

RELATÓRIO TÉCNICO

Nº: DP/DEA - 37884/10 Nº DE PÁGINAS: 45 ANEXOS: -

TÍTULO: REVISÃO BIBLIOGRÁFICA - PROGRAMAÇÃO ESTOCÁSTICA MULTI-ESTÁGIO E CRITÉRIOS DE PARADA PARA A PROGRAMAÇÃO DINÂMICA DUAL.

DEPARTAMENTO: DEPARTAMENTO DE OTIMIZAÇÃO ENERGÉTICA E MEIO AMBIENTE - DEA

ÁREA DE RESPONSABILIDADE: B200

Nº DO PROJETO: 1345

DESTINATÁRIOS:

Centrais Elétricas Brasileiras S/A - ELETROBRÁS
Av. Presidente Vargas, 409 - 12º andar
20071-003 - Rio de Janeiro / RJ

ATENÇÃO: Sr. Egídio Schoenberger

EQUIPE DE ACOMPANHAMENTO:

Luis Ernesto Areias - ELETROBRÁS
Renata Leite Falcão - ELETROBRÁS

RESUMO: Este Relatório Técnico apresenta uma pesquisa na literatura técnica que teve como objetivo fazer uma revisão bibliográfica e verificar o estado da arte em programação estocástica, especificamente com relação aos critérios de parada e análise de convergência. Essa pesquisa abrangeu trabalhos publicados em diversos periódicos e congressos nas áreas de pesquisa operacional, programação matemática e energia elétrica. Embora não estivesse restrito a esse tema, o foco do estudo foi em problemas de programação estocástica multi-estágio.

AUTORES:

Andre Luiz Diniz Souto Lima - CEPEL
Débora Dias Jardim Penna - CEPEL
Vitor Silva Duarte - CEPEL
Maria Elvira Piñeiro Maceira - CEPEL
Luiz Guilherme Barbosa Marzano - CEPEL
Jorge Machado Damázio - CEPEL
Fernanda da Serra Costa - CEPEL

PALAVRAS-CHAVE:

Programação Estocástica, Decomposição de Benders, Programação Dinâmica Dual, Critério de parada

CLASSIFICAÇÃO: CONTROLADO


GERENTE DO PROJETO
NOME: Maria Elvira Piñeiro Maceira
TEL: (21) 2598-6454 FAX: (21) 2598-6482
EMAIL: elvira@cepel.br


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

APROVAÇÃO


DIRETOR DE PESQUISA, DESENVOLVIMENTO E INOVAÇÃO
ROBERTO PEREIRA CALDAS

22 / 10 / 2010

CEPEL

Centro de Pesquisas de Energia Elétrica

PROJETO NEWAVE

Relatório Técnico

**REVISÃO BIBLIOGRÁFICA
PROGRAMAÇÃO ESTOCÁSTICA MULTI-
ESTÁGIO E CRITÉRIOS DE PARADA PARA A
PROGRAMAÇÃO DINÂMICA DUAL**

ÍNDICE

1. INTRODUÇÃO.....	1
2. CARACTERÍSTICAS GERAIS DOS PROBLEMAS TRATADOS NA LITERATURA.....	3
2.1. Estrutura dos problemas	3
2.1.1. Problemas de 2 estágios.....	4
2.1.2. Problemas multi-estágio	5
2.2. Representação das incertezas	6
2.2.1. Problema “Real”	6
2.2.2. Problema de “Árvore Completa”	7
2.3. Hipóteses consideradas no problema de otimização.....	9
2.3.1. “Solução” do problema.....	9
2.3.2. Conjunto viável poliedral.....	10
2.3.3. Recurso “relativamente completo”.....	10
2.3.4. Valor limitado para a solução ótima.....	11
2.3.5. Convexidade da “função valor” de custo de cada estágio, em relação a variações nos valores da variável aleatória ξ	11
2.3.6. Distribuição de probabilidades com suporte finito	12
2.3.7. Extensão do horizonte de estudo.....	12
2.4. Propriedades da variável aleatória	15
2.4.1. Consideração de interdependência temporal	15
2.4.2. Consideração de correlação cruzada (ou espacial).....	15
2.5. Considerações Finais.....	16
3. TÉCNICAS DE RESOLUÇÃO DO PROBLEMA ESTOCÁSTICO	17
3.1. Abordagem determinística	17
3.1.1. Resolução de forma direta	18
3.1.2. Métodos de decomposição temporal	18
3.1.3. Métodos de decomposição por cenários.....	19
3.1.4. Decomposição de Dantzig & Wolfe:.....	19
3.2. Abordagem estocástica.....	20
3.2.1. Amostragem Explícita	20
3.2.2. Amostragem implícita	23
3.3. Discussão comparativa entre abordagens.....	25
3.3.1. Abordagem Determinística X Estocástica	25
3.3.2. Amostragem implícita X explícita	26
3.4. Tamanho dos problemas resolvidos na literatura	27
4. CRITÉRIOS DE PARADA	29
4.1. Análise da qualidade da solução.....	29

4.1.1.	O grau de otimalidade da solução	30
4.1.2.	Verificação da robustez da solução em relação à variação dos cenários.	30
4.2.	Intervalos de Confiança para o valor da função objetivo.....	31
4.3.	Critérios de Parada e Intervalos de Confiança para a PDDE	33
4.3.1.	Adoção de critérios não estatísticos	34
5.	CONCLUSÕES.....	35
6.	REFERÊNCIAS.....	38

1. INTRODUÇÃO

Este Relatório Técnico apresenta um resumo de uma extensa pesquisa bibliográfica que foi realizada na literatura técnica sobre programação estocástica. Foi dada uma ênfase em problemas que, de uma forma ou de outra, apresentam alguma similaridade com o tipo de problema estocástico resolvido pela cadeia de modelos de planejamento desenvolvida pelo CEPEL e utilizada pelo ONS para o planejamento e operação do sistema elétrico brasileiro [1]. O objetivo desse trabalho foi consolidar os desenvolvimentos já realizados e verificar qual o estado da arte nesse campo, com um foco específico na análise de convergência e critérios de parada para os algoritmos que resolvem o problema estocástico utilizando técnicas de amostragem, como a metodologia de programação dinâmica dual (PDD) [2], [3].

Desta forma, o contexto em que se insere esse trabalho é o do planejamento da operação energética do sistema interligado nacional (SIN), com um foco específico no modelo NEWAVE [1]. Neste modelo, o horizonte temporal é de médio prazo (5 anos), e a representação da incerteza nas afluências aos reservatórios é feita por meio de geração de cenários sintéticos de afluência, através de um modelo autorregressivo periódico de ordem p (PAR(p)) [4].

Essa pesquisa engloba trabalhos publicados nos principais periódicos de circulação internacional que envolvem a área de pesquisa operacional, tais como:

- Annals of Operations Research;
- Computational Optimization and Applications;
- European Journal of Operations Research;
- Journal of Optimization Theory and Applications;
- Management Science;
- Mathematical Programming;
- Operations Research e Operations Research Letters;
- Siam Journal on Optimization.

Este trabalho está organizado da seguinte forma: na seção 2, descrevem-se as características e hipóteses gerais que são consideradas para os problemas estocásticos estudados na literatura abrangida por esse documento. Na seção 3, descrevem-se os tipos de técnicas de resolução adotadas, e na seção 4 faz-se uma análise específica dos critérios de parada adotados. Finalmente, na seção 5 relacionam-se as principais conclusões que podem ser extraídas a partir do estudo bibliográfico realizado.

2. CARACTERÍSTICAS GERAIS DOS PROBLEMAS TRATADOS NA LITERATURA

Nesta seção, descrevem-se as características gerais dos problemas envolvidos nessa pesquisa bibliográfica, para definir de forma mais precisa o escopo do trabalho. Além disso, apresentam-se algumas definições utilizadas ao longo desse documento, e descrevem-se as principais hipóteses consideradas na maioria dos trabalhos. Em cada seção, discute-se o contexto em que se insere o problema de otimização considerado pelos modelos NEWAVE e DECOMP.

2.1. Estrutura dos problemas

Neste relatório, a ênfase será dada em problemas estocásticos onde a variável aleatória encontra-se no termo independente¹ das restrições do problema de otimização. Portanto, a cada estágio, a matriz de restrições do problema estocástico é fixa para todos os cenários, assim como os termos da função objetivo. Desta forma, o problema de programação estocástica que é objeto deste estudo pode ser representado, de forma bem geral e simplificada, pela seguinte expressão recursiva:

$$\varphi^{t,i}(x^{t-1}, \xi_i^t) = \begin{cases} \min_{x^t} \left[f^t(x^t, \xi_i^t) + \mathbb{E}_{\xi^{t+1}} [\varphi^{t+1}(x^t, \xi^{t+1})] \right] \\ \text{s.a.} \quad Ax^{t+1} \geq b(\xi_i^t) - dx^t, \end{cases} \quad [P^{i,t}]$$

onde

$$\varphi^{t+1,i}(x^t, \xi_i^{t+1}) = \begin{cases} \min_{x^{t+1}} \left[f^{t+1}(x^{t+1}, \xi_i^{t+1}) + \mathbb{E}_{\xi^{t+2}} [\varphi^{t+2}(x^{t+1}, \xi^{t+2})] \right] \\ \text{s.a.} \quad Ax^{t+2} \geq b(\xi_i^{t+1}) - dx^{t+1} \end{cases} \quad [P^{i,t+1}] \quad (2.1)$$

e $\xi^t \equiv \{\xi_i^t, i = 1, \dots, N^t\}$, $t = 1, \dots, T$,

Segundo essa notação, $\varphi^{t,i}(x^{t-1}, \xi_i^t)$ corresponde ao subproblema resolvido no cenário i do período t , tendo como condição inicial o valor x^{t-1} do período anterior; N^t é o número de cenários existentes para o período t , $f^t(.,.)$ é a função de custo para o

¹ é o termo situado à direita nas restrições de um problema de programação linear, que corresponde aos recursos disponíveis.

período t , onde t representa cada estágio de decisão em que o problema é subdividido, e $[P^{i,t}]$ é o problema para cada estágio t e cenário $\xi^{t,i}$. Alguns trabalhos, como [5], discutem as condições para a formulação de um problema estocástico multi-estágio, para que a recursão entre os estágios esteja bem definida, assim como o valor de sua solução ótima.

Note que, em (2.1), as decisões x^{t,ξ^i} são diferentes para cada estágio t e cenário $\xi^{t,i}$, mas a função de custo futuro (FCF) é a mesma para todos os nós (cenários) de um dado estágio t . Esse conceito é válido para o modelo NEWAVE. Entretanto, não é aplicado para o modelo DECOMP [6], onde se considera um problema de árvore completa, resolvido por meio de uma amostragem do tipo explícita (vide seção 3.3.2). Assim, para um modelo DECOMP tem-se uma FCF para cada nó da árvore.

