temporal do processo estocástico, a utilização da informação dual para construção de uma extrapolação da função de custo futuro e utilização de um processo de amostragem de cenários para criar o algoritmo da PDDE. O algoritmo PDDE busca pela solução ótima, iterativamente, explorando a árvore de cenários de maneira *implícita*, ou seja, sem a necessidade de percorrer todos os seus nós. Além disto, evita o problema da “maldição de dimensionalidade” pois utiliza a informação dual para extrapolar a função de custo futuro no entorno dos estados considerados “interessantes”. Estes estados, por sua vez, são gerados a partir de um processo de amostragem de cenários, o que permite estimar com a precisão desejada a qualidade da solução encontrada.

Atualmente, a PDDE é uma metodologia consolidada e considerada o estado-da-arte na solução de problemas de planejamento da operação hidrotérmica. É extensivamente utilizada em estudos de operação e centros de despacho em mais de 30 países em todo o mundo, nos cinco continentes, incluindo América do Sul, Centroamérica e América do Norte, Áustria, Espanha, Noruega, Turquia, Nova Zelândia e China. Em 2008, MACEIRA *et al.* [39] apresentaram a história de 10 anos de sucesso da metodologia no sistema brasileiro.

A próxima seção descreve a necessidade da hipótese de independência temporal para o algoritmo da PDDE e como a formulação pode ser alterada para considerar dependência temporal do processo estocástico. A seção seguinte descreve brevemente o processo de amostragem de cenários. Finalmente, é apresentado o algoritmo de PDDE baseado no problema de programação estocástica multiestágio linear, definido na seção 2.2.2.

2.4.1 Dependência temporal

De acordo com a seção 2.3.2, a solução numérica de um problema estocástico multiestágio implica na discretização do processo de incertezas, o que leva à representação de uma árvore de cenários. Uma estratégia de solução interessante para abordar este problema é a utilização de técnicas para decomposição do problema original em subproblemas menores, para cada um dos estágios da árvore, acoplados pela função de

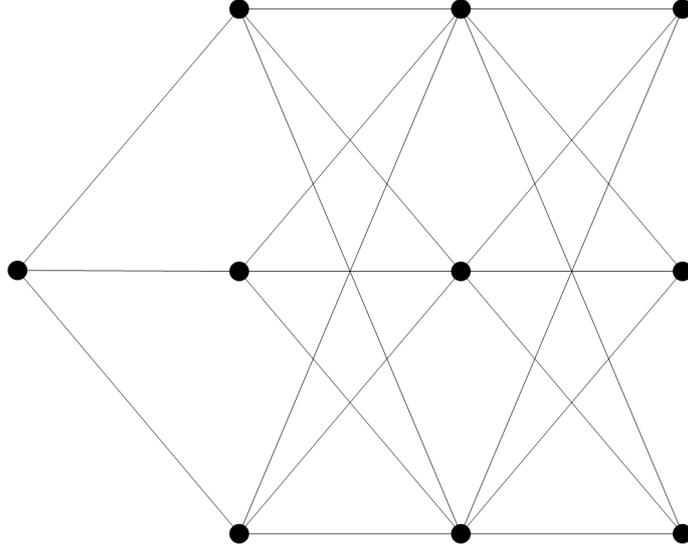


Figura 2.2: Árvore de cenários com independência temporal: exemplo de 4 níveis (estágios) com 3 aberturas em cada nível.

custo futuro.

No caso geral, considera-se o processo estocástico possui dependência temporal. Com isto, além de depender dos estados x_t , a FCF também depende do histórico de realizações das variáveis aleatórias $\xi_{[t]}$. Esta dependência exige que a FCF tenha que ser calculada de maneira condicional ao histórico do processo estocástico, como pode ser visto na equação (2.10). Como o número de realizações do processo (e da árvore de incertezas) cresce exponencialmente com o número de estágios, a representação da FCF torna-se uma difícil tarefa.

Por outro lado, considere o caso particular onde o processo estocástico tem a propriedade de *independência temporal*, ou seja, o vetor aleatório ξ_{t+1} não depende do processo $\xi_{[t]} = (\xi_1, \xi_2, \dots, \xi_t)$, para $t = 1, \dots, T - 1$. Neste caso, o cálculo da FCF torna-se bastante mais simples pois a mesma deixa de ser expressa em função dos estados anteriores e do histórico do processo estocástico $\mathcal{Q}_{t+1}(x_t, \xi_{[t]})$ e pode ser expressa apenas como $\mathcal{Q}_{t+1}(x_t)$. Isto permite que uma única função, calculada para um determinado estágio t , seja *compartilhada* por todos os nós da árvore de incertezas deste estágio.

Na prática, esta hipótese permite *colapsar* a árvore de cenários e isto implica que o número de nós não cresce exponencialmente com o número de estágios. Para cada estágio, existem tantos nós quantos forem os números de realizações L , conforme ilustrado na figura 2.2. Como assumimos que o processo tem independência temporal, a FCF pode ser calculada visitando-se apenas os poucos nós existentes na árvore de cenários a cada estágio.

Em muitos casos, entretanto, é necessária a representação da dependência tem-

poral do processo estocástico em questão. PEREIRA e PINTO [48] perceberam que a dependência temporal poderia ser representada de forma *indireta* em um ambiente de programação estocástica multiestágio linear considerando, por exemplo, a utilização de modelos auto-regressivos periódicos [45]. Em particular, considere o caso em que (c_t, B_t, A_t) tem independência temporal e que b_t tem dependência temporal. Para um modelo auto regressivo de primeira ordem, AR(1), temos que

$$b_t = \Phi b_{t-1} + \epsilon_t, \quad (2.15)$$

onde Φ é uma matriz de acoplamento temporal apropriada.

Considerando esta dependência temporal explicitamente na formulação, o conjunto viável do problema (2.12) pode ser reescrito como

$$b_t - \Phi b_{t-1} = \epsilon_t \quad (2.16a)$$

$$B_t x_{t-1} - \Phi b_{t-1} + A_t x_t = \epsilon_t \quad (2.16b)$$

$$x_t \geq 0 \quad (2.16c)$$

para $t = 2, \dots, T$. Desta forma, o parâmetro incerto e independente temporalmente passa a ser ϵ_t . Se substituirmos x_t por (x_t, b_t) e o processo estocástico por $\xi_t = (c_t, B_t, A_t, \epsilon_t)$, $t = 2, \dots, T$ é possível voltar à formulação onde o processo tem independência temporal. Nesta formulação, a dependência temporal é considerada de maneira *implícita*.

Esta reformulação é fundamental para o algoritmo da PDDE. Como será visto adiante, a PDDE é uma metodologia de decomposição que baseia-se na construção de uma FCF representada por hiperplanos cortantes e esta característica de consideração da dependência temporal através de modelos autorregressivos, ou de forma mais geral modelos lineares, permite o compartilhamento dos cortes que representam a FCF. O trabalho de DE QUEIROZ e MORTON [26] apresenta de forma detalhada o compartilhamento de cortes no ambiente de decomposição de problemas de programação dinâmica estocástica multiestágio.

Considere que, a cada estágio t , é realizada uma aproximação das incertezas através do sorteio de L amostras do vetor de variáveis aleatórias ξ_t , $t = 2, \dots, T$. Desta forma, temos que as incertezas passam a ser representadas por uma aproximação discreta dada por

$$\tilde{\xi}_t^l = (\tilde{c}_t^l, \tilde{B}_t^l, \tilde{A}_t^l, \tilde{b}_t^l), \quad l = 1, \dots, L \quad (2.17)$$

Este processo e sua distribuição de probabilidade podem ser representados por uma árvore onde, em um dado estágio $t - 1$, todos os nós da árvore tem as *mesmas ramificações* (aberturas) correspondendo à $\tilde{\xi}_t^1, \tilde{\xi}_t^2, \dots, \tilde{\xi}_t^L$. Note que esta característica ocorre devido à hipótese de independência temporal do processo.

O número total de cenários desta árvore é dado pelas diferentes combinações de realizações que representam diferentes histórias do processo estocástico. Neste sentido, ainda existe um número bastante elevado de cenários, que pode ser calculado como:

$$\mathcal{L} = L^{T-1}. \quad (2.18)$$

É importante notar que, mesmo para um número pequeno de aberturas L a cada estágio t , o número de cenários pode ser intratável. Por exemplo, para um caso com $T = 60$ estágios e $L = 2$ aberturas a cada estágio temos que o número de cenários é da ordem de 10^{18} . Na próxima seção, será exibida a técnica de amostragem de cenários utilizada no algoritmo de PDDE para contornar este problema.

2.4.2 Amostragem por Monte Carlo e *Sample Average Approximation*

Uma abordagem comum tratar um problema e reduzir o conjunto de cenários para um tamanho “tratável” é através do uso da técnica de simulação Monte Carlo [38]. Suponha que o número total de cenários é muito grande ou até mesmo infinito e considere também que é possível gerar uma amostra ξ^1, \dots, ξ^S de tamanho S do vetor de variáveis aleatórias ξ . Em geral, também se assume que a amostra é independente e identicamente distribuída.

Para o exemplo da função do valor esperado, seu valor “real” pode ser estimado a partir da média da amostra (“*sample average*”). Logo, se

$$\mathcal{Q}_{t+1}(x_t, \xi_{[t]}) := \mathbb{E}[\mathcal{Q}_{t+1}(x_t, \xi_{t+1})], \quad (2.19)$$

então temos que a aproximação da média de $\mathcal{Q}_{t+1}(x_t, \xi_{[t]})$ é dada por

$$\hat{\mathcal{Q}}_{t+1}^S(x_t, \xi_{[t]}) = \frac{1}{S} \sum_{s=1}^S \mathcal{Q}_{t+1}(x_t, \xi_{t+1}^s). \quad (2.20)$$

A formulação de um problema de otimização considerando uma amostra de cenários é conhecida por “*Sample Average Approximation*” (SAA). O problema obtido pela SAA é uma função do número considerado de amostras e, neste sentido, pode ser considerado que também é aleatório. Dada uma amostra de tamanho S , o problema SAA pode ser formulado com a mesma estrutura do problema obtido pela discretização de cenários com probabilidade $p_s = \frac{1}{S}$.

A extensão para o caso multiestágio pode ser obtida de maneira natural, contruindo-se cada um dos cenários amostrados a partir de uma sequência de sorteios do vetor de variáveis aleatórias em cada estágio.

2.4.3 Formulação matemática

O algoritmo de PDDE utiliza a técnica de decomposição de Benders para separar o problema original em problemas de um estágio e iterativamente construir aproximações das funções de custo futuro como funções lineares por partes. Esta metodologia torna dinâmico o processo de discretização do espaço de estados e possibilita a geração de limites para a solução ótima do problema, eliminando assim todas as desvantagens associadas aos métodos de PDE.

Considere o problema estocástico multiestágio linear (2.12), assumindo a hipótese de independência temporal apresentada na seção 2.4.1.

$$\tilde{Q}_t^l(x_{t-1}) = \min \quad \tilde{c}_t^{l\top} x_t + \tilde{Q}_{t+1}(x_t) \quad (2.21a)$$

$$\text{s.a. :} \quad \tilde{B}_t^l x_{t-1} + \tilde{A}_t^l x_t = \tilde{b}_t^l, \quad \pi_t^l \quad (2.21b)$$

$$x_t \geq 0 \quad (2.21c)$$

onde, neste caso,

$$\tilde{Q}_{t+1}(x_t) := \sum_{l=1}^L p_l \tilde{Q}_{t+1}^l(x_t), \quad t = T, \dots, 2 \quad (2.22)$$

onde $\tilde{Q}_{t+1}^l(x_t)$ é o problema associado à realização $\tilde{\xi}_t^l$, $p_t^l = \Pr(\xi_t^l)$ é a probabilidade de ocorrência de $\tilde{\xi}_t^l$ e π_t^l é a variável dual associada à restrição (2.21b). Por definição, assume-se que $\tilde{Q}_{T+1}(\cdot) = 0$.

Como, de acordo com (2.17), o número de realizações das incertezas é finito e os problemas são convexos, é possível demonstrar que as funções de custo futuro $\tilde{Q}_t(\cdot)$ são funções lineares por partes. Sendo assim, é possível representar esta função por um conjunto de hiperplanos cortantes, também conhecidos como cortes de Benders.

Seja x_t^{l*} uma solução ótima primal e π_t^{l*} uma solução ótima dual para $\tilde{Q}_t^l(\bar{x}_{t-1})$ para $l = 1, \dots, L$, é possível calcular um hiperplano suporte de $\tilde{Q}_t(\cdot)$ no ponto \bar{x}_{t-1} como

$$\partial \tilde{Q}_t(x_{t-1}) := \tilde{Q}_t(\bar{x}_{t-1}) + \tilde{g}_t^\top (x_{t-1} - \bar{x}_{t-1}) \quad (2.23)$$

onde

$$\tilde{Q}_t(\bar{x}_{t-1}) = \sum_{l=1}^L p_t^l \tilde{c}_t^l x_t^{l*} \quad \text{e} \quad \tilde{g}_t^\top = - \sum_{l=1}^L p_t^l \tilde{B}_t^{l\top} \pi_t^{l*}. \quad (2.24)$$

Sendo assim, é possível reescrever o problema (2.21) de modo equivalente, substituindo a função de custo futuro $\tilde{Q}_{t+1}(x_t)$ por uma representação linear por partes

dada pelo conjunto de hiperplanos definidos em (2.23).

$$\tilde{Q}_t^l(x_{t-1}) = \min \quad \tilde{c}_t^{l\top} x_t + \theta_{t+1} \quad (2.25a)$$

$$\text{s.a. :} \quad \tilde{B}_t^l x_{t-1} + \tilde{A}_t^l x_t = \tilde{b}_t^l, \quad \pi_t^l \quad (2.25b)$$

$$\theta_{t+1} \geq \tilde{Q}_{t+1}^p(x_t^p) + \tilde{g}_{t+1}^{p\top}(x_t - x_t^p), \quad p \in \bar{P}_{t+1} \quad (2.25c)$$

$$x_t \geq 0 \quad (2.25d)$$

No problema (2.25), \bar{P}_{t+1} é o conjunto de todos os hiperplanos que definem a função $\tilde{Q}_{t+1}(x_t)$. Como, na solução ótima, apenas alguns destes hiperplanos estarão ativos não é necessário representar todos os cortes do conjunto \bar{P}_{t+1} . O algoritmo de PDDE utiliza esta característica em seu favor e baseia-se na relaxação das restrições (2.25c) e aplica um esquema iterativo, onde os problemas são solucionados recursivamente, de trás para frente, considerando a melhor aproximação da função $\tilde{Q}_{t+1}(x_t)$ até a iteração corrente. A cada iteração, são calculados novos cortes que são utilizados para melhorar a representação da função de custo futuro.

2.4.4 Algoritmo

O algoritmo da PDDE é composto por duas fases: *simulação forward* e *recursão backward*, descritas a seguir.

Na fase de simulação forward, é utilizada a técnica de Monte Carlo para amostrar um subconjunto $\mathcal{M} = \{1, \dots, M\}$ de *cenários forward* dentre os possíveis cenários \mathcal{L} . Para cada cenário, os problemas são solucionados sucessivamente no sentido correto do tempo, $t = 1, \dots, T$, utilizando uma aproximação da função de custo futuro a cada estágio. Durante este processo, são armazenadas as soluções x_t^m , $m \in \mathcal{M}$. Nesta fase, obtém-se uma solução que atende às restrições de viabilidade (2.25b) para todas as etapas e, portanto, consiste em uma política operativa viável.

A figura 2.3 ilustra a amostragem de cenários utilizada na fase *forward* aplicada à um exemplo de árvore de cenários. Em azul, destacam-se os nós e arcos da árvore correspondentes aos cenários amostrados.

Na fase de recursão backward os problemas de cada etapa são solucionado de trás para frente, $t = T, \dots, 2$, para cada uma das amostras do subconjunto \mathcal{M} e para cada uma das L aberturas considerando como soluções iniciais os valores armazenados x_{t-1}^m na fase simulação forward. A cada estágio t e cenário forward m é calculado um corte de Benders que é, então, adicionado à representação da função de custo futuro do estágio t .

A figura 2.4 ilustra os problemas que são solucionados para o cálculo de novos cortes da FCF no penúltimo estágio. Em verde, destacam-se os nós e arcos do último estágio, associados aos nós filhos dos nós amostrados na fase *forward* para o

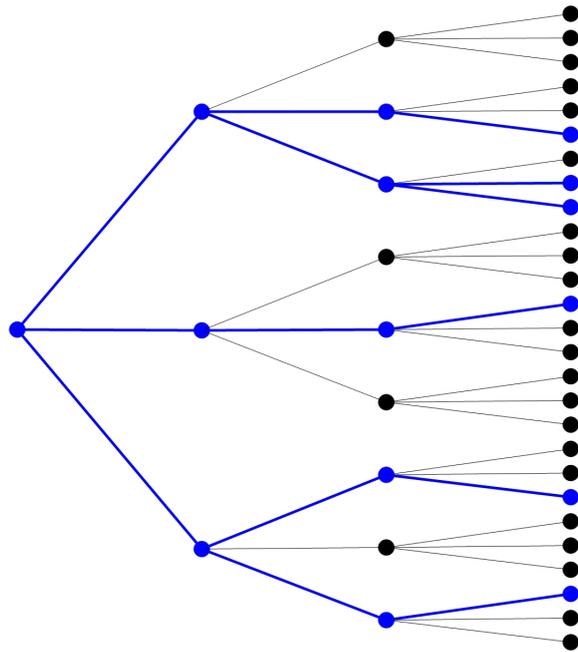


Figura 2.3: Amostragem de cenários na fase *forward* do algoritmo de PDDE

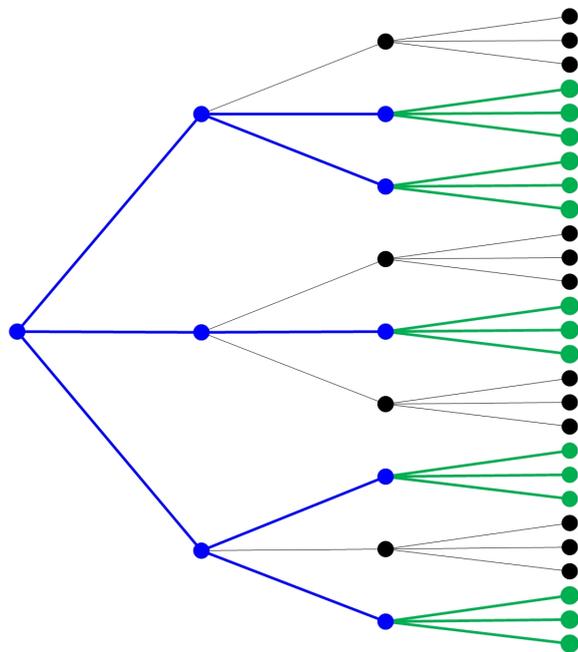


Figura 2.4: Problemas solucionados para o cálculo dos cortes da FCF do penúltimo estágio na fase *backward* do algoritmo de PDDE

penúltimo estágio (em azul).

O algoritmo de PDDE pode ser descrito pelos seguintes passos:

1. Inicialize o número de iterações $k = 0$ e o conjunto de cortes $P = \{\}$;
2. Simulação Forward

Para cada $t = 1, \dots, T$ e $m = 1, \dots, M$:

Resolva o problema associado a este estágio t e cenário forward m :

$$\tilde{Q}_t(x_{t-1}) = \min \tilde{c}_t^{m\top} x_t + \theta_{t+1} \quad (2.26a)$$

$$\text{s.a. : } \tilde{A}_t^m x_t = \tilde{b}_t^m - \tilde{B}_t^m x_{t-1} \quad (2.26b)$$

$$\theta_{t+1} \geq \tilde{Q}_{t+1}^p(x_t^p) + \tilde{g}_{t+1}^{p\top} (x_t - x_t^p), \quad p \in P \quad (2.26c)$$

$$x_t \geq 0 \quad (2.26d)$$

Armazene o vetor de soluções ótimas em $x_t^m = x_t$.

3. Teste de convergência

Calcule o limite superior do custo operativo total para a solução encontrada na simulação forward como o valor esperado do custo total de cada cenário m :

$$\bar{z} = \frac{1}{M} \sum_{m=1}^M z_m = \frac{1}{M} \sum_{m=1}^M \sum_{t=1}^T \tilde{c}_t^{m\top} x_t \quad (2.27)$$

e seu desvio padrão:

$$\bar{\sigma} = \sqrt{\frac{1}{M-1} \sum_{m=1}^M (z_m - \bar{z})^2} \quad (2.28)$$

Obtenha o limite inferior \underline{z} do custo operativo total como a solução do problema aproximado do primeiro estágio $\tilde{Q}_1(x_0)$.

Se o limite inferior \underline{z} pertencer a um intervalo de confiança especificado ϵ centrado no limite superior \bar{z} :

$$\bar{z} - \frac{\epsilon}{\sqrt{M}} \bar{\sigma} < \underline{z} < \bar{z} + \frac{\epsilon}{\sqrt{M}} \bar{\sigma} \quad (2.29)$$

pare;

caso contrário, siga para o passo 4.

4. Recursão Backward

Para cada $t = T, \dots, 2$, $m = 1, \dots, M$ e $l = 1, \dots, L$

Resolva o problema (2.25) considerando como ponto inicial a solução ótima x_{t-1}^m calculada na simulação forward;

Calcule um corte de Benders (2.23) associado ao problema solucionado e adicione-o à função de custo futuro da etapa $t - 1$; incremente o conjunto de cortes P .

5. Faça $k \leftarrow k + 1$ e volte ao passo 2.

Capítulo 3

Planejamento da operação de sistemas hidrotérmicos

O objetivo do planejamento ótimo da operação de sistemas hidrotérmicos é determinar a programação da geração para cada usina hidrelétrica e térmica de forma a minimizar o valor esperado dos custos de operação necessários para o atendimento da demanda ao longo de um dado horizonte, garantindo também a confiabilidade do suprimento. Tipicamente, os custos de operação incluem custos de combustível térmico, custos de importação de energia de sistemas vizinhos e penalidades por falhas no fornecimento de energia [46].

O problema em geral tem horizonte de médio a longo prazo com intervalo de decisões semanais ou mensais¹. Como mencionado anteriormente, este horizonte requer que sejam consideradas incertezas no problema, em especial quanto às vazões afluentes às usinas hidrelétricas. A seguir, será apresentada uma formulação básica para um problema de despacho hidrelétrico onde serão modelados apenas os aspectos necessários para introduzi-lo no contexto da formulação de problemas de programação estocástica multiestágio.

3.1 Usinas termelétricas

As usinas termoelétricas produzem energia através da conversão de energia térmica em energia elétrica. Em geral, é utilizado o vapor como força motriz: a água é aquecida, se transforma em vapor e sua pressão gira uma turbina a vapor que aciona um gerador elétrico. A maior variação no desenho das centrais térmicas é devido aos diferentes recursos que são utilizados para aquecer a água. Dentre elas, podemos citar as usinas que utilizam materiais fósseis (carvão, óleo combustível, gás natural) e as nucleares (que utilizam combustíveis físséis como urânio natural ou enriquecido).

¹No caso do sistema brasileiro, o Operador Nacional do Sistema (ONS) realiza estudos mensais com horizonte de 10 anos.

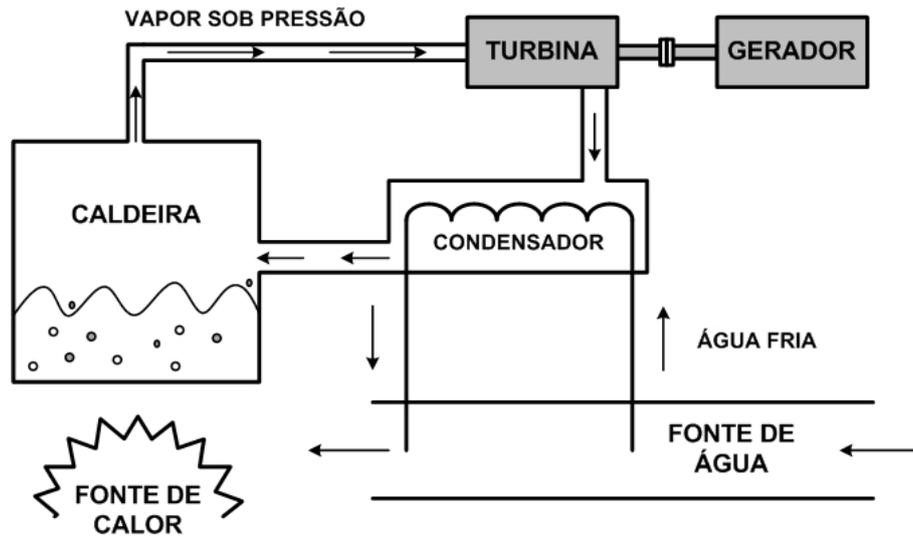


Figura 3.1: Esquema simplificado de geração de energia de uma usina térmica (a vapor)

Apesar do uso de diferentes combustíveis, é possível uma representação esquemática simplificada das usinas que usam a queima de matéria-prima para a obtenção de vapor, como apresentado na figura 3.1. Além das usinas à vapor, também existem usinas térmicas baseadas em outros princípios como, por exemplo, turbinas e motores à gás, entre outras.

Qualquer seja o tipo de usina termoelétrica, sua representação nos estudos de planejamento da operação é feita através de suas características físicas e restrições operativas, tais como potência máxima, combustível utilizado, consumo específico, taxa de tomada de carga e nível mínimo operativo. Dependendo do horizonte de planejamento, algumas características se tornam menos relevantes e podem ser aproximadas com o objetivo de simplificar sua representação matemática².

3.1.1 Custo de produção

Para o planejamento da operação de médio e longo prazo, conforme considerado nesse trabalho, as usinas termoelétricas serão representadas pelo seu custo operativo unitário direto, que depende basicamente do combustível utilizado, e é representado por $c_j, j = 1, \dots, J$ (em \$/MWh). Sendo assim, o custo operativo térmico a cada

²Para a horizontes de curto prazo, é necessária uma representação mais detalhada que leve em consideração restrições dinâmicas para as centrais térmicas tais como tempos mínimos de operação e desligamento, máximo intervalo para a variação da produção entre duas horas sucessivas, máximo número de arranques num certo intervalo de tempo e possibilidade de decisões tipo *commitment*.

estágio é definido pela seguinte expressão:

$$\sum_{j=1}^J c_j \times g_{t,j}, \quad t \in \mathbb{T}, \quad (3.1)$$

onde j é o índice das usinas térmicas, J é o número total de usinas térmicas e $g_{t,j}$ é variável de decisão que define a energia produzida pela usina j no estágio t , $\mathbb{T} = 1, \dots, T$ é o conjunto de etapas e T representa a última etapa do horizonte de planejamento.

3.1.2 Limites operativo

O limite operativo das usinas térmicas podem ser decorrentes de diferentes fatores, ligados às características físicas das usinas como manutenção da estabilidade do ciclo de termodinâmica, restrições associadas à disponibilidade de combustível ou, até mesmo, restrições associadas à limites elétricos ou energéticos. Em geral, este limite não depende de outras usinas ou da estratégia de operação da usina e pode ser determinado diretamente através da desigualdade (3.2),

$$g_{t,j} \leq \bar{g}_j, \quad j \in \mathbb{J}, \quad (3.2)$$

onde \bar{g}_j é o limite máximo de geração da usina j e \mathbb{J} é o conjunto de usinas termelétricas.

3.2 Usinas hidrelétricas

Um aproveitamento hidráulico para geração de energia elétrica é composto basicamente por: uma barragem formadora de um reservatório que represa um curso d'água e condutos forçados que levam a água do reservatório até a casa de força, situada em um nível mais baixo; a casa de força, onde estão instalados os grupos turbina-gerador e outros equipamentos auxiliares; além de um canal de restituição, através do qual a água é reconduzida ao rio ou a um outro curso d'água. A figura 3.2 apresenta o diagrama simplificado de uma usina hidrelétrica.

No processo de geração de energia elétrica, a energia potencial da água armazenada no reservatório é transformada em energia cinética e energia de pressão dinâmica pela passagem da água pelos condutos forçados. Ao fazer o acionamento da turbina, essa energia é convertida em energia mecânica, por sua vez transmitida pelo eixo ao gerador. Neste último, a energia mecânica é transformada em energia elétrica, que passa por uma subestação elevadora de tensão, sendo então injetada no sistema de transmissão até chegar aos centros consumidores.

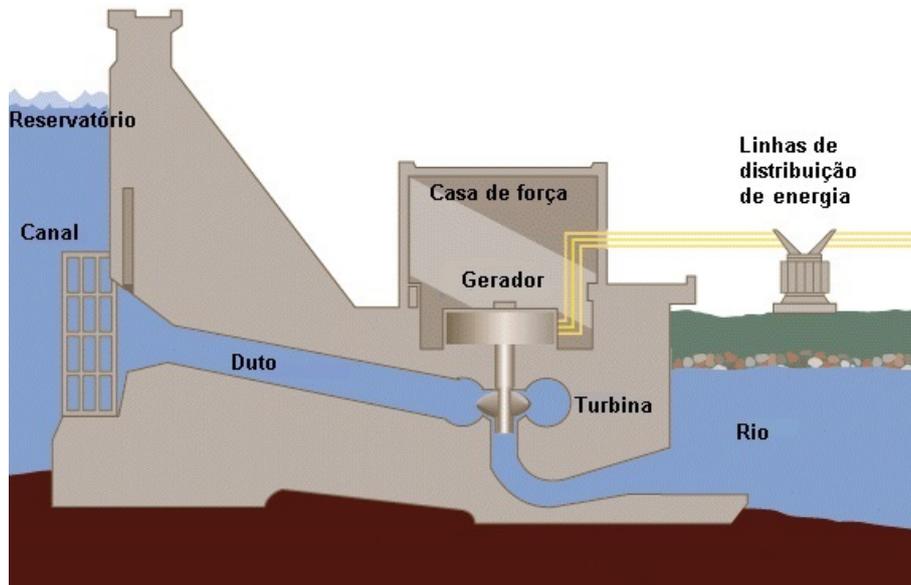


Figura 3.2: Perfil típico de uma usina hidrelétrica

A barragem tem como papel fundamental a criação de uma diferença de níveis entre o espelho d'água do reservatório (montante) e o canal de fuga (jusante). Além disto, ela forma um reservatório onde a água, constantemente renovada, é captada para a produção de energia elétrica.

Uma boa compreensão do funcionamento dos componentes de uma usina é fundamental para estabelecer as relações matemáticas que regem o seu comportamento e, assim como as usinas termelétricas, é necessário definir qual é o modelo matemático mais adequado para as necessidades das diferentes fases dos estudos de planejamento.

As principais características das usinas hidrelétricas são o seu baixo custo de produção de energia e, devido à presença dos reservatórios, a sua capacidade de armazenar energia (água), para utilização futura. Como comentado anteriormente, isto faz com que o problema de planejamento da operação fique acoplado no tempo e exige que as usinas hidrelétricas sejam avaliadas não apenas pelo seu custo direto de produção, mas pelo seu *custo de oportunidade*³.

A seguir, serão descritas as equações matemáticas consideradas neste trabalho para descrever o funcionamento das usinas hidrelétricas.

3.2.1 Balanço hídrico

As usinas hidrelétricas são construídas de forma a maximizar o aproveitamento dos recursos existentes em um determinado rio ou bacia hidrográfica e, para isso, é

³O custo de oportunidade da água representa o *trade-off* entre evitar um custo de produção de energia no presente ou no futuro.

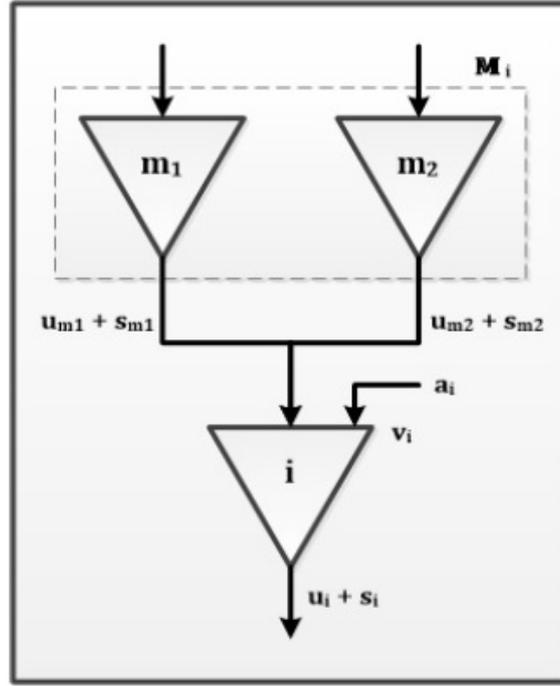


Figura 3.3: Balanço hídrico de usinas em cascata

comum que várias usinas sejam dispostas em sequência ou cascata. Este arranjo cria um acoplamento espacial entre as usinas ao passo que a presença de reservatórios cria um acoplamento temporal, devido à relação entre os volumes armazenados em cada reservatório entre estágios consecutivos.

A operação hidráulica de um aproveitamento ao longo de um período de tempo obedece ao princípio de conservação da água. Assim, o volume de água que aflui ao aproveitamento no período deve ser igual à soma dos volumes correspondentes à variação do nível do reservatório e liberados pelas turbinas e/ou vertedores. Esta relação está indicada na equação (3.3).

$$v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}), \quad i \in \mathbb{I}, \quad t \in \mathbb{T}, \quad (3.3)$$

onde i indexa as usinas hidrelétricas, I é o número total de usinas, $v_{t,i}$ é o volume final da usina i no estágio t , $a_{t,i}$ é a vazão incremental afluyente à usina, $u_{t,i}$ e $s_{t,i}$ são respectivamente o volume turbinado e vertido pela usina i no estágio t , $m \in M_i$ é o conjunto de usinas diretamente à montante da usina i e, finalmente, \mathbb{I} é o conjunto de usinas hidrelétricas. A equação de balanço hídrico para uma usina em cascata é ilustrada esquematicamente na figura 3.3.

3.2.2 Limites no armazenamento e na defluência

Representam, para cada usina i e em cada estágio de tempo t , as restrições operativas relacionadas à operação do reservatório, de acordo com o conjunto de equações (3.4).

$$v_{t,i} \leq \bar{v}_i, \quad i \in \mathbb{I}, \quad t \in \mathbb{T} \quad (3.4a)$$

$$u_{t,i} \leq \bar{u}_i, \quad i \in \mathbb{I}, \quad t \in \mathbb{T} \quad (3.4b)$$

onde \bar{v}_i e \bar{u}_i são os limites máximos de armazenamento (volume) e turbinamento da usina i , respectivamente.

3.2.3 Produção de energia

As usinas hidroelétricas convertem a energia potencial da água armazenada em energia cinética, a qual se usa para mover as turbinas acopladas a geradores. A energia produzida em uma determinada hidroelétrica i pelo turbinamento de $u_{t,i}$ se calcula por:

$$e_{t,i} = \rho_i \times u_{t,i}, \quad i \in \mathbb{I}, \quad t \in \mathbb{T} \quad (3.5)$$

onde $e_{t,i}$ é a variável de decisão que representa a energia produzida pela usina hidroelétrica i no estágio t e ρ_i é o coeficiente de produção da usina i . O coeficiente de produção é uma constante que depende de características físicas da usina tais como a eficiência do conjunto turbina-gerador e altura de queda líquida (diferença entre os níveis de montante e justante do reservatório), além de constantes físicas como o fator gravitacional e a massa específica da água.