Ressalta-se que, embora o foco tenha sido em problemas com esta estrutura, a revisão bibliográfica procurou abranger também artigos que tenham desenvolvido técnicas em outros tipos de problemas, mas que possam ser utilizadas para o problema (2.1). Por exemplo, muitos dos resultados existentes na literatura para problemas do tipo (2.1) podem ser estendidos para problemas onde há aleatoriedade nos termos da função objetivo $f(x^{t,\xi^i})$.

Em relação ao número de estágios de decisão t considerados no problema, podem-se distinguir dois tipos de problema básicos:

2.1.1. Problemas de 2 estágios

A grande maioria dos trabalhos considera um problema de dois estágios, onde o primeiro estágio é determinístico, ou seja, envolvendo um único vetor de decisão x , e o segundo estágio apresenta incertezas no termo independente (“RHS”) das restrições. Esse problema é chamado de “problema estocástico de dois estágios com recurso”, cuja descrição mais aprofundada pode ser encontrada em alguns livros-texto como [7]. A Figura 2.1 mostra um esquema da árvore de decisões para esse problema.

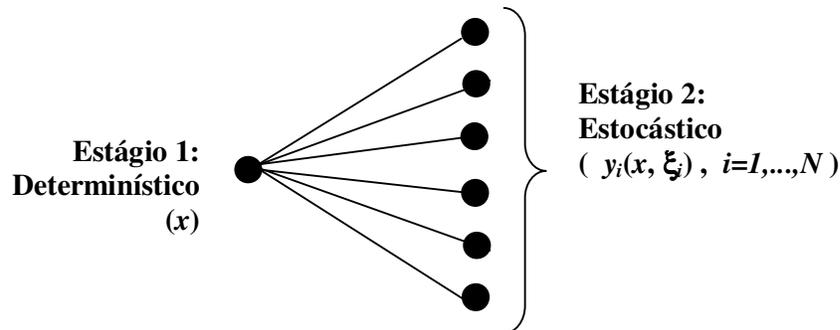


Figura 2.1 - Exemplo de um problema estocástico de 2 estágios.

2.1.2. Problemas multi-estágio

Uma pequena proporção dos trabalhos trata de um problema multi-estágio, ou seja, onde o horizonte de estudo compreende diversos intervalos de tempo. Como se considera a incerteza em todos os estágios de tempo a partir do segundo intervalo, a representação da árvore de cenários assume um caráter exponencial, conforme ilustrado na Figura 2.2.

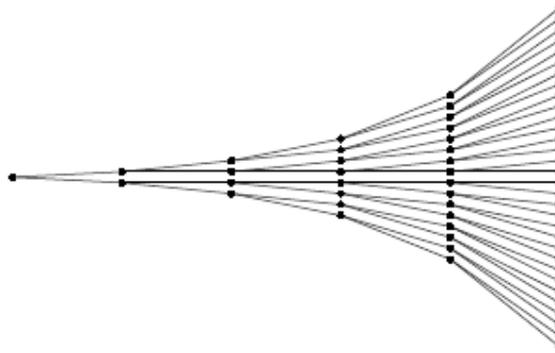


Figura 2.2 - Exemplo de um problema estocástico de 2 estágios.

Considerações

Observa-se que na literatura, tanto no problema de 2 estágios como no problema multi-estágio o problema do primeiro estágio quase sempre é determinístico, onde a solução x (única) desse estágio é um parâmetro de entrada para os estágios seguintes, onde se consideram diversos possíveis cenários.

Quando um problema estocástico de dois ou vários estágios é representado por uma árvore com um conjunto finito de possíveis cenários (“suporte finito”), pode-se formulá-lo como um grande problema de otimização. Por exemplo, se o problema for formulado como um problema de programação linear, todas as restrições e variáveis de todos os períodos e cenários poderiam ser incluídas simultaneamente em um mesmo problema de programação linear (PPL). Esta formulação é denominada de “equivalente determinístico” [8], [9].

As características do modelo NEWAVE em relação à estrutura do problema são as seguintes:

- o problema considerado pelo MODELO NEWAVE é multi-estágio, e compreende um horizonte de 180 períodos mensais¹.
- a árvore de incertezas não envolve apenas o segundo estágio em diante, mas também inclui o primeiro estágio.

É importante ressaltar que, tanto do ponto de vista teórico como principalmente do ponto de vista prático, nem sempre as técnicas de resolução aplicadas ao problema de dois estágios pode ser estendida de forma imediata para o problema multi-estágio. No segundo aspecto, isto ocorre principalmente devido ao imenso tamanho da árvore de cenários para o problema multi-estágio.

2.2. Representação das incertezas

Em relação à representação das incertezas, algumas definições fazem-se necessárias.

2.2.1. Problema “Real”

O “problema real” corresponde ao problema de engenharia que se deseja modelar. Neste caso, quando a variável aleatória assumir um número real ao longo de um determinado intervalo, sua distribuição de probabilidades será contínua. Desta forma, a avaliação de alguns parâmetros dessa variável (como o valor esperado, por exemplo) envolve uma

¹ Em aplicações de planejamento da expansão, onde o horizonte é de 10 anos mais 5 anos para eliminação da condição final, onde a FCF seria nula, ou seja, o valor da água seria zero para qualquer condição de armazenamento e tendência hidrológica.

integração multi-dimensional, cuja representação numérica em um problema de otimização é uma tarefa bastante difícil e cara do ponto de vista computacional.

2.2.2. Problema de “Árvore Completa”

Para a implementação prática de um problema de programação estocástica e sua resolução por técnicas de programação matemática, o problema real com distribuição contínua deve ser aproximado por um problema com distribuição de probabilidades discreta, no qual se diz que a variável aleatória possui suporte finito. Desta forma, o conjunto de possíveis cenários para o problema estocástico é representado através de uma árvore de cenários. Neste documento, o termo “árvore completa” será utilizado para denotar a árvore total obtida após a discretização da árvore aleatória. Portanto, se o problema estocástico possui T estágios, com uma discretização de N valores por estágio a partir do segundo período¹, essa árvore completa terá tamanho N^{T-1} . Do ponto de vista matemático, é o problema de árvore completa que se está resolvendo na prática, embora tendo sempre a preocupação de que esse represente suficientemente bem o problema real contínuo, descrito na seção 2.2.1.

A Figura 2.3 ilustra a aproximação de uma distribuição contínua por uma distribuição discreta, considerando o caso de uma variável aleatória univariada com distribuição Normal.

¹ Considerando-se a hipótese usual na literatura de representar o subproblema do primeiro período como determinístico.

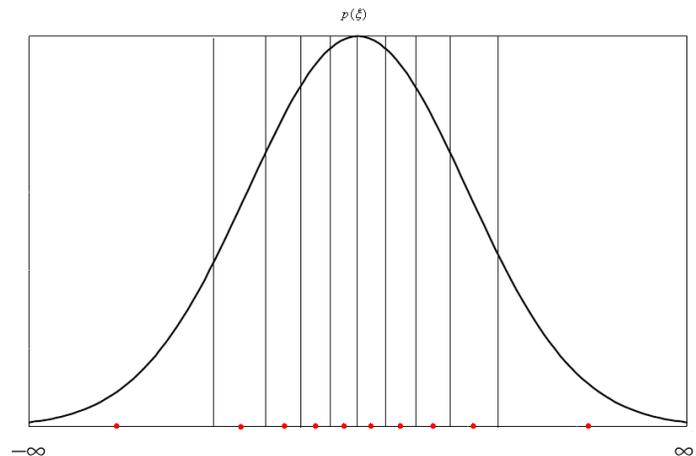


Figura 2.3 - Ilustração da aproximação da distribuição contínua por uma distribuição discreta – caso de apenas uma variável aleatória.

Na prática, quando se trabalha com múltiplos estágios e se deseja uma boa representação do problema real, a otimização considerando a representação do problema contínuo por uma árvore completa torna-se intratável computacionalmente. Em [10] e [11] são feitas revisões bibliográficas sobre as diversas técnicas de amostragem que podem ser empregadas com o objetivo de se reduzir o problema original. Esse trabalho apresenta ainda uma proposta, denominada de Amostragem Seletiva, que pode ser aplicada diretamente ao modelo NEWAVE, a qual consiste na aplicação de técnicas de agregação de cenários para construção da sub-árvore de cenários de forma a torná-la mais robusta com relação a variações da amostra utilizada e ao tamanho dessas amostras

Refinando a discretização para melhorar a aproximação do problema contínuo

O trabalho [12] estabelece uma expressão teórica para os erros incorridos na aproximação da distribuição contínua pela finita, em relação aos valores da função objetivo. Prova-se que, à medida que a discretização vai sendo refinada, o valor da função objetivo tende para o valor do problema real, com distribuição contínua. Entretanto, na prática, o cálculo exato do nível ideal de discretização é inviável do ponto de vista computacional. Em [13], procura-se resolver o problema contínuo pela técnica de *L-shaped*, descrita posteriormente na seção 3.1.2, porém alterando iterativamente a distribuição finita para obter limites inferiores e superiores para o problema contínuo. O algoritmo para quando o resultado do problema discreto estiver "suficientemente

próximo" do resultado do problema contínuo. O trabalho [14] apresenta uma discussão geral de esquemas de aproximação do problema contínuo sob dois aspectos: aproximações da distribuição real contínua e aproximações da função de custo futuro real do primeiro estágio.

Considerações

No caso do modelo NEWAVE, pode-se dizer, a princípio, que o problema de árvore completa possui 20^{120} cenários, em uma aplicação típica com 120 estágios e 20 aberturas por estágio. Entretanto, o principal resultado do modelo não é o valor em si das variáveis para o primeiro estágio, mas o conjunto de funções de custo futuro (FCF) construídas ao longo do processo. Desta forma, pode-se dizer que o problema de árvore completa possui 20^{119} cenários, uma vez que a FCF ao final do primeiro estágio independe de quantas aberturas são feitas no primeiro estágio.

2.3. Hipóteses consideradas no problema de otimização

Em geral, os trabalhos da literatura supõem uma série de hipóteses para o problema estocástico em questão, seja de dois estágios ou multi-estágio. Em seguida, discutem-se essas hipóteses e sua compatibilidade com o problema resolvido pelo modelo NEWAVE. Ao final dessa seção, apresenta-se uma tabela-resumo dessas hipóteses.

2.3.1. “Solução” do problema

Na grande maioria dos trabalhos da literatura, considera-se que a “solução” do problema corresponde ao vetor x de solução do problema do primeiro estágio, já que, do ponto de vista conceitual, essa é a “ação” que deve ser tomada no instante atual diante da árvore de incertezas no futuro. As decisões dos estágios subseqüentes serão revisadas futuramente, após a realização da variável aleatória para os estágios subsequentes, quando se resolvem novos problemas estocásticos, onde o estágio 2 passará a ser o estágio 1, e assim por diante.