3.3 Demanda

A demanda de energia elétrica é dada pela soma da demanda de consumidores residenciais, comerciais, industriais e rurais em um determinado intervalo de tempo.

3.3.1 Atendimento à demanda

O atendimento à demanda é imposto através de uma equação que determina que, em cada intervalo de tempo, a soma da energia produzida (pelas usinas termelétricas e hidroelétricas) deve ser igual à demanda, conforme expresso na equação (3.6).

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} = d_t, \quad t \in \mathbb{T} \quad (3.6)$$

onde d_t é a demanda de energia elétrica no intervalo de tempo t .

3.3.2 Déficit de energia

A falha de suprimento ocorre quando o sistema de geração não consegue produzir energia suficiente para atender a demanda. Neste caso, a equação de atendimento à demanda (3.6) ficaria inviável e a política não poderia ser determinada. Este problema pode ser contornado através da introdução de uma nova variável (de folga), r_t , que representa o corte de carga, ou déficit, no estágio. Desta forma, a equação (3.6) é reescrita como:

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t, \quad t \in \mathbb{T} \quad (3.7)$$

Obviamente, déficits de energia não são desejados e, tradicionalmente, a maneira utilizada para evitá-los é considerar uma penalização associada à sua ocorrência na função objetivo, através do termo $\delta \times r_t, t \in \mathbb{T}$, onde δ representa um custo de déficit pré-determinado. No capítulo 4 serão apresentadas e propostas outras abordagens para evitar a ocorrência de déficits de energia.

Vale a pena ressaltar que, sob o ponto de vista matemático, o déficit r_t é similar a uma nova “usina termelétrica” de “capacidade máxima” igual a d_t (pois no pior caso toda a demanda seria cortada) e “custo operativo” δ . Em outras palavras, o modelo de otimização deverá decidir entre atender a demanda com usinas termelétricas reais ou com o déficit.

3.4 Formulação do problema de uma etapa

Como visto anteriormente, o problema de planejamento da operação de sistemas hidrotérmicos é acoplado no tempo devido à presença dos reservatórios das usinas hidrelétricas. Uma maneira bastante conveniente de representar este problema é através das equações de programação dinâmica, onde descreve-se o problema associado a cada estágio de tempo e seu acoplamento com os estágios anteriores através das variáveis de estado que definem as decisões do estágio anterior e a função de custo futuro, que define o impacto das decisões do estágio corrente na operação futura.

O problema de planejamento da operação tradicional é definido então como uma

sequência de problemas de um estágio, conforme o problema descrito a seguir.

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}) = \\ & \min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \tilde{Q}_{t+1}(v_t, a_{[t]}) \end{aligned} \quad (3.8a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad i \in \mathbb{I} \quad (3.8b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad i \in \mathbb{I} \quad (3.8c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (3.8d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (3.8e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (3.8f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (3.8g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (3.8h)$$

Capítulo 4

Representação da aversão ao risco

4.1 Introdução

Como foi visto anteriormente, o objetivo básico do planejamento da operação de sistemas elétricos é atender a demanda dentro do horizonte pré-estabelecido a um custo mínimo. Este custo de operação é composto pela soma atualizada dos custos incorridos pelo uso de combustíveis para geração térmica mais penalizações devido a falhas no atendimento da demanda (custo de déficit). Como o futuro está sujeito à incertezas, em geral avalia-se a qualidade de uma política operativa através de seu custo médio.

Entretanto, além dos valor esperado dos custos, outro aspecto de importante relevância é a avaliação da confiabilidade de suprimento de energia, ou seja, a exposição destas decisões à ocorrência de déficits de energia. Desta maneira, é desejável também garantir que a política escolhida atenda a algum critério de segurança pré-estabelecido.

Como foi visto na seção 3.3.2, o déficit de energia pode ser representado por uma variável adicional no problema de planejamento da operação e é possível informar o peso relativo das entre falhas de suprimento (déficits) com relação aos custos operativos. Uma maneira de calcular este peso relativo é calcular o *custo econômico* do déficit. Isto pode ser feito através de uma matriz insumo-produto da economia, onde basicamente se calcula a perda da produção agregada (PIB) no caso de um racionamento. O resultado deste estudo é uma função custo x profundidade do déficit. Se a função for linear, ela é caracterizada por um único parâmetro, o custo unitário do déficit¹. Uma vez conhecida a função de custo de déficit, o objetivo do planejamento da operação passa a ser o de minimizar o valor esperado da soma dos custos operativo e de déficit.

No entanto, o uso do custo econômico do déficit tem duas limitações importantes:

¹Por exemplo, o custo unitário do déficit utilizado atualmente no Brasil pela Empresa de Pesquisa Energética (EPE) em seus estudos é de R\$ 3.100/MWh.

- (i) incerteza nos parâmetros da matriz insumo-produto, que impacta a precisão do cálculo deste custo; e
- (ii) a confiabilidade de suprimento é um resultado do processo de otimização e não um parâmetro de entrada para o mesmo.

Isto significa que, se por alguma razão, houver uma mudança na aversão ao risco da sociedade, não seria possível representá-la diretamente no cálculo da política operativa.

Classicamente, o modelo econômico utilizado para capturar o comportamento de aversão ao risco em problemas de decisão sob incerteza é o modelo de utilidade esperada [44]. Porém, em função da dificuldade de se encontrar uma utilidade prática, que expresse a aversão ao risco do tomador de decisão com base em parâmetros práticos, de forma mais isenta a uma subjetividade pessoal e que simplesmente expressassem a “a aversão ao risco da entidade”, neste caso, o planejador do sistema, a busca por medidas de risco que fossem capazes de controlar o processo decisório foi bastante impulsionada.

Neste sentido, buscou-se atacar o problema “risco” de resultados indesejáveis de forma mais direta e pragmática, onde a distribuição de probabilidade dos resultados seria diretamente monitorada e “controlada” através da imposição de limites, previamente estipulados, à medidas estatísticas das variáveis incertas. Assim sendo, a atribuição subjetiva de um valor de utilidade ou satisfação para cada resultado, caracterizado pela função utilidade foi deixada de lado e o objetivo passou a ser, em mercados financeiros, a maximização do valor esperado dos resultados futuros controlados por restrições de risco [60].

No caso do planejamento da operação de sistemas hidrotérmicos, passa a ser interessante analisar a possibilidade de determinar uma política operativa onde o objetivo é minimizar o valor esperado do custo operativo, sujeito a uma *restrição na confiabilidade* de suprimento.

O objetivo deste capítulo é descrever como lidar com o déficit, avaliar diferentes medidas de risco e suas propriedades e, finalmente, descrever o modelo matemático e algoritmo proposto para a solução do problema de planejamento da operação considerado restrições de risco, em especial, para o caso da medida de risco *Conditional Value at Risk* (CVaR).

4.2 Risco

Considere uma variável aleatória R e seja Ω o conjunto de todas as possíveis realizações desta variável.

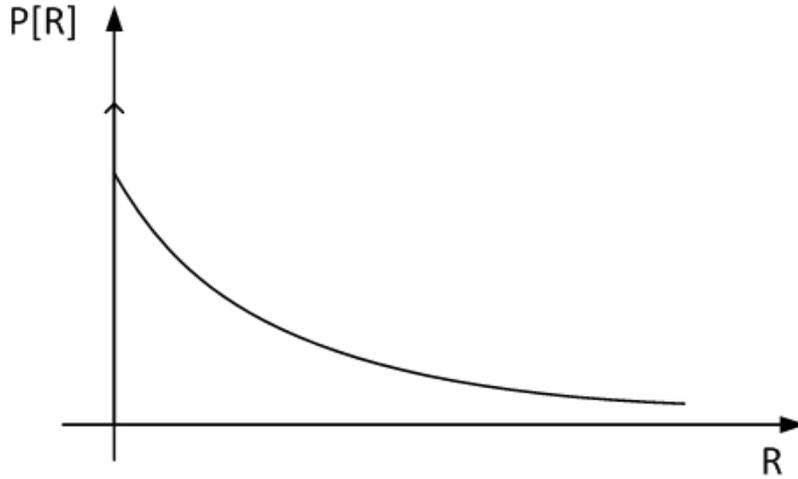


Figura 4.1: Distribuição de probabilidade

A abordagem mais direta para avaliar o desempenho de um sistema é identificar, dentro do conjunto de possíveis estados, o número de estados que levam a algum problema. Além do custo de produção de energia, um indicador de qualidade de uma política operativa tipicamente utilizada no contexto do planejamento da operação é o déficit de energia, representado pela variável aleatória (v.a.) R .

O risco de déficit corresponde à probabilidade de haver algum déficit energia, ou seja, $\text{Risco} = \Pr(R > 0)$. A figura 4.2 ilustra a função de distribuição de probabilidade de R e o risco de déficit.

Para cada estado ξ_s está associada uma realização da v.a. corte de carga r_s e, desta forma, o risco de déficit pode ser calculado como

$$\text{Risco} = \sum_{s \in \bar{\Omega}} p_s \quad (4.1)$$

onde $\bar{\Omega} = \{s \in \Omega | r_s > 0\}$ é o conjunto dos cenários insuficientes e p_s é a probabilidade de ocorrência de cada cenário.

Seja $u(R)$ a função indicadora (ou degrau), definida a seguir e ilustrada na figura 4.3.

$$u(R) = \begin{cases} 1, & \text{se } R > 0 \\ 0, & \text{caso contrário} \end{cases} \quad (4.2)$$

Alternativamente, o risco de déficit pode ser calculado como o valor esperado da função indicadora

$$\text{Risco} = \sum_{s \in \Omega} p_s u(r_s) \quad (4.3)$$

Apesar de o risco de déficit ser uma estatística associada a um quantil da função

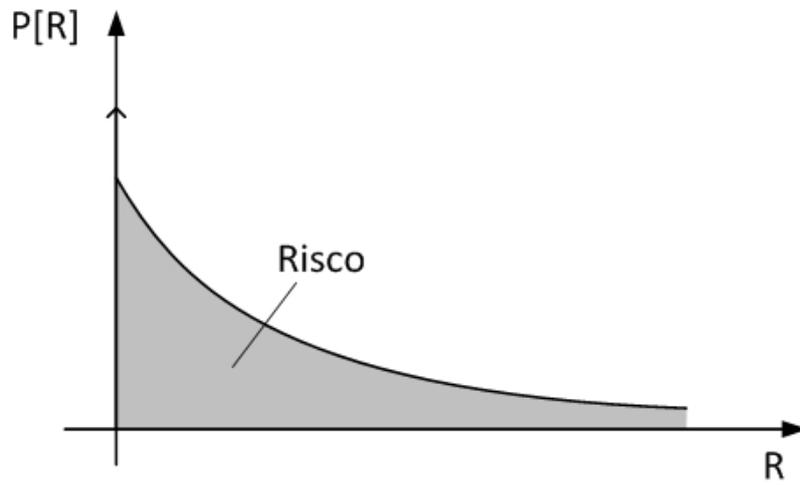


Figura 4.2: Risco de déficit de energia

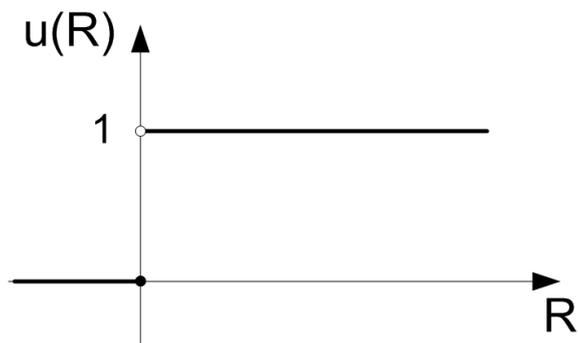


Figura 4.3: Função indicadora

de distribuição de probabilidade da variável aleatória corte de carga, obtida a partir da solução do problema de operação, é possível formular seu cálculo em um problema de otimização linear inteira mista. O objetivo desta abordagem é obter um modelo que possa ser inserido no problema de planejamento da operação.

$$\text{Risco} = \text{Min} \quad \sum_{s \in \Omega} p_s \phi_s \quad (4.4a)$$

$$\text{sujeito a} \quad \phi_s \geq \frac{1}{D} r_s \quad \forall s \in \Omega \quad (4.4b)$$

$$\phi_s \in \{0, 1\} \quad \forall s \in \Omega \quad (4.4c)$$

onde ϕ_s é uma variável binária que desempenha o mesmo papel da função indicadora.

Observe que a restrição (4.4b) assegura que, para os cenários com corte de carga, ϕ_s deverá ser maior que 0 e como trata-se de uma variável binária, assumirá o valor 1. Adicionalmente, o corte de carga r_s não pode ser maior que a demanda D , o que garante que o lado direito da restrição (4.4b) estará sempre no intervalo $[0, 1]$.

O problema (4.4) representa o cálculo de um quantil da distribuição de R que é, neste caso, a probabilidade de a variável aleatória R ser maior que zero. Apesar de sua simplicidade, o interessante de se formular o cálculo do risco como um problema de otimização é notar que seu cálculo envolve a utilização de uma variável inteira para cada cenário.

Além disto, é possível notar que esta medida não é capaz de diferenciar os cenários com corte de carga, sendo indiferente para situações onde há um corte de carga pequeno ou um corte de carga mais severo (em termos de potência/energia) e esta pode ser considerada uma das maiores críticas na utilização do risco como critério de planejamento.

4.3 Valor esperado

O valor esperado do déficit $\mathbb{E}[R]$ corresponde ao valor esperado da energia não suprida e consiste em avaliar a média do mínimo corte de carga, para todos os possíveis estados do sistema, conforme ilustrada na figura 4.4. Este valor é calculado como:

$$\mathbb{E}[R] = \sum_{s \in \Omega} p_s r_s \quad (4.5)$$

Observe que, como o corte de carga é uma variável não negativa, sua representação é dada implicitamente pela função truncamento

$$[R]^+ = \begin{cases} R, & \text{se } R > 0 \\ 0, & \text{caso contrário} \end{cases} \quad (4.6)$$

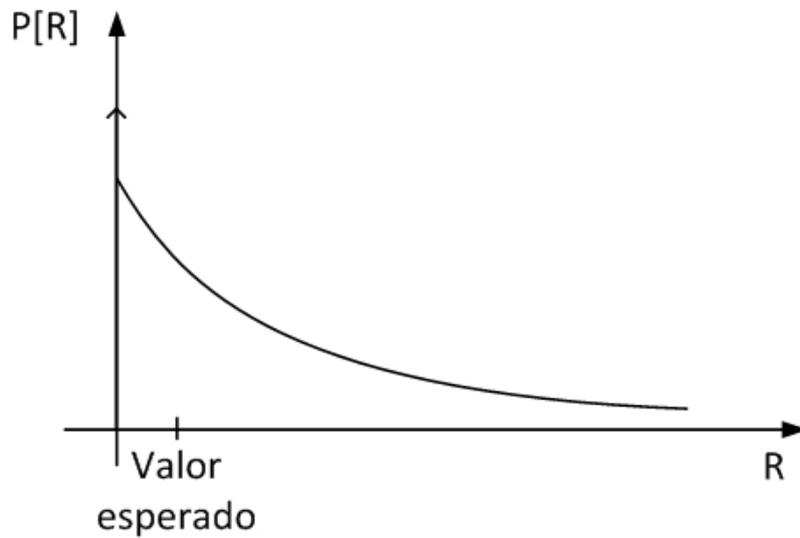


Figura 4.4: Valor esperado da corte de carga

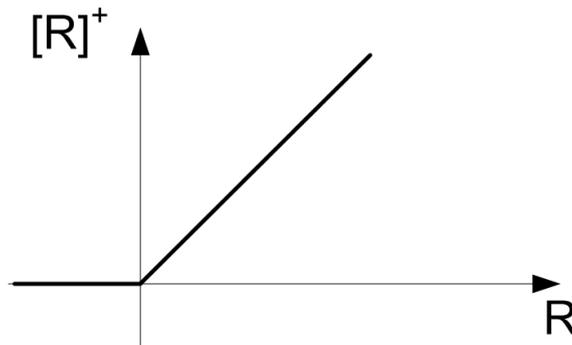


Figura 4.5: Função truncamento

que é ilustrada na figura 4.5.

A EENS reflete o valor esperado do corte de carga capturando, em média, uma medida de o risco do sistema. Por outro lado, como trata-se da média de todos os cenários, esta medida acaba ficando “diluída”, podendo não revelar a presença de eventos desastrosos sendo esta uma crítica feita à esta medida.

4.4 Value-at-Risk (VaR)

4.4.1 Definição

O *Value-at-Risk* é uma medida de risco de utilização bastante comum em mercados financeiros para a quantificação de riscos e desempenho associados a portfólios e passou a ser amplamente difundido após a ocorrência de desastres que levaram diversas instituições financeiras à falência [37]. Considerando sua aplicação no contexto de

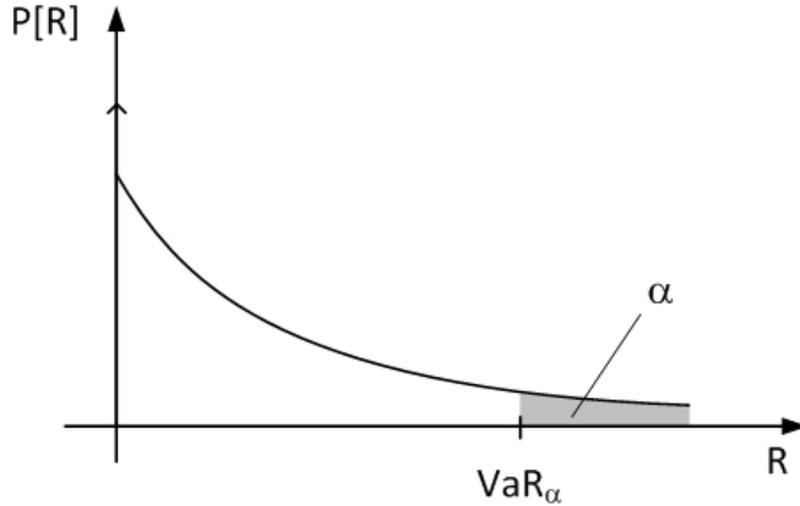


Figura 4.6: VaR_α : *Value-at-Risk* associado ao corte de carga

cálculo de confiabilidade, esta medida procura responder à seguinte questão:

“Qual é o máximo corte de carga previsto dentro de um intervalo de confiança estabelecido?”

Considerando uma variável aleatória R qualquer, o valor de risco é definido como

$$\text{VaR}_\alpha(R) = \inf\{r : \Pr(R > r) \leq \alpha\} \quad (4.7)$$

onde α é um determinado nível de confiança. Se R representa o corte de carga, $\text{VaR}_\alpha(R)$ é o máximo corte de carga a um nível de confiança de α , conforme ilustrado na figura 4.6.

Em outras palavras, $\text{VaR}_\alpha(R) = \overline{\text{VaR}}$ é o quantil mínimo associado a $\Pr(R > \overline{\text{VaR}}) \leq \alpha$, ou seja, a probabilidade de a variável aleatória R exceder o valor $\overline{\text{VaR}}$ ser menor ou igual a α .

É possível definir um problema de otimização para o cálculo do VaR, conforme ilustrado no problema (4.8).

$$\text{VaR}_\alpha = \text{Min } y \quad (4.8a)$$

$$\text{sujeito a } y \geq r_s - D\phi_s \quad \forall s \in \Omega \quad (4.8b)$$

$$\sum_{s \in S} p_s \phi_s \leq \alpha \quad (4.8c)$$

$$\phi_s \in \{0, 1\} \quad \forall s \in \Omega \quad (4.8d)$$

A equivalência entre o problema (4.8) e a definição (4.7) pode ser demonstrada. No problema (4.8), a variável binária ϕ_s tem o objetivo de “selecionar” alguns estados para que não influenciem no valor de y . Quando o estado é selecionado ($\phi_s = 1$) o

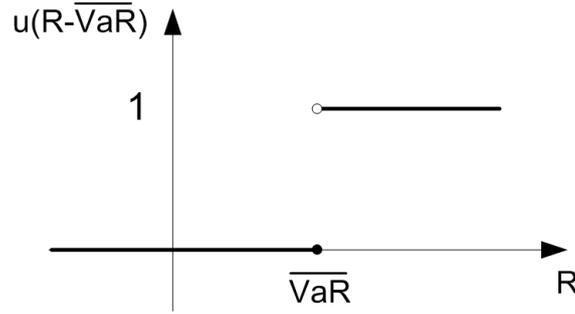


Figura 4.7: Função indicadora transladada

lado direito da restrição (4.8b) é relaxado, uma vez que o máximo corte de carga é igual à demanda total D . Entretanto, o número de estados selecionados está limitado, pois a restrição (4.8c) restringe a soma das probabilidades dos estados selecionados ao nível de confiança definido α .

Como a função objetivo do problema é a minimização da variável y e y deve ser maior que o lado direito da restrição (4.8b) para todos os estados, y assumirá o valor do maior lado direito entre os estados não selecionados ($\phi_s = 0$).

Portanto, o VaR_α define o maior valor do corte de carga dentro de um intervalo de confiança, entretanto, assim como o risco de déficit, não é sensível aos estados que excedem este valor, ou seja, que compõe a cauda da distribuição.

4.4.2 Relação entre o risco de déficit e VaR

De modo similar ao risco de déficit, é possível notar através da restrição (4.8b) que a representação do VaR_α também é realizada implicitamente pela função indicadora, porém transladada de $\overline{\text{VaR}}$, como ilustrado na figura 4.7.

Este fato sugere uma relação entre o critério VaR_α e o risco de déficit. Na realidade, se for definido $\overline{\text{VaR}} = 0$ e $\alpha = \overline{\text{Risco}}$, tem-se que o problema torna-se equivalente ao problema onde tem-se o risco como critério de confiabilidade, ou seja

$$\Pr(R > 0) \leq \alpha \Leftrightarrow \text{Risco} \leq \alpha \tag{4.9}$$

Isto demonstra que o critério de planejamento utilizando a medida VaR_α é, na verdade, uma generalização do critério do risco de déficit.

4.5 Coerência das medidas de risco

Durante anos, tornaram-se claras as diferenças entre as práticas do mercado financeiro e os poucos avanços teóricos obtidos na área de gestão de risco. Desde então,

diversos trabalhos [2] surgiram com o intuito de definir, de maneira clara e objetiva, quais características um índice estatístico deveria ter para ser considerado uma medida de risco razoável.

Em 1997, ARTZNER *et al.* [6] introduziram o conceito de *medida de risco coerente* (MRC) impondo, via formulação axiomática, condições matemáticas específicas que forçam alguns princípios básicos que uma medida de risco razoável deve sempre satisfazer. Além deste trabalho, muitos outros esforços [7] foram realizados no sentido de estabelecer tais propriedades. Este novo “pilar” da matemática financeira é considerado por muitos especialistas como a primeira tentativa de uma definição precisa do próprio risco (financeiro) através de uma abordagem teórica.

Segundo ARTZNER *et al.* [6], se χ é um espaço linear de funções mensuráveis que definem as variáveis aleatórias de um espaço de probabilidades bem definido, então $\rho : \chi \mapsto \mathbb{R}$ é uma medida de risco coerente se satisfaz as propriedades 1-4, definidas a seguir:

1. Sub-aditividade: $\rho(X + Y) \leq \rho(X) + \rho(Y), \forall X, Y \in \chi$,
 2. Monotonicidade: $X \leq Y \Rightarrow \rho(X) \leq \rho(Y), \forall X, Y \in \chi$,
 3. Homogeneidade Positiva: $\rho(\lambda X) = \lambda\rho(X), \forall X \in \chi$ e $\lambda > 0$,
 4. Invariância por translação: $\rho(X + a) = \rho(X) + a, \forall X \in \chi$ e $a \in \mathbb{R}$,
- Note que, em conjunto, as propriedades 1 e 3 implicam em:
5. Convexidade: $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y), \forall X, Y \in \chi, \forall \lambda \in [0, 1]$.

onde $\rho(\cdot)$ representa uma medida de risco e X, Y duas variáveis aleatórias.

Alguns autores como ACERBI e TASCHE [2] consideram o adjetivo “coerente” redundante e classificam como medida de risco qualquer medida $\rho(\cdot)$ que satisfaça as propriedades 1-4. A importância destas propriedades são discutidas a seguir:

- A propriedade 1 da sub-aditividade implica que a medida de risco leva em consideração o efeito “portfólio”, ou seja, incentiva a diversificação. A medida de risco de dois portfólios em conjunto não pode ser maior que a soma das medidas de risco dos portfólios em separado.
- A propriedade 2 da monotonicidade implica que, dadas duas políticas operativas A e B , se o corte de carga associado a A é menor que o corte de carga associado a B , para todos os estados, então a medida de risco para a política A deve ser menor que para B .

- A propriedade 3 que define que a medida deve ser homogênea positiva implica que um aumento na variável aleatória deve aumentar linearmente sua medida de risco.
- Finalmente, a propriedade 4 da equivariância por translação implica que, adicionando ou subtraindo uma quantidade certa a à variável aleatória X , a medida de risco aumenta ou diminui do mesmo valor a . Em finanças, isto pode ser visto como a adição de uma renda certa ao portfólio, ou seja, livre de risco. A medida de risco do portfólio é transladada exatamente deste valor, o que significa que o risco do portfólio não se altera.

Dentre os axiomas de coerência, é possível destacar o axioma da subaditividade. A razão pela qual este axioma vem sendo amplamente discutido [1] é provavelmente pelo fato de o VaR - a medida de risco mais popular em termos de adequação de capital e que sempre fora uma medida de risco adotada como uma “boa prática” pela maioria dos bancos e reguladores - não garantir esta característica e conseqüentemente ser não coerente [3].

Desde esta “descoberta” o desenvolvimento da teoria de MRCs deu-se em paralelo à discussão de *quando* e *como* o VaR deve ser abandonado pelos gestores de risco. Aliado a isto, o fato de muitas medidas não apresentarem tais características e ainda assim terem sido utilizadas durante anos e o fato de o VaR ainda possuir diversas vantagens (simplicidade, grande aplicabilidade, universalidade, etc.) levou muitos a julgar este conjunto de propriedades como opcional, dando-se pouca atenção à não coerência do VaR [2].

Entretanto, o não cumprimento dos axiomas pode conduzir a paradoxos e conclusões erradas, como é o caso do VaR [4]. Para medidas como esta é possível encontrar exemplos onde um portfólio mais diversificado é caracterizado como mais arriscado. Do ponto de vista do planejamento da operação, uma medida com tal característica poderia classificar, por exemplo, uma determinada política com *mais* recursos disponíveis *menos* confiável do que uma outra com uma operação *mais* arriscada. Portanto, a utilização de uma medida como esta pode levar a distorções no processo de seleção da política mais adequada, prejudicando o planejamento da operação.

4.6 Conditional Value-at-Risk (CVaR)

4.6.1 Definição

Dentro deste contexto de incoerência do VaR, diversas medidas de risco alternativas foram estudadas e propostas para substituí-lo. Em [2], os autores introduzem o

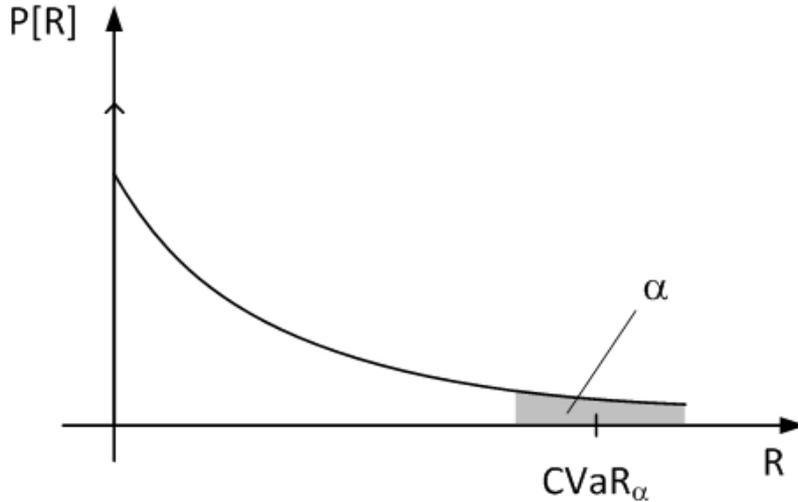


Figura 4.8: CVaR_α : *Conditional Value-at-Risk* associado ao corte de carga

CVaR como a alternativa natural para o VaR em problemas de portfólio, evidenciando suas propriedades de coerência e convexidade, como será visto adiante.

O *Conditional Value-at-Risk* de uma variável aleatória R qualquer é definido como

$$\text{CVaR}_\alpha(R) = \mathbb{E}[R : R \geq \text{VaR}_\alpha(R)] \quad (4.10)$$

que representa o valor esperado condicionado aos eventos maiores que o VaR_α e pode ser visto como um quantil de R , conforme ilustrado na figura 4.8.

Se R representa o corte de carga, o $\text{CVaR}_\alpha(R)$ procura responder à seguinte questão:

“Qual é o valor esperado do corte de carga condicionado aos $\alpha\%$ piores cenários?”

Embora a definição (4.10) seja bastante intuitiva, à primeira vista o cálculo do CVaR_α está condicionado ao cálculo do VaR_α , herdando assim as dificuldades de sua representação.

Entretanto, após o trabalho publicado por [54], o CVaR foi impulsionado frente às demais medidas de risco devido às características de sua formulação matemática. Essencialmente, o CVaR_α que anteriormente era calculado como um valor esperado condicionado ao valor do VaR_α , conseqüentemente tornado-se dependente da implementação deste último, passou a ser formulado como

$$\text{CVaR}_\alpha(R) = \inf_{b \in \mathbb{R}} \{ \sigma + \alpha^{-1} \mathbb{E}[[R - b]^+] \} \quad (4.11)$$

onde $[\cdot]^+$ é a função truncamento apresentada na seção 4.3, neste caso transladada em b , conforme ilustrado na figura 4.9.

A formulação (4.11) possibilitou a utilização do CVaR_α em problemas de oti-

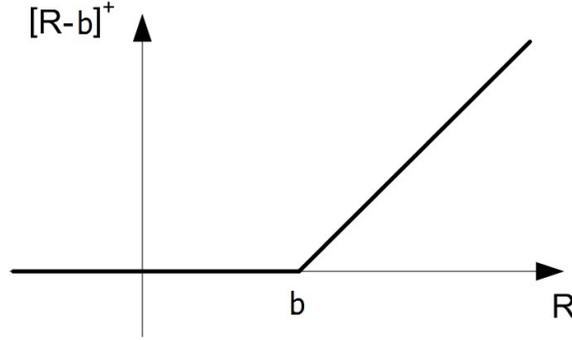


Figura 4.9: Função truncamento transladada

mização linear de forma direta, com a simples adição de um conjunto de restrições lineares ao problema original.

No contexto da confiabilidade, o CVaR_α da variável aleatória corte de carga, que tem distribuição discreta de suporte finito, pode ser representado pelo problema de programação linear (4.12).

$$\text{CVaR}_\alpha(R) = \text{Min} \quad b + \alpha^{-1} \sum_{s \in \Omega} p_s y_s \quad (4.12a)$$

$$\text{sujeito a} \quad y_s \geq r_s - b \quad \forall s \in \Omega \quad (4.12b)$$

Como y_s é uma variável não negativa, a restrição (4.12b) representa a função truncamento deslocada em b , ou seja, y_s representa a diferença positiva entre o corte de carga r_s e b . Portanto, o CVaR_α define a soma ponderada pela probabilidade dos cortes de carga superiores a b .

A métrica de risco CVaR tem sido vastamente utilizada em problemas de decisão de portfólios em função da sua capacidade de capturar a presença de eventos de alta profundidade (catastróficos) na função de distribuição. Esta tem sido uma crítica recorrente feita ao VaR que, como visto anteriormente, não diferencia distribuições com diferentes profundidades.

4.6.2 Relação entre o valor esperado do corte de carga e o CVaR

Assim como o valor esperado do corte de carga, o CVaR_α também utiliza implicitamente a função truncamento em sua definição, porém transladada, sugerindo alguma relação entre estas medidas. É possível notar que o CVaR_α é uma generalização do conceito de valor esperado, uma vez que, por definição

$$\text{CVaR}_\alpha(R) = \mathbb{E}[R], \quad \text{para } \alpha = 1 \quad (4.13)$$

ou seja, o valor esperado do corte de carga é igual ao valor esperado corte de carga condicionado aos 100% estados mais severos.

Além disto, quando $\alpha \geq \text{Risco}$ tem-se que todos os estados com corte de carga irão contribuir para o cálculo do CVaR_α . Neste caso é possível estabelecer outra relação entre o CVaR_α e a $\mathbb{E}[R]$:

$$\text{CVaR}_\alpha = \frac{1}{\alpha} \sum_{s \in \Omega} p_s r_s = \frac{1}{\alpha} \mathbb{E}[R] \quad (4.14)$$

o que ilustra o fato de o CVaR_α ser uma medida mais geral que $\mathbb{E}[R]$, lembrando que $0 \leq \alpha \leq 1$.

4.6.3 Relação entre o VaR e o CVaR

Como visto anteriormente, tanto o risco de déficit como o VaR são representados implicitamente pela função indicadora e, em modelo de otimização, deseja-se que a solução satisfaça à seguinte restrição:

$$\text{VaR}_\alpha(R) \leq \overline{\text{VaR}} \quad (4.15)$$

onde R é o corte de carga que depende diretamente do vetor de decisões x .

Como $\Pr(R > \overline{\text{VaR}}) = \mathbb{E}[u(R - \overline{\text{VaR}})]$, é possível reescrever a equação (4.15) como a seguinte restrição em valor esperado:

$$\mathbb{E}[u(R - \overline{\text{VaR}})] \leq \alpha \quad (4.16)$$

A fonte de dificuldades da utilização de restrições em probabilidade como a (4.16) é que a função indicadora $u(\cdot)$ é não convexa e, pior ainda, descontínua em 0. Como resultado, este tipo de restrição é em geral não-convexa mesmo no caso em que a função $x \mapsto R$ é convexa². Uma alternativa para contornar esta dificuldade é construir uma aproximação convexa para o valor esperado apresentado no lado esquerdo de (4.16).

Seja $\psi : \mathbb{R} \mapsto \mathbb{R}$ uma função não-negativa, não-decrescente e convexa de tal forma que $\psi(y) \geq u(y)$ para todo $y \in \mathbb{R}$. Observando que $u(\sigma y) = u(y)$ para qualquer $\sigma > 0$ e $y \in \mathbb{R}$, tem-se que $\psi(\sigma y) \geq u(y)$ e então a seguinte inequação é válida

$$\inf_{\sigma > 0} \mathbb{E}[\psi(\sigma y)] \geq \mathbb{E}[u(y)] \quad (4.17)$$

²além disto, a utilização de variáveis inteiras na sua formulação para modelar a descontinuidade introduz dificuldades para solução

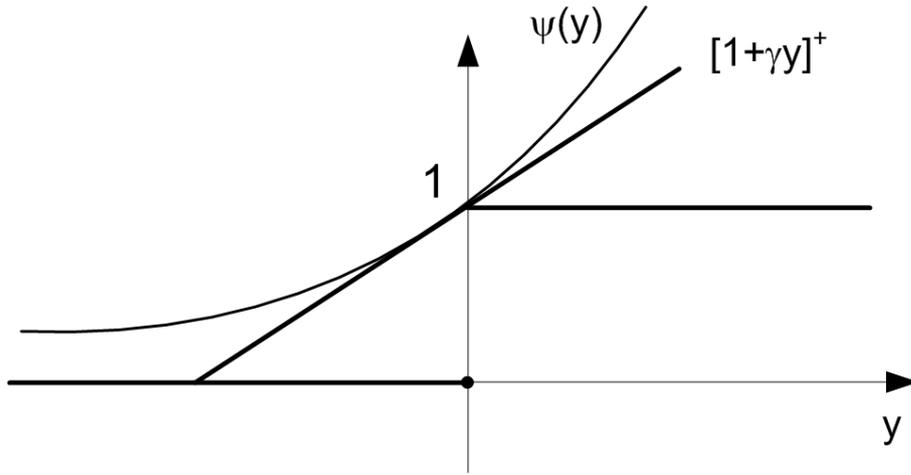


Figura 4.10: Aproximação convexa da função indicadora

Conseqüentemente, a restrição

$$\inf_{\sigma > 0} \mathbb{E}[\psi(\sigma R)] \leq \alpha \quad (4.18)$$

é uma aproximação *conservadora* da restrição em probabilidade (4.15) no sentido de que a região viável definida por (4.18) está contida na região viável definida por (4.15).

Obviamente, quanto menor for a função $\psi(\cdot)$ melhor será a aproximação da restrição (4.15). Por esta razão, a melhor escolha para $\psi(\cdot)$ é a função linear por partes $\psi(y) = [1 + \gamma y]^+$, para algum $\gamma > 0$. Como foi visto que a restrição (4.18) é invariante com respeito à escala, tem-se que $\psi(y) = [1 + y]^+$ é a melhor escolha para esta função, conforme ilustrado na figura 4.10. Note que esta é exatamente a função truncamento (transladada) apresentada na seção 4.3.

Dada esta escolha para função $\psi(\cdot)$, tem-se que a restrição (4.18) é equivalente a

$$\inf_{\sigma > 0} \{ \sigma \mathbb{E}[\sigma^{-1} + R]^+ - \alpha \} \leq 0 \quad (4.19)$$

ou equivalentemente

$$\inf_{\sigma > 0} \{ \alpha^{-1} \mathbb{E}[R + \sigma^{-1}]^+ - \sigma^{-1} \} \leq 0 \quad (4.20)$$

Substituindo $b := -\sigma^{-1}$, encontra-se a forma

$$\inf_{b < 0} \{ b + \alpha^{-1} \mathbb{E}[R - b]^+ \} \leq 0 \quad (4.21)$$

que corresponde à expressão do CVaR para um nível de confiança α , conforme pode

ser observado abaixo

$$\text{CVaR}_\alpha(R) = \inf_{b \in \mathbb{R}} \{b + \alpha^{-1} \mathbb{E}[R - b]^+\} \quad (4.22)$$

Desta forma, é possível aproximar a restrição em probabilidade do VaR por

$$\text{CVaR}_\alpha(R) \leq \overline{\text{VaR}} \quad (4.23)$$

demonstrando que a medida de risco CVaR_α é uma aproximação convexa conservadora de (4.15). Mais precisamente, (4.23) é a *melhor* aproximação convexa conservadora de (4.15). A demonstração original pode ser encontrada no trabalho de NEMIROVSKI e SHAPIRO [43].

É possível encontrar diversos trabalhos na literatura [4, 60] que estudam as características do CVaR e que demonstram que esta medida de risco satisfaz as propriedades 1-4 que definem uma medida de risco (coerente).

Capítulo 5

Planejamento da operação com restrições de CVaR com formulação em árvore

5.1 Introdução

O objetivo desta seção é mostrar como incorporar restrição de risco do tipo CVaR em problemas de programação linear estocástica multiestágio. A restrição de CVaR será introduzida considerando a formulação do capítulo 2 e, em seguida, será apresentada a dualização da restrição e o algoritmo proposto para obtenção da solução.

5.2 Formulação geral

Considere o problema representado pela formulação em árvore definida na seção 2.3.3, expresso novamente abaixo:

$$\min \quad c_1^\top x_1 + \sum_{t=2}^T \sum_{n \in \Omega_t} p^{(n)} c_t^\top x_t^{(n)} \quad (5.1a)$$

$$\text{s.a.:} \quad A_1 x_1 = b_1 \quad (5.1b)$$

$$B_t x_{t-1}^{a(n)} + A_t x_t^{(n)} = b_t^{(n)} \quad n \in \Omega_t, \quad t = 2, \dots, T \quad (5.1c)$$

$$x_t^{(n)} \geq 0 \quad n \in \Omega_t, \quad t = 1, \dots, T \quad (5.1d)$$

Seja $h_\tau(x_\tau^{(n)})$ uma função linear definida para um determinado estágio $\tau \in \mathbb{T}$ e calculada em função do vetor de variáveis de decisão $x_\tau^{(n)}$, associado à cada realização do vetor de incertezas, ou seja, para nó da árvore de cenários $n \in \Omega_\tau$. Como a função objetivo do problema busca a minimização dos custos em termos de valor esperado, é possível que a função $h_\tau(x_\tau^{(n)})$ obtida para um determinado nó n apresente valores

indesejados ou, até mesmo, considerados “arriscados” para o problema em questão.

Um abordagem direta para restringir tais valores seria representar uma restrição de forma a impor diretamente um limite para esta função *individualmente* para todos os estados possíveis, ou seja,

$$h_\tau(x_\tau^{(n)}) \leq \eta, \quad n \in \Omega_\tau. \quad (5.2)$$

Contudo, esta abordagem pode ser demasiadamente conservadora, levando à uma solução com custos bastante elevados ou até mesmo inviável. Uma abordagem alternativa é a imposição de uma restrição que limite a função $h_\tau(\cdot)$ levando em consideração, *conjuntamente*, todas as possíveis ocorrências de $x_\tau^{(n)}$. Isto pode ser realizado através da inclusão de uma restrição de limite para o CVaR_α , com um nível de confiança α , associado à $h_\tau(\cdot)$:

$$\text{CVaR}_\alpha(h_\tau(x_\tau)) \leq \eta. \quad (5.3)$$

Na prática, como $h_\tau(\cdot)$ é calculada em função da variável aleatória x_τ , esta também é uma variável aleatória. O objetivo da restrição (5.3) não é controlar o valor da função para cada realização das incertezas mas o seu comportamento.

No caso do planejamento da operação de sistemas hidrotérmicos esta função pode representar, por exemplo, o déficit de energia em um determinado instante de tempo ou o déficit acumulado ao longo de um período estabelecido, conforme será apresentado adiante.

De acordo com a formulação linear para o CVaR (4.12), podemos escrever a restrição (5.3) como

$$b + \alpha^{-1} \sum_{n \in \Omega_\tau} p^{(n)} y_\tau^{(n)} \leq \eta \quad (5.4a)$$

$$y_\tau^{(n)} \geq h_\tau(x_\tau^{(n)}) - b, \quad n \in \Omega_\tau \quad (5.4b)$$

$$y_\tau^{(n)} \geq 0, \quad n \in \Omega_\tau. \quad (5.4c)$$

Na formulação (5.4) é adicionada a variável b , que representar o VaR associado ao CVaR_α . Esta variável é única e, por esta razão, é considerada uma variável de decisão relacionada ao nó raiz da árvore de incertezas. Além disto, são adicionadas variáveis auxiliares $y_\tau^{(n)}$ associadas à cada nó da árvore $n \in \Omega_\tau$ do estágio τ . A restrição (5.4a) é única e envolve todas as variáveis $y_\tau^{(n)}$, $n \in \Omega_\tau$. Por outro lado, existem $|\Omega_\tau|$ restrições auxiliares individuais para (5.4b) e (5.4c).

Rearrmando o problema, é possível escrever as restrições (5.4b) e (5.4c) na forma das restrições (5.1c) do problema original. Entretanto, o mesmo não ocorre para a restrição (5.4a) devido ao acoplamento direto entre todos os nós do estágio

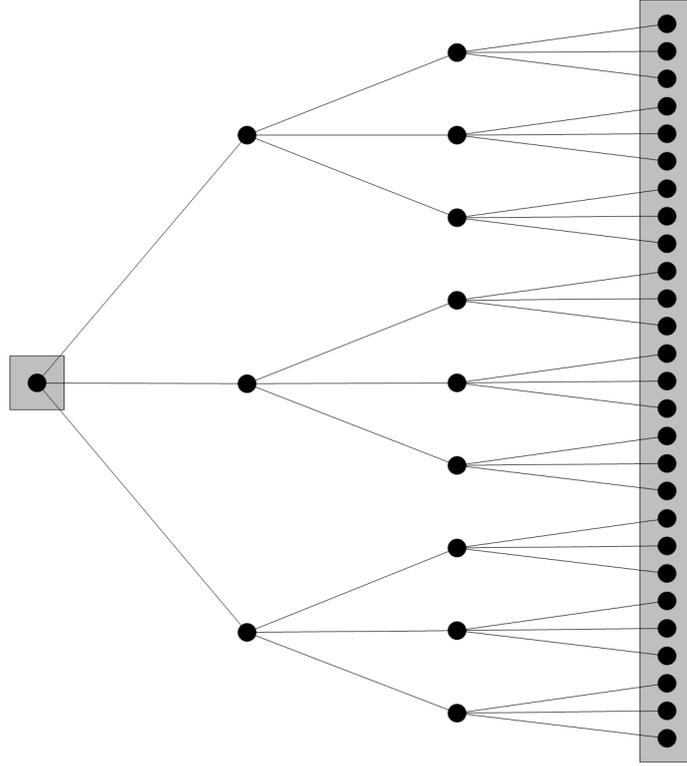


Figura 5.1: Árvore de cenários: exemplo de formulação com restrição de CVaR associada ao estágio 4.

τ , devido à presença de todas as variáveis $y_\tau^{(n)}$, conforme ilustrado na figura 5.1. Os nós destacados em cinza representam os nós que possuem variáveis envolvidas na restrição de limite de CVaR.

Finalmente, definimos como P o seguinte problema, composto pelo problema (5.1) considerando as variáveis e restrições adicionais (5.4) para representação do limite de CVaR.

$$P = \min \quad c_1^\top x_1 + \sum_{t=2}^T \sum_{n \in \Omega_t} p^{(n)} c_t^\top x_t^{(n)} \quad (5.5a)$$

$$\text{s.a.:} \quad A_1 x_1 = b_1 \quad (5.5b)$$

$$B_t x_{t-1}^{a(n)} + A_t x_t^{(n)} = b_t^{(n)} \quad n \in \Omega_t, \quad t = 2, \dots, T \quad (5.5c)$$

$$x_t^{(n)} \geq 0 \quad n \in \Omega_t, \quad t = 1, \dots, T \quad (5.5d)$$

$$b + \alpha^{-1} \sum_{n \in \Omega_\tau} p^{(n)} y_\tau^{(n)} \leq \eta \quad (5.5e)$$

$$y_\tau^{(n)} \geq h_\tau(x_\tau^{(n)}) - b, \quad n \in \Omega_\tau \quad (5.5f)$$

$$y_\tau^{(n)} \geq 0, \quad n \in \Omega_\tau \quad (5.5g)$$

A formulação P pode ser solucionada facilmente como um problema equivalente

determinístico, porém esta abordagem está limitada à solução de problemas de pequeno porte. Por outro lado, devido ao acoplamento da restrição (5.5e), esta formulação não pode ser solucionada por algoritmos baseados em decomposição como a PDDE.

5.3 Relaxação Lagrangeana

Para contornar esta dificuldade e permitir que o problema possa ser decomposto, propõe-se a dualização da restrição (5.5e). Neste caso, podemos definir a seguinte relaxação do problema P.

$$P(\mu) = \min \quad c_1^\top x_1 + \sum_{t=2}^T \sum_{n \in \Omega_t} p^{(n)} c_t^\top x_t^{(n)} - \mu \left(\eta - b - \alpha^{-1} \sum_{n \in \Omega_\tau} p^{(n)} y_\tau^{(n)} \right) \quad (5.6a)$$

$$\text{s.a.:} \quad A_1 x_1 = b_1 \quad (5.6b)$$

$$B_t x_{t-1}^{a(n)} + A_t x_t^{(n)} = b_t^{(n)} \quad n \in \Omega_t, \quad t = 2, \dots, T \quad (5.6c)$$

$$x_t^{(n)} \geq 0 \quad n \in \Omega_t, \quad t = 1, \dots, T \quad (5.6d)$$

$$y_\tau^{(n)} \geq h_\tau(x_\tau^{(n)}) - b \quad n \in \Omega_\tau \quad (5.6e)$$

$$y_\tau^{(n)} \geq 0 \quad n \in \Omega_\tau. \quad (5.6f)$$

Podemos observar que nesta formulação a restrição “complicada” é controlada através de sua adição na função objetivo. É considerado um novo termo que é penalizado por μ , onde μ é a variável dual ou *multiplicador de Lagrange* associado à restrição (5.5e).

Rearranjando a função objetivo e as restrições, temos a seguinte formulação

$$P(\mu) = \min \quad c_1^\top x_1 + \mu b - \mu \eta + \sum_{t=2}^T \sum_{n \in \Omega_t} p^{(n)} c_t^\top x_t^{(n)} + \mu \alpha^{-1} \sum_{n \in \Omega_\tau} p^{(n)} y_\tau^{(n)} \quad (5.7a)$$

$$\text{s.a.:} \quad A_1 x_1 = b_1 \quad (5.7b)$$

$$B_t x_{t-1}^{a(n)} + A_t x_t^{(n)} = b_t^{(n)} \quad n \in \Omega_t, \quad t = 2, \dots, T \quad (5.7c)$$

$$y_\tau^{(n)} \geq h_\tau(x_\tau^{(n)}) - b \quad n \in \Omega_\tau \quad (5.7d)$$

$$x_t^{(n)} \geq 0 \quad n \in \Omega_t, \quad t = 1, \dots, T \quad (5.7e)$$

$$y_\tau^{(n)} \geq 0 \quad n \in \Omega_\tau. \quad (5.7f)$$

Observe que se substituirmos x_1 por (x_1, b) e $x_\tau^{(n)}$ por $(x_\tau^{(n)}, y_\tau^{(n)})$ para $n \in \Omega_\tau$ é possível voltar à formulação original (5.1) onde o problema pode ser decomposto em subproblemas menores associados à cada nó da árvore de incertezas. Esta formulação é bastante atraente pois permite que $P(\mu)$ seja solucionado por algoritmos de decomposição como a PDDE.

O problema $P(\mu)$ é chamado de uma *relaxação Lagrangeana* (subproblema) de P , com parâmetro μ . Como $P(\mu)$ é uma relaxação de P , $P(\mu) \leq P$, ou seja, é um limite inferior para P . Para obter o melhor (maior) limite inferior considerando todos os possíveis valores de μ , devemos solucionar o problema denominado *Dual Lagrangeano*:

$$LD = \max_{\mu} \{P(\mu) : \mu \geq 0\}. \quad (5.8)$$

5.4 Algoritmo de solução

Para a solução do problema original (5.1) é necessária a solução do problema (5.8), explorando os possíveis valores de μ enquanto busca-se a maximização do resultado encontrado pelo problema interior $P(\mu)$. Cada valor encontrado para $P(\mu)$ é um limite inferior candidato.

A PDDE pode ser utilizada a solução do problema $P(\mu)$ em conjunto a um algoritmo de atualização da penalização para a solução dos problemas intermediários, até que a solução do problema original P (com restrição de CVaR) seja obtida.

Propõe-se utilizar um método de busca binária para o ajuste da penalização μ onde, a cada iteração, é solucionado um problema de PDDE até a otimalidade. A seguir, ilustraremos a idéia proposta para a atualização de μ de forma a encontrar a solução ótima do problema.

1. Inicialização:

a) Defina um limite inferior $\underline{\mu} = 0$, solucione o problema $P(\underline{\mu})$ através do algoritmo de PDDE e obtenha o $CVaR(\underline{\mu})$ através de uma simulação. Se $CVaR(\underline{\mu}) < \eta$, então pare, pois a restrição já está atendida para o menor valor de penalidade considerado;

b) Defina um limite superior $\bar{\mu}$, solucione o problema $P(\bar{\mu})$ através do algoritmo de PDDE e obtenha o $CVaR(\bar{\mu})$ através de uma simulação. Se $CVaR(\bar{\mu}) > \eta$, então pare e aumente a penalidade $\bar{\mu}$ e reinicie o processo;

c) Inicialize o número de iterações $k \leftarrow 0$;

2. Verificação da convergência: se $\bar{\mu} - \underline{\mu} \leq \epsilon$ então pare;

3. Faça o número de iterações $k \leftarrow k + 1$;

4. Faça $\mu \leftarrow 0.5(\underline{\mu} + \bar{\mu})$;

5. Solucione o problema $P(\mu)$ através do algoritmo de PDDE e obtenha o $CVaR(\mu)$ através de uma simulação.

6. Verifique se o $\text{CVaR}(\mu)$ da solução ótima encontrada viola o limite η :

a) Se viola, $\text{CVaR}(\mu) > \eta$, atualize o limite inferior: $\underline{\mu} \leftarrow \mu$

b) Se não viola, $\text{CVaR}(\mu) < \eta$, atualize o limite superior: $\bar{\mu} \leftarrow \mu$

7. Volte para o passo 2;

O algoritmo proposto acima pode ser visto como uma variação dos algoritmos tradicionais de atualização dos multiplicadores para relaxação Lagrangeana para o caso particular onde apenas uma restrição foi dualizada e, conseqüentemente, o vetor de multiplicadores de Lagrange é unidimensional. Nesta classe de algoritmos, a cada iteração são necessários basicamente dois passos: (i) cálculo do vetor (normalizado) que indica a direção de atualização do vetor de multiplicadores e (ii) cálculo do passo de atualização nesta direção. No caso de problemas lineares, a direção de atualização é dada pela folga da restrição dualizada.

No passo 6 é definida a direção para o passo de atualização do multiplicador, baseado na comparação do $\text{CVaR}(\mu)$ e do η . O $\text{CVaR}(\mu)$ é exatamente o lado esquerdo da equação dualizada (5.5e) enquanto que o η define o lado direito, ou seja, a comparação destes dois termos é equivalente à calcular a folga da restrição dualizada para obtenção da direção de atualização.

Entretanto, o cálculo do tamanho do passo é realizado de forma diferente, baseado no algoritmo da bisseção. Caso a restrição seja violada, o que indicaria que a folga é igual a zero, o limite inferior $\underline{\mu}$ é aumentado, o que significa que o multiplicador da próxima iteração será maior. Por outro lado, se a restrição não for violada, indicando uma folga maior que zero, o limite superior $\bar{\mu}$ é reduzido e, por consequência, o multiplicador da próxima iteração será menor.

Capítulo 6

Planejamento da operação com restrições de CVaR em um estágio com PDDE

6.1 Introdução

Como foi visto no capítulo anterior, é possível representar uma restrição de risco em problema de otimização com incertezas representadas através de uma estrutura de árvore de cenários. Nesta formulação, a representação direta da restrição de risco envolve variáveis de decisão de diferentes cenários, o que limita a aplicação de algoritmos baseados em decomposição, como é o caso da PDDE. Entretanto, também foi visto que é possível obter um problema separável através da dualização da restrição de risco e esta técnica se torna atrativa pois torna possível a aplicação da PDDE no problema dualizado.

O objetivo deste capítulo é desenvolver a formulação do problema dualizado para a aplicação de algoritmos de decomposição. A seção 6.2 apresenta a formulação das equações recursivas de programação dinâmica para o problema dualizado, enquanto a seção 6.3 apresenta o desenvolvimento das equações que definem os cortes de Benders, necessários para a aplicação do algoritmo de PDDE. A seção 6.5 apresenta uma contribuição conceitual importante baseada na interpretação do efeito que a restrição de CVaR exerce sobre os déficits de energia. Finalmente, a seção 6.6 descreverá um estudo de caso que ilustrará os conceitos abordados neste capítulo.

6.2 Programação Dinâmica

Considere que será definida uma restrição de CVaR aplicada ao déficit de energia em uma determinado etapa $\tau \in \mathbb{T}$. Como foi visto anteriormente, a dualização da res-

trição de CVaR permite que o problema seja representada de maneira desacoplada. Além disto, a estrutura dos problemas associados à cada estágio torna-se diferente. Teremos 4 tipos de problema, definidos por suas formulações P_A , P_B , P_C e P_D , e que são descritos a seguir:

P_A : Problema associado ao estágio de decisão “estratégica”. Com a dualização da restrição de CVaR, o VaR, representado pela variável de decisão b_τ , passa a ser considerado uma variável de estado e seu valor é compartilhado por todos os problemas associados ao estágio da etapa $t = \tau$. Sendo assim, a decisão do VaR é realizada no primeiro estágio de decomposição, considerado um *estágio de decisão estratégica*.

P_B : Problemas associados aos estágios anteriores à etapa da restrição de CVaR ($1 \leq t < \tau$). Neste grupo de problemas, as variáveis de decisão e restrições consideradas são idênticas às restrições do problema original (3.8) e a diferença consiste na presença da variável b_τ que aparece como novo argumento da função de custo futuro. Isto ocorre pois as decisões futuras são influenciadas pela decisão tomada no problema P_A .

P_C : Problemas associados ao estágio da etapa da restrição do CVaR ($t = \tau$). Neste problema, além das variáveis e restrições tradicionais do problema original (3.8), teremos também a variável auxiliar e restrição adicional relativa à representação linear do CVaR (4.12).

P_D : Problemas associados aos estágios posteriores à etapa da restrição de CVaR ($t > \tau$). Estes problemas são idênticos ao problema original (3.8).

A figura 6.1 apresenta uma representação do esquema de decomposição do problema com restrição de CVaR dualizado. Observe que a variável associada ao VaR, b_τ , é uma decisão associada ao problema P_A e seu valor é “repassado” recursivamente até o estágio da etapa $t = \tau$ através do acoplamento existente na função de custo futuro.

A seguir, a formulação do problema de planejamento da operação será detalhada, através da definição dos problemas decompostos, em função do estágio associado à restrição de CVaR. Os problemas serão apresentados no sentido inverso da decomposição, seguindo o processo de recursão (D, C, B, A).

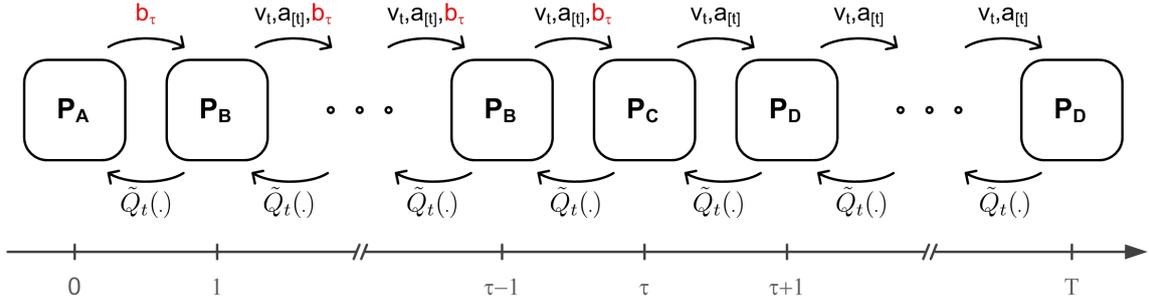


Figura 6.1: Diagrama do esquema de decomposição para o problema dualizado com restrição de CVaR no estágio τ

6.2.1 Estágios posteriores ao estágio da restrição de CVaR

Equação recursiva

Para estágios posteriores ao estágio da restrição de CVaR, ou seja, $\forall t, \tau < t \leq T$, temos o seguinte problema P_D de um estágio:

$$\begin{aligned} \tilde{Q}_t(v_{t-1}, a_{[t-1]}) = \\ \min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \tilde{Q}_{t+1}(v_t, a_{[t]}) \end{aligned} \quad (6.1a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (6.1b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (6.1c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (6.1d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (6.1e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (6.1f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (6.1g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (6.1h)$$

Neste caso, o problema (6.1) é idêntico ao problema (3.8) apresentado no capítulo 3 pois a variável de estado b_τ e as restrições e variáveis auxiliares para a representação do limite de CVaR estão presentes apenas nos problemas associados aos estágios onde $t \leq \tau$. Para $t > \tau$, o impacto da restrição de CVaR no estágio τ ocorre indiretamente através das demais variáveis de estado, neste caso, o volume armazenado nas hidrelétricas, pois a operação dos reservatórios pode ser alterada devido à restrição de CVaR.

Função de custo futuro

A função de custo futuro, representada por $\tilde{Q}_{t+1}(v_t, a_{[t]})$, é definida de modo similar à equação (2.22) como o valor esperado das realizações das incertezas futuras:

$$\tilde{Q}_{t+1}(v_t, a_{[t]}) = \sum_{l=1}^L p_l \times \tilde{Q}_{t+1}^l(v_t, a_{[t]}), \quad \forall t, \tau \leq t < T \quad (6.2)$$

onde p_l é a probabilidade de ocorrência do cenário l e, por definição, a função de custo futuro para o último estágio é definida como $\tilde{Q}_{T+1}(v_T, a_{[T]}) := 0$.

6.2.2 Estágio da restrição de CVaR

Equação recursiva

De modo similar, é possível definir o problema de otimização de um estágio P_C para o estágio da restrição de CVaR, $t = \tau$, como:

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau) = \\ & \min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \frac{\mu}{\alpha} y_\tau + \tilde{Q}_{t+1}(v_t, a_{[t]}) \end{aligned} \quad (6.3a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (6.3b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (6.3c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (6.3d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (6.3e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (6.3f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (6.3g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (6.3h)$$

$$y_\tau \geq r_t - b_\tau \quad \pi_t^b \quad (6.3i)$$

$$y_\tau \geq 0 \quad (6.3j)$$

Neste problema, além das variáveis e restrições definidas no problema original (3.8), é adicionada a variável auxiliar y_τ que representa o “excesso” do déficit de energia r_t (variável controlada pela restrição de CVaR) com relação à variável de estado b_τ , que representa a “franquia” para o déficit de energia. Também são adicionadas as restrições auxiliares (6.3i) e (6.3j) para o cálculo da função truncamento conforme apresentado na seção 4.6. Observe também que a variável de estado b_τ é um argumento da função para este estágio, indicando que a solução deste problema

está condicionada ao valor de b_τ .

Função de custo futuro

A função de custo futuro para este problema é calculada de forma idêntica à dos problemas dos estágios posteriores à restrição de CVaR, conforme indicado na equação (6.2). Isto ocorre pois as variáveis de estado para os estágios $\tau \leq t \leq T$ são as mesmas (volume e afluências).

6.2.3 Estágios anteriores ao estágio da restrição de CVaR

Equação recursiva

Para os estágios anteriores à restrição de CVaR, $1 \leq t < \tau$, o problema P_B é definido pelo problema (6.4), a seguir.

$$\begin{aligned} \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau) = \\ \min \quad & \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \tilde{Q}_{t+1}(v_t, a_{[t]}, b_\tau) \end{aligned} \quad (6.4a)$$

$$\text{s.a.:} \quad v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (6.4b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (6.4c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (6.4d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (6.4e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (6.4f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (6.4g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (6.4h)$$

Função de custo futuro

Observe que (6.4) é similar ao problema original de um estágio (3.8) sendo que a diferença consiste unicamente na dimensão aumentada da função de custo, devido à presença da variável de estado b_τ no custo do problema corrente, $\tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau)$, e também no custo futuro $\tilde{Q}_{t+1}(v_t, a_{[t]}, b_\tau)$.

A função de custo futuro é calculada então como

$$\tilde{Q}_{t+1}(v_t, a_{[t]}, b_\tau) = \sum_{l=1}^L p_l \times \tilde{Q}_{t+1}^l(v_t, a_{[t]}, b_\tau), \quad \forall t, 0 \leq t < \tau. \quad (6.5)$$

O impacto direto da variável de estado b_τ no problema do estágio τ (6.3) é

passado recursivamente no sentido inverso do tempo, desde o estágio τ até o estágio estratégico.

6.2.4 Estágio de decisão estratégica

Equação recursiva

Finalmente temos o estágio inicial, que não representa um estágio no tempo mas um estágio de “decisão estratégica”, onde está associada a decisão do valor da “franquia” b_τ e é formulado através do problema a seguir.

$$P_A : \min \quad \mu b_\tau - \mu \eta_\tau + \tilde{Q}_1(b_\tau) \quad (6.6a)$$

Como é possível observar, há um custo direto igual a μ para o incremento da “franquia” b_τ e sabemos que quanto maior o valor de b_τ menos restrito e, conseqüentemente, menor será o custo do problema (6.3). A decisão ótima para este problema consiste em encontrar o *trade-off* entre o termo μb_τ e a economia obtida nos estágios futuros, representada pela função de acomplamento $\tilde{Q}_1(b_\tau)$.

Adicionalmente, há uma parcela constante $-\mu \eta_\tau$ que está associada ao limite da restrição de CVaR. Apesar de não afetar diretamente na obtenção da solução ótima, este termo constante de ajuste é importante para o cálculo da função Lagrangeana associada do problema dualizado.

Função de custo futuro

Apesar de matematicamente idêntica à função de custo futuro, a função de acomplamento para o estágio estratégico representa o custo de operação total em função de decisão da franquia ótima b_τ . Sendo assim, pode ser calculada como

$$\tilde{Q}_1(b_\tau) = \sum_{l=1}^L p_l \times \tilde{Q}_1^l(b_\tau), \quad \forall t, 0 \leq t < \tau. \quad (6.7)$$

6.3 Programação Dinâmica Dual Estocástica

Como foi visto na seção 2.4, o algoritmo de PDDE baseia-se na construção, através de um procedimento iterativo, das funções de custo futuro associadas à cada estágio de decisão. As funções de custo futuro são calculadas na fase denominada recursão *backward* e são representadas como hiperplanos (cortes de Benders) calculados para determinados pontos de interesse das variáveis de estado (particularmente, para os pontos obtidos na fase de simulação *forward*).

Uma vez estabelecidos os subproblemas associados à cada estágio de decisão ((6.1), (6.3), (6.4) e (6.6)) e a expressão para o cálculo das funções de custo futuro

associadas a cada tipo de estágio, é possível determinar a formulação das funções de custo futuro $\tilde{Q}_{t+1}(\cdot)$. Para isto, é necessário definir os hiperplanos de maneira mais precisa, determinando-se as expressões que descrevem os coeficientes associados a cada variável de estado na FCF, conforme descrito a seguir.

6.3.1 Estágios posteriores ao estágio da restrição de CVaR

Equação recursiva

Considere inicialmente o problema (6.1), associado aos estágios posteriores à restrição de CVaR. Para estes problemas, a FCF é uma função das variáveis de estado v_t (volume final) e $a_{[t]}$ (sequência de afluências). Sendo assim, a FCF $\tilde{Q}_{t+1}(v_t, a_{[t]})$ para estes estágios pode ser substituída por um conjunto de hiperplanos e o problema pode ser reescrito como:

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}) = \\ & \min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \theta_{t+1} \end{aligned} \quad (6.8a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (6.8b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (6.8c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (6.8d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (6.8e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (6.8f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (6.8g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (6.8h)$$

$$\begin{aligned} \theta_{t+1} \geq & \theta_{t+1}^k + \sum_{i \in \mathbb{I}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]})}{\partial v_{t,i}^k} (v_{t,i} - v_{t,i}^k) + \\ & + \sum_{\substack{i \in \mathbb{I}, \\ p \in P_i}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]})}{\partial a_{t-p,i}^k} (a_{t-p,i} - a_{t-p,i}^k) \quad \pi_{t,k}^\theta \quad k \in \mathbb{K} \end{aligned} \quad (6.8i)$$

A equação (6.8i) descreve os cortes de Benders que representam a função de custo futuro considerando o conjunto de cortes \mathbb{K} , obtidos durante o processo iterativo para a aproximação corrente da FCF.

Função de custo futuro

A partir do problema (6.8), é possível extrair os coeficientes dos cortes para a FCF do estágio anterior. Estes coeficiente são calculados como a média das L realizações de incertezas associadas ao estágio imediatamente seguinte, conforme descrito a seguir.

O primeiro termo da expressão que compõe o corte é dado pelo termo independente θ_{t+1} , que representa o custo futuro total a partir do estágio $t + 1$. Este termo é calculado como

$$\theta_{t+1} = \sum_{l=1}^L p_l \times \tilde{Q}_{t+1}^l(v_t, a_{[t]}), \quad (6.9)$$

onde p_l é a probabilidade de ocorrência do cenário associado à abertura l .

O segundo termo é relativo ao coeficientes associados aos volumes finais e é calculado a partir da derivada do problema com relação à esta variável de estado. A derivada de (6.8) com relação ao volume inicial v_{t-1} consiste na variável dual da equação de balanço hídrico $\pi_{t,i}^h$ e, sendo assim, o coeficiente do corte é dado por

$$\frac{\partial \tilde{Q}_{t+1}(v_t, a_{[t]})}{\partial v_t} = \sum_{l=1}^L p_l \times \pi_{t+1,i}^h, \quad i \in \mathbb{I}. \quad (6.10)$$

Finalmente o último termo, associado às afluências anteriores, é calculado como a derivada de (6.8) com relação às afluências anteriores é composta de duas parcelas: a primeira corresponde à equação de transição das afluências (6.8c) e é dada pelo produto do coeficiente da afluência na equação e a variável dual correspondente; e a segunda parcela, associada aos cortes, é dada pelo produto dos coeficientes das afluências anteriores no corte e a variável dual das restrições dos cortes, conforme indicado na equação (6.11).

$$\begin{aligned} \frac{\partial \tilde{Q}_{t+1}(v_t, a_{[t]})}{\partial a_{t-p}} = \sum_{l=1}^L p_l \times & \left(\phi_{t+1,i,p} \pi_{t+1,i}^a + \right. \\ & \left. + \sum_{k \in \mathbb{K}} \sum_{p \in P_i} \frac{\partial \tilde{Q}_t^k(v_{t+1}, a_{[t+1]})}{\partial a_{t+1-p}^k} \pi_{t+1,k}^\theta \right), \quad i \in \mathbb{I} \quad (6.11) \end{aligned}$$

Pode-se ressaltar que a estrutura da FCF para os estágios $\tau < t \leq T$ é idêntica, composta das variáveis de estado associadas aos volume e afluências.