Conforme já mencionado anteriormente, no processo de utilização do modelo NEWAVE para o planejamento a médio prazo da operação, o mais importante não é a

solução fornecida pelo modelo para o primeiro mês¹, mas sim o conjunto de funções de custo futuro obtidas para todos os estágios. Essas funções são importantes não só para a realização de estudos estatísticos, como por exemplo, a avaliação dos riscos de déficit para um determinado conjunto de cenários, mas também para fornecer uma função de custo futuro para o modelo de planejamento a curto prazo, DECOMP. Portanto, pode-se dizer que a solução x do modelo NEWAVE corresponde não apenas às variáveis de decisão do problema de despacho, mas também ao conjunto de cortes de Benders das funções de custo futuro de todos os estágios.

Ressalta-se que essa diferença em relação à definição do que é a solução do problema é fundamental para se analisar a aplicabilidade de várias das técnicas propostas na literatura para avaliar a qualidade da solução obtida para o problema estocástico (vide seção 4.1). Por exemplo, enquanto na abordagem usual da literatura é relativamente fácil comparar os desvios em relação a duas soluções alternativas x_1 e x_2 , no caso do modelo NEWAVE é mais difícil comparar os desvios entre funções de custo futuro diferentes.

2.3.2. Conjunto viável poliedral

Observa-se que o conjunto viável considerado pelos modelos NEWAVE e DECOMP é poliedral, já que todas as restrições do problema são representadas por meio de equações/inequações lineares², configurando um problema de programação linear. Essa propriedade atende à hipótese da grande maioria dos trabalhos encontrados na literatura.

2.3.3. Recurso “relativamente completo”

Em programação estocástica, há dois conceitos bastante difundidos:

- **Recurso completo:** diz-se que um problema apresenta recurso completo quando os subproblemas de todos os estágios são viáveis, independentemente se os valores

¹ Ressalta-se ainda que essa solução não é única, visto que, no modelo NEWAVE, o primeiro estágio já apresenta várias aberturas.

² Eventuais não linearidades são representadas por aproximações lineares ou, se for necessário, aproximações lineares por partes, como a função de produção das usinas hidroelétricas no modelo DECOMP.

para as variáveis dos estágios anteriores são viáveis para as suas respectivas restrições;

- **Recurso relativamente completo:** diz-se que um problema apresenta recurso relativamente completo quando os subproblemas de todos os estágios são viáveis, quando os valores obtidos para as variáveis dos estágios anteriores são viáveis para as suas respectivas restrições.

Na literatura, assume-se em geral o problema como sendo de recurso relativamente completo. Essa hipótese também é atendida no problema concebido pelo modelo NEWAVE, já que as restrições são modeladas com variáveis de folga (como, por exemplo, o custo de déficit), que fazem com que os subproblemas de todos os estágios sempre apresentem uma solução viável¹.

O trabalho [15] discute as condições necessárias e suficientes de otimalidade para um problema de recurso relativamente completo. Conclui-se que, nesta situação, as condições de estacionariedade de Karush-Kuhn-Tucker (KKT) continuam válidas, se os multiplicadores estiverem restritos a determinado espaço associado às restrições do problema.

2.3.4. Valor limitado para a solução ótima

Uma hipótese necessária na grande maioria dos trabalhos é que o valor da solução ótima para todos os subproblemas seja sempre limitada, ou seja, $< \infty$ e $> -\infty$. Pelas características do problema e da modelagem das restrições e variáveis, essa condição é claramente satisfeita pelo problema resolvido pelo modelo NEWAVE.

2.3.5. Convexidade da “função valor” de custo de cada estágio, em relação a variações nos valores da variável aleatória ξ .

Em um problema de otimização, define-se por “função valor” de b a função que retorna o valor da solução ótima do problema em relação a variações no valor de uma variável de estado b presente no termo independente de uma ou mais restrições.

¹ Na verdade, o problema do modelo NEWAVE atende ainda a hipótese mais forte de um problema com recurso completo.

Na literatura de programação estocástica, supõe-se que, para cada estágio t , o valor da solução ótima do problema do estágio $t+1$ até o final do horizonte de estudo é uma função convexa dos valores da variável aleatória ξ para o estágio t , como ilustrado na Figura 2.4 a seguir.

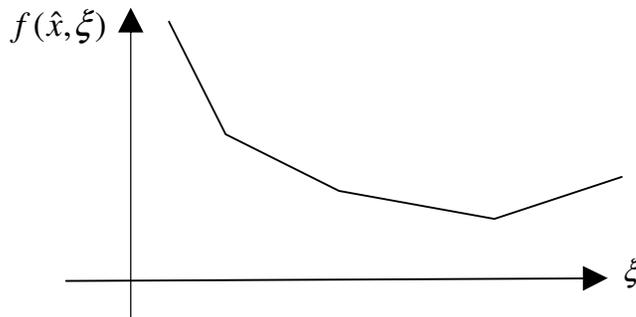


Figura 2.4 - Exemplo de uma função de custo futuro convexa

O problema estocástico resolvido pelo modelo NEWAVE é concebido como um problema de programação linear multi-estágio. Como consequência, consiste em um problema de programação convexa, e desta forma, garante-se a convexidade dessa função valor de cada estágio, em relação aos valores da variável aleatória ξ .

2.3.6. Distribuição de probabilidades com suporte finito

Essa hipótese requer que o conjunto de possíveis valores para a variável aleatória ξ seja finito, ou seja, a cardinalidade de \mathcal{N} na expressão (2.1) deve ser menor que ∞ . Embora no problema real descrito na seção 2.2.1 as vazões afluentes possam assumir um conjunto infinito de valores¹, a implementação computacional leva à definição do problema de árvore completa descrito em 2.2.2, no qual a distribuição de probabilidades é discreta.

2.3.7. Extensão do horizonte de estudo

Em teoria, um problema real de otimização estocástica possui horizonte de análise infinito, uma vez que o atendimento à demanda do sistema deve ser satisfeito de forma

¹ A distribuição adotada para representar a aleatoriedade das vazões é a distribuição log-Normal três parâmetros

contínua, ao longo do tempo. Entretanto, na prática é suficiente considerar um horizonte finito de análise, cuja extensão depende das características do sistema e dos propósitos do estudo. Do ponto de vista matemático, é também necessário estabelecer um horizonte finito de análise para que o problema possa ser resolvido por técnicas de programação matemática aplicada a problemas estocásticos.

O modelo NEWAVE é utilizado pelo Operador Nacional do Sistema Elétrico (ONS) para determinação de uma estratégia de operação mais robusta para o planejamento energético de médio prazo e pela Câmara de Comercialização de Energia Elétrica (CCEE) para o cálculo do preço de liquidação no mercado de curto prazo (horizonte de 5 anos). No contexto do planejamento da expansão energética, o Ministério das Minas e Energia (MME) e a Empresa de Pesquisa Energética (EPE) utilizam o modelo NEWAVE na elaboração dos planos de planejamento da expansão, nos leilões de energia (horizonte de 10 anos) e no cálculo da garantia física para os empreendimentos de geração de energia elétrica (horizonte de 5 anos). Em todos esses estudos, acrescentam-se mais 5 anos (estáticos, utilizando a mesma configuração do último mês do planejamento) para minimizar o efeito de esvaziamento dos reservatórios no final do horizonte. Em particular, nos casos de garantia física são adicionados mais 10 anos (estáticos, utilizando a mesma configuração do primeiro mês do planejamento) no início do estudo com intuito de eliminar o efeito das condições iniciais.

Há vários trabalhos que discutem a similaridade entre o problema de horizonte infinito e o problema de horizonte finito com uma extensão suficientemente grande. Por exemplo, em [16] discute-se como determinar a extensão do horizonte de análise (T), de forma que seja suficiente para representar bem o problema multi-estágio. Em [17], mostra-se que soluções de horizonte finito podem, sob certas condições, representar bem o problema com horizonte infinito, e que métodos baseados em decisões *steady-state* e *resource values* são apresentados de forma a diminuir o efeito do final do horizonte de estudo.

Considerações

Pode-se considerar que, em termos de propriedades matemáticas, o problema concebido pelo modelo NEWAVE atende às hipóteses usuais consideradas na literatura. O único ponto que merece mais atenção é a definição do que se considera como solução x do problema, a qual difere da maioria dos trabalhos, principalmente dos que consideram um problema de dois estágios. Entretanto, é possível encontrar alguns trabalhos, como [18], [19] onde o interesse maior está na obtenção das funções de custo futuro de cada estágio em um problema multi-estágio .

Na Tabela 2.1, apresenta-se um resumo das hipóteses analisadas nessa seção, e a sua validade para o problema considerado no modelo NEWAVE.

Tabela 2.1 - Hipóteses consideradas pelos problemas da literatura e que se aplicam ao modelo NEWAVE.

Hipótese	Descrição	Justificativa para ser válida no modelo NEWAVE
Distribuição com “suporte finito”	O conjunto de valores que podem ser sorteados é limitado	O número de aberturas consideradas na recursão <i>backward</i> é limitado
Conjunto de soluções viáveis é poliedral	O espaço viável é definido por inequações lineares	O problema é formulado como um programa linear
Conjunto de soluções ótimas é limitado e não vazio	Há pelo menos uma solução ótima e qualquer componente dessa solução é limitado, ou seja, não pode assumir um valor “ $+\infty$ ” ou “ $-\infty$ ”	Pela própria característica do problema, e pelo fato de todas as variáveis serem limitadas no PL
Recurso relativamente completo	Para qualquer solução viável para o subproblema de primeiro estágio, existe solução viável para os outros estágios	Pelo fato do não atendimento das restrições ser tratado por penalização, com a utilização de variáveis de folga
Convexidade da chamada “função-valor” $f(\tilde{x}, \xi)$,	\tilde{x} é uma solução para o problema do primeiro estágio, e ξ é a variável estocástica presente no vetor b do PL do segundo (ou demais) estágios.	Em particular, essa função é linear por partes, pelo fato do problema ser formulado por programação linear e da distribuição da variável aleatória ser discretizada para o problema de “árvore completa”. ¹ .

¹ Para o problema “real”, onde a distribuição da variável aleatória é contínua, a FCF seria não linear.

2.4. Propriedades da variável aleatória

Na grande maioria dos trabalhos encontrados na literatura, considera-se uma variável aleatória multidimensional onde suas componentes são independentes, ou seja, não há correlação temporal entre os valores de uma componente ao longo do tempo, nem correlação espacial entre os valores das diferentes componentes em um determinado estágio específico.