6.3.2 Estágio da restrição de CVaR

Equação recursiva

Por sua vez, o problema (6.3) associado ao estágio da restrição do CVaR, ou seja, onde $t = \tau$, pode ser reescrito como:

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau) = \\ & \min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \frac{\mu}{\alpha} y_\tau + \theta_{t+1} \end{aligned} \quad (6.12a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (6.12b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (6.12c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (6.12d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (6.12e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (6.12f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (6.12g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (6.12h)$$

$$y_\tau \geq r_t - b_\tau \quad \pi_t^b \quad (6.12i)$$

$$y_\tau \geq 0 \quad (6.12j)$$

$$\begin{aligned} \theta_{t+1} \geq & \theta_{t+1}^k + \sum_{i \in \mathbb{I}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]})}{\partial v_t^k} (v_{t,i} - v_{t,i}^k) \\ & + \sum_{\substack{i \in \mathbb{I}, \\ p \in P_i}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]})}{\partial a_{t-p}^k} (a_{t-p,i} - a_{t-p,i}^k) \end{aligned} \quad \pi_{t,k}^\theta \quad k \in \mathbb{K} \quad (6.12k)$$

Função de custo futuro

Além das variáveis de estado associadas aos volumes e afliências, o problema referente ao estágio da restrição do CVaR (6.12) também possui variável de estado associada à “franquia” b_τ . Esta variável aparece no lado direito da equação (6.12i) e deve ser considerada para o cálculo da função de custo futuro do estágio τ (utilizada no estágio $\tau - 1$). É possível obter a parcela associada à b_τ a partir do problema (6.12), conforme descrito na expressão a seguir:

$$\frac{\partial \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau)}{\partial b_\tau} = \sum_{l=1}^L p_l \times (-\pi_t^b), \quad t = \tau. \quad (6.13)$$

Este coeficiente corresponde à variação marginal do custo a partir do estágio $t = \tau$ com relação à um incremento da variável de estado b_τ . Observe que o sinal do coeficiente é negativo, indicando que haverá uma diminuição do custo caso b_τ seja incrementado. Os demais coeficientes, relacionados ao volume e às afluências são calculados pelas expressões (6.10) e (6.11), definidas anteriormente.

6.3.3 Estágios anteriores à restrição de CVaR

Equação recursiva

O problema (6.4) associado aos estágios anteriores ao estágio da restrição do CVaR é definido como:

$$\begin{aligned} \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau) = \\ \min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \theta_{t+1} \end{aligned} \quad (6.14a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (6.14b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (6.14c)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (6.14d)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (6.14e)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (6.14f)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (6.14g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (6.14h)$$

$$\begin{aligned} \theta_{t+1} \geq \theta_{t+1}^k + \sum_{i \in \mathbb{I}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]}, b_\tau)}{\partial v_t^k} (v_{t,i} - v_{t,i}^k) \\ + \sum_{\substack{i \in \mathbb{I}, \\ p \in P_i}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]}, b_\tau)}{\partial a_{t-p}^k} (a_{t-p,i} - a_{t-p,i}^k) \quad \pi_{t,k}^\theta \quad k \in \mathbb{K} \quad (6.14i) \\ + \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]}, b_\tau)}{\partial b_\tau^k} (b_\tau - b_\tau^k) \end{aligned}$$

Função de custo futuro

Como pode ser observado na restrição (6.14i), que define o corte de Benders nos problemas (6.14), há um novo termo relacionado à variável de estado b_τ . O coeficiente associado à este termo para a função do custo futuro em τ (utilizada no estágio $\tau - 1$) é calculado de acordo com a expressão (6.13).

Para os demais estágios anteriores, a variável de estado b_τ aparece apenas no

lado direito da expressão do corte (6.14i) e, sendo assim, o coeficiente associado à este termo para a FCF dos estágios $1 \leq t < \tau$ é definido como:

$$\frac{\partial \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau)}{\partial b_\tau} = \sum_{l=1}^L p_l \times \sum_{k \in \mathbb{K}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]}, b_\tau)}{\partial b_\tau^k} \pi_{t,k}^\theta \quad (6.15)$$

6.3.4 Estágio de decisão estratégica

Equação recursiva

Finalmente, a função de custo futuro do problema associado ao estágio estratégico pode ser substituída pelos cortes e o problema (6.6) pode ser reescrito como:

$$\min \quad \mu b_\tau - \mu \eta_\tau + \theta_1 \quad (6.16a)$$

$$\text{s.a.:} \quad \theta_1 \geq \theta_1^k + \frac{\partial \tilde{Q}_1^k(b_\tau)}{\partial b_\tau^k} (b_\tau - b_\tau^k) \quad \pi_{1,k}^\theta \quad k \in \mathbb{K} \quad (6.16b)$$

Função de custo futuro

A função de custo futuro utilizada no estágio estratégico representa, na verdade, o custo total de operação. Como foi visto, este custo é uma função apenas da variável de decisão b_τ e, por esta razão, a equação que representa o corte (6.16b) é composta por apenas uma parcela.

O cálculo deste coeficiente é realizado de maneira similar à (6.15), ou seja:

$$\frac{\partial \tilde{Q}_1(b_\tau)}{\partial b_\tau} = \sum_{l=1}^L p_l \times \sum_{k \in \mathbb{K}} \frac{\partial \tilde{Q}_1^k(b_\tau)}{\partial b_\tau^k} \pi_{1,k}^\theta. \quad (6.17)$$

6.4 Algoritmo de solução

Uma vez determinadas as expressões que definem os cortes de Benders associados à cada estágio do problema de planejamento da operação com restrição de CVaR dualizada, é possível aplicar o algoritmo de Programação dinâmica Dual Estocástica tradicional para a obtenção da política de operação ótima para uma determinada penalização μ . Para a obtenção da solução do problema original, propõe-se a utilização do algoritmo de atualização da penalização baseado em busca binária, proposto na seção 5.4.

6.5 Interpretação econômica da restrição de CVaR de um estágio

6.5.1 Análise direta

Após o processo de ajuste do Lagrangeano descrito na seção 6.4, a solução ótima determinará a política operativa que minimiza o custo total de operação e garante o critério de confiabilidade imposto através da restrição de limite do CVaR no estágio τ . Além disto, será determinado o valor ótimo da “franquia” b_τ^* e também a penalidade ótima μ^* .

Podemos interpretar o impacto da representação da restrição de CVaR analisando diretamente o problema P_C (6.12). Na solução ótima, observa-se que o custo associado à variável y_τ será de $\frac{\mu^*}{\alpha}$. Sabemos que as restrições (6.12i) e (6.12j) impõem que $y_\tau = \max(r_t - b_\tau^*, 0)$, ou seja, y_τ será maior que zero somente se o déficit de energia r_t exceder a franquia b_τ^* . Em resumo, quando o déficit energia r_t for menor que b_τ^* o custo aplicado ao déficit é igual a δ e quando o déficit exceder b_τ^* é aplicado adicionalmente um custo igual $\frac{\mu^*}{\alpha}$ sobre o excedente $r_t - b_\tau^*$.

Este resultado mostra um aspecto interessante que é a relação direta que a representação de uma restrição de CVaR em um estágio tem com função de custo de déficit de *dois* segmentos. O primeiro segmento representa o custo δ que é aplicado à qualquer nível de déficit e é denominado *custo econômico* do déficit. Por sua vez, o segundo segmento está associado ao critério de confiabilidade imposto pela restrição de CVaR e custo associado à este segmento pode ser interpretado como um *prêmio de risco*¹, ou seja, o custo adicional incorrido devido à decisão por uma operação com déficits considerados “mais arriscados” (acima da franquia).

A função de custo de déficit resultante é dada pela soma dos segmentos associados ao custo econômico e ao prêmio de risco e é denominada curva de custo de déficit implícito, conforme ilustrado na figura 6.2.

É possível estender a interpretação da curva de custo de déficit para um número maior de segmentos. Neste caso, cada segmento adicional corresponde à uma restrição de CVaR com um nível de confiabilidade α , limite máximo η_τ e penalização μ diferentes. Estes segmentos sobrepostos formam uma função de custo de déficit linear por partes, o que significa que é possível mapear a aversão a risco da sociedade através da imposição de limites máximos para o CVaR do déficit de energia através de uma função de custo de déficit linear por partes, como já é utilizado pelo setor elétrico.

¹ O termo *prêmio de risco* é utilizado no setor financeiro para designar a quantia dada pela diferença entre o rendimento de um investimento arriscado em comparação com um investimento seguro.

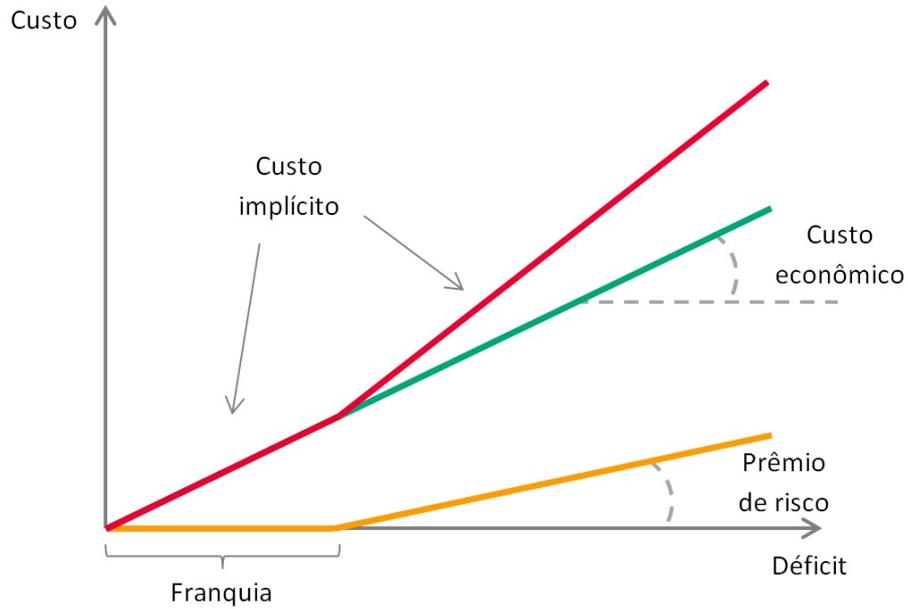


Figura 6.2: Curva de custo de déficit implícito

6.5.2 Reformulação do problema

Outra maneira de interpretar a aplicação da restrição de CVaR de um estágio é através de uma reformulação do problema. Para isto, considere a seguinte transformação de variáveis aplicada ao problema associado ao estágio da restrição de CVaR (6.12)

$$r_t = r_t^1 + r_t^2 \quad (6.18a)$$

$$y_\tau = r_t^2. \quad (6.18b)$$

Ou seja, o déficit de energia r_t é dividido em duas parcelas, r_t^1 e r_t^2 , onde a segunda parcela é definida igual ao excesso com relação à franquia y_τ .

O problema (6.12) pode ser reescrito como:

$$\tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_\tau) =$$

$$\min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t^1 + \left(\delta + \frac{\mu}{\alpha} \right) r_t^2 + \theta_{t+1} \quad (6.19a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (6.19b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (6.19c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (6.19d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (6.19e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (6.19f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (6.19g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (6.19h)$$

$$r_t = r_t^1 + r_t^2 \quad (6.19i)$$

$$r_t^1 \leq b_\tau \quad \pi_t^b \quad (6.19j)$$

$$r_t^2 \geq 0 \quad (6.19k)$$

$$\theta_{t+1} \geq \theta_{t+1}^k + \sum_{i \in \mathbb{I}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]})}{\partial v_t^k} (v_{t,i} - v_{t,i}^k)$$

$$+ \sum_{\substack{i \in \mathbb{I}, \\ p \in P_i}} \frac{\partial \tilde{Q}_{t+1}^k(v_t, a_{[t]})}{\partial a_{t-p}^k} (a_{t-p,i} - a_{t-p,i}^k) \quad \pi_{t,k}^\theta \quad k \in \mathbb{K} \quad (6.19l)$$

No problema resultante (6.19), a variável associada ao primeiro segmento de déficit r_t^1 tem custo na função objetivo igual à δ e é limitada ao valor máximo igual b_τ . Como a franquia b_τ é menor que o máximo corte de carga que, por sua vez, é menor que a demanda, sabemos que b_τ também é menor que a demanda total. Sendo assim, este limite pode ser visto como uma fração da demanda total que determina a profundidade do primeiro segmento de déficit e o objetivo da metodologia proposta pode ser interpretado como a determinação “profundidade ótima” b_τ^* desta curva.

Já a variável que representa o custo do segundo segmento de déficit têm valor na função objetivo igual à $(\delta + \frac{\mu}{\alpha})$, que corresponde exatamente à soma do custo econômico e o prêmio de risco e o prêmio de risco ótimo também é resultado da metodologia proposta, como foi visto anteriormente.

6.6 Resultados numéricos

Para exemplificar a metodologia proposta na seção anterior, foram realizados os primeiros experimentos através da aplicação algoritmo à uma instância do problema de planejamento da operação do sistema elétrico da Costa Rica considerando a configuração de Janeiro de 2013².

6.6.1 Descrição do sistema

O sistema elétrico da Costa Rica possui um parque gerador existente composto por 35 usinas hidrelétricas, 15 usinas termelétricas e 5 plantas eólicas, com capacidade instalada total de aproximadamente 4.255 MWmed (MW médios). A tabela 6.1 detalha, por fonte de geração, a capacidade instalada e a geração prevista para 2013 (em MWmed).

Tabela 6.1: Sistema da Costa Rica - capacidade instalada e geração prevista (2013)

Fonte	Capacidade Instalada	Geração
Hidrelétrica	3.245 (76%)	1.797 (83%)
Termelétrica	772 (18%)	261 (12%)
Eólica	238 (6%)	102 (5%)
TOTAL	4.255	2.160

A demanda média prevista para o ano de 2013 é de aproximadamente 2.160 MWmed e é representada por apenas 1 patamar. Adicionalmente, considera-se um custo econômico de déficit de apenas 1 segmento (linear) com valor igual a 200 \$/MWh.

Para cada estágio e cenário, são consideradas restrições de (i) atendimento à demanda total, (ii) balanço hídrico por usina hidrelétrica (considera-se uma representação individualizada das usinas) e (iii) limites operativos para hidrelétricas (volume e turbinamento) e termelétricas (geração). A função objetivo do problema é a minimização do valor esperado do custo de operação, composto pelos custos de geração térmica e penalizações devido à déficits de energia.

As incertezas associadas às afluições às usinas hidrelétricas são representadas por uma árvore de cenários implícita com $L = 5$ (aberturas backward a cada estágio), obtidas de um modelo independente (modelo autorregressivo periódico de ordem zero) ajustado a partir de dados históricos (1965 a 2010). Por simplicidade, a geração das plantas eólicas é suposta conhecida e é abatida da demanda, ou seja, a demanda considerada é a demanda líquida.

²Os dados de entrada deste problema foram modificados para ilustrar os conceitos abordados.

Para fins ilustrativos, será considerado um horizonte reduzido de apenas 3 etapas mensais³. O objetivo do horizonte reduzido é ter um problema de menor porte onde o número total de cenários da árvore de cenários é “tratável”. Neste caso, temos um número total de cenários $\mathcal{L} = 5 \times 5 \times 5 = 125$. Desta forma, é possível substituir o processo de geração de cenários por simulação Monte Carlo pela enumeração de todos os cenários, fazendo $M = \mathcal{L} = 125$, e obtendo assim a solução exata para esta instância da árvore de cenários.

6.6.2 Caso BASE

Primeiramente, foi avaliado o caso de referência (BASE) onde não se considera nenhuma restrição de CVaR. Sendo assim, a única medida adotada para conter déficits de energia é a penalização do déficit na função objetivo com um custo de 200 \$/MWh.

A tabela 6.2 apresenta os resultados médios de geração e déficit para as 3 etapas. Observa-se que nas etapas 2 e 3 não foi possível atender a demanda com o uso de geração hidrelétrica ou térmica e foi necessário “acionar” o déficit de energia. O déficit *médio* foi de 2,8 MWmed na etapa 2, o que representa aproximadamente 0,13% da demanda da etapa 2, e 24,7 MWmed na etapa 3, o que representa aproximadamente 1,18% da demanda desta etapa.

Tabela 6.2: Sistema da Costa Rica - Caso BASE: geração média (MWmed)

Etapa	Hidro	Térmica	Déficit	Demanda
1	1.774,9	231,3	-	2.006,2
2	1.858,1	250,6	2,8	2.111,5
3	1.764,2	300,3	24,7	2.089,2

Se realizarmos uma análise mais detalhada do déficit para os 125 cenários da etapa 3, temos que 103 cenários (82,4%) não levam a corte de carga, 19 cenários (15,2%) levam a um corte de carga menor que 10% da demanda e 3 cenários mais extremos (2,4%) levam a um corte de carga maior que 10% da demanda, conforme ilustrado no histograma da figura 6.3.

Vale ressaltar a grande diferença entre os valores de déficit e o valor médio do déficit de energia na etapa 3. O valor esperado do déficit é de 1,18% da demanda, enquanto os valores extremos da cauda da distribuição indicam que o déficit pode ser de 5 a 15 vezes maior que o valor médio. Isso ocorre devido a presença de muitos

³É considerada uma função de custo futuro para a terceira etapa obtida de um estudo de 12 anos com etapas mensais com o propósito de considerar um custo para utilização da água na ultimo etapa (caso contrário a política operativa ótima seria esvaziar todos os reservatórios).

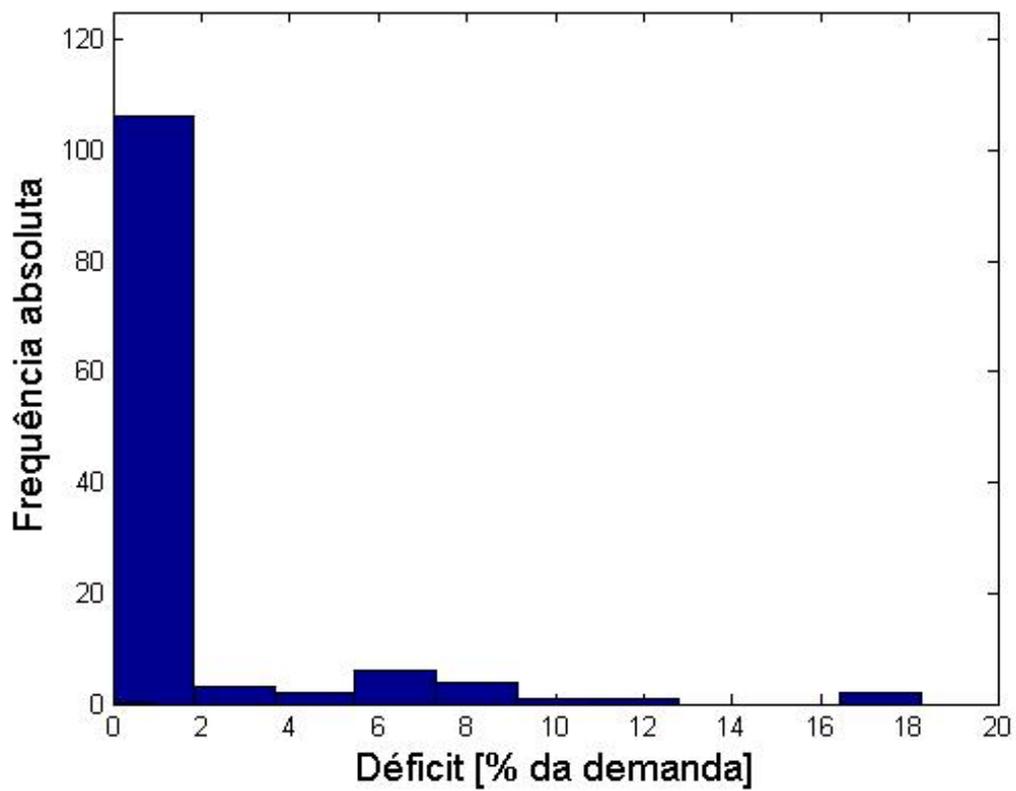


Figura 6.3: Sistema da Costa Rica (caso base): histograma do déficit de energia para a etapa 3

cenários com déficit igual a zero e que “compensam” (e diluem) os valores extremos. A tabela 6.3 apresenta os valores das medidas risco, EENS (valor médio) e CVaR para os níveis de confiança α iguais a 4% e 8%, por etapa. Observe que, como esperado, o CVaR reflete de maneira mais adequada o valor dos déficit extremos.

Tabela 6.3: Sistema da Costa Rica - Caso BASE: medidas de risco

Medida	Etapa 1	Etapa 2	Etapa 3
Risco	-	4%	17,6%
EENS	-	0,13%	1,18%
CVaR _{8%}	-	1,64%	10,27%
CVaR _{4%}	-	3,29%	13,16%

6.6.3 Caso com restrição de CVaR de um estágio

Em seguida, foi adicionada ao problema do caso base uma restrição que limita o CVaR_{8%} associado ao déficit de energia da etapa 3 à um valor máximo de 3,0% da demanda da etapa 3, ou seja, $\eta_3 = 3,0\%$. Este caso é denominado de caso CVaR(3).

Foram definidos como intervalos iniciais do método da bisseção um valor mínimo para penalização $\underline{\mu} = 0$ \$/MWh e um valor máximo igual a $\bar{\mu} = 100$ \$/MWh. O algoritmo encontrou a solução ótima após 12 iterações e o valor da penalização ótima foi igual a $\mu^* = 6,93$ \$/MWh e CVaR_{8%} = 2,89% da demanda da etapa 3, abaixo do limite máximo imposto de 3,0%, como esperado.

A tabela 6.4 mostra o processo de convergência do método da bisseção para a obtenção da penalidade ótima e o gráfico da figura 6.4 ilustra a evolução da penalidade a cada iteração.

O gráfico da figura 6.5 ilustra a variação da franquia (VaR_{8%}) e do CVaR_{8%} em função da penalidade. A curva em azul ilustra a variação do CVaR_{8%} com relação à penalidade e a curva em verde o VaR_{8%}. Para facilitar a visualização, os pontos foram interligados por uma linha tracejada apesar de apenas os pontos referentes às penalidades obtidas no processo iterativo terem sido explicitamente calculados. Em outras palavras, as funções VaR_{8%}(μ) e CVaR_{8%}(μ) podem não ter comportamento linear por partes como sugerido no gráfico. Como esperado, é possível observar que tanto a franquia como o CVaR diminuem com o aumento da penalidade.

A partir dos valores da solução ótima para VaR_{8%} e para penalidade μ , é possível obter a curva de custo de déficit de 2 segmentos. O primeiro segmento da curva tem inclinação igual ao custo econômico do déficit $\delta = 200$ \$/MWh e uma profundidade igual à franquia VaR_{8%} = 0,72% da demanda. Já o segundo segmento tem inclinação igual à soma do custo econômico e do prêmio de risco, ou seja, $\delta + \frac{\mu}{\alpha} = 200 + \frac{6,93}{0,08} =$

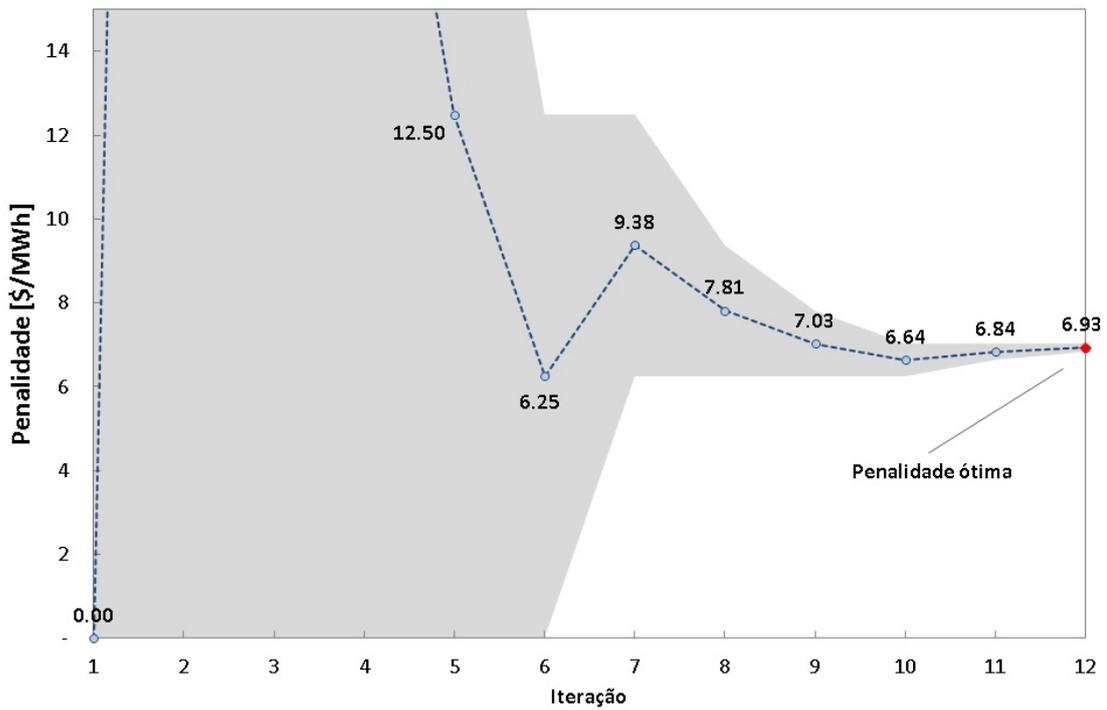


Figura 6.4: Sistema da Costa Rica - Caso CVaR(3): penalidade por iteração

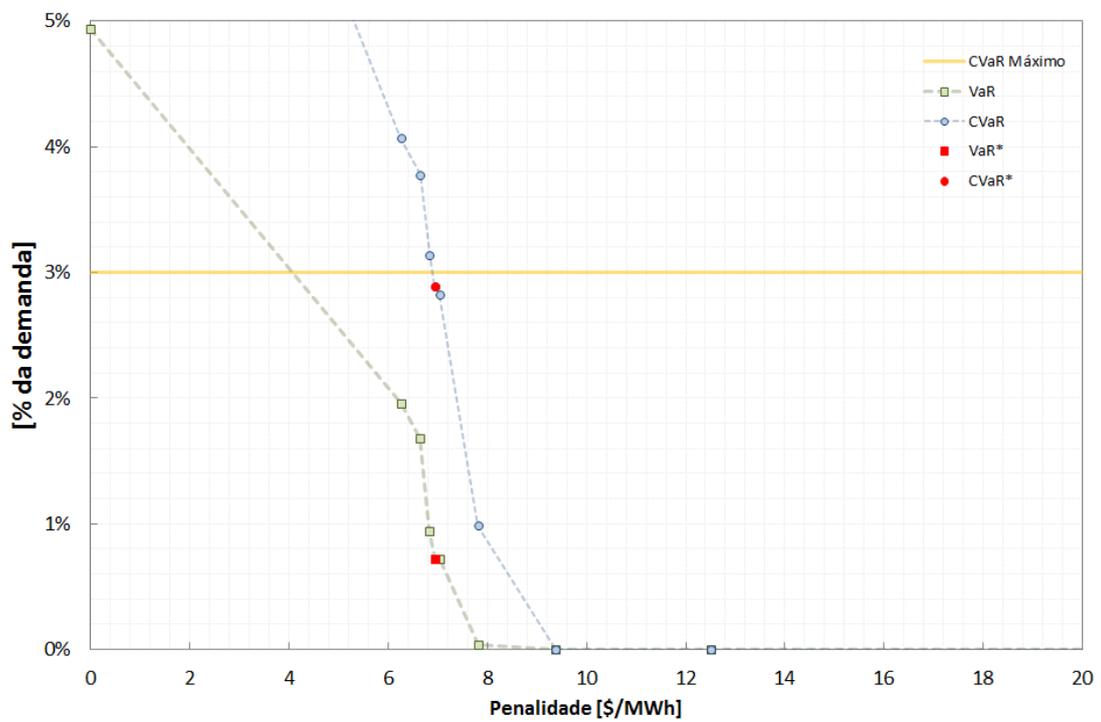


Figura 6.5: Sistema da Costa Rica: $VaR_{8\%}$ e $CVaR_{8\%}$ em função da penalidade

Tabela 6.4: Sistema da Costa Rica - Caso CVaR(3): tabela de convergência do método da bisseção

Iteração	Penalidade	VaR _{8%}	CVaR _{8%}	Custo total
1	0	4,94%	10,27%	17.372,34
2	100	0 %	0 %	13.540,90
3	50	0 %	0 %	15.828,80
4	25	0 %	0 %	16.972,75
5	12,5	0 %	0 %	17.544,73
6	6,25	1,96%	4,07%	17.766,70
7	9,375	0 %	0 %	17.687,72
8	7,8125	0,04%	0,99%	17.758,87
9	7,0313	0,72%	2,83%	17.773,94
10	6,6406	1,68%	3,78%	17.772,66
11	6,8359	0,94%	3,14%	17.774,18
12	6,9336	0,72%	2,89%	17.774,18

286,625 \$/MWh. A tabela 6.5 consolida os valores da curva de custo de déficit enquanto a figura 6.6 a ilustra graficamente.

Tabela 6.5: Sistema da Costa Rica - Caso CVaR(3): curva de custo de déficit de 2 segmentos

Segmento (% da demanda)	Custo (\$/MWh)
0,00% - 0,72%	200,0
0,72% - 100,00%	286,625

Além da variação do CVaR_{8%} em função da penalidade, também é possível observar o custo total como função da penalidade, ilustrado no gráfico da figura 6.7. Esta função corresponde ao Lagrangeano e é possível observar que a solução ótima é aquela que maximiza esta função.

A tabela 6.6 apresenta os resultados, em termos médios, para as 3 etapas para a penalidade ótima. É possível observar que houve um acréscimo na geração de energia total na etapa 3, o que acarretou na diminuição do déficit de energia da etapa 3. Entretanto, este aumento foi impulsionado pelo aumento da geração hidrelétrica, que ocorreu às custas da transferência da energia hidrelétrica da etapa 2 para etapa 3, o que levou à um efeito colateral negativo que é a elevação dos déficits de energia na etapa 2.

A tabela 6.7 mostra as medidas de risco obtidas por etapa. Primeiramente, é possível observar que houve uma redução significativa das medidas de risco na etapa 3, indicando que há uma menor ocorrência de déficits (menor risco) e que a

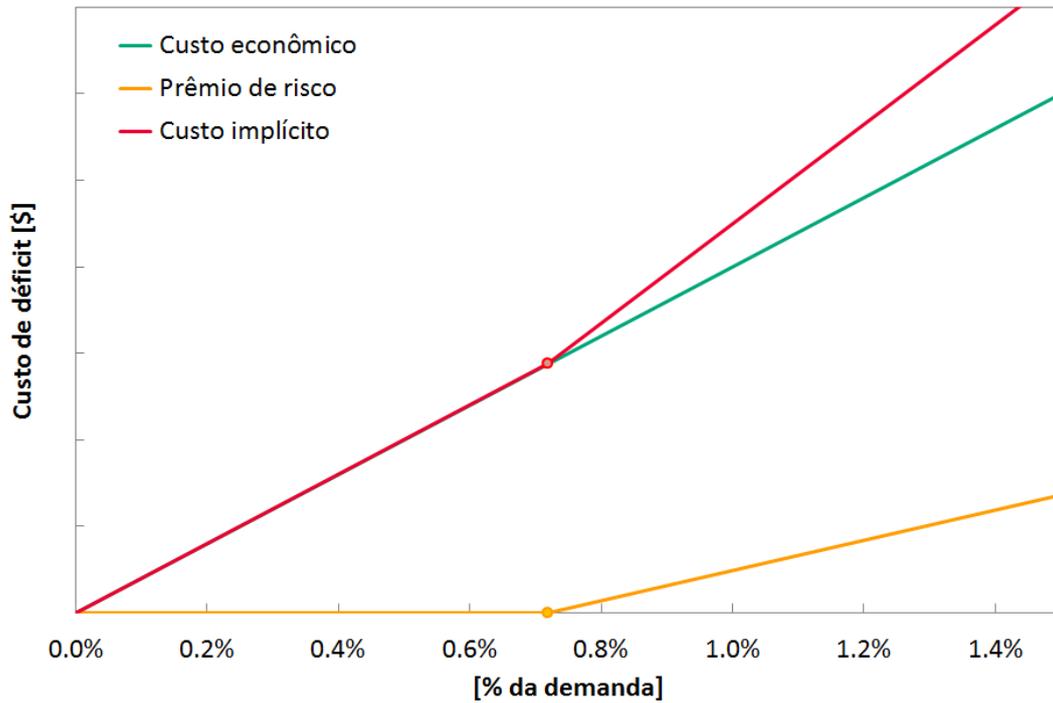


Figura 6.6: Sistema da Costa Rica - gráfico da curva de custo de déficit de 2 segmentos

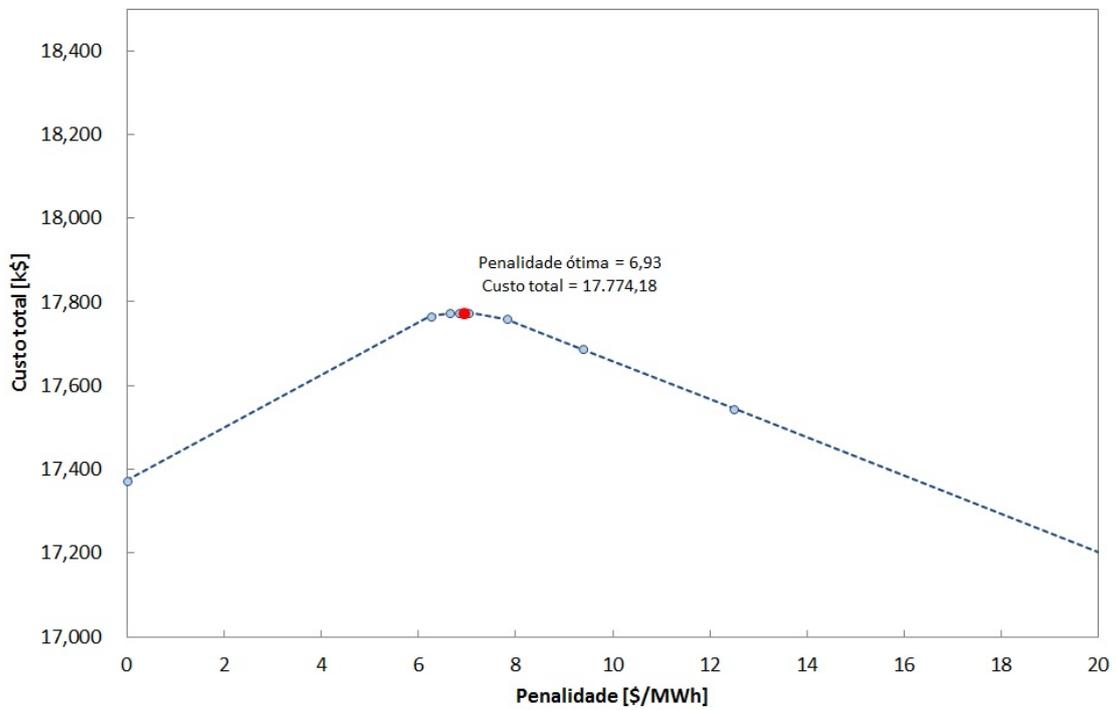


Figura 6.7: Sistema da Costa Rica - Caso CVaR(3): custo total em função da penalidade (Lagrangeano)

Tabela 6.6: Sistema da Costa Rica - Caso CVaR(3): geração média (MWmed)

Etapa	Hidro	Térmica	Déficit	Demanda
1	1.761,1	245,1	-	2.006,2
2	1.827,6	274,1	9,8	2.111,5
3	1.790,6	293,4	5,2	2.089,2

severidade destes déficits é também menor. Entretanto, exceto o risco de déficit da etapa 2 que permaneceu igual a 4%, todos os demais índices apresentaram valores mais severos, indicando que houve uma piora na confiabilidade de suprimento nesta etapa.

Tabela 6.7: Sistema da Costa Rica - Caso CVaR(3): medidas de risco

Medida	Etapa 1	Etapa 2	Etapa 3
Risco	-	4%	11,2%
EENS	-	0,46%	0,25%
CVaR _{8%}	-	5,78%	2,89%
CVaR _{4%}	-	11,56%	5,06%

Finalmente, os gráficos das figuras 6.8 e 6.9 apresentam o histograma do déficit nas etapas 2 e 3, respectivamente.

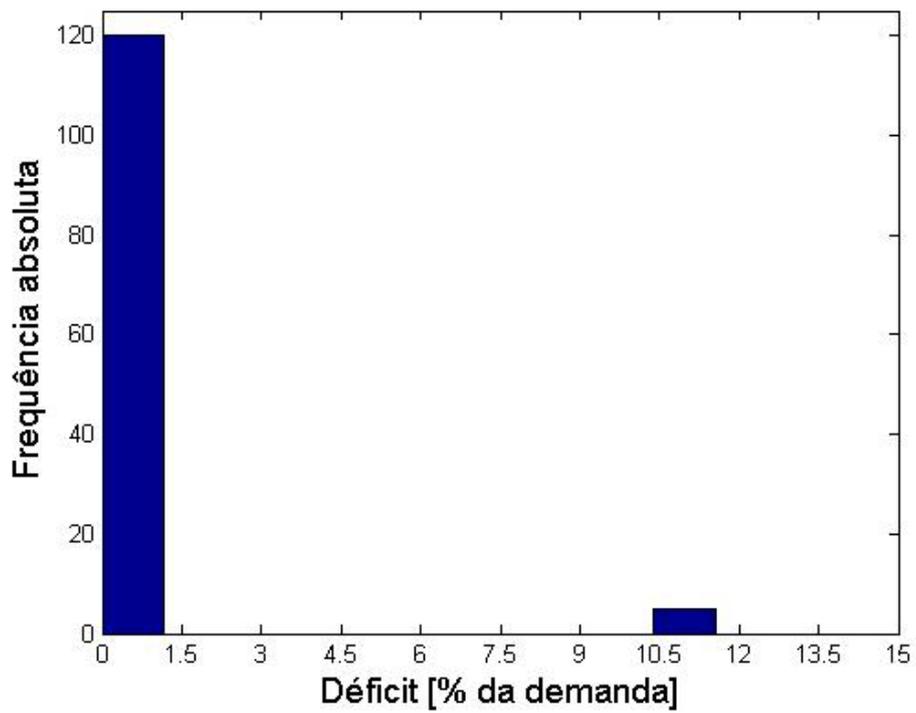


Figura 6.8: Sistema da Costa Rica - Caso CVaR(3): histograma do déficit de energia para a etapa 2

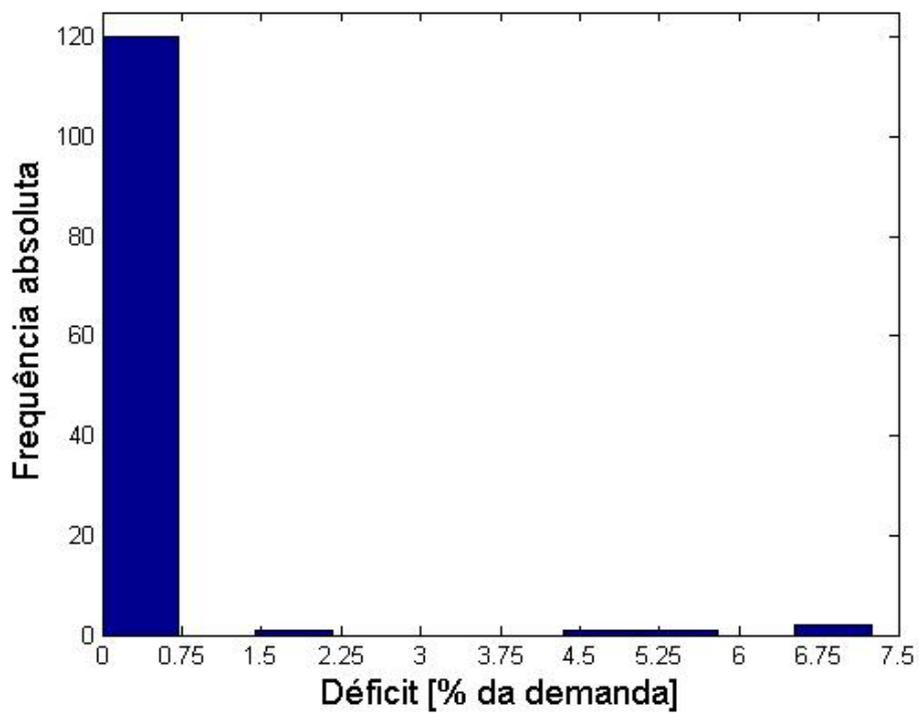


Figura 6.9: Sistema da Costa Rica - Caso CVaR(3): histograma do déficit de energia para a etapa 3

Capítulo 7

Planejamento da operação com restrições de CVaR multiestágio com PDDE

7.1 Introdução

No capítulo anterior foi apresentada uma metodologia para solução de problemas de planejamento da operação com restrição de CVaR aplicada em um único estágio. Apesar de conseguir limitar o risco no estágio de interesse, conforme esperado, foi visto que esta formulação pode levar a efeitos colaterais indesejados como a diminuição da confiabilidade no suprimento de energia dos demais estágios.

Neste capítulo, será apresentada uma extensão da formulação apresentada no capítulo anterior que tem como objetivo superar estas deficiências. A metodologia proposta consiste em considerar um limite de CVaR aplicado não ao déficit de energia em um determinado estágio, mas ao *déficit acumulado* ao longo de um período de interesse. Em termos matemáticos, será mostrado que a metodologia proposta neste capítulo é uma extensão da metodologia proposta no capítulo anterior.

Primeiramente, será apresentada uma formulação que é a extensão natural para o caso da aplicação da restrição de limite de CVaR para o déficit acumulado multiestágio. Em seguida, será proposta uma reformulação para um modelo equivalente que tem propriedades mais atraentes e será a considerada neste trabalho.

Finalmente, a seção 7.6 apresentará os resultados numéricos com a aplicação da metodologia proposta para o sistema da Costa Rica modificado, mesmo caso considerado na seção 6.

7.2 Formulação matemática para o déficit acumulado

7.2.1 Equivalente determinístico

Seja $\mathcal{T} = [\tau_1, \dots, \tau_2]$, $\tau_1 \in \mathbb{T}$, $\tau_2 \in \mathbb{T}$, $\tau_1 \leq \tau_2$ um intervalo de tempo para o qual estamos interessados em calcular o déficit de energia acumulado e limitar a sua ocorrência.

Primeiramente, é necessário formulá-lo matematicamente. O déficit acumulado no intervalo \mathcal{T} , representado por $R_{[\mathcal{T}]}$, pode ser expresso como:

$$R_{[\mathcal{T}]} = \sum_{t \in \mathcal{T}} r_t. \quad (7.1)$$

ou seja, pode ser representado diretamente como a soma dos déficit ocorridos em cada estágio no intervalo $t \in \mathcal{T}$.

Em um modelo equivalente determinístico, bastaria inserir a equação (7.1) na formulação para calcular o déficit acumulado em função das decisões operativas. Nesta formulação, a adição de um limite de CVaR seria direta, sendo necessária apenas definir a função que deseja-se controlar ($h_\tau(\cdot)$) como o déficit acumulado na restrição que calcula o excesso com relação ao VaR (5.4b).

Neste caso, logicamente, as variáveis que representam o excesso y_τ e a franquia b_τ passam a ter outro significado (se referem ao déficit acumulado), porém a formulação matemática do CVaR se mantém.

7.2.2 Programação dinâmica

Como visto anteriormente, a programação dinâmica permite decompor o problema em problemas menores e a partição do problema de planejamento da operação ocorre de acordo com os estágios de tempo. Esta característica faz com que não seja possível representar diretamente a equação (7.1), pois a mesma envolve variáveis de déficit r_t de estágios diferentes.

A formulação do déficit acumulado pode ser obtida então com a adição de uma variável de estado auxiliar que contabiliza o déficit acumulado desde a etapa τ_1 até a etapa corrente t e será representada por R_t . No estágio τ_1 esta variável é inicializada em zero e, a cada estágio $t \in \mathcal{T}$, além da nova variável de estado será adicionada uma restrição de “balanço de déficit”, que tem como objetivo acumular o déficit, conforme descrito a seguir.

$$R_t = R_{t-1} + r_t, \quad t \in \mathcal{T} \quad (7.2)$$

onde, por definição, fazemos $R_{\tau_1-1} := 0$.

7.3 Planejamento da operação com restrições de CVaR no déficit acumulado - formulação direta

Similarmente ao problema de planejamento da operação com restrição de CVaR aplicada ao déficit de apenas um estágio, apresentada no capítulo 6, a sequência de problemas solucionados varia em função do estágio.

A principal diferença com relação à formulação anterior consiste no problema P_C , associado ao estágio da restrição de CVaR, que será estendido e substituído por uma sequência de problemas para os estágios $t \in \mathcal{T}$. Os problemas P_A , P_B e P_D , que definem respectivamente os problemas para os estágios de planejamento estratégico, anterior e posterior ao estágio da restrição de CVaR (neste caso, anterior e posterior ao intervalo do déficit acumulado), têm a sua formulação mantida.

Estes problemas são descritos a seguir:

P_A : Problema associado ao estágio de decisão “estratégica”, com formulação idêntica à formulação do problema P_A no caso do CVaR de um estágio.

P_B : Problemas associados aos estágios anteriores ao início do intervalo da restrição de CVaR ($1 \leq t < \tau_1$). Sua formulação permanece inalterada.

\tilde{P}_C : Grupo de problemas associados aos períodos pertencentes ao intervalo $t \in \mathcal{T}$, análogos ao estágio da restrição do CVaR na formulação anterior. Dentro desta sequência de problemas, será considerada a variável de estado adicional que representa o déficit acumulado e será aplicado o limite de CVaR sobre esta variável. Os subproblemas associados a este grupo são dados por:

\tilde{P}_C^1 : Problemas associados ao estágio inicial do intervalo \mathcal{T} ($t = \tau_1$). Neste estágio, é inicializada a variável de estado que contabilizará o déficit acumulado no intervalo o efeito da decisão de déficit neste estágio nos estágios futuros e na restrição de CVaR dualizada será indicado através do acoplamento com a função de custo futuro.

\tilde{P}_C^i : Problemas associados aos estágios intermediários do intervalo, $\tau_1 < t < \tau_2$. Para estes problemas, a variável de estado de déficit acumulado é atualizada como a soma do déficit acumulado até o estágio anterior e o déficit do estágio corrente, conforme indicado na equação 7.2, além do acoplamento com a função de custo futuro.

\tilde{P}_C^2 : Problemas associados aos estágio final do intervalo, ou seja, $t = \tau_2$. São adicionadas a variável auxiliar y_{τ_2} e equações que representam o “excesso”

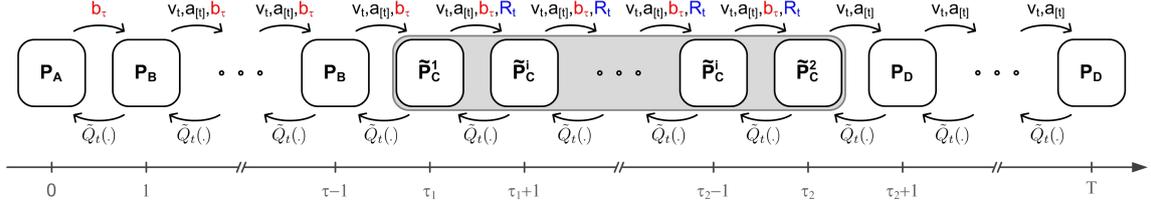


Figura 7.1: Diagrama do esquema de decomposição para o problema dualizado com restrição de CVaR para o déficit acumulado no intervalo \mathcal{T}

do CVaR. Além disto, este estágio é o último para o qual se contabiliza o déficit acumulado e, por esta razão, esta variável não está acoplada aos estágios seguintes através da função de custo futuro.

P_D : Problemas associados aos estágios posteriores ao intervalo da restrição de CVaR ($t > \tau_2$). Estes problemas têm sua formulação mantida e são idênticos ao problema original (3.8).

O diagrama da figura 7.1 ilustra a sequência de problemas de um estágio que define a estrutura do problema de planejamento com restrição de CVaR no déficit acumulado, além de ilustrar as variáveis de estados e seu acoplamento no tempo. O grupo de problemas \tilde{P}_C é representado em cinza e é possível observar, dentro deste grupo, os diferentes problemas associados aos estágios no intervalo $t \in \mathcal{T}$. É possível notar também a presença da variável de estado adicional R_t que representa o déficit acumulado e esta variável é considerada uma variável de estado “interna” ao grupo de problemas \tilde{P}_C .

A seguir, serão apresentadas as formulações dos problemas que descrevem o planejamento da operação com restrições de CVaR que limitam o déficit acumulado em um intervalo de tempo \mathcal{T} .

7.3.1 Estágio de decisão estratégica

O problema associado ao estágio inicial possui mesma a formulação que a apresentada no problema (6.6). A diferença ocorre no significado da franquia b_{τ_2} , que passa a ser interpretada como a “franquia” de déficit acumulado ao longo do intervalo \mathcal{T} .

7.3.2 Estágios anteriores ao intervalo de déficit acumulado

Os problemas associados aos estágios associados ao $1 \leq t < \tau_1$ também têm sua formulação mantida, de acordo com o problema (6.4).

7.3.3 Estágio inicial do intervalo do déficit acumulado

No estágio inicial do intervalo para o qual desejamos controlar o déficit acumulado ($t = \tau_1$), é adicionada uma nova variável R_t com relação ao problema original (3.8), com o objetivo de contabilizar o déficit total no período. Neste estágio, também é adicionada uma equação que “inicializa” o déficit acumulado com valor igual ao déficit do estágio t .

$$\tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_{\tau_2}) =$$

$$\min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \tilde{Q}_{t+1}(v_t, a_{[t]}, b_{\tau_2}, R_t) \quad (7.3a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (7.3b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (7.3c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (7.3d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (7.3e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (7.3f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (7.3g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (7.3h)$$

$$R_t = r_t \quad (7.3i)$$

Observe na função objetivo (7.3a) que a função de custo futuro agora tem a sua dimensão aumentada pois existe um impacto direto nos problemas futuros em função desta decisão.

7.3.4 Estágio intermediário do intervalo do déficit acumulado

O problema definido para um estágio intermediário do intervalo do déficit acumulado P_C^i é semelhante ao problema do estágio inicial P_C^1 , porém considera o acoplamento

com a variável de estado do déficit acumulado até o estágio anterior R_{t-1} .

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_{\tau_2}, R_{t-1}) = \\ \min & \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \tilde{Q}_{t+1}(v_t, a_{[t]}, b_{\tau_2}, R_t) \end{aligned} \quad (7.4a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (7.4b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (7.4c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (7.4d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (7.4e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (7.4f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (7.4g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (7.4h)$$

$$R_t = R_{t-1} + r_t \quad \pi_t^R \quad (7.4i)$$

A função de custo futuro utilizada nos problemas P_C^i assim como a utilizada no problema P_C^1 possuem um termo adicional associado à variável de estado R_t . Esta variável está presente na equação (7.4i) e este termo é calculado a partir da variável dual π_t^R .

7.3.5 Estágio final do intervalo do déficit acumulado

O problema de um estágio definido para último estágio do intervalo do déficit acumulado P_C^2 é análogo e bastante similar ao problema (6.3), referente ao estágio da restrição de CVaR aplicado ao déficit de um único estágio. Este problema é apre-

sentado em (7.5).

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_{\tau_2}, R_{t-1}) = \\ \min & \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \frac{\mu}{\alpha} y_{\tau_2} + \tilde{Q}_{t+1}(v_t, a_{[t]}) \end{aligned} \quad (7.5a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (7.5b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (7.5c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (7.5d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (7.5e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (7.5f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (7.5g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (7.5h)$$

$$R_t = R_{t-1} + r_t \quad \pi_t^R \quad (7.5i)$$

$$y_{\tau_2} \geq R_t - b_{\tau_2} \quad \pi_t^b \quad (7.5j)$$

$$y_{\tau_2} \geq 0 \quad (7.5k)$$

Como pode ser observado, as diferenças consistem na adição da variável de estado do déficit acumulado R_t e na substituição da variável déficit no estágio r_t por esta variável na equação do cálculo do excesso com relação franquia y_{τ_2} . Esta alteração no cálculo do excesso é equivalente à alterar a função $h_{\tau}(\cdot)$ controlada pelo função que calcula o déficit acumulado na restrição de CVaR do problema original (5.5), antes da dualização da restrição.

7.3.6 Estágios posteriores ao intervalo de déficit acumulado

Os estágios posteriores ao intervalo de déficit acumulado têm formulação idêntica à formulação apresentada no problema (6.1). Assim como no caso anterior, a influência da restrição de CVaR nestes estágios ocorre indiretamente através de alterações na política operativa dos volumes nos reservatórios.

7.4 Planejamento da operação com restrições de CVaR no déficit acumulado - formulação final

Na formulação apresentada na seção anterior, a variável que representa o excesso y_{τ_2} só é penalizada no estágio final do intervalo do déficit acumulado, mesmo que o

déficit acumulado já ultrapasse o valor de b_{τ_2} nos estágios intermediários, anteriores ao estágio τ_2 . Entretanto, sabemos que todos os cortes de carga que ocorrem no intervalo \mathcal{T} contribuem igualmente para o cálculo do déficit acumulado R_{τ_2} .

Esta penalização no estágio τ_2 sinaliza que este excesso não é desejável e, durante o processo de busca da solução ótima, esta indicação ocorrerá através da parcela adicional referente à variável de estado R_t nos cortes da função de custo futuro. A partir do conhecimento do algoritmo de PDDE, sabemos que este sinal será levado recursivamente ao longo das iterações, indicando aos estágios anteriores à τ_2 que o excesso de déficit será penalizado no futuro e, por esta razão, deve ser evitado.

Nesta seção, será proposta uma formulação alternativa à apresentada na seção anterior onde a representação do CVaR será alterada para uma representação equivalente e que também é uma extensão da reformulação proposta na seção 6.5.2 para o caso do CVaR de um estágio. Esta reformulação tem dois objetivos principais:

- Definir uma interpretação econômica para o modelo do CVaR aplicado ao déficit acumulado a partir da unificação dos conceitos já apresentados, e
- Sinalizar antecipadamente para o algoritmo de PDDE que há uma penalização associada em caso de o déficit acumulado exceder a franquia. A ideia é penalizar este excesso já a partir do estágio que ele ocorre, ao invés de “esperar” que este sinal seja trazido através dos cortes de Benders desde o estágio final do intervalo.

De forma análoga com o problema de um estágio, o problema multiestágio será reformulado, estendendo a separação da variável de déficit r_t em duas parcelas r_t^1 e r_t^2 para a variável de déficit acumulado R_t . Sendo assim, consideraremos a seguinte transformação de variáveis:

$$r_t = r_t^1 + r_t^2 \quad t \in \mathcal{T} \quad (7.6a)$$

$$R_t = R_t^1 + R_t^2 \quad t \in \mathcal{T} \quad (7.6b)$$

$$y_{\tau_2} = R_{\tau_2}^2 \quad (7.6c)$$

Além da separação do déficit do estágio em duas parcelas, como definido na seção 6.5.2, onde o déficit acumulado também é separado e passa a ser representado por duas parcelas: R_t^1 , que representa o déficit acumulado *abaixo* da franquia b_{τ_2} e R_t^2 , que representa o déficit acumulado que *excede* a franquia.

Estas novas variáveis acumulam o déficit ao longo dos estágios e, assim como o déficit acumulado, também são variáveis de estado. Como é possível notar, a segunda parcela acumulada R_t^2 corresponde exatamente à variável de excesso para $t = \tau_2$, conforme indicado na equação (7.6c).

Sendo assim, é possível reformular o problema apresentado na seção 7.3. A estrutura de problemas apresentada no diagrama da figura 7.1 segue inalterada e a formulação problemas P_A , P_B e P_D , referentes ao período fora do intervalo do déficit acumulado permanecerá a mesma. Entretanto, os problemas pertencentes ao intervalo do déficit acumulado \mathcal{T} , \tilde{P}_C^1 , \tilde{P}_C^i e \tilde{P}_C^2 serão redefinidos como \bar{P}_C^1 , \bar{P}_C^i e \bar{P}_C^2 , de acordo com o apresentado a seguir.

7.4.1 Estágio final do intervalo do déficit acumulado

Considere primeiramente o problema do último estágio do intervalo do déficit acumulado. Aplicando a transformação de variáveis especificada em (7.6), temos:

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_{\tau_2}, R_{t-1}^1, R_{t-1}^2) = \\ \min & \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \frac{\mu}{\alpha} R_t^2 + \tilde{Q}_{t+1}(v_t, a_{[t]}) \end{aligned} \quad (7.7a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (7.7b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (7.7c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (7.7d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (7.7e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (7.7f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (7.7g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (7.7h)$$

$$r_t = r_t^1 + r_t^2 \quad (7.7i)$$

$$R_t^1 = R_{t-1}^1 + r_t^1 \quad (7.7j)$$

$$R_t^2 = R_{t-1}^2 + r_t^2 \quad (7.7k)$$

$$R_t^1 \leq b_{\tau_2} \quad (7.7l)$$

Como pode ser observado, existe um acoplamento no cálculo das parcelas de déficit acumulado R_t^1 e R_t^2 com o déficit acumulado até o estágio anterior, representado pelas variáveis de estado R_{t-1}^1 e R_{t-1}^2 . Estas variáveis definem a condição inicial do déficit acumulado, conforme indicado nas equações (7.7j) e (7.7k). A variável correspondente ao primeiro segmento de déficit acumulado R_t^1 é limitada pelo valor da franquia na equação (7.7l), o que garante que qualquer excesso será contabilizado na variável r_t^2 e, conseqüentemente, em R_t^2 . Finalmente, podemos observar que a variável R_t^2 é penalizada na função objetivo representando o excesso de déficit acumulado ao final do intervalo.

De acordo com a equação de acoplamento (7.7k), R_t^2 pode ser escrita em função do déficit acumulado até o estágio anterior R_{t-1}^2 e da segunda parcela do déficit ocorrido no estágio corrente r_t^2 . Se substituirmos esta definição na função objetivo (7.7a), temos:

$$\min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \frac{\mu}{\alpha} (R_{t-1}^2 + r_t^2) + \tilde{Q}_{t+1}(v_t, a_{[t]}) \quad (7.8)$$

Observe no termo $\frac{\mu}{\alpha} (R_{t-1}^2 + r_t^2)$ que a parcela $\frac{\mu}{\alpha} R_{t-1}^2$ é constante pois R_{t-1}^2 não é uma variável de decisão do estágio t , mas uma variável de estado do estágio anterior. Sendo assim, podemos fazer uma outra manipulação e contabilizar este termo na função objetivo do estágio anterior. Aplicando esta manipulação recursivamente do estágio τ_2 até o estágio inicial do intervalo τ_1 , concluímos que é possível eliminar esta variável de estado do problema.

A função objetivo resultante pode então ser redefinida como:

$$\min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \frac{\mu}{\alpha} r_t^2 + \tilde{Q}_{t+1}(v_t, a_{[t]}) \quad (7.9)$$

e a equação (7.7k) pode ser eliminada do problema.

Por fim, substituindo a variável r_t pela soma de suas parcelas na função objetivo, temos:

$$\min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t^1 + \left(\delta + \frac{\mu}{\alpha} \right) r_t^2 + \tilde{Q}_{t+1}(v_t, a_{[t]}) \quad (7.10)$$

o que remete à mesma forma da curva de custo de déficit implícito apresentada na seção 6.5. O primeiro segmento de déficit acumulado é penalizado por um custo econômico δ e o segundo segmento, dado pelo excesso de déficit acumulado com relação à franquia, é penalizado pelo mesmo custo econômico somado à um “prêmio de risco” $\frac{\mu}{\alpha}$.

Levando em conta estas transformações podemos, finalmente, definir o problema

\bar{P}_C^2 para o ultimo estágio do intervalo \mathcal{T} como:

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_{\tau_2}, R_{t-1}^1) = \\ \min & \quad \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t^1 + \left(\delta + \frac{\mu}{\alpha} \right) r_t^2 + \tilde{Q}_{t+1}(v_t, a_{[t]}) \end{aligned} \quad (7.11a)$$

$$\text{s.a.:} \quad v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (7.11b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (7.11c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (7.11d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (7.11e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (7.11f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (7.11g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (7.11h)$$

$$r_t = r_t^1 + r_t^2 \quad (7.11i)$$

$$R_t^1 = R_{t-1}^1 + r_t^1 \quad \pi_t^R \quad (7.11j)$$

$$R_t^1 \leq b_{\tau_2} \quad \pi_t^b \quad (7.11k)$$

7.4.2 Estágio intermediário do intervalo do déficit acumulado

O problema referente aos estágios intermediários do intervalo \mathcal{T} , \bar{P}_C^i , é definido como:

$$\tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_{\tau_2}, R_{t-1}^1) =$$

$$\min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t^1 + \left(\delta + \frac{\mu}{\alpha} \right) r_t^2 + \tilde{Q}_{t+1}(v_t, a_{[t]}, b_{\tau_2}, R_t^1) \quad (7.12a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (7.12b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (7.12c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (7.12d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (7.12e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (7.12f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (7.12g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (7.12h)$$

$$r_t = r_t^1 + r_t^2 \quad (7.12i)$$

$$R_t^1 = R_{t-1}^1 + r_t^1 \quad \pi_t^R \quad (7.12j)$$

$$R_t^1 \leq b_{\tau_2} \quad \pi_t^b \quad (7.12k)$$

7.4.3 Estágio inicial do intervalo do déficit acumulado

O problema referente ao primeiro estágio do intervalo \mathcal{T} , \bar{P}_C^1 , é definido como:

$$\begin{aligned} & \tilde{Q}_t(v_{t-1}, a_{[t-1]}, b_{\tau_2}) = \\ & \min \sum_{j \in \mathbb{J}} c_j g_{t,j} + \delta r_t + \tilde{Q}_{t+1}(v_t, a_{[t]}, b_{\tau_2}, R_t^1) \end{aligned} \quad (7.13a)$$

$$\text{s.a.: } v_{t,i} = v_{t-1,i} + a_{t,i} - u_{t,i} - s_{t,i} + \sum_{m \in M_i} (u_{t,m} + s_{t,m}) \quad \pi_{t,i}^h \quad i \in \mathbb{I} \quad (7.13b)$$

$$a_{t,i} = \sum_{p \in P_i} \phi_{t,i,p} a_{t-p,i} + \xi_{t,i} \quad \pi_{t,i}^a \quad i \in \mathbb{I} \quad (7.13c)$$

$$e_{t,i} = \rho_i u_{t,i} \quad i \in \mathbb{I} \quad (7.13d)$$

$$v_{t,i} \leq \bar{v}_i \quad i \in \mathbb{I} \quad (7.13e)$$

$$u_{t,i} \leq \bar{u}_i \quad i \in \mathbb{I} \quad (7.13f)$$

$$g_{t,j} \leq \bar{g}_j \quad j \in \mathbb{J} \quad (7.13g)$$

$$\sum_{i \in \mathbb{I}} e_{t,i} + \sum_{j \in \mathbb{J}} g_{t,j} + r_t = d_t \quad (7.13h)$$

$$r_t = r_t^1 + r_t^2 \quad (7.13i)$$

$$R_t^1 = r_t^1 \quad (7.13j)$$

$$R_t^1 \leq b_{\tau_2} \quad \pi_t^b \quad (7.13k)$$

7.5 Algoritmo de solução

Da mesma forma que a formulação apresentada para a representação do CVaR aplicado ao déficit de um estágio, a formulação com CVaR associado ao déficit multiestágio também pode ser solucionada pelo algoritmo geral proposto na seção 5.4.

7.6 Resultados numéricos

Para ilustrar a aplicação da metodologia proposta, será utilizado o mesmo caso modificado da Costa Rica apresentado na seção 6.6. Este caso possui um horizonte de 3 etapas, que será dividido em apenas um intervalo de cálculo do déficit acumulado, ou seja, $\tau_1 = 1$, $\tau_2 = 3$ e, conseqüentemente, $\mathcal{T} = \{1, 2, 3\}$.

O cálculo do déficit acumulado é realizado a partir da soma dos déficits das 3 etapas para cada uma das séries consideradas, de forma a obter uma distribuição de 125 amostras de déficit acumulado. Além disto, como o déficit acumulado é calculado em termos percentuais cabe ressaltar que a base do cálculo considera a soma da demanda das 3 etapas.

Antes de avaliar a metodologia proposta, é interessante avaliar o déficit acumulado do caso base, assim como o seu comportamento após a inclusão do um limite de 3% aplicado ao CVaR do déficit da etapa 3, como será apresentado nas próximas seções.

7.6.1 Caso base

A tabela 7.1 apresenta novamente as medidas de risco calculadas para as etapas 1, 2 e 3 para o caso BASE da seção 6.6.2, sem nenhuma restrição de CVaR. Foi adicionada uma nova coluna a esta tabela que contém o cálculo das medidas de risco para o déficit acumulado.

Tabela 7.1: Sistema da Costa Rica - Caso BASE: medidas de risco

Medida	Etapa 1	Etapa 2	Etapa 3	Acumulado
Risco	-	4%	17,6%	18,4%
EENS	-	0,13%	1,18%	0,45%
CVaR _{8%}	-	1,64%	10,27%	3,95%
CVaR _{4%}	-	3,29%	13,16%	5,21%

De acordo com a tabela, o risco do déficit acumulado ser maior do que zero, ou seja, ocorrer de déficit em qualquer uma das três etapas, é igual a 18,4%. Comparando com o mesmo índice para as etapas 2 e 3, percebe-se que o risco acumulado é 0,8% maior (probabilidade de 1 série = $\frac{1}{125} = 0,8\%$) e pode-se concluir que há uma série em que ocorreu déficit na etapa 2 mas não ocorreu na etapa 3.

Vale ressaltar que a grandeza do índice de risco acumulado é sempre maior ou igual do que o risco individual de cada etapa pois ele representa a união de todos os eventos de déficit de cada etapa. Para que um cenário seja marcado como um cenário “com déficit” e contribua com o índice de risco, basta ocorrer um déficit de qualquer valor em pelo menos uma das etapas.

Por outro lado, as demais medidas de risco apresentadas são dadas em valores percentuais da demanda de referência. Para os índices acumulados, a demanda utilizada é a soma das demandas das 3 etapas e como não há déficits na etapa 1, o valor numérico dos índices acumulados é menor, como pode ser observado.

É importante observar esta característica na definição do critério de confiabilidade para o planejamento da operação que definirá os limites da restrição que deseja-se impor.

7.6.2 Caso com restrição de CVaR em um estágio

Tomando o caso base como referência, foram apresentados na seção 6.6.3 os resultados da aplicação da metodologia que considera uma restrição para limitar o $\text{CVaR}_{8\%}$ do déficit da etapa 3 a um valor máximo de 3% da demanda.

A tabela 7.2 apresenta as medidas de risco calculadas para as etapas 1, 2 e 3. Como apresentado, foi obtida uma política operativa com $\text{CVaR}_{8\%}$ associado ao déficit da etapa 3 igual a 2,89%. É possível observar adicionalmente uma nova coluna nesta tabela que contém o cálculo das medidas de risco para o déficit acumulado.

Tabela 7.2: Sistema da Costa Rica - Caso $\text{CVaR}(3)$: medidas de risco

Medida	Etapa 1	Etapa 2	Etapa 3	Acumulado
Risco	-	4%	11,2%	12,8%
EENS	-	0,46%	0,25%	0,23%
$\text{CVaR}_{8\%}$	-	5,78%	2,89%	2,79%
$\text{CVaR}_{4\%}$	-	11,56%	5,06%	4,7%

7.6.3 Caso com restrição de CVaR multiestágio

Foi adicionada ao problema do caso base uma restrição que limita o $\text{CVaR}_{8\%}$ associado ao déficit de energia acumulado à um valor máximo de 1,0% da demanda total, ou seja, $\eta_T = 1,0\%$. Este caso é denominado caso $\text{CVaR}(1-3)$.

Foram definidos como intervalos iniciais do método da bisseção um valor mínimo para penalização $\underline{\mu} = 0$ e um valor máximo igual a $\bar{\mu} = 100$ \$/MWh. O algoritmo encontrou a solução ótima após 11 iterações e o valor da penalização ótima foi igual a $\mu^* = 7,23$ \$/MWh e $\text{CVaR}_{8\%} = 0,99\%$ da demanda total, abaixo do limite máximo imposto de 1,0%, como esperado.

A tabela 7.3 mostra o processo de convergência do método da bisseção para a obtenção da penalidade ótima e o gráfico da figura 7.2 ilustra a evolução da penalidade a cada iteração.

O gráfico da figura 7.3 ilustra a variação da franquia ($\text{VaR}_{8\%}$) e do $\text{CVaR}_{8\%}$ em função da penalidade. A curva em azul ilustra a variação do $\text{CVaR}_{8\%}$ com relação à penalidade e a curva em verde o $\text{VaR}_{8\%}$. Assim como no gráfico do estudo de caso da seção 6.6.3, os pontos obtidos a cada iteração foram interligados por uma linha tracejada para facilitar a visualização, apesar de apenas os pontos associados às penalidades obtidas no processo iterativo terem sido explicitamente calculados. Em outras palavras, as funções $\text{VaR}_{8\%}(\mu)$ e $\text{CVaR}_{8\%}(\mu)$ podem não ter comportamento linear por partes como sugerido. Como esperado, é possível observar que tanto a franquia como o CVaR diminuem com o aumento da penalidade.