Nesse aspecto, há uma diferença teórica com relação à formulação do problema no modelo NEWAVE, onde as afluições aos reservatórios (que, nesse contexto, são as variáveis aleatórias), são correlacionadas com as afluições dos períodos anteriores. O processo estocástico é modelado a partir de um modelo autorregressivo periódico de ordem p (PAR(p)) e, portanto, há correlação temporal (ou interdependência temporal) entre as variáveis aleatórias. Além disso, com a existência de múltiplos reservatórios no sistema, há também uma correlação entre as suas afluições, caracterizando também uma correlação espacial entre as variáveis. Nas seções seguintes, apresenta-se uma discussão geral de como esses dois aspectos têm sido considerados na literatura.

2.4.1. Consideração de interdependência temporal

Muitos dos resultados apresentados na literatura utilizando a hipótese de independência temporal podem ser aplicados aos ruídos do modelo PAR(p), já que estes ruídos são variáveis aleatórias independentes¹. Esta questão é discutida em [20]-[22]. Entretanto, devem-se tomar os devidos cuidados antes de aplicar esses resultados diretamente aos valores de vazão afluente utilizados pelo modelo NEWAVE.

Alguns aspectos mais específicos da modelagem do problema estocástico são discutidos para o caso de existência de interdependência temporal: o trabalho [21] estuda condições para compartilhamento de cortes para problemas com interdependência de ordem (*lag*) 1 no lado direito das restrições (RHS), e com ordens maiores em outros tipos de modelos.

2.4.2. Consideração de correlação cruzada (ou espacial)

¹ Em particular, esses ruídos apresentam distribuição $N(0,1)$.

A questão da correlação espacial é tratada em alguns poucos trabalhos. Em [23], considera-se o caso em que a variável estocástica ξ é dada por $H\eta$, onde η são variáveis independentes e H é a matriz de carga (portanto, ξ tem correlação cruzada). Discute-se como aplicar as inequações de *Jensen* para refinar progressivamente o espaço de probabilidades Ξ , trabalhando-se com a discretização no espaço de η . Propõe-se uma forma para achar o η equivalente a $\xi = Tx$, pois esse valor de ξ é o mais indicado para se refinar a discretização. Em [24], utilizam-se momentos cruzados no cálculo de limites inferiores e superiores para um problema estocástico de 2 estágios. Entretanto, percebe-se que, até hoje, ainda não há uma difusão consolidada do uso de correlação cruzada na variável aleatória ξ .

2.5. Considerações Finais

Nos capítulos que se seguem, mencionam-se os trabalhos mais relevantes que foram encontrados em relação ao desenvolvimento de técnicas para a resolução de problemas de otimização estocástica. Os trabalhos estão organizados conforme o assunto principal que abordam. Revisões bibliográficas mais exaustivas sobre o tema podem ser encontradas em [25]-[31]. Por exemplo, em [27] apresenta-se uma revisão dos trabalhos existentes até o ano de 1999, e no site [31] há uma listagem atualizada de trabalhos nesse assunto.

3. TÉCNICAS DE RESOLUÇÃO DO PROBLEMA ESTOCÁSTICO

As técnicas de resolução de um problema estocástico podem ser subdivididas, de forma geral, em duas formas de abordagem, denominadas de **determinística** e **estocástica**, apresentadas a seguir. Para cada uma destas formas, descrevem-se os principais métodos e algoritmos apresentados até então na literatura.

3.1. Abordagem determinística

Neste tipo de abordagem enumeram-se *a priori* todas as possíveis combinações para as variáveis aleatórias do problema, e representa-se explicitamente essa “árvore completa” no problema de otimização, através de uma formulação matemática conhecida como equivalente determinístico [8]. Esta forma de representação do problema estocástico será referenciada ao longo desse documento como “problema de árvore completa”.

Nesta formulação, todas as restrições e variáveis de decisão para os estágios de tempo estocásticos são replicadas tantas vezes quantos forem os cenários considerados no problema. De forma a representar adequadamente os diversos níveis hierárquicos (estágios) de tomada de decisões, restrições de não antecipatividade devem ser inseridas ao problema. Tais restrições impõem que, se dois ou mais cenários são indistinguíveis até o estágio t (em termos das realizações passadas das variáveis aleatórias), então todas as variáveis de decisão replicadas até o estágio t devem ter o mesmo valor para esses cenários.

Uma vez formulado o problema matemático, este pode ser resolvido aplicando-se uma determinada técnica de otimização, que pode envolver ou não a decomposição do problema em vários subproblemas. Independentemente da técnica utilizada, **o critério de parada é sempre determinístico**, ou seja, a tolerância adotada refere-se apenas ao grau de otimalidade desejada para a solução obtida.

As principais técnicas de otimização utilizadas na literatura que se enquadram nessa forma de abordagem são relacionadas a seguir, divididas em quatro tipos, descritos nas seções a seguir.

3.1.1. Resolução de forma direta

Uma vez formulado o problema de equivalente determinístico do problema estocástico de “árvore completa”, pode-se resolvê-lo como um grande problema de programação linear. Devido ao porte do problema resultante, foram propostos alguns aprimoramentos ao método simplex, tradicionalmente empregado para resolver problemas de programação linear [25]. Em [32], apresenta-se uma revisão da aplicação de alguns desses métodos em problemas de programação estocástica.

Em [33], propõe-se um algoritmo de programação não linear, baseado em *dynamic splitting*, que utiliza Lagrangeanos aumentados, e são apresentados problemas práticos da PG&E, de até 24 meses e 27 cenários.

3.1.2. Métodos de decomposição temporal

Nos métodos desse tipo, o problema é decomposto em diversos estágios, um para cada instante de tempo na discretização temporal adotada. Em cada estágio, há diversos subproblemas, de acordo com os cenários correspondentes a cada estágio, na estrutura de árvore utilizada para representar o problema estocástico.

O método que tem sido mais empregado na literatura é o chamado de “L-shaped”, que foi proposto por Slyke e Wets [34] para problemas de dois estágios. Este método foi uma extensão, para programação contínua, da decomposição de Benders proposta inicialmente para problemas inteiros [35], adicionando-se também cortes de viabilidade no processo de resolução do problema. Posteriormente, este método foi estendido para problemas multi-estágio por Birge [16], recebendo também a denominação de *Nested decomposition*, que tem sido traduzido como “decomposição aninhada”. Há muitas variantes do método L-shaped, que surgiram em diversas épocas em função de diversos aprimoramentos que foram propostos ao método original de Slyke e Wets, seja na forma de construção dos cortes ou na forma de percorrer a árvore de cenários.

Em [36], desenvolveu-se uma versão do *L-shaped* com cortes múltiplos (*multicut*) para problemas de dois estágios, a qual foi estendida para problemas multi-estágios por [37]. Em [38], propõe-se a utilização de cortes inexatos para construção da função de custo futuro em cada nó. Em [37] também são discutidas diversas formas de se caminhar pela

árvore no algoritmo L-shaped, de forma a tornar mais eficiente a resolução do problema multi-estágio. O modelo desenvolvido nesse artigo, denominado de MSLiP, foi aplicado para problemas com até 7 estágios, sendo estendido para até 9 estágios em [39].

Foram propostos também algoritmos para regularização da função de planos cortantes gerada pelo método L-shaped, através da adoção de termos quadráticos pelo método proximal, dando origem ao chamado método de decomposição regularizada [32]. Em [40] são propostos / utilizados 4 aprimoramentos em relação ao algoritmo tradicional de Benders (L-shaped) tradicional: *warm starts*, procedimento de geração de cortes preliminares, utilização de cortes múltiplos e estratégias mais eficientes de percorrer a árvore (*tree transversing strategies*).

3.1.3. Métodos de decomposição por cenários

Outra forma de decomposição do problema de “árvore completa” é relaxar as restrições que impõem as condições de não antecipatividade das decisões ao longo dos estágios do problema estocástico, obtendo-se assim diversos cenários “em pente”. Em seguida, aplica-se algum método de otimização primal-dual para, ao mesmo tempo em que se otimizam os subproblemas resultantes da relaxação, forçar o atendimento das restrições relaxadas, a fim de se buscar o ótimo do problema original. Dois desses métodos são o *progressive hedging* [41] e o Lagrangeano aumentado [42]. Uma revisão da aplicação dos métodos de decomposição por cenários em programação estocástica é feita em [32].

3.1.4. Decomposição de Dantzig & Wolfe:

Este algoritmo consiste em dividir as restrições do problema em dois subconjuntos: um representado explicitamente e outro representado de forma iterativa, como uma combinação linear dos pontos e raios extremos que representam o poliedro que define o conjunto viável [43]. Este método também é chamado de “geração de colunas”, já que, no problema mestre, adicionam-se gradativamente colunas referentes aos pontos / raios extremos que vão sendo encontrados ao longo do processo de resolução do problema. Pode-se mostrar que esta forma de decomposição pode ser vista como um algoritmo dual ao método de Benders / L-shaped discutido na seção 3.1.2.

Uma revisão da aplicação de métodos de Dantzig & Wolfe para problemas estocásticos é feita em [32]. De forma geral, percebe-se que a aplicação desta técnica se dá em uma menor escala se comparada aos métodos descritos anteriormente.

3.2. Abordagem estocástica

Nesta forma de abordagem, assume-se que não é possível resolver o problema através de enumeração direta de todas as possíveis combinações das realizações das variáveis aleatórias, representando-se então apenas uma sub-árvore dessas possíveis realizações. A natureza estocástica do problema pode ser tratada de duas formas básicas: resolvendo-se o problema diversas vezes com diferentes sub-árvores (amostragem explícita), ou resolvendo-se o problema uma vez porém amostrando novas sub-árvores na medida em que o problema vai sendo resolvido (amostragem implícita). Estas abordagens tem sido chamadas na literatura de *exterior approach* e *interior approach*, respectivamente em [22] e [44].

Em ambos os tipos de amostragem, descritos na seqüência, o critério de parada é estocástico, e leva em consideração o nível de significância estatístico de sua otimalidade.

3.2.1. Amostragem Explícita

Na estratégia de amostragem explícita, os procedimentos de construção da árvore de cenários e resolução do problema se dão de forma separada: primeiro amostra-se uma determinada sub-árvore de todas as possíveis realizações das variáveis aleatórias, e em seguida resolve-se o problema como se fosse determinístico, utilizando uma das estratégias descritas na seção 3.1.

Embora haja diversas variantes dessa abordagem na literatura, o procedimento geral é feito da seguinte maneira: amostra-se uma sub-árvore inicial do processo estocástico, e resolve-se o problema para essa árvore, obtendo-se uma solução $x^{(1)}$ que, tipicamente, se refere aos valores da solução para o primeiro estágio. Como não se tem uma garantia, *a priori*, de que a solução obtida para a sub-árvore amostrada é “boa” o suficiente para o problema original, o processo de geração de uma sub-árvore e resolução do problema é repetido diversas vezes, obtendo-se soluções $x^{(2)}$, $x^{(3)}$, ..., $x^{(N)}$, para N sub-árvores

diferentes. Aplica-se então uma métrica para avaliar o quão próximas estão essas diferentes soluções, o que indicaria que as sub-árvores amostradas representam bem o processo estocástico. Caso haja diferença sensível entre as soluções, aumenta-se o tamanho da árvore, e reinicia-se todo o processo novamente.

Considere que o problema estocástico de “árvore completa” tem solução ótima cujo valor é Z^* . Em cada repetição i do processo descrito no parágrafo anterior, resolve-se o problema com uma sub-árvore de tamanho N , e tem-se um valor ótimo $Z_N^{(i)*}$ para a sub-árvore i . Se esse processo for repetido n vezes, teríamos n valores aleatórios $Z_N^{(i)*}$, $i = 1, \dots, n$, cuja média (que é uma variável aleatória), denotada por $Z_{N,n}^*$, é dada por:

$$Z_{N,n}^* = \frac{\sum_{i=1}^n Z_N^{(i)*}}{n} \quad (3.1)$$

Quando o número de repetições n tender ao infinito, obtemos o valor esperado Z_N^* para o problema estocástico representado através de sub-árvores de tamanho N :

$$Z_N^* = E[Z_{N,n}^*] = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n Z_N^{(i)*}}{n} \quad (3.2)$$

Alguns resultados gerais teóricos deste tipo de abordagem apresentados na literatura são relacionados a seguir, extraídos de [44].

- Z_N^* é um limite inferior absoluto para Z^* ;
- Ao realizarmos n repetições, a média obtida $Z_{N,n}^*$ é um limite inferior estatístico para Z^* . Isto quer dizer que, apesar de $Z_{N,n}^*$ poder assumir um valor superior a Z^* , seu valor ótimo (que corresponde ao valor ótimo médio obtido resolvendo-se n vezes o problema com cenários de tamanho N) tem um viés para baixo, em relação ao valor ótimo Z^* do problema original

- À medida que aumentamos o número de cenários, o valor esperado da solução do problema de sub-árvore tende assintoticamente para o valor ótimo do problema original:

$$\lim_{N \rightarrow \infty} Z_N^* = Z^* \quad (3.3)$$

As propriedades assintóticas dos estimadores para os problemas de programação estocástica, em particular do valor ótimo do problema e da solução ótima, são estudados em [45]-[47].

Outros resultados importantes são:

- para um ponto viável qualquer x fixo, o valor da função objetivo para esse ponto para uma dada árvore i de tamanho N (denotado por $Z_N^{(i)}(x)$) é um limite superior estatístico para Z_N^* . Ou seja, apesar de $Z_N^{(i)}(x)$ poder ser inferior a Z_N^* (se a árvore escolhida for, por exemplo, “muito otimista”), em média ele será superior a Z_N^* e, em decorrência do resultado apresentado no primeiro item da página anterior, também será um limite superior para Z^* ;
- Seja x_N^* a solução ótima do problema de primeiro estágio obtida com uma árvore de tamanho N . Em geral, x_N^* se aproxima assintoticamente da solução ótima (x^*) do problema original. Isto quer dizer que, para obter uma solução com qualidade aceitável, pode-se esperar que o tamanho de amostra N necessário para que isso aconteça seja bem menor do que o tamanho da população;
- Diz-se que um problema é mal condicionado quando o valor de tamanho de amostra N necessário para se ter uma solução de boa qualidade é muito grande. Teoricamente, pode-se definir um número (denominado na literatura de *condition number*) que indicaria se um problema é bem ou mal condicionado [48]. Entretanto, conforme mencionado pelos autores, esse número, na grande maioria dos casos, é impossível de ser calculado na prática;
- Como condição geral, pode-se dizer que se o ponto de ótimo for bem definido próximo de sua vizinhança (ou seja, a região próxima a solução ótima é dita *sharp*),

então em princípio o problema é bem condicionado. Entretanto, se a região próxima à solução ótima for “flat” (ou seja, há um grande número de soluções com valores ótimos muito próximos de Z^*), então o problema é mal condicionado.

- Métodos de amostragem baseados na técnica de hipercubo latino em geral necessitam de um valor menor de tamanho de amostra N para obter soluções de boa qualidade, quando comparados com métodos de amostragem Monte Carlo com amostragem aleatória simples .

Um dos principais métodos apresentados na literatura associados à abordagem estocástica com amostragem explícita das sub-árvores de cenários é o *Sample Average Approximation* (SAA), que tem originado uma série de artigos [49]-[51]. Entretanto, a aplicação deste método tem se limitado a problemas com dois estágios. Em [22], a amostragem explícita foi aplicada a problemas multi-estágio, e os problemas associados a cada sub-árvore são resolvidos a partir de aprimoramentos do método *L-shaped*.

Conforme será discutido posteriormente na seção 3.3, a grande desvantagem da amostragem explícita é a necessidade de resolução do problema diversas vezes, o que encarece muito a sua utilização do ponto de vista prático. Este é um dos principais motivos pelo qual ele tem sido aplicado a problemas de dois estágios, cujo porte é duas ordens de grandeza menor do que o problema de planejamento hidrotérmico de médio prazo, que pode ter 120 estágios¹.

3.2.2. Amostragem implícita

Na estratégia de amostragem implícita, os cenários são amostrados na medida em que o problema vai sendo resolvido, o que é feito em geral por técnicas de otimização que envolvem a decomposição do problema. Os limites inferiores e superiores estatísticos para a solução ótima do problema original vão sendo obtidos ao longo da resolução do problema, e o processo para quando um determinado critério estatístico de otimalidade é atingido. Para a definição do critério de parada, que é discutido com mais detalhes na seção 4, utilizam-se alguns dos conceitos apresentados ao longo da seção 3.2.1.

¹ ressaltando-se que os problemas de planejamento da expansão têm 180 estágios e os de cálculo da energia garantida têm 240 estágios.

As principais técnicas apresentadas na literatura relacionadas a essa forma de abordagem são relacionadas a seguir:

- ***Stochastic decomposition***: é um método desenvolvido por Hagle e Sen [52], concebido inicialmente para problemas de dois estágios [53] e estendido recentemente para problemas multi-estágio [54]. Este método consiste em resolver o problema pelo método de *L-shaped*, inicialmente com um número muito reduzido de cenários, e adicionando um novo cenário em cada iteração. Os cortes de Benders construídos em iterações anteriores são multiplicados por um fator relacionado ao grau de acréscimo no número de cenários por iteração, para que esses cortes permaneçam válidos para as próximas iterações. Em [55], esse algoritmo foi aprimorado, introduzindo-se um termo quadrático de regularização para suprir algumas das deficiências do método de planos cortantes. O critério de parada desse artigo é baseado na avaliação da qualidade da solução final x obtida
- **Decomposição de Benders com *importance sampling***: aplica decomposição de Benders de dois estágios, mas sem percorrer todos os cenários do segundo estágio. Neste caso, os cortes (denominados de “cortes probabilísticos”) não são válidos como uma aproximação inferior da função. Considerando uma distribuição normal para o termo independente (RHS) dos cortes, deriva-se uma expressão para a variância dos limites inferior e superior para a solução ótima, construindo-se assim um intervalo de confiança que é utilizado como critério de parada. Um dos diferenciais dessa técnica é a forma como os cenários são amostrados, que se baseia na técnica de *importance sampling* [56]. As aplicações também se iniciaram com problemas de dois estágios, e depois se estenderam para problemas multi-estágio [39], [57], [58].
- **Programação dinâmica dual estocástica (PDDE)**: é o método de decomposição de Benders multi-estágio, desenvolvido em [3] e utilizado atualmente no modelo NEWAVE, empregado para o planejamento da geração a médio prazo do sistema brasileiro [2]. Este método também tem sido utilizado para o planejamento do sistema elétrico na Noruega [59], [60] e estudado por outros pesquisadores [18];

- **Outros algoritmos para problemas multi-estágio:** no que diz respeito à aplicação para problemas com mais de dois estágios, podemos citar os trabalhos [61], [62], que propõem metodologias intermediárias entre resolver o problema de árvore completa (inviável do ponto de vista prático) e o problema com uma árvore reduzida. A idéia geral é construir uma árvore com muitas ramificações, porém, para evitar a explosão do número de nós (subproblemas a serem resolvidos), aglutinam-se as soluções de diversos nós de um mesmo estágio em um único nó. Esse tipo de procedimento é denominado em [61] de “Programação Linear por Eventos Multi-estágio”, onde se aglutinam os nós através da média dos resultados de vários cenários. Em [62], esse procedimento é denominado de “Recombinação de Árvore de Cenários” e faz-se a aglutinação através de uma agregação dos resultados obtidos, selecionando apenas alguns resultados de cada estágio. Outro algoritmo, proposto em [63], consiste em realizar apenas uma abertura por estágio, construindo cortes que, embora possam ser bem otimistas, continuam válidos.

O trabalho [57] menciona que os métodos baseados em amostragem Monte Carlo são muito lentos, e propõe técnicas de redução de variância para melhorar sua performance. Em [18], é apresentada a prova de convergência da PDDE em um número infinito de iterações. Neste artigo, os autores não analisam a robustez dos algoritmos com relação à árvore completa de cenários, restringindo-se a analisar a convergência. Na formulação apresentada no artigo, pressupõe-se que o processo estocástico é um processo discreto com um número finito de possíveis realizações, e que a sub-árvore de cenários é sorteada novamente a cada iteração *forward* do algoritmo. Adicionalmente, o artigo sugere que as amostras utilizadas nos processos *forward* e *backward*, da qual são extraídas as sub-árvores de cenários, devem ser as mesmas.

3.3. Discussão comparativa entre abordagens

3.3.1. Abordagem Determinística X Estocástica

A abordagem determinística apresenta duas desvantagens importantes:

- devido à necessidade de enumeração completa de todas as combinações para as variáveis aleatórias do problema, dificilmente consegue-se representar problemas

com mais de dois estágios de decisão estocásticos. Além disso, o número de variáveis aleatórias não pode ser muito grande, pois o problema de otimização pode se tornar intratável do ponto de vista computacional;

- não se obtém, ao final da resolução do problema, uma avaliação estatística da qualidade da solução final obtida. Ou seja, apesar da solução ser ótima para a árvore que foi amostrada, não se tem uma análise de sensibilidade da qualidade dessa solução se outra árvore tivesse sido amostrada.

Devido a esses dois fatores, a aplicação da abordagem determinística na literatura se restringe basicamente a problemas de dois estágios, ou, quando se utiliza uma formulação multi-estágio, a problemas com um pequeno número de aberturas por estágios. Para aplicações reais como o problema do modelo NEWAVE, torna-se necessário então a adoção da abordagem estocástica, cuja aplicação é discutida na seção seguinte

3.3.2. Amostragem implícita X explícita

Uma vez identificada que a abordagem estocástica é a única factível de ser aplicada na prática para o modelo NEWAVE, devem-se discutir as vantagens e desvantagens de ambos os tipos de amostragem apresentados na seção 3.2.