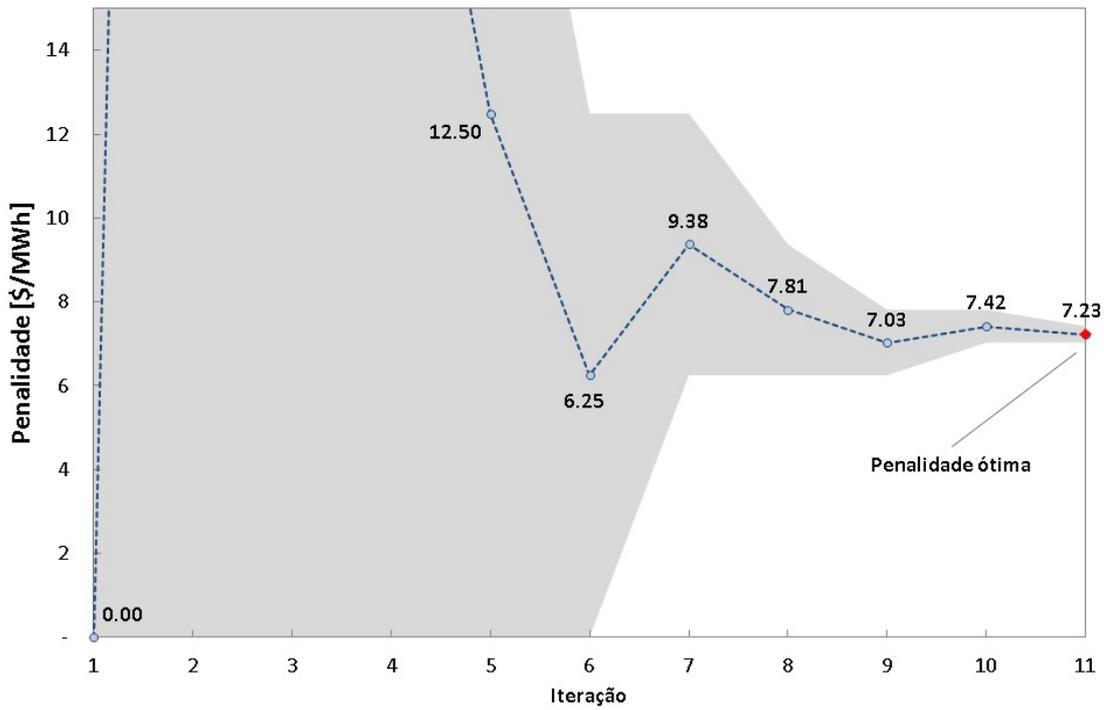


Figura 7.2: Sistema da Costa Rica - Caso CVaR(1-3): penalidade por iteração

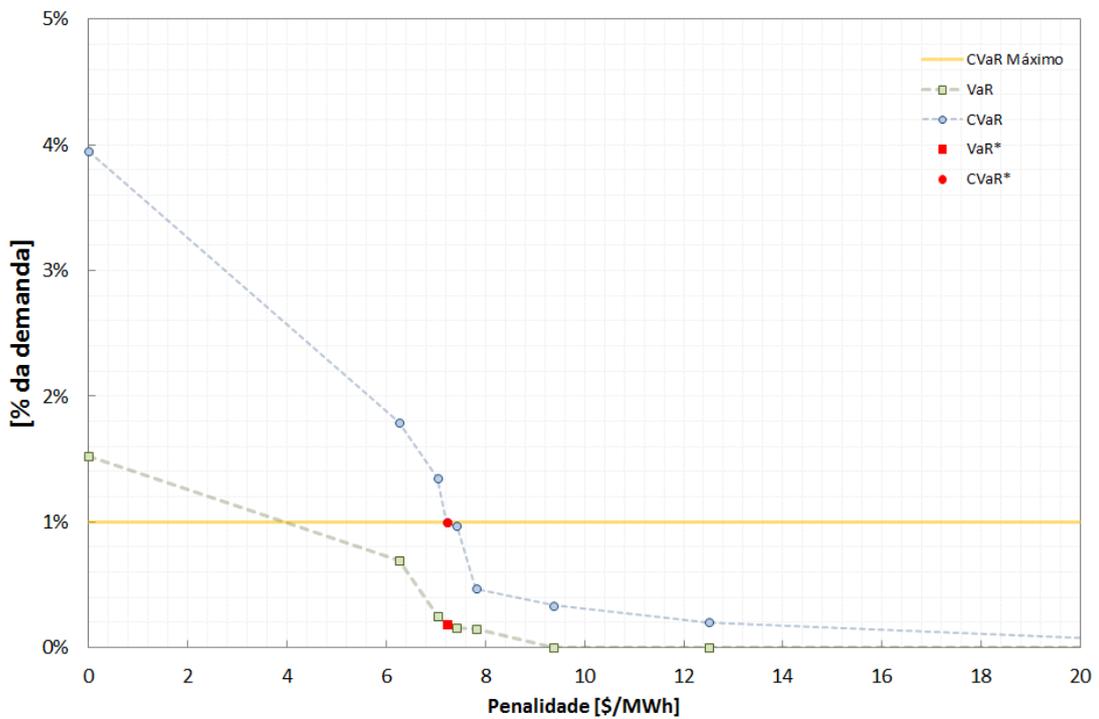


Figura 7.3: Sistema da Costa Rica - Caso CVaR(1-3): $VaR_{8\%}$ e $CVaR_{8\%}$ do déficit acumulado em função da penalidade

Tabela 7.3: Sistema da Costa Rica - Caso CVaR(1-3): tabela de convergência do método da bisseção

Iteração	Penalidade	VaR _{8%}	CVaR _{8%}	Custo total
1	0	1,53%	3,95%	17.372,34
2	100	0%	0%	13.998,57
3	50	0%	0%	16.231,49
4	25	0%	0%	17.347,94
5	12,5	0%	0,20%	17.868,64
6	6,25	0,70%	1,79%	17.989,95
7	9,375	0%	0,33%	17.966,32
8	7,8125	0,14%	0,47%	18.011,94
9	7,0313	0,25%	1,35%	18.013,42
10	7,4219	0,16%	0,97%	18.015,25
11	7,2266	0,19%	0,99%	18.015,42

A partir dos valores da solução ótima para VaR_{8%} e para penalidade μ , é possível obter a curva de custo de déficit de 2 segmentos. O primeiro segmento da curva tem inclinação igual ao custo econômico do déficit $\delta = 200$ \$/MWh e uma profundidade igual à franquia VaR_{8%} = 0,19% da demanda total. Já o segundo segmento tem inclinação igual à soma do custo econômico e do prêmio de risco, ou seja, $\delta + \frac{\mu}{\alpha} = 200 + \frac{7,23}{0,08} = 290,3$ \$/MWh. A tabela 7.4 consolida os valores da curva de custo de déficit enquanto a figura 7.4 a ilustra graficamente.

Tabela 7.4: Sistema da Costa Rica - Caso CVaR(1-3): curva de custo de déficit acumulado de 2 segmentos

Segmento (% da demanda)	Custo (\$/MWh)
0,00% - 0,19%	200,0
0,19% - 100,00%	290,3

Além da variação do CVaR_{8%} em função da penalidade, também é possível observar o custo total como função da penalidade (Lagrangeano), ilustrado no gráfico da figura 7.5. Assim como no caso do CVaR aplicado ao déficit de uma etapa, é possível observar que a solução ótima também é aquela que maximiza a função.

A tabela 7.5 apresenta os resultados para a penalidade ótima, em termos médios, para as 3 etapas. Observa-se que, em comparação com o caso CVaR(3), a geração de energia hidrelétrica foi reduzida na etapa 1 e armazenada para aumento de produção nas etapas 2 e 3. Esta energia armazenada foi suficiente eliminar completamente os déficits de energia na etapa 2 e ainda reduzir o déficit na etapa 3.

A tabela 7.6 mostra as medidas de risco obtidas por etapa. Primeiramente,

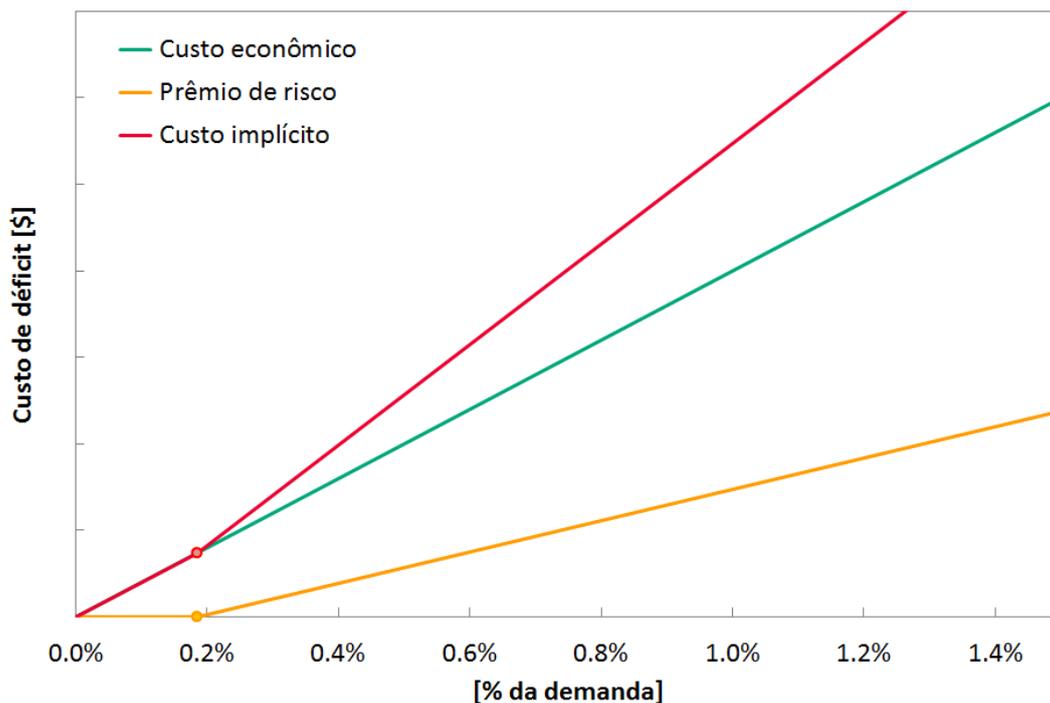


Figura 7.4: Sistema da Costa Rica - Caso CVaR(1-3): gráfico da curva de custo de déficit acumulado de 2 segmentos

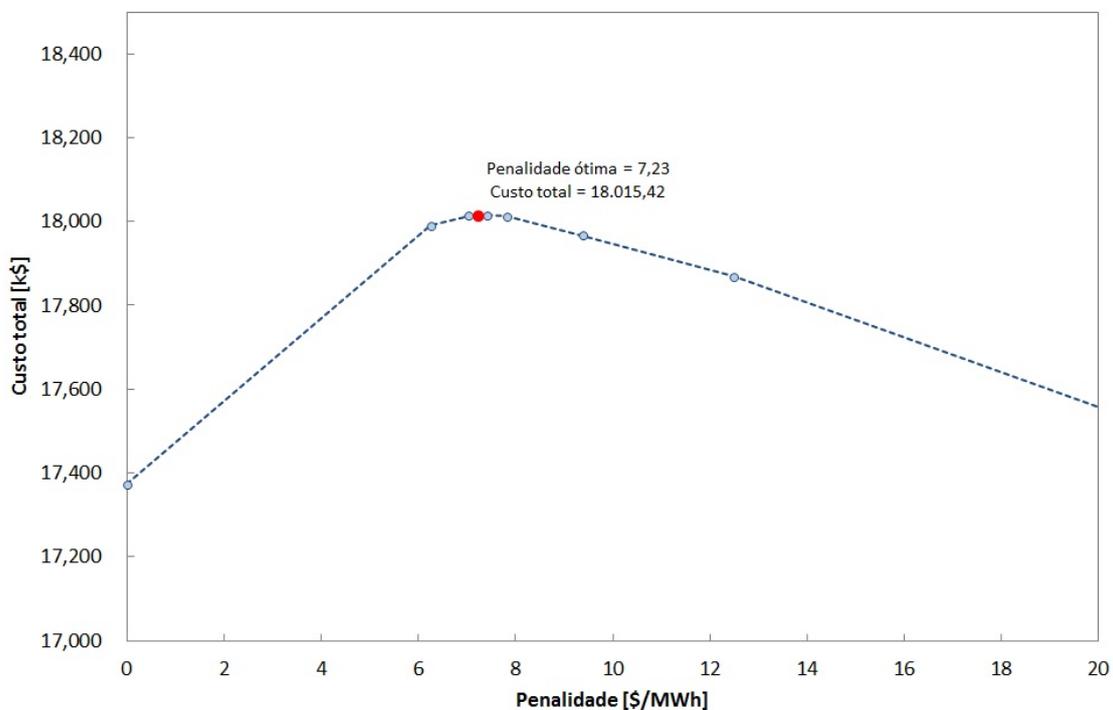


Figura 7.5: Sistema da Costa Rica - Caso CVaR(1-3): custo total em função da penalidade (Lagrangeano)

Tabela 7.5: Sistema da Costa Rica - Caso CVaR(1-3): geração média (MWmed)

Etapa	Hidro	Térmica	Déficit	Demanda
1	1.742,3	263,9	-	2.006,2
2	1.842,0	269,5	-	2.111,5
3	1.792,6	291,6	4,9	2.089,2

é possível observar que os índices da etapa 2 ficaram iguais à zero pois os déficits foram eliminados e que os índices associados à etapa 3 foram reduzidos. Finalmente, destaca-se que o $CVaR_{8\%}$ associado ao déficit acumulado foi igual à 0,99% e ficou abaixo do limite imposto de 1%, como esperado.

Tabela 7.6: Sistema da Costa Rica - Caso CVaR(1-3): medidas de risco

Medida	Etapa 1	Etapa 2	Etapa 3	Acumulado
Risco	-	-	9,6%	9,6%
EENS	-	-	0,24%	0,08%
$CVaR_{8\%}$	-	-	2,86%	0,99%
$CVaR_{4\%}$	-	-	5,18%	1,8%

7.7 Análise comparativa entre as metodologias

Como foi observado, a aplicação das metodologias proposta no capítulo 6 e 7, apesar de utilizarem o mesmo arcabouço matemático, têm aspectos diferentes sob o ponto de vista do planejamento da operação.

Em comparação com metodologia de controle do CVaR associado ao déficit de apenas uma etapa, podemos considerar que a metodologia de controle do CVaR associado ao déficit acumulado apresenta uma evolução pois evita que os déficits sejam realocados nas etapas anteriores para cumprir a meta estabelecida para uma determinada etapa.

Este comportamento pode ser observado no gráfico da figura 7.8. Comparando os resultados obtidos para comportamento do déficit associado à etapa 2 nos casos BASE e CVaR(3) observa-se que, apesar de o risco ter se mantido igual a 4%, todos os demais índices aumentaram. Isto indica que o número de cenários com ocorrência de déficit se manteve igual porém os déficits resultantes foram mais severos. Por outro lado, o caso CVaR(1-3) não apresentou déficits para esta etapa.

A avaliação dos resultados para a etapa 3, mostra que ambas metodologias, CVaR(3) e CVaR(1-3), foram capazes de reduzir consideravelmente a ocorrência e a severidade dos déficits para esta etapa em comparação com o caso BASE. Esta

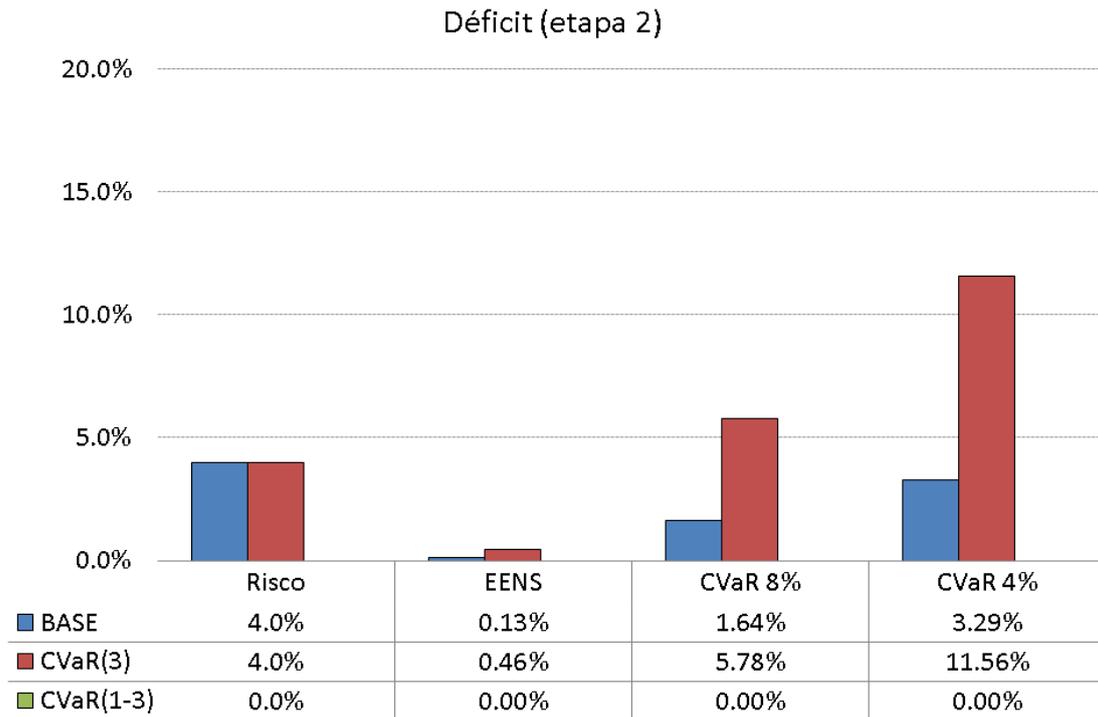


Figura 7.6: Sistema da Costa Rica: comparação das medidas de risco associadas ao déficit da etapa 2

redução pode ser observada no gráfico da figura 7.7. Embora a restrição do caso CVaR(1-3) não limite diretamente o CVaR_{8%} do déficit da etapa 3, o resultado obtido foi inferior à esta meta, mas esta característica depende diretamente do parâmetro que define o limite máximo para a medida de risco.

A figura 7.8 mostra as medidas de risco associadas ao déficit acumulado nas 3 etapas.

Finalmente, a figura 7.9 ilustra o comportamento do CVaR_{8%} para as 3 etapas e para o déficit acumulado. Pode-se observar que o resultante déficit associado à etapa 3 tem valor bastante semelhante nos casos CVaR(3) e CVaR(1-3) mas que, por outro lado, o caso CVaR(1-3) foi capaz de eliminar a ocorrência de déficits na etapa 2.

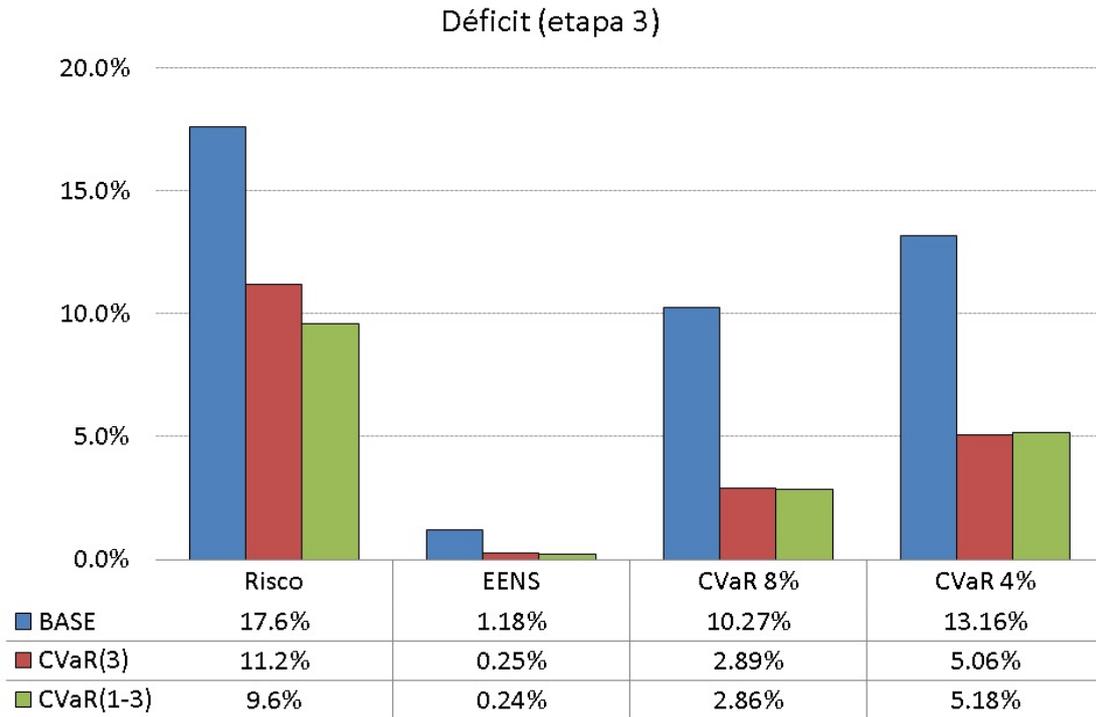


Figura 7.7: Sistema da Costa Rica: comparação das medidas de risco associadas ao déficit da etapa 3

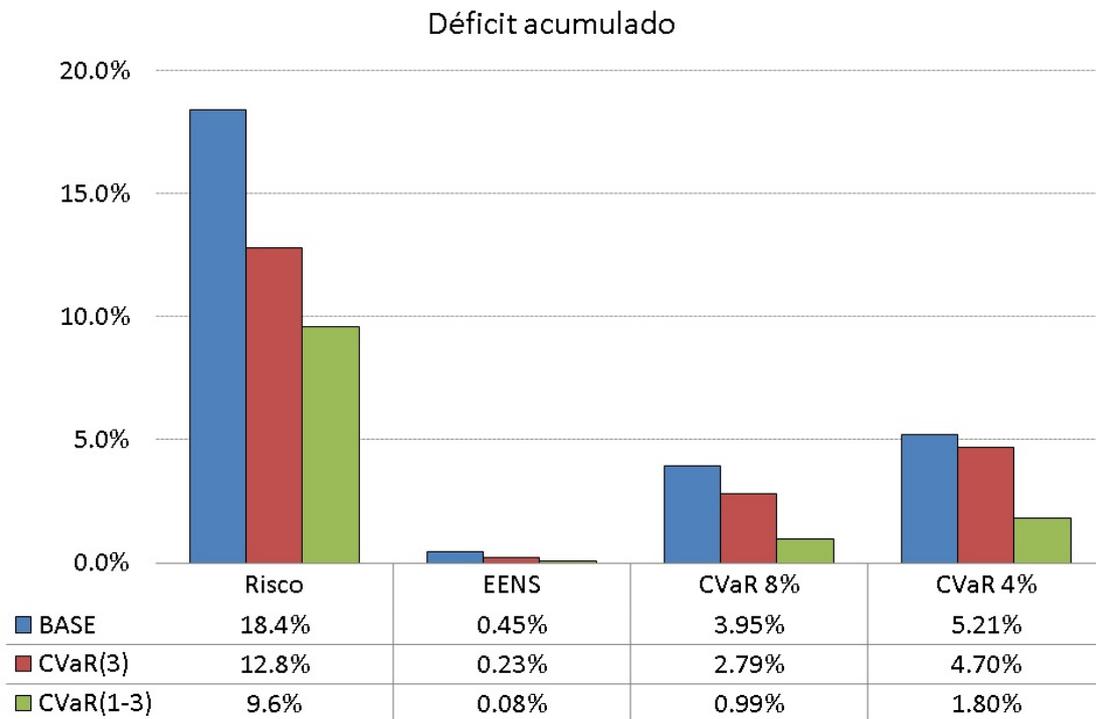


Figura 7.8: Sistema da Costa Rica: comparação das medidas de risco associadas ao déficit acumulado

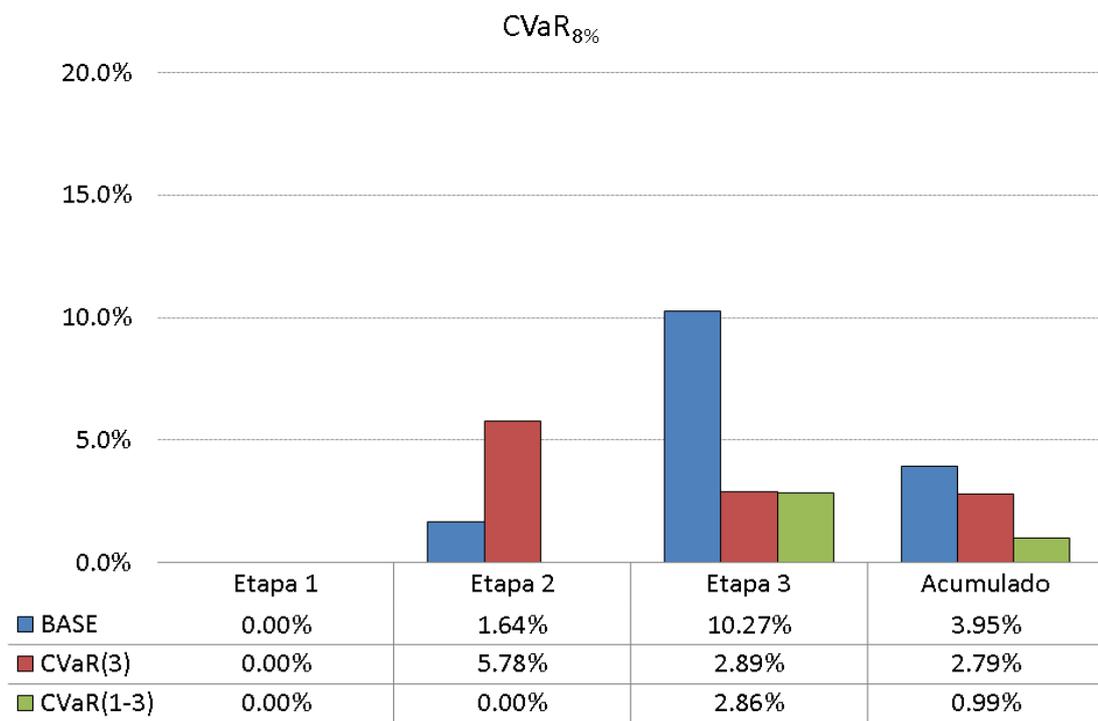


Figura 7.9: Sistema da Costa Rica: comparação do CVaR_{8%} associado ao déficit de cada etapa e ao déficit acumulado

Capítulo 8

Estudo de caso

Este capítulo apresenta os resultados obtidos com a aplicação da metodologia proposta no capítulo 7 à um estudo de caso considerando o sistema elétrico brasileiro. À seguir, serão descritas a implementação computacional, características do sistema e os resultados e conclusões obtidas.

8.1 Implementação computacional

A implementação foi realizada utilizando-se como base o modelo computacional SDDP[®], desenvolvido pela PSR [53]. O modelo é desenvolvido em FORTRAN/C e a solução dos problemas lineares é obtida através do pacote de otimização XPRESS[®] 7.1 [10].

Os casos foram executados em um ambiente de computação paralela Linux com um *cluster* composto por 15 computadores Intel[®] Xeon[®] E5-2670, cada um com 16 cores de 2.60 GHz e memória RAM de 60 GB.

8.2 Descrição do sistema

O sistema elétrico do Brasil é o maior da América Latina, com capacidade instalada ao final de 2012 de 124 GW - 70% dos quais correspondem à usinas hidrelétricas - e um pico de demanda de 72.7 GW. Em termos de estrutura de negócio, a participação privada é de 85% no setor de distribuição e de 15% no setor de geração. Esta é uma consequência do processo de privatização no Brasil, que começou pelo setor de distribuição e foi parado por oposição política quando havia alcançado o segmento de geração de energia.

Quase 90% da energia produzida é oriunda de usinas hidrelétricas, a geração restante é um *mix* que inclui gás natural, carvão, nuclear e óleo. Bioeletricidade (cogeração a partir do bagaço da cana de açúcar) surgiu como uma nova fonte com-

petitiva e a energia eólica parece seguir o mesmo caminho. O Brasil tem um mercado com crescimento acelerado e as taxas crescimento do consumo de energia flutuando por cerca de 5% ao ano. Para atender a esta demanda, é necessária a entrada de aproximadamente 3.200 MW médios (cerca de 5.000 MW de capacidade de pico) por ano. Os consumidores cativos ainda representam 70% do mercado e os 30% restantes são consumidores livres, ou seja, podem decidir livremente quem será seu fornecedor de energia.

O sistema hidrelétrico principal é composto de 200 usinas, com capacidades que variam desde 30 MW até 14.000 MW (usinas de Itaipú binacional, na fronteira com o Paraguai), localizadas em um dúzia de bacias hidrográficas ao longo do país. Devido à grande área ocupada pelo Brasil (equivalente à área continental dos Estados Unidos mais metade do Alasca), as bacias apresentam uma grande variedade de climas e padrões de afluências. Por exemplo, durante a ocorrência do conhecido fenômeno climático “El Niño”, a região Nordeste enfrenta secas, enquanto a região Sudeste apresenta um aumento das chuvas. Para tirar vantagem desta diversidade, o Operador Nacional do Sistema (ONS) despacha todo o sistema hidrelétrico como um “portfólio”, com transferências de grandes blocos de energia das regiões “úmidas” para as regiões “secas”. O Sistema Interligado Nacional (SIN) é formado pelas empresas das regiões Sul, Sudeste, Centro-Oeste, Nordeste e parte da região Norte. Apenas 3,4% da capacidade de produção de eletricidade do país encontra-se fora do SIN, em pequenos sistemas isolados localizados principalmente na região amazônica.

As usinas hidrelétricas são despachadas baseadas em seus custos de oportunidade esperados (“valores da água”), calculados por um modelo computacional que soluciona um problema de programação estocástica multiestágio que utiliza o algoritmo de PDDE e representa com detalhe a operação das usinas hidrelétricas e as incertezas nas afluências.

Este tipo de despacho com amplitude nacional requer a existência de uma rede de transmissão integrada e robusta, capaz de acomodar diferentes padrões de exportação e importação entre as regiões. O sistema principal (rede básica) tem 90 mil quilômetros de extensão, com tensões entre 230 kV e 765 kV em corrente alternada. Existem também dois circuitos de corrente contínua de 600 kV: uma linha de 900 km, que é parte do sistema de transmissão de Itaipú e um sistema *back-to-back* de 2.200 MW para conexão com a Argentina. A figura 8.1 apresenta um diagrama esquemático do sistema de transmissão brasileiro.

8.3 Caso Base

Como caso BASE deste estudo, foi considerado o caso do Programa Mensal de Operação (PMO) de Janeiro de 2013. O horizonte do estudo é de 5 anos representados

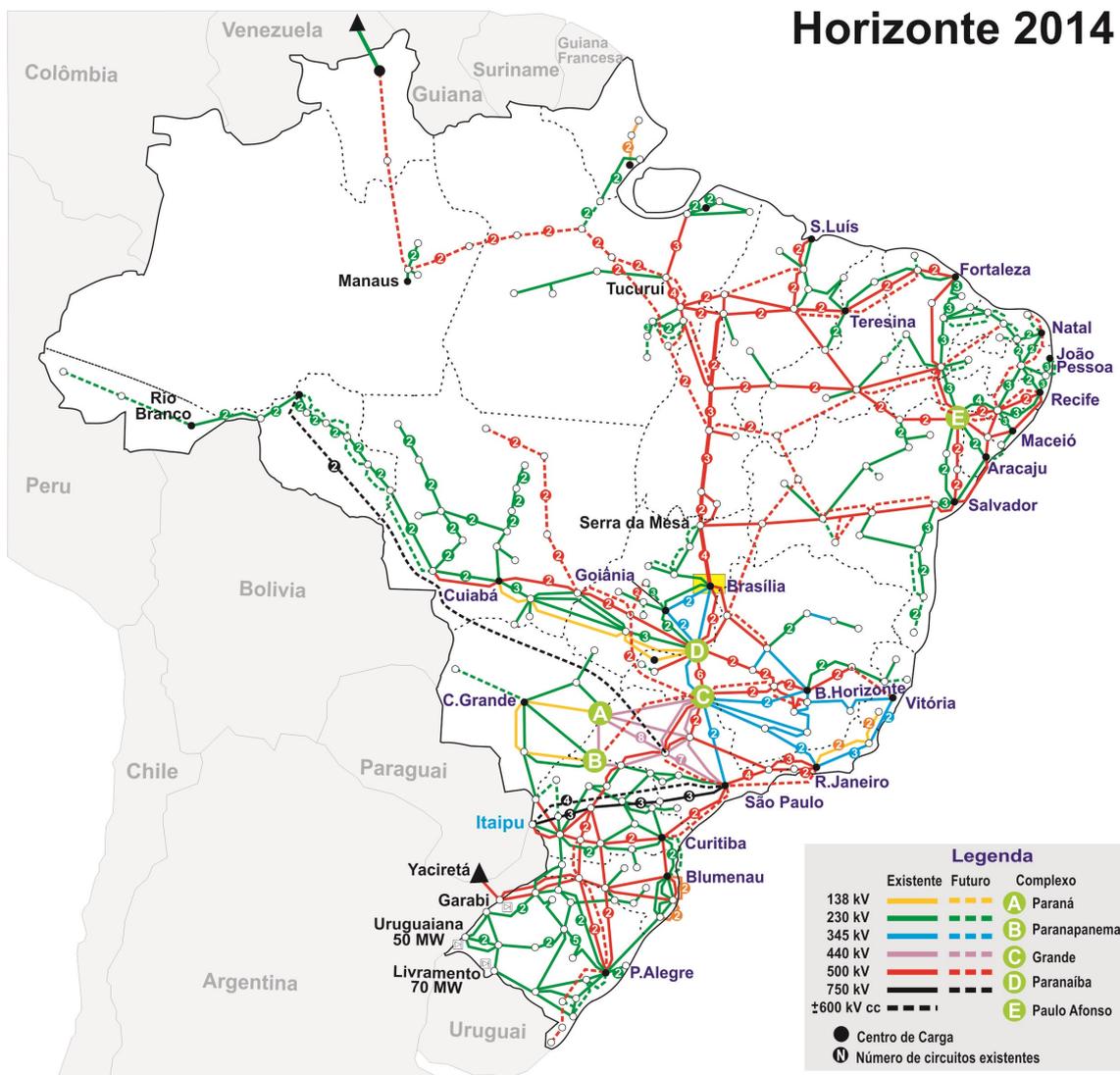


Figura 8.1: Sistema Interligado Nacional - diagrama esquemático do sistema de transmissão (Fonte: ONS)

em etapas mensais com 5 anos adicionais para evitar o esvaziamento dos reservatórios ao final do quinto ano, totalizando 120 etapas.

O SIN está representado por 4 subsistemas principais (Sudeste/Centro-Oeste, Sul, Nordeste e Norte) e o sistema de transmissão através de 5 interligações regionais, incluindo o sistema fictício de Imperatriz que representa um importante nó do sistema de transmissão. O parque gerador representado é composto por 157 usinas hidrelétricas representadas individualmente e 119 usinas térmicas.

Como a metodologia proposta nesta tese também consiste em um mecanismo de aversão a risco, destaca-se que este estudo foi elaborado desconsiderando a Curva de Aversão ao Risco (CAR) e os Procedimentos Operativos de Curto Prazo (POCP), assim como os mecanismos de aversão ao risco estabelecidos na Resolução CNPE n° 3, de 6 de março de 2013. Esta premissa tem como objetivo principal permitir que os impactos da metodologia proposta sejam capturados nas análises realizadas neste estudo.

As incertezas associada às afluências às usinas hidrelétricas são representadas através de uma árvore de cenários implícita com 30 aberturas em cada estágio e a política operativa é avaliada considerando 1.200 cenários hidrológicos no processo de simulação forward. A árvore de cenários completa teria $\mathcal{L} = 30^{120} \approx 10^{178}$ nós e torna-se crucial a utilização do método de PDDE para obtenção da solução do problema.

É considerada a função de custo de déficit de patamar único utilizada atualmente pela Empresa de Planejamento Energético (EPE) nos estudos de planejamento, cujo valor é de de R\$ 3.100,00/MWh. A taxa de desconto adotada é de 12% ao ano.