A amostragem explícita tem como vantagem o fato de que o problema associado a determinada sub-árvore pode ser resolvido como um problema de otimização determinístico, aplicando-se, por exemplo, uma das técnicas de decomposição descritas na seção 3.1. Desta forma, para cada sub-árvore amostrada ao longo das repetições do algoritmo, obtém-se uma solução com um critério de otimalidade bem definido.

Entretanto, a grande desvantagem deste tipo de amostragem é que se deve resolver o problema de otimização diversas vezes, uma para cada sub-árvore. As aplicações na literatura mostram que isso só é possível na prática para problemas de poucos estágios, razão pela qual o método de *Sample Average Approximation* (SAA) só tem sido aplicado para problemas de dois estágios. Para se ter uma idéia da dificuldade de aplicação dessa abordagem, se considerarmos apenas duas aberturas por estágio no

problema do modelo NEWAVE, já teremos um número de cenários igual a 2^{120} , que já está muito além da capacidade do computador comercial mais potente.

Por esse motivo, se conclui que a única forma factível de se tratar o problema do modelo NEWAVE é através da **abordagem estocástica com amostragem implícita dos cenários**, que é o tipo de estratégia em que se classifica o método de PDDE, adotado atualmente.

3.4. Tamanho dos problemas resolvidos na literatura

A fim de reforçar a conclusão feita no final da seção anterior, fez-se um levantamento do tamanho dos problemas resolvidos pelas várias técnicas de otimização discutidas ao longo deste capítulo. A Tabela 3.1 apresenta um resumo das maiores aplicações encontradas na literatura, considerando-se apenas problemas multi-estágio.

Tabela 3.1 – Algumas das maiores aplicações na literatura de técnicas de otimização estocástica para problemas multi-estágio.

Trabalho	Método	Abordagem	Número de estágios	Observação
[32]	Benders regularizado	Determ.	8	
[37]	Aprimor. <i>L-shaped</i>	Determ.	7	
[22]	Aprimor. <i>L-shaped</i>	Estocástica / Amost. Explic.	10	-
[40]	Aprimor. <i>L-shaped</i>	Determ.	24	Sistema da PG&E, EUA
[38]	Aprimor. <i>L-shaped</i>	Estocástica / Amost. Explic.	6	Sistema da Nova Zelândia, 22UHE, 6UTE
[39]	Benders <i>c/ Importance sampling</i>	Estocástica / Amost. Explic.	9	-
[59], [60]	PDDE	Estocástica / Amost. Implic.	24	Sistema da Noruega 35 UHE
[33]	<i>Dynamic splitting</i> (PNL)	Determ.	24	Sistema da PG&E, EUA 27 cenários
[1]	PDDE	Estocástica / Amost. Implic.	120	SIN – Brasil (NEWAVE)
[1]	<i>L-shaped</i>	Determ.	3	SIN – Brasil (DECOMP)

Trabalho	Método	Abordagem	Número de estágios	Observação
[54]	<i>Stochastic Decomp.</i>	Estocástica / Amost. Explic.	5	-

Percebe-se uma predominância das abordagens estocásticas em relação às determinísticas. As maiores aplicações foram:

- os trabalhos [33] e [40], que utilizam uma abordagem determinística e foram utilizados para resolver o problema de coordenação hidrotérmica da PG&E, com horizonte de até 2 anos e discretização mensal (24 estágios). Deve-se ressaltar, entretanto, que estes trabalhos trataram problemas com apenas 27 cenários no total
- o trabalho [59], que utiliza a mesma técnica de PDDE aplicada ao modelo NEWAVE, para resolver o problema de coordenação hidrotérmica do sistema norueguês, que é subdividido em dois subsistemas.

Ressalta-se que nenhuma das aplicações mostradas na Tabela 3.1 se aproxima, em termos de tamanho, ao problema tratado pelo modelo NEWAVE, razão pela qual se torna difícil fazer uma comparação direta da performance dos algoritmos na literatura em relação à performance do modelo NEWAVE.

4. CRITÉRIOS DE PARADA

Conforme mencionado anteriormente, as abordagens determinísticas descritas na seção 3.1 utilizam como critério de parada somente uma determinada tolerância numérica para a otimalidade da melhor solução já encontrada. Entretanto, as abordagens estocásticas descritas na seção 3.2 não podem considerar um critério de parada semelhante ao caso determinístico, já que há uma aleatoriedade intrínseca no valor da solução ótima, uma vez que não é possível resolver o problema de “árvore completa” de uma só vez.

Portanto, ao se utilizar uma abordagem estocástica, surge uma importante questão: se o problema original está sendo resolvido de forma aproximada, através de amostragem, como avaliar a qualidade da solução encontrada até então? Torna-se necessário, desta forma, desenvolver critérios de parada “satisfatórios”, que ofereçam uma determinada “garantia de qualidade” da solução obtida. Tais critérios devem ser elaborados, necessariamente, lançando mão de conceitos estatísticos, seja de forma direta ou indireta.

Os critérios de parada para os métodos que utilizam abordagem estocástica podem ser aplicados analisando-se os valores das componentes do vetor solução x encontrado ou o valor “ Z ” da função objetivo correspondente a essa solução. A primeira alternativa, discutida na seção 4.1, é aplicada em geral para os métodos de amostragem explícita (seção 3.2.1), onde o interesse maior está na avaliação da qualidade do vetor solução x para o problema de primeiro estágio. A segunda alternativa, discutida na seção 4.2, é utilizada em ambos os tipos de amostragem, e o interesse maior nesse caso é avaliar se os limites inferior e superior obtidos para a solução ótima estão “suficientemente próximos”, de forma a serem indistinguíveis do ponto de vista estatístico. Finalmente, na seção 4.3.1 discutem-se alguns critérios não estatísticos que foram propostos na literatura.

4.1. Análise da qualidade da solução

Os critérios que se baseiam na análise da qualidade da solução são utilizados principalmente nos métodos de amostragem explícita, nos quais se resolve diversas vezes o problema, uma para cada sub-árvore da amostra. A solução analisada

corresponde, em geral, ao ponto x^k obtido para o subproblema do primeiro estágio, onde k indica a iteração do algoritmo. Dois aspectos importantes dessa solução, descritos nas seções seguintes, tem sido analisados:

4.1.1. O grau de otimalidade da solução

Em relação a esse aspecto, a preocupação maior tem sido a avaliação da satisfação das condições de KKT para o ponto x^k . Em [44], no contexto do método SAA de amostragem explícita (seção 3.2.1), os autores afirmam que essa avaliação é inviável do ponto de vista prático, além de ter alguns problemas de consistência do ponto de vista teórico. Apesar disso, em [50] a investigação da convergência do método SAA é baseada nos critérios de KKT à medida que o tamanho da amostra cresce. É feita uma derivação matemática e apresentam-se alguns resultados numéricos para um caso de duas variáveis e apenas um estágio estocástico. Entretanto, o referido trabalho não apresenta conclusões contundentes.

Já no contexto dos métodos de amostragem implícita, mais especificamente para o método de *stochastic decomposition* de Higel & Sen (seção 3.2.2), a análise é focada nos testes das variáveis duais das restrições de não antecipação. Baseado na geração de L árvores de tamanho M , verificam-se a aplicação das condições de otimalidade dos multiplicadores para a maioria dessas árvores [52].

4.1.2. Verificação da robustez da solução em relação à variação dos cenários

Neste caso, avalia-se a estabilidade da solução ao se alterar o tamanho e/ou os cenários da sub-árvore amostrada. No contexto do método de *stochastic decomposition* [55], é feita uma verificação se a solução “ótima” obtida até então (denominada de *incumbent solution*) também é ótima para uma reamostragem da árvore por *bootstrap*. Para o método de SAA, o trabalho [44] avalia uma determinada solução candidata \hat{x} variando a amostra e verificando se a nova solução ótima é similar a \hat{x} , considerando-se uma dada tolerância. Caso isso não se verifique, aumenta-se o tamanho da amostra N , resolve-se novamente o problema e obtém-se um novo ponto \hat{x} .

Para o modelo NEWAVE, esse processo de avaliação iterativa da solução é muito caro do ponto de vista computacional. Além disso, ressalta-se que a solução x fornecida por

esse modelo não é um ponto de operação para um determinado estágio, mas sim o conjunto de funções de custo futuro para todos os estágios, que são utilizados em diversos estudos e para acoplamento com o modelo DECOMP. Percebe-se então que, neste caso, a avaliação da solução x se torna bem mais complicada.

4.2. Intervalos de Confiança para o valor da função objetivo

Um critério de parada bastante utilizado para algoritmos de programação estocástica é avaliar o valor da melhor solução obtida, em relação a limites inferior e superior, em geral estatísticos, para o valor da solução ótima. Estes limites definem o que se chama de intervalo de confiança para o valor ótimo do problema, o qual em princípio é desconhecido. Algumas das principais conclusões apresentadas na literatura em relação a esse aspecto são discutidas a seguir:

- critérios de parada baseados apenas nas diferenças no valor da solução de uma iteração para outra (ou seja, $f(x^k) - f(x^{k+1})$), são bem mais fáceis de serem implementados, entretanto devem ser evitados por não trazerem informações sobre a qualidade da solução [53];
- É mais interessante estimar a diferença $f(x^k) - f(x^*)$, que mede de fato o grau de otimalidade do ponto x^k , em relação à solução ótima x^* . Como não se conhece $f(x^*)$, devem-se obter cotas inferiores e superiores para o valor da solução ótima, que serão denominadas ao longo desse documento de ZINF e ZSUP, respectivamente;
- Os valores de ZINF e ZSUP são estimadores estatísticos para o valor da solução ótima. Desta forma, além de estabelecer valores máximos para a diferença entre ZINF e ZSUP, é importante utilizar como critério de parada adicional a exigência de que a variância desses estimadores esteja abaixo de determinado valor desejado [49], [58], [64]. Ou seja, deseja-se que ZINF seja estatisticamente igual a ZSUP mas também que a incerteza nesses dois valores seja satisfatória;
- Um fator de grande relevância na avaliação desses limites é o tamanho e forma de obtenção das amostras de cenários (sub-árvores) utilizadas para resolver o problema. Em [65], sugere-se que o tamanho da amostra aumente ao longo do processo de resolução do problema; os trabalhos [58], [66], [67], mencionam que, para o teste

ser realizado de forma adequada, as amostras utilizadas para avaliar ZSUP e ZINF devem ser independentes. Para tal, sugere-se uma reamostragem no processo de avaliação de ZSUP.

Diversos critérios de parada têm sido estabelecidos na literatura com base na definição de intervalos de confiança. Em [66], propõe-se para um problema de 2 estágios um critério de parada onde se verifica se o valor de $f(x^k)$ é “suficientemente próximo” de $f(x^{k+1})$, com base em um intervalo de confiança de 95%, com a hipótese nula de que esses estimadores sejam iguais (ou seja, o critério de parada é atendido quando não se rejeita essa hipótese). Em seguida, verificam-se as “condições de otimalidade estatísticas” de x^k .

Nos trabalhos [22], [67], [68], propõem-se intervalos de confiança para um problema de dois estágios, onde o limite inferior é construído com base em valores de função Z^n obtidas com diferentes amostras de tamanho n , e que se constituem em limites inferiores estatísticos para Z^* . O limite superior é obtido a partir de valores de $Z^n(x)$ para alguns pontos viáveis x , os quais são limites superiores estatísticos para Z^* . Duas formas de estimar o intervalo de confiança são propostas nestes trabalhos:

- a primeira é baseada no cálculo de ZINF e ZSUP de forma independente. Neste caso, o *gap* pode ser inferior a zero e devem-se utilizar amostras independentes para estimar ZINF e ZSUP;
- a segunda é baseada nas diferenças entre ZSUP e ZINF. Neste caso, garante-se que o *gap* é sempre maior do que zero e pode-se utilizar a mesma amostra.

O trabalho [22] menciona que o primeiro caso é mais caro do ponto de vista computacional, pois são necessárias amostras diferentes para ZINF e ZSUP, porém o nível de confiança é de $(1-\alpha)^2$. O segundo caso é mais barato computacionalmente, mas o seu nível de confiança é de $(1-2\alpha)$, portanto inferior ao do primeiro caso.

Em [44], propõem-se os seguintes procedimentos na obtenção dos estimadores:

- **para ZINF:** gerar M árvores de tamanho N , e resolver o problema de forma determinística para cada árvore, obtendo-se M valores de custo. O valor de ZINF é o

valor médio dessas M árvores, e calcula-se o intervalo de confiança com base na variância obtida nessas M árvores;

- **para ZSUP:** gerar T árvores de tamanho N , escolher uma solução (por exemplo, obtida por SAA em uma determinada árvore) e testar o seu valor ótimo para cada uma das T árvores. O estimador para ZSUP será o valor médio dessas T árvores, com intervalo de confiança calculado com base na variância obtida nessas T árvores

O trabalho menciona ainda que a diferença entre ZSUP e ZINF é tendenciosa, superestimando o verdadeiro valor do *gap*. Para reduzir o vício, sugere-se aumentar a amostra ou utilizar a técnica de hipercubo latino na geração da amostra.

É importante ressaltar que, na grande maioria dos trabalhos, discute-se a construção do intervalo de confiança, mas não se estabelece claramente como deve ser o critério de parada para os algoritmos que resolvem o problema estocástico.

4.3. Critérios de Parada e Intervalos de Confiança para a PDDE

Nesta seção, discutem-se critérios de parada e intervalos de confiança aplicados especificamente ao algoritmo de programação dinâmica dual estocástica, que é a técnica aplicada no modelo NEWAVE.

Em sua proposta original [3], aplicada ao problema multi-estágio de planejamento hidrotérmico, propõe-se que o algoritmo de PDDE pare quando o limite inferior ZINF estiver dentro do intervalo de confiança para ZSUP, cujo valor adotado usualmente é de 95%. Em [60], onde o algoritmo também é aplicado a um problema de planejamento hidrotérmico multi-estágio, para ao sistema norueguês, diz-se que o algoritmo converge quando os valores de ZINF e ZSUP são “suficientemente próximos”. Como há incerteza no cálculo de ZSUP, devido à amostragem, considera-se o desvio padrão de ZSUP nesta comparação, caracterizando um critério de parada bastante semelhante ao utilizado no Brasil.

Em [69] é feita uma análise do critério de parada adotado na PDDE e em uma variante do algoritmo, denominado de ReSa, onde se realiza uma reamostragem a cada iteração. Neste trabalho, prova-se a convergência da PDDE, mas defende-se a reamostragem da

sub-árvore de cenários a cada iteração. O limite inferior ZINF deve ser estatisticamente igual ao limite superior ZSUP, assumindo-se que esse último possui distribuição de probabilidades aproximada por uma normal. Este critério também se assemelha ao adotado no modelo NEWAVE.

4.3.1. Adoção de critérios não estatísticos

Em [69], menciona-se a possibilidade de se utilizar outros critérios de parada para o algoritmo da PDDE, que não sejam estritamente estatísticos, tais como:

- verificar se o valor de ZINF não se altera em k iterações;
- verificar se essa alteração não é significativa, em termos percentuais em relação ao valor de ZINF nas últimas k iterações;
- analisar a variação percentual do *gap* entre ZSUP e ZINF;

Os autores concluem que o desempenho do critério depende do tamanho da amostra considerada.

Além desse trabalho, não foram encontrados, na literatura abrangida por esse documento, outros trabalhos que tenham proposto, de forma sólida e bem fundamentada, critérios não estatísticos para a parada do algoritmo de PDDE.

5. CONCLUSÕES

Este Relatório Técnico teve como objetivo levantar o estado da arte em otimização estocástica, de forma a auxiliar a análise das funcionalidades existentes e eventuais aprimoramentos nos modelos NEWAVE e DECOMP, atualmente utilizados pelo ONS para o planejamento da operação energética a médio e curto prazos do sistema Interligado Nacional (SIN). Em particular, focou-se na análise de convergência e critérios de parada para os algoritmos que resolvem o problema estocástico utilizando técnicas de amostragem, como a metodologia de programação dinâmica dual (PDD).

As principais conclusões que podem ser feitas após a realização desta revisão bibliográfica são:

Em relação aos problemas tratados na literatura

- Há alguns aspectos particulares do modelo NEWAVE que não são comumente estudados na literatura, como as correlações temporais e espaciais entre as diversas variáveis aleatórias consideradas no problema;
- O tamanho dos problemas estudados na literatura é muito inferior ao do problema tratado pelo modelo NEWAVE (até 120 períodos para o planejamento da operação), e a grande maioria dos trabalhos considera problemas de apenas dois estágios;

Em relação às técnicas de otimização aplicadas para resolver o problema:

- as técnicas que mais se assemelham ao modelo NEWAVE são as que consideram uma abordagem estocástica, com amostragem do tipo implícita, onde a geração dos cenários é realizada ao longo das iterações do algoritmo;
- as abordagens estocásticas com amostragem do tipo explícita, onde o problema é resolvido várias vezes com diversas sub-árvores diferentes, são bastante caras do ponto de vista computacional e são proibitivas de serem aplicadas na prática em um problema como o do modelo NEWAVE.

Em relação às técnicas de amostragem utilizadas para problemas estocásticos

- identificou-se a necessidade de utilizar diversas amostras da árvore de cenários a fim de se ter uma robustez da solução. Para tal, utiliza-se uma das duas estratégias descritas nos dois parágrafos anteriores;
- para problemas de dois estágios, as técnicas de reamostragem, em geral realizadas de forma aleatória simples, tem proporcionado robustez à solução final encontrada para o subproblema do primeiro estágio. Entretanto, para problemas multi-estágio, a obtenção de tal robustez é uma questão ainda em aberto na literatura, com poucas contribuições;
- o desafio da aplicação de técnicas de reamostragem ao problema de planejamento da operação energética brasileiro, onde o horizonte de estudo é bastante extenso, é escolher, a cada iteração do processo de solução por PDDE, uma nova amostra que de fato agregue informações ao processo estocástico.

Em relação aos critérios de parada para os algoritmos de otimização estocástica

- é importante observar não somente a estabilidade do valor da função ao longo das iterações, mas principalmente ter alguma medida de sua otimalidade. Assim, é bastante importante a definição de intervalos de confiança para o valor da solução ótima do problema;
- apesar de muitos trabalhos discutirem a necessidade de consideração de intervalos de confiança, poucos detalham como utilizá-lo de forma prática como critério de parada nos algoritmos de otimização estocástica;
- para a técnica de PDDE, adotada atualmente no modelo NEWAVE, o critério estatístico ainda é o utilizado como padrão para aplicações práticas de tomada de decisão na operação de sistemas elétricos, como são os casos dos sistemas brasileiro e norueguês.

Finalmente, é necessário avaliar qual o impacto que a funcionalidade de Amostragem Seletiva, implementada recentemente no modelo NEWAVE [10] causa no processo de convergência / critério de parada do modelo NEWAVE. A adoção de critérios não

estatísticos pode ser útil como um critério complementar aos critérios estatísticos. Assim, pode-se parar o algoritmo quando os resultados já estiverem suficientemente estáveis, porém sem perder em vista uma medida estatística de otimalidade da solução encontrada pelo modelo.

6. REFERÊNCIAS

- [1] M.E.P. Maceira, L.A. Terry, F.S. Costa, J. M. Damazio, A C. G. Melo, “Chain of optimization models for setting the energy dispatch and spot price in the Brazilian system”, *Proceedings of the Power System Computation Conference - PSCC’02*, Sevilla, Spain, June 2002
- [2] M.E.P. Maceira, V. S. Duarte, D.D.J. Penna, L.A.M. Moraes, A.C.G. Melo, “Ten years of application of stochastic dual dynamic Programming in official and agent studies in Brazil – Description of the NEWAVE program”, *16th Power Systems Computation Conference - PSCC*, Glasgow, SCO, July 2008.
- [3] M. V. F. Pereira, L. M. V. G. Pinto, “Multi-stage stochastic optimization applied to energy planning”, *Mathematical Programming*, v. 52, n.1-3, pp. 359-375, May 1991.
- [4] Maceira, M.E.P., Bezerra, C.V., 1997, “Stochastic Streamflow model for Hydroelectric Systems” In: *Proceedings of 5th International Conference on Probabilistic Methods Applied to Power Systems*, pp. 305-310, Vancouver, Canada, Sep.
- [5] P. Olsen, “When is a multistage stochastic programming well-defined?”, *Siam Journal on Control and Optimization*, v.14, n.3, pp. 0518-0527, 1976.
- [6] Centro de Pesquisas de Energia Elétrica, “Manual de Referência do modelo DECOMP”, versão 16, Fevereiro de 2010.
- [7] J. R. Birge, F. Louveaux, “Introduction to stochastic programming”, *Mathematical Programming*, v.79, pg. 333-353, 1997.
- [8] R. J-B. Wets, “Stochastic programs with fixed recourse: The equivalent deterministic program”, *SIAM Review*, v.16, n.3, pp. 0309-0339, 1974.
- [9] P. Olsen, “Multistage stochastic programming with recourse: the equivalent deterministic problem”, *Siam Journal on Control and Optimization*, v.14, n.3, pp. 0495-0517, 1976.
- [10] D.D.J.Penna, 2009, “Definição da Árvore de Cenários de Afluências para o Planejamento da Operação Energética de Médio Prazo”, Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro.
- [11] D.D.J. Penna, M.E.P. Maceira, J.M. Damázio, “Streamflow Scenario Tree: Selective Sampling in the Long-Term Operation Planning”, *10th International Symposium on Stochastic Hydraulics & 5th International Conference on Water Resources and Environment Research*, Quebec, Canadá, 2010.
- [12] P. Kall, “Approximations to stochastic programs with complete fixed recourse”, *Numerische Mathematik*, v.22, pp. 333-339, 1974.