O caso foi executado no ambiente de computação paralela descrito acima e a execução demorou um 4.615 segundos (1 hora e 17 minutos).

Primeiramente, algumas medidas de risco associadas à solução encontrada foram calculadas e avaliadas. A tabela 8.1 apresenta o risco de qualquer déficit anual por subsistema e para o SIN. Observe que o risco do SIN será sempre maior ou igual ao maior risco de déficit dos sistemas e, neste caso, é igual ao risco do sistema Sul.

Tabela 8.1: SIN - Caso BASE: risco de déficit anual

Ano	Sudeste	Sul	Nordeste	Norte	SIN
2013	6,3%	7,0%	2,3%	2,6%	7,0%
2014	5,2%	5,8%	2,0%	2,3%	5,8%
2015	3,5%	4,6%	0,7%	1,2%	4,6%
2016	1,9%	2,1%	0,9%	1,0%	2,1%
2017	2,6%	2,9%	0,7%	0,5%	2,9%

Em seguida, foram calculadas outras medidas de risco considerando o déficit

total do SIN acumulado para cada ano, como pode ser observado na tabela 8.2. Os valores apresentados para EENS e CVaR estão expressos em percentual da demanda total do ano correspondente. Observa-se que, enquanto o ano de 2013 apresenta o maior risco (probabilidade de ocorrer um déficit igual 7,0%), os déficits referentes ao ano de 2014 são mais severos. A média dos déficits nos 4% piores cenários (CVaR_{4%}) é igual a 5,3% e chega a quase 11% da demanda se considerarmos os 1% casos mais extremos (CVaR_{1%}).

Tabela 8.2: SIN - Caso BASE: medidas de risco para o déficit acumulado anual

Ano	Risco	EENS	CVaR _{4%}	CVaR _{2%}	CVaR _{1%}
2013	7,0%	0,18%	4,05%	5,88%	7,64%
2014	5,8%	0,22%	5,30%	7,98%	10,95%
2015	4,6%	0,11%	2,75%	4,95%	7,36%
2016	2,1%	0,07%	1,84%	3,68%	6,02%
2017	2,9%	0,07%	1,70%	3,28%	5,03%

8.4 CVaR associado ao déficit acumulado anual

Foi utilizado como critério a inclusão de uma restrição que limita o pior CVaR anual (entre os anos de 2013 e 2017). Em termos práticos, os déficits acumulado são calculados para cada um dos 5 anos do horizonte durante o processo de otimização e a penalização foi ajustada de forma a minimizar o maior déficit ocorrido. Além disto, foi considerada apenas uma franquia para todos os anos. A intuição por trás deste critério é que, como foi visto, a restrição de CVaR define indiretamente uma curva de déficit de 2 segmentos e acredita-se que esta curva não deve variar ao longo do horizonte, ou seja, a curva de custo de déficit não deve se adaptar à conjuntura de cada ano. Em outras palavras, o critério de planejamento deve permanecer constante em todo o horizonte do estudo.

Um importante aspecto de qualquer metodologia de aversão a risco é a definição dos parâmetros que serão considerados. Vale ressaltar que o objetivo deste estudo de caso é ilustrar a aplicação da metodologia que permite a inclusão de restrição de CVaR no planejamento da operação e não como os parâmetros nível de confiança α e limite η associados à restrição de confiabilidade devem ser estabelecidos. A definição dos parâmetros deve seguir critérios de planejamento e devem ser realizados estudos para assegurar que os parâmetros selecionados são adequados para o sistema em questão.

O nível de confiança α indica o percentual dos piores cenários serão contabilizados no cálculo do CVaR. Apesar de qualquer valor entre 0% e 100% ser considerado

válido (onde $\alpha = 100\%$ corresponde ao critério de limite EENS), é importante notar que caso seja selecionado um valor de α maior que o risco de déficit haverão cenários sem ocorrência de déficit incluídos no cálculo do CVaR. Neste estudo de caso, optou-se pelo valor de $\alpha = 4\%$, que é menor que o risco no ano de 2014 (ano que apresenta o pior $\text{CVaR}_{4\%}$).

Como limite máximo foi adotado o valor de $\eta = 5\%$ da demanda, que é inferior ao valor $\text{CVaR}_{4\%} = 5,3\%$ obtido no ano de 2014. Desta forma, é possível avaliar o impacto da inclusão de uma restrição “ativa” que implicará na alteração da política operativa para que esta seja satisfeita.

Sendo assim, foi imposta uma restrição que limita a pior ocorrência de $\text{CVaR}_{4\%}$ associado ao déficit acumulado anual para os 5 anos do horizonte à um valor máximo de 5% da demanda anual. Os limites iniciais mínimo e máximo para a busca da penalização através do método da bisseção foram definidos como $\underline{\mu} = 0$ \$/MWh e $\bar{\mu} = 100$ \$/MWh, respectivamente.

O algoritmo convergiu em 12 iterações e o tempo total de execução foi de 45.347 segundos (12 horas e 35 minutos). A penalização ótima encontrada foi de $\mu^* = 11,621$ \$/MWh e foi obtido $\text{CVaR}_{4\%} = 4,86\%$ da demanda para o pior ano do estudo. A tabela 8.3 apresenta a tabela de convergência e o gráfico da figura 8.2 ilustra a evolução da penalização em cada iteração. Em cinza, destaca-se o intervalo entre os limites mínimo e máximo da penalidade durante o processo de convergência.

Tabela 8.3: SIN - Caso CVaR: tabela de convergência do método da bisseção

Iteração	Penalidade	VaR _{4%}	CVaR _{4%}	Custo total
1	100	1,46%	4,06%	46,962,822
2	0	1,60%	5,51%	55,775,413
3	50	1,21%	4,37%	53,076,619
4	25	1,55%	4,94%	56,273,699
5	12,5	1,23%	4,90%	57,163,533
6	6,25	1,50%	5,58%	56,963,933
7	9,375	1,32%	5,54%	57,012,748
8	10,938	1,57%	5,58%	57,183,552
9	11,719	1,28%	4,86%	57,173,967
10	11,328	1,39%	5,58%	57,224,786
11	11,523	1,67%	5,51%	57,048,949
12	11,621	1,28%	4,86%	57,195,502

A figura 8.3 apresenta a evolução do $\text{CVaR}_{4\%}$ em função da penalidade. Diferentemente do ocorrido com os casos apresentados para o sistema da Costa Rica, observa-se um comportamento não decrescente em alguns trechos da função. Isto pode ser explicado devido à incerteza amostral inerente à aproximação realizada da

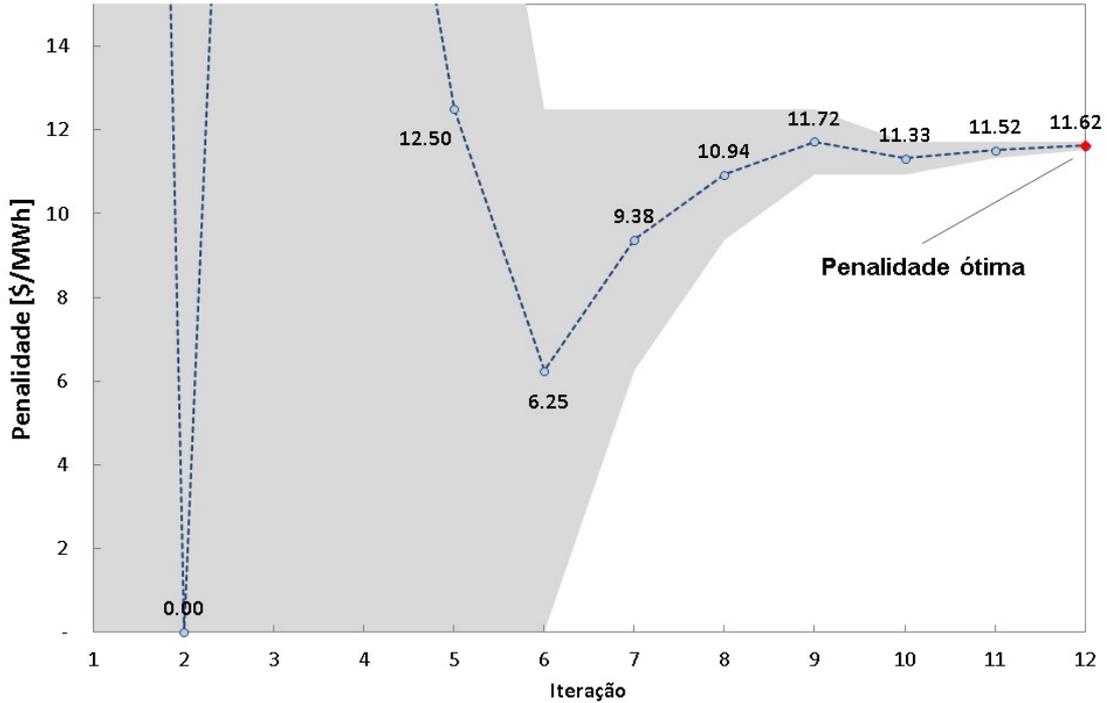


Figura 8.2: SIN - Penalidade por iteração

árvore de incertezas considerada no caso brasileiro. Como esperado, o comportamento esperado para CVaR não é alterado e é possível observar que este diminui com o aumento da penalidade.

A função custo total (função Lagrangeana) associada à este caso é ilustrada na figura 8.4 e podemos observar que a solução ótima é atingida no ponto que maximiza o custo total obtido.

As mesmas medidas de risco calculadas para o caso BASE são calculadas para o caso CVaR e são apresentadas na tabela 8.4. Pode-se notar que, além do cumprimento do limite, houve uma melhora em todas as demais medidas de risco.

Tabela 8.4: SIN - Caso CVaR: medidas de risco para o déficit acumulado anual

Ano	Risco	EENS	CVaR _{4%}	CVaR _{2%}	CVaR _{1%}
2013	6,5%	0,16%	3,70%	5,39%	7,04%
2014	5,8%	0,19%	4,86%	7,16%	9,96%
2015	4,0%	0,09%	2,33%	4,29%	6,46%
2016	1,9%	0,06%	1,55%	3,11%	5,14%
2017	3,0%	0,06%	1,54%	3,00%	4,61%

A figura 8.5 apresenta a evolução da energia armazenada no SIN, em termos percentuais da energia máxima armazenável. As linhas cheias no gráfico representam

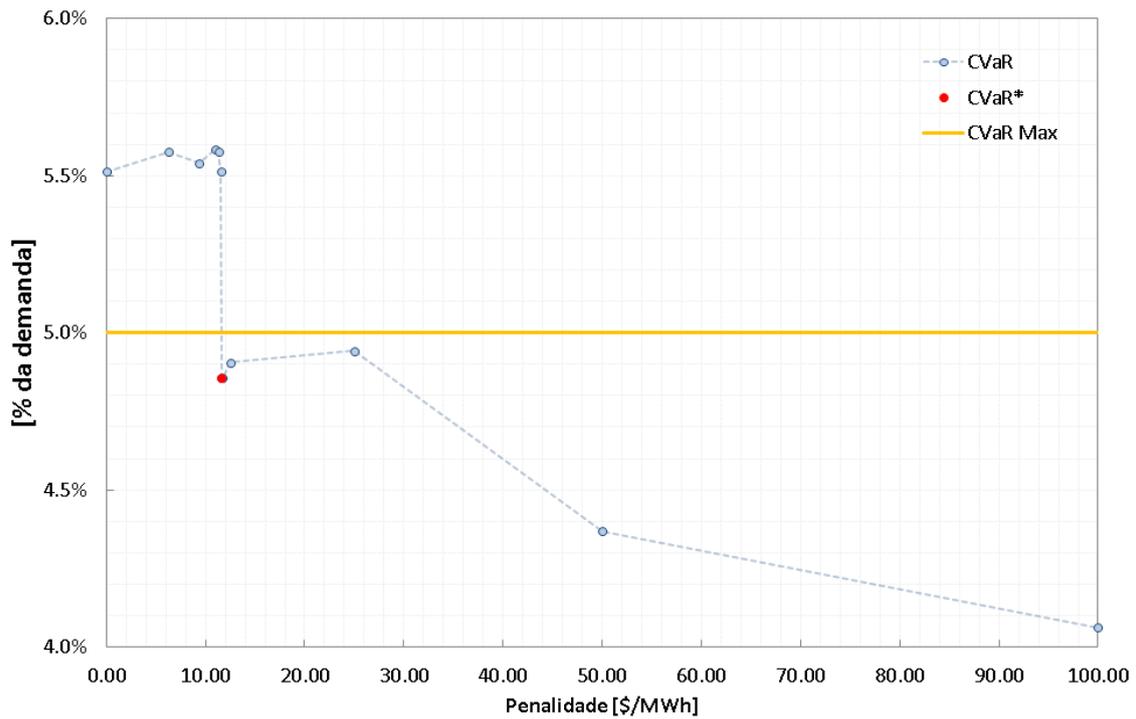


Figura 8.3: SIN - $CVaR_{5\%}$ associado ao déficit acumulado em função da penalidade

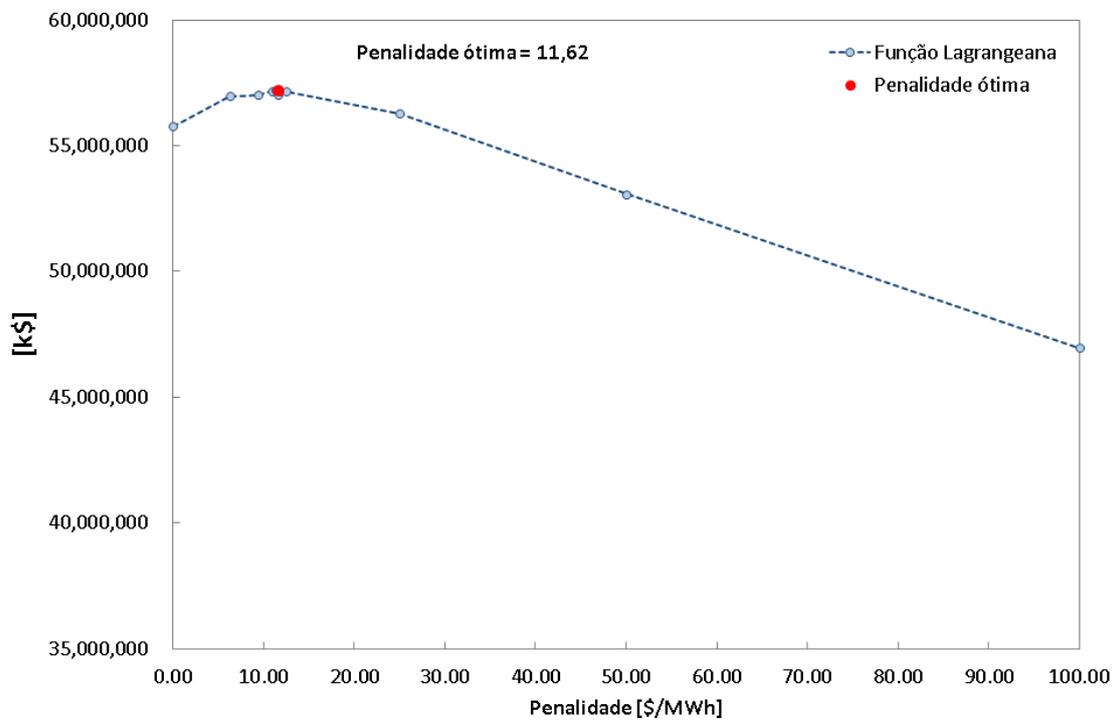


Figura 8.4: SIN - Custo total em função da penalidade (função Lagrangeana)

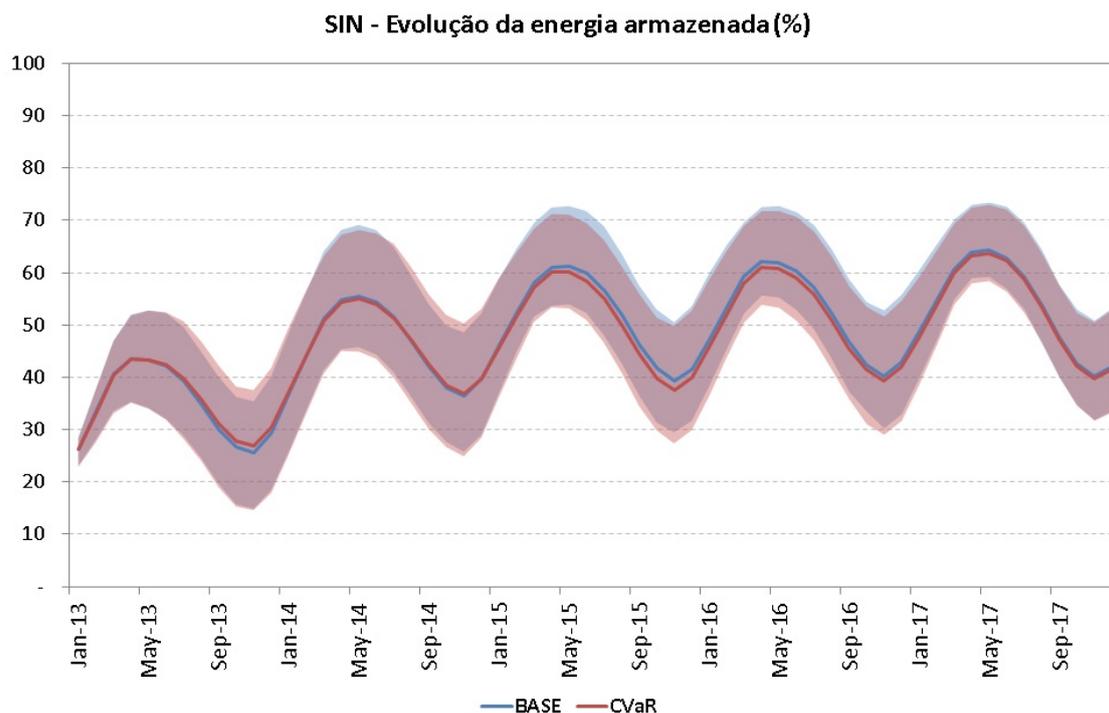


Figura 8.5: SIN - Evolução da energia armazenada para os casos BASE de CVaR

os valores médios, enquanto as regiões sombreadas representam a dispersão da média para os quantis de 25% e 75%. Pode-se observar que, em comparação com o caso BASE, os resultados obtidos no caso CVaR apresentaram um ligeiro aumento na energia armazenada no ano de 2013, indicando um aumento no armazenamento que foi capaz de garantir maior segurança no suprimento no ano de 2014. Por outro lado, como os anos seguintes não violavam a restrição de confiabilidade, a operação do caso CVaR permitiu uma ligeira redução da energia armazenada.

O gráfico da figura 8.6 compara as funções densidade de probabilidade aproximada associadas ao déficit acumulado no ano de 2014 para os casos BASE e CVaR¹. É possível observar que a inclusão da restrição de CVaR, em comparação com o caso BASE, reduziu a probabilidade de ocorrência de déficits mais severos porém aumentou a probabilidade de ocorrência de déficits de menor valor. Este efeito garante a redução do CVaR e o atendimento do limite estabelecido, sendo esta uma característica da medida de risco CVaR (foco nos cenários mais extremos).

Finalmente, é possível obter a curva de custo de déficit de 2 segmentos a partir dos valores da solução ótima para $VaR_{4\%}$ e para penalidade μ . O primeiro segmento da curva tem inclinação igual ao custo econômico do déficit $\delta = 3.100 \text{ \$/MWh}$ e uma profundidade igual à franquia $VaR_{4\%} = 1,28\%$ da demanda acumulada. Já o

¹ É importante destacar que os valores de déficit exibidos no gráfico são apenas aqueles correspondente aos cenários com ocorrência de déficit, ou seja, a área abaixo da curva corresponde apenas ao risco de déficit de cada caso (aproximadamente 6%).

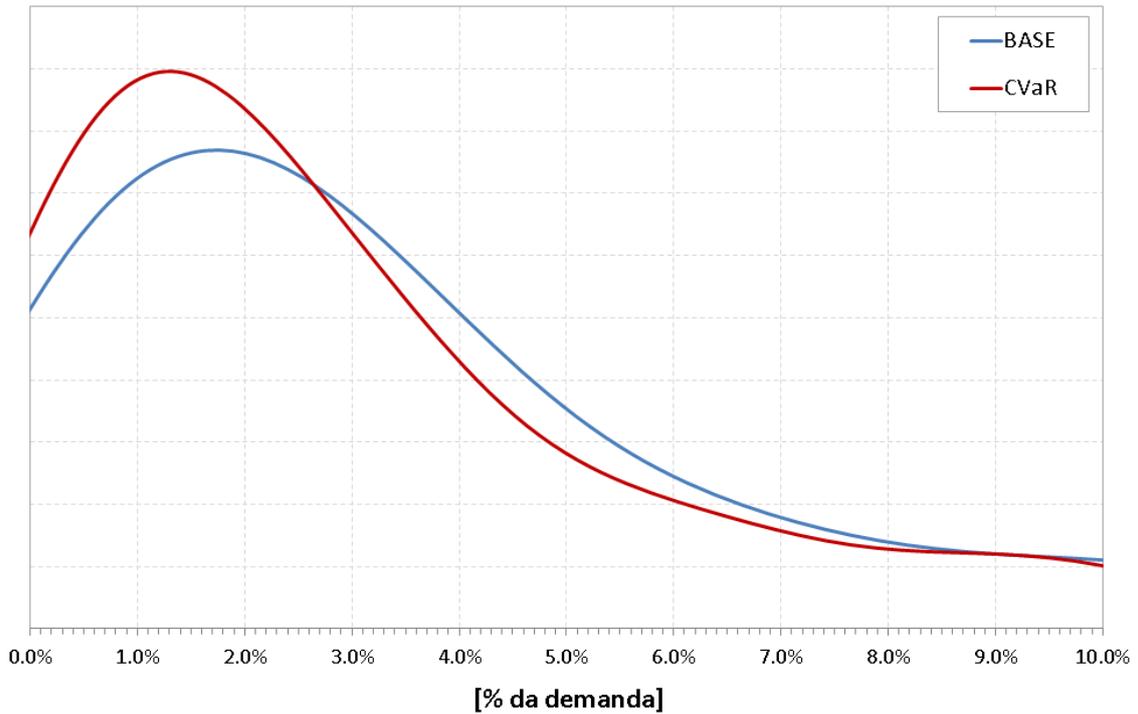


Figura 8.6: SIN - Comparação das funções densidade (aproximada) associadas ao déficit acumulado para os casos BASE e CVaR (ano 2014)

segundo segmento tem inclinação igual à soma do custo econômico e do prêmio de risco, ou seja, $\delta + \frac{\mu}{\alpha} = 3.100 + \frac{11,621}{0,04} = 3.390,5$ \$/MWh. A tabela 8.5 consolida os valores da curva de custo de déficit enquanto a figura 8.7 a ilustra graficamente.

Tabela 8.5: SIN - Caso CVaR: curva de custo de déficit acumulado de 2 segmentos

Segmento (% da demanda)	Custo (\$/MWh)
0,00% - 1,28%	3.100,0
1,28% - 100,00%	3.390,5

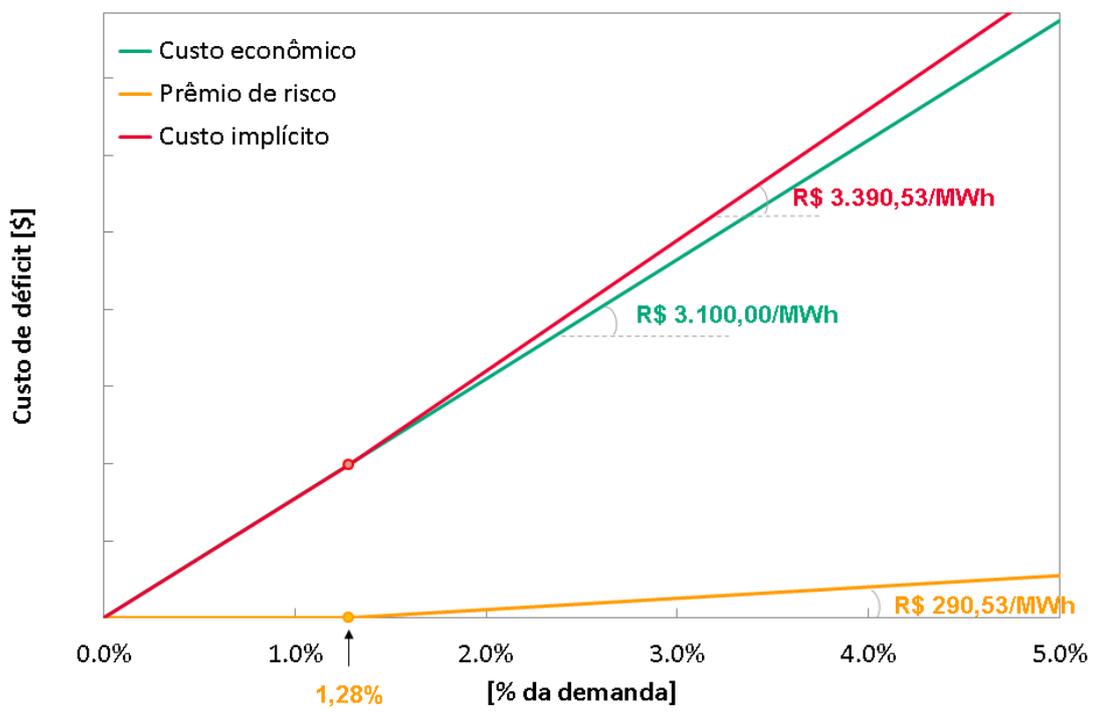


Figura 8.7: SIN - curva de custo de déficit acumulado de 2 segmentos

Capítulo 9

Considerações finais e desenvolvimentos futuros

9.1 Conclusões

As características de sistemas hidrotérmicos interligados, como é o exemplo do setor elétrico brasileiro, obrigam que as decisões quanto ao despacho das usinas sejam tomadas com antecedência e sob um cenário de incerteza quanto à disponibilidade dos recursos futuros, principalmente as aflúências. Adicionalmente, foi visto que os modelos computacionais tipicamente utilizados no setor atualmente não são capazes de garantir, em sua política operativa, decisões que protejam o sistema quando da ocorrência de cenários futuros menos favoráveis. Para solucionar este problema, comumente são propostas regras e procedimentos operativos adicionais que visam complementar as decisões operativas, com o objetivo de aumentar a segurança no suprimento do sistema.

A justificativa para isto consiste no fato de os modelos estocásticos de otimização tradicionais buscarem o despacho que minimiza os custos de suprimento da demanda sobre a média dos cenários hidrológicos considerados. Esta premissa baseia-se na Teoria de Grandes Números e é razoável para estudos de longo prazo, quando não se está interessado em uma sequência hidrológica específica. Tais modelos são caracterizados como “neutros com relação ao risco” e, como sua influência é refletida no valor esperado dos custos operativos, não há garantia de proteção para os eventos de maior arrependimento que correspondem à ocorrência de séries críticas e à violação de critérios de segurança previamente estabelecidos.

Por sua vez, foi visto que o custo econômico do déficit procura refletir o impacto de um racionamento nas atividades econômicas do país. Embora esta redução seja obviamente indesejável, a experiência mostra que a mesma não captura plenamente o valor que a sociedade atribui à energia elétrica e, conseqüentemente, a aversão à

possibilidade, ou risco, de que ocorra um racionamento. Na prática, verifica-se que custos diretos e sócio-políticos de um racionamento são inaceitáveis pela sociedade, tornando imprescindível a adoção pelo operador de medidas para evitar a elevação dos riscos de déficit.

Diferentes estratégias podem ser adotadas para lidar com aversão ao risco em problemas de planejamento da operação. A primeira, abordada neste trabalho, consiste em considerar uma restrição de confiabilidade explícita associada aos déficits de energia, que são os eventos que se deseja evitar. A segunda abordagem é alterar a política operativa para criar uma reserva energética de segurança que evite a ocorrência de déficits em cenários desfavoráveis e a terceira consiste em atribuir um peso relativo maior aos cenários de afluências desfavoráveis no cálculo da função de custo futuro e tentar minimizar os custos associados.

O objetivo da abordagem proposta neste trabalho foi considerar um critério de confiabilidade explícito através da representação de restrição que impõe que a medida de risco CVaR (Conditional Value-at-Risk) associada aos déficits de energia deve estar limitada à um máximo valor, de acordo com critério de segurança pré-determinado pelo planejador do sistema. A inclusão da restrição de CVaR se mostrou interessante pois permitiu controlar diretamente a exposição do sistema à ocorrência de déficits mais severos sendo possível, por exemplo, a inclusão de uma restrição que impôs que o valor esperado da energia não suprida nos 4% piores cenários estivesse limitado à um percentual da demanda do sistema.

Do ponto de vista matemático, esta formulação necessitaria da representação de uma restrição que envolvesse explicitamente as variáveis de déficit de todos os cenários em uma única restrição. Apesar de esta representação ser bastante simples e direta, o problema deveria ser representado na forma de uma árvore de cenários e solucionado como um único problema de otimização. Entretanto, sabe-se que o tamanho do problema de otimização está diretamente relacionado ao número de nós da árvore de incertezas, que cresce exponencialmente com o aumento do número de etapas do problema. Desta forma, esta formulação limitaria a sua aplicação a problemas de poucos estágios.

Para contornar este problema, foi proposto um modelo onde esta restrição é lagrangeada e representada através de sua penalização na função objetivo do problema. O uso desta técnica permitiu considerar a solução do problema em um ambiente de Programação Dinâmica Dual Estocástica cuja estratégia é solucionar o problema original através da solução iterativa de um conjunto de subproblemas menores relacionados recursivamente e associados aos nós da árvore de decisão. Sendo assim, determinada uma penalização, o problema pôde ser solucionado por pelo algoritmo de PDDE considerando a extensão do modelo proposta. Para a solução do problema Dual Lagrangeano, foi proposto um algoritmo de ajuste iterativo da

penalização baseado no método da bisseção.

A primeira aplicação da metodologia proposta foi apresentada através da formulação do problema de planejamento da operação de sistemas hidrotérmicos considerando uma restrição de risco de CVaR associada ao déficit de energia em um determinado estágio. Mostra-se que, ao considerar a restrição de CVaR dualizada, os problemas solucionados a cada um dos estágios deixam de ser semelhantes pois são adicionadas variáveis e restrições auxiliares para a representação do CVaR.

Além disto, um novo estágio de decomposição é adicionado antes do primeiro estágio de operação, denominado “estágio de decisão estratégica”, que tem como objetivo determinar o valor da variável associada ao VaR (franquia). Diferentemente das variáveis de estados tradicionais do problema de operação, como o volume final dos reservatórios em cada etapa e afluências anteriores, esta variável de estado é única e influencia à todos os cenários da etapa da restrição de CVaR.

Como a PDDE baseia-se na solução iterativa de problemas recursivos de um estágio, buscou-se apresentar em detalhe a formulação do problema de otimização solucionado em cada um dos estágios assim como as expressões que definem os coeficientes dos cortes de Benders associados às variáveis de estado na função de custo futuro.

Através de uma análise da formulação do problema, é estabelecida uma interpretação econômica para este modelo onde ressalta-se a sua equivalência com o problema tradicional com representação da curva de custo de déficit de dois patamares. Observa-se que o VaR (franquia) corresponde à profundidade do primeiro segmento da curva de custo de déficit e a penalização corresponde à inclinação adicional aplicada à variável do segundo segmento de déficit. Esta importante contribuição conceitual permitiu um entendimento claro a respeito dos efeitos da restrição de CVaR, aliada ao fato de o parâmetro que define o limite corresponder à uma medida/grandeza do déficit para o qual se quer proteger.

Desta forma, pode-se interpretar o algoritmo proposto como uma ferramenta de programação matemática para o ajuste ótimo de uma curva de déficit implícito que representa, na verdade, a aversão ao risco do operador do sistema segundo um critério de segurança de suprimento baseado no índice de risco do CVaR associado ao déficit de energia.

Esta metodologia foi aplicada à um caso de estudo simplificado, baseado no sistema real da Costa Rica, com o objetivo de ilustrar a validade do algoritmo proposto assim como os principais resultados obtidos. A simplificação adotada foi a redução no número de estágios (3 estágios) e das aberturas da árvore de incertezas consideradas (5 aberturas), para que fosse possível a enumeração e solução de todos os cenários da árvore de incertezas (125 cenários). O objetivo desta simplificação foi permitir que o número de nós da árvore fosse “tratável” de modo que o algoritmo de

PDDE pudesse considerar todos os nós da árvore no processo de solução. Sendo assim, excluíram-se quaisquer desvios nos resultados causados por incertezas amostrais que pudessem ser obtidas devido à aproximação da árvore.

Neste estudo de caso, o objetivo foi a minimização dos custos esperados de operação sujeito à uma restrição de risco associada ao déficit de energia da última etapa (etapa 3). Os custos de operação foram compostos por custos de combustível para geração térmica e penalizações associadas ao custo econômico de redução da demanda (déficit de energia). A restrição de risco considerada foi o $\text{CVaR}_{8\%}$, ou seja, a média dos 8% piores cenários de déficit de energia.

Os resultados obtidos mostraram que o algoritmo proposto para o ajuste da penalização converge para a solução ótima e verificou-se que a solução coincide com o valor máximo da função Dual Lagrangeana. Também verificou-se, como esperado, que a função controlada durante o processo iterativo, que corresponde ao CVaR associado ao déficit do último estágio, tem comportamento não crescente com o aumento da penalização.

Do ponto de vista do planejamento da operação, os resultados obtidos mostraram que foi possível controlar o CVaR dos déficits de energia da última etapa abaixo do limite imposto. Além disto, a imposição desta restrição teve o efeito prático de diminuir outras medidas de risco calculadas para a última etapa, como o risco de déficit, valor esperado da energia não suprida e o CVaR associado à outros níveis de confiança.

Os resultados obtidos como solução ótima para a franquia associada ao CVaR e o multiplicador de Lagrange da restrição (penalização) foram utilizados para ilustrar a curva de déficit implícito associada à restrição de CVaR imposta, contribuindo para o entendimento da metodologia.

Entretanto, mostrou-se também que os índices de risco associados à penúltima etapa (etapa 2) sofreram uma elevação significativa. A explicação para este comportamento consiste no fato de que, com o objetivo de minimizar os custos em termos esperados, a solução mais econômica consistiu em deixar de gerar energia na etapa 2 e armazenar energia nos reservatórios das usinas hidrelétricas para gerar energia na etapa 3 e, desta forma, atender a restrição imposta. Sendo assim, fica evidente uma deficiência desta formulação específica, que considera uma restrição envolvendo os déficits de apenas uma etapa.

Em seguida, foi proposta a extensão do modelo anterior para a consideração de restrições de risco CVaR que envolvessem os déficits de um conjunto de etapas. Na abordagem adotada, é proposta a utilização do déficit acumulado ao longo de um período e é proposta a inclusão de uma variável de estado adicional para sua representação no esquema de programação dinâmica.

Esta nova variável de estado pode ser vista como um “reservatório” de déficit,

onde a cada etapa do problema os déficits que ocorrem no estágio corrente são somados ao déficit acumulado que ocorreu nas etapas anteriores. Na primeira formulação proposta a extensão do modelo anterior baseia-se em simplesmente alterar a função controlada pelo CVaR, substituindo a variável déficit no estágio pela função déficit acumulado. Sendo assim, o modelo acumula os déficits nos estágios e, no final do período, o “excesso” de déficit acumulado com relação à franquia do CVaR é penalizado na função objetivo. O sinal econômico associado ao benefício da redução dos déficits nos estágios anteriores é indicado através de termos adicionais na função de custo futuro e o algoritmo de PDDE se encarrega de obter a solução ótima ao longo do processo iterativo.

Entretanto, observou-se que seria possível reformular um problema matematicamente equivalente para distribuir a penalização sobre o excesso de déficit acumulado em todas as etapas do período de interesse, diferentemente de aplicar a penalização apenas na última etapa do período do CVaR. Esta formulação foi então adotada pois se mostrou atraente, com um potencial de reduzir o número de iterações do algoritmo da PDDE para a obtenção da solução ótima.

Utilizou-se o mesmo caso simplificado da Costa Rica para ilustrar esta formulação, considerando como objetivo a minimização dos custos operativos sujeitos à uma restrição que limita o $\text{CVaR}_{8\%}$ associado ao déficit acumulado ao longo dos 3 períodos do estudo. O algoritmo encontrou a solução ótima e foi capaz de restringir o déficit acumulado abaixo do limite estipulado.

Observou-se que, em contraste com a metodologia anterior, a consideração do déficit acumulado foi capaz de eliminar os déficits ocorridos na etapa 2, concentrando os déficits na última etapa de forma a ainda assim atender à restrição imposta. Como a restrição imposta limita o déficit acumulado, independente da etapa que ele ocorra, este comportamento é esperado e desejado pois em um problema de otimização sujeito à incertezas não faz sentido antecipar a “certeza” de uma decisão de déficit enquanto ainda é possível postergar esta decisão indesejada para uma etapa futura. Outro efeito obtido foi a redução das demais medidas de risco calculadas.

Uma interpretação conceitual semelhante pode ser extraída desta formulação na consideração da franquia e penalização ótimas na construção de uma curva de custo de déficit implícito. Porém, cabe notar que neste modelo o segundo segmento da curva de custo de déficit se aplica ao déficit acumulado e não à ocorrência de déficits individuais ao longo das etapas do horizonte. A justificativa para este tipo de modelo encontra-se no fato de que estamos interessados em incluir uma restrição de segurança energética no planejamento da operação, que permita gerenciar de maneira mais eficiente a distribuição dos recursos hidrelétricos no tempo, através do uso dos reservatórios das usinas hidrelétricas.

Finalmente, a metodologia proposta foi aplicada à um estudo de caso real brasi-

leiro (Sistema Interligado Nacional), com dados provenientes do Programa Mensal da Operação de Janeiro de 2013, disponibilizados pelo Operador Nacional do Sistema (ONS). O estudo é composto por um horizonte de 5 anos de interesse (2013 a 2017) mais 5 anos adicionais, representados em etapas mensais. O sistema elétrico é representado por 5 subsistemas e interligações regionais. Foi considerada uma árvore de cenários implícita com 30 aberturas a cada estágio o que equivale a uma árvore completa com um total de 30^{120} cenários. Para o processo amostral de simulação forward do algoritmo de PDDE, foram considerados 1.200 cenários.

Foi realizada uma análise inicial do caso base, onde mostrou-se a presença de índices de risco elevados para o horizonte de estudo, principalmente nos 2 primeiros anos. Adotou-se como estratégia a inclusão de uma restrição que limita o máximo CVaR associado ao déficit total acumulado do SIN para cada um dos 5 anos do horizonte. Foi escolhido o índice de risco $\text{CVaR}_{4\%}$ limitado a um valor máximo de 5% da demanda anual. Mostrou-se que o algoritmo foi capaz de obter a solução ótima do problema em 12 iterações para o ajuste da penalização e solução do problema Dual Lagrangeano.

Como resultados obtidos, mostrou-se que a solução encontrada de fato respeitou o limite imposto para o CVaR, limitando a ocorrência de déficits mais severos no ano de 2014, que é o ano com piores índices de risco no horizonte considerado. Em uma comparação das funções densidade associadas às distribuições dos déficits acumulados para o ano de 2014, foi visto que o efeito prático da restrição de CVaR foi a diminuição da probabilidade de ocorrência de déficits na cauda da distribuição, o que ilustra a ótica do CVaR que tem o objetivo de impor uma proteção aos eventos mais severos. Em compensação, esta mesma análise mostrou que houve uma ocorrência maior de déficits de menor valor. Outro efeito prático foi a moderada elevação dos níveis dos reservatórios nos dois primeiros anos do estudo.

A curva de custo de déficit implícito também foi calculada para ilustrar a interpretação econômica da metodologia, ilustrando a profundidade do primeiro segmento de déficit (franquia) e o “prêmio de risco” (penalização) que deve ser pago pelo operador caso a solução ótima decida por uma operação mais arriscada.

Uma análise mais detalhada da evolução do CVaR em função da penalização mostrou que, devido à característica amostral do algoritmo de PDDE, o seu comportamento está sujeito a pequenas oscilações com o aumento da penalização. O mesmo comportamento é observado com a função Lagrangeana.

Em resumo, mostrou-se que os resultados são facilmente interpretados pois há uma relação direta biunívoca entre a restrição de limite de CVaR e uma função de penalização linear por partes associada ao déficit de energia. Uma função semelhante a esta já é utilizada nos estudos de planejamento da operação e a metodologia proposta permite estabelecer os coeficientes da curva já ajustados para o nível de risco

desejado. Um aspecto interessante é que o algoritmo proposto pode ser utilizado em um estudo de planejamento estratégico onde são definidos os critérios de aversão a risco e, nos estudos de planejamento da operação seguintes, é suficiente a utilização da curva de custo de déficit implícito que foi calculada.

Esta característica é bastante atraente pois permitiria conciliar os estudos de planejamento da expansão, que utiliza um critério de segurança de suprimento, com os estudos de planejamento da operação unificando as duas abordagens utilizadas no setor elétrico brasileiro.

Vale ressaltar que, como todos os modelos são uma simplificação da realidade, as soluções fornecidas por modelos computacionais (modelos de programação estocástica ou outros modelos em geral) devem ser embasadas por um conjunto de análises criteriosas para verificação e viabilidade de implantação da solução obtida.

Finalmente, cabe dizer que todas as metodologias necessitam de premissas e parâmetros pré-estabelecidos. Neste trabalho é apresentada uma metodologia que baseia-se no controle direto da grandeza déficit de energia e, por esta razão, acredita-se que é possível traduzir os critérios de planejamento e requisitos de segurança energética de maneira mais direta e intuitiva. Entretanto, a aversão a risco varia com o perfil do tomador de decisão e recomenda-se que sejam feitas análises mais aprofundadas a respeito dos parâmetros necessários para a correta representação a risco para cada sistema em questão.

Por fim, ressalta-se que o trabalho desenvolvido nesta tese de doutorado foi apresentado em duas conferências internacionais, um artigo nos anais de um congresso nacional [23] e um artigo técnico [20] submetido ao jornal científico IEEE Transactions on Power Systems, da sociedade americana de pesquisa IEEE Power Engineering Society.

Propostas de estender o modelo apresentado para representar outros aspectos são discutidas a seguir.

9.2 Desenvolvimentos futuros

Três linhas de trabalho distintas podem ser investigadas em futuros trabalhos: extensão da metodologia apresentada para o planejamento da expansão de sistemas elétricos, ajuste dos cortes da função de custo futuro, e aplicação da metodologia à problemas de outras áreas.

9.2.1 Planejamento da expansão

O problema do planejamento da expansão de sistemas elétricos (PPE) origina-se das mudanças necessárias no sistema devido ao crescimento da demanda de energia

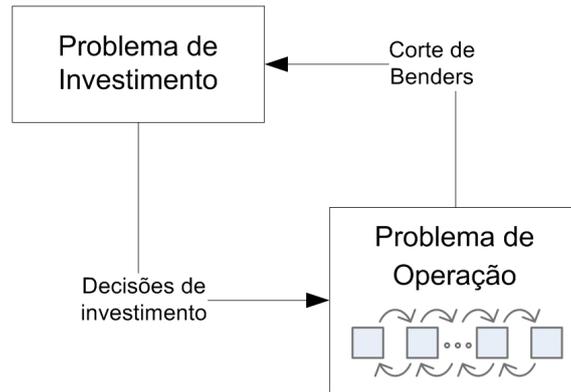


Figura 9.1: Planejamento da expansão: esquema de decomposição tradicional

com o passar dos anos. Para tanto, novos geradores devem ser construídos com o objetivo de satisfazer as novas necessidades do sistema e as decisões do processo de planejamento estão associadas à seleção das melhores usinas geradoras. Este processo de decisão dá origem a um problema de otimização de grande porte, onde o objetivo é planejar o sistema elétrico futuro minimizando os custos de investimento e operação sujeito a critérios mínimos de segurança pré-estabelecidos.

Tipicamente, os esquemas de solução tiram proveito da estrutura deste problema, pois os conjuntos de restrições de investimento e operação são acoplados apenas pelo vetor de decisões de investimento. O problema é decomposto em um problema de investimento (subproblema mestre), que tem como objetivo obter uma proposta de plano de investimentos com base nos custos de investimento e na representação aproximada da função de custo operativo; e um problema de operação (subproblema escravo) onde, dado um plano de investimentos candidato, é calculada a política operativa ótima que determina seu custo de operação e é extraída informação dual para a geração de um novo corte que melhora a representação da função de custo de operação aproximada utilizada no problema mestre. Este esquema segue iterativamente em um processo de decomposição de Benders até a obtenção da solução ótima do problema [52].

Como foi visto neste trabalho, o problema de operação é bastante complexo no caso de sistemas hidrotérmicos e requer o uso de técnicas como a PDDE para obtenção de sua solução. Sendo assim, a cada iteração é solucionado o subproblema de investimento e o subproblema da operação é solucionado por PDDE iterativamente, conforme ilustrado na figura 9.1.

No modelo proposto neste trabalho, a cada iteração algoritmo de ajuste da penalização μ , é solucionado um problema de planejamento da operação $P(\mu)$. Este problema, além de calcular a política operativa ótima também define qual é o valor da franquia (VaR) ótima associada ao CVaR. Esta franquia pode ser interpretada

como uma usina térmica candidata à construção no PPE. Na verdade, seria possível utilizar a formulação do problema de planejamento da expansão para solucionar o problema $P(\mu)$ considerando um par adicional de usinas térmicas: (i) uma térmica candidata com custo de investimento igual à μ , capacidade máxima igual à b (variável de decisão) e custo de geração igual ao custo de déficit δ ; e (ii) uma segunda térmica, também candidata, porém com custo de investimento igual à zero, capacidade infinita e custo de operação igual ao custo de déficit mais o prêmio de risco $\delta + \frac{\mu}{\alpha}$ ¹.

Analisando em detalhe, pode-se observar que o problema de planejamento da expansão tem uma estrutura próxima à da formulação apresentada neste trabalho: o estágio de decisão estratégica, presente na formulação do problema de operação com restrição de CVaR dualizada, é análogo ao subproblema mestre de investimento do problema de planejamento da expansão. A principal e fundamental diferença entre o problema proposto neste trabalho e o PPE consiste no fato de que, a cada iteração o subproblema de operação é solucionado *até a otimalidade* e é extraído *apenas um corte* com relação às variáveis de investimento para o subproblema mestre. Por sua vez, o algoritmo proposto neste trabalho gera um corte para o problema de decisão estratégica para cada iteração do problema de operação. Além de este esquema ser potencialmente mais eficiente, é sabido que o corte gerado pelo subproblema de operação só é válido se o problema for solucionado até a solução ótima².

Sendo assim, sugere-se como extensão deste trabalho a sua aplicação ao problema de planejamento da expansão, incluindo a representação das variáveis e restrições do problema mestre de investimento no problema de decisão estratégica. Esta modelagem visa permitir que seja gerado um corte de Benders para o problema de investimento a cada iteração do problema de operação dentro do algoritmos de PDDE, visando assim uma oportunidade de acelerar o processo de convergência para a solução do problema de planejamento da expansão.

9.2.2 Ajuste da Função de Custo Futuro

Os cortes de Benders calculados para representar função de custo futuro para uma determinada penalização μ não são válidos para a próxima iteração do algoritmo de solução do problema Dual Lagrangeano, onde será utilizada uma nova penalização μ' . Isto ocorre pois esta penalização influencia diretamente os custos na função objetivo de algumas das variáveis de decisão do problema e, ao alterar esta penalização, não é possível garantir que o novo problema tem viabilidade dual preservada.

¹ Nesta formulação, as térmicas representam o primeiro e o segundo segmentos de déficit.

² Isto ocorre pois a informação dual utilizada para construção dos cortes ser obtida da simulação final (passo *forward*) do algoritmo de PDDE e os cortes são extraídos através da soma dos benefícios obtidos em cada estágio do problema da operação.

Seguindo a analogia com o problema de planejamento da expansão, suponha um problema com dois pares de térmicas candidatas, onde o segundo par de térmicas é composto por: (i) uma térmica candidata com custo de investimento igual à μ' , capacidade máxima igual à b' (variável de decisão) e custo de geração igual ao custo de déficit δ ; e (ii) uma segunda térmica, também candidata, porém com custo de investimento igual à zero, capacidade infinita e custo de operação igual ao custo de déficit mais o prêmio de risco $\delta + \frac{\mu'}{\alpha}$. Adicionalmente, considera-se uma restrição que indica que apenas um dos pares de térmicas pode ser selecionado.

Este problema é análogo ao problema onde são representadas duas restrições de CVaR dualizadas, uma associada à penalidade μ e outra associada à penalidade μ' , porém apenas uma delas é ativa. A solução ótima deste problema encontra a restrição ativa e, além disto, os cortes de Benders considerarão termos adicionais que indicarão o benefício de se alterar a restrição que está ativa. Se desejarmos alterar a solução encontrada, ou seja, tornar a outra restrição ativa, os cortes poderão ser adaptados e continuarão válidos.

Se considerarmos um número maior de restrições de CVaR (pares de térmicas) com penalizações variando entre $\underline{\mu}$ e $\bar{\mu}$, teremos novos termos adicionados à FCF e a penalidade poderia ser obtida a partir de uma busca entre os valores discretizados das penalizações. Entretanto, é possível mostrar que não é necessária a representação explícita das variáveis de estado associadas às restrições de CVaR inativas para calcular os coeficientes dos termos adicionais no corte de Benders. De acordo com o trabalho de THOMÉ *et al.* [65], estes coeficientes podem ser extraídos a partir da informação dual *implícita* presente em cada problema. Ao alterar a penalização os termos de ajuste podem ser calculados a partir das informações dual já coletadas e a alteração é traduzida diretamente em correções aplicadas aos coeficientes dos cortes da FCF já existentes e que, com isso, continuam válidos.

Este desenvolvimento tem dois objetivos principais:

- Desenvolver e explorar uma metodologia que permita a reutilização dos cortes da função de custo futuro calculada para iterações anteriores, ao invés de reiniciar o processo a cada iteração. Com este procedimento de ajuste dos cortes, o método de solução descrito neste trabalho pode ser aperfeiçoado para um aumento na eficiência computacional. Duas propostas podem ser avaliadas:

- O reaproveitamento simples dos cortes, onde para cada penalização candidata o problema é solucionado até a otimalidade e, nas iterações seguintes, os cortes já calculados são reaproveitados, visando acelerar o processo de convergência para a nova penalização;

- Avaliar uma metodologia de ajuste da penalização a cada iteração do

algoritmo da PDDE. Com isso, não seria necessário solucionar o problema até a otimalidade para o ajuste da penalização. Entretanto, o passo de ajuste da penalização deve ser mais elaborado a fim de garantir a convergência do processo.

- Desenvolver uma metodologia para o planejamento da operação que permita alterações de coeficientes da função objetivo no esquema de PDDE e, com isso, a representação de outros tipos de incertezas como, por exemplo, nos custos de combustível.

9.2.3 Aplicação da metodologia proposta em problemas relacionados de outras áreas

A metodologia proposta nesta tese para a solução de problemas com representação de restrições de CVaR no ambiente de Programação Dinâmica Dual Estocástica (PDDE) é genérico e pode ser aplicado para a representação de aversão ao risco em outros problemas de programação estocástica multiestágio.

Apesar de suas especificidades, o problema de planejamento da operação de sistemas hidrotérmicos tem características semelhantes a uma outra classe de problemas que é a otimização de portfólios. O desenvolvimento de modelos de otimização de portfólio tem origem na área econômico-financeira e busca como deve ser utilizado o princípio da diversificação para otimizar as suas carteiras de investimentos, levando em conta o risco de cada um dos ativos. Neste contexto, esta metodologia torna-se uma ferramenta poderosa para obter, através de uma formulação linear, a solução ótima para um problema de maximização de renda sujeito à restrições de risco.

Referências Bibliográficas

- [1] ACERBI, C., 2007, “Coherent measures of risk in everyday market practice”, *Quantitative Finance*, v. 7, n. 4 (August), pp. 359–364.
- [2] ACERBI, C., TASCHE, D., 2001, “Expected Shortfall: a natural coherent alternative to Value at Risk”, *arXiv*, (May).
- [3] ACERBI, C., TASCHE, D., 2002, “On the coherence of Expected Shortfall”, *Journal of Banking & Finance*, v. 26, n. 7, pp. 1487–1503.
- [4] ACERBI, C., NORDIO, C., SIRTORI, C., 2001, “Expected Shortfall as a Tool for Financial Risk Management”, *arXiv*, (February).
- [5] ARARIPE NETO, T. A., PEREIRA, M. V. F., KELMAN, J., 1985, “A Risk-Constrained Stochastic Dynamic Programming Approach to the Operation Planning of Hydrothermal Systems”, *Power Apparatus and Systems, IEEE Transactions on*, v. PAS-104, n. 2 (February), pp. 273 –279. ISSN: 0018-9510. doi: 10.1109/TPAS.1985.319040.
- [6] ARTZNER, P., DELBAEN, F., EBER, J.-M., et al., 1997, “Thinking coherently”, *The Journal of Risk*, v. 10, n. 11, pp. 68–71.
- [7] ARTZNER, P., DELBAEN, F., EBER, J.-M., et al., 1999, “Coherent measures of risk”, *Mathematical Finance*, v. 9, n. 3, pp. 203–228. Disponível em: <<http://econpapers.repec.org/RePEc:bla:mathfi:v:9:y:1999:i:3:p:203-228>>.
- [8] ARTZNER, P., DELBAEN, F., EBER, J.-M., et al., 2007, “Coherent multiperiod risk adjusted values and Bellman’s principle”, *Annals of Operations Research*, v. 152, pp. 5–22. ISSN: 0254-5330. Disponível em: <<http://dx.doi.org/10.1007/s10479-006-0132-6>>. 10.1007/s10479-006-0132-6.
- [9] ASKEW, A. J., 1974, “Optimum Reservoir Operating Policies and the Imposition of a Reliability Constraint”, *Water Resources Research*, v. 10, n. 1, pp. 51–

56. ISSN: 0043-1397. Disponível em: <<http://dx.doi.org/10.1029/WR010i001p00051>>.

- [10] ASSOCIATES, F., 2013, *FICO Xpress Optimization Suite, version 7.1 - User Guide and Reference Manual*. FICO Associates. Disponível em: <<http://www.fico.com/en/Products/DMTools/xpress-overview/Pages/Xpress-and-Optimization.aspx>>.
- [11] BEALE, E. M. L., 1955, “On minimizing a convex function subject to linear inequalities”, *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 17, n. 2, pp. 173–184.
- [12] BELLMAN, R. E., 1957, *Dynamic Programming*. Princeton, NJ, Princeton University Press.
- [13] BEN-TAL, A., NEMIROVSKI, A., 1998, “Robust convex optimization”, *Mathematics of Operations Research*, v. 23, n. 4, pp. 769–805.
- [14] BEN-TAL, A., NEMIROVSKI, A., 2002, “Robust optimization - methodology and applications”, *Mathematical Programming*, v. 92, pp. 453–480. ISSN: 0025-5610.
- [15] BENDERS, J. F., 1962, “Partitioning procedures for solving mixed-variables programming problems”, *Numerische Mathematik*, v. 4, pp. 238–252. ISSN: 0029-599X. Disponível em: <<http://dx.doi.org/10.1007/BF01386316>>. 10.1007/BF01386316.
- [16] BIRGE, J. R., 1985, “Decomposition and partitioning methods for multi-stage stochastic linear programs”, *Operations Research*, v. 33, n. 5 (September-October), pp. 989–1007. doi: DOI:10.1287/opre.33.5.989.
- [17] BIRGE, J. R., LOUVEAUX, F., 1997, *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. 1st ed. New York, Springer.
- [18] CÂMARA DE GESTÃO DA CRISE DE ENERGIA ELÉTRICA - GCE, 2002. “Resolução No. 109”. Jan.
- [19] COSTA, J. P., 2007, *Árvore reduzida de cenários para o planejamento da operação de médio prazo*. Tese de Doutorado, Federal University of Rio de Janeiro - COPPE/UFRJ, Dec.
- [20] COSTA JR., L. C., PEREIRA, M. V. F., GRANVILLE, S., et al., 2013, “Representing CVaR risk aversion constraints on Stochastic Dual Dynamic

Programming with application in the hydrothermal operation planning (submetido)”, *IEEE Transactions on Power Systems*.

- [21] COSTA JR., L. C., 2008, *Incorporação de Restrições de Confiabilidade ao Problema de Planejamento Ótimo da Expansão de Sistemas Elétricos*. Tese de Mestrado, Federal University of Rio de Janeiro - COPPE/UFRJ, May.
- [22] COSTA JR., L. C., VIEIRA, B. B., BARROSO, L. A. N., et al., 2009, “Nível meta: avaliação da metodologia e dos impactos econômicos para o consumidor”. In: *XX Seminário Nacional de Produção e Transmissão de Energia Elétrica*, November.
- [23] COSTA JR., L. C., PEREIRA, M. V. F., GRANVILLE, S., et al., 2013, “Planejamento da operação de sistemas hidrotérmicos com aversão ao risco: uma abordagem para controlar o risco de déficit e limitar a energia não suprida”. In: *XXII Seminário Nacional de Produção e Transmissão de Energia Elétrica*, November.
- [24] DANTZIG, G. B., 1955, “Linear Programming Under Uncertainty”, *Management Science*, v. 1, n. 3/4 (Apr.-Jul.), pp. 197–206. Disponível em: <<http://www.jstor.org/stable/2627159>>.
- [25] DE FRANCO, N., VENTURA, F., KELMAN, J., et al., 1982, “Impact of construction delays on the reliability of a hydro system. A Brazilian planning experience”. In: *CIGRE Conference*, Paris, France, September.
- [26] DE QUEIROZ, A. R., MORTON, D. P., 2013, “Sharing cuts under aggregated forecasts when decomposing multi-stage stochastic programs”, *Operations Research Letters*, v. 41, n. 3, pp. 311 – 316. ISSN: 0167-6377. doi: <http://dx.doi.org/10.1016/j.orl.2013.03.003>. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167637713000321>>.
- [27] DINIZ, A. L. S., TCHEOU, M., MACEIRA, M. E. P., 2012, “Uma abordagem direta para consideração do CVaR no problema de planejamento da operação hidrotérmica”. In: *XII SYMPOSIUM OF SPECIALISTS IN ELECTRIC OPERATIONAL AND EXPANSION PLANNING*, May.
- [28] DUPACOVA, J., CONSIGLI, G., WALLACE, S. W., 2000, “Scenarios for Multistage Stochastic Programs”, *Annals of Operations Research*, v. 100, pp. 25–53. ISSN: 0254-5330. Disponível em: <<http://dx.doi.org/10.1023/A:1019206915174>>. 10.1023/A:1019206915174.

- [29] GUIGUES, V., RÖMISCH, W., 2010, “Sampling-based decomposition methods for risk-averse multistage stochastic programs”, *Optimization Online*, (October). Disponível em: <http://www.optimization-online.org/DB_HTML/2010/10/2763.html>.
- [30] GUIGUES, V., SAGASTIZÁBAL, C., 2010, “The value of rolling horizon policies for risk-averse hydro-thermal planning”, *Optimization Online*, (October). Disponível em: <http://www.optimization-online.org/DB_HTML/2010/10/2756.html>.
- [31] HONGLING, L., CHUANWEN, J., YAN, Z., 2008, “A review on risk-constrained hydropower scheduling in deregulated power market”, *Renewable and Sustainable Energy Reviews*, v. 12, n. 5, pp. 1465 – 1475. ISSN: 1364-0321. doi: DOI:10.1016/j.rser.2007.01.018. Disponível em: <<http://www.sciencedirect.com/science/article/B6VMY-4NG5MRV-2/2/8722bc4372e5ab14f71a3b5767b7afef>>.
- [32] HØYLAND, K., WALLACE, S. W., 2001, “Generating Scenario Trees for Multistage Decision Problems”, *Management Science*, v. 47, n. 2 (February), pp. 295–307. Disponível em: <<http://www.jstor.org/stable/2661576>>.
- [33] HØYLAND, K., KAUT, M., WALLACE, S. W., 2003, “A Heuristic for Moment-Matching Scenario Generation”, *Computational Optimization and Applications*, v. 24, pp. 169–185. ISSN: 0926-6003. Disponível em: <<http://dx.doi.org/10.1023/A:1021853807313>>. 10.1023/A:1021853807313.
- [34] ILIADIS, N. A., PEREIRA, M. V. F., GRANVILLE, S., et al., 2007, “Portfolio Optimization of Hydroelectric Assets Subject to Financial Indicators”. In: *Power Engineering Society General Meeting, IEEE*, pp. 1 –8, June. doi: 10.1109/PES.2007.385726.
- [35] JABR, R. A., 2005, “Robust self-scheduling under price uncertainty using Conditional Value-at-Risk”, *Power Systems, IEEE Transactions on*, v. 20, n. 4 (nov.), pp. 1852 – 1858. ISSN: 0885-8950. doi: 10.1109/TPWRS.2005.856952.
- [36] JACOBS, J., FREEMAN, G., GRYGIER, J., et al., 1995, “SOCRATES: A system for scheduling hydroelectric generation under uncertainty”, *Annals of Operations Research*, v. 59, pp. 99–133. ISSN: 0254-5330. Disponível em: <<http://dx.doi.org/10.1007/BF02031745>>. 10.1007/BF02031745.
- [37] JORION, P., 2000, *Value at Risk*. New York, NY [u.a.], McGraw Hill.

- [38] KLEYWEGT, A. J., SHAPIRO, A., HOMEM-DE MELLO, T., 2002, “The Sample Average Approximation Method for Stochastic Discrete Optimization”, *SIAM J. on Optimization*, v. 12 (February), pp. 479–502. ISSN: 1052-6234. doi: <http://dx.doi.org/10.1137/S1052623499363220>. Disponível em: <http://dx.doi.org/10.1137/S1052623499363220>.
- [39] MACEIRA, M. E. P., DUARTE, V. S., PENNA, D. D. J., et al., 2008, “Ten years of application of Stochastic Dual Dynamic Programming in official and agent studies in Brazil - description of the Newave program”. In: *16th Power Systems Computation Conference (PSCC)*, July.
- [40] MASSÉ, P., 1946, *Les reserves et la regulation de l'avenir dans la vie economique*, v. I and II. Paris, Hermann.
- [41] MELO, J. C. O., 1994, *Determinação do valor da confiabilidade em sistemas compostos de geração e transmissão*. Tese de Doutorado, PUC/RJ, Rio de Janeiro.
- [42] MORTON, D. P., 1996, “An enhanced decomposition algorithm for multistage stochastic hydroelectric scheduling”, *Annals of Operations Research*, v. 64, pp. 211–235. ISSN: 0254-5330. Disponível em: <http://dx.doi.org/10.1007/BF02187647>. 10.1007/BF02187647.
- [43] NEMIROVSKI, A., SHAPIRO, A., 2007, “Convex approximations of chance constrained programs”, *SIAM Journal on Optimization*, v. 17, n. 4, pp. 969–996. doi: 10.1137/050622328. Disponível em: <http://dx.doi.org/10.1137/050622328>.
- [44] NEUMANN, J. V., MORGENSTERN, O., 1944, *Theory of Games and Economic Behavior*. Princeton, NJ, Princeton University Press. ISBN: 0691119937.
- [45] NOAKES, D. J., MCLEOD, A. I., HIPEL, K. W., 1985, “Forecasting monthly riverflow time series”, *International Journal of Forecasting*, v. 1, n. 2, pp. 179–190. Disponível em: [http://dx.doi.org/10.1016/0169-2070\(85\)90022-6](http://dx.doi.org/10.1016/0169-2070(85)90022-6).
- [46] PEREIRA, M. V. F., 1985, “Optimal scheduling of hydrothermal systems - an overview”. In: *IFAC Symposium on Planning and Operation of Electric Energy Systems*, pp. 1–9, July.
- [47] PEREIRA, M. V. F., PINTO, L., 1985, “Stochastic Optimization of a Multireservoir Hydroelectric System: A Decomposition Approach”, *Water Resources Research*, v. 21, n. 6, pp. 779–792. doi: 10.1029/

WR021i006p00779. Disponível em: <<http://www.agu.org/journals/ABS/1985/WR021i006p00779.shtml>>.

- [48] PEREIRA, M. V. F., PINTO, L., 1991, “Multi-stage stochastic optimization applied to energy planning”, *Mathematical Programming*, v. 52, pp. 359–375.
- [49] PEREIRA, M. V. F., CAMPODÓNICO, N. M., KELMAN, R., 1998, “Long-term Hydro Scheduling based on Stochastic Models”. In: EPSOM (Ed.), *Proceedings of EPSOM Conference*, September.
- [50] PFLUG, G. C., 2001, “Scenario tree generation for multiperiod financial optimization by optimal discretization”, *Mathematical Programming*, v. 89, pp. 251–271. ISSN: 0025-5610. Disponível em: <<http://dx.doi.org/10.1007/PL00011398>>. 10.1007/PL00011398.
- [51] PHILPOTT, A. B., DE MATOS, V., 2010, “Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion”, *Optimization Online*, (December). Disponível em: <http://www.optimization-online.org/DB_HTML/2010/12/2861.html>.
- [52] PSR, 2013, *OptGen methodology manual, version 7.0*. PSR, . Disponível em: <<http://www.psr-inc.com>>.
- [53] PSR, 2013, *SDDP methodology manual, version 12.0*. PSR, March. Disponível em: <<http://www.psr-inc.com>>.
- [54] ROCKAFELLAR, R. T., URYASEV, S., 2000, “Optimization of Conditional Value-at-Risk”, *Journal of Risk*, v. 2, pp. 21–41.
- [55] ROSSMAN, L. A., 1977, “Reliability-constrained dynamic programming and randomized release rules in reservoir management”, *Water Resources Research*, v. 2, n. 13, pp. 247–255.
- [56] ROTTING, T., GJELSVIK, A., 1992, “Stochastic dual dynamic programming for seasonal scheduling in the Norwegian power system”, *Power Systems, IEEE Transactions on*, v. 7, n. 1 (feb), pp. 273 –279. ISSN: 0885-8950. doi: 10.1109/59.141714.
- [57] SAMSA, M. E., 1984, “Value of reliability”. In: *Expansion planning for electrical generating systems: a guidebook*, International Atomic Energy Agency, cap. 6.3, Vienna.

- [58] SHAPIRO, A., 2009, “On a time consistency concept in risk averse multistage stochastic programming”, *Operations Research Letters*, v. 37, n. 3, pp. 143 – 147. ISSN: 0167-6377. doi: DOI:10.1016/j.orl.2009.02.005. Disponível em: <<http://www.sciencedirect.com/science/article/B6V8M-4VSB147-3/2/e5c0b0ede2763d1e7d811a439b64736d>>.
- [59] SHAPIRO, A., 2011, “Analysis of stochastic dual dynamic programming method”, *European Journal of Operational Research*, v. 209, n. 1, pp. 63 – 72. ISSN: 0377-2217. doi: DOI:10.1016/j.ejor.2010.08.007. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0377221710005448>>.
- [60] SHAPIRO, A., DENTCHEVA, D., RUSZCZYNSKI, A., 2009, *Lectures on Stochastic Programming*. Philadelphia, PA, Society for Industrial and Applied Mathematics. ISBN: 9780898718751. doi: DOI:10.1137/1.9780898718751. Disponível em: <<http://link.aip.org/link/doi/10.1137/1.9780898718751>>.
- [61] SHAPIRO, A., TEKAYA, W., COSTA, J. P., et al., 2013, “Risk neutral and risk averse Stochastic Dual Dynamic Programming method”, *European Journal of Operational Research*, v. 224, pp. 375–391.
- [62] SHERKAT, V., CAMPO, R., MOSLEHI, K., et al., 1985, “Stochastic Long-Term Hydrothermal Optimization for a Multireservoir System”, *Power Apparatus and Systems, IEEE Transactions on*, v. PAS-104, n. 8 (aug.), pp. 2040 –2050. ISSN: 0018-9510. doi: 10.1109/TPAS.1985.318779.
- [63] SNIEDOVICH, M., 1979, “Reliability-constrained reservoir control problems: 1. Methodological issues”, *Water Resources Research*, v. 15, n. 6, pp. 1574–1582. ISSN: 0043-1397. Disponível em: <<http://dx.doi.org/10.1029/WR015i006p01574>>.
- [64] TERRY, L. A., PEREIRA, M. V. F., ARARIPE NETO, T. D. A., et al., 1986, “Coordinating the Energy Generation of the Brazilian National Hydrothermal Electrical Generating System”, *INTERFACES*, v. 16, n. 1, pp. 16–38. doi: 10.1287/inte.16.1.16. Disponível em: <<http://interfaces.journal.informs.org/cgi/content/abstract/16/1/16>>.
- [65] THOMÉ, F. S., BINATO, S., PEREIRA, M. V. F., et al., 2013, “Decomposition approach for generation and transmission expansion planning with implicit multipliers evaluation”, *Pesquisa Operacional*, v. 33, pp. 1–17.

- [66] TURGEON, A., 1980, “Optimal operation of multireservoir power systems with stochastic inflows”, *Water Resources Research*, v. 16, n. 2, pp. 275–283. ISSN: 0043-1397. Disponível em: <<http://dx.doi.org/10.1029/WR016i002p00275>>.
- [67] WALLACE, S. W., FLETEN, S.-E., 2003, “Stochastic Programming Models in Energy”. In: Ruszczyński, A., Shapiro, A. (Eds.), *Stochastic Programming*, v. 10, *Handbooks in Operations Research and Management Science*, Elsevier, pp. 637 – 677. doi: DOI:10.1016/S0927-0507(03)10010-2. Disponível em: <<http://www.sciencedirect.com/science/article/B7P6G-4FM0G3F-D/2/d37a4156cb8506982813efaec41b556>>.
- [68] WETS, R. J.-B., 1974, “Stochastic programs with fixed recourse: the equivalent deterministic program”, *SIAM review*, v. 16, n. 3 (July), pp. 309–339.
- [69] YAKOWITZ, S., 1982, “Dynamic Programming Applications in Water Resources”, *Water Resources Research*, v. 18, n. 4, pp. 673–696.