- [13] J. R. Birge, “Using sequential approximations in the L-shaped and generalized programming algorithms for stochastic linear programs”, Tech Report 83-12 -Dept of IE & OR, Univ., Michigan, 1983.
- [14] J. R. Birge, R. J-B. Wets, “Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse”, *Mathematical Programming Study*, v. 27, pp. 54-102, 1986
- [15] R. T. Rockafellar, R. J-B.Wets, “Stochastic convex programming: relatively complete recourse and induced feasibility”, *Siam Journal on Control and Optimization*, v.14, n.3, pp. 574-589, 1976
- [16] Birge, J.R., “Decomposition and partitioning methods for multistage stochastic linear programs”, *Operations Research*, v.33, n.5, pp. 989-1007, 1985.
- [17] R. C. Grinold, “Model building techniques for the correction of end effects in multistage convex programs”, *Operations Research*, v.31, n.3, pp. 0407-0431, 1983.
- [18] A. B. Philpott, Z. Guan, “On the convergence of stochastic dual dynamic programming and related methods”, *Operations Research Letters*, v.36, n.4, pp. 450-455, 2008.
- [19] D. P. Morton, “An enhanced decomposition algorithm for multistage stochastic hydroelectric scheduling”, *Annals of Operations Research*, v. 64, pp. 211-235, 1996.
- [20] K. Frauendorfer, “Solving SLP Recourse problems with arbitrary multivariate distributions - the dependent case”, *Mathematics of Operations Research*, v.13, n.3, pp. 377-394, 1988.
- [21] G. Infanger, D. P. Morton, “Cut sharing for multistage stochastic linear programs with interstage dependency”, *Mathematical Programming*, v.75, n.2, pp. 241-256, 1996.
- [22] A. Chiralaksanakul, D. P. Morton, “Assessing policy quality in multi-stage stochastic programming”, Technical Report, The University of Texas at Austin, 2003.
- [23] J. R. Birge, S. W. Wallace, “Refining bounds for stochastic linear programs with linearly transformed independent random variables”, *Operations Research Letters*, v.5, n.2, pp. 73-77.
- [24] N. C. P. Edirisinghe, W. T. Ziemba, “Bounds for two-stage stochastic programs with fixed recourse”, *Mathematics of Operations Research*, v.19, n.2, pp. 292-313, 1994.
- [25] Ermoliev, Yu, Wets, R. J-B., “Numerical techniques for stochastic optimization”, Springer Series in Computational Mathematics, v.10, Springer-Verlag, 1988.
- [26] P. Kall, S. W. Wallace, “Stochastic programming”, John Wiley & Sons, 1994.

- [27] J. Dupacova, “Multistage stochastic programs: The state-of-the-art and selected bibliography”, *Kybernetika*, v.31, n. 2, pp. 151-174, 1995.
- [28] J. R. Birge, “Stochastic programming computation and applications”, *INFORMS Journal on Computing*, v.9, n. 2, pp. 0111-133, 1997.
- [29] D. P. Morton, “Overview of stochastic programming applications”, Technical Report, Dash Optimization, 2002.
- [30] J. Dupacova, “Applications of stochastic programming: Achievements and questions”, *European journal of operational research*, v.140, n.2, pp. 281-290, 2002
- [31] M. H. Vlerk, , Stochastic Programming Bibliography, disponível em <http://mally.eco.rug.nl/index.html?spbib.html>
- [32] A. Ruszczynski, “Regularized decomposition of stochastic programs: algorithmic techniques and numerical results”, Technical Report, International Institute for Applied Analysis.
- [33] D. H. Salinger, R. T. Rockafellar, “Dynamic splitting: an algorithm for deterministic and stochastic multiperiod optimization”, Stochastic Programming E-print series (SPEPS), 2003.
- [34] R. Van Slyke, R. J-B Wets, “L-shaped linear programs with application to optimal control and stochastic programming”, *SIAM Journal on applied mathematics*, v.17, pp. 638-663, 1969.
- [35] J. F. Benders, “Partitioning procedures for solving mixed-variables programming problems”, *Numerische Mathematik*, v. 4, n.1, pp. 238-252, 1962.
- [36] J. R. Birge, F. V. Louveaux, “A multicut algorithm for two-stage stochastic linear programs”, *European Journal of Operational Research*, v.34, n.3, pp. 384-392, 1988.
- [37] H. I. Gassmann, “MSLiP: A computer code for multistage stochastic linear programming problem”, *Mathematical Programming*, v.47, pp. 407-423, 1990.
- [38] G. Zakeri, A. B. Philpott, D. M. Ryan, “Inexact cuts in Benders decomposition”, *Siam Journal on Optimizaton*, v.10, n.3, pp. 6430657, 2000.
- [39] M. Dempster, R. T. Thompson, “EVPI-based importance sampling solution procedures for multistage stochastic linear programmes on parallel MIMD architectures”, *Annals of Operations Research*, v. 90, pp. 161-184, 1999.
- [40] J. Jacobs, G. Freeman, J. Grygier, “SOCRATES: A system for scheduling hydroelectric generation under uncertainty”, *Annals of Operations Research*, v. 59, n. 1, pp. 99-133, Dec. 1995.

- [41] R. T. Rockafellar, R. J-B. Wets, “Scenarios and policy aggregation in optimization under certainty”, *Mathematics of Operations Research*, v.16, n.1, pp. 119-147, 1991.
- [42] A. Ruszczyński, “Decomposition Methods In Stochastic Programming”, *Mathematical Programming*, v.79, pg. 333-353, 1997.
- [43] G. B. Dantzig, P. Wolfe, “The decomposition principle for linear programs”, *Operations Research*, v.8, n.1, pp. 1010111, 1960.
- [44] J. Linderoth, A. Shapiro, S. Wright, “The empirical behavior of sampling methods for stochastic programming”, *Annals of Operations Research*, v.142, pp. 215-241, 2006.
- [45] J. Dupacova, R. J-B. Wets, “Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems”, *Annals of Statistics*, v16, pp. 1517-1549, 1988.
- [46] A. Shapiro, “Asymptotic analysis of stochastic programs”, *Annals of Operations Research*, v30, pp. 169-186, 1991.
- [47] A. J. King, R. T. Rockafellar, “Asymptotic theory for solutions in statistical estimation and stochastic programming”, *Mathematics of Operations Research*, v.18, pp. 148-162.
- [48] A. Shapiro, T. Homem-de-Mello, “Conditioning of convex piecewise linear stochastic programs”, *Mathematical Programming Series A*, v.94, pp. 1-19, 2002.
- [49] A. Kleywegt, A. Shapiro, T. Homem-de-Mello, “The sample average approximation method for stochastic discrete optimization”, *Siam Journal on Optimizaton*, v.12, pp. 479-502, 2001.
- [50] H. Xu, F. Meng, “Convergence analysis of sample average approximation methods for a class of stochastic mathematical programs with equality constraints”, *Mathematics of Operations Research*, v. 32, n. 3. pp. 648-668, 2007.
- [51] B. K. Pagnoncelli, S. Ahmed, A. Shapiro, “Sample average approximation method for chance constrained programming: theory and applications”, *Journal of Optimization Theory and Applications*, v. 142, n. 2, pp. 399-416, 2009.
- [52] J.L. Hige, S. Sen, “Stochastic decomposition , a statistical method for large scale stochastic linear programming”, *Nonconvex optimization and its applications* 8, 1996.
- [53] J.L. Hige, S. Sen, “Stochastic decomposition: an algorithm for two-stage linear programs with recourse”, *Mathematics of Operations Research*, v.16, n.03, pp. 650-669, 1991.
- [54] J.L. Hige, B. Rayco, S. Sen, “Stochastic scenario decomposition for multistage stochastic programs”, *IMA Journal of Management Mathematics*, v.21, pp. 39-66, 2010.

- [55] J.L. Hige, S. Sen, “Statistical approximations for stochastic linear programming problems”, *Annals of Operations Research*, v.85, pp. 173-192, 1999.
- [56] P. W. Glynn, D. L. Iglehart, “Importance sampling for stochastic simulation”, *Management Science*, v.35, pp. 1367-1392, 1989.
- [57] G. B. Dantzig, P. W. Glynn, “Parallel processors for planning under uncertainty”, *Annals of Operations Research*, v.22, n1, pp. 1-21, 1990.
- [58] G. Infanger, “Monte Carlo (importance) sampling within a Benders decomposition algorithm for stochastic linear programs”, *Annals of Operations Research*, v.39, pp. 69-95, 1992.
- [59] T. A. Rotting, A. Gjelsvik, “Stochastic dual dynamic programming for seasonal scheduling in the Norwegian power system”, *IEEE Transactions on Power Systems*, v. 7, n. 1, pp. 273-279, Feb. 1992.
- [60] O. B. Fosso, A. Gjelsvik, A. Haugstad, B. Mo, I. Wangensteen, “Generation scheduling in a deregulated system. The norwegian case”, *IEEE Transactions on Power Systems*, v. 14, n. 1, pp. 75-81, Feb. 1999.
- [61] C. Beltran, L. F. Escudero, R. E. Rodrigues-Ravines, “Multi-stage stochastic linear programming: an approach by events”, *Optimization Online*, 2008.
- [62] C. Kuchler, S. Vigerske, “Decomposition of multistage stochastic programs with recombining scenario trees”, Technical Report, Humboldt University, 2007.
- [63] Z. L. Chen, W. B. Powell, “Convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse”, *Journal of Optimization Theory and Applications*, v.102, n. 3, pp. 497-524.
- [64] G. Bayraksan, D. P. Morton, “Sequential sampling for solving stochastic programs”, *Winter Simulation Conference*, 2007.
- [65] D. P. Morton, “Stopping rules for a class of sampling-based stochastic programming algorithms”, *Operations Research*, v.46, n. 5, pp. 710-718, 1998.
- [66] A. Shapiro, T. Homem-de-Mello, “A simulation-based approach to two-stage stochastic programming with recourse”, *Mathematical Programming*, v.81, pp. 301-325, 1998.
- [67] W. K. Mak, D. P. Morton, R. K. Wood, “Monte Carlo bounding techniques for determining solution quality in stochastic programs”, *Operations Research Letters*, v.24, pp. 47-56, 1999.
- [68] G. Bayraksan, D. P. Morton, “Assessing solution quality in stochastic programs”, *Mathematical Programming Series B*, v.108, pp. 495-514, 2006.

[69] M. Hindsberger, “Interconnected hydro-thermal systems - Models, methods, and applications”, Phd Thesis, Tech. Univ. Denmark, 2002.